VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ
ÚSTAV MATEMATIKY

Pravděpodobnost, statistika a operační výzkum

(Elektrotechnika, elektronika, komunikační a řídicí technika)

Jaromír Baštinec
Břetislav Fajmon
Jan Koláček

evropský sociální fond v ČR
EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Ústav matematiky FEKT VUT v Brně, 2014

http://www.umat.feec.vutbr.cz

Tento text byl vytvořen v rámci realizace projektu CZ.1.07/2.2.00/15.0156,
Inovace výuky matematických předmětů v rámci studijních programů FEKT a FIT VUT v Brně.

Součástí tohoto učebního textu jsou odkazy na tzv. maplety, tj. programy vytvořené v prostředí Maple. Tyto odkazy jsou v textu zvýrazněny barvou, příp. uvozeny slovem *maplet*. Maplety ke svému běhu nevyžadují software Maple – je však nutné mít na klientském počítači nainstalováno prostředí Java a nastavenou vhodnou úroveň zabezpečení prohlížeče i prostředí Java. Po kliknutí na odkaz mapletu se v závislosti na softwarovém prostředí klientského počítače zobrazí různá hlášení o zabezpečení – všechny dialogy je třeba povolit a spouštění požadovaných prvků neblokovat.

Doplňující součástí tohoto učebního textu jsou příklady zpracované v elektronické bance příkladů.

Obsah

Úvod

Motto:
Učitel Vám může pootevřít dvéře,
vstoupit už musíte sami.

Čínské přísloví

Předmluva

Studijní materiál, který máte v rukou, je přepracovanou verzí elektronické studijní opory [13], jejíž poslední dva autoři už na naší fakultě nepůsobí. Text je určen studentkám a studentům předmětu *Pravděpodobnost, statistika a operační výzkum*, který je zařazen do prvního ročníku navazujícího magisterského studia na FEKT VUT v Brně.

Na základě zkušeností z předchozích ročníků bylo do textu mj. nově zařazeno důkladné opakování potřebného matematického aparátu a především teorie pravděpodobností. Až po vybudování těchto základů bylo přistoupeno k výkladu jednotlivých částí předmětu. Celý text byl upraven a sjednocen.

Rozsah studijního textu vyplývá z obsáhlosti osnovy předmětu. Při inovaci textu jsme byli vedeni snahou zahrnout do textu vše potřebné nejen pro studium daného předmětu, ale i pro případné pozdější aplikace a použití pro řešení konkrétních technických problémů.

Uvítáme jakékoli vaše připomínky a návrhy.

V Brně 31. 3. 2014 Autoři

Označení

$\mathbb{N}$	množina přirozených čísel
$\mathbb{Z}$	množina celých čísel
$\mathbb{R}$	množina reálných čísel
$\mathbb{Q}$	množina racionálních čísel
$\mathbb{I}$	množina iracionálních čísel
$\mathbb{C}$	množina komplexních čísel
$P_n(x)$	polynom n-tého stupně proměnné x
$A_{m,n}$	matice typu m, n (s m řádky a n sloupci)
$A = (a_{ij})$	matice s prvky a_{ij}
I	jednotková matice
$\mathcal{O}$	nulová matice
$det\, A = \|A\|$	determinant matice A
A^{-1}	matice inverzní k matici A
$adj\, A$	matice adjungovaná k matici A
A_{ks}	algebraický doplněk prvku a_{ks}
$hod\,(A)$	hodnost matice A
$(\mathbb{R}^n, +, .)$	vektorový prostor všech uspořádaných n-tic
$\dim P$	dimenze prostoru P
$\square$	konec důkazu, konec řešení
$A \times B$	kartézský součim množin A, B
$\\|A\\|$	míra množiny A
$\{x_n\}_{n=1}^{\infty}$	posloupnost prvků x_n
$A \cap B$	průnik množin A, B
$A \cup B$	sjednocení množin A, B
$\emptyset$	prázdná množina, jev nemožný
I	jev jistý
$\bar{A}$	jev opačný k jevu A
μ	průměr základního souboru
σ^2	rozptyl základního souboru
$\bar{x}$	výběrový průměr
s^2	výběrový rozptyl
s	výběrová směrodatná odchylka
m_k	obecný moment k-tého řádu
M_k	centrální moment k-tého řádu
C	kovariace
R	koeficient korelace

1 Pravděpodobnost

Průvodce studiem

Se základy teorie pravděpodobností jste se seznámili v předmětu Matematika 3. Protože budeme potřebovat pravděpodobnost v průběhu celeho kurzu, zařadili jsme zde důkladné opakování.

Nejdříve si řekneme co jsou to jevy, ketré budeme studovat, zavedeme si klasickou i axiomatickou pravděpodobnost, ukážeme si použití při výpočtech pravděpodobností.

Potom si nadefinujeme náhodnou veličinu a budeme dále studovat její vlastnosti.

Uvedeme si nejčastěji používaná rozdělění a to jak diskrétní, tak i spojitá. V závěru se seznámíme s limitními větami, které nám mohou dobře posloužit při nahrazení jednoho rozdělení jiným, se ketrým se bude lépe počítat.

Cíle

Po prostudování této kapitoly budete schopni:

- Řešit základní úlohy z pravděpodobnosti.
- Pracovat s náhodnou veličinou a jejími charakteristikami.
- Zvládát základní rozdělení a práci s nimi.
- Pracovat s tabulkami hodnot vybraných náhodných veličin.

1.1 Jevy a jejich vlastnosti

Definice 1.1. Na neprázdné množině B definujme operace $\cap$ (průnik) a $\cup$ (sjednocení), které splňují podmínky:

1.

a)	$a \cap b = b \cap a$,	$a \cup b = b \cup a$
b)	$a \cap a = a$,	$a \cup a = a$
c)	$(a \cap b) \cap c = a \cap (b \cap c)$,	$(a \cup b) \cup c = a \cup (b \cup c)$
d)	$a \cup (a \cap b) = a$,	$a \cap (a \cup b) = a$
e)	$a \cup (b \cap c) = (a \cup b) \cap (a \cup c)$,	$a \cap (b \cup c) = (a \cap b) \cup (a \cap c)$

2. V množině B existuje největší prvek I a nejmenší prvek $\emptyset$, pro které platí:
   a) $a \cap \emptyset = \emptyset, \quad a \cup \emptyset = a$
   b) $a \cap I = a, \quad a \cup I = I$

3. Ke každému prvku $a \in B$ existuje *komplement* $\bar{a}$, pro který platí:

   $a \cap \bar{a} = \emptyset, \quad a \cup \bar{a} = I$

Potom $\mathcal{B} = (B, \cap, \cup, \emptyset, I, \bar{\ })$ nazveme *Booleovou algebrou*

Definice 1.2. *Jevem* nazveme výsledek provedeného pokusu.

Jevy mohou být jisté, náhodné, nemožné. Je to dělení relativní, vždy vztažené na daný soubor podmínek.

Příklad 1.3. Při hodu hrací kostkou bude:

jev jistý	-	padne kladný počet bodů,
jev náhodný	-	počet bodů,
jev nemožný	-	padne záporný počet bodů.

1.2 Definice, základní vlastnosti, příklady

Definice 1.4. Jev B je *následkem* jevu A (jev A je částí jevu B), jestliže při nastoupení jevu A nastupuje vždy i jev B.

Označení: $A \subset B$.

Věta 1.5. *Pro libovolné jevy A, B, C platí:*

1. $A \subset A$.

2. *Jestliže* $A \subset B, B \subset C$, *potom* $A \subset C$.

Příklad 1.6. A - při hodu kostkou padne čtyřka.
B - při hodu kostkou padne sudé číslo,
potom je jev $A \subset B$.

Definice 1.7. Jestliže současně platí $A \subset B$ a $B \subset A$, pak jsou jevy A, B *ekvivalentní* a píšeme $A = B$.

Věta 1.8. *Pro libovolné jevy A, B, C platí:*

1. $A = A$.

2. $A = B$ *tehdy a jen tehdy, když* $B = A$.

3. *Jestliže* $A = B, B = C$, *potom* $A = C$.

Definice 1.9. *Průnik* C jevů A, B se nazývá jev ekvivalentní se současným nastoupením jevů A, B. Značíme $C = A \cap B$.

Věta 1.10. *Pro libovolné jevy A, B, C platí:*

1. $A \cap A = A$.
2. $A \cap B = B \cap A$.
3. $(A \cap B) \cap C = A \cap (B \cap C)$.
4. $A \subset B \Rightarrow A \cap B = A$.
5. $(A \cap B) \subset A$.
6. $(A \cap B) \subset B$.
7. *Jestliže* $C \subset A, C \subset B \Rightarrow C \subset (A \cap B)$.

Definice 1.11. *Sjednocením* jevů A, B se nazývá jev C, ekvivalentní s nastoupením alespoň jednoho z jevů A, B. Označení $C = A \cup B$.

Věta 1.12. *Pro libovolné jevy A, B, C platí:*

1. $A \cup A = A$.
2. $A \cup B = B \cup A$.
3. $(A \cup B) \cup C = A \cup (B \cup C)$.
4. $A \subset B \Rightarrow A \cup B = B$.
5. $A \subset A \cup B$.
6. $B \subset A \cup B$.
7. $A \subset C, B \subset C \Rightarrow (A \cup B) \subset C$.

Definice 1.13. *Jistým jevem* nazveme jev, kterě za daného systému podmínek musí vždy nastat. Značíme jej I. *Nemožným jevem* nazveme jev, který za daného systému podmínek nastat nemůže. Značíme jej $\emptyset$.

Věta 1.14. *Pro libovolný jev A platí:*

1. $\emptyset \subset A, A \subset I$.
2. $\emptyset \cap A = \emptyset, I \cap A = A$.
3. $\emptyset \cup A = A, I \cup A = I$.

Definice 1.15. *Jevem opačným* k jevu A rozumíme jev ekvivalentnˇ s tím, že jev A nenastoupí. Označujeme jej $\bar{A}$.

Věta 1.16. *Pro libovolný jev A platí:*

1. $\overline{(\bar{A})} = A$.
2. $A \cup \bar{A} = I, A \cap \bar{A} = \emptyset$.

Věta 1.17. *Pro libovolné jevy A, B platí de Morganovy vzorce:*

1. $\overline{A \cup B} = \bar{A} \cap \bar{B}$.
2. $\overline{A \cap B} = \bar{A} \cup \bar{B}$.

1.3 Elementární jevy

Definice 1.18. *Elementární jevy* jsou takové jevy, různé od jevu $\emptyset$, které se nedají rozložit na další jevy různé od jevu $\emptyset$.
Jev E je elementární, když ze vztahu $E = A \cup B$ plyne $E = A$ nebo $E = B$.
Jestliže máme množinu elementárních jevů $\{E_i, i \in J\}$, pro kterou platí $\bigcup\limits_{i \in J} E_i = I$, pak mluvíme o *úplném systému elementárních jevů.*

Definice 1.19. Jevy A, B se nazývají *disjunktní* (navzájem neslučitelné), nemohou-li nastat současně, tzn. $A \cap B = \emptyset$.

Věta 1.20. *Dva různé elementární jevy jsou navzájem disjunktní.*

Důkaz. Mějme dva různé elementární jevy E_1, E_2. Označme

$$X = E_1 \cap E_2,$$

$$Y = E_1 \cap \bar{E}_2,$$

$$Z = \bar{E}_1 \cap E_2.$$

Potom

$$X \cup Z = E_2,$$

$$X \cup Y = E_1$$

a podle definice 1.18 proto platí, že buď $X = E_1$ a nebo $Y = E_1$.
a) Nechť $Y = E_1$, potom

$$E_1 = E_1 \cap \bar{E}_2,$$

$$X = E_1 \cap E_2,$$

dosazením dostaneme

$$X = (E_1 \cap \bar{E}_2) \cap E_2 = E_1 \cap \emptyset = \emptyset.$$

b) Nechť $X = E_1$, potom

$$E_1 = E_1 \cap E_2 \Rightarrow E_1 \subset E_2.$$

A současně

$$X \cap Z = \emptyset \Rightarrow X = \emptyset.$$

□

1.4 Axiomatická definice pravděpodobnosti

Autorem axiomatické teorie pravděpodobnosti, přijaté dnes na cekém světě, byl sovětský matematik A. N. Kolmogorov. Jeho teorie byla poprvé publikována v roce 1933 a je budována na základě teorie množin a teorie míry.

Věta 1.21. *Nechť M je množina jevů, pro kterou platí:*

1. $A, B \in M \Rightarrow A \cup B \in M, A \cap B \in M,$
2. $\emptyset \in M, I \in M,$
3. $A \in M \Rightarrow \bar{A} \in M.$

Potom $(M, \cup, \cap, \emptyset, I, ^{-})$ je Booleova algebra.

Definice 1.22. *σ-algebrou* nazýváme Booleovu algebru jevů v případě, že jevů může být nekonečně mnoho a platí, že ke kažgdé posloupnosti jevů $\{A_1, A_2, \dots\}$ existuje jejich sjednocení $\bigcup\limits_i A_i$ a průnik $\bigcap\limits_i A_i$.

Definice 1.23. Označme M nějakou σ-algebru jevů. *Pravděpodobnost*, že za určité situace nastane jev $A \in M$ vyjadřuje hodnota funkce $p(A)$, která splňuje podmínky:

1. $p(A) \geq 0 \ \forall A \in M,$
2. $p(I) = 1,$
3. pro množinu $\{A_i\}, i \in J$ navzájem disjunktních jevů platí

$$p\left(\bigcup_{i \in J} A_i\right) = \sum_{i \in J} p(A_i).$$

Dvojici (M, p) budeme nazývat *pravděpodobnostním prostorem*.

Věta 1.24. *Nechť (M, p) je pravděpodobnostní prostor. Potom platí:*

1. *Jestliže $A, B \in M, A \cap B = \emptyset \Rightarrow p(A \cup B) = p(A) + p(B).$*
2. $p(\emptyset) = 0.$
3. $p(A) \leq 1.$
4. $p(A) = 1 - p(\bar{A}).$
5. $p(A \cup B) = p(A) + p(B) - p(A \cap B),$
6. *Jestliže je $A \subset B$, potom $p(A) \leq p(B)$.*

1.5 Klasická pravděpodobnost

Věta 1.25. *Nechť (M, p) je pravděpodobnostní prostor, M je konečná σ-algebra, která obsahuje n elementárních jevů, tvořících úplný systém elementárních jevů*

$$\{E_i, i = 1, \ldots, n\} \text{ takově, §e } p(E_1) = p(E_2) = \cdots = p(E_n) = \frac{1}{n}.$$

Nechť jev $A \in M$ lze rozložit na m navzájem různých elementárních jevů. Potom platí

$$p(A) = \frac{m}{n}.$$

Věta 1.26. *Věta o geometrické pravděpodobnosti.*
Nechť (M, p) je pravděpodobnostní prostor. Nechť σ-algebra jevů je systém podmnožin A množiny I, které mají míru $\|A\|$. Potom

$$p(A) = \frac{\|A\|}{\|I\|}.$$

1.6 Podmíněná pravděpodobnost

Definice 1.27. Nechť (M, p) je pravděpodobnostní prostor. Nechť nastoupil jev $B \in M, p(B) \neq$ $= 0$. *Podmíněnou pravděpodobností* jevu $A \in M$ za předpokladu, že jev B nastal nazveme výraz

$$p_B(A) = \frac{p(A \cap B)}{p(B)}.$$

Věta 1.28. *Vlastnosti podmíněné pravděpodobnosti.*

1. $p_B(\emptyset) = 0.$
2. $p_B(B) = 1.$
3. $A \in M, A \cap B = \emptyset \Rightarrow p_B(A) = 0.$
4. $A \in M, B \subset A \Rightarrow p_B(A) = 1.$
5. $A \in M, B \subset A \Rightarrow p_A(B) = \frac{p(B)}{p(A)}$, *kde* $p_A(B) = \frac{p(A \cap B)}{p(A)}$.
6. $p_I(A) = p(A).$

Věta 1.29. *Nechť (M, p) je pravděpodobnostní prostor, $A, B \in M$. Potom pro pravděpodobnost průniku jevů A, B platí*

$$p(A \cap B) = p_B(A) \cdot p(B).$$

Věta 1.30. *Nechť (M, p) je pravděpodobnostní prostor, $A, B \in M, p(A) \neq 0, p(B) \neq 0$. Potom platí*

$$p_B(A) \cdot p(B) = p_A(B) \cdot p(A).$$

Věta 1.31. *Nechť $M, p)$ je pravděpodobnostní prostor, $A_i \in M, i = 1, \dots, n$. Pro výpočet pravděpodobnosti současného nastoupení n jevů platí*

$$p(A_1 \cap A_2 \cap \ldots \cap A_n) = p(A_1) \cdot p_{A_1}(A_2) \cdot p_{A_1 \cap A_2}(A_3) \ldots p_{A_1 \cap \ldots \cap A_{n-1}}(A_n).$$

1.7 Nezávislé jevy

Definice 1.32. Nechť (M, p) je pravděpodobnostní prostor. Jevy $A, B \in M$ jsou *nezávislé*, jestliže platí aspoň jedna z podmínek:

1. $p(B) = 0$ nebo $p_B(A) = p(A)$,
2. $p(A) = 0$ nebo $p_A(B) = p(B)$,

Podmínky 1. a 2. definice 1.32 jsou ekvivalentní. Při důkazech stačí proto prověřit platnost jen jedné z nich.

Věta 1.33. *Jevy $A, B \in M$ jsou nezávislé, právě tehdy, když platí*

$$p(A \cap B) = p(A)p(B). \tag{1.1}$$

Důkaz.
a) Nechť platí vztah (1.1).
Potom je-li $p(B) = 0$, jsou podle definice 1.32 jevy A, B nezávislé.
Je-li $p(B) \neq 0$, potom

$$p_B(A) = \frac{p(A \cap B)}{p(B)} = \frac{p(A)p(B)}{p(B)} = p(A).$$

a podle definice 1.32 jsou jevy A, B nezávislé.
b) Nechť jsou jevy A, B nezávislé, potom je-li $p(B) = 0$ vztah (1.1) platí. Je-li $p(B) \neq 0$, potom

$$p(A) = p_B(A) = \frac{p(A \cap B)}{p(B)} \Rightarrow p(A \cap B) = p(A)p(B).$$

I v tomto případě vztah (1.1) platí. □

Definice 1.34. Nechť (M, p) je pravděpodobnostní prostor, $M_1 \subset M$. Jevy množiny M_1 jsou *navzájem nezávislé*, jestliže pro každý jev $A \in M_1$ platí, že je nezávislý na libovolném jevu podmnožiny $M_2 \subset \{M_1 \setminus \{A\}\} = M_1 \cap \bar{A}_1$.

Věta 1.35. *Nechť (M, p) je pravděpodobnostní prostor, $A_i \in M, i = 1, \dots, n$. Jestliže jevy množiny $\{A_1, \dots, A_n\}$ jsou navzájem nezávislé, potom*

$$p(A_1 \cap A_2 \cap \ldots \cap A_n) = p(A_1)p(A_2) \ldots p(A_n).$$

Příklad 1.36. Házíme dvěma kostkami. Jev A_1 - padne liché číslo na první kostce, jev A_2 - padne liché číslo na druhé kostce, jev A_3 - součet na obou kostkách je sudé. Určete pravděpodobnosti jevů A_1, A_2, A_3 a jejich nezávislost.

Řešení.

$$p(A_1) = \frac{1}{2}, \quad p(A_2) = \frac{1}{2}, \quad p(A_3) = \frac{1}{2}.$$

$$p_{A_2}A_1 = \frac{p(A_1 \cap A_2)}{p(A_2)} = \frac{\frac{9}{36}}{\frac{1}{2}} = \frac{1}{2} = p(A_1).$$

Jevy A_1, A_2 jsou nezávislé.

$$p_{A_3}A_1 = \frac{p(A_1 \cap A_3)}{p(A_3)} = \frac{\frac{9}{36}}{\frac{1}{2}} = \frac{1}{2} = p(A_1).$$

Jevy A_1, A_3 jsou nezávislé.

$$p_{A_1 \cap A_2}A_3 = 1 \neq p(A_3).$$

Jevy A_1, A_2, A_3 jsou závislé. □

Věta 1.37. *Nechť (M, p) je pravděpodobnostní prostor. Pro každý jev $A \in M$ platí*

1. *$\emptyset, A$ jsou nezávislé jevy,*
2. *I, A jsou nezávislé jevy.*

Věta 1.38. *Jevy $A, B \in M$ jsou nezávislé, jsou-li nezávislé jevy $A, \bar{B}$ či $\bar{A}, B$ či $\bar{A}, \bar{B}$.*

1.8 Úplná pravděpodobnost

Věta 1.39. *O úplné pravděpodobnosti.*
Nechť (M, p) je pravděpodobnostní prostor a $\{B_1, \ldots, B_n\}$ je úplný systém navzájem disjunktních jevů ze σ-algebry M, pro které platí $p(B_j) \neq 0, j = 1, \ldots, n$. Potom pro libovolný jev $A \in M$ platí

$$p(A) = \sum_{j=1}^{n} p_{B_j}(A) \cdot p(B_j).$$

Příklad 1.40. Máme n klobouků a v každém je a bílých a b černých kuliček. Z prvního klobouku náhodně vyjmeme jednu kuličku a přendáme ji do druhého, poté z druhého klobouku přendáme jednu kuličku do třetího, atd., z $n-1$ klobouku přendáme jednu kuličku do posledního klobouku. Z posledního klobouku vyjmeme jednu kuličku. Určete pravděpodobnost, že bude bílá.

Řešení. Pravděpodobnost vytažení bílé kuličky z prvního klobouku je

$$p(B1) = \frac{a}{a+b}.$$

Pravděpodobnost vytažení bílé kuličky z druhého klobouku je podle věty 1.39 o úplné pravděpodobnosti

$$p(B2) = \left(\frac{a+1}{a+b+1}\right)\left(\frac{a}{a+b}\right) + \left(\frac{a}{a+b+1}\right)\left(\frac{b}{a+b}\right) =$$

$$\left(\frac{a}{a+b}\right)\left(\frac{a+1}{a+b+1}+\frac{b}{a+b+1}\right)=\frac{a}{a+b}.$$

Máme tedy $p(B1) = p(B2)$. Pokračujeme dále po indukci a dostaneme

$$p(B1) = p(B2) = \cdots = p(Bn).$$

□

1.9 Bayesova věta

Věta 1.41. *Bayesova věta.*
Nechť (M, p) je pravděpodobnostní prostor a $\{B_1, \dots, B_n\}$ je úplný systém navzájem disjunktních jevů ze σ-algebry M, pro které platí $p(B_j) \neq 0, j = 1, \dots, n$. Potom pro libovolný jev $A \in M$, pro kterě platí $p(A) \neq 0$, platí Bayesův vzorec pro $k = 1, \dots, n$

$$p_A(B_k) = \frac{p_{B_k}(A) \cdot p(B_k)}{\sum\limits_{j=1}^{n} p_{B_j}(A) \cdot p(B_j)}.$$

Příklad 1.42. Ve skupině 10 studentů, kteří se dostavili ke zkoušce, jsou 3 připraveni výborně, 4 dobře , 2 průměrně a 1 špatně. Materiál ke zkoušce obsahuje 20 otázek. Výborně připravený student odpoví na všechny otázky, dobře připravený na 16, průměrně připravený na 10 a špatně připravený na 5. Náhodně vybraný student odpověděl správně na všechny tři náhodn zadané otázky. Určete pravděpodobnost, že šlo o špatně připraveného studenta.

Řešení. Použijeme Bayesův vzorec. Jev A – student odpověděl na všechny tři zadané otázky. Úplný systém disjunktních jevů je
H_1 – výborně připravený student,
H_2 – dobře připravený student,
H_3 – průměrně připravený student,
H_4 – špatně připravený student.

$$p_{H_1}(A) = 1,$$

$$p_{H_2}(A) = \frac{16}{20}\frac{15}{19}\frac{14}{18} \doteq 0.491,$$

$$p_{H_3}(A) = \frac{10}{20}\frac{9}{19}\frac{8}{18} \doteq 0.105,$$

$$p_{H_4}(A) = \frac{5}{20}\frac{4}{19}\frac{3}{18} \doteq 0.009.$$

$$p_A(H_4) = \frac{p_{H_4}(A)p(H_4)}{\sum_{j=1}^{4} p_{H_j}(A)p(H_j)} \doteq 0.002.$$

□

1.10 Opakované pokusy

Věta 1.43. *Bernoulliova posloupnost nezávislých pokusů.*
Provedeme n po sobě jdoucích pokusů, přičemž při každém pokusu může nebo nemusí nastat jev A. Nechť jsou výsledky pokusů na sobě nazávislé a nechť dále v každém z pokusů platí, že $p(A) = p, p(\bar{A}) = q = 1-p$, neboli pravděpodobnost nastoupení jevu A je v kkaždém pokusu stále stejná. Potom pravděpodobnost, že jev A nastane během n pokusů právě k-krát, $k \leq n$ je rovna

$$b(n,p,k) = \binom{n}{k} p^k q^{n-k}.$$

Věta 1.44. *Věta o závislých pokusech*
Nechť je dán soubor N prvků, z nichž M vykazuje sledovaný znak a $N-M$ prvků tento znak nemá. Vybereme postupně náhodně n prvků, z nichž žádný nevracíme zpět. Pravděpodobnost toho, že vybereme právě k prvků majících sledovaný znak a $n-k$ prvků, které tento znak nemají (jev A) je rovna

$$p(A) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}.$$

1.11 Náhodná veličina

Věta 1.45. *Nechť U je σ-algebra číselných množin, generovan intervaly $(-\infty, a), a \in \mathbb{R}$. Potom platí:*

1. $\langle a, +\infty) \in U \ \forall a \in \mathbb{R}$,
2. $(a, +\infty) \in U, \ (-\infty, a\rangle \in U$,
3. $\langle a, b\rangle \in U, \ (a, b) \in U$,
4. $\{x\} \in U \ \forall x \in \mathbb{R}$.

Definice 1.46. *Náhodná veličina* $\mathbb{X}$ je reálná funkce $\mathbb{X}(\omega)$ definovaná na pravděpodobnostním prostoru $(M, p), \omega \in M$, taková, že pro každé $x \in \mathbb{R}$ je množina

$$\{\omega \in M | \mathbb{X}(\omega) < x\}$$

náhodněm jevem. T.j. hodnota náhodné veličiny je jednoznačně určena pokusem a $\forall x \in \mathbb{R}$ můžeme určit pravděpodobnost $p = p(\mathbb{X} < x)$.

1.12 Distribuční funkce

Definice 1.47. Nechť $\mathbb{X}$ je náhodná veličina definovaná na (U, p). Funkci $F(x)$ definovanou vztahem

$$F(x) = p(\mathbb{X} \in (-\infty, x)) = p(\mathbb{X} < x)$$

nazveme *distribuční funkcí* náhodné veličiny $\mathbb{X}$.

Věta 1.48. *Vlastnosti distribuční funkce:*
Nechť $F(x)$ je distribuční funkce. Potom pro všechna $x \in \mathbb{R}$ platí:

1. $0 \leq F(x) \leq 1$.
2. *$F(x)$ je neklesající funkce.*
3. *$F(x)$ je spojitá zleva* $(\lim_{h \to 0^-} F(x+h) = F(x)$.
4. $\lim_{x \to -\infty} F(x) = 0$. $\lim_{x \to +\infty} F(x) = 1$.
5. $p(\mathbb{X} = x) = \lim_{h \to 0^+} F(x+h) - F(x)$.
6. $p(x_1 \leq \mathbb{X} < x_2) = F(x_2) - F(x_1)$.

1.13 Diskrétní a spojitá náhodná veličina

Definice 1.49. Nechť $F(x)$ je stupňovitá funkce, tzn. že existuje posloupnost $\{x_n\}, n = 1, 2, \ldots$ taková, že na intervalech $(x_i, x_{i+1}]$ je $F(x)$ konstantní. Potom se $\mathbb{X}$ nazývá *diskrétní* náhodnou veličinou.
Nechť $F(x)$ je spojitá a po částech hladká na množině $\mathbb{R}$. Potom se $\mathbb{X}$ nazývá *spojitou* náhodnou veličinou.

Důsledek 1.50. *V bodě x, kde je $F(x)$ spojitá platí $p(\mathbb{X} = x) = 0$.*

Věta 1.51. *Nechť je dána funkce $F(x)$ definovaná na $\mathbb{R}$. Splňuje-li $F(x)$ podmínky $1 - 4$ věty 1.48, pak existuje náhodná veličina $\mathbb{X}$, definovaná na (U, p) s distribuční funkcí $F(x)$.*

1.14 Vlastnosti náhodné veličiny

Definice 1.52. Nechť $\mathbb{X}$ je diskrétní náhodná veličina nabývající hodnot z konečné a nebo spočetné číselné množiny S (množina S je množina bodů nespojitosti distribuční funkce $F(x)$). Na množině S definujeme funkci $f(x_i)$ vztahem

$$f(x_i) = p(\mathbb{X} = x_i), \ \forall x_i \in S.$$

Potom funkci $f(x)$ nazveme *frekvenční funkcí* náhodné veličiny $\mathbb{X}$.

Věta 1.53. *Pro distribuční funkci $F(x)$ diskrétní náhodné veličiny $\mathbb{X}$ nabývající hodnot z množiny S a frekvenční funkci $f(x_i)$ platí*

$$F(x) = \sum_{x_i \in S, \, x_i < x} f(x_i).$$

Věta 1.54. *Pro frekvenční funkci $f(x)$ diskrétní náhodné veličiny $\mathbb{X}$ nabývající hodnot z množiny S platí*

$$\sum_{x_i \in S} f(x_i) = 1.$$

Příklad 1.55. Sestavte frekvenční funkci pro Bernoulliovu posloupnost nezávislých jevů pro $n = 5$ a $p = 0.2$.

Řešení. Podle věty 1.43 máme

$$b(n, p, k) = \binom{n}{k} p^k q^{n-k}.$$

Po dosazení dostaneme

k	0	1	2	3	4	5
$f(k)$	0.32768	0.40960	0.20480	0.05120	0.00640	0.00032

Tím máme určenou hledanou frekvenční funkci. □

Definice 1.56. Nechť $\mathbb{X}$ je spojitá náhodná veličina s distribuční funkcí $F(x)$. *Funkcí hustoty pravděpodobnosti* náhodné veličiny $\mathbb{X}$ nazveme reálnou funkci $f(x)$ pro kterou platí:

1. $f(x) \geq 0 \; \forall x \in \mathbb{R}$,
2. $\int\limits_{-\infty}^{+\infty} f(x)dx = 1$,
3. $F(x) = \int\limits_{-\infty}^{x} f(t)dt$.

Důsledek 1.57. *1. V bodě x, kde je $F(x)$ spojitá, platí $f(x) = F'(x)$.*

2. Funkce hustoty $f(x)$ je buď spojitá a nebo po částech spojitá.

3. Platí $\lim\limits_{x \to \pm\infty} f(x) = 0$.

Příklad 1.58. Máme zadanou funkci

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in < a, b >, \\ 0, & \text{jinak.} \end{cases}$$

Ověřte, že se jedná o funkci hustoty některé náhodné veličiny a určete její distribuční funkci.

Řešení.

$$\int_{-\infty}^{+\infty} f(x)dx = \int_a^b \frac{1}{b-a}dx = \frac{1}{b-a}x \Big|_a^b = \frac{1}{b-a}(b-a) = 1.$$

Protože funkce $f(x)$ je nezáporná, jedná se o funkci hustoty.

Pro distribuční funkci platí

$$F(x) = \int_{-\infty}^{x} f(t)dt.$$

Po integraci dostaneme

$$F(x) = \begin{cases} 0, & x < a, \\ \dfrac{x-a}{b-a}, & a \le x \le b, \\ 1, & x > b. \end{cases}$$

□

Věta 1.59. *Pro spojitou náhodnou veličinu $\mathbb{X}$ s distribuční funkcí $F(x)$ a pro libovolná reálná čísla $x_1, x_2, x_1 < x_2$ platí*

$$p(x_1 < \mathbb{X} < x_2) = p(x_1 \le \mathbb{X} \le x_2) = p(x_1 < \mathbb{X} \le x_2) = p(x_1 \le \mathbb{X} < x_2) = F(x_2) - F(x_1).$$

Věta 1.60. *Pro spojitou náhodnou veličinu $\mathbb{X}$ s funkcí hustoty $f(x)$ a pro libovolná reálná čísla $x_1, x_2, x_1 < x_2$ platí*

$$p(x_1 \le \mathbb{X} \le x_2) = \int_{x_1}^{x_2} f(t)dt.$$

Důsledek 1.61. *Funkce hustoty $f(x)$ je buď spojitá a nebo po částech spojitá funkce. Dále platí $\lim\limits_{x \to -\infty} f(x) = 0$, $\lim\limits_{x \to +\infty} f(x) = 0$.*

1.15 Vícerozměrná náhodná veličina

Definice 1.62. Nechť $\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_n$ jsou náhodné veličiny definované na pravděpodobnostních prostorech $(U_1, p_1), (U_2, p_2), \ldots, (U_n, p_n)$. Nechť všechny veličiny jsou stejného typu (buď diskrétní a nebo spojité). Potom n-tici $(\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_n)$ nazveme *n-rozměrnou náhodnou veličinou.*

V případě $n = 2$ dostaneme dvourozměrnou náhodnou veličinu, kterou budeme označovat jako dvojici $(\mathbb{X}, \mathbb{Y})$.

Definice 1.63. Nechť $(\mathbb{X}, \mathbb{Y})$ je dvourozměrná náhodná veličina. *Simultání distribuční funkcí* $F(x, y)$ budeme nazývat funkci definovanou vztahem

$$F(x, y) = p(\mathbb{X} < x \wedge \mathbb{Y} < y).$$

Věta 1.64. *Simultání distribuční funkce $F(x, y)$ dvourozměrné náhodné veličiny $(\mathbb{X}, \mathbb{Y})$ má následující vlastnosti:*

1. $0 \le F(x, y) \le 1$,
2. *jestliže* $x_1 \le x_2, y_1 \le y_2$, *potom* $F(x_1, y_1) \le F(x_2, y_2)$,
3. $\lim\limits_{x \to x_0^-, y \to y_0^-} F(x, y) = F(x_0, y_0)$,
4. $\lim\limits_{x \to -\infty} F(x, y) = \lim\limits_{y \to -\infty} F(x, y) = \lim\limits_{x \to -\infty, y \to -\infty} F(x, y) = 0$,
5. $\lim\limits_{x \to +\infty, y \to +\infty} F(x, y) = 1$.

Definice 1.65. Jsou-li u dvourozměrné náhodné veličiny $(\mathbb{X}, \mathbb{Y})$ obě veličiny $\mathbb{X}$ i $\mathbb{Y}$ diskrétní, pak hovoříme o diskrétní dvourozměrné náhodné veličině, která nabývá nejvýše spočetně mnoha hodnot tvořících množinu T. Na T definujeme *simultání frekvenční funkci*

$$f(x_i, y_j) = p(\mathbb{X} = x_i, \mathbb{Y} = y_j).$$

Věta 1.66. *Pro dvourozměrnou diskrétní náhodnou veličinu platí*

$$F(x, y) = \sum_{x_i < x, y_j < y} f(x_i, y_j),$$

$$\sum_{(x_i, y_j) \in T} f(x_i, y_j) = 1.$$

Definice 1.67. Jsou-li u dvourozměrné náhodné veličiny $(\mathbb{X}, \mathbb{Y})$ obě veličiny $\mathbb{X}$ i $\mathbb{Y}$ spojité, pak hovoříme o spojité dvourozměrné náhodné veličině. *Simultání funkce hustoty* $f(x, y)$ je funkce pro kterou platí

1. $f(x, y) \geq 0 \ \forall x, y \in \mathbb{R}$,
2. $\int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} f(x, y) dx dy = 1$,
3. $F(x, y) = \int\limits_{-\infty}^{y} \left(\int\limits_{-\infty}^{x} f(u, v) du \right) dv$.

Věta 1.68. *Pro dvourozměrnou spojitou náhodnou veličinu platí*

1. $p(\mathbb{X} = x, \mathbb{Y} = y) = 0 \ \forall x, y \in \mathbb{R}$,
2. *V bodech spojitosti funkce hustoty je* $f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y)$.

1.16 Marginální rozložení

Věta 1.69. *Mějme dvourozměrnou náhodnou veličinu* $(\mathbb{X}, \mathbb{Y})$. *Každá z veličin* $\mathbb{X}, \mathbb{Y}$ *má svoje rozdělení. Nechť* $\mathbb{X}$ *má distribuční funkci* $F_1(x)$, $\mathbb{Y}$ *má distribuční funkci* $F_2(y)$ *a* $(\mathbb{X}, \mathbb{Y})$ *má simultání distribuční funkci* $F(x, y)$.*Pak platí*

1. $p(\mathbb{X} < x) = p(\mathbb{X} < x, \mathbb{Y} \in (-\infty, +\infty))$,
2. $p(\mathbb{Y} < y) = p(\mathbb{X} \in (-\infty, +\infty), \mathbb{Y} < y)$,
3. $F_1(x) = \lim\limits_{y \to +\infty} F(x, y)$,
4. $F_2(y) = \lim\limits_{x \to +\infty} F(x, y)$.

Definice 1.70. Distribuční funkce $F_1(x), F_2(y)$ se nazývají *marginální* (okrajové) distribuční funkce.

Definice 1.71. Pro diskrétní dvourozměrnou náhodnou veličinu $(\mathbb{X}, \mathbb{Y})$ se simultání frekvenční funkcí $f(x_i, y_j)$ definujeme *marginální* (okrajové) frekvenční funkce $f_1(x_i)$, $f_2(y_j)$ následovně: Nechť $\mathbb{X}$ nabývá hodnot z množiny T_1, $\mathbb{Y}$ nabývá hodnot z množiny T_2, potom

$$f_1(x_i) = \sum_{y_j \in T_2} f(x_i, y_j), \;\; f_2(y_j) = \sum_{x_i \in T_1} f(x_i, y_j).$$

Pro spojitou dvourozměrnou náhodnou veličinu $(\mathbb{X}, \mathbb{Y})$ se simultání funkcí hustoty $f(x, y)$ definujeme *marginální* (okrajové) funkce hustoty $f_1(x), f_2(y)$ následovně:

$$f_1(x) = \int_{-\infty}^{+\infty} f(x, y)dy,$$

$$f_2(y) = \int_{-\infty}^{+\infty} f(x, y)dx.$$

1.17 Nezávislé náhodné veličiny

Definice 1.72. Nechť $\mathbb{X}$ je náhodná veličina definovaná na pravděpodobnostním prostoru (U, p_1), $\mathbb{Y}$ je náhodná veličina definovaná na pravděpodobnostním prostoru (U, p_2). Jestliže pro všechny množiny $A, B \in U$ jsou jevy $\mathbb{X} \in A, \mathbb{Y} \in B$ navzájem nezávislé (t.j. platí $p(\mathbb{X} \in A, \mathbb{Y} \in B) = p_1(\mathbb{X} \in A) \cdot p_2(\mathbb{Y} \in B)$), pak nazýváme $\mathbb{X}, \mathbb{Y}$ *nezávislými náhodnými veličinami.*

Věta 1.73. *Nechť $(\mathbb{X}, \mathbb{Y})$ je dvourozměrná náhodná veličina se simultání distribuční funkcč $F(x, y)$. Nechť $F_1(x), F_2(y)$ jsou marginální distribuční funkce. Potom $\mathbb{X}, \mathbb{Y}$ jsou nezávislé náhodné veličiny právě tehdy, když*

$$F(x, y) = F_1(x) \cdot F_2(y).$$

Věta 1.74. *Nechť $(\mathbb{X}, \mathbb{Y})$ je diskrétní dvourozměrná náhodná veličina se simultání frekvenční funkcí $f(x_i, y_j)$. Nechť $f_1(x_i), f_2(y_j)$ jsou marginální frekvenční funkce. Potom $\mathbb{X}, \mathbb{Y}$ jsou nezávislé náhodné veličiny právě tehdy, když*

$$f(x_i, y_j) = f_1(x_i) \cdot f_2(y_j).$$

Věta 1.75. *Nechť $(\mathbb{X}, \mathbb{Y})$ je spojitá dvourozměrná náhodná veličina se simultání funkcí hustoty $f(x, y)$. Nechť $f_1(x), f_2(y)$ jsou marginální funkce hustoty. Potom $\mathbb{X}, \mathbb{Y}$ jsou nezávislé náhodné veličiny právě tehdy, když*

$$f(x, y) = f_1(x) \cdot f_2(y).$$

Důkaz.
a) Nechť $\mathbb{X}, \mathbb{Y}$ jsou nezávislé náhodné veličiny, potom platí $F(x, y) = F_1(x)F_2(y)$.

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} = \frac{\partial^2 (F_1(x)F_2(y))}{\partial x \partial y} = \frac{\partial F_1(x)}{\partial x}\frac{\partial F_2(y)}{\partial y} = f_1(x)f_2(y).$$

b) Platí-li $f(x,y) = f_1(x)f_2(y)$, potom

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(x,y)dxdy = \int_{-\infty}^{x}\int_{-\infty}^{y} f_1(x)f_2(y)dxdy =$$

$$\int_{-\infty}^{x} f_1(x)dx \int_{-\infty}^{y} f_2(y)dy = F_1(x)F_2(y).$$

Veličiny $\mathbb{X}, \mathbb{Y}$ jsou nezávislé. □

Příklad 1.76. Nechť $(\mathbb{X}, \mathbb{Y})$ je spojitá náhodná veličina se simultání funkcí hustoty

$$f(x,y) = \begin{cases} \frac{1}{4}, & x \in < -1, 1 >, y \in < -1, 1 >, \\ 0, & \text{jinak.} \end{cases}$$

Rozhodněte, zda jsou $(\mathbb{X}, \mathbb{Y})$ nezávislé.

Řešení. Pro marginální funkce hustoty platí

$$f_1(x) = \int_{-\infty}^{+\infty} f(x,y)dy, f_2(y) = \int_{-\infty}^{+\infty} f(x,y)dx.$$

Po integraci dostaneme

$$f_1(x) = \begin{cases} 0, & x < -1, \\ \frac{1}{2}, & x \in < -1, 1 >, \\ 0, & x > 1. \end{cases}$$

Analogicky pro druhou proměnnou dostaneme

$$f_2(y) = \begin{cases} 0, & y < -1, \\ \frac{1}{2}, & y \in < -1, 1 >, \\ 0, & y > 1. \end{cases}$$

Takže platí $f(x,y) = f_1(x)f_2(y)$ a podle věty 1.75 jsou $\mathbb{X}, \mathbb{Y}$ nezávislé. □

1.18 Transformace náhodných veličin

Definice 1.77. Mějme pravděpodobnostní prostor (U, p). Nechť $y = y(x)$ je funkce definovan na $(-\infty, +\infty)$ přiřazující každé množině $A \in U$ množinu $B = \{x|y(x) \in A\}$, pro kterou platí $B \in U$. Nechť $\mathbb{X}$ je náhodná veličina. Funkcí náhodné veličiny $\mathbb{X}$ rozumíme náhodnou veličinu $\mathbb{Y} = y(\mathbb{X})$, která nabývá hodnoty y právě když náhodná veličina $\mathbb{X}$ nabude takové hodnoty x, že platí $y = y(x)$. Pro libovolnou množinu $A \in U$ a $B = \{x|y(x) \in A\}$ platí

$$p(\mathbb{Y} \in A) = p(\mathbb{X} \in B).$$

Věta 1.78. *Nechť $\mathbb{Y} = y(\mathbb{X})$ je fukce náhodné veličiny $\mathbb{X}$. Nechť náhodná veličina $\mathbb{Y}$ má distribuční funkci $G(y)$. Potom platí*

$$G(y_0) = p(\mathbb{Y} < y_0) = p(\mathbb{X} \in \{x|y(x) < y_0\}).$$

Příklad 1.79. Nechť náhodná veličina $\mathbb{X}$ má distribuční funkci $F(x)$. Určete distribuční funkci náhodné veličiny $\mathbb{Y} = \mathbb{X}^2$.

Řešení. Postupujeme přesně podle předchozí věty.
1) Nechť $y_0 \leq 0$. Potom $\{x|x^2 < y_0\} = \emptyset$, a proto

$$G(y_0) = p(\mathbb{Y} < y_0) = p(\mathbb{X} \in \{x|x^2 < y_0 \leq 0\}) = 0.$$

2) Nechť $y_0 > 0$. Potom $\{x|x^2 < y_0\} = (-\sqrt{y_0}, \sqrt{y_0})$ a tedy

$$G(y_0) = p(\mathbb{Y} < y_0) = p(-\sqrt{y_0} < \mathbb{X} < \sqrt{y_0}) = F(\sqrt{y_0}) - F(-\sqrt{y_0}).$$

Proto

$$G(y) = \begin{cases} 0, & y \leq 0, \\ F(\sqrt{y}) - F(-\sqrt{y}), & y > 0. \end{cases}$$

□

Věta 1.80. *Nechť $g(x)$ je monotonně rostoucí funkce. Jestliže $\mathbb{X}$ je náhodná veličina s distribuční funkcí $F(x)$. Potom $\mathbb{Y} = g(\mathbb{X})$ je náhodná veličina s distribuční funkcí*

$$G(y) = F(g^{-1}(y)).$$

Věta 1.81. *Nechť $g(x)$ je monotonně klesající funkce. Jestliže $\mathbb{X}$ je náhodná veličina s distribuční funkcí $F(x)$. Potom $\mathbb{Y} = g(\mathbb{X})$ je náhodná veličina s distribuční funkcí*

$$G(y) = 1 - F(g^{-1}(y)).$$

Věta 1.82. *Nechť $g(x)$ je ryze monotonní funkce. Jestliže $\mathbb{X}$ je náhodná veličina s distribuční funkcí $F(x)$ a funkcí hustoty $f(x)$. Potom $\mathbb{Y} = g(\mathbb{X})$ je náhodná veličina s funkcí hustoty*

$$h(y) = f\left(g^{-1}(y)\right) \left| \frac{d}{dy}(g^{-1}(y)) \right|.$$

Důsledek 1.83.

$$y = ax + b, a \neq 0 \Rightarrow g(y) = \frac{1}{|a|} f\left(\frac{y-b}{a}\right).$$

1.19 Charakteristiky náhodných veličin

Definice 1.84. Nechť $\mathbb{X}$ je diskrétní náhodná veličina s frekvenční funkcí $f(x_i)$ definovaná na množině S. *Obecným momentem k-tého řádu $m_k(\mathbb{X})$* náhodné veličiny $\mathbb{X}$ nazveme

$$m_k(\mathbb{X}) = \sum_{x_i \in S} x_i^k f(x_i), \quad k = 1, 2, \ldots.$$

Definice 1.85. Nechť $\mathbb{X}$ je spojitá náhodná veličina s funkcí hustoty $f(x)$. *Obecným momentem k-tého řádu $m_k(\mathbb{X})$* náhodné veličiny $\mathbb{X}$ nazveme

$$m_k(\mathbb{X}) = \int_{-\infty}^{+\infty} x^k f(x) dx, \quad k = 1, 2, \ldots,$$

pokud integrál vpravo existuje.

Definice 1.86. *Střední hodnotou* náhodné veličiny $\mathbb{X}$ nazveme číslo $E(\mathbb{X}) = m_1(\mathbb{X})$.

Důsledek 1.87. *Aritmetický průměr je speciálním případem střední hodnoty v případě, že $f(x_i) = c > 0, \forall i$.*

Důkaz. Mějme diskrétní náhodnou veličinu $\mathbb{X}$ s frekvenční funkcí $f(x_i) = \frac{1}{n}, i = 1, 2, \ldots, n$ Potom střední hodnota $E(\mathbb{X})$ je

$$E(\mathbb{X}) = \sum_{i=1}^{n} x_i \frac{1}{n} = \frac{1}{n}(x_1 + x_2 + \cdots + x_n) = \bar{x}.$$

□

Definice 1.88. Nechť $\mathbb{X}$ je diskrétní náhodná veličina s frekvenční funkcí $f(x_i)$ definovaná na množině S. *Centrálním momentem k-tého řádu* $M_k(\mathbb{X})$ náhodné veličiny $\mathbb{X}$ nazveme

$$M_k(\mathbb{X}) = \sum_{x_i \in S} (x_i - E(\mathbb{X}))^k f(x_i), \;\; k = 1, 2, \ldots.$$

Definice 1.89. Nechť $\mathbb{X}$ je spojitá náhodná veličina s funkcí hustoty $f(x)$. *Centrálním momentem k-tého řádu* $M_k(\mathbb{X})$ náhodné veličiny $\mathbb{X}$ nazveme

$$M_k(\mathbb{X}) = \int_{-\infty}^{+\infty} (x - E(\mathbb{X}))^k f(x) dx, \;\; k = 1, 2, \ldots,$$

pokud integrál vpravo existuje.

Důsledek 1.90. *Pro libovolnou náhodnou veličinu $\mathbb{X}$ platí $M_1(\mathbb{X}) = 0$*

Důkaz.

$$M_1(\mathbb{X}) = \int_{-\infty}^{+\infty} (x - E(\mathbb{X})) f(x) dx = \int_{-\infty}^{+\infty} (x f(x) - E(\mathbb{X}) f(x)) \, dx =$$

$$\int_{-\infty}^{+\infty} x f(x) dx - E(\mathbb{X}) \int_{-\infty}^{+\infty} f(x) dx = E(\mathbb{X}) - E(\mathbb{X}) = 0.$$

□

Definice 1.91. *Rozptylem* (disperzí) náhodné veličiny $\mathbb{X}$ nazveme číslo $D(\mathbb{X}) = M_2(\mathbb{X})$.

Věta 1.92. *Nechť $\mathbb{X}$ je náhodná veličina. Potom pro náhodnou veličinu $\mathbb{Y} = a\mathbb{X} + b, a \neq 0$ platí*

$$E(\mathbb{Y}) = aE(\mathbb{X}) + b,$$

$$D(\mathbb{Y}) = a^2 D(\mathbb{X}).$$

Důkaz. Podle důsledku 1.83 věty 1.82 platí

$$y = ax + b, a \neq 0 \Rightarrow g(y) = \frac{1}{|a|} f\left(\frac{y - b}{a}\right).$$

Dále

$$E(\mathbb{Y}) = \int_{-\infty}^{+\infty} y g(y) dy = \int_{-\infty}^{+\infty} y \frac{1}{|a|} f\left(\frac{y-b}{a}\right) dy = (*).$$

Zavedeme si substituci

$$y = ax + b,$$

$$dy = adx.$$

Pro $a > 0$ zůstavají hranice beze změny, pro $a < 0$ se změní na opačné.

$$(*) = \int_{-\infty}^{+\infty} \frac{ax+b}{|a|} f(x)|a|dx = \int_{-\infty}^{+\infty} (axf(x) + bf(x))\, dx = aE(\mathbb{X}) + b.$$

Pro rozptyl máme analogicky

$$D(\mathbb{Y}) = \int_{-\infty}^{+\infty} (\mathbb{Y} - E(\mathbb{Y}))^2 g(y) dy = \int_{-\infty}^{+\infty} (ax + b - aE(\mathbb{X}) - b)^2 \frac{1}{|a|} f(x)|a|dx =$$

$$\int_{-\infty}^{+\infty} a^2 (x - E(\mathbb{X}))^2 f(x) dx = a^2 D(\mathbb{X}).$$

□

Věta 1.93. *Nechť $\mathbb{X}$ je náhodná veličina. Potom pro náhodnou veličinu $\mathbb{Y} = (\mathbb{X} - E(\mathbb{X}))^2$ platí*

$$D(\mathbb{X}) = E(\mathbb{Y}).$$

Věta 1.94. *Nechť $\mathbb{X}$ je náhodná veličina. Potom platí*

$$D(\mathbb{X}) = E(\mathbb{X}^2) - (E(\mathbb{X}))^2 .$$

Důkaz.

$$D(\mathbb{X}) = \int_{-\infty}^{+\infty} (x - E(\mathbb{X}))^2 f(x) dx = \int_{-\infty}^{+\infty} \left(x^2 - 2E(\mathbb{X})x + (E(\mathbb{X}))^2\right) f(x) dx =$$

$$\int_{-\infty}^{+\infty} x^2 f(x) dx - 2E(\mathbb{X}) \int_{-\infty}^{+\infty} x f(x) dx + (E(\mathbb{X}))^2 \int_{-\infty}^{+\infty} f(x) dx =$$

$$E(\mathbb{X}^2) - 2E(\mathbb{X})E(\mathbb{X}) + (E(\mathbb{X}))^2 = E(\mathbb{X}^2) - (E(\mathbb{X}))^2.$$

□

Definice 1.95. Nechť $\mathbb{X}$ je náhodná veličina. Číslo $\sigma(\mathbb{X}) = \sqrt{D(\mathbb{X})}$ nazveme *směrodatnou odchylkou* náhodné veličiny $\mathbb{X}$.

Definice 1.96. *Normovaná náhodná veličina* k náhodné veličině $\mathbb{X}$ je

$$\mathbb{U} = \frac{\mathbb{X} - E(\mathbb{X})}{\sigma(\mathbb{X})}.$$

Věta 1.97. *Pro normovanou náhodnou veličinu $\mathbb{U}$ platí*

$$E(\mathbb{U}) = 0, \;\; D(\mathbb{U}) = 1.$$

Důkaz.
$$\mathbb{U} = \frac{\mathbb{X} - E(\mathbb{X})}{\sigma(\mathbb{X})} = \frac{1}{\sigma(\mathbb{X})}\mathbb{X} - \frac{E(\mathbb{X})}{\sigma(\mathbb{X})}.$$
Podle věty 1.92
$$E(\mathbb{U}) = \frac{1}{\sigma(\mathbb{X})}E(\mathbb{X}) - \frac{E(\mathbb{X})}{\sigma(\mathbb{X})} = 0.$$
Dále podle téže věty
$$D(\mathbb{U}) = \frac{1}{(\sigma(\mathbb{X}))^2}D(\mathbb{X}) = \frac{1}{D(\mathbb{X})}D(\mathbb{X}) = 1.$$
□

Definice 1.98. Nechť $\mathbb{X}$ je náhodná veličina. Číslo
$$k_1 = \frac{M_3(\mathbb{X})}{(\sigma(\mathbb{X}))^3}$$
nazveme *koeficientem šikmosti* náhodné veličiny $\mathbb{X}$.

Definice 1.99. Číslo
$$k_2 = \frac{M_4(\mathbb{X})}{(\sigma(\mathbb{X}))^4} - 3$$
nazveme *koeficientem špičatosti* náhodné veličiny $\mathbb{X}$.

Věta 1.100. *Nechť $\mathbb{X}$ je náhodná veličina. Koeficient šikmosti i koeficient špičatosti se nemění při lineární transformaci. T.j. $\forall a, b \in \mathbb{R}, a \neq 0$:*
$$k_1(a\mathbb{X} + b) = k_1(\mathbb{X}),$$
$$k_2(a\mathbb{X} + b) = k_2(\mathbb{X}),$$

Věta 1.101. *Platí:*

1. *Pro symetrickou náhodnou veličinu $\mathbb{X}$ je $k_1(\mathbb{X}) = 0$.*
2. *Pro náhodnou veličinu s rozdělením protáhlejším vpravo je $k_1 > 0$ a pro náhodnou veličinu s rozdělením protáhlejším vlevo je $k_1 < 0$.*
3. *Pro normální rozdělení je $k_2 = 0$.*
4. *Má-li náhodná veličina symetrické rozdělení a je-li $k_2 > 0$, potom je funkce hustoty pro $x \to \pm\infty$ větší než funkce hustoty normálního rozdělení se stejnou střední hodnotou.*
5. *Má-li náhodná veličina symetrické rozdělení a je-li $k_2 < 0$, potom je funkce hustoty pro $x \to \pm\infty$ menší než funkce hustoty normálního rozdělení se stejnou střední hodnotou.*

Definice 1.102. Nechť $0 < p < 1$. p-*kvantil* náhodné veličiny $\mathbb{X}$ je číslo x_p takové, že $F(x_p) \leq$ $\leq p$, $F(x_p + 0) > p$.

Poznámka 1.103. p-kvantil není určen jednoznačně. Je-li $\mathbb{X}$ spojitá náhodná veličina je $F(x_p) = p$.

Definice 1.104. 0.5-kvantil ve nazývá *medián.*

Definice 1.105. *Modus* $\tilde{x}$ spojité náhodné veličiny $\mathbb{X}$ je bod pro který platí $f(\tilde{x}) \geq f(x) \forall x \in \mathbb{R}$.

Příklad 1.106. Náhodná veličina má funkci hustoty

$$f(x) = \begin{cases} 0, & \text{pro } x < 0, \\ a(3x - x^2), & \text{pro } 0 \leq x \leq 3, \\ 0, & \text{pro } x > 3. \end{cases}$$

Určete parametr a, střední hodnotu a rozptyl.

Řešení. Musí platit

$$\int_{-\infty}^{+\infty} f(x)dx = 1.$$

$$\int_{-\infty}^{+\infty} f(x)dx = \int_0^3 a(3x - x^2)dx = a\left[\frac{3}{2}x^2 - \frac{1}{3}x^3\right]_0^3 = a\left[\frac{27}{2} - 9\right] = 1 \to a = \frac{2}{9}.$$

$$E(x) = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^3 \left(\frac{2}{3}x^2 - \frac{2}{9}x^3\right) dx = \left[\frac{2}{9}x^3 - \frac{1}{18}x^4\right]_0^3 = 6 - \frac{9}{2} = 1.5.$$

$$D(x) = \int_{-\infty}^{+\infty} (x - E(x))^2 f(x)dx = E(x^2) - (E(x))^2 = \int_0^3 \left(\frac{2}{3}x^3 - \frac{2}{9}x^4\right) dx - 1.5^2 =$$

$$\left[\frac{1}{6}x^4 - \frac{2}{45}x^5\right]_0^3 - 1.5^2 = 2.7 - 2.25 = 0.45.$$

□

Příklad 1.107. Graf funkce hustoty náhodné veličiny X tvoří na intervalu $[0, a]$ přímka spojující body $(0, 2/a)$ a $(a, 0)$. Mimo interval $[0, a]$ je funkce hustoty nulová. Urźete: hodnotu parametru a, střední hodnotu a rozptyl.

Řešení.

$$f(x) = \begin{cases} 0 & x < 0, \\ -\frac{2}{a^2}x + \frac{2}{a} & x \in [0, a], \\ 0 & , x > a. \end{cases}$$

Musí platit

$$\int_{-\infty}^{+\infty} f(x)dx = 1 \Rightarrow \int_0^a \left(-\frac{2}{a^2}x + \frac{2}{a}\right) dx = -\frac{2}{a^2}\frac{x^2}{2} + \frac{2}{a}x\Big|_0^a = -\frac{2}{a^2}\frac{a^2}{2} + \frac{2}{a}a = 1.$$

Každá hodnota $a > 0$ vyhovuje.

$$E(x) = \int_0^a \left(-\frac{2}{a^2}x^2 + \frac{2}{a}x\right) dx = -\frac{2}{3}a + a = \frac{a}{3}.$$

$$D(x) = \frac{a^2}{18}.$$

□

1.20 Číselné charakteristiky dvourozměrných náhodných veličin

Definice 1.108. *Centrálním bodem* dvourozměrné náhodné veličiny $(\mathbb{X},\mathbb{Y})$ nazveme bod $(E(\mathbb{X}),E(\mathbb{Y}))$.

Definice 1.109. Nechť $\mathbb{X},\mathbb{Y}$ jsou náhodné veličiny. Hodnotu

$$K(\mathbb{X},\mathbb{Y})=E(\mathbb{X}\mathbb{Y})-E(\mathbb{X})E(\mathbb{Y})$$

nazýváme *kovariací* náhodných veličin $\mathbb{X},\mathbb{Y}$.

Věta 1.110. *Platí:*

1. $K(\mathbb{X},\mathbb{Y})=K(\mathbb{Y},\mathbb{X})$.

2. $K(\mathbb{X},\mathbb{Y})=0$ *pro nezávislé náhodné veličiny* $\mathbb{X},\mathbb{Y}$.

3. $K(\mathbb{X},\mathbb{X})=D(\mathbb{X})$.

4. $K(\mathbb{X},\mathbb{X})=E(\mathbb{X}^2)-[E(\mathbb{X})]^2$.

Definice 1.111. Nechť $(\mathbb{X},\mathbb{Y})$ jsou náhodné veličiny. Hodnotu

$$R(\mathbb{X},\mathbb{Y})=\frac{K(\mathbb{X},\mathbb{Y})}{\sigma(\mathbb{X})\sigma(\mathbb{Y})}$$

nazýváme *korelačním koeficientem* náhodných veličin $\mathbb{X},\mathbb{Y}$.

Věta 1.112. *Platí:*

1. $-1\le R(\mathbb{X},\mathbb{Y})\le 1$.
2. $R(\mathbb{X},\mathbb{Y})=R(\mathbb{Y},\mathbb{X})$.
3. $R(\mathbb{X},\mathbb{Y})=0$ *pro nezávislé* $\mathbb{X},\mathbb{Y}$.
4. *Je-li* $\mathbb{Y}=a\mathbb{X}+b, a\neq 0$, *potom* $R(\mathbb{X},\mathbb{Y})=$ *sign* a.

Příklad 1.113. Dvourozměrná náhodná veličina $(\mathbb{X},\mathbb{Y})$ má funkci hustoty

$$f(x,y)=\begin{cases} a(x^2+y^2), & \text{pro } x^2+y^2\le r^2,\\ 0, & \text{pro } x^2+y^2>r^2.\end{cases}$$

Určete koeficient a.

Řešení. Koeficient a určíme z rovnice

$$a\int\int (x^2+y^2)dxdy = 1,$$

kde integrujeme přes kružnici $x^2+y^2=r^2$. Přejdeme k polárním souřadnicím a dostaneme

$$a\int_0^{2\pi}\int_0^r \varrho^3 d\varrho d\varphi = 1,$$

$$\frac{r^4}{4}2\pi a = 1,$$

$$a = \frac{2}{\pi r^4}.$$

□

Příklad 1.114. Dvourozměrná náhodná veličina $(\mathbb{X},\mathbb{Y})$ má funkci hustoty

$$f(x,y) = \begin{cases} a\sin(x+y), & \text{v oblasti } D,\\ 0, & \text{mimo oblast } D,\end{cases}$$

kde oblast D je urźena nerovnostmi $0\le x\le \pi/2, 0\le y\le \pi/2$.
Určete

1. Koeficient a,
2. Střední hodnoty $E(x), E(y)$,
3. Směrodatné odchylky $\sigma(x), \sigma(y)$,
4. Koeficient korelace.

Řešení. 1. Musí platit

$$a\int_0^{\pi/2}\int_0^{\pi/2}\sin(x+y)dydx = 1.$$

Odtud

$$a\int_0^{\pi/2}\int_0^{\pi/2}\sin(x+y)dydx = -a\int_0^{\pi/2} cos(x+y)|_0^{\pi/2}\,dx =$$

$$= -a\int_0^{\pi/2}(\sin x+\cos x)dx = a\;(\sin x - \cos x)|_0^{\pi/2} = 2a.$$

$$a = \frac{1}{2}.$$

2.

$$E(x) = \frac{1}{2}\int_0^{\pi/2}\int_0^{\pi/2} x\sin(x+y)dydx = \frac{1}{2}\int_0^{\pi/2} xdx\int_0^{\pi/2}\sin(x+y)dy =$$

$$= -\frac{1}{2}\int_0^{\pi/2}\cos(x+y)|_0^{\pi/2}\,xdx = -\frac{1}{2}\int_0^{\pi/2}\left[\cos(x+\frac{\pi}{2})-\cos x\right]xdx =$$

$$= \frac{1}{2}\int_0^{\pi/2} x\,[\sin x + \cos x]\,dx = \frac{1}{2}x\,[\sin x - \cos x]\Big|_0^{\pi/2} - \frac{1}{2}\int_0^{\pi/2}[\sin x - \cos x]\,dx =$$

$$= \frac{\pi}{4} + \frac{1}{2}\,(\cos x + \sin x)\Big|_0^{\pi/2} = \frac{\pi}{4}.$$

Stejným způsobem dostaneme

$$E(y) = \frac{\pi}{4}.$$

3.

$$D(x) = E(x^2) - (E(x))^2 = \frac{1}{2}\int_0^{\pi/2}\int_0^{\pi/2} x^2\sin(x+y)dydx - \frac{\pi^2}{16} =$$

$$-\frac{1}{2}\int_0^{\pi/2} x^2\cos(x+y)\Big|_0^{\pi/2} dx - \frac{\pi^2}{16} = \frac{1}{2}\int_0^{\pi/2} x^2(\sin x + \cos x)dx - \frac{\pi^2}{16} =$$

$$\frac{1}{2}\,x^2(\sin x - \cos x)\Big|_0^{\pi/2} - \int_0^{\pi/2} x(\sin x - \cos x)dx - \frac{\pi^2}{16} =$$

$$\frac{\pi^2}{8} + x(\sin x + \cos x)\Big|_0^{\pi/2} - \int_0^{\pi/2}(\sin x + \cos x)dx - \frac{\pi^2}{16} =$$

$$= \frac{\pi^2}{8} + \frac{\pi}{2} + (\sin x - \cos x)\Big|_0^{\pi/2} - \frac{\pi^2}{16} = \frac{\pi^2}{16} + \frac{\pi}{2} - 2.$$

Stejným způsobem dostaneme $D(y) = D(x)$.

$$\sigma(x) = \sigma(y) = \sqrt{\frac{\pi^2}{16} + \frac{\pi}{2} - 2}.$$

4.

$$K(x,y) = E(xy) - E(x)E(y) = \frac{1}{2}\int_0^{\pi/2}\int_0^{\pi/2} xy\sin(x+y)dydx - \frac{\pi}{4}\frac{\pi}{4} =$$

$$\frac{1}{2}\int_0^{\pi/2} x\left(\int_0^{\pi/2} y\sin(x+y)dy\right)dx - \frac{\pi^2}{16} =$$

$$= \frac{8\pi - 16 - \pi^2}{16}.$$

$$R(x,y) = \frac{K(x,y)}{\sigma(x)\sigma(y)} = \frac{8\pi - 16 - \pi^2}{\pi^2 + 8\pi - 32} \approx -0.2454$$

□

1.21 Regresní koeficient a regresní přímka

Nechť $\mathbb{X}, \mathbb{Y}$ jsou náhodné veličiny. Nechť náhodná veličina $\mathbb{Z}$ je lineární funkcí náhodných veličin $\mathbb{Y}, \mathbb{X}$, tj. $\mathbb{Z} = \mathbb{Y} - k\mathbb{X}$. Hledáme takové k, aby $D(\mathbb{Z})$ bylo minimální.

$$D(\mathbb{Z}) = D(\mathbb{Y} - k\mathbb{X}) = E((\mathbb{Y} - k\mathbb{X})^2) - (E(\mathbb{Y} - k\mathbb{X}))^2 =$$

$$= E(\mathbb{Y}^2) - 2kE(\mathbb{Y}\mathbb{X}) + k^2E(\mathbb{X}^2) - (E(\mathbb{Y}))^2 + 2kE(\mathbb{X})E(\mathbb{Y}) - k^2(E(\mathbb{X}))^2 =$$
$$= D(\mathbb{Y}) + k^2D(\mathbb{X}) - 2kK(\mathbb{X},\mathbb{Y})$$
$$\frac{d}{dk}D(\mathbb{Z}) = 2kD(\mathbb{X}) - 2K(\mathbb{X},\mathbb{Y}) = 0$$
$$k = \frac{K(\mathbb{X},\mathbb{Y})}{D(\mathbb{X})} = R(\mathbb{X},\mathbb{Y})\frac{\sigma(\mathbb{Y})}{\sigma(\mathbb{X})}.$$

Definice 1.115. *Koeficientem regrese* náhodné veličiny nazýváme číslo

$$k = \frac{K(\mathbb{X},\mathbb{Y})}{D(\mathbb{X})} = R(\mathbb{X},\mathbb{Y})\frac{\sigma(\mathbb{Y})}{\sigma(\mathbb{X})}.$$

Důsledek 1.116. *Hodnota k nám dává minimum $D(\mathbb{Z})$.*

Důkaz. Protože

$$\frac{d}{dk}D(\mathbb{Z}) = 0$$

a

$$\frac{d^2}{dk^2}D(\mathbb{Z}) = 2D(\mathbb{X}) > 0$$

dává nám hodnota k minimum $D(\mathbb{Z})$. □

Definice 1.117. Přímku

$$y - E(\mathbb{Y}) = k(x - E(\mathbb{X}))$$

nazýváme *regresní přímkou* náhodné veličiny $\mathbb{Y}$ vzhledem k náhodné veličině $\mathbb{X}$.

Příklad 1.118. Nechť $(\mathbb{X},\mathbb{Y})$ je diskrétní náhodná veličina se simultání frekvenční funkcí

$x_i \backslash y_i$	0	1	2	$f(x_i)$
0	0.2	0.1	0.05	0.35
1	0	0.3	0	0.3
2	0	0.2	0.15	0.35
$f(y_j)$	0.2	0.6	0.2	1

Určete regresní přímku.

Řešení. Určíme postupně $E(\mathbb{X}), E(\mathbb{Y}), D(\mathbb{X}), D(\mathbb{Y})$.

$$E(\mathbb{X}) = 0 \cdot 0.35 + 1 \cdot 0.3 + 2 \cdot 0.35 = 1.$$

$$E(\mathbb{Y}) = 0 \cdot 0.2 + 1 \cdot 0.6 + 2 \cdot 0.2 = 1.$$
$$D(\mathbb{X}) = (0-1)^2 0.35 + (1-1)^2 0.3 + (2-1)^2 0.35 = 0.7.$$
$$D(\mathbb{Y}) = (0-1)^2 0.2 + (1-1)^2 0.6 + (2-1)^2 0.2 = 0.4.$$

Dále

$$E(\mathbb{X}\cdot\mathbb{Y}) = 0 \cdot (0.2 + 0.1 + 0.05) + 1 \cdot 03 + 2 \cdot 0.2 + 4 \cdot 0.15 = 1.3.$$

Kovariace

$$K(\mathbb{X},\mathbb{Y}) = 1.3 - 1 \cdot 1 = 0.3.$$

Koeficient korelace

$$R(\mathbb{X}, \mathbb{Y}) = \frac{0.3}{\sqrt{0.4}\sqrt{0.7}} = 0.567.$$

Regresní koeficient

$$k = 0.567\frac{\sqrt{0.4}}{\sqrt{0.7}} = 0.429.$$

Regresní přímka má potom tvar

$$y - 1 = 0.429(x - 1),$$

$$y = 0.429x + 0.571.$$

□

1.22 Nejužívanější rozložení diskrétních náhodných veličin

1. **Klasické rozložení**. Náhodná veličina $\mathbb{X}$ má klasické rozložení s parametrem $n \in \mathbb{N}$, jestliže

$$f(x) = p(\mathbb{X} = x) = \begin{cases} \frac{1}{n} & \text{pro } x = 1, 2, \ldots, n, \\ 0 & \text{pro zbývající případy.} \end{cases}$$

Věta 1.119. *Má-li náhodná veličina $\mathbb{X}$ klasické rozložení s parametrem $n \in \mathbb{N}$, potom platí:*

$$E(\mathbb{X}) = \frac{n+1}{2},$$

$$D(\mathbb{X}) = \frac{n^2 - 1}{12}.$$

2. **Binomické rozložení** $Bi(n, p)$. Náhodná veličina $\mathbb{X}$ má binomické rozložení s parametry $n \in \mathbb{N}$ a $p \in (0, 1)$, jestliže

$$f(i) = p(\mathbb{X} = i) = \begin{cases} \binom{n}{i} p^i (1-p)^{n-i} & \text{pro } i = 0, 1, 2, \ldots, n, \\ 0 & \text{pro zbývající případy.} \end{cases}$$

Věta 1.120. *Má-li náhodná veličina $\mathbb{X}$ binomické rozložení $Bi(n, p)$, potom platí:*

$$E(\mathbb{X}) = np,$$

$$D(\mathbb{X}) = np(1-p),$$

$$k_1 = \frac{1 - 2p}{\sqrt{np(1-p)}},$$

$$k_2 = \frac{1 - 6p(1-p)}{np(1-p)}.$$

Modus $\tilde{x}$ je určen nerovnicí $(n+1)p - 1 \le \tilde{x} \le (n+1)p$.

Důsledek 1.121. *V případě, že $(n+1)p$ je celé číslo, budeme mít pro modus dvě hodnoty.*

Důsledek 1.122. *$Bi(n, \frac{1}{2})$ má symetrickou frekvenční funkci.*

3. **Alternativní rozložení** $A(p)$. Náhodná veličina $\mathbb{X}$ má alternativní rozložení s parametrem $p \in (0,1)$, jestliže
$$f(1) = p, \quad f(0) = 1 - p.$$

Věta 1.123. *Má-li náhodná veličina $\mathbb{X}$ alternativní rozložení $A(p)$, potom platí:*
$$\begin{aligned} E(\mathbb{X}) &= p, \\ D(\mathbb{X}) &= p(1-p), \\ k_1 &= \frac{1-2p}{\sqrt{p(1-p)}}, \\ k_2 &= \frac{1-6p(1-p)}{p(1-p)}, \\ A(p) &\equiv Bi(1,p). \end{aligned}$$

4. **Poissonovo rozložení** $Po(\lambda)$. Náhodná veličina $\mathbb{X}$ má Poissonovo rozložení s parametrem $\lambda \in \mathbb{R}^+$, jestliže
$$f(i) = p(\mathbb{X} = i) = \begin{cases} \frac{\lambda^i}{i!} \mathrm{e}^{-\lambda} & \text{pro } i = 0, 1, 2, \dots, \\ 0 & \text{pro zbývající případy.} \end{cases}$$

Věta 1.124. *Má-li náhodná veličina $\mathbb{X}$ Poissonovo rozložení $Po(\lambda)$, potom platí:*
$$\begin{aligned} E(\mathbb{X}) &= \lambda, \\ D(\mathbb{X}) &= \lambda, \\ k_1 &= \lambda^{-\frac{1}{2}}, \\ k_2 &= \lambda^{-1}. \end{aligned}$$

Důsledek 1.125. *$E(\mathbb{X}) = D(\mathbb{X}) = \lambda$ je charakteristickou vlastností Poissonova rozložení.*

Příklad 1.126. Určete střední hodnotu diskrétní náhodné veličinz s frekvenční funkci
$$p(\mathbb{X} = m) = \frac{\lambda^m e^{-\lambda}}{m!}, \quad m = 0, 1, 2, \dots.$$

Řešení. Jde o náhodnou veličinu s Poissonověm rozložením.
$$E(\mathbb{X}) = \sum_{m=0}^{\infty} m \frac{\lambda^m e^{-\lambda}}{m!} = \sum_{m=1}^{\infty} m \frac{\lambda^m e^{-\lambda}}{m!} = \lambda e^{-\lambda} \sum_{m=1}^{\infty} \frac{\lambda^{m-1}}{(m-1)!}.$$
Protože
$$\sum_{m=1}^{\infty} \frac{\lambda^{m-1}}{(m-1)!} = e^{\lambda},$$
po dosazení dostaneme
$$E(\mathbb{X}) = \lambda.$$

□

5. **Negativní binomické rozložení** $Nbi(n,p)$. Kolik pokusů při Bernoulliovské posloupnosti nezávislých pokusů je třeba udělat, aby nastal jev A po n-té? Jestliže náhodná veličina $\mathbb{X}$ nabývá hodnoty počtu pokusů, při nichž jev A nenastal předtím než nastal po n-té a $\mathbb{X} = x_i$, pak jev A nastal po n-té v $(x_i + n)$-tém pokusu.

   Náhodná veličina $\mathbb{X}$ má negativní binomické rozložení s parametry $n \in \mathbb{N}$ a $p \in (0,1)$, jestliže

$$f(x_i) = p(\mathbb{X} = x_i) = \begin{cases} \binom{x_i+n-1}{n-1} p^n (1-p)^{x_i} & \text{pro } x_i = 0,1,2,\dots, n = 1,2,\dots, \\ 0 & \text{pro zbývající případy.} \end{cases}$$

   Věta 1.127. *Má-li náhodná veličina $\mathbb{X}$ negativní binomické rozložení $Nbi(n,p)$, potom platí:*

$$E(\mathbb{X}) = n\frac{1-p}{p},$$
$$D(\mathbb{X}) = n\frac{1-p}{p^2}.$$

6. **Geometrické rozložení** $Ge(p)$. Jde o rozložení $Nbi(1,p)$. Počet nezávislých pokusů, které končí při prvním úspěšném pokusu. Náhodná veličina $\mathbb{X}$ má geometrické rozložení s parametrem $p \in (0,1)$, jestliže

$$f(x_i) = p(\mathbb{X} = x_i) = \begin{cases} p(1-p)^{x_i} & \text{pro } x_i = 0,1,2,\dots, \\ 0 & \text{pro zbývající případy.} \end{cases}$$

   Věta 1.128. *Má-li náhodná veličina $\mathbb{X}$ geometrické rozložení $Ge(p)$, potom platí:*

$$E(\mathbb{X}) = \frac{1-p}{p},$$
$$D(\mathbb{X}) = \frac{1-p}{p^2}.$$

7. **Hypergeometrické rozložení** $H(N,M,n)$. Počet prvků vykazujících sledovanou vlastnost v souboru n prvků, vybraných bez vracení ze souboru N prvků, z nichž celkem M má dannou vlatnost.

   Náhodná veličina $\mathbb{X}$ má hypergeometrické rozložení s parametry $N, M, n \in \mathbb{N}, N > M$, jestliže

$$f(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}},$$

   Věta 1.129. *Má-li náhodná veličina $\mathbb{X}$ hypergeometrické rozložení $H(N,M,n)$, potom platí:*

$$E(\mathbb{X}) = n\frac{M}{N},$$
$$D(\mathbb{X}) = n\frac{M}{N}\left(1 - \frac{M}{n}\right)\frac{N-n}{N-1}.$$

1.23 Nejužívanější rozložení spojitých náhodných veličin

1. **Rovnoměrné rozložemní**. Náhodná veličina $\mathbb{X}$ má rovnoměrné rozložení, jestliže má konstantní hustotu pravděpodobnosti na celém intervalu hodnot, kterých může nabýt. Funkce hustoty je

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{pro } x \in [a,b], \\ 0 & \text{pro } x \notin [a,b]. \end{cases}$$

Distribuční funkce je

$$F(x) = \begin{cases} 0 & x < a, \\ \frac{x-a}{b-a} & \text{pro } x \in [a,b], \\ 1 & \text{pro } x > b. \end{cases}$$

Věta 1.130. *Má-li náhodná veličina $\mathbb{X}$ rovnoměrné rozložení, potom platí:*

$$E(\mathbb{X}) = \frac{1}{2}(a+b),$$
$$D(\mathbb{X}) = \frac{1}{12}(b-a)^2,$$
$$k_1 = 0,$$
$$k_2 = \frac{(b-a)^4}{80} - 3.$$

2. **Normální rozložení** $No(\mu, \sigma)$. Náhodná veličina $\mathbb{X}$ má normální rozložení s parametry $\mu \in \mathbb{R}, \sigma > 0$, jestliže má funkci hustoty

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

a distribuční funkci

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

Věta 1.131. *Má-li náhodná veličina $\mathbb{X}$ normální rozložení $No(\mu, \sigma)$, potom platí:*

$$E(\mathbb{X}) = \mu,$$
$$D(\mathbb{X}) = \sigma^2.$$

Důkaz:

$$E(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

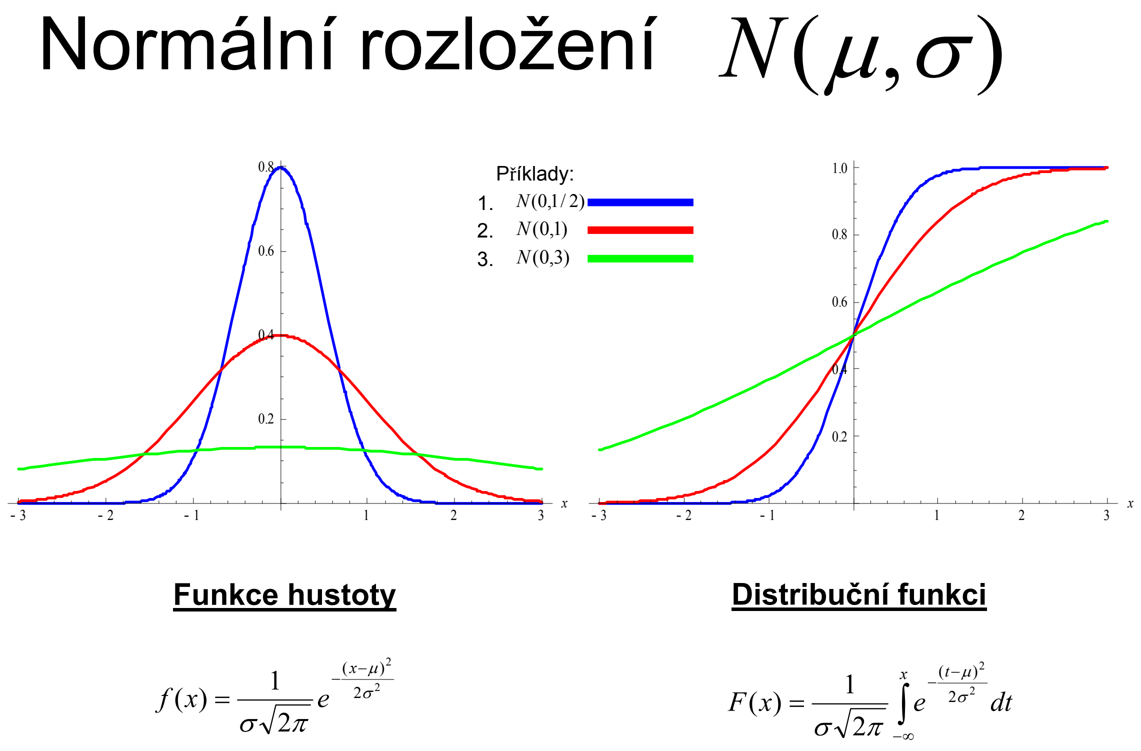
Zavedeme si substituci

$$\frac{x-\mu}{\sigma} = t \Rightarrow x = \sigma t + \mu \Rightarrow dx = \sigma dt.$$

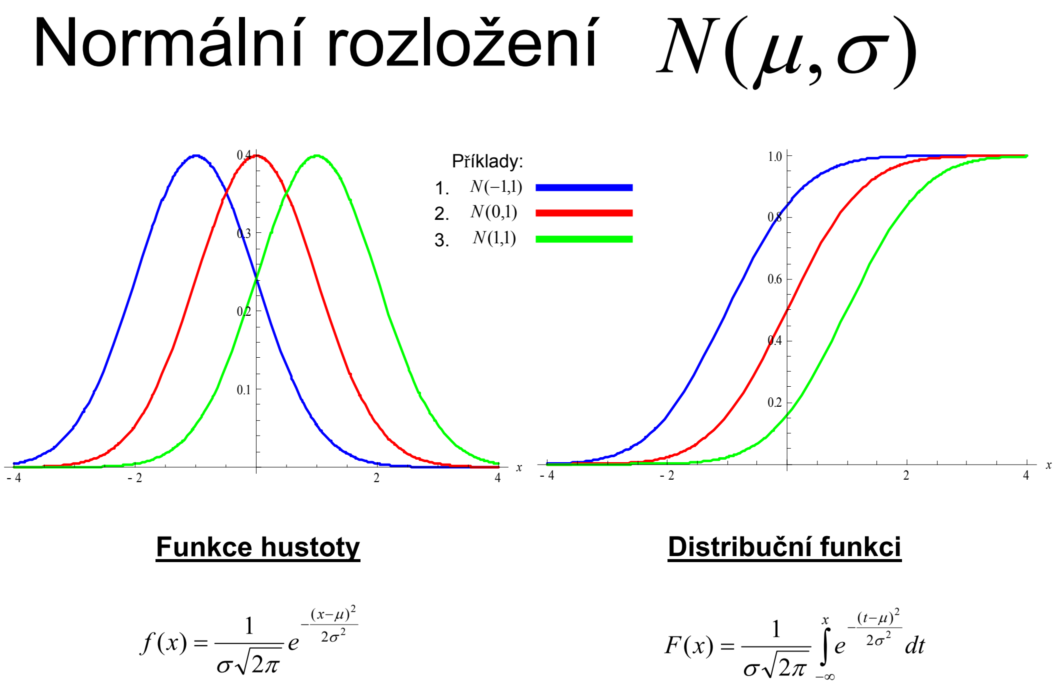
Po dosazení máme

$$E(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma t + \mu)\mathrm{e}^{-\frac{t^2}{2}} dt = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t\mathrm{e}^{-\frac{t^2}{2}} dt + \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{t^2}{2}} dt = \mu.$$

Protože první integrál je integrálem z liché funkce, který je roven nule, a druhý integrál je roven $\sqrt{2\pi}$.



Obr. 1.1: Normální rozložení 1



Obr. 1.2: Normální rozložení 2

3. **Exponenciální rozložení** $E(A,d)$. Náhodná veličina $\mathbb{X}$ má exponenciální rozložení s parametry $A \in \mathbb{R}, d > 0$, jestliže má funkci hustoty

$$f(x) = \begin{cases} 0 & x \leq A, \\ \frac{1}{d}\mathrm{e}^{-\frac{x-A}{d}} & x > A, \end{cases}$$

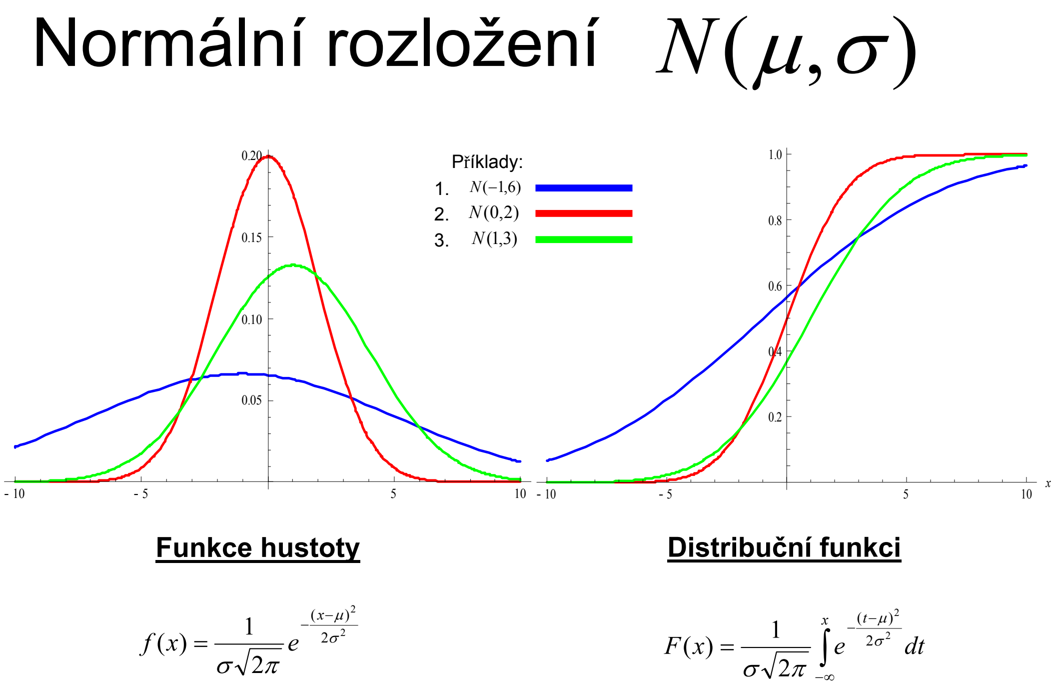
a distribuční funkci

$$F(x) = \begin{cases} 0 & x \leq A, \\ 1-\mathrm{e}^{-\frac{x-A}{d}} & x > A. \end{cases}$$

Věta 1.132. *Má-li náhodná veličina* $\mathbb{X}$ *exponenciální rozložení* $E(A,d)$*, potom platí:*

$$E(\mathbb{X}) = A + d,$$

$$D(\mathbb{X}) = d^2.$$



Obr. 1.3: Normální rozložení 3

4. **Gama rozložení** $\Gamma(m,d)$. Náhodná veličina $\mathbb{X}$ má gama rozložení s parametry $m>0, d>$ >0, jestliže má funkci hustoty

$$f(x)=\begin{cases} 0 & x\le 0,\\ \frac{1}{\Gamma(m)d^m}\mathrm{e}^{-\frac{x}{d}}x^{m-1} & x>0,\end{cases}$$

kde

$$\Gamma(m)=\int_0^m \mathrm{e}^{-t}t^{m-1}dt.$$

Věta 1.133. *Má-li náhodná veličina $\mathbb{X}$ gama rozložení $\Gamma(m,d)$, potom platí:*

$$E(\mathbb{X})=md,$$

$$D(\mathbb{X})=md^2.$$

Důsledek 1.134. *Pro $m=1$ je $\Gamma(1,d)\equiv E(0,d)$.*

5. **Beta rozložení** $Be(p,q)$. Náhodná veličina $\mathbb{X}$ má beta rozložení s parametry $p>0, q>0$, jestliže má funkci hustoty

$$f(x)=\begin{cases} 0 & x\notin(0,1),\\ \frac{1}{B(p,q)}x^{p-1}(1-x)^{q-1} & x\in(0,1),\end{cases}$$

kde

$$B(p,q)=\int_0^1 x^{p-1}(1-x)^{q-1}dx.$$

Věta 1.135. *Má-li náhodná veličina $\mathbb{X}$ beta rozložení $Be(p,q)$, potom platí:*

$$E(\mathbb{X})=\frac{p}{p+q},$$

$$D(\mathbb{X})=\frac{pq}{(p+q)^2(p+q+1)},$$

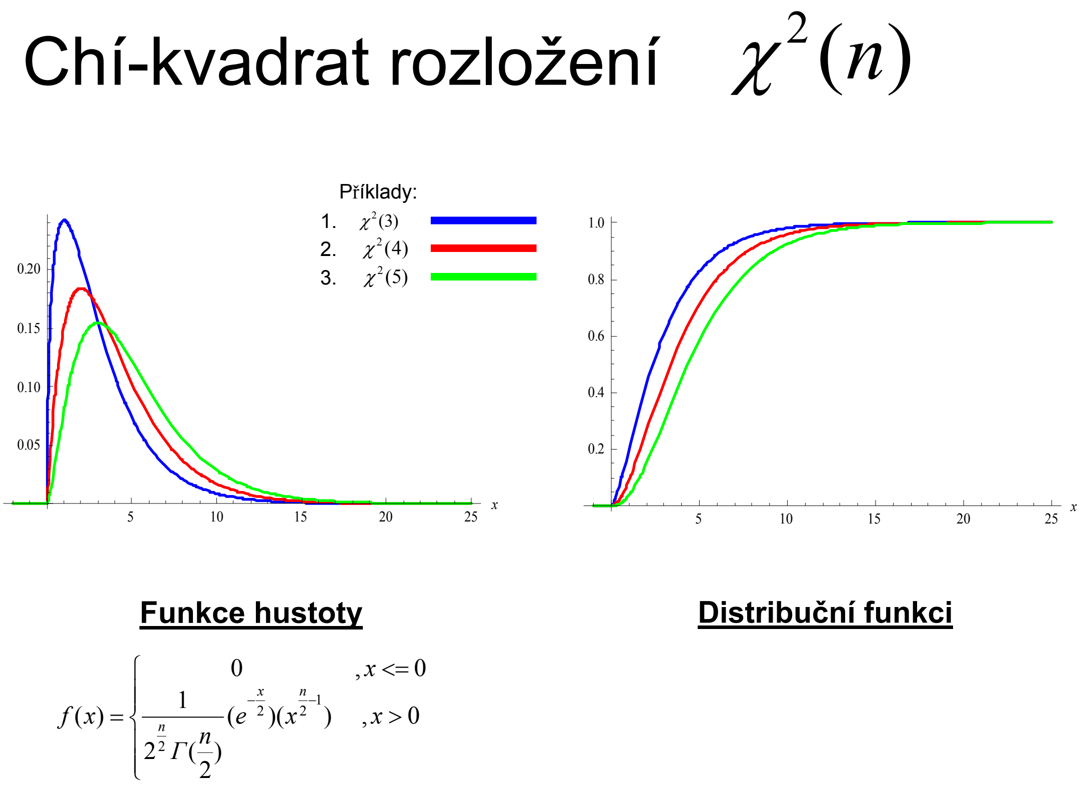
$$B(p,q)=\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}.$$

6. **Pearsonovo rozložení** χ^2.(čti chí-kvadrát) Náhodná veličina $\mathbb{X}$ má rozložení χ^2 s k stupni volnosti, jestliže má funkci hustoty

$$f(x)=\begin{cases} 0 & x\le 0,\\ \frac{1}{2^{k/2}\Gamma(\frac{k}{2})}\left(\mathrm{e}^{-\frac{x}{2}}\right)\left(x^{\frac{k}{2}-1}\right) & x>0.\end{cases}$$

Věta 1.136. *Má-li náhodná veličina $\mathbb{X}$ rozložení χ^2 s k-stupni volnosti, potom platí:*

$$E(\mathbb{X})=k,$$



Obr. 1.4: Pearsonovo rozložení

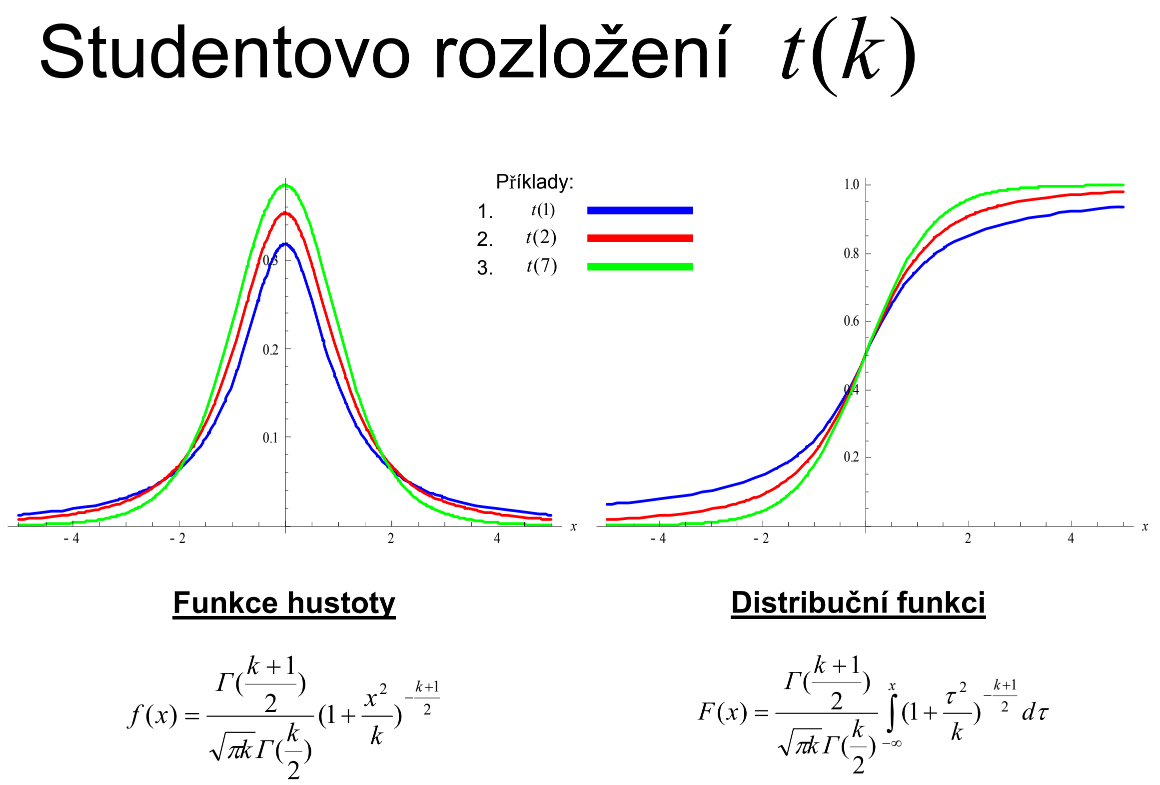
$$D(\mathbb{X}) = 2k,$$
$$k_1 = \frac{4}{\sqrt{2k}},$$
$$k_2 = \frac{12}{k},$$
$$\Gamma(\frac{k}{2}, 2) \equiv \chi^2(k).$$

7. **Studentovo rozložení** t. Mějme dvě nezávislé náhodné veličiny $\mathbb{U}, \mathbb{Y}$. Nechť má $\mathbb{U}$ rozložení $No(0,1)$ a $\mathbb{Y}$ má rozložení χ^2. Utvořme náhodnou veličinu

$$\mathbb{T} = \frac{\mathbb{U}}{\sqrt{\frac{\mathbb{Y}}{k}}}.$$

Náhodná veličina $\mathbb{T}$ má rozložení t o k stupních volnosti s funkcí hustoty pro $k \in \mathbb{N}, x \in \mathbb{R}$

$$f(x) = \frac{1}{\sqrt{k}B(\frac{1}{2}, \frac{k}{2})}\left(1 + \frac{x^2}{k}\right)^{-\frac{k+1}{2}} =$$



Obr. 1.5: Studentovo rozložení

$$= \frac{\Gamma(\frac{k+1}{2})}{\sqrt{\pi k}\Gamma(\frac{k}{2})}\left(1+\frac{x^2}{k}\right)^{-\frac{k+1}{2}}.$$

Pro $k \to +\infty$ konverguje t-rozložení k rozložení $No(0,1)$.

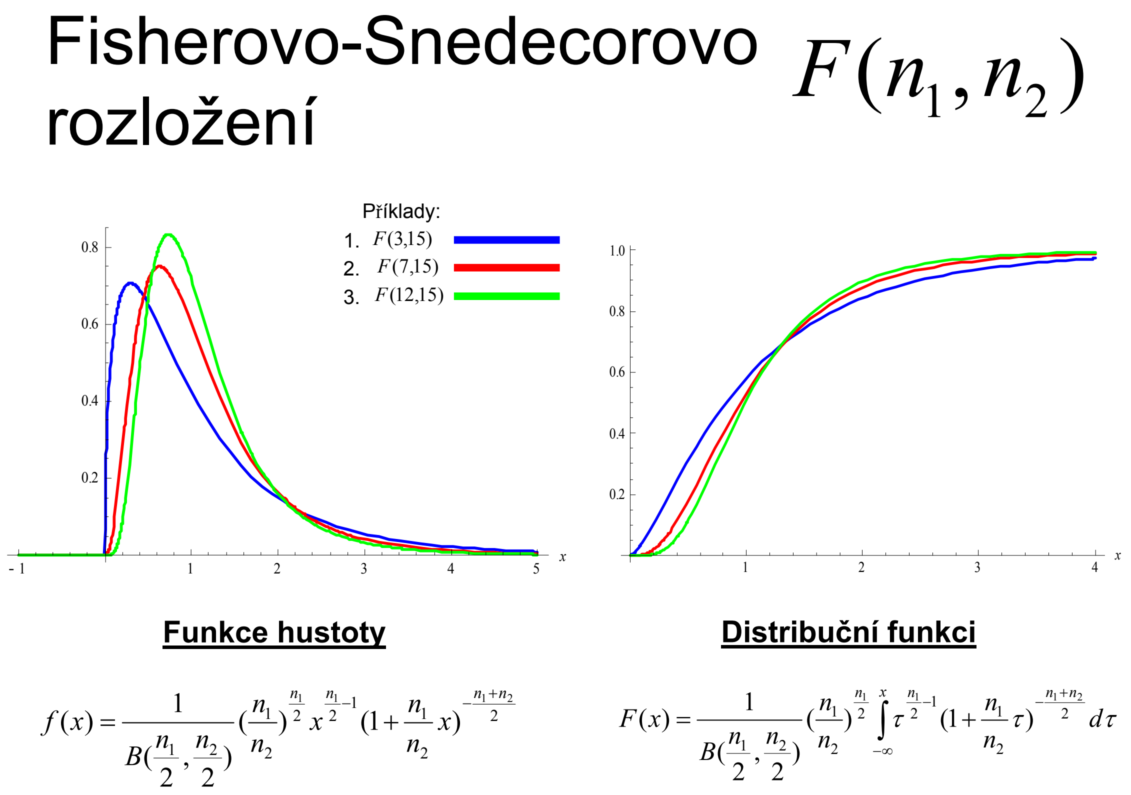
Věta 1.137.

$$k_1 = 0,$$

$$k_2 = \frac{6}{k-4}.$$

8. **Fisherovo - Snedecorovo rozložení** F. Mějme dvě nezávislé náhodné veličiny $\mathbb{X}, \mathbb{Y}$. Nechť má $\mathbb{X}$ rozložení $\chi^2(n_1)$ a $\mathbb{Y}$ má rozložení $\chi^2(n_2)$. Utvořme náhodnou veličinu

$$\mathbb{F} = \frac{\frac{\mathbb{X}}{n_1}}{\frac{\mathbb{Y}}{n_2}}.$$



Obr. 1.6: Fisherovo - Snedecorovo rozložení 1

Náhodná veličina $\mathbb{F}$ má rozložení $F(n_1, n_2)$ o n_1 a n_2 stupních volnosti s funkcí hustoty pro $n_1, n_2 \in \mathbb{N}, x > 0$

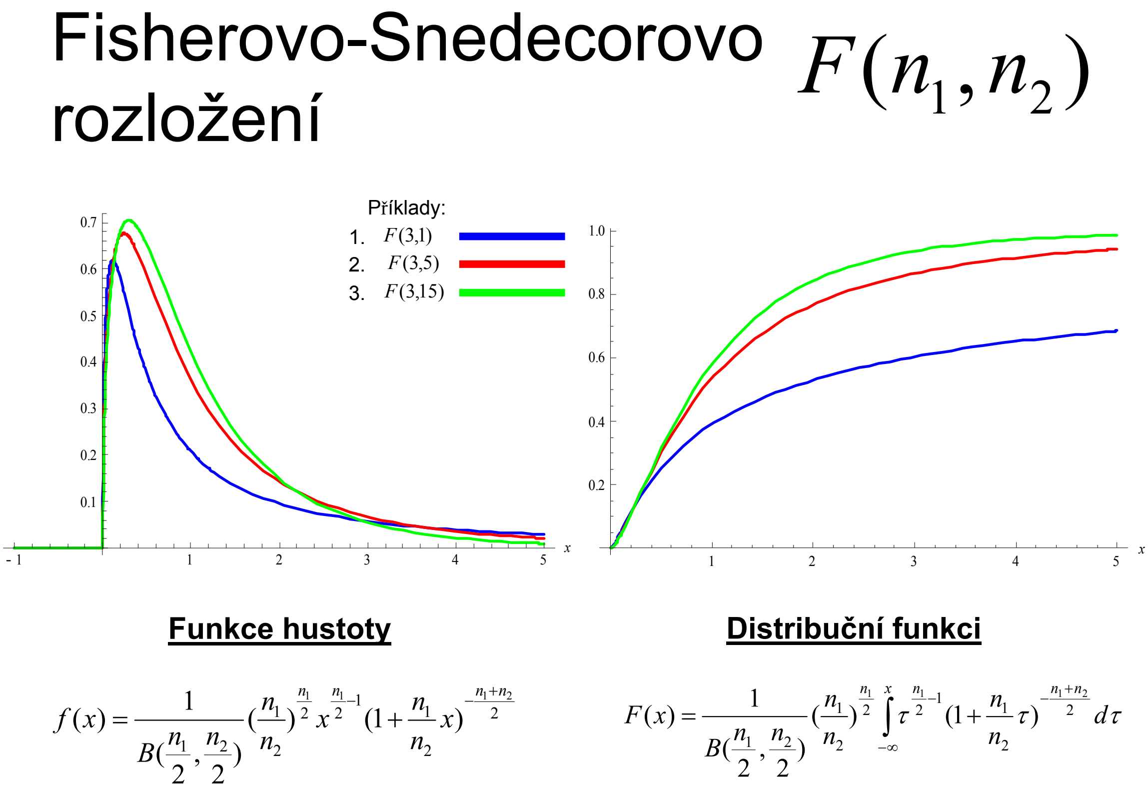
$$f(x) = \frac{1}{B(\frac{n_1}{2}, \frac{n_2}{2})} \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} x^{\frac{n_1}{2}-1} (1 + \frac{n_1}{n_2} x)^{-\frac{n_1+n_2}{2}},$$

$$f(x) = 0, \qquad \text{pro} \qquad x < 0.$$

9. **Weibullovo rozložení** $W(\delta, b, k)$. Náhodná veličina $\mathbb{X}$ má rozložení $W(\delta, b, k)$ o $\delta >$ $> 0, k > 0, b \in \mathbb{R}$, jestliže má funkci hustoty

$$f(x) = \begin{cases} \frac{k}{\delta}(x-b)^{k-1} \exp\left(-\frac{(x-b)^k}{\delta}\right) & \text{pro } x \geq b \\ 0 & \text{pro } x < b. \end{cases}$$

Důsledek 1.138. $W(\delta, b, 1) \equiv E(b, \delta)$.



Obr. 1.7: Fisherovo - Snedecorovo rozložení 2

Weibulovo rozložení pro $k > 1$ modeluje životnost zařízení podléhajícího opotřebení a nebo únavě materiálu.

Weibulovo rozložení pro $k < 1$ modeluje životnost zařízení, kde dochází k poruchám v důsledku skrytých vad, nikoliv opotřebením.

1.24 Vlastnosti normálního rozložení

Náhodná veličina $\mathbb{X}$ má normální rozložení s parametry $\mu \in \mathbb{R}, \sigma > 0$, jestliže má funkci hustoty

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\dfrac{(x-\mu)^2}{2\sigma^2}}$$

a distribuční funkci

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-\dfrac{(t-\mu)^2}{2\sigma^2}} dt.$$

Je to nejdůležitější rozdělení spojité náhodné veličiny. Je použitelné všude tam, kde kolísání náhodné veličiny je způsobeno součtem velkho počtu nepatrných a vzájemně nezávislých jevů a vlivů.

Modus $\tilde{x} = \mu$, $f(\mu) = (\sigma\sqrt{2\pi})^{-1}$.

Inflexní body $x = \mu \pm \sigma$.

Věta 1.139. *Pro náhodnou veličinu $\mathbb{X}$ s normálním rozložením $No(\mu, \sigma)$ platí*

$$E(\mathbb{X}) = \mu, \quad D(\mathbb{X}) = \sigma^2.$$

Věta 1.140. *Je-li $\mathbb{X}$ náhodná veličina s normálním rozložením $No(\mu, \sigma)$, potom normovaná náhodná veličina $\mathbb{U}$ k náhodné veličině $\mathbb{X}$ má normální rozložení $No(0, 1)$.*

Distribuční funkce náhodné veličiny $\mathbb{X}$ s normálním rozložením $No(\mu, \sigma)$ je tvaru

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

Tento integrál nelze vyjádřit pomocí elementárních funkcí.

Definice 1.141. *Laplaceovou funkcí* nazveme

$$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} \mathrm{e}^{-\frac{t^2}{2}} dt.$$

Laplaceova funkce je distibuční funkce normálního normovaného rozložení.

Hodnoty funkce $\Phi(u)$ jsou uvedeny v tabulce 1.1 a 1.2.

Věta 1.142. *Nechť $F(x)$ a $\Phi(u)$ jsou distribuční funkce normálního rozložení a normovaného normálního rozložení. Potom*

$$F(x) \equiv \Phi\left(\frac{x-\mu}{\sigma}\right).$$

Důsledek 1.143. *Buď $\mathbb{X}$ náhodná veličina s normálním rozložením $No(\mu, \sigma)$. Potom pro libovolné $x_1, x_2 \in \mathbb{R}, x_1 < x_2$ platí*

$$p(x_1 \leq \mathbb{X} \leq x_2) = \Phi\left(\frac{x_2 - \mu}{\sigma}\right) - \Phi\left(\frac{x_1 - \mu}{\sigma}\right).$$

Věta 1.144. *Pro Laplaceovu funkci platí*

$$\Phi(0) = 0.5,$$

$$\Phi(-t) = 1 - \Phi(t), \quad \forall t \in \mathbb{R}.$$

Věta 1.145. Zákon tří sigma
Pro náhodnou veličinu s rozložením $No(\mu, \sigma)$ platí,

$$P(\mu - 3\sigma < \mathbb{X} < \mu + 3\sigma) > 99.7\%.$$

Důkaz. Náhodná veličina má rozložení $No(\mu, \sigma)$, potom

$$P(\mu - 3\sigma < \mathbb{X} < \mu + 3\sigma) = \Phi\left(\frac{\mu + 3\sigma - \mu}{\sigma}\right) - \Phi\left(\frac{\mu - 3\sigma - \mu}{\sigma}\right) =$$

$$= \Phi\left(\frac{3\sigma}{\sigma}\right) - \Phi\left(\frac{-3\sigma}{\sigma}\right) = \Phi(3) - \Phi(-3) = 2\Phi(3) - 1 = 0.997...$$

□

Tab. 1.1: **Hodnoty Laplaceovy funkce $\Phi(u)$ I.**

u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$
0,00	0,5000000	0,30	0,6179114	0,60	0,7257469	0,90	0,8159399	1,20	0,8849303
0,01	0,5039894	0,31	0,6217195	0,61	0,7290691	0,91	0,8185887	1,21	0,8868606
0,02	0,5079783	0,32	0,6255158	0,62	0,7323711	0,92	0,8212136	1,22	0,8887676
0,03	0,5119665	0,33	0,6293000	0,63	0,7356527	0,93	0,8238145	1,23	0,8906514
0,04	0,5159534	0,34	0,6330717	0,64	0,7389137	0,94	0,8263912	1,24	0,8925123
0,05	0,5199388	0,35	0,6368307	0,65	0,7421539	0,95	0,8289439	1,25	0,8943502
0,06	0,5239222	0,36	0,6405764	0,66	0,7453731	0,96	0,8314724	1,26	0,8961653
0,07	0,5279032	0,37	0,6443088	0,67	0,7485711	0,97	0,8339768	1,27	0,8979577
0,08	0,5318814	0,38	0,6480273	0,68	0,7517478	0,98	0,8364569	1,28	0,8997274
0,09	0,5358564	0,39	0,6517317	0,69	0,7549029	0,99	0,8389129	1,29	0,9014747
0,10	0,5398278	0,40	0,6554217	0,70	0,7580363	1,00	0,8413447	1,30	0,9031995
0,11	0,5437953	0,41	0,6590970	0,71	0,7611479	1,01	0,8437524	1,31	0,9049021
0,12	0,5477584	0,42	0,6627573	0,72	0,7642375	1,02	0,8461358	1,32	0,9065825
0,13	0,5517168	0,43	0,6664022	0,73	0,7673049	1,03	0,8484950	1,33	0,9082409
0,14	0,5556700	0,44	0,6700314	0,74	0,7703500	1,04	0,8508300	1,34	0,9098773
0,15	0,5596177	0,45	0,6736448	0,75	0,7733726	1,05	0,8531409	1,35	0,9114920
0,16	0,5635595	0,46	0,6772419	0,76	0,7763727	1,06	0,8554277	1,36	0,9130850
0,17	0,5674949	0,47	0,6808225	0,77	0,7793501	1,07	0,8576903	1,37	0,9146565
0,18	0,5714237	0,48	0,6843863	0,78	0,7823046	1,08	0,8599289	1,38	0,9162067
0,19	0,5753454	0,49	0,6879331	0,79	0,7852361	1,09	0,8621434	1,39	0,9177356
0,20	0,5792597	0,50	0,6914625	0,80	0,7881446	1,10	0,8643339	1,40	0,9192433
0,21	0,5831662	0,51	0,6949743	0,81	0,7910299	1,11	0,8665005	1,41	0,9207302
0,22	0,5870604	0,52	0,6984682	0,82	0,7938919	1,12	0,8686431	1,42	0,9221962
0,23	0,5909541	0,53	0,7019440	0,83	0,7967306	1,13	0,8707619	1,43	0,9236415
0,24	0,5948349	0,54	0,7054015	0,84	0,7995458	1,14	0,8728568	1,44	0,9250663
0,25	0,5987063	0,55	0,7088403	0,85	0,8023375	1,15	0,8749281	1,45	0,9264707
0,26	0,6025681	0,56	0,7122603	0,86	0,8051055	1,16	0,8769756	1,46	0,9278550
0,27	0,6064199	0,57	0,7156612	0,87	0,8078498	1,17	0,8789995	1,47	0,9292191
0,28	0,6102612	0,58	0,7190427	0,88	0,8105703	1,18	0,8809999	1,48	0,9305634
0,29	0,6140919	0,59	0,7224047	0,89	0,8132671	1,19	0,8829768	1,49	0,9318879

Tab. 1.2: **Hodnoty Laplaceovy funkce $\Phi(u)$ II.**

u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$
1,50	0,9331928	1,80	0,9640697	2,10	0,9821356	2,40	0,9918025	4,50	0,9999966
1,51	0,9344783	1,81	0,9648521	2,11	0,9825708	2,41	0,9920237	5,00	0,9999997
1,52	0,9357445	1,82	0,9656205	2,12	0,9829970	2,42	0,9922397	5,50	0,9999999
1,53	0,9369916	1,83	0,9663750	2,13	0,9834142	2,43	0,9924506		
1,54	0,9382198	1,84	0,9671159	2,14	0,9838226	2,44	0,9926564		
1,55	0,9394392	1,85	0,9678432	2,15	0,9842224	2,45	0,9928572		
1,56	0,9406201	1,86	0,9685572	2,16	0,9846137	2,46	0,9930531		
1,57	0,9417924	1,87	0,9692581	2,17	0,9849966	2,47	0,9932443		
1,58	0,9429466	1,88	0,9699460	2,18	0,9853713	2,48	0,9934309		
1,59	0,9440826	1,89	0,9706210	2,19	0,9857379	2,49	0,9936128		
1,60	0,9452007	1,90	0,9712834	2,20	0,9860966	2,50	0,9937903		
1,61	0,9463011	1,91	0,9719334	2,21	0,9864474	2,51	0,9939634		
1,62	0,9473839	1,92	0,9725711	2,22	0,9867906	2,52	0,9941323		
1,63	0,9484493	1,93	0,9731966	2,23	0,9871263	2,53	0,9942969		
1,64	0,9494974	1,94	0,9738102	2,24	0,9874545	2,54	0,9944574		
1,65	0,9505285	1,95	0,9744119	2,25	0,9877755	2,55	0,9946139		
1,66	0,9515428	1,96	0,9750021	2,26	0,9880894	2,56	0,9947664		
1,67	0,9525403	1,97	0,9755808	2,27	0,9883962	2,57	0,9949151		
1,68	0,9535213	1,98	0,9761482	2,28	0,9886962	2,58	0,9950600		
1,69	0,9544860	1,99	0,9767045	2,29	0,9889893	2,59	0,9952012		
1,70	0,9554345	2,00	0,9772499	2,30	0,9892759	2,60	0,9953388		
1,71	0,9563671	2,01	0,9777844	2,31	0,9895559	2,70	0,9965330		
1,72	0,9572838	2,02	0,9783083	2,32	0,9898296	2,80	0,9974449		
1,73	0,9581849	2,03	0,9788217	2,33	0,9900969	2,90	0,9981342		
1,74	0,9590705	2,04	0,9793248	2,34	0,9903581	3,00	0,9986501		
1,75	0,9599408	2,05	0,9798178	2,35	0,9906133	3,20	0,9993129		
1,76	0,9607961	2,06	0,9803007	2,36	0,9908625	3,40	0,9996631		
1,77	0,9616364	2,07	0,9807738	2,37	0,9911060	3,60	0,9998409		
1,78	0,9624620	2,08	0,9812372	2,38	0,9913437	3,80	0,9999277		
1,79	0,9632730	2,09	0,9816911	2,39	0,9915758	4,00	0,9999683		

1.25 Limitní věty

Věta 1.146. První Čebyševova nerovnost
Nechť $\mathbb{X}$ je náhodná veličina, která nabývá pouze nezáporných hodnot. Potom

$$p(\mathbb{X} \geq 1) \leq E(\mathbb{X}).$$

Důkaz. Nechť $\mathbb{X}$ je diskrétní náhodná veličina s frekvenční funkcí $f(x)$. Potom pro nezáporné x_i platí

$$p(\mathbb{X} > 1) = \sum_{x_i \geq 1} f(x_i) \leq \sum_{x_i \geq 1} x_i f(x_i) \leq \sum_{x_i \geq 0} x_i f(x_i) = E(\mathbb{X}).$$

□

Věta 1.147. Druhá Čebyševova nerovnost
Pro každou náhodnou veličinu $\mathbb{X}$ a pro každé $\varepsilon > 0$ platí

$$p\left(|\mathbb{X} - E(\mathbb{X})| < \varepsilon\right) \geq 1 - \frac{D(\mathbb{X})}{\varepsilon^2}$$

Důkaz.

$$p(|\mathbb{X} - E(\mathbb{X})| \geq \varepsilon) = 1 - p(|\mathbb{X} - E(\mathbb{X})| < \varepsilon)$$
$$p(|\mathbb{X} - E(\mathbb{X})| \geq \varepsilon) = p\left(\frac{|\mathbb{X} - E(\mathbb{X})|}{\varepsilon} \geq 1\right) = p\left(\frac{|\mathbb{X} - E(\mathbb{X})|^2}{\varepsilon^2} \geq 1\right).$$

Podle první Čebyševovy nerovnosti pro náhodnou veličinu s nezápornými hodnotami platí

$$p\left(\frac{|\mathbb{X} - E(\mathbb{X})|^2}{\varepsilon^2} \geq 1\right) \leq E\left(\frac{|\mathbb{X} - E(\mathbb{X})|^2}{\varepsilon^2}\right) = \frac{D(\mathbb{X})}{\varepsilon^2}.$$

Je tedy

$$1 - p(|\mathbb{X} - E(\mathbb{X})| < \varepsilon) \leq \frac{D(\mathbb{X})}{\varepsilon^2}.$$

Po úpravě dostaneme tvrzení věty. □

Příklad 1.148. Pomocí druhé Čebyševovy nerovnosti odhadněte pravděpodobnost

$$p(|\mathbb{X} - E(\mathbb{X})| < 0.1)$$

pro náhodnou veličinu $\mathbb{X}$ s rozložením $E(3, \frac{1}{20})$.

Řešení. Pro náhodnou veličinu s exponenciálním rozložením $E(A, d)$ platí

$$E(\mathbb{X}) = A + d \Rightarrow E(\mathbb{X}) = 3 + \frac{1}{20} = 3.05,$$

$$D(\mathbb{X}) = d^2 \Rightarrow D(\mathbb{X}) = 0.0025.$$

Dosadíme

$$p(|\mathbb{X} - E(\mathbb{X})| < \varepsilon) = p(|\mathbb{X} - 3.05| < 0.1) \geq 1 - \frac{0.0025}{0.01} = 1 - 0.25 = 0.75.$$

□

Věta 1.149. Bernouliova věta, Zákon velkých čísel
Nechť náhodná veličina $\mathbb{X}$ má rozložení $Bi(n,p)$, potom pro libovolné $\varepsilon > 0$ platí

$$p\left(\left|\frac{\mathbb{X}}{n} - p\right| < \varepsilon\right) \geq 1 - \frac{p(1-p)}{n\varepsilon^2}.$$

Důkaz. Podle druhé Čebyševovy nerovnosti platí

$$p\left(\left|\frac{\mathbb{X}}{n} - E\left(\frac{\mathbb{X}}{n}\right)\right| < \varepsilon\right) \geq 1 - \frac{D\left(\frac{\mathbb{X}}{n}\right)}{\varepsilon^2}.$$

Náhodná veličina $\mathbb{X}$ má binomické rozložení $Bi(n,p)$ a proto

$$E(\mathbb{X}) = np, \quad D(\mathbb{X}) = np(1-p).$$

Podle věty 1.92 máme

$$E\left(\frac{\mathbb{X}}{n}\right) = E\left(\frac{1}{n}\mathbb{X}\right) = \frac{1}{n}E(\mathbb{X}) = \frac{1}{n}np = p,$$

$$D\left(\frac{\mathbb{X}}{n}\right) = \frac{1}{n^2}D(\mathbb{X}) = \frac{1}{n^2}np(1-p) = \frac{p(1-p)}{n}.$$

Po dosazení dostaneme

$$p\left(\left|\frac{\mathbb{X}}{n} - E\left(\frac{\mathbb{X}}{n}\right)\right| < \varepsilon\right) = \left(\left|\frac{\mathbb{X}}{n} - p\right| < \varepsilon\right) \geq 1 - \frac{D\left(\frac{\mathbb{X}}{n}\right)}{\varepsilon^2} = 1 - \frac{p(1-p)}{n\varepsilon^2}.$$

□

Důsledek 1.150. *Nechť náhodná veličina $\mathbb{X}$ má rozložení $Bi(n,p)$, potom pro libovolné $\varepsilon > 0$ platí*

$$\lim_{n\to+\infty} p\left(\left|\frac{\mathbb{X}}{n} - p\right| < \varepsilon\right) = 1.$$

Příklad 1.151. Mějme náhodnou veličinu s binomickým rozložením $Bi(1000; 0.514)$. Určete:
a) Pravděpodobnost, že relativní četnost se bude od pravděpodobnosti lišit o méně než 0.02.
b) Kolik musíme udělat pokusů, aby jsme s pravděpodobností alespoň 0.95 mohli očekávat, že se relativní četnost bude lišit od pravděpodobnosti o méně než 0.02?

Řešení. Náhodná veličina $\mathbb{X}$ má binomické rozložení $Bi(n,p)$ a proto

$$E(\mathbb{X}) = np, \quad D(\mathbb{X}) = np(1-p).$$

Relativní četnost je veličina $\frac{\mathbb{X}}{n}$.
a) Podle Bernoulliovy věty máme

$$p\left(\left|\frac{\mathbb{X}}{n} - p\right| < \varepsilon\right) \geq 1 - \frac{p(1-p)}{n\varepsilon^2}.$$

Dosazením dostaneme

$$p\left(\left|\frac{\mathbb{X}}{1000} - 0.514\right| < 0.02\right) \geq 1 - \frac{0.514(1-0.514)}{1000\cdot 0.02^2} = 0.3755.$$

b) Opět podle Bernouliovy věty máme

$$p\left(\left|\frac{\mathbb{X}}{n}-0.514\right|<0.02\right)\geq 1-\frac{0.514(1-0.514)}{n\cdot 0.02^2}\geq 0.95.$$

Vyřešíme poslední nerovnost

$$1-\frac{0.514(1-0.514)}{n\cdot 0.02^2}\geq 0.95$$

a dostaneme

$$n\geq 12490.$$

□

Věta 1.152. Lindebergova - Levyho věta, Centrální limitní věta
Nechť $\mathbb{X}_1,\mathbb{X}_2,\ldots,\mathbb{X}_n$ jsou navzájem nezávislé náhodné veličiny, které mají všechny stejnou střední hodnotu μ a stejný rozptyl σ^2. Potom pro dostatečně velké n má náhodná veličina

$$\mathbb{Y}=\frac{\mathbb{X}_1+\mathbb{X}_2+\cdots+\mathbb{X}_n-n\mu}{\sigma\sqrt{n}}$$

přibližně normální rozložení $No(0,1)$.

Důsledek 1.153. *Nechť $\mathbb{X}_1,\mathbb{X}_2,\ldots,\mathbb{X}_n$ jsou navzájem nezávislé náhodné veličiny, které mají všechny stejnou střední hodnotu μ a stejný rozptyl σ^2. Potom pro dostatečně velké n pro náhodnou veličinu $\mathbb{Y}=\mathbb{X}_1+\mathbb{X}_2+\cdots+\mathbb{X}_n$ a pro libovolné $x_1,x_2\in\mathbb{R},x_1<x_2$ platí*

$$p\left(x_1\leq\mathbb{Y}\leq x_2\right)\doteq\Phi\left(\frac{x_2-n\mu}{\sigma\sqrt{n}}\right)-\Phi\left(\frac{x_1-n\mu}{\sigma\sqrt{n}}\right).$$

$$E(\mathbb{Y})=n\mu,\quad D(\mathbb{Y})=n\sigma^2.$$

Věta 1.154. Moivreova - Laplaceova věta
Nechť náhodná veličina $\mathbb{X}$ má binomické rozložení $Bi(n,p)$. Jestliže je n dostatečně velké a p není blízké ani k nule ani k jedné, potom lze toto binomické rozložení aproximovat normálním rozložením $No(\mu,\sigma)$, kde $\mu=np,\sigma=\sqrt{np(1-p)}$.

Důsledek 1.155. *Nechť $\mathbb{X}$ má binomické rozložení $Bi(n,p)$. Potom pro dostatečně velké n, v praxi $n>30$, p které není blízké ani k nule ani k jedné a pro libovoln $x_1,x_2\in\mathbb{R},x_1<x_2$ platí*

$$p\left(x_1\leq\mathbb{X}\leq x_2\right)\doteq\Phi\left(\frac{x_2-np}{\sqrt{np(1-p)}}\right)-\Phi\left(\frac{x_1-np}{\sqrt{np(1-p)}}\right).$$

$$E(\mathbb{X})=n\mu,\quad D(\mathbb{X})=n\sigma^2.$$

Aproximace se považuje za vyhovující pro

$$np(1-p)>9\quad a\quad \frac{1}{n+1}<p<\frac{n}{n+1}.$$

Příklad 1.156. Pravděpodobnost nastoupení jevu A v každém pokuse je rovna 0.7. Pomocí Moivre-Laplaceovy věty odhadněte, kolikrát musíme opakovat pokus, abychom s pravděpodobností 0.9 mohli očekávat, že se relativní četnost bude odlišovat od pravděpodobnosti o méně než 0.05?

Řešení. Máme náhodnou veličinu s binomickým rozložením, které můžeme aproximovat normálním rozložením.

$$\left|\frac{X}{n} - 0.7\right| < 0.05,$$

$$0.65n < X < 0.75n.$$

$$p(0.65n < X < 0.75n) = \Phi\left(\frac{0.75n - 0.7n}{\sqrt{n \cdot 0.7 \cdot 0.3}}\right) - \Phi\left(\frac{0.65n - 0.7n}{\sqrt{n \cdot 0.7 \cdot 0.3}}\right) =$$

$$\Phi\left(\frac{0.05n}{\sqrt{n \cdot 0.7 \cdot 0.3}}\right) - \Phi\left(-\frac{0.05n}{\sqrt{n \cdot 0.7 \cdot 0.3}}\right) = 2\Phi\left(\frac{0.05n}{\sqrt{n \cdot 0.7 \cdot 0.3}}\right) - 1 = 0.9$$

$$\Phi\left(\frac{0.05n}{\sqrt{n \cdot 0.7 \cdot 0.3}}\right) = 0.95$$

$$\frac{0.05\sqrt{n}}{\sqrt{21}} = 1.645$$

$$n = 228.65$$

Musíme provést alespoň 229 pokusů.
Podle přesnosti Vámi použitých hodnot Laplaceovy funkce se může numerický výsledek odlišovat

□

Příklad 1.157. Bylo provedeno 100 nezávislých pokusů. Pravděpodobnost nastoupení jevu A je v každém pokuse rovna 0.2. Určete pravděpodobnost, že počet nastoupení jevu A bude větší než 15 a menší než 30.

Řešení. Náhodná veličina udávající počet nastoupení jevu A má binomické rozložení $Bi(100; 0.2)$. Budeme je aproximovat normálním rozložením $No(\mu, \sigma)$, kde

$$\mu = np = 100 \cdot 0.2 = 20,$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{100 \cdot 0.2 \cdot 0.8} = \sqrt{16} = 4.$$

Proto

$$p\,(15 < \mathbb{X} < 30) = \Phi\left(\frac{30 - 20}{4}\right) - \Phi\left(\frac{15 - 20}{4}\right) =$$

$$\Phi(2.5) - \Phi(-1.5) = 0.99379 - (1 - 0.89435) = 0.88814.$$

□

Příklad 1.158. Pravděpodobnost nastoupení jevu A je v každém pokusu rovna p. Jestliže je n dostatečně velké ($n > 100$), jak je pravděpodobnost, že počet nastoupení jevu A bude od α do β.

Řešení. Pomocí Moivre-Laplaceovy věty

$$P(\alpha < x < \beta) = P\left(\frac{\alpha - np}{\sqrt{np(1-p)}} < \frac{x - np}{\sqrt{np(1-p)}} < \frac{\beta - np}{\sqrt{np(1-p)}}\right) =$$

$$= \Phi\left(\frac{\beta - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{\alpha - np}{\sqrt{np(1-p)}}\right).$$

□

Věta 1.159. Poissonova
Mějme posloupnost $\{\mathbb{X}_n\}$ náhodných veličin s rozložením $Bi(n, p_n), n = 1, 2 \ldots$. Nechť $p_n \to 0$ pro $n \to +\infty$ a $\lim\limits_{n\to+\infty} np_n = \lambda$, kde $\lambda > 0$ je konstanta. Potom

$$\lim_{n\to+\infty} p(\mathbb{X}_n = i) = \lim_{n\to+\infty} \binom{n}{i} p^i (1-p)^{n-i} = \frac{\lambda^i}{i!} \mathrm{e}^{-\lambda}.$$

Neboli binomické rozložení v limitě přechází v Poissonovo.
V praxi pro $n > 30$ a $p < 0.1$ můžeme $Bi(n, p)$ aproximovat rozložením $Po(\lambda = np)$ s chybou menší ne§ 10^{-2}.

Pojmy k zapamatování

- Definovali jsme si pravděpodobnost a odvodili si její základní vlastnosti.
- Zavedli jsme si náhodnou veličinu a ulázali si, jak se s ní pracuje.
- Uvedli sme si některé nejužívanější rozložení diskrétních i spojitých náhodných veličin.
- Speciální pozornost jsme věnovali normálnímu rozložení.
- Seznámili jsme se s limitními větami.

Kontrolní otázky

1. Co rozumíme pojmem rozptyl?
2. K čemu nám slouží náhodná veličina?
3. Kde se všude můžete setkat s normálním rozložením?

Cvičení

1. Obrazovka radaru je kruhová o poloměru r. Při zapnutí se na ní náhodně objeví svítící bod znamenající letící objekt. Určete pravděpodobnost, že svítící bod bude od středu obrazovky vzdálen o méně než $\frac{r}{2}$.

2. Náhodná veličina $\mathbb{X}$ má funkci hustoty

$$f(x) = \begin{cases} \dfrac{a}{\sqrt{a^2 - x^2}}, & |x| < a, \\ 0, & |x| \geq a. \end{cases}$$

   Určete 1) parametr a, 2) pravděpodobnost, že náhodná veličina $\mathbb{X}$ bude ležet v intervalu $\left(\frac{a}{2}, a\right)$.

3. Může být pro některou spojitou náhodnou veličinu $\mathbb{X}$
   a) distribuční funkce větší než 1?
   b) funkce hustoty větší než 1?
   c) distribuční funkce záporná?
   d) funkce hustoty záporná?

4. Náhodná veličina $\mathbb{X}$ má Laplaceovo rozložení s funkcí hustoty
$$f(x) = a \cdot e^{-\lambda|x|},$$
kde $\lambda > 0$ je konstanta příslušná k danému rozložení. Určete a) paramatr a, b) distribuční funkci, c)$E(\mathbb{X})$, d) $D(\mathbb{X})$.

5. Továrna vyrobí za směnu 20 000 diod. Pravděpodobnost výroby vadné diody je 0.02. Jaká je pravděpodobnost, že počet vadných diod za směnu bude nejvýše 450.

6. Továrna vyrobí za směnu 15 000 čipů. Pravděpodobnost výroby vadného čipu je 0.03. Jak je pravděpodobnost, že počet vadných čipů za směnu bude nejvýše 475.

7. Je dána funkce
$$f(x) = \begin{cases} 0, & x < 0, \\ ax^2, & 0 \le x < 1, \\ a(2-x)^2, & 1 \le x < 2, \\ 0, & x \ge 2. \end{cases}$$
Určete: a) parametr a tak, aby funkce $f(x)$ byla funkcí hustoty, b) střední hodnotu, rozptyl a směrodatnou odchylku, c) koeficienty šikmosti a špičatosti.

8. Pravděpodobnost poruchy stroje za dobu T je rovna 0.2. Určete pravděpodobnost, že ze 100 strojů stejného typu, které pracují nezávisle na sobě, bude mít poruchu 14 až 26 strojů. Řešte a) pomocí binomického rozložení, b) pomocí Čebyševovy nerovnosti, c) pomocí Moivre-Laplaceovy věty.

9. Odhadněte pomocí Moivre-Laplaceovy věty, kolik je třeba provést nezávislých pokusů, abychom s pravděpodobností 0.8 získali alespoň pětkrát kladný výsledek, jestliže při každém pokusu nastane kladný výsledek s pravděpodobnostˇ 0.05.

10. Náhodná veličina $\mathbb{X}$ má pro $x > 0$ funkci hustoty $f(x) = Axe^{-h^2x^2}$, kde $h > 0$ je parametr, a pro $x \le 0$ je $f(x) = 0$. Určete 1) koeficient A, 2) modus $\mathcal{M}$, 3) střední hodnotu a rozptyl 4) pravděpodobnost, že náhodná veličina bude menší než $\mathcal{M}$.

11. Odhadněte pomocí Moivre-Laplaceovy věty, kolik je třeba provést nezávislých pokusů, abychom s pravděpodobnostˇ 0.9 získali alespoň šestkrát kladný výsledek, jestliže při každém pokusu nastane kladný výsledek s pravděpodobností 0.04.

12. Pro náhodnou veličinu X, která má rozložení $F(2, k)$ určete 1)$P(X \ge x)$, 2) distribuční funkci.

13. Náhodný pokus spočívá ve vytažení 4 karet z důkladně promíchané sady 32 karet. Určete pravděpodobnost, že budou vytaženy karty červeá sedma, zelená desítka, žaludský král a kulové eso v uvedeném pořadí.

14. Máme 10 krabic. V každé je 10koulí. V i-té krabici je i černých a $10-i$ bílých koulí. Náhodně vybereme krabici a z ní vyjmeme jednu kouli. Jak je pravděpodobnost, že bude černá?

Výsledky

1. Použijeme geometrickou pravděpodobnost

$$p(A) = \frac{S_A}{S_\Omega} = \frac{\pi\frac{r^2}{4}}{\pi r^2} = \frac{1}{4}.$$

2. 1) Parametr a určíme z podmínky pro funkci hustoty

$$\int_{-\infty}^{+\infty} f(x)dx = 1.$$

Proto

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-a}^{a} \frac{a}{\sqrt{a^2-x^2}}dx = 1 \Rightarrow a = \frac{1}{\pi}.$$

2)$p\left(\frac{\frac{1}{\pi}}{2} < \mathbb{X} < \frac{1}{\pi}\right) = \frac{1}{3}$.

3. Přímo z definice plyne: a)ne, b) ano, c)ne, d) ne.

4.

$$a = \frac{\lambda}{2},$$

$$F(x) = \begin{cases} \frac{1}{2}e^{-\lambda x}, & x \geq 0 \\ 1 - \frac{1}{2}e^{-\lambda x}, & x > 0 \end{cases}$$

$$E(\mathbb{X}) = 0,$$

$$D(\mathbb{X}) = \frac{2}{\lambda^2}.$$

5. Máme $n = 20000, p = 0.02$. Potom

$$\bar{x} = np = 20000 \cdot 0.02 = 400, \sigma = \sqrt{np(1-p)} = 19.79898987.$$

$$p(\mathbb{X} \leq 450) = F(450) = \Phi\left(\frac{450-400}{19.79898987}\right) = \Phi(2.52538136179) = 0.99413.$$

6. Máme $n = 15000, p = 0.03$, potom

$$\bar{x} = np = 15000 \cdot 0.03 = 450, \sigma = \sqrt{np(1-p)} = 20.893.$$

$$p(\mathbb{X} \leq 475) = F(475) = \Phi\left(\frac{475-450}{20.893}\right) = \Phi(1.196573) = 0.884268.$$

7.

$$1 = \int_{-\infty}^{+\infty} f(x)dx = a\int_0^1 x^2dx + a\int_1^2 (2-x)^2dx = a\frac{x^3}{3}\Big|_0^1 - a\frac{(2-x)^3}{3}\Big|_1^2 = \frac{a}{3} + \frac{a}{3}.$$

Odtud $a = \frac{3}{2}$.

$$m_1 = E(X) = \frac{3}{2}\int_0^1 x^3dx + \frac{3}{2}\int_1^2 x(2-x)^2dx = \frac{3}{2}\frac{1}{4} + \frac{3}{2}\left(6 - \frac{28}{3} + \frac{15}{4}\right) = 1,$$

$$m_2 = \frac{3}{2}\int_0^1 x^4 dx + \frac{3}{2}\int_1^2 x^2(2-x)^2 dx = \frac{3}{10} + 14 + \frac{45}{2} = 1.1,$$
$$m_3 = \frac{3}{2}\int_0^1 x^5 dx + \frac{3}{2}\int_1^2 x^3(2-x)^2 dx = \frac{1}{4} + \frac{45}{2} - \frac{186}{5} + \frac{63}{4} = 1.3,$$
$$m_4 = \frac{3}{2}\int_0^1 x^6 dx + \frac{3}{2}\int_1^2 x^4(2-x)^2 dx = \frac{3}{14} + \frac{186}{5} - 63 + \frac{381}{14} = 1\frac{22}{35}.$$
$$M_1 = 0,$$
$$M_2 = m_2 - (m_1)^2 = 0.1,$$
$$M_3 = m_3 - 3m_1m_2 + 2(m_1)^3 = 0,$$
$$M_4 = m_4 - 4m_1m_3 + 6(m_1)^2m_2 - 3(m_1)^4 = \frac{1}{35}.$$

Potom
$$E(X) = m_1 = 1, D(X) = M_2 = 0.1, \sigma = \sqrt{D(X)}, k_1 = \frac{M_3}{\sigma^3} = 0,$$
$$k_2 = \frac{M_4}{\sigma_4} - 3 = \frac{\frac{1}{35}}{0.01} - 3 = -\frac{1}{7}.$$

8.
$$p(14 \le X \le 26) = p(|X-20| \le 6) = p(|X-20| < 7).$$

a) Přímý výpožet pomocí biomického rozložení je velmi zdlouhavý!
$$p(14 \le X \le 26) = \sum_{k=14}^{26} \binom{100}{k} \cdot 0.2^k 0.8^{100-k}$$

b) Podle Čebyševovy nerovnosti máme
$$p(|X-20| < 7) \ge 1 - \frac{np(1-p)}{\varepsilon^2} = 1 - \frac{100 \cdot 0.2 \cdot 0.8}{7^2} = 1 - \frac{16}{49} = 0.6735...$$

c) Podle Moivre-Laplaceovy věty máme:
$$p(14 \le X \le 26) = p\left(-1.5 \le \frac{x-20}{4} \le 1.5\right) = 2\Phi(1.5) - 1 = 0.8664...$$

9. Máme $Bi(n, 0.05)$ a máme určit n.
$$P(5 \le X < n) = \Phi\left(\frac{n - 0.05n}{\sqrt{n0.05 \cdot 0.95}}\right) - \Phi\left(\frac{5 - 0.05n}{\sqrt{n0.05 \cdot 0.95}}\right) = 0.8,$$
$$\Phi\left(\sqrt{n} \cdot \sqrt{19}\right) - \Phi\left(\frac{5 - 0.05n}{\sqrt{n0.05 \cdot 0.95}}\right) = 0.8,$$
protože je $n > 5$ je $\sqrt{n} \cdot \sqrt{19} > 4$ a proto můžeme předpokládat, že $\Phi(\sqrt{n} \cdot \sqrt{19}) = 1$. Potom
$$1 - \Phi\left(\frac{5 - 0.05n}{\sqrt{n0.05 \cdot 0.95}}\right) = 0.8,$$
$$\Phi\left(\frac{5 - 0.05n}{\sqrt{0.0475n}}\right) = 0.2,$$
$$\frac{5 - 0.05n}{\sqrt{0.0475n}} = -0.8416,$$
$$n \ge 144.$$

10. 1)

$$\int_{-\infty}^{+\infty} f(x)dx = 1 \Rightarrow A = 2h^2.$$

2)

$$\mathcal{M} = \frac{1}{h\sqrt{2}}$$

3)

$$E(X) = \frac{\sqrt{\pi}}{2h}, \quad D(X) = \frac{4-\pi}{4h^2}$$

4)

$$P(X < \mathcal{M}) \approx 0.393.$$

11. Máme $Bi(n, 0.04)$ a máme určit n.

$$P(6 \le X < n) = \Phi\left(\frac{n - 0.04n}{\sqrt{n0.04 \cdot 0.96}}\right) - \Phi\left(\frac{6 - 0.04n}{\sqrt{n0.04 \cdot 0.96}}\right) = 0.9,$$

$$\Phi\left(\sqrt{n} \cdot \sqrt{24}\right) - \Phi\left(\frac{6 - 0.04n}{\sqrt{n0.04 \cdot 0.96}}\right) = 0.9,$$

protože je $n > 6$ je $\sqrt{n} \cdot \sqrt{24} > 5$ a proto můžeme předpokládat, že $\Phi(\sqrt{n} \cdot \sqrt{24}) = 1$. Potom

$$1 - \Phi\left(\frac{6 - 0.04n}{\sqrt{n0.04 \cdot 0.96}}\right) = 0.9,$$

$$\Phi\left(\frac{6 - 0.04n}{\sqrt{0.0384n}}\right) = 0.1,$$

$$\frac{6 - 0.04n}{\sqrt{0.0384n}} = -1.289,$$

$$n \ge 247.$$

12. Prostým dosazením dostaneme
1)$P(X \ge x) = (1 + \frac{2x}{k})^{-k/2}$
2)

$$F(x) = 1 - \left(1 - \frac{2x}{k}\right)^{-\frac{k}{2}}.$$

13.

$$P = \frac{1}{32} \cdot \frac{1}{31} \cdot \frac{1}{30} \cdot \frac{1}{29} = 1.1587 \cdot 10^{-6}.$$

14. Použijeme větu o úplné pravděpodobnosti.
Označme A vytažení černé koule. Pravděpodobnost výběru i-té krabice je

$$p(B_i) = \frac{1}{10}.$$

Jevy B_i, $i = 1, 2. \ldots, 10$ tvoří úplný systém jevů. Proto

$$P(A) = \sum_{i=1}^{10} p_{B_i}(A)p(B_i) = \sum_{i=1}^{10} \frac{i}{10}\frac{1}{10}.$$

Maplety

V následujících mapletech si můžete některé studované pojmy přiblížit, případně si sestavit vlastní zadání příkladů.

1. Opakování pojmů z kombinatoriky
2. Binomické rozložení pravděpodobnosti
3. Aproximace binomického rozložení normálním
4. Exponenciální rozložení pravděpodobnosti
5. Geometrické rozložení pravděpodobnosti
6. Hypergeometrické rozložení pravděpodobnosti
7. Normální rozložení pravděpodobnosti
8. Kvantily normálního rozložení pravděpodobnosti
9. Poissonovo rozložení pravděpodobnosti
10. Fisher-Snedocorovo rozložení pravděpodobnosti
11. Studentovo rozložení pravděpodobnosti
12. Distribuční a frekvenční (pravděpodobnostní) funkce
13. Určitý integrál
14. Integrování
15. Derivování

2 Statistika

Průvodce studiem

Předmět Pravděpodobnost, statistika a operační výzkum je po předmětu Matematika 3 už druhým předmětem, jehož část se zabývá statistikou. Zatímco v předmětu Matematika 3 byla těžištěm zájmu pravděpodobnost, v této části předmětu MPSO bude těžištěm zejména statistika.

Protože základní principy statistiky byly už probrány v předmětu Matematika 3, doporučujeme příslušné oddíly projít ještě jednou. Uvedené opakování se bude určitě hodit. Také tabulku hodnot distribuční funkce Φ normovaného normálního rozdělení U budete potřebovat i v této části. Hodnoty funkce $\Phi(u)$ jsou uvedeny v tabulce 1.1 a 1.2 v předchozí kapitole.

Cíle

Po prostudování této kapitoly budete schopni:

- Pracovat s nejčastěji prováděným testem – t-testem a také s intervaly spolehlivosti.
- Budou zopakovány některé vzorce z předmětu BMA3, zejména vzorce pro $\overline{X}$, S^2 – a dále je vysvětlen a odvozen vzorec pro $\overline{S^2}$ jako nejlepší nestranný odhad neznámého rozptylu σ^2 normálního rozdělení.
- Naučíte se práci s testy typu $\mu_1 = \mu_2$.
- Seznámíte se s rozdíly mezi párovým a nepárovým testem.

2.1 Úvod

Ze všech existujících výroků o statistice si připomeneme jen dva:

1. *Existují tři druhy lží:*
   (a) *lež obyčejná*
   (b) *předvolební sliby*
   (c) *statistika*
2. V pohádce *Princové jsou na draka* zazněl refrén písně od Zdeňka Svěráka:

Statistika nuda je,
má však cenné údaje,
neklesejte na mysli,
ona vám to vyčíslí.

Někde mezi těmito dvěma výroky leží podstata statistiky.

Statistika nám může pomoci odhalit skrytou strukturu procesu, který máme popsán dosti velkým počtem měření, údajů, ... Pokud je hodnot příliš mnoho, tak se můžeme v nich „stratit“. Statistika nás učí, jak postupovat a jakým způsobem hledat podtatné údaje o každém souboru hodnot.

Protože jste se už se základy statistiky setkali, ale od té doby už uběhly minimálně tři semestry, tak si v rychlosti zopakujeme základní pojmy.

2.2 Zpracování statistického materiálu.

S těmito pojmy jste se stkali už v předmětu Matematika 3. Provedeme si opakování a zavedeme si jejich označení:
Základní soubor = množina hodnot se kterou pracujeme. Výběr= $\{x_1, x_2, \dots, x_n\}$.
n je rozsah výběru.
$[x_{\min}, x_{\max}]$ variační obor.
$R = x_{\max} - x_{\min}$ variační rozpětí.
Pokud pracujeme s rozsáhlejším výběrem, můžeme jej rozdělit do tříd.
Třídy - navzájem disjunktní množiny, jednoznačná zařaditelnost prvků.
Délka třídy - většinou se volí stejná délka h.
Počet tříd k bývá většinou mezi 7 až 15, popřípadě roven $\sqrt{n}$.

$$k = \frac{x_{\max} - x_{\min}}{h}.$$

Jinak využijeme empirické vztahy

$$k = 0.08R,$$

$$h < \frac{1}{12}R < 2h.$$

Každému prvku dané třídy přiřazujeme stejnou hodnotu - třídní znak.
Absolutní četnost n_i = počet prvků v i-té třídě.

$$\sum_i n_i = n.$$

Relativní četnost $f_i = \dfrac{n_i}{n}$ - často bývá uváděna v procentech.
Platí

$$\sum_i f_i = 1$$

Důkaz.

$$\sum_i f_i = \sum_i \frac{n_i}{n} = \frac{1}{n} \sum_i n_i = \frac{1}{n} n = 1.$$

□

Kumulativní absolutní četnost $N_i = \sum\limits_{s \leq i} n_s$.
Kumulativní relativní četnost $F_i = \sum\limits_{s \leq i} f_s$.

$$N_k = n, \qquad F_k = 1.$$

Relativní četnost f_i představuje empirickou hodnotu pravděpodobnosti, že náhodná veličina $\mathbb{X}$ nabude hodnoty z i-té třídy.
F_i je empirická hodnota distribuční funkce.
Variační řada = roztřídění do dvojic (x_i, f_i) a nebo (x_i, n_i).
Grafické způsoby zobrazení:
Graf - množina bodů $\{(x_i, n_i), i = 1, \ldots, k\}$.
Polygon - lomená čára spojující body (x_i, n_i).
Histogram - stupňovitá čára ohranizující obdelníky, jejichž základnou jsou rozsahy tříd a výškou odpovídající třídní četnosti.
Úsečkový diagram - množina úseček rovnoběžných s osou y, vycházejících ze středu třídy a s délkou n_i.

2.3 Výběrové charakteristiky a jejich vlastnosti

Definice 2.1. 1. Výběrový průměr

$$\bar{x} = \frac{1}{n} \sum_{i-1}^{n} x_i.$$

2. Výběrový rozptyl

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2.$$

3. Výběrová směrodatná odchylka

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2}.$$

4. Výběrové rozpětí

$$v = \max_{i=1,\ldots,n} \{x_i\} - \min_{i=1,\ldots,n} \{x_i\}.$$

Věta 2.2. *Nechť $X = (x_1, \ldots, X_n)$ je náhodný výběr z rozložení, které má střední hodnotu μ a rozptyl σ^2. Potom*

$$E(\bar{X}) = \mu,$$

$$D(\bar{X}) = \frac{\sigma^2}{n},$$

$$E(s^2) = \frac{n-1}{n} \sigma^2.$$

Důkaz.

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(x_i) = \frac{1}{n}n\mu = \mu,$$

$$D(\bar{X}) = D\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} D\left(\frac{x_i}{n}\right) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n} = \frac{D(X)}{n}.$$

□

Věta 2.3. *Nechť* $\mathbb{X}$ *má rozložení* $No(\mu, \sigma)$. *Potom*

$$\bar{x} \text{ má rozložení } No\left(\mu, \frac{\sigma}{\sqrt{n}}\right),$$

$$\frac{\bar{x}-\mu}{\sigma}\sqrt{n} \text{ má rozložení } No\,(0,1)\,,$$

$$\frac{s^2}{\sigma^2} \text{ má rozložení } \chi^2(n-1),$$

$$\frac{\bar{x}-\mu}{s}\sqrt{n-1} \text{ má rozložení } t(n-1).$$

2.4 Základní bodové odhady.

Požadavky na odhad:
Nechť G je výběrová charakteristika parametru a. Potom požadujeme:
a) Nestranost: $a = E(G)$.
b) Konzistentnost: $\lim\limits_{n\to+\infty} P(|E(G) - a| < \varepsilon) = 1$.
Se zvětšujícím se rozsahem výběru se zmenšuje počet chyb.

2.4.1 Bodové odhady

Pro paramety souboru se zhruba normálním rozdělením platí:

$$E(\bar{X}) = \mu,$$

$$D(\bar{X}) = \frac{\sigma^2}{n},$$

$$E(s^2) = \frac{n-1}{n}\sigma^2.$$

Proto

$$\mu := \bar{X},$$

$$\sigma^2 := \frac{n}{n-1}s^2,$$

$$\sigma := \sqrt{\frac{n}{n-1}}s.$$

2.5 Odhad parametrů, t-test, intervaly spolehlivosti

2.5.1 Nestranný a konzistentní odhad parametru rozdělení

Ve statistických testech v předmětu Matematika 3 jsme tiše předpokládali, že rozptyl σ^2 je známý. To ale ve skutečnosti většinou není pravda a my jej musíme odhadnout, tj. přibližně určit. Proto se nyní pustíme do trochy teorie a praxe v odhadování parametrů.

Příklad 2.4. Pět sad součástek o dvaceti kusech bylo podrobeno zkouškám extrémních teplot. U každé sady je v tabulce uveden počet součástek z daných dvaceti, které v teplotní zkoušce obstály:

z 20 obstálo x_i	$x_i - \overline{x}$	$(x_i - \overline{x})^2$
13	0	0
11	-2	4
12	-1	1
15	2	4
14	1	1

V tabulce už byla využita hodnota průměru $\frac{1}{5}\sum x_i = 13$. Ve třetím sloupci tabulky jsou uvedeny čtverce odchylek od průměru, odkud spočteme empirický rozptyl (= průměr čtverců odchylek od průměru ... :-)):

$$s^2 = \frac{1}{5}\sum(x_i - \overline{x})^2 = \frac{10}{5} = 2.$$

Jedná se o měření hodnot náhodné veličiny, kterou je možné popsat střední hodnotou μ a rozptylem σ^2. Ovšem tyto hodnoty neznáme - pokusíme se je odhadnout.

Otázka zní: Jak dobrým odhadem pro μ je průměr $\overline{x}$? Jak dobrým odhadem pro σ^2 je empirický rozptyl s^2?

Hodnoty $\overline{x}$, s^2 jsou různé pro různé soubory měření, při jejich popisu užíváme náhodné veličiny $X_1, X_2, \ldots, X_N$ označované jako náhodný výběr. Zde v teorii odhadů je potřeba tento i následující pojmy uvést přesně.

Definice 2.5. Říkáme, že veličiny $X_1, X_2, \ldots, X_N$ tvoří **náhodný výběr rozsahu** N z rozdělení pravděpodobnosti o distribuční funkci $F(x)$, pokud

a) jsou navzájem nezávislé;

b) mají stejné rozdělení pravděpodobnosti zadané distribuční funkcí $F(x)$.

Rozlišujeme malá a velká písmena. Naměřenou hodnotu označíme např. $x_1 = 13$, tuto náhodnost prvního měření reprezentuje náhodná veličina X_1, které nepřiřazujeme žádnou konkrétní hodnotu, pouze jsme si vědomi, že pod (velkým písmenem) X_1 se mohou skrývat různé hodnoty. Podobně se mohou skrývat různé hodnoty pod veličinami $X_2, \ldots, X_N$. Takže malá písmena budou vždy označovat konkrétní realizaci a velká budou označovat náhodné veličiny.

Definice 2.6. Libovolnou funkci $T_N := T(X_1, X_2, \ldots, X_N)$ nad náhodným výběrem X_1, $X_2, \ldots, X_N$ nazveme **statistikou**. Speciálně

a) Statistiku

$$\overline{X} = \frac{1}{N} \cdot \sum_{i=1}^{N} X_i \tag{2.1}$$

nazveme **výběrovým průměrem**;

b) Statistiku

$$\overline{S^2} = \frac{1}{N-1} \cdot \sum_{i=1}^{N} (X_i - \overline{X})^2 \tag{2.2}$$

nazveme **výběrovým rozptylem**.

Dále pokud do statistiky T_N dosadíme konkrétní naměřené hodnoty $x_1, x_2, \ldots, x_N$, dostaneme hodnotu $t_N = t(x_1, x_2, \ldots, x_N)$, která se nazývá ***realizací statistiky*** *T_N.*

Například pokud dosadíme do vzorců 2.1, 2.2 konkrétní hodnoty měření x_i, dostaneme realizaci $\overline{x}$ výběrového průměru $\overline{X}$ a realizaci $\overline{s^2}$ výběrového rozptylu $\overline{S^2}$.

Nyní nás bude zajímat, jak dobrým odhadem neznámé střední hodnoty μ veličiny X je realizace $\overline{x}$ výběrového průměru $\overline{X}$ (respektive jak dobrým odhadem neznámého rozptylu σ jsou realizace s^2, $\overline{s^2}$ veličin S^2, $\overline{S^2}$).

Definice 2.7. Statistiku T_N nazveme **nestranným odhadem** parametru γ, pokud $ET_N = \gamma$ (střední hodnota veličiny T_N je rovna hodnotě parametru γ).

Vysvětlení pojmu nestrannosti pomocí konkrétních hodnot měření: pokud budeme opakovaně měřit hodnoty $x_{1,i}, x_{2,i}, \ldots, x_{N,i}$ a opakovaně počítat realizace $t_{N,i} = t(x_{1,i}, x_{2,i}, \ldots, x_{N,i})$ nestranného odhadu T_N parametru γ pro $i = 1, 2, \ldots, n, \ldots$, bude platit vztah

$$\lim_{n\to\infty} P\left(\left(\frac{1}{n}\sum_{i=1}^{n} t_{N,i}\right) \in (\gamma - \varepsilon; \gamma + \varepsilon)\right) = 1 \tag{2.3}$$

platí pro každé malé pevně zvolené reálné kladné ε.

Laicky řečeno, pokud T_N je nestranným odhadem parametru γ rozdělení veličiny X, tak **pro rostoucí n (= rostoucí počet realizací $t_{N,i}$)** je průměr realizací $\frac{1}{n}\sum_{i=1}^{n} t_{N,i}$ **skoro jistě (= s pravděpodobností rovnou jedné)** stále blíže hodnotě γ.

Jinými slovy, nestrannost zaručuje takovou konstrukci vzorce pro T_N, která bere v úvahu všechny možné dostupné realizace $t_{N,i}$ a **nestraní** žádné z nich – pro rostoucí počet realizací se aritmetický průměr těchto realizací skoro jistě blíží neznámé hledané hodnotě γ.

Definice 2.8. Statistiku T_N nazveme **konzistentním odhadem** parametru γ, pokud posloupnost náhodných veličin $(T_N)_{N=1}^{\infty}$ konverguje k hodnotě parametru γ podle pravděpodobnosti, tj.

$$\lim_{N\to\infty} P\left(T_N \in (\gamma-\varepsilon;\gamma+\varepsilon)\right) = 1 \tag{2.4}$$

platí pro každé malé pevně zvolené reálné kladné ε.

Jinak řečeno, pokud T_N je konzistentním odhadem parametru γ rozdělení veličiny X, tak pro rostoucí N (= **rostoucí počet měření pro výpočet jedné realizace**) je hodnota t_N **skoro jistě (= s pravděpodobností rovnou jedné)** stále blíže hodnotě γ.

Jinými slovy, konzistence zaručuje takovou konstrukci vzorce pro T_N, která je **konzistentní (= česky: důsledná)** v tom ohledu, že pro rostoucí počet měření při výpočtu jedné realizace se tato realizace skoro jistě blíží neznámé hledané hodnotě γ.

Uvedené definice nyní osvětlíme konkrétně při hledání odhadu střední hodnoty μ a rozptylu σ^2 veličiny X s normálním rozdělením pravděpodobnosti.

Věta 2.9. *Pokud náhodná veličina X má konečnou střední hodnotu μ, tak výběrový průměr $\overline{X}$ je nestranným a konzistentním odhadem μ.*

Důkaz. a) $\mu_{\overline{X}} = E\overline{X} = E\frac{1}{N}\sum X_i = \mu$, tj. střední hodnota náhodné veličiny $\overline{X}$ je rovna parametru μ; tedy odhad $\overline{X}$ je nestranným odhadem střední hodnoty μ.

b) $\sigma^2_{\overline{X}} = D\overline{X} = D\frac{1}{N}\sum X_i = \frac{\sigma^2}{N}$ a platí

$$\lim_{N\to\infty} D\overline{X} = \lim_{N\to\infty} \frac{\sigma^2}{N} = 0,$$

tj. limita rozptylu odhadu $\overline{X}$ pro rostoucí počet měření je rovna nule – jinými slovy, realizace odhadu $\overline{X}$ se pro rostoucí počet měření skoro jistě blíží hodnotě μ, čili $\overline{X}$ je konzistentním odhadem hodnoty μ.

$\square$

Čili vzorec pro průměr hodnot $\overline{x}$ funguje přesně tak, jak potřebujeme – "nestraní" konkrétnímu měření a "skoro jistě" směřuje přímo k určení střední hodnoty μ, a dále je

konzistentní (= důsledný) v tom smyslu, že průměr tisíce hodnot je "skoro jistě" lepší odhadem μ než průměr stovky hodnot.

Otázkou nyní je najít nejvhodnější odhad pro neznámý rozptyl σ^2 veličiny X. Máme k dispozici hodnotu S^2 jako míru vychýlení od průměru – je ona tím nejvhodnějším odhadem hodnoty σ^2?

Věta 2.10. *Pokud náhodná veličina X má konečnou střední hodnotu μ a konečný rozptyl σ^2, tak nestranným a konzistentním odhadem rozptylu σ^2 veličiny X je výběrový rozptyl*

$$\overline{S^2} := \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \overline{X})^2 = \frac{N-1}{N}\cdot S^2.$$

Při vysvětlení obsahu předchozí věty začněme nejprve u odhadu S^2. Víte, že empirický rozptyl s^2 lze vyjádřit buď z definice jako

$$s^2 = \frac{1}{N}\cdot\sum_{i=1}^{N}(x_i - \overline{x})^2, \tag{2.5}$$

nebo po úpravách ve tvaru praktičtějším pro výpočet (= průměr čtverců minus čtverec průměru)

$$s^2 = \left(\frac{1}{N}\sum_{i=1}^{N}x_i^2\right) - \overline{x}^2. \tag{2.6}$$

Při matematickém popisu nyní musíme konkrétní měřené hodnoty ve vzorcích nahradit náhodnými veličinami s velkými písmeny a dostáváme

$$S^2 = \frac{1}{N}\cdot\sum_{i=1}^{N}(X_i - \overline{X})^2, \tag{2.7}$$

respektive

$$S^2 = \left(\frac{1}{N}\sum_{i=1}^{N}X_i^2\right) - \overline{X}^2. \tag{2.8}$$

Vyjdeme z toho, že $\sigma^2 = EX_i^2 - \mu^2$, a dále platí $\sigma^2_{\overline{X}} = E\overline{X}^2 - \mu^2$. Vypočtěme střední hodnotu veličiny S^2:

$$\begin{aligned}
ES^2 &= E\left(\frac{1}{N}\sum X_i^2 - \overline{X}^2\right) = \frac{1}{N}\left(\sum EX_i^2\right) - E\overline{X}^2 = \\
&= \frac{1}{N}\left(\sum(\sigma^2+\mu^2)\right) - (\sigma^2_{\overline{X}} + \mu^2) = \\
&= \sigma^2 + \mu^2 - \sigma^2_{\overline{X}} - \mu^2 = \sigma^2 - \sigma^2_{\overline{X}} = \sigma^2 - \frac{\sigma^2}{N} = \frac{N-1}{N}\cdot\sigma^2.
\end{aligned}$$

Čili střední hodnota statistiky S^2 není rovna odhadovanému parametru σ^2, to znamená, že S^2 není nestranným odhadem rozptylu σ^2. Zkrátka a dobře vzoreček 2.7 není dobře zkonstruován, protože jeho realizace s^2 jsou vždy trochu menší než neznámá hledaná hodnota σ^2

$$\left(s^2 \doteq \frac{N-1}{N} \cdot \sigma^2 < \sigma^2\right),$$

až na patologické případy $\sigma^2 = 0$ a $\sigma^2 = \infty$, které nás nezajímají (matematicky jsou takové náhodné veličiny zkonstruovatelné, ale v praxi měřené veličiny mají vždy konečný kladný rozptyl). Ale k nalezení nestranného odhadu už máme jen krok – můžeme se totiž poučit z výpočtu ES^2. Pokud vynásobíme S^2 konstantou $\frac{N}{N-1}$ (označme novou veličinu $\overline{S^2}$):

$$\overline{S^2} := \frac{N}{N-1} \cdot S^2, \tag{2.9}$$

tak dostaneme

$$E\overline{S^2} = \frac{N}{N-1} \cdot ES^2 = \frac{N}{N-1} \cdot \frac{N-1}{N} \cdot \sigma^2 = \sigma^2,$$

čili $\overline{S^2}$ už je nestranným odhadem hodnoty σ^2.

Dále budeme jako nestranný a konzistentní odhad rozptylu σ^2 veličiny X užívat tedy statistiku $\overline{S^2}$. Má tedy ještě nějaký smysl "stará a překonaná" hodnota S^2? Vrátíme-li se k příkladu 2.4, kde $s^2 = 2$, lze vypočíst $\overline{s^2} = \dfrac{5}{4} \cdot 2 = 2{,}5$.

1. Hodnota $s^2 = 2$ má svou váhu – vyjadřuje rozptyl souboru uvedených pěti měření. Jedná se o empirický rozptyl – rozptyl naměřených hodnot.

2. Skutečný rozptyl σ^2 veličiny přeživších součástek je větší než rozptyl měření u pěti sad – proto je $\overline{s^2} = 2{,}5$ jeho vhodnějším odhadem.

V příkladu 2.4 můžeme uzavřít, že počet přeživších součástek ze sady dvaceti při extrémní teplotní zátěži lze popsat normálním rozdělením s parametry $\mu \doteq \overline{x} = 13$, $\sigma^2 \doteq \overline{s^2} = 2{,}5$.

Jednoduše řečeno, rozptyl s^2 vypočtený z několika naměřených hodnot je vždy o něco menší než skutečný rozptyl σ^2 měřené veličiny. Proto při odhadu σ^2 musíme vypočtené s^2 "trochu zvětšit" vynásobením členem $\dfrac{N}{N-1}$. Je zřejmé, že pro rostoucí N bude tento člen rychle konvergovat k 1.

Další možný tvar vzorce pro $\overline{S^2}$ lze získat po úpravách s využitím 2.8 a vztahu $\overline{X} = \frac{1}{N} \cdot \sum X_i$:

$$\overline{S^2} = \frac{N}{N-1} \cdot S^2 = \frac{N}{N-1} \cdot \left(\frac{1}{N}\sum X_i^2 - \frac{1}{N^2}\left(\sum X_i\right)^2\right),$$

a tedy po vykrácení konstantou N a roznásobení závorky dostaneme

$$\overline{S^2} = \left(\frac{1}{N-1} \cdot \sum_{i=1}^{N} X_i^2\right) - \frac{1}{N(N-1)} \cdot \left(\sum_{i=1}^{N} X_i\right)^2. \tag{2.10}$$

Definice 2.11. Výraz $ss := \sum(x_i - \overline{x})^2$ budeme nazývat **součet čtverců** měření veličiny X.

Při tomto označení platí

$$S^2 = \frac{1}{N} \cdot SS \tag{2.11}$$

a zejména

$$\overline{S^2} = \frac{1}{N-1} \cdot SS. \tag{2.12}$$

S pojmem součtu čtverců budeme ještě pracovat zejména v kapitolách 3 a 5. Podle okolností budeme při výpočtu $\overline{S^2}$ užívat vzorec 2.2, 2.9, 2.10 nebo 2.12.

Zbývá ještě vyjádření k takzvanému počtu stupňů volnosti odhadu.

Příklad 2.12. Kdybych vám řekl, abyste mi nadiktovali pět reálných čísel, a nedal žádnou další podmínku nebo omezení, mohli byste říct čísla, jaká chcete, například

$$0, 70, 314, 32, 100.$$

Máte svobodu volby, která čísla vybrat. Jinými slovy, máte pět stupňů volnosti, protože si zcela svobodně a volně vybíráte pět čísel.

Kdyby ale úkol zněl: Nadiktujte mi pět čísel, jejichž průměr je 25, trochu bych vaši volnost omezil – první čtyři čísla byste mohli volit libovolně, například

$$0, 70, 314, 32,$$

ale páté číslo je mým požadavkem jednoznačně určeno. Aby průměr pěti čísel byl 25, jejich součet musí být roven $5 \cdot 25 = 125$, tj. páté číslo musí být rovno $125 - 416 = -291$. Tuto druhou úlohu lze charakterizovat tím, že stupeň její volnosti je 4.

Čili obecně pro N hodnot, u kterých je předem dán jejich průměr, zbývá $N - 1$ stupňů volnosti.

Podobná situace se objevuje i při odhadování rozptylu populace: Uvažujeme-li soubor měření N hodnot jisté veličiny X, z nich lze určit průměr $\overline{x}$. Chceme-li dále odhadnout rozptyl pro tuto konkrétní (už určenou) hodnotu $\overline{x}$, máme už jen $N - 1$ stupňů volnosti (ve vzorci pro $\overline{s^2}$ hodnotu $\overline{x}$ potřebujeme znát, protože při určení $\overline{s^2}$ počítáme totiž míru vychýlení měření právě od hodnoty $\overline{x}$).

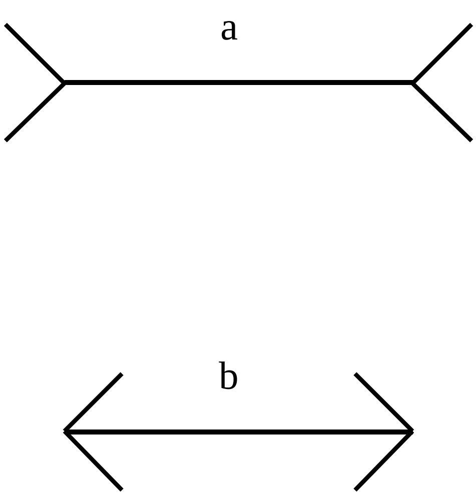
Tedy třeba v příkladu 2.4 má odhad rozptylu $\overline{s^2} = 2{,}5$ čtyři stupně volnosti.

Tento přístup určení počtu stupňů volnosti lze užít i na některé další situace v tomto textu. Obecně platí:

Počet stupňů volnosti odhadu = počet měření minus počet parametrů odhadnutých již dříve.

2.5.2 t-test typu „μ =konstanta"

Příklad 2.13. Je známa následující grafická iluze (viz obr. 2.1), že totiž úsečka **a** se zdá delší než úsečka **b** (počítá se délka bez koncových šipek), i když ve skutečnosti jsou obě úsečky stejně dlouhé.



Obr. 2.1: K př. 2.13: Grafická iluze: délky úseček a, b jsou stejné.

Chceme nyní experimentem ověřit, že úsečka typu **a** se zdá pozorovateli delší sama o sobě, bez porovnání s úsečkou **b**. Náhodně vybraným pěti lidem jsme na deset sekund ukázali úsečku **a** dlouhou 6 cm. Poté jsme je požádali, aby úsečku dané délky nakreslili, a změřili jsme její délku. Byla získána data $x_1 = 8$, $x_2 = 11$, $x_3 = 9$, $x_4 = 5$, $x_5 = 7$.

Průměr těchto dat je $\overline{x} = 8$ cm. Chceme testovat hypotézu, že střední hodnota μ celé lidské populace při ohodnocení délky úsečky je významně větší než její skutečná délka 6 cm. V této situaci neznáme rozptyl σ^2 měřené veličiny a musíme jej odhadnout pomocí $\overline{s^2}$. Pak testovacím kritériem bude tzv. rozdělení t s hodnotou vypočtenou podle vztahu

$$t := \frac{\overline{x} - 6}{\frac{\overline{S}}{\sqrt{N}}}, \text{ kde } \overline{S} = \sqrt{\overline{S^2}}.$$

Toto t-rozdělení odvodil jistý pan William Sealy Gosset – ovšem příslušný článek uveřejnil nikoli pod svým vlastním jménem, ale pod jménem Student, a rozdělení je tedy známo pod názvem Studentovo t-rozdělení.

Hustota t-rozdělení závisí na počtu ν stupnů volnosti, se kterými se určí odhad rozptylu $\overline{S^2}$, a má tvar

$$f_\nu(x) = \frac{1}{\sqrt{\nu} \cdot \beta(\frac{1}{2}, \frac{\nu}{2}) \cdot \left(1 + \frac{x^2}{\nu}\right)^{\frac{1+\nu}{2}}} = \frac{\Gamma(\frac{1+\nu}{2})}{\sqrt{\pi \cdot \nu} \cdot \Gamma(\frac{\nu}{2}) \cdot \left(1 + \frac{x^2}{\nu}\right)^{\frac{1+\nu}{2}}},$$

podle toho, zda se čtenáři více líbí funkce

$$\beta(p, q) = \int_0^1 u^{1-p} \cdot (1-u)^{1-q} \, \mathrm{d}u$$

(tzv. β-funkce), nebo funkce

$$\Gamma(r) = \int_0^\infty u^{r-1} \cdot e^{-u} \, \mathrm{d}u$$

(tzv. Γ-funkce). Přechod mezi jednotlivými verzemi vzorce hustoty plyne ze vztahu mezi β-funkcí a Γ-funkcí

$$\beta(p, q) = \frac{\Gamma(p) \cdot \Gamma(q)}{\Gamma(p+q)}$$

a z toho, že platí $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ (odvození viz například v učebnici [30]).

Funkce $f_\nu(x)$ je opět jedna z funkcí, kterou by člověk nerad potkal pozdě v noci v lese, ale ukazuje se, že i takové funkce jsou užitečné.

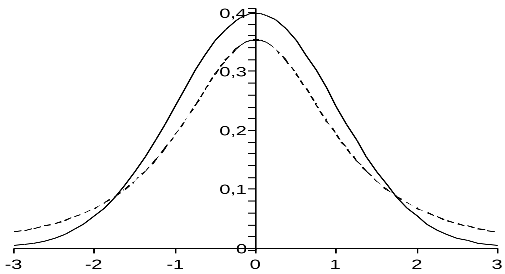
Uveďme zde některé vlastnosti t-rozdělení, které budeme využívat:

a) Hustota f_ν je symetrickou funkcí vzhledem k ose y, tj. je sudá: platí $f_\nu(x) = f_\nu(-x)$.

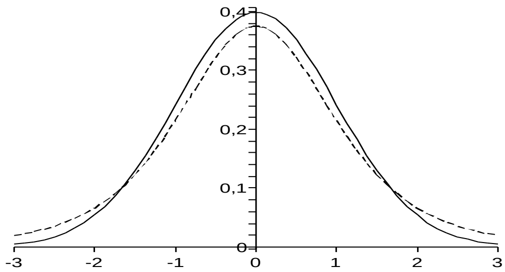
b) Kritická t-hodnota je dále od počátku než kritická U-hodnota (kritická hodnota normovaného normálního rozložení), protože t-rozdělení je odvozeno při neznámém rozptylu, tj. zahrnuje větší míru náhodnosti a nejistoty než U-rozdělení (veličina t má větší rozptyl než veličina U). Hustota rozdělení t je *nižší a širší* než hustota rozdělení U.

c) Čím lepší je náš odhad neznámého rozptylu σ měřené veličiny, tím více se t-rozdělení bude podobat U-rozdělení. A odhad rozptylu bude tím lepší, čím vyšší je počet měření N (tj. čím vyšší je počet stupňů volnosti ν).

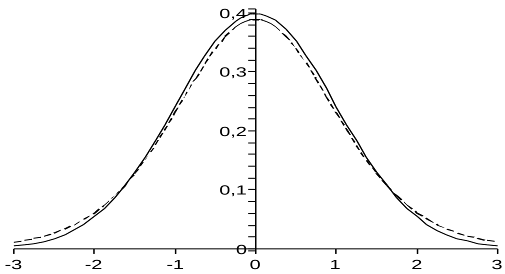
Vlastnosti b), c) lze znázornit graficky porovnáním U-rozdělení s t-rozdělením o různém počtu ν stupňů volnosti – (hustota U-rozdělení je v obrázcích znázorněna plnou čarou, hustota t-rozdělení čárkovanou čarou):



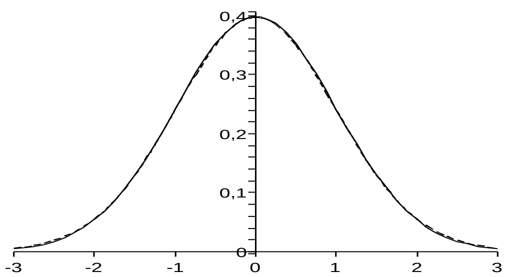
$\nu = 2$



$\nu = 4$



$\nu = 10$



$\nu = 60$

Z obrázků je vidět, že pro rostoucí počet stupňů volnosti se hustota t-rozdělení (v obrázku její graf vyznačen slabě) svým tvarem stále více blíží ke tvaru funkce hustoty U-rozdělení, a pro $\nu = 60$ je hustota t-rozdělení prakticky totožná s hustotou U-rozdělení (omlouvám se za malou tloušťku čar, ale při silnější tloušťce nebyl na obrázku patrný rozdíl mezi čárkovanou a plnou čarou).

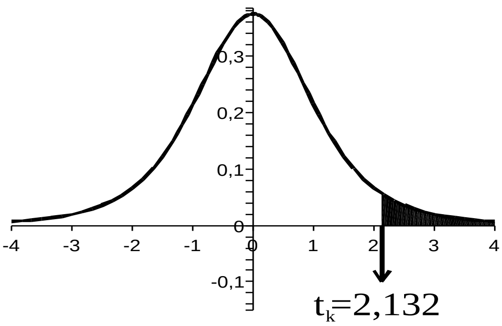
d) Pro určení kritických hodnot t_k budeme potřebovat hodnoty integrálů

$$P_\nu(t \leq x) = F_\nu(x) = \int_{-\infty}^{x} f_\nu(u)\mathrm{d}u, \quad P_\nu(-x \leq t \leq x) = \int_{-x}^{x} f_\nu(u)\underset{.}{u}.$$

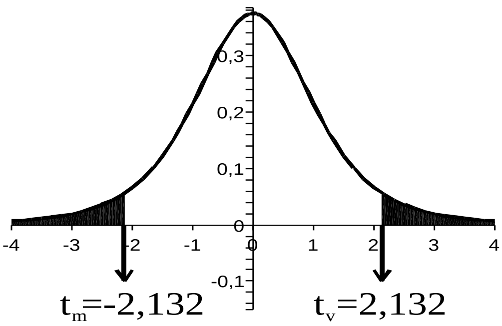
Tyto integrály se nepočítají vždy znovu a znovu, poněvadž jejich výpočet je složitý, ale jednou provždy byly spočteny a sestaveny do tabulky. Protože pro jednu hodnotu ν lze sestavit tabulku tak velkou jako je tabulka funkce Φ (BMA3), měli bychom pro 35 různých hodnot ν také 35 různých tabulek. Toto množství dat je zredukováno jen na několik hodnot v závislosti na hladině významnosti α.

A vůbec, pro statistické testy je nejužitečnější místo hodnot distribuční funkce přímo tabulka kritických hodnot pro různé α a ν – jedná se o tabulku 2.1. S touto tabulkou budeme pracovat tak, že vybereme řádek s daným počtem stupňů volnosti ν, sloupec s danou hodnotou $\alpha = q$ u pravostranného ($\alpha = 2q$ u oboustranného) testu. Na průsečíku řádku se sloupcem se pak nachází požadovaná kritická hodnota testu.Tabulku lze volit takto úsporně, protože mezi jednostranným a oboustranným testem je následující vztah:

- t_k u levostranného testu se liší od pravostranného pouze znaménkem.
- t_k u pravostraného testu pro $\alpha = q$ je stejné jako $|\pm t_k|$ u oboustranného testu pro $\alpha = 2q$. Je to vidět i na srovnání následujících dvou obrázků:



Pravostranný t-test pro $\alpha = q = 0{,}05$, $\nu = 4 \ldots t_k = 2{,}132$. Šrafovaná část tvoří 5% obsahu celého podgrafu.



Tab. 2.1: Kritické hodnoty t-testu.

ν	q=0,4 2q=0,8	0,25 0,5	0,1 0,2	0,05 0,1	0,025 0,05	0,01 0,02	0,005 0,01	0,001 0,002
1	0,325	1,000	3,078	6,314	12,704	31,821	63,657	318,31
2	0,289	0,816	1,886	2,920	4,303	6,965	9,925	22,326
3	0,277	0,765	1,638	2,353	3,182	4,541	5,841	10,213
4	0,271	0,741	1,533	2,132	2,776	3,747	4,604	7,173
5	0,267	0,727	1,476	2,015	2,571	3,365	4,032	5,893
6	0,265	0,718	1,440	1,943	2,447	3,143	3,707	5,208
7	0,263	0,711	1,415	1,895	2,365	2,998	3,499	4,785
8	0,262	0,706	1,397	1,860	2,306	2,896	3,355	4,501
9	0,261	0,703	1,383	1,833	2,262	2,821	3,250	4,297
10	0,260	0,700	1,372	1,812	2,228	2,764	3,169	4,144
11	0,260	0,697	1,363	1,796	2,201	2,718	3,106	4,025
12	0,259	0,695	1,356	1,782	2,179	2,681	3,055	3,930
13	0,259	0,694	1,350	1,771	2,160	2,650	3,012	3,852
14	0,258	0,692	1,345	1,761	2,145	2,624	2,977	3,787
15	0,258	0,691	1,341	1,753	2,131	2,602	2,947	3,733
16	0,258	0,690	1,337	1,746	2,120	2,583	2,921	3,686
17	0,257	0,689	1,333	1,740	2,110	2,567	2,898	3,646
18	0,257	0,688	1,330	1,734	2,101	2,552	2,878	3,610
19	0,257	0,688	1,328	1,729	2,093	2,539	2,861	3,579
20	0,257	0,687	1,325	1,725	2,086	2,528	2,845	3,552
21	0,257	0,686	1,323	1,721	2,080	2,518	2,831	3,527
22	0,256	0,686	1,321	1,717	2,074	2,508	2,819	3,505
23	0,256	0,685	1,319	1,714	2,069	2,500	2,807	3,485
24	0,256	0,685	1,318	1,711	2,064	2,492	2,797	3,467
25	0,256	0,684	1,316	1,708	2,060	2,485	2,787	3,450
26	0,256	0,684	1,315	1,706	2,056	2,479	2,779	3,435
27	0,256	0,684	1,314	1,703	2,052	2,473	2,771	3,421
28	0,256	0,683	1,313	1,701	2,048	2,467	2,763	3,408
29	0,256	0,683	1,311	1,699	2,045	2,462	2,756	3,396
30	0,256	0,683	1,310	1,697	2,042	2,457	2,750	3,385
40	0,255	0,681	1,303	1,684	2,021	2,423	2,704	3,307
60	0,254	0,679	1,296	1,671	2,000	2,390	2,660	3,232
120	0,254	0,677	1,289	1,658	1,98	2,358	2,617	3,160
∞	0,253	0,674	1,282	1,645	1,960	2,326	2,576	3,090

Oboustranný t-test pro $\alpha = 2q = 0{,}1$, $\nu = 4 \ldots t_m = -2{,}132$, $t_v = 2{,}132$. Šrafovaná část tvoří celkem 10% obsahu celého podgrafu (na každé straně je vyšrafováno 5% obsahu podgrafu).

– Ze symetrie hustoty t-rozdělení je patrné, že pokud budeme provádět levostranný t-test, stačí najít kritickou hodnotu pravostranného testu a změnit její znaménko na záporné.

Vraťme se nyní zpět k příkladu 2.13 a proveďme statistický t-test:

(K1) H_0: $\mu = 6$ *(střední hodnota délky úsečky není ovlivněna iluzí prodloužení).* H_1: $\mu > 6$.

(K2) *Testovým kritériem bude výběrový průměr $\overline{X}$, jehož realizaci $\overline{x} = 8$ převedeme na t-hodnotu*

$$\frac{8-6}{est\sigma_{\overline{X}}}.$$

*(*est *označuje odhad, z anglického* estimate *[estimit] – protože v dalším textu budeme odhadovat odchylku i pomocí jiných funkcí než $\overline{X}$, bude výhodné si tímto označením připomenout, u které veličiny vlastně odchylku nebo rozptyl odhadujeme).*

(K3) $\overline{s^2} = 5$, *a tedy*

$$est\sigma^2_{\overline{X}} = \frac{\overline{s^2}}{N} = \frac{5}{5} = 1,$$

přičemž počet stupňů volnosti odhadu je $N - 1 = 5 - 1 = 4$, tj. za předpokladu platnosti H_0 má veličina $\frac{\overline{X}-6}{1}$ Studentovo t-rozdělení pro $\nu = 4$.

(K4) *Příslušná kritická hodnota t_k je na průsečíku řádku $\nu = 4$ a sloupce pro $\alpha = 0{,}05$ (u pravostranného testu), tj. $t_k = 2{,}132$.*

(K5) $\frac{8-6}{1} = 2 < t_k = 2{,}132 \Rightarrow H_0$ *nezamítáme, nenašli jsme dostatek důkazů pro potvrzení iluze větší délky.*

2.5.3 Několik poznámek ke statistickému testu

Logika formulace *nezamítáme* H_0

V právě dokončeném příkladu bylo odpovědí *hypotézu H_0 nezamítáme*. Logiku této formulace snad osvětlí následující příklad.

Příklad 2.14. Když něco někde nenajdu, neznamená to, že to tam není.

S nástupem lyžařské sezóny jsem začal oprašovat svou výstroj a zjistil, že nemohu najít své lyžařské brýle, i když jsem prohledal celý byt. Usoudil jsem, že se asi ztratily v průběhu posledního lyžování nebo přes léto, a se zoufalstvím v očích oznámil manželce, že budu muset utratit 800 Kč za nové. Manželka byla vyděšena touto poznámkou a požádala mne, abych provedl hledání ještě jednou a důkladněji. Dosti neochotně jsem tak učinil, ale nakonec jsem brýle našel v zadním rohu své skříně.

Tento příklad je ilustrací jednoho celkem logického principu: pokud naleznu hledaný předmět, určitě vím, že tam je. Ale pokud jej nenaleznu, může to znamenat buď že tam není, nebo že tam je, ale mé hledání nebylo dost důkladné (nemělo dostatečnou sílu).

Testování hypotéz je také takovým hledáním – hledáme vliv jedné veličiny na druhou veličinu. Pokud nalezneme tento vliv (zamítáme H_0), víme „s jistotou“ (na hladině významnosti α), že existuje. Pokud vliv nenalezneme, může to znamenat buď že tento vliv neexistuje, nebo že existuje, ale síla testu nebyla dostatečná pro jeho nalezení.

Výsledek *zamítáme* H_0 má docela pevný logický základ (pro $\alpha = 0{,}05$ platí s pravděpodobností 95%). Ale říct v případě, kdy nebyla překročena kritická hodnota, že H_0 *přijímáme* nebo H_0 *platí*, je příliš ukvapené, protože při důkladnějším hledání by se mohlo ukázat, že určitý vliv existuje, tj. H_0 neplatí. Ustálilo se tedy rčení *nezamítáme* H_0, které odpovídá jisté opatrnosti v učinění konečného závěru.

Fráze *nezamítáme* H_0 je tedy opatrným vyjádřením, které je zcela na místě. Znamená to, že říkáme: „Výsledky testu nám neposkytují dostatečné důkazy k závěru, že H_0 neplatí.“

K příkladu 2.14: Co by to znamenalo, kdybych provedl tak důkladné hledání, že bych obrátil celý byt naruby, ale přesto brýle nenašel? Stále by existovala jistá malá šance, že jsem je přehlédl v nějaké zapadlé škvíře, ale protože jsem je nenašel, bylo by rozumné koupit nové.

Podobně i experiment provedený v úžasné síle a rozsahu, pokud neprokáže vliv nezávislé proměnné na závislou proměnnou, nás vede k závěru, že „je rozumné přijmout H_0“.

Příklad 2.15. Chceme otestovat kvalitu jisté techniky pamatování. Vybereme náhodně dvě skupiny lidí. Oběma skupinám je předložen jistý počet navzájem nesouvisejících slov k zapamatování (například 20 slov). Poté jsou lidé z první skupiny okamžitě vyzkoušeni, kolik slov si zapamatovali. Lidé ze druhé skupiny jsou vyzkoušeni až o týden později. Čili nezávislou proměnnou je doba pamatování, závislou proměnnou je počet zapamatovaných slov z daných dvaceti. Stanovíme hypotézy testu:

H_0: počet zapamatovaných slov nezávisí na době pamatování (technika zapamatování byla tak dobrá, že po týdnu si pamatuji stejně dobře jako po bezprostředním naučení slov).

H_1**:** počet zapamatovaných slov závisí na době pamatování (= době držení slov v paměti).

Pokud v testovaných skupinách bude například jen v každé pět lidí a vliv se neprokáže, můžeme uzavřít, že H_0 platí, tj. že technika učení je vynikající? Asi ne, protože právě zde bychom se mohli dopustit té chyby, že jsme vliv nenašli, i když je možné, že existuje. Na druhé straně pokud by v každé z testovaných skupin bylo 10000 lidí a stále by se neprokázala existence závislosti, závěr „H_0 platí" by asi byl docela rozumný.

Snížení rozptylu zvyšuje sílu testu

V příkladu 2.15 lze vidět jednu celkem přirozenou skutečnost, že totiž výpovědní síla statistického testu se zvyšuje se zvýšením počtu měření. Pro připomenutí **síla testu** je pozitivní pojem – je to pravděpodobnost, se kterou správně zamítneme H_0, pokud platí H_1. Je to tedy jakási síla nalezení vlivu mezi proměnnými testu. Na tuto sílu má především vliv rozptyl $\sigma^2_{\overline{X}}$ neboli

$$\sigma^2_{\overline{X}} = \frac{\sigma^2}{N}. \tag{2.13}$$

Cokoli, co snižuje rozptyl $\sigma^2_{\overline{X}}$, zvyšuje současně i sílu testu. Jedním činitelem je počet měření N – je vidět ze vzorce, že pokud zvyšujeme N, zvyšuje se hodnota jmenovatele ve zlomku, a tím klesá hodnota rozptylu. Další možností, jak snížit rozptyl, je snížit přímo hodnotu rozptylu σ^2 celé populace v čitateli zlomku. Pokud si říkáte, že σ^2 je dáno a nelze je přece měnit, máte pravdu. Ale stejně některé vlivy na rozptyl σ^2 celé populace můžeme vyloučit vhodným naplánováním experimentu.

Příklad 2.16. Experimentem se má zjistit, zda alkohol v krvi má vliv na reakční dobu řidiče. Náhodně byly vybrány dvě skupiny lidí. První skupině je nabídnut alkohol ve formě ginu s tonikem, druhé skupině pouze tonik. Pak jsou všichni podrobeni měření reakční doby. Testovaný člověk sedí u stolu a před sebou má lampu s červenou žárovkou a tlačítko. Lampa se rozsvěcuje v různých nepravidelných intervalech – jakmile se rozsvítí, je úkolem testovaného co nejrychleji stisknout tlačítko. Je změřena jeho reakční doba (v milisekundách). U každého člověka se měření několikrát opakuje, a pak je vypočten průměr jeho reakční doby.

Samozřejmě uvnitř každé ze skupin se projeví určitá variabilita v průměrné době reakce. Ta je dána různými faktory, z nichž některé nemůžeme ovlivnit, ale jiné ano. Mezi **neovlivnitelné faktory** patří:

1. Nálada člověka během měření (špatná nálada = delší doba reakce).
2. Osobní předpoklady – někdo má prostě schopnost rychlejší reakce než ostatní.
3. Postoj člověka vůči experimentu (znuděný postoj = delší doba reakce).

Mezi **ovlivnitelné faktory** lze zahrnout

1. Teplotu v místnosti, kde se provádí měření (extrémní teploty $\Rightarrow$ delší doba reakce).
2. Vlhkost v místnosti, kde se provádí měření (vyšší vlhkost $\Rightarrow$ delší doba reakce).
3. Pohlaví (u žen … kratší doba reakce).
4. Věk (vyšší věk … kratší doba reakce).
5. Čas dne (měření pozdě odpoledne … delší doba reakce).

Některé faktory rozptylu hodnot můžeme významně ovlivnit – jak výběrem sledované populace (omezení se na stejné pohlaví a věk eliminuje rozdíly způsobené těmito faktory), tak zajištěním stejných podmínek měření (stálá vlhkost a teplota místnosti, měření u všech ve stejnou denní dobu). Tímto způsobem plánování a provedení experimentu pak následný statistický test získá větší výpovědní sílu v těch parametrech, které jsou pro nás důležité, a není zkreslen rozdíly v těch ovlivnitelných faktorech, které nás nezajímají.

2.5.4 Interval spolehlivosti pro střední hodnotu μ

Kromě statistických testů jsou při zpracování měření často užitečnější tzv. **intervaly spolehlivosti**. V této fázi výkladu se seznámíme s intervalem spolehlivosti pro střední hodnotu $E\overline{X} = \mu$ průměru z normálního rozdělení.

Interval spolehlivosti pro μ při známém rozptylu

Střední hodnotu μ měřené veličiny obyčejně neznáme. Kdybychom ji znali, nemusíme provádět ani měření, ani statistický test Určit μ je v podstatě naším cílem.

Jakýmsi odhadem střední hodnoty je výběrový průměr $\overline{X}$. Ovšem tento průměr (zejména při nižším počtu měření N) není přesně roven střední hodnotě μ – vlastně víme, že platí $P(\overline{X} = \mu) = 0$. Potřebovali bychom spíše najít nějaký interval $(\overline{X} - a; \overline{X} + a)$ pro nějaké vhodné a „ne příliš velké“ $a > 0$. A to bude vlastně interval spolehlivosti pro střední hodnotu μ. Přesněji řečeno, **r%-ní interval spolehlivosti pro střední hodnotu μ průměru $\overline{X}$ je takový interval, který obsahuje μ s pravděpodobností $\frac{r}{100}$.**

Příklad 2.17. Životnost 75-wattové žárovky má normální rozdělení se směrodatnou odchylkou $\sigma = 25$ hodin. U náhodně vybraného vzorku dvaceti žárovek byla naměřena průměrná životnost $\overline{x} = 1024$ hodin. Utvořte 95%-ní interval spolehlivosti pro střední hodnotu μ průměrné životnosti.

Řešení tohoto příkladu je instruktivní, proto si dovolím vypnout kurzívu (:-)). Každý interval spolehlivosti je velmi úzce svázán se statistickým testem. Budeme se zabývat zejména oboustrannými intervaly spolehlivosti, proto v našem příkladu je zde vazba na oboustranný U-test s hypotézami

H_0: $\mu = \mu_0$ (skutečnou střední hodnotu μ_0 neznáme).

H_1: $\mu \neq \mu_0$.

Pokud hledáme 95%-ní interval spolehlivosti, je příslušná hladina významnosti $\alpha = 0{,}05$. Kritériem testu je výběrový průměr $\overline{X}$. Dále vezmeme v úvahu počet měření pro výpočet průměru:

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N}} = \frac{25}{\sqrt{20}} = 5{,}59.$$

Jak je dobře známo z předmětu Matematika 3, příslušné kritické hodnoty v tomto případě jsou

$$u_{\frac{\alpha}{2}} = -1{,}96, \qquad u_{1-\frac{\alpha}{2}} = 1{,}96.$$

Obě tyto hodnoty lze převést na kritické hodnoty vzhledem k veličině $\overline{X}$:

$$-1{,}96 = \frac{\overline{x}_m - \mu_0}{5{,}59} \quad \Rightarrow \quad \overline{x}_m = \mu_0 - 1{,}96 \cdot 5{,}59 = \mu_0 - 10{,}96;$$

$$1{,}96 = \frac{\overline{x}_v - \mu_0}{5{,}59} \quad \Rightarrow \quad \overline{x}_v = \mu_0 + 1{,}96 \cdot 5{,}59 = \mu_0 + 10{,}96.$$

Interval pro „nezamítnutí H_0“ je pak

$$\overline{X} \in (\mu_0 - 10{,}96; \mu_0 + 10{,}96)\,. \tag{2.14}$$

Ovšem hodnotu μ_0 neznáme. Ale pokud ze vzorce 2.14 vyjádříme místo $\overline{X}$ hodnotu μ_0, dostaneme vztah pro interval spolehlivosti:

$$\mu_0 \in \left(\overline{X} - 10{,}96; \overline{X} + 10{,}96\right)\,. \tag{2.15}$$

Odtud pro jednu realizaci $\overline{x} = 1024$ dostaneme

$$\mu_0 \in (1024 - 10{,}96; 1024 + 10{,}96) = (1013{,}04; 1034{,}96),$$

se spolehlivostí 0,95, což je odpověď příkladu 2.17.

Obecně s využitím symetrie

$$u_{\frac{\alpha}{2}} = -u_{1-\frac{\alpha}{2}}$$

lze $(1-\alpha)\cdot 100$%-ní interval spolehlivosti pro střední hodnotu μ při známém rozptylu psát ve tvaru

$$\mu \in \left(\overline{X} - u_{1-\frac{\alpha}{2}} \cdot \sigma_{\overline{X}}; \overline{X} + u_{1-\frac{\alpha}{2}} \cdot \sigma_{\overline{X}}\right)\,. \tag{2.16}$$

V některé literatuře se objevuje kromě pojmu „interval spolehlivosti“ ještě pojem „konfidenční interval“ – to je jen nesprávný (nebo jiný) překlad anglického termínu „confidence interval“, ale jedná se o totéž (překladatelé se zase jednou nedomluvili … :-)).

Pojem intervalu spolehlivosti úzce souvisí se silou příslušného statistického testu – jak už to vyplývá ze vzorce 2.16, čím větší je síla příslušného statistického testu, tím menší je rozptyl $\sigma^2_{\overline{X}}$, a tím užší je interval spolehlivosti.

Interval spolehlivosti pro μ při neznámém rozptylu

Příklad 2.18. Uvažujme stejnou situaci jako v příkladu 2.17 ($N = 20$, $\overline{x} = 1024$), pouze odchylka σ není známá a musíme ji odhadnout – tedy z měření životnosti u dvaceti žárovek byl vypočten rozptyl $\overline{s^2} = 625$. Odtud

$$est\sigma^2_{\overline{X}} = \frac{\overline{s^2}}{N} = \frac{625}{20} = 31{,}25.$$

Tedy $est\sigma_{\overline{X}} = \sqrt{31{,}25} = 5,59$ – toto číslo je stejné jako v příkladu 2.17, ovšem nyní se nejedná o přesnou hodnotu, ale odhad, čili zde bude $\nu = N - 1 = 19$ stupňů volnosti odhadu. Nalezněte 95%-ní interval spolehlivosti pro střední hodnotu μ výběrového průměru $\overline{X}$.

V příslušném statistickém testu použijeme nyní t-rozdělení – nalezneme kritické hodnoty pro oboustanný test ($\alpha = 2q = 0{,}05$) a pro $\nu = 19$ stupňů volnosti (tabulka 2.1): $t_v = 2{,}093$, odtud $t_m = -2{,}093$.

Hypotézy H_0, H_1 a kritérium oboustranného testu zde zůstávají stejné jako v 2.17: $H_0 : \mu = \mu_0$, $H_1 : \mu \neq \mu_0$, kritériem je výběrový průměr $\overline{X}$. Přepočteme nyní kritické hodnoty vzhledem k veličině $\overline{X}$:

$$-2{,}093 = \frac{\overline{x}_m - \mu_0}{5{,}59} \quad \Rightarrow \quad \overline{x}_m = \mu_0 - 2{,}093 \cdot 5{,}59 = \mu_0 - 11{,}7;$$

$$2{,}093 = \frac{\overline{x}_v - \mu_0}{5{,}59} \quad \Rightarrow \quad \overline{x}_v = \mu_0 + 2{,}093 \cdot 5{,}59 = \mu_0 + 11{,}7.$$

Nyní interval pro „nezamítnutí H_0“ je

$$\overline{X} \in (\mu_0 - 11{,}7; \mu_0 + 11{,}7)\,, \tag{2.17}$$

nás ovšem zajímá spíše tvar (rychle jej naň převedem)

$$\mu_0 \in \left(\overline{X} - 11{,}7; \overline{X} + 11{,}7\right). \tag{2.18}$$

Odtud pro příklad 2.18 se spolehlivostí 0,95 a realizaci (= konkrétní měření) průměru $\overline{x} = 1024$ dostaneme

$$\mu_0 \in (1024 - 11{,}7; 1024 + 11{,}7) = (1012{,}3; 1035{,}7),$$

Vidíme, že při větší míře nejasnosti, kdy rozptyl neznáme a musíme jej odhadnout, je interval spolehlivosti širší (ostatní hodnoty v příkladech 2.17 a 2.18 jsou stejné, takže lze opravdu provést porovnání).

Obecně s využitím symetrie

$$t_{\frac{\alpha}{2}} = -t_{1-\frac{\alpha}{2}}$$

lze $(1-\alpha)\cdot 100\%$-ní interval spolehlivosti pro střední hodnotu μ při neznámém rozptylu psát ve tvaru

$$\mu \in \left(\overline{X} - t_{1-\frac{\alpha}{2}} \cdot est\sigma_{\overline{X}}; \overline{X} + t_{1-\frac{\alpha}{2}} \cdot est\sigma_{\overline{X}}\right). \tag{2.19}$$

Několik důležitých poznámek k intervalům spolehlivosti

Následuje několik poznámek k intervalům spolehlivosti, každá z nich je důležitější než ty ostatní ... :-)

a) **Vztah mezi intervalem spolehlivosti a pravděpodobností.** Je potřeba říci něco velmi důležitého k formulaci „95%-ní interval spolehlivosti pro střední hodnotu μ obsahuje tuto hodnotu s pravděpodobností 0,95“. Věta je vyslovena správně ale může být zavádějící – a problém pochopení souvisí s pojmem **statistika** a **realizace statistiky** (viz definice 2.6). Interval spolehlivosti je totiž vlastně intervalem, jehož mezemi jsou náhodné veličiny. A pak pro konkrétní měření dosadíme do vzorců 2.16, 2.19 konkrétní realizaci $\overline{x}$. Třeba v příkladu 2.17 interval $(1013{,}04; 1034, 96)$ neznámou střední hodnotu μ buď obsahuje (stoprocentně), nebo neobsahuje (stoprocentně).

   Souvislost s pravděpodobností se projeví pouze při opakovaném měření a opakované konstrukci realizace intervalu spolehlivosti: pokud bychom například tisíckrát zopakovali experiment změření životnosti u dvaceti žárovek, spočetli tisíc různých realizací $\overline{x}_1$, $\overline{x}_2 \ldots$, $\overline{x}_{1000}$ a sestrojili tisíc různých realizací intervalu spolehlivosti podle vzorce 2.16, tak přibližně 95% těchto intervalů (tj. asi 950 z nich) bude neznámou hodnotu μ obsahovat, zbylých pět procent ji obsahovat nebude.

   V dalších výpočtech realizací intervalů spolehlivosti pro $\overline{x}$ budu slovo „realizace“ vynechávat, takže pojem **interval spolehlivosti** se středem v $\overline{X}$ (středem intervalu je nikoli konkrétní hodnota, ale náhodná veličina) bude splývat s pojmem **realizace intervalu spolehlivosti** se středem v $\overline{x}$ (středem intervalu je konkrétní číslo). Matematicky je mezi těmito dvěma pojmy rozdíl, ale v praxi pod intervalem spolehlivosti máme na mysli vždy konkrétní interval vypočtený na základě měření. !!!! :-)

b) **Jednostranné intervaly spolehlivosti.** Pozorný čtenář by možná mohl položit otázku: no dobrá, oboustranný interval spolehlivosti odpovídá jistému oboustrannému statistickému testu (příklad 2.17) – pokud se ale týká jednostranných testů (například test grafické iluze 2.5.3), existuje také u nich nějaký analogický jednostranný interval spolehlivosti? Odpověď zní „ano, existuje“. Jeho odvození je analogické, proveďme jej například pro data testu grafické iluze 2.5.3:

$H_0 : \mu = \mu_0$, $H_1 : \mu > \mu_0$, kritériem je $\overline{X}$, kritickou hodnotu použijeme přesně tu stejnou jako v daném pravostranném testu: $t_k = 2{,}132$. Převodem normovaného tvaru kritické hodnoty do tvaru aktuálního vzhledem k veličině $\overline{X}$ dostaneme

$$2{,}132 = \frac{\overline{x}_k - \mu_0}{1} \quad \Rightarrow \overline{x}_k = \mu_0 + 2{,}132 \cdot 1 = \mu_0 + 2,132;$$

tedy interval pro „nezamítnutí H_0“ je

$$\overline{X} \in (-\infty; \mu_0 + 2{,}132), \tag{2.20}$$

ale protože nás zajímá spíše interval pro μ_0, tak z tohoto vztahu vyjádříme ohraničení pro μ_0 a index 0 vypustíme:

$$\mu \in (\overline{X} - 2{,}132; \infty) = (5{,}868; \infty) \tag{2.21}$$

a to je hledaný interval se spolehlivostí 0,95. Vidíme, že **oboru 2.20 pravostranného testu odpovídá levostranný interval spolehlivosti 2.21**.

I když jsou tedy jednostranné intervaly spolehlivosti zcela přirozené a možné, v dalším textu se na ně nebudeme příliš zaměřovat – spíše nás bude zajímat vymezení pro μ na intervalu konečné délky. Tedy i když budeme provádět jednostranné statistické testy, intervaly spolehlivosti budeme hledat na základě příslušného oboustranného testu se stejnou hladinou významnosti.

Tak tedy i v příkladu 2.5.3 lze sestrojit oboustranný interval spolehlivosti: Pro $\nu = 4$ je kritická hodnota rovna

$$t_{1-\frac{0{,}05}{2}}(4) = 2{,}776;$$

Pak (vzorec 2.19):

$$\mu \in (8 - 2{,}776 \cdot 1; 8 + 2{,}776 \cdot 1) = (5{,}224; 10{,}776).$$

c) **Vztah mezi intervalem spolehlivosti a statistickým testem.** Mezi výsledkem statistického testu a intervalem spolehlivosti existuje jednoznačný vztah:

Statistický test zamítne hypotézu $H_0 : \mu = \mu_0$ právě tehdy, když μ_0 nenáleží do příslušného intervalu spolehlivosti

Třeba pokud bychom v situaci příkladu 2.17 testovali hypotézu $H_0 : \mu = 1000$, místo abychom prováděli test, stačí se podívat na příslušný interval spolehlivosti oboustranného testu: $1000 \notin (1013{,}04; 1034{,}96)$, tj. to znamená, že příslušný statistický test zamítne hypotézu H_0.

Nebo podíváme-li se na příklad pravostranného testu 2.5.3, kde $H_0 : \mu = 6$, vidíme, že $6 \in (5{,}868; \infty)$ (levostranný interval spolehlivosti vypočten v předchozí poznámce b)), tj. to znamená, že hypotéza H_0 příslušného jednostranného testu nebude zamítnuta.

d) **Interval spolehlivosti uvádí více informací než statistický test.** Omlouvám se za další členění, ale budu prezentovat tuto poznámku ve čtyřech myšlenkách:

1. Výsledek statistického testu není přesně to, co bychom chtěli znát. Ve skutečnosti chceme znát míru platnosti hypotézy H_1, je-li dán výsledek experimentu – ovšem místo toho se ze statistického testu dovídáme pravděpodobnost výsledku experimentu za předpokladu platnosti hypotézy H_0 (a stále neznáme **míru platnosti** H_1, a to ani v případě, kdy je H_0 zamítnuto).
2. Statistický test nám ve svém výsledku nedává informaci o rozsahu studovaného vlivu, kdežto interval spolehlivosti ano. Informace o umístění střední hodnoty μ je ve statistických testech, které jsme dosud prováděli, skryta, ale interval spolehlivosti činí tuto informaci zjevnou a překládá ji do srozumitelného měřítka.

   Například test 2.5.3 pouze prohlásí, že se nenašlo dostatek důkazů pro vliv grafické iluze na μ. Ale interval spolehlivosti (nyní už spíše ten oboustranný, tj. pro $\alpha = 0{,}05$ byl odvozen v poznámce b)) $(5{,}224; 10{,}776)$ navíc naznačuje, že se spolehlivostí 0,95 by střední hodnota místo šesti mohla být stejně dobře rovna i osmi nebo deseti, ale už ne jedenácti nebo dvanácti.
3. Statistický test zdůrazňuje chybu prvního druhu, ale říká velmi málo o chybě druhého druhu. Na druhé straně interval spolehlivosti naznačuje i chybu druhého druhu – pokud α necháme pevné a zmenšujeme β, tak interval spolehlivosti zmenšuje svou délku.
4. Závěr: Z uvedených důvodů je lepší používat intervaly spolehlivosti spíše než jen pouhé testování hypotéz. Někdy statistické zpracování v literatuře příliš zdůrazňuje testy a zapomíná dodat užitečné informace navíc, které lze snadno vyčíst z intervalů spolehlivosti.

2.5.5 t-test typu „$\mu_1 = \mu_2$"

Párový test

V tomto oddílu se budeme zabývat statistickými testy při experimentech, kde získáváme dva soubory měření. Zde je potřeba si dát pozor na vztah mezi těmito dvěma soubory (= skupinami) měření, na základě tohoto vztahu rozlišujeme totiž dva typy statistického testu – párový a nepárový test. Párovým testem se budeme zabývat nejdříve – spočívá v tom, že sice získáme dvě skupiny (= dva soubory) měření, ale tyto soubory jsou navzájem těsně svázány v tom smyslu, že ke každé hodnotě v prvním souboru měření lze jednoznačně přiřadit tzv. párovou hodnotu měření ze druhého souboru. Zejména to taky znamená, že počet měření v obou souborech je stejný – a v podstatě bychom mohli říct, že místo dvou souborů měření máme jediný soubor, ve kterém jedna položka je reprezentována uspořádanou dvojicí hodnot.

Párový test tedy užijeme v situaci, kdy sice máme k dispozici dva soubory měření, ale tyto dva soubory měření jsou spolu těsně svázány – obyčejně tak, že v obou skupinách jsou hodnoceni stejní jedinci; nejprve provedeme měření vybrané skupiny jedinců za systému podmínek A, a pak provedeme měření téže skupiny jedinců za systému podmínek B. Proto se tomuto typu experimentů také říká **experiment opakovaného měření**. Další vhodný název je zde experiment typu „**jedna skupina dvakrát**", protože jedna skupina jedinců je podrobena měření při dvou různých situacích.

Příklad 2.19. Chceme experimentem zjistit, jak se změní počet úderů srdce člověka za minutu po vypití šálku kávy (studujeme vliv kofeinu na činnost srdce). Kdybychom k tomuto experimentu přistupovali „nepárovým" přístupem a vybrali náhodně dvě skupiny lidí, z nichž jedna by vypila kávu s kofeinem a druhá kávu bez kofeinu, do měření by byl zanesen jistý rozptyl způsobený tím, že tempo srdečních úderů se liší u různých lidí.

Mnohem vhodnější je zde experiment opakovaného měření, kdy jedna a táž skupina vybraných lidí je vystavena měření po kávě bez kofeinu, a pak po kávě s kofeinem. Potom se vypočte rozdíl obou hodnot vždy u téhož člověka a testuje se, zda je tento rozdíl významný. Byla získána následující data počtu srdečních úderů za minutu u devíti lidí (na každém řádku jsou hodnoty měření jednoho člověka):

x_{i1} bez kofeinu	x_{i2} s kofeinem	rozdíl $x_{i2} - x_{i1}$
70	76	6
60	61	1
49	52	3
72	71	-1
70	81	11
66	70	4
55	55	0
54	61	7
80	89	9

První dva sloupečky tabulky představují dva soubory měření párového testu. My ovšem budeme dále pracovat jen s jednorozměrným souborem, a sice s vektorem rozdílů měření ve třetím sloupci. To znamená, že párový test vlastně odpovídá situaci jednorozměrného souboru měření (= oddílu 2.5.2).

Spočteme průměr $\overline{x} = 4{,}44$ a odhadneme rozptyl pomocí $\overline{s^2}$. Aby čtenář neměl pocit, že v oddílu 2.5.5 nebude vzhledem k 2.5.2 nic nového, vypočteme $\overline{s^2}$ pomocí pátého možného vzorce, který jsme si ještě neuváděli:

$$\overline{S^2} = \frac{SS}{\nu} \tag{2.22}$$

(ono se vlastně jedná o vzorec 2.12, ovšem ve jmenovateli vzorce je ν místo $N-1$). Dále do čitatele za SS dosadíme ze vzorce 2.11, který zkombinujeme s nejpohodlnějším způsobem výpočtu 2.8:

$$SS = N \cdot S^2 = \left(\sum_{i=1}^{N} X_i^2\right) - \frac{\left(\sum_{i=1}^{N} X_i\right)^2}{N},$$

čili pak pro realizaci ss platí

$$ss = \left(\sum_{i=1}^{N} x_i^2\right) - \frac{\left(\sum_{i=1}^{N} x_i\right)^2}{N}, \tag{2.23}$$

a po dosazení

$$ss = 314 - \frac{40^2}{9} = 136{,}22.$$

Počet stupňů volnosti pro odhad rozptylu je $\nu = N - 1 = 8$, a tedy

$$\overline{s^2} = \frac{ss}{\nu} = 17{,}03.$$

a) Sestrojme nyní 95**%-ní interval spolehlivosti** pro μ: t_k najdeme jako průsečík řádku $\nu = 8$ a sloupce $2q = \alpha = 0{,}05$ (oboustranný interval spolehlivosti vychází z oboustranného testu): $t_k = 2{,}306$. Pak interval pro μ se spolehlivostí 0,95 je

$$\overline{x} \pm \sqrt{\frac{\overline{s^2}}{N}} \cdot t_k(\nu = 8) = 4{,}44 \pm \sqrt{\frac{17{,}03}{9}} \cdot 2{,}306 = (1{,}27; 7{,}61).$$

Z tohoto intervalu spolehlivosti se dovídáme, že kofein zvyšuje činnost srdce, a sice o 1,27 až o 7,61 úderů za minutu.

b) Provedeme i statistický t-test:

(K1) H_0: $\mu = 0$ (rozdíl hodnot je nulový, tj. počet úderů srdce za minutu je stejný s kofeinem i bez kofeinu – srdeční tep nezávisí na kofeinu);
H_1: $\mu \neq 0$.

(K2) Testovým kritériem bude výběrový průměr $\overline{X}$, respektive jeho normovaná hodnota $\frac{\overline{X}-0}{est\ \sigma_{\overline{X}}}$.

(K3) $\overline{s^2} = 17{,}03 \Rightarrow$

$$est\ \sigma^2_{\overline{X}} = \frac{\overline{s^2}}{N} = \frac{17{,}03}{9} = 1{,}89 \Rightarrow est\ \sigma_{\overline{X}} = \sqrt{1{,}89} = 1{,}37.$$

Tj. za předpokladu platnosti H_0 má veličina $\frac{\overline{X}-0}{1{,}37}$ Studentovo t-rozdělení s $\nu = 8$ stupni volnosti.

(K4) $t_k = \pm 2{,}306$ je u oboustranného testu stejné jako u intervalu spolehlivosti a).

(K5) Příslušná t-hodnota je

$$\frac{4{,}44 - 0}{1{,}37} = 3{,}24 > 2{,}306,$$

a tedy zamítáme H_0 o nezávislosti, je potvrzen vliv kofeinu na zvýšení srdečního tepu.

Hypotézy H_0 nejsou pravdivé téměř nikdy (pokud uvažujeme větší počet desetinných míst), tj. zamítnutí H_0 nám nic podstatného neříká. Spíše nás zajímalo, jak velký je vliv kofeinu, a to jsme se dozvěděli z intervalu spolehlivosti.

Nepárový test

Nyní se pojďme věnovat **nepárovému testu** neboli zpracování dat při experimentu typu „**dvě skupiny jednou**“.

Příklad 2.20. Chceme zjistit kvalitu jisté techniky pamatování. Náhodně jsme vybrali deset lidí a rozdělili do dvou skupin po pěti lidech. Skupina 1 (tzv. experimentální skupina) se naučila 100 zadaných slov novou technikou, skupina 2 (kontrolní skupina … zažitý nesprávný překlad anglického „control group“, správný význam překladu je „řízená skupina“, protože „control“= řídit, vést, nikoli kontrolovat) použila obyčejnou klasickou techniku zapamatování. Po týdnu se vyzkouší, kolik si kdo pamatuje z daných 100 slov – jsou získána data

experimentální skupina	kontrolní skupina
43	16
37	22
51	24
27	30
32	18

Nyní data na jednom řádku spolu nijak nesouvisí, jedná se o měření u dvou různých lidí. Zdá se, že experimentální skupina má lepší výsledky, ale musíme statisticky prokázat, že to není způsobeno pouze náhodnými vlivy.

V každé ze skupin vypočteme průměr a odhadneme rozptyl.

skupina 1: $\overline{x}_1 = 38$, podle vzorce 2.23 $ss_1 = 352$, $\nu_1 = 4$, a tedy podle 2.22

$$est_1\, \sigma^2 = \overline{s_1^2} = \frac{352}{4} = 88.$$

skupina 2: $\overline{x}_2 = 22$, podle vzorce 2.23 $ss_2 = 120$, $\nu_2 = 4$, a tedy podle 2.22

$$est_2\, \sigma^2 = \overline{s_2^2} = \frac{120}{4} = 30.$$

No a teď přicházíme k úvahám, které jsou pro další vývoj (i další kapitolu) důležité. V situaci experimentu typu „dvě skupiny jednou“ je potřeba se vypořádat se dvěma typy rozptylu, kterými jsou – **vnitřní rozptyl** a **vnější rozptyl**.

vnitřní rozptyl: Jedná se o rozptyl představující vzájemnou rozdílnost jedinců v celé populaci (např. rozdílnost lidí, rozdílnost různých součástek stejného typu, apod.). V tomto textu se budeme převážně zabývat situacemi, kdy je splněn tzv. **princip homogenního vnitřního rozptylu: jakýkoliv experiment nemá vliv na rozptyl rozdělení celé populace, z níž byla náhodně vybrána skupina jedinců pro měření.**

Slovy našeho příkladu, rozdílnost výsledků est_1 způsobená růzností lidí ve skupině 1 je přibližně stejná jako rozdílnost výsledků est_2 způsobená růzností lidí ve skupině 2. Jinak řečeno, ať už měříte daný soubor měření za jakékoli podmínky, „rozmanitost“ těchto měření je v dané skupině přibližně stejná.

Z Tohoto principu homogenního rozptylu tedy plyne, že odhady est_1, est_2 jsou odhady jednoho a stejného vnitřního rozptylu σ^2 (díky tomu jsem u písmenka σ o pár řádků výše už nepsal žádný index), který vypočteme jako aritmetický průměr obou odhadů:

$$est\ \sigma^2 = \frac{est_1\ \sigma^2 + est_2\ \sigma^2}{2} = \frac{88 + 30}{2} = 59.$$

Tedy nejlepší možný odhad vnitřního rozptylu σ^2 je roven 59 a v dalším budeme pracovat s ním. Počet stupňů volnosti je $\nu = 4 + 4 = 8$, protože jsme tento odhad získali pomocí dvou jiných odhadů o počtu stupňů volnosti 4 – stupně volnosti ve výsledném odhadu sečítáme.

vnější rozptyl: Jedná se o **rozptyl vyjadřující rozdílnost mezi danými podmínkami měření**. Tento rozptyl v případě dvou souborů měření lze odhadnout na základě průměrů $\overline{x}_1 = 38$, $\overline{x}_2 = 22$. Způsob je následující: Vypočteme rozptyl těchto dvou průměrů podle vzorců 2.23 a 2.22: Protože $\nu = 2 - 1 = 1$, máme

$$est\ r_N = \frac{ss}{1} = \frac{38^2 + 22^2 - \frac{60^2}{2}}{1} = 128.$$

Ale to ještě není všechno – tento odhad je odhadem rozptylu průměru pěti hodnot ($N = 5$). Abychom získali odhad rozptylu jediné hodnoty měření, musíme v souladu se vzorcem $est\ r_N = \frac{est\ r}{N}$ (analogie vzorce 2.13) počítat

$$est\ r = N \cdot est\ r_N = 5 \cdot 128 = 640.$$

celkový rozptyl: Aby byla plejáda přehledu rozptylů úplná, je možné si položit následující otázku – co se vlastně spočítá, když budeme považovat všech deset hodnot za měření jediné veličiny a vypočteme $\overline{s^2}$ ze všech deseti měření? Podle logiky výpočtu by to měl být jakýsi celkový rozptyl – a skutečně je tomu tak.

Průměr všech deseti hodnot je $\overline{x} = 30$, což je mimochodem aritmetický průměr hodnot $\overline{x}_1$, $\overline{x}_2$. Bude tedy podobně hodnota $\overline{s^2}$ průměrem hodnot $\overline{s_1^2}$, $\overline{s_2^2}$?

Podle vzorce 2.23 $ss = 1112$, a tedy podle 2.22

$$\overline{s^2} = \frac{ss}{\nu} = \frac{1112}{9} \doteq 123{,}56,$$

což není hodnota rovná součtu $\overline{s_1^2} + \overline{s_2^2} = 88 + 30 = 118$. Ještě méně se zdá, že by odhad celkového rozptylu 123,56 byl součtem odhadu vnitřního rozptylu 59 a odhadu vnějšího rozptylu 640 – ale přesto zde platí jisté součtové vzorce:

a) Celkový počet stupňů volnosti 9 u odhadu celkového rozptylu je součtem volnosti 8 u odhadu vnitřního rozptylu a volnosti 1 u odhadu vnějšího rozptylu.

b) Celkový součet čtverců 1112 je součtem součtu čtverců 472 vnitřního rozptylu (který vznikl součtem $ss_1 = 352$ a $ss_2 = 120$) a 640 u vnějšího rozptylu.

Tolik přehled a jemné intro do problematiky rozptylu – více na to téma bude řečeno v další kapitole. Nyní se vraťme k řešení našeho příkladu.

Odhad vnitřního rozptylu je est $\sigma^2 = 59$, *odtud odhad rozptylu průměru*

$$est\sigma^2_{\overline{X}} = \frac{59}{5} = 11{,}8.$$

Nyní lze užitím 2.19 určit např. 95 %*-ní interval spolehlivosti pro střední hodnotu průměru v každé ze skupin měření: Pro* $t_k(\nu = 8; \alpha = 2q = 0{,}05) = 2{,}306$

$$\mu_1 \in 38 \pm \sqrt{11{,}8} \cdot 2{,}306 \doteq (30{,}08; 45{,}92),$$

$$\mu_2 \in 22 \pm \sqrt{11{,}8} \cdot 2{,}306 \doteq (14{,}08; 29{,}92).$$

V souvislosti se statistickým testem tohoto příkladu nás ovšem spíše zajímá interval spolehlivosti pro střední hodnotu veličiny $\overline{X_1} - \overline{X_2}$*. Střed intervalů spolehlivosti bude v tomto případě* $\overline{x_1} - \overline{x_2} = 38 - 22 = 16$, *odhad rozptylu je*

$$est\ \sigma^2_{\overline{X_1}-\overline{X_2}} = est\ \sigma^2_{\overline{X_1}} + est\ \sigma^2_{\overline{X_2}} = \frac{59}{5} + \frac{59}{5} = 23{,}6.$$

Tedy pro $t_k(\nu = 8; \alpha = 2q = 0{,}05) = 2{,}306$

$$\mu_1 - \mu_2 \in 16 \pm \sqrt{23{,}6} \cdot 2{,}306 \doteq 16 \pm 11{,}2 = (4{,}8; 27{,}2).$$

Protože výsledný interval spolehlivosti pro rozdíl středních hodnot neobsahuje nulu, víme také, že příslušný statistický test zamítne hypotézu $H_0 : \mu_1 = \mu_2$*. A skutečně, pojďte se přesvědčit:*

(K1) H_0*:* $\mu_1 = \mu_2$ *(střední hodnoty obou skupin ohodnocení jsou stejné, tj. nová technika zapamatování ovlivňuje výsledek přibližně stejně jako ta dřívější).*
H_1*:* $\mu_1 \neq \mu_2$ *(nová technika zapamatování má přibližně stejné výsledky jako dřívější technika).*

(K2) *Testovým kritériem bude rozdíl* $\overline{X_1} - \overline{X_2}$.

(K3) *Při platnosti* H_0 *má veličina*

$$\frac{\overline{X_1} - \overline{X_2} - 0}{\sqrt{est\ \sigma^2_{\overline{X_1}-\overline{X_2}}}} = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{23{,}6}}$$

Studentovo t-rozdělení pro $\nu = 8$.

(K4) *Pro* $\alpha = 0{,}05$ *příslušná kritická hodnota* t_k *je na průsečíku řádku* $\nu = 8$ *a sloupce pro* $\alpha = 2q = 0{,}05$, *tj.* $t_k = 2{,}306$.

(K5) $\frac{38-22}{4{,}858} = 3{,}29 \notin (-t_k; t_k)$. *$H_0$ tedy zamítáme – nová technika významně zvyšuje úroveň zapamatování.*

Příklad 2.21. Psychologové fyziologie chtějí experimentem ověřit, že podvěsek mozkový je hlavním řídícím centrem sexuálního chování. Rozhodli se získat dvacet dobrovolníků z řad studentů FEKT, kteří by se chtěli podrobit operaci mozku. Protože žádní dobrovolníci se nepřihlásili (zejména do experimentální skupiny), byly náhodně vybrány dvě skupiny po deseti krysách.

Krysám z experimentální skupiny byl operací odebrán podvěsek mozkový. Krysám z kontolní skupiny byla pouze otevřena lebka, ale nic nebylo odebráno (aby byl snížen rozptyl způsobený otevřením lebky).

Protože operaci prováděl nezkušený student, některé z krys zahynuly přímo na operačním stole. V experimentální skupině přežilo pět krys, v kontrolní skupině sedm. Psycholové byli rozmrzelí nad nezkušeným studentem, ale pokračovali dále v experimentu. Byla získána data o počtu pohlavních spojení v průběhu jistého časového intervalu. V experimentální skupině ($N_1 = 5$): $0, 1, 4, 4, 1$. Odtud $\overline{x_1} = 2$, $ss_1 = \sum x_i^2 - \frac{(\sum x_i)^2}{N_1} = 14$, $\nu_1 = N_1 - 1 = 4$, pak $\overline{s_1^2} = \frac{ss_1}{\nu_1} = \frac{14}{4} = 3{,}5$.

V kontrolní skupině ($N_2 = 7$): $5, 7, 4, 3, 4, 6, 6$. Odtud $\overline{x_2} = 5$, $ss_2 = \sum x_i^2 - \frac{(\sum x_i)^2}{N_2} = 12$, $\nu_2 = N_2 - 1 = 6$, pak $\overline{s_2^2} = \frac{ss_2}{\nu_2} = \frac{12}{6} = 2$.

Odhad vnitřního rozptylu: Podobně jako u předchozího příkladu (vyjdeme z platnosti principu homogenního rozptylu), i nyní vypočteme jakýsi jeden odhad rozptylu jako průměr odhadů $\overline{s_1^2}$, $\overline{s_2^2}$ – nebude se jednat ovšem o aritmetický průměr, ale o tzv. **vážený průměr**. Protože odhad $\overline{s_2^2}$ byl sestaven na základě většího počtu stupňů volnosti (většího počtu měření), budeme mu přikládat větší váhu:

$$est\ \sigma^2 = \frac{\nu_1}{\nu_1 + \nu_2} \cdot \overline{s_1^2} + \frac{\nu_2}{\nu_1 + \nu_2} \cdot \overline{s_2^2} \tag{2.24}$$

V našem příkladu $est\ \sigma^2 = \frac{4}{10} \cdot 3{,}5 + \frac{6}{10} \cdot 2 = 2{,}6$. Pak intervaly spolehlivosti pro jednotlivé střední hodnoty a $t_k(\nu = 10, \alpha = 2q = 0{,}05) = 2{,}228$ jsou

$$\mu_1 \in 2 \pm \sqrt{\frac{2{,}6}{5}} \cdot 2{,}228 \doteq 2 \pm 1{,}61 = (0{,}39; 3{,}61);$$

$$\mu_2 \in 5 \pm \sqrt{\frac{2{,}6}{7}} \cdot 2{,}228 \doteq 2 \pm 1{,}36 = (3{,}64; 6{,}36).$$

Intervaly nemají společný průnik, což znamená, že testovaná hypotéza o rovnosti středních hodnot bude zamítnuta. Dále je možné si všimnout, že interval spolehlivosti pro μ_2 má menší délku než interval pro μ_1 – to je dáno větším počtem měření ve druhé skupině, pak totiž ve druhé skupině při odhadu rozptylu průměru dělíme sedmi, nikoli pěti.

V testu, který bude následovat, budeme používat rozdělení $\overline{X_1} - \overline{X_2}$. Pokud bychom chtěli najít 95%-ní interval spolehlivosti pro rozdíl $\mu_1 - \mu_2$, použijeme střed $\overline{x_1} - \overline{x_2} = 2 - 5 = -3$

a odhad rozptylu

$$est\ \sigma^2_{\overline{X_1}-\overline{X_2}} = est\ \sigma^2_{\overline{X_1}} + est\ \sigma^2_{\overline{X_2}} = \frac{2{,}6}{5} + \frac{2{,}6}{7} \doteq 0{,}891.$$

Pak

$$\mu_1 - \mu_2 \in -3 \pm \sqrt{0{,}891} \cdot 2{,}228 \doteq -3 \pm 2{,}1 = (-5{,}1; -0{,}9).$$

Příslušný statistický test:

(K1) H_0: $\mu_1 = \mu_2$ (odstranění podvěsku mozkového nemá vliv na sexuální chování);
H_1: $\mu_1 < \mu_2$ (odstranění podvěsku mozkového povede ke snížení sexuální aktivity).

(K2) Testovým kritériem bude veličina $\overline{X_1} - \overline{X_2}$.

(K3) Při platnosti H_0 má veličina

$$\frac{\overline{X_1} - \overline{X_2} - 0}{\sqrt{est\ \sigma^2_{\overline{X_1}-\overline{X_2}}}} = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{0{,}891}}$$

Studentovo t-rozdělení pro $\nu = 10$.

(K4) Pro $\alpha = 0{,}05$ příslušná kritická hodnota t_k je na průsečíku řádku $\nu = 10$ a sloupce pro $\alpha = q = 0{,}05$, tj. $t_k = 1{,}812$. **Pozor, intervaly spolehlivosti budeme vždy konstruovat pro $\alpha = 2q$, ovšem u jednostranného statistického testu musíme vzít hodnotu t_k pro $\alpha = q$!!**

(K5) Odpovídající t-hodnota kritéria je $\frac{-3}{\sqrt{0{,}891}} = -3{,}178 \notin (-t_k; t_k)$. H_0 tedy zamítáme.

2.5.6 Předpoklady použitelnosti parametrických testů

Při odvozování testů (zejména t-testu – odvození je mimo rámec tohoto kursu) muselo být učiněno několi předpokladů:

1. Naměřené hodnoty x_i jsou navzájem nezávislé (= předpoklad nezávislosti) – například předpoklad nezávislosti v příkladu 2.20 je porušen, pokud členové skupiny podvádějí a opisují jeden od druhého.

2. Měřená veličina má normální rozdělení (= předpoklad normality).

3. Rozptyl uvnitř experimentální skupiny je stejný jako rozptyl uvnitř kontrolní skupiny, tj. oba rozptyly jsou odhadem stejného (vnitřního) rozptylu σ^2 celé populace (= princip homogenního rozptylu).

Testy, které splňují uvedené tři předpoklady, se nazývají **parametrické testy**. Pokud některý z předpokladů není splněn, kritérium použité v testu nemá rozdělení t (nebo U, pokud známe rozptyl σ^2), a tudíž nelze určit kritické hodnoty – pro srovnání středních hodnot se v tom případě užívají tzv. **neparametrické testy**, o nichž bude blíže řeč v kapitole 7.

Úplná diskuse použitelnosti parametrických testů U, t (a v následující kapitole testu F) je mimo rámec tohoto kursu. Ovšem následující poznámky mohou být důležitým vodítkem, který pro běžného „uživatele“ statistiky dostačuje:

a) Mějte se na pozoru pouze tehdy, když t-hodnota kritéria je blízká hodnotě t_k. Předpoklady by musely být porušeny velmi hrubě, aby například při jednostanném testu pro $\alpha = 0{,}05$ kritická hodnota t_k „usekla“ ve skutečnosti významně více než 5% obsahu podgrafu hustoty t – i při porušených předpokladech tato kritická hodnota většinou usekne 6 nebo 7 procent, a nikoli třeba 15 procent obsahu. Proto **pokud t-hodnota kritéria přesáhne hodnotu t_k výrazně, je rozhodnutí zamítnout H_0 celkem bezpečné**.

b) Zkontrolujte své rozdělení. Nákres histogramu rozdělení je užitečný jak pro ověření předpokladu normality, tak ověření principu homogenního rozptylu.

c) Porovnejte $\overline{s_1^2}$ a $\overline{s_2^2}$. pokud se hodnoty obou odhadů liší výrazně, znamená to porušení předpokladu homogenního rozptylu. Ale pokud podíl těchto odhadů je menší než 4 při přibližně stejné délce obou souborů měření, není třeba dělat paniku.

d) Porušení předpokladů nemá takový dopad, pokud $N_i \geq 20$ a $N_1 \doteq N_2$.

e) Pozor na porušení měřítka. Některé proměnné jsou bezproblémové, např. X = průměrná rychlost (v km za hodinu), protože rozsah stupnice 10 až 20 km má stejnou váhu jako rozsah 70km až 80km.

Ale existují pochybné proměnné v tom smyslu, že měřítko porušují – například Y = ohodnocení otázky v dotazníku počtem bodů ze stupnice 1 až 7. Zde totiž rozsah 3 až 4 body nemusí být ekvivalentní rozsahu stupnice 6 až 7 bodů. Taková stupnice je hodně subjektivní, nemá pevně vymezený absolutní přírůstek, a proto může vést ke zkreslenému tvaru rozdělení.

f) Pokud některý z předpokladů je porušen do té míry, že U-test nebo t-test nelze použít, je možné zkusit neparametrické testy (viz 7).

Pojmy k zapamatování

– Tato kapitola je klíčovou kapitolou statistické části této studijní opory, protože obsahuje nejčastěji prováděný test porovnání dvou souborů měření – t-test – a také odvozuje a ilustruje význam intervalů spolehlivosti.

– Úvodem v oddílu 2.5.1 jsou zopakovány některé vzorce z předmětu Matematika 3, zejména vzorce pro $\overline{X}$, S^2 – a dále je vysvětlen a odvozen vzorec pro $\overline{S^2}$ jako nejlepší nestranný odhad neznámého rozptylu σ^2 normálního rozdělení.

– t-test používáme v situacích analogických U- testu (= u veličiny s normálním rozdělením) s tím jediným rozdílem, že totiž není znám rozptyl $\sigma^2_{\overline{X}}$ a musíme jej odhadnout pomocí $\frac{S^2}{N}$. Pro odhad rozptylu průměru je důležitý zejména vzorec 2.13, na který je potřeba nezapomenout.
– Celá kapitola předpokládá, že čtenář už ví, co je to statistický test, a zvládl kapitoly 13,14 textu Matematika 3 – přesto ty důležité informace ohledně statistických testů opakuje (poznámky 2.5.3 objasňují terminologii a otázku přístupu ke statistickému testu).
– Některé pojmy (hladina významnosti testu, síla testu) nebyly zopakovány, ale jsou procvičovány v otázkách a příkladech následujících za tímto shrnutím.
– Protože téměř všechno se vším souvisí, zamítne statistický test, který je dostatečně silný, hypotézu H_0 prakticky vždy – z toho důvodu je někdy lepší hledat informace nikoli ve statistickém testu, ale v intervalu spolehlivosti (blíže viz poslední poznámku v 2.5.4).
– Vyvrcholením kapitoly jsou testy typu $\mu_1 = \mu_2$. Ovšem věcí novou je představení rozdílu mezi párovým a nepárovým testem. Tato otázka je klíčová pro zapamatování z celé kapitoly.

Kontrolní otázky

1. U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.
   a) Výběrový průměr $\overline{X}$ je nestranným a konzistentním odhadem parametru μ měření veličiny s normálním rozdělením.
   b) Statistika S^2 je nestranným a konzistentním odhadem parametru σ^2 měření veličiny s normálním rozdělením.
   c) Kritická t- hodnota je blíže počátku než odpovídající (= pro stejnou hladinu významnosti sestrojená) kritická U-hodnota.
   d) Pro rostoucí počet stupňů volnosti graf hustoty t-rozdělení stále více splývá s grafem hustoty normovaného normálního rozdělení.
   e) Je možné, že interval spolehlivosti pro střední hodnotu μ průměru $\overline{X}$ sestrojený na základě konkrétní realizace $\overline{x}$ tuto střední hodnotu vůbec neobsahuje.
   f) Pokud hodnota μ_0 nepadne do oboustranného intervalu spolehlivosti pro střední hodnotu μ průměru $\overline{X}$, znamená to, že oboustranný test pro $H_0 : \mu = \mu_0$ tuto hypotézu H_0 nezamítne.
   g) Párovým testem je každý test typu $\mu_1 = \mu_2$, který porovnává střední hodnoty dvou náhodných veličin.
   h) Vnitřní rozptyl popisuje rozmanitost konkrétních jedinců v populaci – tato rozmanitost přitom nezávisí na podmínkách, při kterých je měření prováděno (je stejná u kontrolní skupiny i experimentální skupiny).
   i) Celkový rozptyl je součtem vnitřního rozptylu a vnějšího rozptylu.
   j) Při nestejných délkách obou souborů měření lze vnitřní rozptyl odhadnout jako aritmetický průměr rozptylů $\overline{s_1^2}$, $\overline{s_2^2}$ jednotlivých souborů měření.
   k) Existují situace, kdy neplatí princip homogenního rozptylu (rozptyl měření v kontrolní skupině je nesrovnatelně větší než rozptyl měření v experimentální skupině).

Odpovědi na otázky
1a) – A, 1b) – N, 1c) – N, 1d) – A, 1e) – A, 1f) – N, 1g) – N, 1h) – A, 1i) – N, 1j) – N (průměr se

bere nikoli aritmetický, ale vážený), 1k) – A (pokud se oba rozptyly liší o více než čtyřnásobek, je vhodnější místo paramatrického testu použít test neparametrický).

Cvičení

1. Honza Kovář je testérem motocyklů FONDA. S jistým prototypem provede šest jízd na 400m, získává časy (v sekundách): $10, 11, 10, 9, 10, 12$.

   a) Jaký je rozptyl těchto šesti hodnot?

   b) Jaký je odhad rozptylu měření u daného prototypu?

2. Honza nyní dostal pět různých strojů FONDA a s každým jede jízdu dlouhou 400 metrů. Získal data: $9, 12, 15, 8, 10$.

   a) Jaký je nyní nejvhodnější odhad rozptylu času na 400 metrů u strojů FONDA?

   b) Jaký je odhad rozptylu průměrného času všech pěti strojů?

3. Náhodně byly vybrány tři hodnoty z jisté populace: $x_1 = 2$, $x_2 = 4$, $x_3 = 4$.

   a) Vypočtěte rozptyl tohoto vzorku hodnot.

   b) Odhadněte rozptyl celé populace.

   c) Odhadněte rozptyl teoretického rozdělení průměru $\overline{X}$.

4. Při průzkumu zájmu o nový výrobek odpovědělo ze 400 dotázaných zákazníků supermarketu 80 zákazníků kladně na otázku, zda o výrobek mají zájem. Vycházejte z tohoto průzkumu a určete intervalový odhad procenta zájemců se spolehlivostí 0,95. (Návod: určete nejprve intervalový odhad počtu zákazníků ze 400, kteří si výrobek skutečně koupí - binomické rozdělení aproximujte normálním se stejnou střední hodnotou a rozptylem; pak vydělte obě meze v intervalu spolehlivosti číslem 400 (dostanete pravděpodobnosti) a vynásobte číslem 100 (dostanete procenta))

5. Aneta Šedá čte v novinách, že průměrný čech vidí za týden v televizi deset mrtvých. Zeptá se čtyř svých spolužaček, kolik mrtvých viděla každá z nich minulý týden v televizi. Jejich odpovědi jsou: $9, 8, 10, 9$.

   a) Jaký je průměr a 95%-ní interval splehlivosti pro tyto hodnoty?

   b) Proveďte t-test pro tato data (odpovídá měření čtyř hodnot novinové zprávě?).

6. Hokejové týmý Kometa Brno a Draci Brno mají v posledních pěti letech výsledky vzájemný zápasů (Kometa: Draci) $3:2$, $3:3$, $3:5$, $4:5$, $6:4$, $6:2$, $4:1$, $1:0$, $1:2$, $3:1$. $2:0$, $3:5$, $2:1$, $3:0$, $2:1$. Má trenér Komety právo říkat, že jeho tým je významně lepší? Odpovězte

   a) znaménkovým testem (viz BMA3), který bere v úvahu pouze výhru nebo prohru;

   b) t-testem, který bere v úvahu skóre v jednotlivých zápasech

7. Sociologové chtějí provést experiment, který má zjistit, zda počet schůzek chlapce s děvčetem závisí na tom, zda je chlapec prvorozený nebo ne. Získá náhodně vybraných šest náctiletých chlapců prvorozených a šest jiných druhorozených a zjistí, kolik schůzek měl každý z nich během jednoho měsíce. Prvorození: $6, 4, 5, 7, 3, 5$; druhorození: $2, 5, 4, 2, 1, 4$.

   a) Vypočtěte příslušné intervaly spolehlivosti (pro střední hodnotu průměru měření) se spolehlivostí 95% u každé z obou skupin chlapců.

   b) Proveďte t-test.

8. Chceme testem ověřit, zda kvalita reakcí člověka je stejná za denního i za umělého světla. U náhodně vybrané skupiny deseti lidí byly získány výsledky zkoušky „denní světlo“: $9, 2, 7, 12, 14, 10, 6, 7, 12, 10$ a hodnoty zkoušky „umělé osvětlení“: $7, 2, 4, 13, 13, 7, 4, 6, 8, 9$ (oba soubory jsou uspořádané, tj. data od i-tého respondenta jsou v obou případech na i-té pozici). Za použití t-testu rozhodněte, zda soubor „denní světlo“ nabývá významně vyšších hodnot než soubor „umělé osvětlení“.

9. Ve skupině A (292 studentů) je průměr hodnocení semestrální písemky roven $\overline{x_1} = 35{,}5$ bodů, ve skupině B (260 studentů) $\overline{x_2} = 41{,}6$ bodů. Byly spočteny také rozptyly ohodnocení v každé skupině písemek $\overline{s_1^2} = 100$, $\overline{s_2^2} = 90$. Studenti si neoficiálně mezi sebou stěžují, že zkoušející ve skupině A je výrazně přísnější. Ověřte testem, zda je rozdíl mezi průměry ohodnocení statisticky významný.

Výsledky

ad 1. a) Podle (2.6) $s^2 \doteq 0{,}889$; b) podle (2.9) je $\overline{s^2} = \frac{N}{N-1} \cdot s^2 = \frac{6}{5} \cdot 0{,}889 \doteq 1{,}0667$.

ad 2. a) Jedná se o veličinu X= čas u jednoho stroje – tedy $\overline{s^2} = 7{,}7$; b) Jedná se o veličinu $\overline{X}$ – proto pro odhad rozptylu průměru máme $\frac{7{,}7}{5} = 1{,}54$.

ad 3. a) 0,889 b) 1,333 c) 0,444

ad 4. Vycházíme z toho, že $p = \frac{80}{400} = 0{,}2$, máme tedy rozdělení binomické (Bi($N = 400$, $p = 0{,}2$)) a rozptyl $\sigma^2 = Np(1-p) = 64$. Pro $\alpha = 0{,}05$ víme z oboustranného testu (BMA3 - kapitola 13), že $u_k = 1{,}96$. Tedy střední hodnota $\mu \in 80 \pm 1{,}96 \cdot \sqrt{64} = (64{,}32; 95{,}68)$. Vydělením 400 dostaneme $(0{,}1608; 0{,}2392)$, tedy procento zájemců je přibližně 16 až 24.

ad 5. a) $\overline{x} = 9$; odhad roztylu $\overline{s^2} = \frac{ss}{\nu} = \frac{2}{3}$, ale musíme ještě vydělit počtem měření, abychom dostali odhad rozptylu průměru $\frac{2}{12}$. Dále t_k pro $\nu = 3$, $\alpha = 2q = 0{,}05$ je rovno 3,182; tedy

$$\mu \in 9 \pm \sqrt{\frac{2}{12}} \cdot 3{,}182 \doteq (7{,}701; 10{,}299).$$

b)

(K1) H_0: $\mu = 10$; H_1: $\mu \neq 10$.

(K2) Testovým kritériem bude veličina $\overline{X}$, respektive $\frac{\overline{X}-9}{\sqrt{212}}$.

(K3) Při platnosti H_0 má veličina $\frac{\overline{X}-9}{\sqrt{212}}$ Studentovo t-rozdělení pro $\nu = 3$.

(K4) Pro $\alpha = 0{,}05$ příslušná kritická hodnota $t_k = 3{,}182$.

(K5) Odpovídající t-hodnota kritéria je $-2{,}44949 \in (-t_k; t_k)$. H_0 tedy nezamítáme, neprokázala se nepravdivost novinové zprávy.

ad 6. a) Kometa získala 10 výher, jednu remízu a 4 prohry; pokud remízu nebudeme brát v úvahu pro žádnou ze stran, ze 14 možností je 10 znamének „+“; jedná se o jednostranný znaménkový test:

(K1) H_0: oba týmy jsou přibližně stejně silné; H_1: Kometa je statisticky vyznamně lepší.

(K2) Testovým kritériem bude veličina T= počet výher Komety ze čtrnácti zápasů.

(K3) Při platnosti H_0 má veličina T rozdělení $Bi(N = 14, p = 0{,}5)$.

(K4) Pro $\alpha = 0{,}05$ příslušná kritická hodnota $T_k = 11$ (blíže viz BMA3).

(K5) Měření $T = 10 < T_k = 11$, čili H_0 nezamítáme.

b) Uvažujme veličinu X= rozdíl skóre Kometa minus Draci, dostaneme hodnoty

$$1, 0, -2, -1, 2, 4, 3, 1, -1, 2, 2, -2, 1, 3, 1$$

(počítáme i remízu jako hodnotu 0).

(K1) H_0: oba týmy jsou přibližně stejně silné ($\mu = 0$); H_1: Kometa je statisticky vyznamně lepší ($\mu > 0$).

(K2) Testovým kritériem bude veličina $\overline{X}$, respektive veličina $\frac{\overline{X}-0}{est\ \sigma_{\overline{X}}}$.

(K3) Při platnosti H_0 má veličina $\frac{\overline{X}-0}{est\ \sigma_{\overline{X}}}$ rozdělení t pro $\nu = 14$. Lze spočítat $\overline{s^2} \doteq 3{,}35238$, tedy odmocnina $\overline{s} \doteq 1{,}83$. Tedy $est\ \sigma_{\overline{X}} = \frac{1{,}83}{\sqrt{15}} \doteq 0{,}47$

(K4) Pro $\alpha = 0{,}05 = q$ příslušná kritická hodnota $t_k(\nu = 14) = 1{,}761$.

(K5) Měření $\frac{0{,}93-0}{0{,}47} = 1{,}97 \notin (-\infty; 1{,}761)$, čili H_0 zamítáme, Kometa je významně lepší.

Je vidět, že t-test má větší statistickou sílu než znaménkový test – prokázal závislost studovaných proměnných, zatímco znaménkový test závislost neprokázal.

ad 7. a) $5 \pm 1{,}35$; $3 \pm 1{,}35$ b) Při oboustranném testu pro $\alpha = 2q = 0{,}05$ a $\nu = 10$ je naměřená hodnota kritéria $2{,}336 \notin (-2{,}228; 2{,}228)$, tj. zamítáme H_0 o rovnosti obou skupin. Prvorození mají statisticky významně více schůzek. Jak to interpretovat? Snad tím, že prvorození více stojí o děvčata :-)

ad 8. Vytvoříme příslušný soubor rozdílů odpovídajících hodnot: $2, 0, 3, -1, 1, 3, 2, 1, 4, 1$ (párový test). Odtud $\overline{x} = 1{,}6$, $\overline{s^2} = 2{,}267$, $est\ \sigma^2_{\overline{X}} = \frac{2{,}26}{10} = 0{,}2267$. Statistický test:

(K1) H_0: $\mu = 0$. H_1: $\mu \neq 0$.

(K2) Testovým kritériem bude veličina $\overline{X}$, respektive podíl $\frac{\overline{X}-0}{\sqrt{0{,}2267}}$.

(K3) Při platnosti H_0 má veličina $\frac{\overline{X}-0}{\sqrt{0{,}2267}}$ rozdělení t pro $\nu = 9$.

(K4) Pro $\alpha = 0{,}05 = 2q$ příslušná kritická hodnota $t_k(\nu = 9) = \pm 2{,}262$.

(K5) Měření $\frac{1{,}6-0}{0{,}4761} = 3{,}361 \notin (-2{,}262; 2{,}262)$, čili H_0 zamítáme, rozdíl mezi denním a umělým světlem je statisticky významný.

ad 9. **(K1)** H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 \neq \mu_2$.

(K2) Testovým kritériem bude veličina $\overline{X_1} - \overline{X_2}$, respektive podíl $\frac{\overline{X_1}-\overline{X_2}-0}{\sqrt{est\ \sigma^2_{\overline{X_1}-\overline{X_2}}}}$.

(K3) Při platnosti H_0 má uvedený podíl rozdělení t s volností tak vysokou ($\nu = 291 + 259 = 550$), že je můžeme beztrestně nahradit normovaným normálním rozdělením U – odhad vnitřního rozptylu pak přímo (protože $\nu > 60$) položíme roven vnitřnímu rozptylu:

$$\sigma^2 \doteq est\sigma^2 = \frac{291}{291+259} \cdot \overline{s_1^2} + \frac{259}{291+259} \cdot \overline{s_2^2} \doteq 95{,}29,$$

a tudíž $\sigma^2_{\overline{X_1}-\overline{X_2}} = \frac{95{,}29}{292} + \frac{95{,}29}{260} = 0{,}69$, tj. $\sigma_{\overline{X_1}-\overline{X_2}} = \sqrt{0{,}69} = 0{,}83$.

(K4) Pro $\alpha = 0{,}05 = 2q$ příslušná kritická hodnota $t_k(\nu = 550) = u_k = \pm 1{,}96$.

(K5) Měření $\frac{35{,}5-41{,}6}{0{,}83} = -7{,}35 \notin (-1{,}96; 1{,}96)$, čili H_0 zamítáme, rozdíl mezi skupinami je významný. Interpretace těchto výsledků je ovšem také náročná – například může být skutečností (a asi to tak i je), že pokud jsou studenti do skupin rozděleni podle oborů, tak rozdílnost výsledků je dána rozdílností studentů přijatých na jednotlivé obory (i kdyby přístup obou zkoušejících byl naprosto stejný) :-)

3 Analýza rozptylu

Průvodce studiem

Tato kapitola je přirozeným pokračováním otázek studovaných v kapitole 2.5, ale jestliže dříve jsme se zabývali porovnáním středních hodnot dvou skupin měření (přitom každá ze skupin měření závislé veličiny byla získána pro jinou hodnotu nezávislé veličiny), tak v této kapitole budeme studovat porovnávání středních hodnot tří a více skupin měření (byly získány tři a více souborů měření závislé veličiny, každá skupina měření byla získána pro jinou hodnotu nezávislé veličiny).

Přesněji řečeno, toto zobecnění počtu hodnot nezávislé veličiny je tématem oddílu 3.1. Oddíl 3.1.2 je zobecněním kapitoly 2.5 v jiném smyslu – zatímco 3.1 studuje vztah jedné nezávislé veličiny a jedné veličiny závislé na té první (tzv. **jednofaktorová analýza rozptylu = jednofaktorová ANOVA***), oddíl 3.1.2 se zabývá studiem vztahu jedné veličiny závislé na dvou (nebo i více) nezávislých veličinách (tzv.* **dvoufaktorová ANOVA nebo vícefaktorová ANOVA***) Jedná se tedy o zobecnění počtu nezávislých veličin. No a do třetice opět situace jiného druhu – zatímco 3.1 a 3.1.2 se zabývaly studiem navzájem nezávislých skupin měření, tak oddíl 3.1.3 (***experiment opakovaného měření***) studuje skupiny měření, které jsou spolu navzájem provázány – stejný jedinec je podroben měření v každé z podmínek (vlastně zobecnění situace párového testu z kapitoly 2.5 pro tři a více podmínek měření).*

Cíle

Po prostudování této kapitoly budete schopni:

- Pracovat s analýzou rozptylu ANOVA.
- Vedle t-testu se naučíte používat i F-test.

3.1 Jednofaktorová analýza rozptylu

3.1.1 Příklad a vzorce

Zatím jsme se zabývali statistickými testy v situaci jednoho nebo dvou souborů měření. V této kapitole začneme mluvit o složitějších experimentálních situacích se třemi a více skupinami měření.

Uvažujme například experiment, ve kterém jsou za třech různých podmínek získány tři soubory měření a chceme testovat, zda se odpovídající střední hodnoty populace v těchto třech případech statisticky významně liší. Jedním ze způsobů, jak toho dosáhnout, je provést tři t-testy: první pro skupiny 1 a 2, druhý pro skupiny 1 a 3, třetí pro skupiny 2 a 3. Z důvodů, které budou uvedeny v kapitole 5, je tento postup poněkud problematický – proto nyní popíšeme test, který porovnání všech tří souborů měření provede najednou.

Jak už bylo předesláno v předchozí kapitole, budeme analyzovat dva typy rozptylu – vnitřní rozptyl neboli **rozptyl uvnitř tříd (... RUT)** a vnější rozptyl neboli **rozptyl mezi třídami (... RMT)**.

Testovým kritériem při analýze rozptylu bude podíl odhadů rozptylů obou typů

$$\frac{est\ RMT}{est\ RUT}$$

Toto kritérium vlastně není ničím novým, protože i u t-testu lze na kritérium

$$\frac{\overline{X_1} - \overline{X_2}}{est\ \sigma_{\overline{X_1}-\overline{X_2}}}$$

pohlížet jako na podíl rozptylů dvojího typu: $RMT \doteq k \cdot (\overline{X_1} - \overline{X_2})$ (vnější rozptyl je jen k-násobkem rozdílu průměrů pro jisté k), zatímco $est\ RUT \doteq est\ \sigma_{\overline{X_1}-\overline{X_2}}$ (odhad vnitřní rozmanitosti je dán odmocninou z odhadu vnitřního rozptylu).

Příklad 3.1. Zaměstnanci UMAT hledají nejlepší metodu výuky matematiky v prvním semestru ze tří možností – cvičení, konzultace, přednáška. Aby zjistily, zda se metody od sebe nějak významně liší, vybrali náhodně 27 studentů a rozdělili je do tří skupin po devíti $N_1 = N_2 = N_3 = 9$, z nichž každá je vyučována jinou metodou. Poté jsou záskána data ohodnocení zkoušky BMA1:

skupina 1 (cvičení)	skupina 2 (konzultace)	skupina 3 (přednáška)
94	83	80
90	86	85
95	89	81
89	87	81
88	85	79
92	86	83
92	85	78
97	81	80
91	83	82
$T_1 = \sum_1^9 x_{1i} = 828$	$T_2 = \sum_1^9 x_{2i} = 765$	$T_3 = \sum_1^9 x_{3i} = 729$
$\overline{x_1} = \frac{T_1}{9} = 92$	$\overline{x_2} = \frac{T_2}{9} = 85$	$\overline{x_3} = \frac{T_3}{9} = 81$

O významnosti rozdílů středního hodnot v různých skupinách měření rozhodněte statistickým testem.

Jedná se o ilustrační příklad, proto si dovoluji pro řešení vypnout kurzívu. Označení T_1, T_2, T_3 z anglického „total"=součet.

a) Z tří skupin měření odhadneme nejprve vnitřní rozptyl $\sigma^2 = RUT$: Stupně volnosti $\nu_1 = \nu_2 = \nu_3 = 8$. Součet čtverců ze vztahu 2.23 s využitím označení pro součet $T = \sum x_i$ budeme psát ve tvaru

$$ss = \left(\sum x_i^2\right) - \frac{T^2}{N}, \tag{3.1}$$

který budeme dále využívat, protože je numericky jednodušší. Užitím 3.1 a 2.12 tedy máme pro každou ze tří skupin měření

$$\overline{s_1^2} = est_1\ RUT = \frac{ss_1}{\nu_1} = \frac{68}{8} = 8{,}5;$$

$$\overline{s_2^2} = est_2\ RUT = \frac{ss_2}{\nu_2} = \frac{46}{8} = 5{,}75;$$

$$\overline{s_3^2} = est_3\ RUT = \frac{ss_3}{\nu_3} = \frac{36}{8} = 4{,}5.$$

Vidíme, že rozptyly měření se liší maximálně o dvojnásobek, takže je rozumné brát v úvahu princip homogenního rozptylu a odhadnout RUT jako aritmetický průměr vypočtených tří hodnot:

$$est\ RUT = \frac{8{,}5 + 5{,}75 + 4{,}5}{3} = 6{,}25.$$

Tento odhad má $\nu_1 + \nu_2 + \nu_3 = 24$ stupňů volnosti. U příkladu 2.20 jsme u t-testu odhadovali vnitřní rozptyl – tento odhad jsme mohli docela dobře určit ze vzorce $est\ \sigma^2 = \frac{ss_1+ss_2}{\nu_1+\nu_2}$ (ověřte, že výsledek je stejný). Podobně i náš aritmetický průměr 6,25 tří odhadů lze počítat tímto způsobem – vzorec

$$est\ \sigma^2 = \frac{\sum_1^J ss_j}{\sum_1^J \nu_j} \tag{3.2}$$

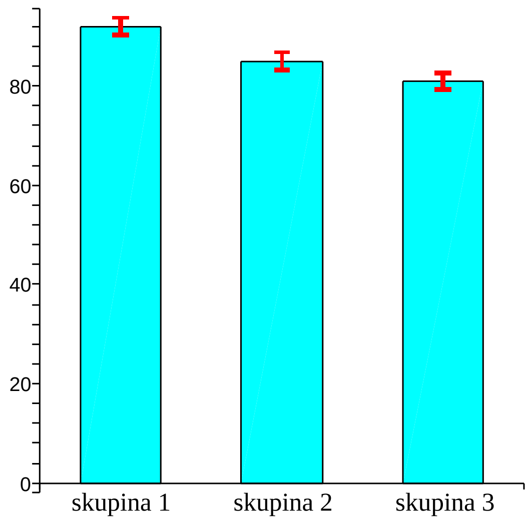
platí nikoli jen pro tři skupiny, měření, ale obecně pro J skupin. Pokud bychom jej použili v našem příkladu,

$$est\ RUT = \frac{68 + 46 + 36}{8 + 8 + 8} = 6{,}25.$$

b) S použitím t-rozdělení nyní můžeme sestavit intervaly spolehlivosti pro střední hodnoty μ_1, μ_2, μ_3: když $t_k(\alpha = 2q = 0{,}05; \nu = 24) = 2{,}064$, pak $\mu_j \in \overline{x_j} \pm \sqrt{\frac{6{,}25}{9}} \cdot t_k$. Protože $N_j = 9$ pro všechna j, budou všechny tři intervaly spolehlivosti stejně dlouhé:

$$\mu_1 \in 92 \pm 1{,}72; \quad \mu_2 \in 85 \pm 1{,}72; \quad \mu_3 \in 81 \pm 1{,}72.$$

Průměry a intervaly spolehlivosti lze znázornit graficky:



Tyto tři intervaly spolehlivosti se vůbec nepřekrývají – to nám mimo jiné říká, že všechny tři střední hodnoty jsou navzájem různé, čili nulová hypotéza $H_0 : \mu_1 = \mu_2 = \mu_3$ bude následujícícm testem zamítnuta.

c) Provedeme nyní statistický test:

(K1) H_0: $\mu_1 = \mu_2 = \mu_3$ (vyučovací metoda nemá vliv na výsledek zkoušky); H_1: neplatí $\mu_1 = \mu_2 = \mu_3$.

Všimněme si, že hypotéza H_1 je docela nejasná. Pokud H_1 platí, může to znamenat hodně věcí: $\mu_1 > \mu_2 = \mu_3$, nebo $\mu_1 > \mu_2 > \mu_3$, nebo $\mu_1 < \mu_2 < \mu_3$, nebo $\mu_1 = \mu_2 > \mu_3$, nebo $\mu_1 = \mu_2 < \mu_3$, atd. Tato nejasnost hypotézy H_1 je hlavní nevýhodou analýzy rozptylu – i když v dalších kapitolách uvidíme, že i s analýzou rozptylu se dá něco dělat.

(K2) Jak už bylo řečeno úvodem, testovým kritériem bude podíl

$$\frac{est\ RMT}{est\ RUT}.$$

Čím větší je hodnota *est RMT* (odhad rozptylu mezi třídami), tím větší důvod bude si myslet, že mezi jednotlivými středními hodnotami průměrů existuje skutečný rozdíl. Čím větší bude *est RUT* (odhad rozptylu uvnitř tříd), tím více máme důvod se domnívat, že rozdíly mezi skupinami měření jsou způsobeny pouze náhodnými vlivy.

$est\ RUT = 6{,}25$; nyní vypočteme ještě *est RMT* – budeme postupovat obdobně jako při určení vnejšího rozptylu v příkladu 2.20. Máme změřeny tři průměry $\overline{x_1} = 92$, $\overline{x_2} = 85$, $\overline{x_3} = 81$ – vypočtěme rozptyl těchto průměrů podle

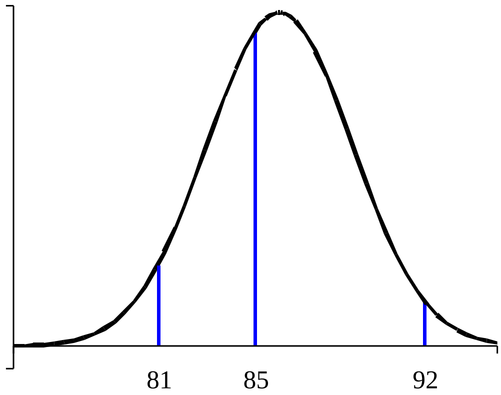
2.23 a 2.22: Protože $\nu = 3 - 1 = 2$, máme

$$est\ r_N = \frac{ss}{2}, \quad ss = \sum_{j=1}^{J} \overline{x_j}^2 - \frac{(\sum \overline{x_j})^2}{J} = 22250 - \frac{258^2}{3} = 62.$$

Tedy $est\ r_N = \frac{62}{2} = 31$. To ale ještě není všechno – protože $est\ r_N$ je odhadem rozptylu průměru devíti hodnot a $r_N = \frac{RMT}{N}$, máme

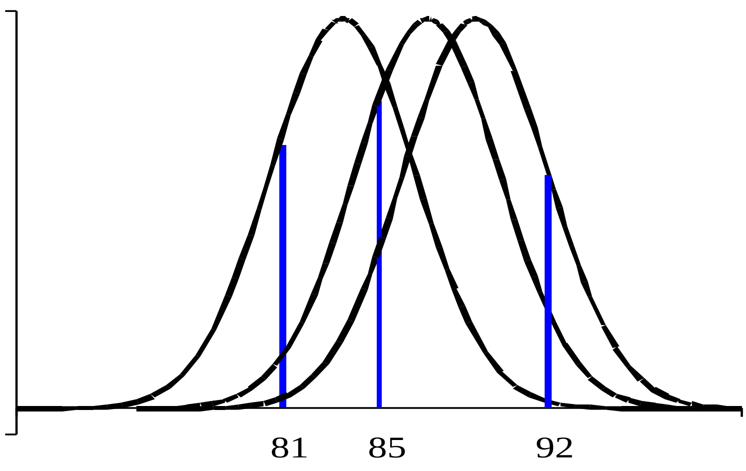
$$est\ RMT = N \cdot r_N = 9 \cdot 31 = 279.$$

Na tomto místě je snad vhodné graficky test „rozebrat“: Pokud platí H_0, tak průměry $\overline{x_1} = 92$, $\overline{x_2} = 85$, $\overline{x_3} = 81$ pocházejí z téhož rozdělení, které má rozptyl $\frac{RMT}{N}$:



To znamená, že $est\ RMT$ i $est\ RUT$ jsou odhady téhož rozptylu σ^2 a rozdíl mezi hodnotami 6,25 a 279 je pouze náhodný.

Pokud ovšem platí H_1, průměry $\overline{x_1} = 92$, $\overline{x_2} = 85$, $\overline{x_3} = 81$ pocházejí z různých rozdělení a rozptyl $\frac{RMT}{N}$ je větší než rozptyl $\frac{RUT}{N}$ z předchozího obrázku:

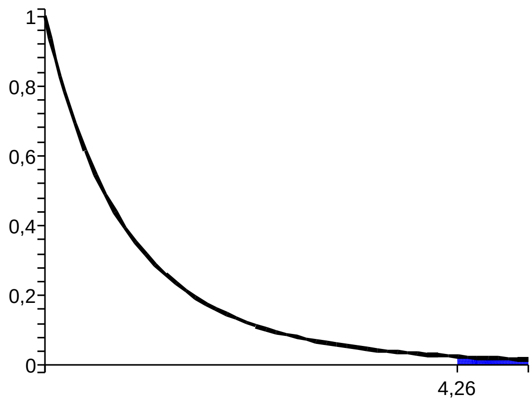


Každý ze tří průměrů na obrázku je průměrem jiné náhodné veličiny, čili $est\ RMV$ je odhadem jiného, většího rozptylu než RUP (rozptýlenost průměrů není popsána jednou veličinou, ale třemi různými veličinami). Obecně lze říci, že čím větší je rozdíl mezi $est\ RMT$ a $est\ RUT$, tím větší je pravděpodobnost, že platí H_1.

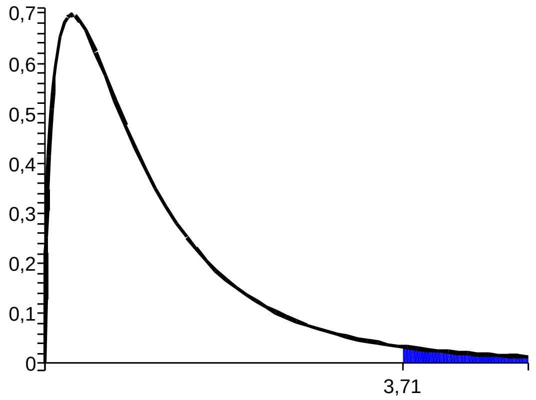
(K3) Jaké je rozdělení veličiny $\frac{est\ RMT}{est\ RUT}$ za předpokladu platnosti hypotézy H_0? Pokud platí H_0, tak $est\ RMT$, $est\ RUT$ jsou odhady téhož rozptylu σ^2. Sir Ronald Fisher odvodil, že pokud máme dva odhady téhož rozptylu, $est_1\ R$ s počtem stupňů volnosti v_1 a $est_2\ R$ s počtem stupňů volnosti v_2, tak podíl těchto odhadů $\frac{est_1\ R}{est_2\ R}$ má tzv. F**-rozdělení (podle svého objevitele nazývané Fisherovo rozdělení) pravděpodobnosti** s hustotou

$$f_{v_1,v_2}(x) = \begin{cases} 0 & \dots \quad x \leq 0; \\ \dfrac{v_2}{v_1} \cdot \dfrac{\frac{x^{\left(\frac{v_1}{2}-1\right)} \cdot e^{-\frac{x}{2}}}{2^{\frac{v_1}{2}} \cdot \Gamma\left(\frac{v_1}{2}\right)}}{\frac{x^{\left(\frac{v_2}{2}-1\right)} \cdot e^{-\frac{x}{2}}}{2^{\frac{v_2}{2}} \cdot \Gamma\left(\frac{v_2}{2}\right)}} = \dfrac{v_2}{v_1} \cdot 2^{\left(\frac{v_2-v_1}{2}\right)} \cdot x^{\left(\frac{v_1-v_2}{2}\right)} \cdot \dfrac{\Gamma\left(\frac{v_2}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right)} & \dots \quad x > 0. \end{cases}$$

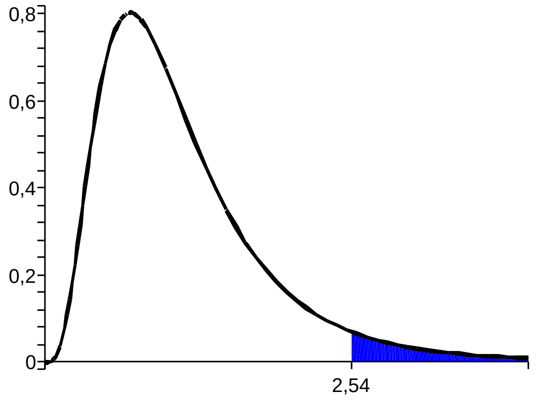
Střední hodnota veličiny popsané F-rozdělením je rovna $\frac{v_2}{v_2-2}$, je tedy blízká jedné, což i odpovídá platnosti hypotézy H_0 – pokud v podílu jsou odhady stejného rozptylu, tak ten podíl by se měl přibližně rovnat jedné. Pro různé hodnoty stupňů volnosti má graf hustoty jiný průběh, tj. jinou kritickou hodnotu (pro $\alpha = 0{,}05$). Například



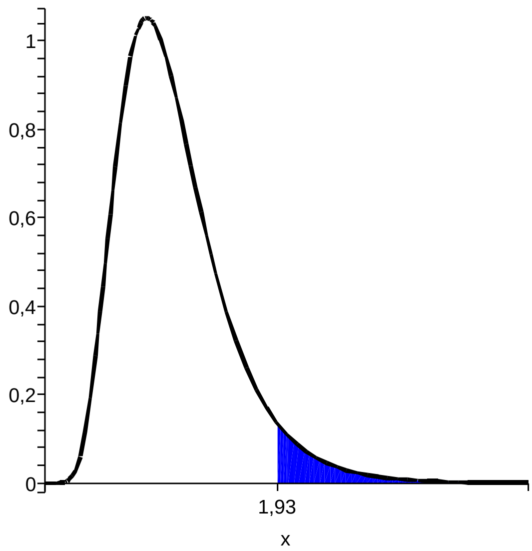
pro $v_1 = 2$, $v_2 = 9$ má hustota $f_{2,9}(x)$ tvar a $F_k = 4{,}26$.



Pro $v_1 = 3$, $v_2 = 10$ má hustota $f_{3,10}(x)$ tvar a $F_k = 3{,}71$.



Pro $v_1 = 10$, $v_2 = 15$ se graf $f_{10,15}(x)$ ještě více liší od nepřímé úměrnosti případu $f_{2,9}$,



graf $f_{20,30}$ už jednoznačně dává čtenáři představu, jaký tvar je pro většinu dvojic stupňů volnosti pro graf hustoty charakteristický. Ve všech čtyřech případech byla vyšrafována oblast na pravé straně mezi grafem a osou x, jejíž obsah je roven $\alpha = 0{,}05$.

Kritické hodnoty F-rozdělení byly také jednou provždy vypočteny a sestaveny do tabulky. Situace je ještě náročnější než u t-rozdělení, protože přibývá jeden parametr navíc – oproti t-rozdělení jsou u F-rozdělení dva stupně volnosti, čili pro každou hladinu významnosti α potřebuje F-rozdělení jednu celou tabulku!!! Tato prostorová náročnost vyžaduje, abychom se omezili pouze na jedinou hladinu významnosti, maximálně na dvě. V tabulce 3.1 jsou uvedeny kritické hodnoty F-rozdělení pro $\alpha = 0{,}05$, v tabulce 3.2 pro $\alpha = 0{,}01$.

Všechny uvedené kritické hodnoty se týkají pravostranného testu – v případě potřeby levostranného F-testu je možné pro nalezení kritické hodnoty užít vzorec

$$F_k^{V_1,V_2}(1-\alpha) = \frac{1}{F_k^{V_2,V_1}(\alpha)} \tag{3.3}$$

(všimněte si, že ve vzorci je na pravé straně přehozeno pořadí stupňů volnosti – pozor na to, tabulka kritických hodnot není symetrická). Dále si prosím všimněte, že volnosti jsou označeny písmenem V a nikoli ν – označení volnosti pomocí písmene V mi připadne jednak více české (V jako „volnost“), jednak vhodnější v této a dalších kapitolách (budeme totiž počítat volnosti odhadů různých typů).

(K4) Pro $\alpha = 0{,}05$ příslušnou kritickou hodnotu určíme z tabulky 3.1, kde volnost odhadu *est RMT* je $V_1 = 2$, volnost odhadu *est RUT* je $V_2 = 24$ – čili na průsečíku sloupce $V_1 = 2$ a řádku $V_2 = 24$ nacházíme kritickou hodnotu $F_k = 3,4$.

Tab. 3.1: Kritické hodnoty F-testu pro $\alpha = 0{,}05$.

$V_1 \rightarrow$ $V_2 \downarrow$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	∞
1	161	199	216	225	230	234	237	239	240	242	244	246	248	249	250	251	252	254
2	18,5	19,0	19,2	19,2	19,3	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5
3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,40
12	4,75	3,90	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,19	2,15	2,10	2,06	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,08	2,01	1,99	1,95	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,07	1,99	1,95	1,90	1,85	1,80	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,13	2,06	1,97	1,93	1,88	1,84	1,79	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,04	1,96	1,91	1,87	1,82	1,77	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,10	2,03	1,94	1,90	1,85	1,81	1,75	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,25
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,00

Tab. 3.2: Kritické hodnoty F-testu pro $\alpha = 0{,}01$.

$V_1 \rightarrow$ $V_2 \downarrow$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	∞
1	4052	4999	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6366
2	98,5	99,0	99,2	99,2	99,3	99,3	99,4	99,4	99,4	99,4	99,4	99,4	99,4	99,5	99,5	99,5	99,5
3	34,1	30,8	29,5	28,7	28,2	27,9	27,7	27,5	27,3	27,2	27,0	26,9	26,7	26,6	26,5	26,4	26,1
4	21,2	18,0	16,7	16,0	15,5	15,2	15,0	14,8	14,7	14,6	14,4	14,2	14,0	13,9	13,8	13,7	13,5
5	16,3	13,3	12,0	11,4	11,0	10,7	10,5	10,3	10,2	10,0	9,89	9,72	9,55	9,47	9,38	9,29	9,02
6	13,8	10,9	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,72	7,56	7,40	7,31	7,23	7,14	6,88
7	12,2	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,47	6,31	6,16	6,07	5,99	5,91	5,65
8	11,3	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,67	5,52	5,36	5,28	5,20	5,12	4,86
9	10,6	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,11	4,96	4,81	4,73	4,65	4,57	4,31
10	10,0	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,71	4,56	4,41	4,33	4,25	4,17	3,91
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,40	4,25	4,10	4,02	3,94	3,86	3,60
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,16	4,01	3,86	3,78	3,70	3,62	3,36
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	3,96	3,82	3,66	3,59	3,51	3,43	3,17
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,80	3,66	3,51	3,43	3,35	3,27	3,00
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,67	3,52	3,37	3,29	3,21	3,13	2,87
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,55	3,41	3,26	3,18	3,10	3,02	2,75
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,46	3,31	3,16	3,08	3,00	2,92	2,65
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,37	3,23	3,08	3,00	2,92	2,84	2,57
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,30	3,15	3,00	2,92	2,84	2,76	2,49
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,23	3,09	2,94	2,86	2,78	2,69	2,42
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	3,17	3,03	2,88	2,80	2,72	2,64	2,36
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,12	2,98	2,83	2,75	2,67	2,58	2,31
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,07	2,93	2,78	2,70	2,62	2,54	2,26
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,03	2,89	2,74	2,66	2,58	2,49	2,21
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	2,99	2,85	2,70	2,62	2,54	2,45	2,17
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	2,96	2,81	2,66	2,58	2,50	2,42	2,13
27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06	2,93	2,78	2,63	2,55	2,47	2,38	2,10
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,90	2,75	2,60	2,52	2,44	2,35	2,06
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00	2,87	2,73	2,57	2,49	2,41	2,33	2,03
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,84	2,70	2,55	2,47	2,39	2,30	2,01
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,66	2,52	2,37	2,29	2,20	2,11	1,80
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,50	2,35	2,20	2,12	2,03	1,94	1,60
120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47	2,34	2,19	2,03	1,95	1,86	1,76	1,38
∞	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,18	2,04	1,88	1,79	1,70	1,59	1,00

(K5) Odpovídající F-hodnota kritéria je

$$F = \frac{est\ RMT}{est\ RUT} = \frac{279}{6{,}25} = 44{,}64,$$

což je více než $F_k = 3,4$, a tedy zamítáme hypotézu H_0 o tom, že střední hodnoty průměrů jsou ve všech třech skupinách měření shodné. Zde budiž také řečeno, že se jedná o pravostranný test, protože rozptyl RMT (vyjadřující rozdílnost mezi všemi měřeními) je vždy velký minimálně tolik jako rozptyl RUT (vyjadřující rozdílnost pouze uvnitř každé ze skupin měření), čili podíl $\frac{est\ RMT}{est\ RUT}$ bude vždy větší než jedna, jedinou otázkou je, kolikanásobně je RMT větší než RUT.

Poznámka 3.2. V případě t-testu v příkladu 2.20 jsme předpokládali, že $est_1\ \sigma^2$ a $est_2\ \sigma^2$ jsou odhady téhož rozptylu – tuto skutečnost je možné nyní ověřit také F-testem: Jako kritérium bychom použili podílu $\frac{est_1\ \sigma^2}{est_2\ \sigma^2}$, který má (při předpokladu H_0, že se jedná o odhady téhož roptylu) rozdělení F_{N_1-1,N_2-1} – další možnost je použít kritéria $\frac{est_2\ \sigma^2}{est_1\ \sigma^2}$, který má (při předpokladu H_0, že se jedná o odhady téhož roptylu) rozdělení F_{N_2-1,N_1-1}. Problémem zde je, že nemáme zajištěno, že by jeden z těchto dvou rozptylů musel být například větší než ten druhý, čili podíl těchto dvou odhadů může být větší i menší než jedna, tedy vychýlen od hodnoty 1 na obě strany – v tomto případě se tedy jedná o oboustranný test, při kterém musíme při určení kritických hodnot na obou stranách podgrafu „useknout“ obsah $\frac{\alpha}{2}$. Protože máme k dispozici jen tabulky 3.2, 3.1 usekávající v pravé části podgrafu jedno nebo pět procent obsahu, můžeme v tomto testu volit pouze $\alpha = 0{,}02$ nebo $\alpha = 0{,}1$, hodnotu F_v najít v jedné z těchto tabulek, a pro určení levé kritické hodnoty F_m užít vzorce 3.3.

Konkrétně v příkladu 2.20 pro $est_1 = 88$ a $est_2 = 30$ (s volnostmi $V_1 = V_2 = 4$), kritérium $\frac{est_1}{est_2}$ a $\alpha = 0{,}02$ dostaneme z tabulky 3.2 $F_v^{4,4} = 16,0$ a ze vzorce 3.3

$$F_m^{4,4} = \frac{1}{F_k^{4,4}} = \frac{1}{16} = 0{,}0625;$$

hodnota kritéria $\frac{88}{30} = 2{,}93$ leží v intervalu $(0{,}0625; 16)$; čili na hladině významnosti $\alpha =$ $= 0{,}02$ nezamítáme hypotézu H_0, že est_1, est_2 jsou odhady téhož rozptylu σ^2.

Shrňme nyní označení a vzorce použité v příkladu 3.1: Vzorec 3.2 lze s využitím označení $SSUT := \sum_1^J SS_i$, $VUT := \sum_1^J \nu_i$ přepsat ve tvaru

$$est\ RUT = \frac{SSUT}{VUT}, \tag{3.4}$$

kde pro VUT platí pro stejný počet měření $N_1 = N_2 = \cdots = N_J = N$ v každé ze skupin $1, 2, \ldots, J$

$$VUT = \sum_1^J \nu_i = J \cdot (N-1) = JN - J = n - J, \tag{3.5}$$

kde n je celkový počet měření ve všech skupinách ($JN = n$). Dále pokud označíme $T_i = \sum_{j=1}^{N} x_{ij}$ součet měření v i-té skupině (třídě), lze upravit vzorec pro $SSUT$ do tvaru

$$SSUT = \sum_{i=1}^{J} SS_i = \sum_{i=1}^{J} \left(\sum_{j=1}^{N} x_{ij}^2 - \frac{T_i^2}{N} \right) = \left(\sum_{i=1}^{J} \sum_{j=1}^{N} x_{ij}^2 \right) - \sum_{i=1}^{J} \frac{T_i^2}{N}. \tag{3.6}$$

Odhad *est RUT* je nejlepším odhadem rozptylu σ^2, tj. budeme jej používat pro sestavení intervalu spolehlivosti pro střední hodnotu průměru měření v dané třídě populace

$$\mu_i \in \overline{x_i} \pm t_k(VUT) \cdot \sqrt{\frac{est\ RUT}{N_i}}. \tag{3.7}$$

Podívejme se nyní na tvar vzorců pro RMT: pro průměr $\overline{x}$ průměrů $\overline{x_i}$ platí (označíme-li $\overline{x_i} = \frac{T_i}{N}$, $\sum T_i = T$ = součet všech měření v experimentu):

$$\overline{x} = \frac{\sum_{i=1}^{J} \overline{x_i}}{J} = \sum_{i=1}^{J} \frac{\frac{T_i}{N}}{J} = \sum_{i=1}^{J} \frac{T_i}{NJ} = \frac{T}{n};$$

dále (podle postupu v příkladu)

$$est\ RMT = N \cdot est\ \sigma_{\overline{X}}^2 = N \cdot \frac{\sum_{i=1}^{J} (\overline{x_i} - \overline{x})^2}{J-1}$$

(pokud označíme $SSMT = N \cdot \sum_1^J (\overline{x_i} - \overline{x})^2$ a $VMT = J - 1$). Upravme nyní ještě vzorec pro $SSMT$ (pro stejné délky měření v každé ze skupin $N_i = N$ lze využít označení $\sum T_i = T$ a faktu $JN = n$):

$$\begin{aligned} SSMT &= N \cdot \sum_1^J (\overline{x_i} - \overline{x})^2 = N \cdot \left[\sum_1^J \overline{x_i}^2 - \frac{\left(\sum_1^J \overline{x_i} \right)^2}{J} \right] = \\ &= N \cdot \sum_1^J \overline{x_i}^2 - \frac{N}{J} \left(\sum_1^J \overline{x_i} \right)^2 = N \cdot \left(\sum_1^J \frac{T_i^2}{N^2} \right) - \frac{N}{J} \cdot \left(\frac{\sum_1^J T_i}{N} \right)^2 = \\ &= \left(\sum_1^J \frac{T_i^2}{N} \right) - \frac{T^2}{n}. \end{aligned}$$

Všechny uvedené vzorce nyní sestavme do přehledné tabulky. Uděláme v ní ale ještě jeden krok navíc – předpoklad $N_i = N$ stejného počtu měření v každé ze skupin lze vypustit, hodnoty N_i mohou být navzájem různé. Příslušné odvození vzorců pro různá N_i ovšem už nebudeme provádět (snad čtenář získal určitou příchuť tohoto typu odvozování aspoň pro stejné hodnoty N_i) – kupodivu se ukazuje, že do všech vzorců právě odvozených lze dosadit místo N obecný počet měření N_i, a jinak vzorce zůstanou beze změny (obecně pro různé délky souborů měření neplatí $n = NJ$, ale platí $n = \sum N_i$)!!!

Z tabulky 3.3 je vidět výhoda přístupu pomocí vzorců pro SS – v případě MT i UT se rozptyl odhadne jako podíl součtu čtverců a počtu stupňů volnosti.

Tab. 3.3: Vzorce pro typy rozpylu u jednofaktorové ANOVA

typ	V (= volnost)	SS (= součet čtverců)	$est\ R$
MT	$VMT = J - 1$	$SSMT = \left(\sum_{i=1}^{J} \frac{T_i^2}{N_i}\right) - \frac{T^2}{n}$	$est\ RMT = \frac{SSMT}{VMT}$
UT	$VUT = n - J$	$SSUT = \left(\sum_{i=1}^{J}\sum_{j=1}^{N_i} x_{ij}^2\right) - \left(\sum_{i=1}^{J} \frac{T_i^2}{N_i}\right)$	$est\ RUT = \frac{SSUT}{VUT}$

Příklad 3.3. Provádíme test účinnosti čtyř značek zubní pasty v prevenci zubního kazu: WHITE, RUBBY, NIGGARD, NOHOLES. U každé značky bylo požádáno šest lidí, aby pastu po dobu jednoho roku používali a zaznamenávali přitom počet zubních kazů. Z každé šestičlenné skupiny během roku někdo odešel, takže byla získána následující data:

skupina 1 (WHITE)	skupina 2 (RUBBY)	skupina 3 (NIGGARD)	skupina 4 (NOHOLES)
$x_{11} = 2$	$x_{21} = 4$	$x_{31} = 6$	$x_{41} = 3$
$x_{12} = 1$	$x_{22} = 1$	$x_{32} = 2$	$x_{42} = 1$
$x_{13} = 3$	$x_{23} = 3$	$x_{33} = 4$	
	$x_{24} = 3$	$x_{34} = 4$	
	$x_{25} = 4$		
$T_1 = 6$	$T_2 = 15$	$T_3 = 16$	$T_4 = 4$
$\overline{x_1} = 2$	$\overline{x_2} = 3$	$\overline{x_3} = 4$	$\overline{x_4} = 2$

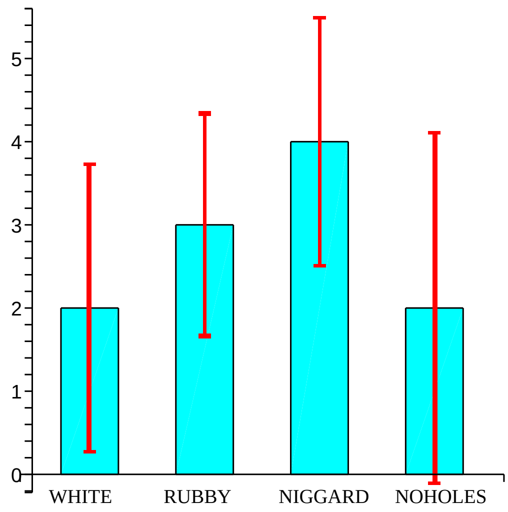
Vypočtěme nejprve intervaly spolehlivosti pro střední hodnotu v každé ze čtyř ($J = 4$) skupin měření (užitím 3.7 a vzorců z tabulky 3.3):

$$est\ RUT = \frac{SSUT}{VUT} = \frac{\sum_{\text{všechno}} x_{ij}^2 - \sum_{\text{skupiny}} \frac{T_i^2}{N_i}}{n - J} = \frac{147 - 129}{14 - 4} = 1{,}8;$$

Dále $VUT = 10$ (sečítá se volnost 2 ve skupině WHITE, volnost 4 ve skupině RUBBY, volnost 3 ve skupině NIGGARD a volnost 1 ve skupině NOHOLES). Z tabulky 2.1 pro $\alpha = 2q = 0{,}05$ a volnost 10 je $t_k(10) = 2{,}228$. Intervaly spolehlivosti mají teď různou délku díky různým počtům měření v každé ze skupin – podle 3.7

$$\mu_i \in \overline{X_i} \pm \sqrt{\frac{1{,}8}{N_i}} \cdot 2{,}228$$

máme $\mu_1 \in 2 \pm 1{,}73$, $\mu_2 \in 3 \pm 1{,}34$, $\mu_3 \in 4 \pm 1{,}49$ a $\mu_4 \in 2 \pm 2{,}11$ (grafické znázornění těchto intervalů viz obrázek):



Z obrázku je možné vyčíst následující informace:

– intervaly spolehlivosti mají neprázdný průnik, což nedává moc důvodů si myslet, že střední hodnoty veličin v jednotlivých situacích měření se liší významně;

– experiment nemá velkou sílu: mezi jednotlivými značkami pasty mohou být značné rozdíly (např. $\mu_1 = 1$, $\mu_3 = 5$), ale statisticky je nejsme schopni prokázat; pomohlo by zopakování experimentu s větším počtem měření v každé ze skupin (tedy s větší silou).

Proveďme nyní test hypotézy:

(K1) H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$;
H_1: Neplatí H_0.

(K2) Testovým kritériem bude veličina $\frac{est\ RMT}{est\ RUT}$. Odhad $est\ RUT = 1{,}8$ už byl nalezen (pro $VUT = 10$), odhadněme ještě RMT (pro $VMT = J - 1 = 4 - 1 = 3$):

$$est\ RMT = \frac{SSMT}{VMT} = \frac{\sum_{i=1}^{J} \frac{T_i^2}{N_i} - \frac{T^2}{n}}{J - 1} = \frac{129 - \frac{41^2}{14}}{4 - 1} = 3.$$

(K3) Při platnosti H_0 má veličina $\frac{est\ RMT}{est\ RUT}$ rozdělení F pro stupně volnosti $V_1 = 3$, $V_2 = 10$.

(K4) Pro $\alpha = 0{,}05$ příslušná kritická hodnota $F_k(3, 10)$ je na průsečíku třetího sloupce a desátého řádku tabulky 3.1 tj. $F_k(3, 10) = 3{,}71$.

(K5) Odpovídající F-hodnota kritéria je

$$\frac{est\ RMT}{est\ RUT} = \frac{3}{1{,}8} = 1{,}67 < F_k = 3{,}71;$$

H_0 tedy nezamítáme, rozdíly mezi pastami nebyly testem odhaleny.

Poznámka 3.4. Někdy se situace analýzy rozptylu popisuje tzv. **lineárním modelem** (tento přístup je podrobně rozpracován ve skriptu [28]): pro j-tou hodnotu i-té skupiny (třídy) platí

$$x_{ij} = \mu + \alpha_i + \varepsilon_{ij},$$

kde μ je střední hodnota celé populace, α_i je vliv (relativní změna) μ způsobený podmínkami měření ve skupině (= třídě) i, ε_{ij} je náhodná chyba, která má rozdělwní s rozptylem σ^2, což je stejný rozptyl jako rozptyl celé populace. Zde platí $\sum_{i=1}^{J} \alpha_i = 0$, $\sum_{j=1}^{N_i} \varepsilon_{ij} = 0$, a H_0, H_1 lze formulovat $H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_J = 0$, H_1: neplatí H_0. V tomto textu se analýze rozptylu ovšem nebudeme věnovat z pohledu lineárního modelu, ale z hlediska již představeného součtu čtverců.

Důležité drobnosti k zapamatování

Vrátíme-li se k situaci J skupin (tříd) měření, i-tá délky N_i, lze z měření ve všech skupinách dohromady spočítat jeden „velký“ součet čtverců

$$SS = \sum_{\text{všechno}} (x_{ij} - \overline{x})^2 = \left(\sum_{\text{všechno}} x_{ij}^2 \right) - \frac{\left(\sum_{\text{všechno}} x_{ij}\right)^2}{n} = \left(\sum_{\text{všechno}} x_{ij}^2 \right) - \frac{T^2}{n}, \tag{3.8}$$

kterému odpovídá $VSS = n-1$ stupňů volnosti (n hodnot závislých na jednom parametru $\overline{x}$). Z tabulky 3.3 je patrno, že

$$SS = SSMT + SSUT; \tag{3.9}$$
$$VSS = VMT + VUT \tag{3.10}$$

(podobné vztahy platily mezi celkovým rozptylem, vnitřním rozptylem a vnejším rozptylem v příkladu 2.20).

Pokud by se někomu zdály vzorec 3.8 a tabulka 3.3 příliš náročné, rád bych uvedl několik skutečností, které čtenáři pomohou si tyto vzorce zamilovat a ocenit krásu analýzy rozptylu :-)

Vztah mezi součtem čtverců a stupni volnosti: Ve vzorcích je přítomna následující symetrie – vzorec pro výpočet počtu stupňů volnosti je klíčem pro určení počtu členů v sumách odpovídajícího součtu čtverců:

- Celkový součet čtverců (3.8): $VSS = n - 1$, a pak při výpočtu SS od sumy n členů (sumy čtverců všech měření experimentu) odečteme jeden člen $\frac{T^2}{n}$.

- $SSUT$: Protože $VUT = n - J$, pak při výpočtu $SSUT$ odečítáme od sumy n členů (čtverců všech měření experimentu) sumu J členů ($\frac{T_i^2}{N_i}$ pro $j = 1, 2, \ldots, J$).
- $SSMT$: Protože $VMT = J - 1$, tak při výpočtu $SSMT$ odečítáme od sumy J členů $\frac{T_i^2}{N_i}$ jeden člen $\frac{T^2}{n}$.

Pravidlo „umocni a poděl“: Kdykoli v probíraných vzorcích umocňujeme určitý součet, dělíme jej počtem členů tohoto součtu. Konkrétně:

- x_{ij}^2: Umocňujeme jediný člen, čili dělíme jej jedničkou :-)
- $\frac{T_i^2}{N_i}$: Umocňujeme součet $T_i = \sum_{j=1}^{N_i} x_{ij}$, čili dělíme jej číslem N_i.
- $\frac{T^2}{n}$: Umocňujeme součet T všech n hodnot, čili dělíme jej číslem n.

3.1.2 Dvoufaktorová analýza rozptylu

Příklady a vzorce

Až dosud jsme se zabývali situacemi, kdy se hledal vliv jedné nezávislé proměnné na jednu závislou proměnnou. Nyní budeme sledovat vliv dvou **nezávislých** proměnných na třetí proměnnou **závislou** na předchozích dvou.

Příklad 3.5. Chceme zkoumat vliv dvou proměnných na výkon paměti. První proměnnou bude finanční motivace (zlepší se výkon paměti, když člověk dostane více zaplaceno?), druhou proměnnou bude délka pamatování (pamatuje si člověk po delší době méně?). Výkon paměti budeme měřit počtem zapamatovaných slov (ze dvaceti uvedených slov).

Při řešení příkladu bychom mohli provést dva oddělené experimenty jednorozměrné analýzy rozptylu: V jednom experimentu zkoumat vliv prvního faktoru (= finanční motivace) v různých podmínkách (podmínka 1: člověk dostane 1 kč za každé správně zapamatované slovo; podmínka 2: člověk dostane 100 kč za každé správně zapamatované slovo), ve druhém experimentu vliv druhého faktoru (= doba pamatování) za různých podmínek (podmínka 1: zkouška paměti ihned po naučení slov; podmínka 2: zkouška paměti hodinu po naučení slov; podmínka 3: zkouška paměti pět hodin po naučení slov).

Ale lepší je oba vlivy zkoumat najednou ve dvourozměrné analýza rozptylu, protože se z ní dovíme více informacíněž ze dvou jednorozměrných experimentů (můžeme studovat tzv. interakci obou nezávislých proměnných - jakýsi jejich vzájemný vliv na třetí proměnnou; za chvíli bude řečeno více).

Skloubením dvou podmínek faktoru 1 a tří podmínek faktoru 2 získáme $2 \cdot 3 = 6$ podmínek dvourozměrné (= dvoufaktorové) analýzy rozptylu. pro takto navržený experiment bylo náhodně vybráno šest skupin po třech lidech a získána data uvedená v tabulce na následující straně.

Vysvětlení označení v této tabulce faktorové analýzy typu $J \times K$:

J ... počet podmínek faktoru 1.

K ... počet podmínek faktoru 2.

T_{ij}, $\overline{x}_{ij}$... součet a průměr příslušný k podmínce ij.

T_{R_i}, N_{R_i} ... součet a počet hodnot příslušný i-té podmínce faktoru 1 (= v i-tém řádku faktorové tabulky).

T_{S_j}, N_{S_j} ... součet a počet hodnot příslušný j-té podmínce faktoru 2 (= v j-tém sloupci faktorové tabulky).

T, n ... celkový součet a počet hodnot v tabulce.

	0 hodin	1 hodina	5 hodin	
1 kč	$x_{111} = 6$ $x_{112} = 4$ $x_{113} = 5$ $T_{11} = 15$ $\overline{x}_{11} = 5$	$x_{121} = 4$ $x_{122} = 4$ $x_{123} = 4$ $T_{12} = 12$ $\overline{x}_{12} = 4$	$x_{131} = 5$ $x_{132} = 4$ $x_{133} = 3$ $T_{13} = 12$ $\overline{x}_{13} = 4$	$T_{R_1} = 39$
100 kč	$x_{211} = 11$ $x_{212} = 8$ $x_{213} = 8$ $T_{21} = 27$ $\overline{x}_{21} = 9$	$x_{221} = 6$ $x_{222} = 10$ $x_{223} = 8$ $T_{22} = 24$ $\overline{x}_{22} = 8$	$x_{231} = 6$ $x_{232} = 4$ $x_{233} = 5$ $T_{23} = 15$ $\overline{x}_{23} = 5$	$T_{R_2} = 66$
	$T_{S_1} = 42$	$T_{S_2} = 36$	$T_{S_3} = 27$	$T = 105$ $n = 18$

Zpracování naměřených dat shrneme do tří bodů:

a) Intervaly spolehlivosti: Potřebujeme vypočítat nejprve *est RUT* – máme k dispozici šest odhadů tohoto „vnitřního rozptylu“ (ze tří měření v každé ze šesti skupin lze určit jeden odhad), lze tedy spočítat jejich aritmetický průměr. Ovšem jednodušší je použít zde analogii vzorce pro součet čtverců rozptylu UT z tabulky 3.3:

$$est\ RUT = \frac{SSUT}{VUT} = \frac{\sum_{\text{všechno}} x_{ijk}^2 - \sum_{\text{skupiny}} \frac{T_{ij}^2}{N_{ij}}}{n - J \cdot K}; \tag{3.11}$$

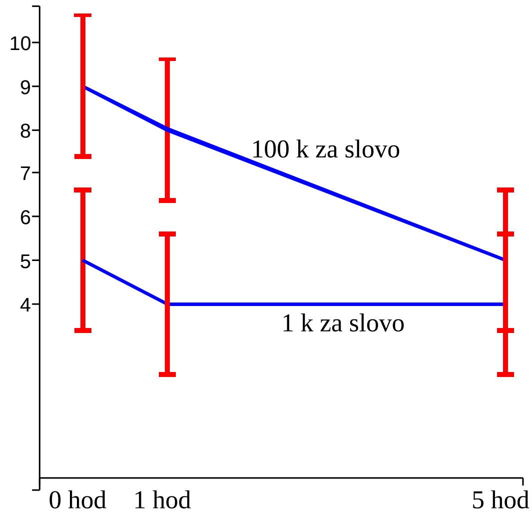
dosazením dostaneme

$$est\ RUT = \frac{701 - 681}{18 - 6} \doteq 1{,}667.$$

Tedy podle vzorce 3.7 máme

$$\mu_{ij} \in \overline{x}_{ij} \pm \sqrt{\frac{1{,}667}{3}} \cdot t_k(12) = \overline{x}_{ij} \pm 1{,}62.$$

Grafické znázornění průměrů a intervalů spolehlivosti bude nyní složitější než dříve: na ose y je vynášena závislá proměnná měřená počtem zapamatovaných slov, na ose x jedna z nezávislých proměnných, například doba pamatování. Druhou nezávislou proměnnou – finanční motivaci – do obrázku graficky znázorníme tím, že pro každou z jejich hodnot (1 kč , 100 kč) se nakreslí jeden graf závislosti počtu zapamatovaných slov na době pamatování. V jednom obrázku je tedy více křivek odpovídajících různým podmínkám faktoru 1:



Z intervalů spolehlivosti lze vyčíst, že oba faktory mají vliv na počet zapamatovaných slov: pro delší dobu pamatování počet zapamatovaných slov klesá (obě křivky), pro vyšší finanční motivaci je počet zapamatovaných slov větší (křivka „100 kč“ je výše než křivka „1 kč“).

b) Jednorozměrná ANOVA pro šest podmínek: Na chvíli zapomeňme, že šest „okének“ v tabulce vzniklo kombinací dvou nezávislých proměnných, a proveďme jednorozměrnou analýzu rozptylu, abychom zjistili, zda mezi uvedenými šesti skupinami (= třídami) existují významné statistické rozdíly:

$$H_0: \quad \mu_{11} = \mu_{12} = \mu_{13} = \mu_{21} = \mu_{22} = \mu_{23};$$

(H_1: neplatí H_0). Už máme spočteno $est\ RUT = 1,667$ (pro $VUT = 12$); dále použijeme analogii vzorce pro součet čtverců rozptylu MT z tabulky 3.3:

$$est\ RMT = \frac{SSMT}{VMT} = \frac{\sum_{\text{podmínky}} \frac{T_{ij}^2}{N_{ij}} - \frac{T^2}{n}}{J-1}; \tag{3.12}$$

po dosazení je (pro $VMT = 5$)

$$est\ RMT = \frac{681 - 612{,}5}{5} = \frac{68{,}5}{5} = 13{,}7.$$

Sestavme výpočty prováděné v našem statistickém testu do tabulky:

typ rozptylu	V	SS	$est\ R$	F-hodnota	F_k
celkový	17	88,5	–	–	–
MT	5	68,5	13,7	8,2	3,11
UT	12	20	1,667	–	–

Protože $\frac{13{,}7}{1{,}667} = 8{,}2 > 3{,}11$ ($F_k^{5,12}(\alpha = 0{,}05) = 3{,}11$), uzavíráme, že rozdíly mezi skupinami jsou statisticky významné – H_0 zamítáme.

Protože nás zajímá i vliv každé z obou nezávislých proměnných zvlášť a vliv interakce obou proměnných (obojí lze zjistit ze dvourozměrné ANOVA), v praxi v této situaci test b) neprovádíme. Je zde uveden pouze z pedagogických důvodů (aby se vidělo, že i tento test je možný, a aby bylo připomenuto rozdělení celkového „koláče rozptylu“, respektive celkového koláče součtu čtverců:

$$SS = 88{,}5 = SSUT + SSMT, \quad VSS = 17 = VUT + VMT).$$

c) **Dvourozměrná ANOVA: C1:** Testujme nejprve vliv faktoru 1 (= finanční motivace) na počet zapamatovaných slov; tento faktor má dvě podmínky vzniklé vždy součtem v příslušném řádku faktorové tabulky:

	6	4	5	
1 kč	4	4	4	$T_{R_1} = 39$
	5	4	3	$N_{R_1} = 9$
	11	6	6	
100 kč	8	10	4	$T_{R_2} = 66$
	8	8	5	$N_{R_2} = 9$
				$T = 105$
				$n = 18$

Rozhodujeme mezi $H_0 : \mu_{R_1} = \mu_{R_2}$ a H_1 : neplatí H_0. Kritériem testu bude opět podíl „vnějšího“ a „vnitřního“ rozptylu, ale roli vnějšího rozptylu bude nyní hrát rozptyl mezi řádky tabulky, proto označení RMR (analogicky $SSMR$ je příslušný součet čtverců, VMR příslušná volnost) – při jeho výpočtu použijeme analogii vzorce pro součet čtverců rozptylu MT z tabulky 3.3:

$$est\ RMR = \frac{SSMR}{VMR} = \frac{\sum_{\text{řádky}} \frac{T_{R_i}^2}{N_{R_i}} - \frac{T^2}{n}}{J - 1}; \tag{3.13}$$

po dosazení

$$est\ RMR = \frac{653 - 612{,}5}{2 - 1} = 40{,}5;$$

je vidět, že $VMR = J - 1 = 2 - 1 = 1$. Pak po dosazení do kritéria

$$\frac{est\ RMR}{est\ RUT} = \frac{40{,}5}{1{,}667} \doteq 24{,}25 > F_k^{1;12}(\alpha = 0{,}05) = 4{,}75.$$

Proto H_0 zamítáme, vliv faktoru 1 na počet zapamatovaných slov se prokázal – je statisticky významný.

C2: Dále testujme vliv faktoru 2 (= doby pamatování) na počet zapamatovaných slov (zase nás zajímá porovnání mezi jednotlivými sloupci faktorové tabulky):

0 hod	1 hod	5 hod	
6	4	5	
4	4	4	
5	4	3	
11	6	6	
8	10	4	
8	8	5	
$T_{S_1} = 42$	$T_{S_2} = 36$	$T_{S_3} = 27$	$T = 105$
$N_{S_1} = 6$	$N_{S_2} = 6$	$N_{S_3} = 6$	$n = 18$

Mezi hypotézami „$H_0 : \mu_{S_1} = \mu_{S_2} = \mu_{S_3}$“ a „$H_1$: Neplatí H_0“ rozhodne nyní kritérium $\frac{est\ RMS}{est\ RUT}$, kde $est\ RMS$ určíme ze vztahu

$$est\ RMS = \frac{SSMS}{VMS} = \frac{\sum \text{sloupce} \frac{T_{S_j}^2}{N_{S_j}} - \frac{T^2}{n}}{K-1}; \tag{3.14}$$

dosazením (pro $VMS = K - 1 = 3 - 1 = 2$) máme

$$est\ RMS = \frac{631{,}5 - 612{,}5}{3-1} = 9{,}5,$$

a tedy

$$\frac{est\ RMS}{est\ RUT} = \frac{9{,}5}{1{,}667} = 5{,}69 > F_k^{2,12}(\alpha = 0{,}05) = 3{,}8,$$

proto H_0 opět zamítáme, vliv faktoru 2 na počet zapamatovaných slov je statisticky významný.

C3: Sestavme dosud vypočtené hodnoty do tabulky:

typ rozptylu	V	SS	$est\ R$	F-hodnota	F_k
celkový	17	88,5	–	–	–
MT	5	68,5	–	–	–
– R	1	40,5	40,5	24,25	4,75
– S	2	19	9,5	5,69	3,8
– I	2	9	4,5	2,7	3,8
UT	12	20	1,667	–	–

Nyní je potřeba vysvětlit, kde se vzal v tabulce červený řádek. Tento řádek odpovídá vlivu **interakce mezi faktorem** 1 **a faktorem** 2 na počet zapamatovaných slov. Tato **interakce** je součástí rozdílnosti mezi různými skupinami měření, a tedy je součástí vnějšího rozptylu, proto budou platit vzorce

$$SSMT = SSMR + SSMS + SSI; \tag{3.15}$$
$$VMT = VMR + VMS + VI, \tag{3.16}$$

kde SSI je součet čtverců (rozptylu) interakce a VI je volnost (rozptylu) interakce. Pomocí těchto vztahů je možné SSI a VI snadno určit:

$$VI = 5 - 1 - 2 = 2, \quad SSI = 68{,}5 - 40{,}5 - 19 = 9.$$

Pak $est\ RI = \frac{SSI}{VI} = \frac{9}{2} = 4{,}5$, a tedy (testové kritérium $\frac{est\ RI}{est\ RUT}$ porovnáme s kritickou hodnotou $F_k^{VI,VUT}$))

$$\frac{est\ RI}{est\ RUT} = \frac{4{,}5}{1{,}667} \doteq 2{,}7 < F_k^{2,12}(\alpha = 0{,}05) = 3{,}8.$$

To znamená, že vliv interakce faktorů 1 a 2 na počet zapamatovaných slov není statisticky významný, hodnota kritéria tohoto pravostranného testu nepřesáhla kritickou hodnotu.

Dvourozměrná ANOVA tedy spočívá v testech C1, C2, C3, ve kterých hledáme vliv

- faktoru 1
- faktoru 2
- interakce faktorů 1 a 2

na závislou proměnnou.

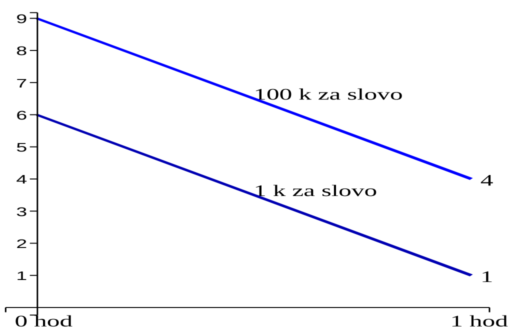
No dobře, ale co to ta **interakce** vlastně je? Podívejme se na příklad:

Příklad 3.6. Modifikujme nyní příklad 3.5 a předpokládejme dvě podmínky faktoru 1 (1 kč, 100 kč) a pouze dvě podmínky faktoru 2 (0 hod, 1 hod), čili dvourozměrnou ANOVA typu 2×2. Uvažme tři různé výsledky získaných dat z hlediska interakce (na vodorovnou osu vynášíme dobu zapamatování, na svislou osu počet zapamatovaných slov):

a) Žádná interakce: **(i)** Při motivaci 100 Kč si lidé pamatují průměrně o tři slova více než při motivaci 1 Kč (bez ohledu na délku pamatování)

(ii) Za jednu hodinu lidé zapomenou průměrně pět slov (bez ohledu na finanční motivaci)

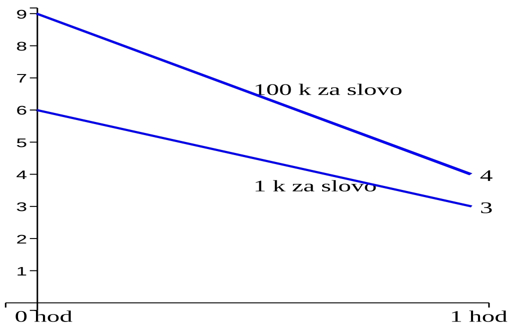
(iii) Shrnutí: vliv jednoho faktoru nezávisí na velikosti druhého faktoru; abychom určili hodnotu vlivu jednoho faktoru, nemusíme se zajímat o druhý faktor – jinými slovy, mezi faktory neexistuje interakce (geometricky: příslušné závislosti na obrázku jsou navzájem rovnoběžné).



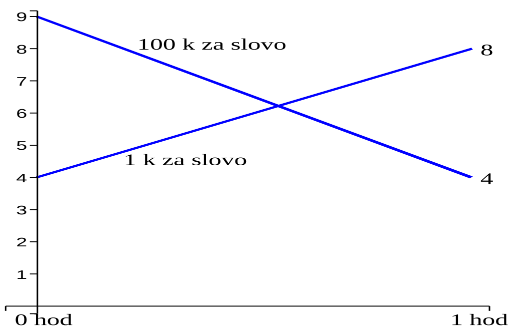
b) Slabá interakce: **(i)** O kolik víc si lidé průměrně pamatují při motivaci 100 Kč než při motivaci 1 Kč? To záleží na délce pamatování – bezprostředně po naučení slov je to průměrně o tři více, hodinu po naučení slov průměrně o jedno slovo více (viz obr.).

(ii) Kolik slov lidé zapomenou průměrně za jednu hodinu? To záleží na motivaci – při 1 Kč průměrně tři slova, při 100 Kč průměrně pět slov (viz obr.).

(iii) Shrnutí: vliv jednoho faktoru závisí na velikosti druhého faktoru; abychom určili vliv jednoho faktoru, musíme uvažovat i hodnotu druhého faktoru – jinými slovy, mezi faktory existuje interakce (geometricky: příslušné závislosti na obrázku jsou různoběžné).



c) Silná interakce: I když tento výsledek je v tomto příkladu nepravděpodobný, popišme, co by znamenala silná interakce:



(i) O kolik víc si lidé průměrně pamatují při motivaci 100 Kč než při motivaci 1 Kč? To záleží na délce pamatování – bezprostředně po naučení slov je to průměrně o pět více, ale hodinu po naučení slov průměrně (světe div se) o čtyři slova méně (viz obr.)!!

(ii) Kolik slov lidé zapomenou průměrně za jednu hodinu? To zcela závisí na motivaci – při 100 Kč zapomenou průměrně pět slov, při 1 Kč si průměrně vzpomenou ještě na čtyři další slova !!.

(iii) Shrnutí: Hodnota jednoho faktoru zcela (a klíčově) závisí na hodnotě druhého faktoru. Tomuto typu interakce říkáme úplná (= silná). příslušné grafy mají opačný sklon – za jedné podmínky je graf rostoucí, za jiné klesající.

Pojem interakce má také svá úskalí: V oddílku 2.5.6 (poznámka e)) jsme mluvili o porušení měřítka, které s otázkou interakce souvisí – **někdy volba hodnoty závislé proměnné není jednoznačná**. V našem příkladu jsme závislou proměnnou (= sílu paměti) měřili počtem zapamatovaných slov. Mohlo by se stát, že bychom tuto „psychologickou“ proměnnou měřili jinou matematickou veličinou (např. logaritmem počtu zapamatovaných slov nebo druhou mocninou počtu zapamatovaných slov) a příslušné křivky na obrázcích by pak měly jiný sklon, což by vedlo k jiné interakci. Protože tato volba veličiny není jednoznačná (více matematických veličin může být k tomuto účelu stejně dobrých), není jednoznačná ani hodnota interakce.

Zpředchozího odstavečku plyne, že **vhodnou transformací závislé proměnné lze někdy interakci odstranit** – zejména případ slabé interakce („mírně různoběžné grafy“) lze převést na případ bez interakce (rovnoběžné grafy), čili nemá smysl tvrdit, že interakce existuje nebo neexistuje. To ovšem neznamená, že nemá cenu interakce studovat – naopak, v některých případech lze interakci jednoznačně prokázat:

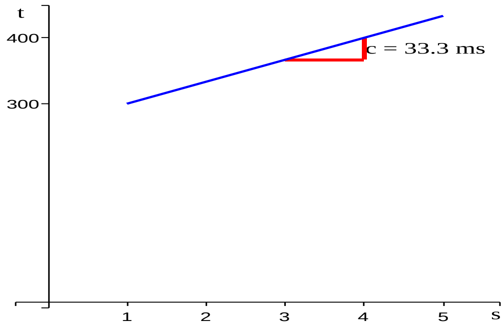
- např. když měřítko závislé proměnné je ekvivalentní měřítku nezávislé proměnné (příkladem takové závislé proměnné je např. doba čekání na výskyt jisté události – např. v našem příkladu by to znamenalo spíše naplánovat experiment tak, že by čas byl závislou proměnnou a vynášel by se na svislé ose, měřili bychom například dobu, po kterou si lidé jsou schopni udržet v paměti aspoň pět ze zapamatovaných slov; nezávislými proměnnými by mohly být např. motivace a věk);

- silnou interakci nelze transformací zachovávající monotonnost (= rostoucí funkce se transformuje na rostoucí funkci) nijak odstranit.

Podívejme se ještě na jeden příklad dvourozměrné analýzy rozptylu:

Příklad 3.7. V roce 1966 ve firmě Bell telephone představil Saul Sternberg způsob, jak testovat rychlost přístupu člověka do krátkodobé paměti: Testovaný má za úkol si zapamatovat určitou množinu písmen, např. $\{A, Q, M, T, G\}$. Potom se podrobí následující zkoušce: jsou mu prezentována různá písmena a on má stisknout jisté tlačítko, pokud dané písmeno pochází ze zapamatované množiny písmen (a jiné tlačítko, pokud se o písmeno z dané množiny nejedná). Měří se doba reakce mezi prezentací písmene a stiskem tlačítka.

Základní nezávislou proměnnou tohoto experimentu je s = počet písmen v „paměťové množině“. Bylo zjištěno, že závislost počtu písmen na reakční době je lineární (s větším počtem písmen předložených k zapamatování se lineárně prodlužuje reakční doba):



Popis obrázku: na vodorovnou osu vynášíme počet s slov v „paměťové množině“, na svislou osu reakční dobu t v milisekundách, přidáním jednoho slova do zapamatované množiny se zvýší reakční doba přibližně o $c = 33{,}3$ milisekund.

Proveďme nyní experiment zjišťující, jak se přístup do krátkodobé paměti liší vzhledem k typu zapamatované jednotky - vytvoříme dvourozměrnou ANOVA typu 2×3, faktor 1 (= typ pamatované jednotky) má dvě podmínky (**písmena** a **slova**); faktor 2 (= počet jednotek v zapamatované množině) má tři podmínky neboli hodnoty (1, 3, 5).

Získala se následující data, proveďte pro ně dvourozměrnou ANOVA:

	1	3	5	
	350	394	456	
	346	392	471	
	361	410	464	
	330	405	460	
písmena	342	400	450	$T_{R_1} = 6031$
	$T_{11} = 1729$	$T_{12} = 2001$	$T_{13} = 2301$	
	$\overline{x}_{11} = 346$	$\overline{x}_{12} = 400$	$\overline{x}_{13} = 460$	
	$est_{11}\ \sigma^2 = 128$	$est_{12}\ \sigma^2 = 56$	$est_{13}\ \sigma^2 = 63$	
	384	466	585	
	371	484	580	
	365	475	560	
	392	470	562	
slova	375	465	570	$T_{R_2} = 7104$
	$T_{21} = 1887$	$T_{22} = 2360$	$T_{23} = 2857$	
	$\overline{x}_{21} = 377$	$\overline{x}_{22} = 472$	$\overline{x}_{23} = 571$	
	$est_{21}\ \sigma^2 = 114$	$est_{22}\ \sigma^2 = 61$	$est_{23}\ \sigma^2 = 120$	
	$T_{S_1} = 3616$	$T_{S_2} = 4361$	$T_{S_3} = 5158$	$T = 13135$
				$n = 30$

Z naměřených dat spočítáme k dalšímu využití:

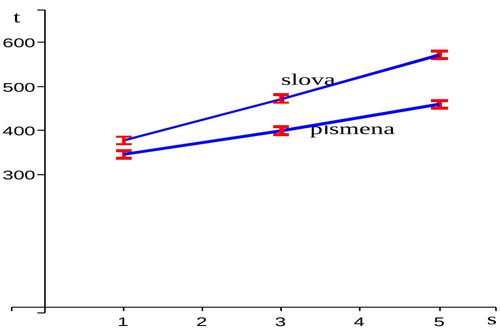
$$
\begin{aligned}
SSUT &= \sum_{\text{vše}} x_{ijk}^2 - \sum_{\text{podmínky}} \frac{T_{ij}^2}{5} = 2169; \quad VUT = 24; \\
SSMT &= \sum_{\text{podmínky}} \frac{T_{ij}^2}{5} - \frac{T^2}{n} = 165231; \quad VMT = 5; \\
SSMS &= \sum_{sloupce} \frac{T_{S_j}^2}{10} - \frac{T^2}{30} = 118933; \quad VMS = 2; \\
SSMR &= \sum_{\text{řádky}} \frac{T_{R_i}^2}{15} - \frac{T^2}{30} = 38377; \quad VMR = 1; \\
SSI &= SSMT - SSMS - SSMR = 7921; \quad VI = 2.
\end{aligned}
$$

a) intervaly spolehlivosti: „Vnitřní“ rozptyl lze vyjádřit jako aritmetický průměr šesti odhadů (protože v každé třídě (= podmínce) se vyskytuje stejný počet měření), nebo jako podíl

$$
est\ RUT = \frac{SSUT}{VUT} = \frac{2169}{24} \doteq 90{,}4;
$$

tj.

$$
\mu_{ij} \in \overline{x}_{ij} \pm \sqrt{\frac{90{,}4}{5}} \cdot t_k(24) = \overline{x}_{ij} \pm 8{,}8,
$$



Intervaly spolehlivosti (vyznačeny na obrázku) jsou malé vzhledem k rozdílům mezi průměry, tj. experiment má dostatečnou sílu (průměry jsou dobrými odhady středních hodnot). Závislosti jsou lineární, je tedy potvrzena platnost původní Sternbergovy teorie. Sklon je větší u slov než u písmen, což znamená, že doba kontroly u slov je obecně delší než doba kontroly u písmen.

b) testy hypotéz: Označme střední hodnoty příslušející jednotlivým podmínkám:

	1	3	5	
písmena	μ_{11}	μ_{12}	μ_{13}	μ_{R_1}
slova	μ_{21}	μ_{22}	μ_{23}	μ_{R_2}
	μ_{S_1}	μ_{S_2}	μ_{S_3}	

Nyní nás zajímají výsledky následujících tří testů:

Test 1: Vliv faktoru 1:
H_0 : $\mu_{R_1} = \mu_{R_2}$ (typ položky nemá vliv na reakční dobu);
H_1: není pravda, že $\mu_{R_1} = \mu_{R_2}$.

Test 2: Vliv faktoru 2:
H_0 : $\mu_{S_1} = \mu_{S_2} = \mu_{S_3}$ (velikost množiny nemá vliv na reakční dobu);
H_1: neplatí H_0.

Test 3: Vliv interakce faktorů 1 a 2:
H_0 : $\mu_{11} - \mu_{21} = \mu_{12} - \mu_{22} = \mu_{13} - \mu_{23}$ (rozdíl mezi typy položek nezávisí na velikosti množiny)
nebo ekvivalentně
H_0 : $\mu_{13} - \mu_{12} = \mu_{23} - \mu_{22}$ a současně $\mu_{12} - \mu_{11} = \mu_{22} - \mu_{21}$ (vliv velikosti množiny nezávisí na typu položky).
H_1: neplatí H_0.

Výsledky těchto testů lze vyčíst z tabulky pro dvourozměrnou ANOVA:

zdroj rozptylu	V	SS	*est R*	F-hodnota	F_k
celkový	29	167398	–	–	–
MT	5	165231	–	–	–
$-R$	1	38377	38377	425 4,26	
$-S$	2	118933	59466,5	658,54	3,4
$-I$	2	7921	3960,5	43,86	3,4
UT	24	2167	90,3	–	–

Ve všech třech testech zamítáme H_0, protože příslušná F-hodnota je vždy podstatně větší než F_k.

Závěry testů nám neřekly tolik, jako intervaly spolehlivosti. V kapitole 5 si řekneme něco o konkrétnějších alternativních hypotézách H_1.

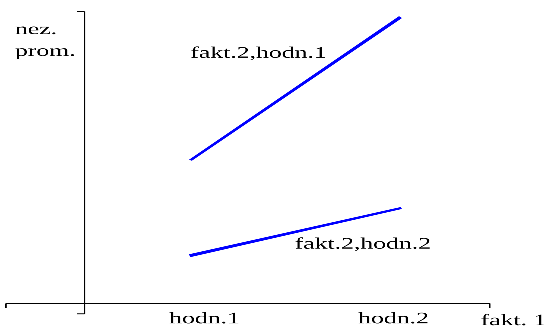
Z údajů tabulky lze potvrdit vztah mezi jednotlivými volnostmi a součty čtverců: celková volnost 29 je součtem „vnitřní“ volnosti 24 a „vnější“ volnosti 5, tu lze dále rozložit na součet volnosti 1 mezi řádky, volnosti 2 mezi sloupci a volnosti 2 interakce (mimo jiné pro volnost interakce vždy platí vzorec $VI = VMR \cdot VMS$). Podobně celkový součet čverců 167398 je součtem „vnitřního“ součtu čtverců 2167 a „vnějšího“ součtu čtverců 165231. Tento vnější součet čtverců lze dále rozložit jako součet $SSR = 38377$, $SSS = 118933$ a $SSI = 7921$.

Obecně lze říci, že **čím větší podíl má dílčí součet čtverců v celkovém koláči součtu čtverců, tím má příslušný zdroj větší vliv na celkový rozptyl**. To znamená pro náš příklad, že největší podíl na rozptylu má rozdílnost ve sloupcích ($SSS = 118933$ tvoří asi 71% celkového SS).

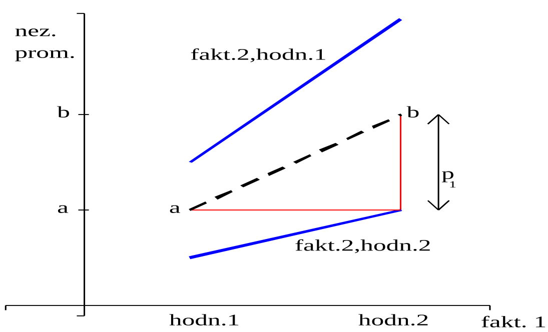
Zde jsme se zatím seznámili s dvourozměrnou ANOVA pouze pro případ stejného počtu měření v každé ze tříd ($N_{ij} = konstanta$). Případ nestejného počtu měření lze vyřešit buď vypuštěním „přesahujících hodnot", nebo tzv. neváženou analýzou rozptylu, kterou se zde nebudeme zabývat.

Dvě poznámky jako bonus

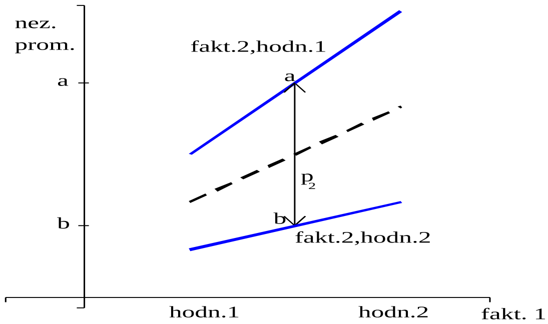
Poznámka 1: Grafická reprezentace. Z grafu popisujícího získaná data lze vyčíst nejen vliv interakce, ale také vliv faktoru 1 a faktoru 2 na závislou proměnnou. Uvažujme situaci slabé interakce popsané následujícím grafem:



U faktoru 1 není důležité číslo 1, ale to, že hodnoty faktoru jsou vynášeny na osu x (tj. faktorem 1 se rozumí faktor vynášený na osu x). Méru vlivu faktoru 1 na závislou proměnnou udává rozdíl p_1:



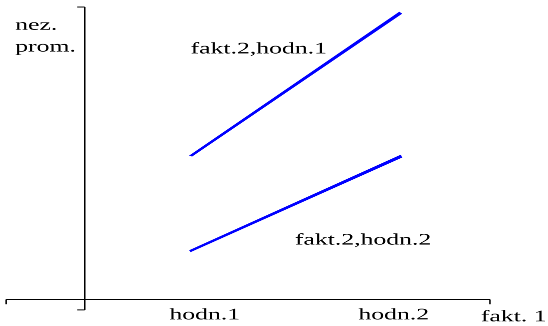
$a \ldots$ průměrný vliv faktoru 1 v podmínce 1; $b \ldots$ průměrný vliv faktoru 1 v podmínce 2; čárkovaná úsečka je grafem průměrného vlivu faktoru 1. Míru vlivu faktoru 2 (= faktor, který není vynášen na osu x) udává rozdíl p_2:



a ... průměrný vliv faktoru 2 v podmínce 1; b ... průměrný vliv faktoru 2 v podmínce 2.

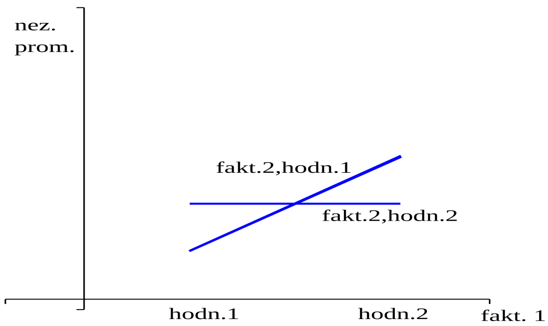
Abychom získali určitý cit na závěry plynoucí z grafické reprezentace, projdeme několik příkladů:

Příklad 3.8. Příklady analýzy typu 2×2:



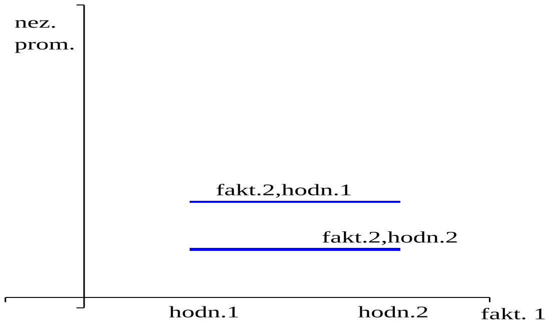
a)

Vliv interakce: ano. Vliv faktoru 1: ano. Vliv faktoru 2: ne.



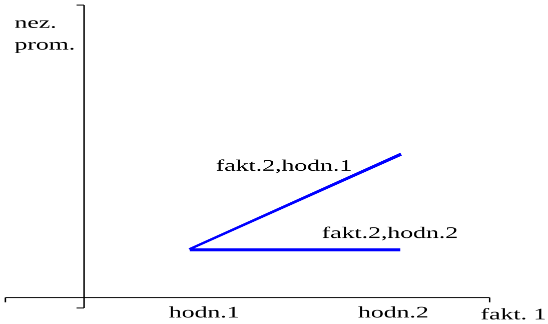
b)

Vliv interakce: ne. Vliv faktoru 1: ne. Vliv faktoru 2: ano.



c)

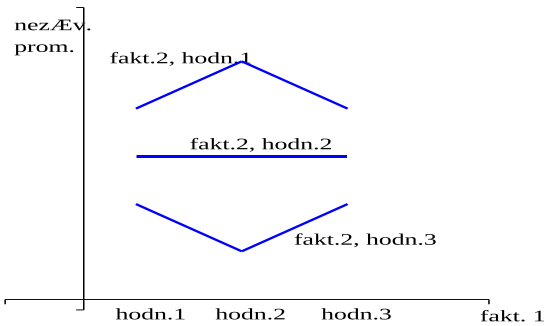
Vliv interakce: ano. Vliv faktoru 1: ne. Vliv faktoru 2: ano.



d)

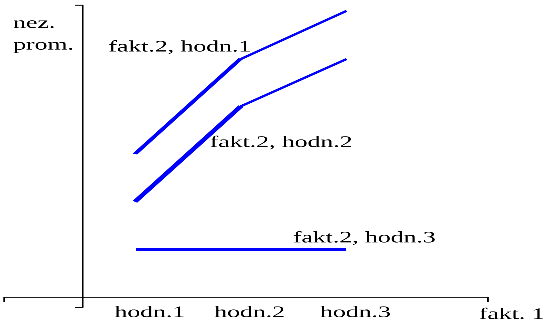
Vliv interakce: ano. Vliv faktoru 1: ano. Vliv faktoru 2: ano.

Příklad 3.9. Příklady analýzy typu 3×3:



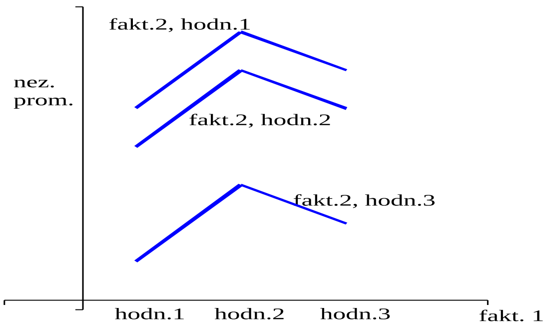
a)

Vliv interakce: ano. Vliv faktoru 1: ne. Vliv faktoru 2: ano.



b)

Vliv interakce: ano. Vliv faktoru 1: ano. Vliv faktoru 2: ano.



c)

Vliv interakce: ne. Vliv faktoru 1: ano. Vliv faktoru 2: ano.

Poznámka 2: Tři a více rozměrů u ANOVA. Uvažujme experiment měření síly paměti z příkladu 3.5. Kdybychom chtěli prokázat, zda se výsledky experimentu liší mezi muži a ženami, přidáním třetí nezávislé proměnné (pohlaví) jako třetího faktoru bychom dospěli k trojrozměrné (= třífaktorové) analýze rozptylu typu $2 \times 3 \times 2$, tj. kombinací faktorových úrovní by vzniklo dvanáct různých podmínek.

Pak bychom potřebovali zpracovat data ze dvanácti skupin měření (např. $N_{ijk} = 3$ je počet měření v každé skupině) a vyplnit faktorovou tabulku (respektive dvojici tabulek)

fakt.3,podm.1: ženy	fakt.2: 0 hodin	fakt.2: 1 hodina	fakt.2: 5 hodin
fakt.1: 1 kč			
fakt.2: 100 kč			

fakt.3,podm.2: muži	fakt.2: 0 hodin	fakt.2: 1 hodina	fakt.2: 5 hodin
fakt.1: 1 kč			
fakt.2: 100 kč			

Tabulka výsledků trojrozměrné ANOVA by pak měla následující tvar (uvedu jen první dva sloupce):

zdroj rozptylu	V
celkový	35
MT	11
– faktor 1	1
– faktor 2	2
– faktor 3	1
– interakce F_1, F_2	2
– interakce F_1, F_3	1
– interakce F_2, F_3	2
– interakce F_1, F_2, F_3	2
UT	24

(platí, že počet stupňů volnosti interakce je roven součinu volností faktorů do této interakce zahrnutých)

V tabulce je vhledem k předchozímu jediná nová veličina, a sice současná interakce faktorů F_1, F_2, F_3 – vyjadřuje míru, do jaké závisí libovolná dvourozměrná interakce na hodnotě třetího faktoru, tj. interakce F_1, F_2 na hodnotě faktoru F_3, nebo ekvivalentně interakce F_1, F_3 na hodnotě faktoru F_2, nebo ekvivalentně interakce F_2, F_3 na hodnotě faktoru F_1.

Více než tři rozměry u ANOVA jsou možné, ale celá situace je už dost nepřehledná, čili se moc neužívá :-)

3.1.3 Experiment opakovaného měření

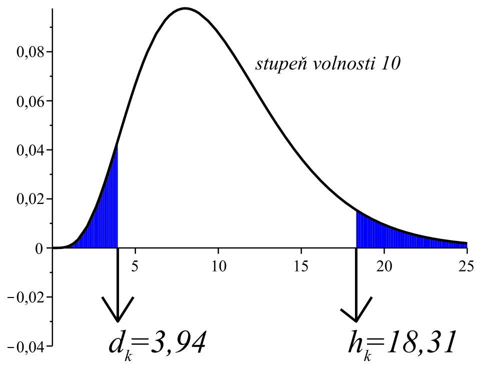
V oddílech 3.1, 3.1.2 jsme se zabývali experimentem typu „**více vzorků jednou**“. Nyní se zaměříme na typ „**jeden vzorek vícekrát**“, neboli tzv. **experiment opakovaného měření** (jedna skupina je zkoumána a měřena za různých podmínek).

Rozdíl mezi pojetím „více vzorků jednou“ a „jeden vzorek vícekrát“

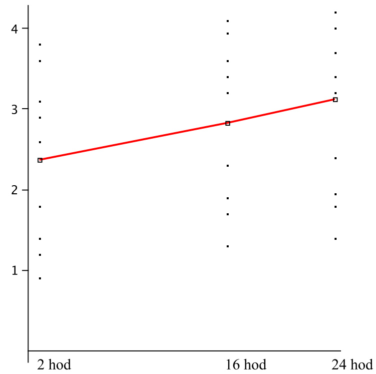
Příklad 3.10. Zajímá nás vliv nedostatku spánku na rychlost řešení problémů. Nezávislou proměnnou je zde doba, která uplynula od posledního spánku (3 podmínky: 2 hod, 16 hod, 24 hod). Závislou proměnnou je doba řešení úlohy jistého typu.

Kdybychom provedli experiment typu „**více vzorků jednou**“ a získali např. následující data (27 lidí je rozděleno do tří skupin po devíti, byla získána měření vyznačená v tabulce – čtvercem je zaznačen průměr v každé ze tří skupin měření).
(na vodorovné ose je vynášena doba od ukončení posledního spánku v hodinách, na svislé doba řešení úlohy jistého typu v minutách). Např. pomocí jednorozměrné ANOVA lze pro tento experiment učinit závěr, že průměry $\overline{x}_i$ se liší pouze málo, zatímco rozptyl měření v každé ze tří skupin je velký – to znamená, že rozdíly mezi $\overline{x}_i$ jsou způsobeny náhodnými

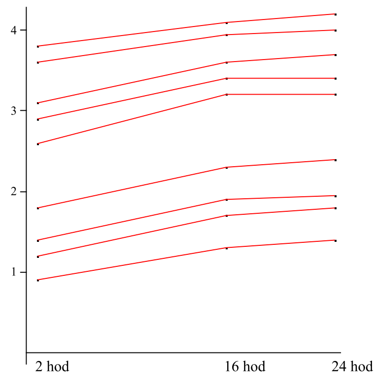
vlivy a statistický test by neodhalil významný rozdíl mezi jednotlivými podmínkami (2, 16 a 24 hodin beze spánku).



Ovšem kdybychom prováděli experiment typu „**jeden vzorek vícekrát**", vybrali bychom třeba devět lidí a s touto skupinou vybraných devíti bychom provedli měření ve třech různých podmínkách (situacích) – a sice 2, 16 a 24 hodin po ukončení posledního spánku. Celkem vzato by výsledkem měření mohly být podobné nebo i stejné hodnoty jako na předchozím obrázku, ale hodně záleží na tom, jaké rozdíly vykazují různé situace měření u jednoho konkrétního člověka – znázorníme toto spojením příslušných tří hodnot měření u každého vybraného. Nyní velký rozdíl spočívá v tom, zda hodnoty měření týchž lidí v různých podmínkách jsou svázány ve stylu



(vliv doby bdělosti na rychlost řešení úlohy je u každého člověka jiný – u někoho se s rostoucí dobou bdělosti rychlost řešení zmenšuje, u jiného zvětšuje, závěr: vliv doby bdělosti na dobu řešení úlohy není významný ani jednoznačný) nebo ve stylu



(u všech zkoumaných lidí se doba řešení úlohy s rostoucí dobou bdělosti zvyšuje, závěr: protože nárůst doby řešení úlohy se zvyšuje u každého pozorovaného člověka, je vliv doby bdělosti významný).

Klasická ANOVA (více vzorků jednou) by zamítla hypotézu o významnosti v obou situacích dokumentovaných na posledních dvou obrázcích, kdežto test „jeden vzorek vícekrát“, který se v tomto oddílku chystáme popsat, samozřejmě bude moci mezi oběma situacemi rozlišit. Snad byl na právě uvedeném příkladu formulován hlavní rozdíl mezi oběma pojetími:

více vzorků jednou: H_0 zamítneme, jestliže rozdíly mezi podmínkami jsou velké vzhledem k rozdílům mezi jedinci v každé skupině měření.

jeden vzorek vícekrát: H_0 zamítneme, jestliže rozdíly mezi podmínkami vykazují podobnou míru změny u každého pozorovaného jedince (bez ohledu na to, jak se od sebe liší měření v jedné skupině).

V právě popsaném příkladu 3.10 bylo tedy rozumné provádět experiment typu „**jeden vzorek vícekrát**“ – ten totiž celkem jistě detekuje vliv doby bdělosti na schopnost soustředění a řešení úloh. Následující příklad je zajímavý svým tvrzením, že v některých situacích je možné navrhnout experiment obou typů.

Příklad 3.11. Společnost EVERGREEN vydává knihu s názvem „V zajetí vlastních představ“ a ráda by zjistila, zda se tato kniha bude lépe prodávat s bílým, nebo černým obalem. Může provést experiment jedním ze dvou způsobů (když předtím vytiskla jistý testovací počet knih s bílým a jistý testovací počet knih s černým obalem):

a) Náhodně vybere dvacet knihkupectví a do každého distribuuje stejný počet knih černých a stejný počet knih bílých. Po určité době provede t-test typu „jedna skupina dvakrát“ pro počty prodaných knih.

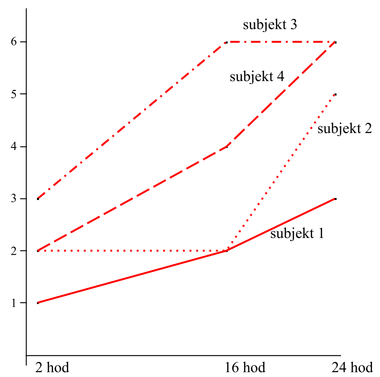
b) Náhodně vybere dvacet knihkupectví a rozdělí je na dvě skupiny po deseti. Do první skupiny distribuuje pouze černé knihy, do druhé skupiny pouze bílé. Po určité době provede párový t-test pro počty prodaných knih.

Případ $N_{ij} = 1$

Příklad 3.12. Předpokládejme, že experimentu typu „jeden vzorek vícekrát" (experiment opakovaného měření) z příkladu 3.10 se účastní čtyři lidé (= čtyři subjekty). Byla získána data sestavená do ANOVA tabulky typu 4 × 3:

	2 hod	16 hod	24 hod	
člověk 1	$x_{11} = T_{11} = \overline{x}_{11} = 1$	$x_{12} = T_{12} = \overline{x}_{12} = 2$	$x_{13} = T_{13} = \overline{x}_{13} = 3$	$T_{R_1} = 6$
člověk 2	$x_{21} = T_{21} = \overline{x}_{21} = 2$	$x_{22} = T_{22} = \overline{x}_{22} = 2$	$x_{23} = T_{23} = \overline{x}_{23} = 5$	$T_{R_2} = 9$
člověk 3	$x_{31} = T_{31} = \overline{x}_{31} = 3$	$x_{32} = T_{32} = \overline{x}_{32} = 6$	$x_{33} = T_{33} = \overline{x}_{33} = 6$	$T_{R_3} = 15$
člověk 4	$x_{41} = T_{41} = \overline{x}_{41} = 2$	$x_{42} = T_{42} = \overline{x}_{42} = 4$	$x_{43} = T_{43} = \overline{x}_{43} = 6$	$T_{R_4} = 12$
	$T_{S_1} = 8$ $\overline{x}_{S_1} = 2$	$T_{S_2} = 14$ $\overline{x}_{S_2} = 3{,}5$	$T_{S_3} = 20$ $\overline{x}_{S_3} = 5$	$T = 42$ $n = 12$ $N_{ij} = 1$

Uvedená data lze graficky znázornit



Na základě příkladu 3.10 a terminologie dvoufaktorové ANOVA lze říci, že všechny subjekty vykazují stejné (obdobné) chování, pokud interakce SUBJEKT–PODMÍNKA je malá, tj. grafické průběhy jednotlivých jedinců (subjektů) jsou „rovnoběžné" ve smyslu posledního obrázku z příkladu 3.9.

Čím menší je tato interakce, tím větší váhu má tvrzení, že rozdíly mezi skupinami jsou způsobeny vlivem nezávislé proměnné PODMÍNKA. Čím větší bude interakce SUBJEKT–PODMÍNKA, tím více budeme mít oprávnění si myslet, že rozdíly mezi podmínkami jsou způsobeny pouze náhodnými vlivy. **Roli kritéria míry důkazů testu tedy místo RUT převezme RI, tj. odpovídající F-hodnota vznikne jako**

$$\frac{est\ RMT}{est\ RI}$$

(člen RI převzal funkci člene RUT ve jmenovateli zlomku kritéria z oddílů 3.1, 3.1.2). Pokračujme nyní v řešení našeho příkladu:

a) Dvourozměrná ANOVA: Od situace „**více vzorků jednou**“ se liší typ „**jedna skupina vícekrát**“ tím, že $N_{ij} = 1$, tj. volnost $V_{ij} = 0$, a tak $VUT = \sum V_{ij} = 0$, a tedy nelze vůbec počítat *est* RUT. Ostatní odhady už počítat lze – vypočteme je podle stejných vzorců jako u dvoufaktorové (=dvourozměrné) ANOVA ve 3.1.2:

$$
\begin{aligned}
SSMT &= \sum \frac{T_{ij}^2}{N_{ij}} - \frac{T^2}{n} = 184 - 147 = 37, \quad VMT = 12 - 1 = 11;\\
SSMS &= \sum_{\text{sloupce}} \frac{T_{S_j}^2}{4} - \frac{T^2}{n} = 165 - 147 = 18, \quad VMS = 3 - 1 = 2;\\
SSMR &= \sum_{\text{řádky}} \frac{T_{R_i}^2}{3} - \frac{T^2}{n} = 162 - 147 = 15, \quad VMR = 4 - 1 = 3;\\
SSI &= SSMT - SSMS - SSMR = 4, \quad VI = VMS \cdot VMR = 6.
\end{aligned}
$$

Výsledky sestavme do tabulky ANOVA:

zdroj rozptylu	V	SS	*est* R	F*-hodnota*	F_k
MT	11	37	–	–	–
$-R$	3	15	5	–	–
$-S$	2	18	9	$\frac{9}{0{,}667} = 13{,}49$	5,14
$-I$	6	4	0,667	–	–

Podle řádku „S“ (sloupce) tabulky tedy zamítáme H_0 a uzavíráme, že doba bdělosti od posledního spánku má vliv na kvalitu řešení problémů. Dále řádek „R“ (tj. rozdíly mezi jedinci = subjekty) nás v situaci opakovaného měření obyčejně nezajímá. To, co nás zajímá, je RMS, tj. rozdíly mezi jednotlivými podmínkami.

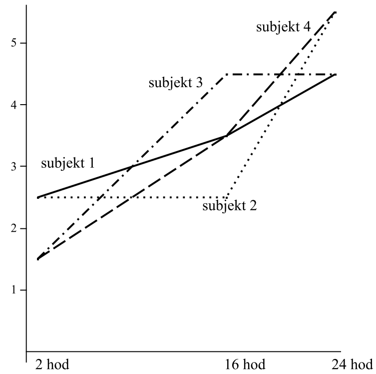
b) Výpočet intervalů spolehlivosti: Odvodíme nejprve vzorec pro interval spolehlivosti v této situaci: Ze získaných dat

subjekt	2 hod	16 hod	24 hod	průměr	odchylka od $\overline{x} = 3{,}5$
člověk 1	1	2	3	2	1,5
člověk 2	2	2	5	3	0,5
člověk 3	3	6	6	5	−1,5
člověk 4	2	4	6	4	−0,5
	$T_{S_1} = 8$	$T_{S_2} = 14$	$T_{S_3} = 20$	$\overline{x} = 3{,}5$	

nejprve odstraníme rozdíly způsobené růzností lidí (protože toto odstranění různosti lidí nemá vliv na experiment typu **jeden vzorek vícekrát**); a sice tak, že k hodnotám u každého člověka přičteme odchylku jeho průměru od celkového průměru – k hodnotám v prvním řádku přičteme 1,5, ve druhém řádku přičteme 0,5, ve třetím odečteme 1,5, ve čtvrtém odečteme 0,5,

subjekt	2 hod	16 hod	24 hod	průměr subjektu
člověk 1	2,5	3,5	4,5	3,5
člověk 2	2,5	2,5	5,5	3,5
člověk 3	1,5	4,5	4,5	3,5
člověk 4	1,5	3,5	5,5	3,5
	$T_{S_1} = 8$	$T_{S_2} = 14$	$T_{S_3} = 20$	

a tedy hodnoty u každého subjektu mají nyní tentýž průměr, kdežto součty T_{S_j} jednotlivých podmínek zůstávají nezměněny. Takto upravená data lze graficky znázornit následovně:



(všechny čtyři křivky mají stejný sklon jako na předchozím obrázku, ale jsou posunuty blíže k sobě). Tím, že jsme k datům přidali požadavek, že průměr každého řádku musí být stejný, klesne počet stupňů volnosti z devíti na šest (požadujeme-li totiž, aby součty řádků i sloupců v tabulce 3×4 byly zachovány, můžeme volně vyplnit pouze šest hodnot, ostatní už jsou jednoznačně určeny).

Když nyní vypočteme „SSUT“ pro upravenou tabulku hodnot, dostaneme

$$„\text{SSUT}“ = \sum x_{ij}^2 - \sum_{\text{podm.}} \frac{T_j^2}{4} = 169 - 165 = 4,$$

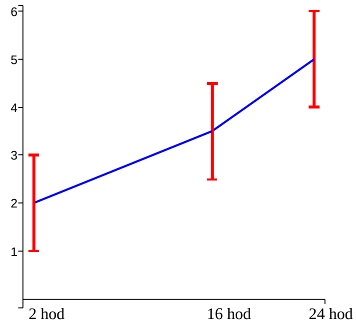
a to je právě i hodnota SSI. Celkem tedy pro odhad intervalu spolehlivosti bereme $est\ R_{\overline{X}} = \frac{est\ R_I}{N} = \frac{0{,}667}{4}$, tj.

$$\mu \in \overline{X_j} \pm t_\alpha^{2q}(VI) \cdot \sqrt{\frac{est\ RI}{N}}. \tag{3.17}$$

V našem příkladu

$$\mu_j \in \overline{X_j} \pm \sqrt{\frac{0{,}667}{4}} \cdot 2{,}447 = \overline{X_j} \pm 1{,}00,$$

a tedy $\mu_1 \in 2 \pm 1$, $\mu_2 \in 3{,}5 \pm 1$ a $\mu_3 \in 5 \pm 1$. Průměry i intervaly spolehlivosti lze znázornit graficky:



Případ $N_{ij} > 1$

Příklad 3.13. Uvažujme experiment z příkladu 3.10, kde každý člověk je podroben každé podmínce dvakrát. Jsou získána data:

	2 hodiny	16 hodin	24 hodin	
člověk 1	$x_{111} = 1$ $x_{112} = 1$ $T_{11} = 2$	$x_{121} = 3$ $x_{122} = 1$ $T_{12} = 4$	$x_{131} = 2$ $x_{132} = 4$ $T_{13} = 6$	$T_{R_1} = 12$
člověk 2	$x_{211} = 1$ $x_{212} = 3$ $T_{21} = 4$	$x_{221} = 2$ $x_{222} = 2$ $T_{22} = 4$	$x_{231} = 6$ $x_{232} = 4$ $T_{23} = 10$	$T_{R_2} = 18$
člověk 3	$x_{311} = 3$ $x_{312} = 3$ $T_{31} = 6$	$x_{321} = 6$ $x_{322} = 6$ $T_{32} = 12$	$x_{331} = 5$ $x_{332} = 7$ $T_{33} = 12$	$T_{R_3} = 30$
člověk 4	$x_{411} = 1$ $x_{412} = 3$ $T_{41} = 4$	$x_{421} = 6$ $x_{422} = 2$ $T_{42} = 8$	$x_{431} = 6$ $x_{432} = 6$ $T_{43} = 12$	$T_{R_4} = 24$
	$T_{S_1} = 16$	$T_{S_2} = 28$	$T_{S_3} = 40$	$T = 84$

V právě uvedeném příkladu už (na rozdíl od situace $N_{ij} = 1$) má smysl počítat $SSUT$:

$$SSUT = \sum_{\text{všichni}} x_{ijk}^2 - \sum_{i,j} \frac{T_{ij}^2}{2} = 20, \quad VUT = \text{počet tříd} = 12.$$

$SSUT$ má zde však jiný význam než v oddílu dvourozměrné (= dvoufaktorové) analýzy rozptylu: nyní se nejedná o rozptyl hodnot pro různé subjekty (= různé lidi), ale rozptyl hodnot jednoho subjektu, podrobeného několikrát za sebou stejným podmínkám.

Další součty čtverců vypočteme přirozeně (VMT = počet „situací" minus jedna = 11, VMR = počet řádků minus jedna = 3, VMS = počet sloupců minus jedna = 2):

$$SSMT = \sum_{i,j} \frac{T_{ij}^2}{2} - \frac{T^2}{n} = 368 - 294 = 74; \quad VMT = 11;$$

$$
\begin{aligned}
SSMR &= \sum_{\text{řádky}} \frac{T_{R_i}^2}{6} - \frac{T^2}{24} = 30; \quad VMR = 3; \\
SSMS &= \sum_{\text{sloupce}} \frac{T_{S_j}^2}{8} - \frac{T^2}{24} = 36; \quad VMS = 2; \\
SSI &= SSMT - SSMR - SSMS = 8; \quad VI = VMS \cdot VMR = 2 \cdot 3 = 6.
\end{aligned}
$$

Intervaly spolehlivosti pro μ_j u jednotlivých podmínek:

$$
\mu_j \in \overline{X}_j \pm t_\alpha(VI) \cdot \sqrt{\frac{est\ RI}{N_j}}, \tag{3.18}
$$

kde $N_j = 8$ (= počet všech měření v daném sloupci). Čili

$$
\mu_j \in \overline{X}_j \pm 2{,}447 \cdot \sqrt{\frac{\frac{8}{6}}{8}} \doteq \overline{X}_j \pm 1.
$$

Výsledek je stejný jako u příkladu 3.10, protože průměry hodnot v každé z dvanácti „buněk“ jsou stejné jako hodnoty měření v př. 3.10.

Testování hypotéz:

a) H_0: doba uplynutá od posledního spánku nemá vliv na dobu řešení problémů;
H_1: doba uplynutá od posledního spánku **má** vliv na dobu řešení problémů.
Testovým kritériem je stejně jako v případě $N_{ij} = 1$ podíl

$$
\frac{est\ RMS}{est\ RI} = \frac{18}{1{,}333} \doteq 13{,}5 > F_k(2; 6) = 5{,}14;
$$

tedy pro $\alpha = 0{,}05$ zamítáme hypotézu H_0.

b) H_0: doba řešení problémů u různých subjektů se neliší;
H_1: doba řešení problémů u různých subjektů **se liší**.
Testovým kritériem je podíl

$$
\frac{est\ RMR}{est\ RUT} = \frac{10}{1{,}666} \doteq 6{,}0 > F_k(3; 12) = 3{,}69;
$$

tedy pro $\alpha = 0{,}05$ zamítáme H_0. To znamená, že pokud se subjekty neliší (ve smyslu rozdílu středních hodnot jednotlivých rozdělení), pravděpodobnost výsledku získaná naším měřením je menší než 0,05.

c) H_0: mezi subjekty a podmínkami není interakce;
H_1: mezi subjekty a podmínkami **dochází** k interakci.
Testovým kritériem je

$$
\frac{est\ RI}{est\ RUT} = \frac{1{,}333}{1{,}666} \doteq 1{,}8 < F_k(6; 12) = 3{,}00;
$$

tedy pro $\alpha = 0{,}05$ hypotézu H_0 nezamítáme, interakce není statisticky významná.

Právě uvedené testy b),c) se používají jen málokdy, ovšem v případě $N_{ij} = 1$ nebyly vůbec proveditelné. Nás ale zajímá především výsledek testu a). Všechna data testů lze opět shrnout do tabulky:

zdroj rozptylu	V	SS	$est\ R$	F-hodnota	F_k
celkový součet	23	94	–	–	–
MT	11	74	–	–	–
$-\ R$	3	30	10	$\frac{10}{1{,}6} \doteq 6{,}0$	3,49
$-\ S$	2	36	18	$\frac{18}{1{,}333} \doteq 13{,}5$	5,14
$-\ I$	6	8	1,333	0,8	3,0
UT	12	20	1,666	–	–

Několik drobných poznámek

Poznámka 1. Tvar kritéria $F-$testu. Vraťme se k příkladu 3.13 a přidejme matematické vysvětlení toho, proč se liší kritéria v testech a), b), c). Označme

- $\sigma_e^2 \ldots$ rozptyl hodnot způsobený výkyvy výkonu jedince;
- $\sigma_R^2 \ldots$ rozptyl hodnot mezi různými jedinci (řádky);
- $\sigma_S^2 \ldots$ rozptyl mezi různými podmínkami (sloupci);
- $\sigma_I^2 \ldots$ rozptyl způsobený interakcí obou faktorů.

Pak můžeme v tabulce doplnit očekávané (= střední) hodnoty pozorovaných odhadů:

zdroj rozptylu	V	SS	$est\ R$	očekávaná hod. $estR$
celkový součet	$n-1$	SS	–	–
MT	$JK-1$	$SSMT$	–	–
$-\ R$	$J-1$	$SSMR$	$est\ RMR$	$\sigma_e^2 + n \cdot J \cdot \sigma_R^2$
$-\ S$	$K-1$	$SSMS$	$est\ RMS$	$\sigma_e^2 + n \cdot K \cdot \sigma_S^2 + n \cdot \sigma_I^2$
$-\ I$	$(J-1)\cdot(K-1)$	SSI	$est\ RI$	$\sigma_e^2 + n \cdot \sigma_I^2$
UT	$n-JK$	$SSUT$	$est\ RUT$	σ_e^2

Z tabulky (jejíž obsah nebudeme dokazovat) vidíme, že $est\ RMS$ a $est\ RI$ jsou odhadu téhož rozptylu právě tehdy, když $\sigma_S^2 = 0$. Ale podmínkou $\sigma_S^2 = 0$ je jen jinak vyjádřena platnost hypotézy H_0. Čili pokud H_0 platí, jsou $est\ RMS$ a $est\ RI$ odhady téhož rozptylu, čili jejich podíl lze popsat rozdělením F.

Dále z tabulky vidíme, že pokud $\sigma_R^2 = 0$ (respektive $\sigma_I^2 = 0$), tak odhady $est\ RUT$, $est\ RMR$ (resp. odhady $est\ RUT$, $est\ RI$) jsou odhady téhož rozptylu σ_e^2 a jejich podíl lze popsat F-rozdělením.

Poznámka 2. Pevné versus náhodné vlivy. Při plánování experimentu se někdy používá terminologie pevných a náhodných vlivů. V příkladech této kapitoly byla pevným vlivem doba bdělosti od posledního spánku, náhodným vlivem byli jednotliví lidé podrobení testu. Jinými slovy, za náhodný vliv či náhodnou proměnnou označujeme ten faktor, jehož výsledky chceme vztáhnout na celou populaci, nikoli jen na vybrané jedince.

Ad příklad 3.11: náhodným vlivem jsou zde knihkupectví, protože výsledky testu chceme vztáhnout nejen na dvacet vybraných, ale na všechna knihkupectví.

Poznámka 3. Třírozměrný experiment. Následující tabulka ukazuje analýzu dvou pevných vlivů A, B a jednoho náhodného vlivu S – náhodným vlivem S jsou pozorovaní jedinci:

zdroj rozptylu	V	SS	testové kritérium
A	$J-1$	SSA	$\frac{est\ RA}{est\ R(A \times S)}$
B	$K-1$	SSB	$\frac{est\ RB}{est\ R(B \times S)}$
S (=subjekty)	$n-1$	SSS	
$A \times S$	$(J-1)(n-1)$	$SS(A \times S)$	
$B \times S$	$(K-1)(n-1)$	$SS(B \times S)$	
$A \times B$	$(J-1)(K-1)$	$SS(A \times B)$	$\frac{est\ R(A \times B)}{est\ R(A \times B \times S)}$
$A \times B \times S$	$(J-1)(K-1)(n-1)$	$SS(A \times B \times S)$	

Jakýkoli pevný vliv (nebo interakce pevných vlivů) je testován pomocí podílu odhadu rozptylu tohoto vlivu a odhadu rozptylu interakce tohoto vlivu se subjekty.

Vyšší než třetí dimenze jsou také zpracovatelné, ovšem už značně nepřehledné.

Také je možné studovat jakési smíšené experimenty, z nichž několik proměnných je pevných a několik náhodných (např. muži/ženy jsou příkladem dvou typů subjektů, tedy dvou náhodných vlivů). Tyto modely jsou ovšem mimo rámec tohoto textu.

Pojmy k zapamatování

– Tato kapitola je přirozeným pokračováním otázek studovaných v kapitole 2.5 – zatímco kap. 2.5 se zabývala porovnáním středních hodnot dvou skupin měření (přitom každá ze skupin měření závislé veličiny byla získána pro jinou hodnotu nezávislé veličiny), kap. 3 studuje porovnání středních hodnot tří a více skupin měření (byly získány tři a více souborů měření závislé veličiny, každá skupina měření byla získána pro jinou hodnotu nezávislé veličiny). Přesněji řečeno, toto *zobecnění počtu hodnot nezávislé veličiny* je tématem oddílu 3.1. Oddíl 3.1.2 je zobecněním kapitoly 2.5 v jiném smyslu – zatímco 3.1 studuje vztah jedné nezávislé veličiny a jedné veličiny závislé na té první (tzv. **jednofaktorová analýza rozptylu = jednofaktorová ANOVA**), oddíl 3.1.2 se zabývá studiem vztahu jedné veličiny závislé na dvou (nebo i více) nezávislých veličinách (tzv. **dvoufaktorová ANOVA nebo vícefaktorová ANOVA**) Jedná se tedy o *zobecnění počtu nezávislých veličin.* No a do třetice opět situace jiného druhu – zatímco 3.1 a 3.1.2 se zabývaly studiem navzájem nezávislých skupin měření, tak oddíl 3.1.3 (**experiment opakovaného měření**) studuje skupiny měření, které jsou spolu navzájem provázány – stejný jedinec je podroben měření v každé z podmínek (vlastně *zobecnění situace párového testu z kapitoly 2.5 pro tři a více podmínek měření*).

– Všechny úvahy v této kapitole souvisí s rozptylem (ANOVA = analysis of variance = analýza rozptylu) – jak už bylo řečeno v příkladu 2.20, existuje jednak vnitřní rozptyl (RUT = rozptyl uvnitř tříd) daný růzností jedinců (neboli subjektů) podrobených měření, jednak vnejší rozptyl (RMT = rozptyl mezi třídami) daný růzností hodnot nezávislé veličiny při měření. V celé této kapitole předpokládáme, že rozptyl veličiny charakterizující každou podmínku měření je stejný, že jednotlivá měření jsou nezávislá a že měřené veličiny mají normální rozdělení pravděpodobnosti (vlastně stejné podmínky jako podmínky použitelnosti t-testu z 2.5.6, používaný F-test je tedy také parametrickým testem).

– Hypotéza H_0 u F-testu prohlašuje, že jakési dva rozptyly jsou stejné, kritériem testu je pak podíl odhadů těchto rozptylů. I když výpočet těchto odhadů z příkladu 3.1 je pedagogicky názorný a vystavěný na poznatcích z kapitoly 2.5, používáme spíše vzorce z tabulky 3.3, zapamatovatelné snadno pomocí oddílku 3.1.1.

– Podobně jako v kapitole 2.5, statistické testy jsou sice krásné, ale více informací než testy podávají intervaly spolehlivosti (vzorce pro ně vycházejí z předpokladu normálního rozdělení měřených veličin při neznámém rozptylu, a proto v intervalech spolehlivosti je použito kritických hodnot rozdělení t, nikoli rozdělení F). Zajímavé informace v situacích ANOVA lze získat z tzv. „plánovaného srovnání“, a možná i z dalších věcí, kterými se zabývá kapitola 5.

– V příkladu 3.6 je při zkoumání vztahu veličiny závislé na dvou nezávislých veličinách představen pojem **interakce**, což je veličina vyjadřující vliv vzájemné souhry obou sledovaných faktorů na závislou veličinu (jak lze rozeznat velikost interakce z grafického znázornění, je vidět v příkladech 3.8, 3.9). Tento pojem je klíčový v oddílu 3.1.3, kde v testech experimentu opakovaného měření rozptyl interakce vystupuje ve jmenovateli kritéria místo vnitřního rozptylu z testů oddílů 3.1, 3.1.2 (příklad 3.12 se zabývá zdůvodněním způsobu zpracování dat, při kterém se snažíme vyloučit rozptyl způsobený růzností jedinců neboli subjektů podrobených opakovanému měření, kdežto příklad 3.13 srovnává oddíly 3.1.2, 3.1.3 pomocí testů b),c), které jsou sice možné, ale v situaci opakovaného měření nás v podstatě nezajímají – klíčovým testem příkladu 3.13 je test a)). V příkladu 3.12 je také vysvětlen počet stupňů volnosti v itervalech spolehlivosti pro střední hodnotu veličin v podmínkách oddílu 3.1.3.

Kontrolní otázky

1. U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.
   a) Při F-testu je důležitá platnost předpokladu, že rozptyly hodnot měření v každé z podmínek jsou přibližně stejné.
   b) Odhad rozptylu RUT lze určit i jako aritmetický průměr odhadů rozptylů v každé z podmínek měření.
   c) Pokud intervaly spolehlivosti pro střední hodnoty při jednofaktorové ANOVA se navzájem významně překrývají, tak příslušný test ANOVA zamítne hypotézu H_0.

d) Pokud interakce faktorů 1, 2 při dvoufaktorové ANOVA má podstatný vliv na závislou proměnnou, příslušné grafické znázornění obsahuje křivky „po částech rovnoběžné“.

e) Vhodnou transformací proměnných lze někdy interakci odstranit.

f) Je možné, že pro stejná data test opakovaného měření pro tři podmínky měření zamítne hypotézu H_0, kdežto analogický test jednofaktorové ANOVA hypotézu H_0 nezamítne.

g) Při experimentu opakovaného měření je podstatnou slabinou fakt, že není možné vyloučit z testu rozptyl způsobený rozdílností mezi různými subjekty.

h) Intervaly spolehlivosti mají různou délku při různých počtech měření v každé z podmínek ANOVA.

i) Všechny testy této kapitoly lze za předpokladu normality veličin vyjádřit jako testy parametru σ^2 jistého typu, jedná se tedy o parametrické testy.

j) V některých situacích je možné navrhnout experiment typu „jeden vzorek vícekrát“ i typu „více vzorků jednou“, a přitom se dovíme z obou podobné informace.

Odpovědi na otázky
1a) – A, 1b) – A (za předpokladu učiněného v celé kapitole, že platí princip homogenního rozptylu), 1c) – N, 1d) – N, 1e) – A, 1f) – A, 1g) – N (je možné vyloučit vliv rozdílnosti subjektů – právě RI tento vliv vylučuje), 1h) – A, 1i) – A (v testu oddílu 3.1 lze H_0 vyjádřit jako $RUT = RMT$, v 3.1.2 jako $RUT = RMR$ (nebo při označení **??** jako $\sigma_R^2 = 0$), $RUT = RMS$ a $RUT = RI$ (nebo při označení **??** jako $\sigma_I^2 = 0$), v 3.1.3 jako $RI = RMS$ (nebo při označení **??** jako $\sigma_S^2 = 0$)), 1j) – A (viz příklad 3.11).

Cvičení

1. Provádíme experiment, který má prokázat vliv věku na schopnost zapamatování. Náhodně bylo vybráno pět lidí ve věku deseti let, pět lidí ve věku 21 let a pět lidí ve věku 75 let. U každého z vybraných bylo zaznamenáno, kolik slov z uvedených dvaceti si po určité době pamatoval. Získala se data:
10iletí: 1,4,5,6,4,
21letí: 9,8,7,10,6,
75iletí: 3,5,7,7,8.

   a) Proveďte test jednofaktorové ANOVA o rovnosti středních hodnot veličin v jednotlivých skupinách (na hladině významnosti $\alpha = 0{,}05$);

   b) určete intervaly se spolehlivostí 95% pro tyto střední hodnoty.

2. Každá z následujících dvou situací obsahuje čísla od 0 do 10 náhodně vygenerovaná počítačem:
Situace 01: 9,6,4,2,6,6,0;
Situace 02: 3,4,7,4,3,0,3.

   a) Ověřte F-testem, zda se rozptyly RUT, RMT v této situaci významně liší.

   b) Ověřte F-testem, zda odhady $est_1\sigma^2$, $est_2\sigma^2$ v jednotlivých skupinách jsou odhady téhož rozptylu.

3. Honzu Kováře zajímá, zda různé značky ústní vody vedou k rozdílům ve svěžesti dechu. Lidem v první skupině dá ústní vodu značky MOPE, druhé ústní vodu LIST a třetí LOVE. Poté změří svěžest dechu dechoměrem (vyšší hodnoty = svěžejší dech). Je známo, že populační rozptyl hodnot dechoměru je $\sigma^2 = 2$ a že hodnoty dechoměru lze popsat normálním rozdělením. Byla získána data:
MOPE: 2,4,4,2;
LIST: 7,10;
LOVE: 5,6,4.

   a) Vypočtěte 95%-ní intervaly spolehlivosti pro střední hodnoty jednotlivých značek;

   b) Proveďte nad daty příslušný test ANOVA – využijte přitom, že populační rozptyl σ^2 je známý.

   c) Pro $\alpha = 0{,}10$ testem ověřte, zda *est RUT* je dobrým odhadem rozptylu σ^2.

   d) Předpokládejte, že σ^2 neznáme, a proveďte standardní ANOVA.

4. Psycholožku Strnadovou zajímá vliv socioekonomické situace a motivace na paměť. Závislá proměnná se měří počtem zapamatovaných slov ze dvaceti uvedených (po jisté stanovené době na naučení slov). Socioekonomický stav (= SES) má dvě úrovně (vysoký, nízký), motivace má dvě úrovně (vysoká = 50 kč za každé zapamatované slovo, nízká = 1 kč za každé zapamatované slovo). Proveďte ANOVA typu 2 × 2 (testy vlivu faktoru 1, vlivu faktoru 2 a vlivu jejich interakce) pro deset náhodně vybraných lidí v každé ze čtyř podmínek experimentu – bylo spočteno $SSUT = 3600$, další informace jsou v tabulce:

	vysoký SES	nízký SES
nízká motivace	$T_{11} = 20$, $\overline{x}_{11} = 2$	$T_{12} = 80$, $\overline{x}_{12} = 8$
vysoká motivace	$T_{21} = 100$, $\overline{x}_{21} = 10$	$T_{22} = 120$, $\overline{x}_{22} = 12$

5. Experimentem zjišťujeme, zda má marihuana vliv na odhad času. Jedné skupině je rozdána v tabletách marihuana, druhé skupině placebo tablety (= bez marihuany). Navíc v obou z dvanáctičlenných skupin je polovina mužů a polovina žen, tj. jedná se o ANOVA typu 2 × 2 (faktor 1: muži, ženy; faktor 2: tablety marihuany, tablety placebo). Poté se všichni zúčastnění podrobí zkoušce odhadu času, kdy sedí a mluví do mikrofonu tak dlouho, až mají pocit, že uběhlo deset minut. Je změřena skutečná délka doby mluvení do mikrofonu (v minutách). Získala se data:

	marihuana	placebo
muži	5,3,4,6,6,7	11,9,8,13,11,10
ženy	8,7,9,6,8,9	11,14,9,13,12,10

   a) Vypočtěte 95%-ní interval spolehlivosti pro střední hodnotu v každé z uvedených čtyř situací;

   b) testujte vliv pohlaví, vliv drogy a vliv jejich interakce na odhad času.

6. Experiment má prokázat vliv verbálního popisu na lidské vnímání. Spočívá v tom, že když klient „čeká v předpokoji na experiment“, jistý experimentátorův kolega popíše experimentátora slovním spojením přídavného jména s příslovcem, přičemž příslovce vybírá z trojice (docela, velmi, neobyčejně), přídavné jméno z trojice (hodný, schopný, nepříjemný). Poté je čekatel uveden k experimentátorovi, ten jej podrobí falešnému experimentu učení se slov. Poté

má testovaný člověk vyplnit dotazník, ve kterém se mimo jiné vyskytuje otázka: Jak se vám líbil experimentátor? Odpovědi se uvádějí v rozsahu 0 (= nemohl jsem ho vystát) až 10 (= zamiloval jsem si ho). Získala se tato data:

	hodný	schopný	nepříjemný
docela	5,5,6,5,4	5,5,6,4,6	4,3,3,6,5
velmi	6,6,9,5,7	5,6,6,7,4	3,1,3,2,2
neobyčejně	9,7,8,8,9	4,7,5,8,4	1,3,2,1,1

a) Vypočtěte 95%-ní interval spolehlivosti pro každou z uvedených situací;

b) existuje významný vliv příslovce, přídavného jména nebo jejich interakce?

7. Psychologa zajímá, jak velkou trému člověk prožívá vzhledem k počtu lidí, kteří jej pozorují. Čtyři lidé jsou požádáni, aby si představili, že musí recitovat báseň před jedním, pěti nebo patnácti lidmi, a pak ať ohodnotí svou trému stupnicí od 0 (= nemám vůbec trému) po 7 (= jsem vyděšen k smrti). Získala se data

	1 člověk v publiku	5 lidí v publiku	15 lidí v publiku
subjekt 1	3	6	5
subjekt 2	1	5	6
subjekt 3	3	6	7
subjekt 4	1	3	5

a) Vypočtěte 95%-ní interval spolehlivosti pro střední hodnotu v každé ze tří podmínek;

b) ověřte testem, že velikost publika má vliv na trému.

8. Je vyvíjena nová metodika likvidace moskytů, která spočívá v tom, že jsou vypouštění sterilní samečci, kteří se sice stýkají se samičkami, ale nemohou mít potomky. Při testování nové metody jsou čtyři národní parky rozděleny vždy na tři části – v první části se nic neděje, ve druhé části je proveden postřik DDT, a ve třetí části jsou vypuštěni sterilní samečci. Získala se následující data (v počtu moskytů na metr čtvereční na sledovaných plochách):

národní park	nic se neděje	DDT	nová metoda
subjekt 1	614	512	123
subjekt 2	320	300	250
subjekt 3	502	500	313
subjekt 4	750	600	430

a) Vypočtěte 95%-ní interval spolehlivosti pro střední hodnotu v každé z daných tří podmínek (= metod);

b) existují významné rozdíly mezi těmito třemi podmínkami? Ověřte testem statistickým.

Je vyvíjena nová metodika likvidace moskytů, která spočívá v tom, že jsou vypouštění sterilní samečci, kteří se sice stýkají se samičkami, ale nemohou mít potomky. Při testování nové metody jsou čtyři národní parky rozděleny vždy na tři části – v první části se nic neděje, ve druhé části je proveden postřik DDT, a ve třetí části jsou vypuštěni sterilní samečci. Získala se následující data (v počtu moskytů na metr čtvereční na sledovaných plochách):

národní park	nic se neděje	DDT	nová metoda
subjekt 1	614	512	123
subjekt 2	320	300	250
subjekt 3	502	500	313
subjekt 4	750	600	430

a) Vypočtěte 95%-ní interval spolehlivosti pro střední hodnotu v každé z daných tří podmínek (= metod);

b) existují významné rozdíly mezi těmito třemi podmínkami? Ověřte testem statistickým.

9. Lékařka má hypotézu, že pravděpodobnost nachlazení se liší v jednotlivých ročních obdobích. Proto u pěti pacientů zaznamenává počet nachlazení v průběhu tří let a výsledky shromáždí do tabulky (v počtech nachlazení – tři čísla v každé buňce představují data ve třech letech po sobě):

subjekt	zima	jaro	léto	podzim
Josef	3,1,2	1,0,1	3,3,2	0,0,0
František	1,1,2	0,1,1	3,4,3	1,1,2
Cyril	1,1,1	0,3,0	2,1,2	1,0,0
Jan	3,4,3	2,2,1	5,2,2	1,2,3
Zdeněk	2,2,3	1,2,1	2,1,0	1,0,1

a) Vypočtěte 95%-ní interval spolehlivosti pro střední hodnotu v každém ročním období;

b) Proveďte ANOVA pro tato data. Existuje významný rozdíl mezi obdobími? Dále zjistěte, zda je významný vliv subjektů a vliv interakce subjekt-období.

Výsledky

ad 1. **ad a)** $H_0 : \mu_1 = \mu_2 = \mu_3$, H_1 : neplatí H_0. *est RUT* $= 3{,}33$, *est RMT* $= 20$, tj.

$$\frac{est\ RMT}{est\ RUT} = 6 > F_k(VMT = 2, VUT = 12) = 3{,}9;$$

a tedy zamítáme H_0, jednotlivé střední hodnoty podmínek se významně liší.

ad b) Příslušný průměr $\pm 1{,}778$.

ad 2. **ad a)** *est RUT* $= 6{,}74$, *est RMT* $= 7{,}14$, tj.

$$\frac{est\ RMT}{est\ RUT} = 1{,}06 < F_k(VMT = 1, VUT = 12, \alpha = 0{,}05) = 4{,}75;$$

tedy není zjištěn významný rozdíl. Je to proto, že podmínky „měření“ v obou situacích jsou vlastně stejné. Tato část ukazuje, že F-test lze provést i v situaci dvou skupin měření.

ad b) $\overline{s_1^2} = 8{,}905$, $\overline{s_2^2} = 4{,}286$, pak test rovnosti obou rozptylů podle **důležité poznámky** za tabulkou 3.2 se děje podle $H_0 : \sigma_1^2 = \sigma_2^2$, kritéria $\frac{\overline{s_1^2}}{\overline{s_2^2}}$ na hladině významnosti $\alpha = 0{,}02$

pro kritické hodnoty $F_v(\frac{\alpha}{2} = 0{,}01, V_1 = 6, V_2 = 6) = 8{,}47$ podle tabulky 3.2 a hodnoty $F_m(\frac{\alpha}{2} = 0{,}01, V_1 = 6, V_2 = 6) = \frac{1}{8{,}47} = 0{,}118$ podle vzorce 3.3, tj.

$$\frac{\overline{s_1^2}}{\overline{s_2^2}} = \frac{8{,}905}{4{,}286} = 2{,}0777 \in (0{,}118; 8{,}47),$$

a tedy H_0 o rovnosti rozptylů nezamítáme. Opět důvodem jsou stejné podmínky měření v obou skupinách. Tato část příkladu ukazuje, že F-test lze užít při testování hypotézy, že máme k dispozici dva odhady téhož rozptylu (tj. že rozptyly v obou populačních měřeních jsou stejné).

ad 3. **ad a)** MOPE: $3 \pm 1{,}386$; LIST: $8{,}5 \pm 1{,}96$; LOVE: $5 \pm 1{,}6$.

ad b) Hodnota kritéria je

$$\frac{est\ RMT}{\sigma^2} \doteq 10{,}1 > F_k(VMT = 2, V_{\sigma^2} = \infty, \alpha = 0{,}05) = 3{,}0,$$

tj. zamítáme H_0 o rovnosti středních hodnot jednotlivých podmínek. Mezi značkami je významný rozdíl.

ad c) $est\ RUT = 1{,}75$, tj.

$$\frac{\sigma^2}{est\ RUT} = \frac{2}{1{,}75} \doteq 1{,}143 < F_v(V_{\sigma^2} = \infty, VUT = 6, \frac{\alpha}{2} = 0{,}05) = 3{,}67,$$

a protože $F_m < 1$, vidíme, že H_0 není zamítnuto, tj. $est\ RUT$ je dobrým odhadem $\sigma^2 = 2$.

ad d)

$$\frac{est\ RMT}{est\ RUT} = \frac{20{,}2}{1{,}75} doteq 11{,}54 > F_k(VMT = 2, VUT = 6) = 5{,}14,$$

a proto zamítáme H_0 a uzavíráme, že mezi středními hodnotami měření na dechoměru jednotlivých značek existují významné rozdíly.

ad 4. $n = 40$, $N_{ij} = 10$, $VUT = 36$, tj. $est\ RUT = 100$. Dále $VMR = 1$, $VMS = 1$, $VI = 1$, tedy $est\ RMR = 360$, $est\ RMS = 160$, a pak $SSI = 40$, tj. $VI = 40$. Statistické testy:

i) Vliv faktoru 1:

$$\frac{est\ RMR}{est\ RUT} = 3{,}6 < F_k(1; 36) \doteq 4{,}1,$$

a tedy H_0 nezamítáme, vliv není významný.

ii) Vliv faktoru 2:

$$\frac{est\ RMS}{est\ RUT} = 1{,}6 < F_k(1; 36) \doteq 4{,}1,$$

a tedy H_0 nezamítáme, vliv není významný.

iii) Vliv interakce faktorů:

$$\frac{est\ RI}{est\ RUT} = 0{,}4 < F_k(1; 36) \doteq 4{,}1,$$

a tedy H_0 nezamítáme, vliv není statisticky významný.

ad 5. ad a) intervaly spolehlivosti: odpovídající průměr $\pm 1{,}35$.
ad b) Testy na hladině významnosti $\alpha = 0{,}05$; vliv pohlaví: hodnota kritéria $8{,}73 > F_k(VMR = 1, VUT = 20) = 4{,}35$; vliv drogy: hodnota kritéria $46{,}35 > F_k(VMS = 1, VUT = 20) = 4{,}75$; vliv interakce: hodnota kritéria $3{,}37 < 4{,}35$, tj. pouze vliv interakce není statisticky významný.

ad 6. ad a) průměr $\pm 1{,}05$; ad b) Testy na hladině významnosti $\alpha = 0{,}05$: vliv příslovce: hodnota kritéria $0{,}42 < F_k(VMR = 2, VUT = 36) \doteq 3{,}3$; vliv přídavného jména: hodnota kritéria $46{,}2 > F_k(VMS = 2, VUT = 36) \doteq 3{,}3$; vliv interakce: hodnota kritéria $8{,}17 > F_k(VI = 4, VUT = 36) \doteq 2{,}66$ (odhadnuto pomocí $F(4, 30)$ a $F(4, 40)$), čili nevýznamný je pouze vliv příslovce.

ad 7. $VMT = 11$, $SSMT = 44{,}25$, $VMS = 2$, $SSMS = 31{,}5$, $VMR = 3$, $SSMR = 8{,}91$, odtud $VI = 6$, $SSI = SSMT - SSMS - SSMR = 3{,}8333$. Nyní můžeme dosadit do vzorců:

a) $\mu_i \in \overline{x}_i \pm 2{,}447 \cdot \sqrt{\frac{3{,}8333}{6 \cdot 4}} = \overline{x}_i \pm 0{,}9779$.

b) H_0 : podmínky se neliší; H_1 : podmínky se liší;

$$\frac{est\ RMS}{est\ RI} = 24{,}67 > F_k(2; 6) = 5{,}14,$$

H_0 zamítáme, rozdíl mezi podmínkami testu opakovaného měření je významný.

ad 8. ad a) průměr $\pm 119{,}5$; ad b) hodnota kritéria $8{,}097 > F_k(2; 6) = 5{,}14$, tj. rozdíly mezi podmínkami (metodami) jsou významné.

ad 9. ad a) průměr $\pm 1{,}11$; ad b) Na hladině významnosti $\alpha = 0{,}05$: Vliv období: hodnota kritéria $5{,}83 > F_k(3; 12) = 3{,}49$, tj. vliv je významný; vliv subjektů: hodnota kritéria $5{,}9 > F_k(4; 40) = 2{,}61$, tj. vliv je významný; vliv interakce: hodnota kritéria $1{,}95 < F_k(12; 40) = 2{,}00$, tj. vliv není významný.

4 Korelační přístup, regresní přímka

Průvodce studiem

V této kapitole se budeme zabývat vztahem mezi dvěma proměnnými. Pro jeho popis budeme používat lineární funkci = přímku. Budeme studovat způsob konstrukce optimální přímky, která nejpřesněji popisuje vztah mezi studovanými veličinami. Naučíte se chápat principy korelace.

Cíle

Po prostudování této kapitoly budete schopni:

- Zkonstruovat regresní přímku.
- Chápat spojitost regresní přímky s metodou nejmenších čtverců.
- Poznat, kdy je popis pomocí lineární funkce vhodný a kdy ne.
- Naučite se tuto vhodnost nebo nevhodnost měřít.

4.1 Predikce

Příklad 4.1. Uvažujme experiment, který už v tomto textu byl popsán (viz příklad 2.16), kdy zkoumáme vliv konzumace alkoholu na dobu reakce člověka. Rozdělme 50 náhodně vybraných osob do pěti skupin po deseti. Každé skupině je „vyrobeno“ jiné množství alkoholu v krvi (0%, 0,01%, 0,02%, 0,03%, 0,04%), a pak je měřena doba reakce mezi rozsvícením světla a stisknutím tlačítka. Jsou získána následující data (reakční doba v milisekundách):

0%	0,01%	0,02%	0,03%	0,04%
192	205	208	231	235
194	198	209	228	230
189	201	220	216	233
178	208	216	220	228
193	216	221	225	237
201	203	210	226	230
199	207	215	220	241
198	200	217	218	242
196	198	208	223	225
190	205	210	229	230
$\overline{x_1} = 193{,}0$	$\overline{x_2} = 204{,}1$	$\overline{x_3} = 213{,}4$	$\overline{x_4} = 223{,}6$	$\overline{x_5} = 233{,}1$

Nyní lze spočítat $estRUT = 30{,}81$, a tedy $\sqrt{estRUT} = 5{,}55$. Dále lze ze všech padesáti měření určit celkový průměr $\overline{x} = 213{,}44$ a odhad celkového rozptylu $estR^2 = 231{,}31$, tj $\sqrt{estR^2} = 15{,}21$.

Zajímá nás teď náhodně vybraný člověk, například Robert, – chceme odhadnout jeho reakční dobu, víme-li, že procento alkoholu v krvi u něj je asi 0,015. K dispozici máme odhady pro lidi s 0%, 0,01%, 0,02%, 0,03% a 0,04% (za daný odhad bereme průměr hodnot reakční doby v dané skupině). Na základě jistých měření tedy známe vztah mezi dvěma proměnnými (v našem případě procento alkoholu v krvi a doba reakce), pak pokud je zadaná hodnota jedné proměnné (té říkáme v tomto textu nezávislá proměnná – v našem případě je to procento alkoholu 0,015), jsme schopni jistým způsobem odhadnout neboli **predikovat** příslušný stav druhé proměnné (které v tomto textu říkáme závislá proměnná – v našem případě je to doba reakce).

V dalším textu (4.3) se budeme zabývat hledáním tzv. **regresní přímky**, což je funkce tvaru $y = a + bx$, do které když za dosadíme $x = 0{,}015$, vypočteme pomocí ní hodnotu y odhadované neboli predikované doby reakce.

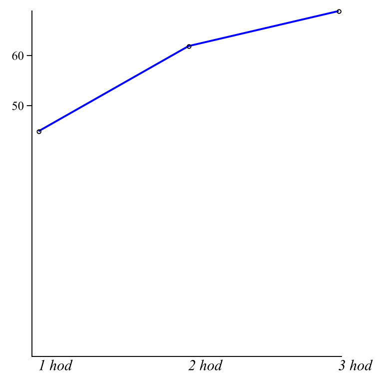
4.2 Korelace

Až dosud jsem při studiu vztahu mezi dvěma proměnnými používali tzv. **experimentální přístup** (= experimentátor ovlivní hodnoty jedné proměnné, a pak změří příslušné hodnoty druhé proměnné). V tomto oddílku a vlastně v celé kapitole se musíme seznámit s tzv. **korelačním přístupem** (= žádná z proměnných není ovlivňována, pouze jsou hodnoty obou proměnných změřeny). Rozdíl mezi experimentálním přístupem a korelačním přístupem ukazuje následující příklad.

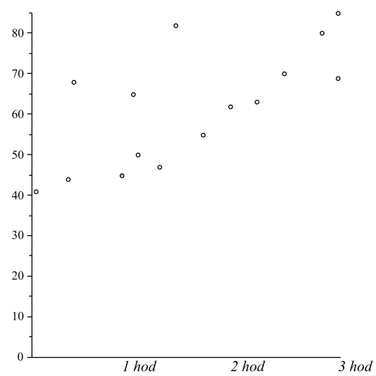
Příklad 4.2. Chceme zjistit, jaký je vztah mezi počtem hodin domácí přípravy za týden a bodovým průměrem zkoušky z předmětu MPSO na konci semestru.

a) Experimentální přístup: Náhodně rozdělíme studenty do tří skupin. Studenty z první skupiny požádáme, aby věnovali domácí přípravě jednu hodinu týdně; studenty ze druhé skupiny požádáme, aby doma studovali dvě hodiny, a studenti ze třetí skupiny tři hodiny týdně.

Koncem semestru se vypočtou průměry výsledků v každé ze tří daných skupin a získáme následující grafické znázornění experimentu:



b) Korelační přístup: Na konci semestru po absolvování zkoušky položíme každému studentovi dvě otázky: Kolik hodin týdně strávil domácí přípravou a jakého výsledku dosáhl nakonec u zkoušky. Studenti odpovídají podle pravdy. Odpovědi patnácti studentů jsou získány (Josef Miklíček studoval 1,5 hodin týdně a dosáhl 82 bodů, Vilém Mrštík jen půl hodiny týdně a získal 68 bodů, atd.) a graficky znázorněny:



Z předchozího příkladu je snad patrný rozdíl mezi experimentálním a korelačním přístupem: na základě experimentálního přístupu uzavíráme, že delší doba přípravy zlepšuje výsledek zkoušky. Na základě korelačního přístupu existuje ovšem více možných vysvětlení:

- **Je možné, že** delší doba přípravy zlepšuje výsledek zkoušky.
- **Je možné, že** je tomu i naopak: Výsledek zkoušky ovlivňuje délku přípravy.
- **Je možné, že** mezi oběma proměnnými neexistuje přímá závislost – možná existuje jakýsi **třetí faktor**, který je zdrojem vztahu mezi oběma proměnnými (např.

motivace – vysoce motivovaní studenti mají dobré výsledky, a současně věnují delší dobu přípravě, kdežto nedostatečně motivovaní studenti nemají chuť se připravovat, ani dělat cokoliv jiného, co by zlepšilo jejich výsledky).

Důležitý rozdíl mezi korelačním a experimentálním přístupem je právě zejména v případě poslední uvedené odrážky – při korelačním přístupu nemusí existovat mezi oběma proměnnými přímá závislost, ale „vliv" jedné proměnné na druhou může být způsoben třetím faktorem. Například:

Příklad 4.3. Konzumace zmrzliny versus úmrtnost: Existuje korelace mezi množstvím zmrzliny zkonzumované v daný den v Brně a úmrtností daný den v Praze. V ty dny, kdy se v Brně sní hodně zmrzliny, je v Praze větší úmrtnost než obyčejně (= ve srovnání se dny, kdy se v Brně sní méně zmrzliny). Znamená to, že musím cítit vinu, když si dám v Brně zmrzlinu, že působím smrt nějakého člověka v Praze? Ne, mezi oběma proměnnými není přímá závislost – spíše zde existuje třetí faktor, a tím je teplota. Díky vyšší teplotě se v Brně zkonzumuje více zmrzliny a v Praze zemře více lidí.

Příklad 4.4. Prostitutky versus ministři: existuje korelace mezi počtem ministrů a počtem prostitutek v daném městě, tj. města s větším počtem prostitutek obvykle mají větší počet ministrů. Opět to neznamená, že nutně musí existovat přímá závislost mezi těmito dvěma proměnnými. Spíše zde hraje roli třetí faktor, kterým je velikost města. Velká města obvykle mají větší počet ministrů i prostitutek.

Korelačního přístupu se hojně užívá v těch oblastech, kde nejsme schopni regulovat hodnotu proměnných. Např. u příkladu 4.2 je korelační přístup mnohem lepší, protože není etické přímým nařízením ovlivňovat dobu přípravy studentů na zkoušku.

Podobně kdyby nás zajímal vztah mezi počtem krádeží vloupáním a počtem prodaných dveřních zámků v daném městě, nebylo by vhodné ovlivňovat ani počet krádeží, ani počet prodaných zámků – nejlepším přístupem je studovat korelaci mezi oběma proměnnými.

Podobně při studiu vztahu teploty během dne a počtem infarktů v ten den nejsme schopni ovlivnit hodnoty žádné z obou proměnných (ve společenských vědách je experimentální přístup spíše výjimkou, než pravidlem – zelenou má ovšem korelační přístup).

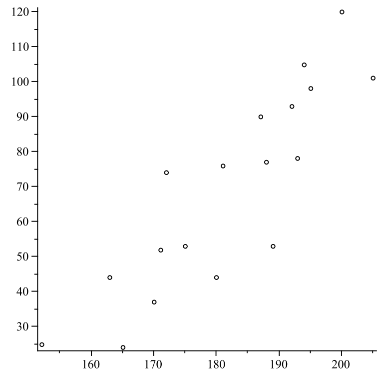
4.3 Regresní přímka

Nyní když bylo řečeno, proč je korelační přístup v mnoha případech důležitý, či dokonce jediný možný přístup, zabývejme se chvíli matematickým popisem této korelace (= vzájemného vztahu) mezi veličinami. Dvě veličiny jsou tedy korelovány (= v korelaci), když na základě hodnoty jedné z nich lze predikovat (= předpovědět, odhadnout) hodnotu té druhé. Podívejme se na tzv. lineární (= přímkový) korelační vztah mezi veličinami x, y: Pokud známe hodnotu proměnné x, tak hodnotu proměnné y můžeme predikovat na základě lineárního vztahu

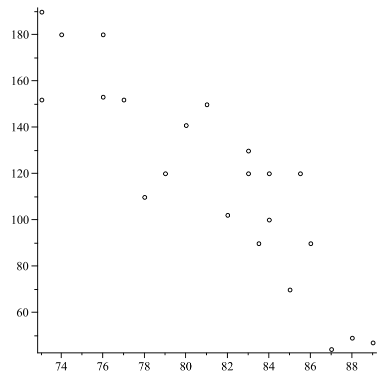
$$y = a + bx,$$

kde a, b jsou reálné konstanty (a grafem funkce $f(x) = a + bx$ je přímka, které říkáme **regresní přímka**).

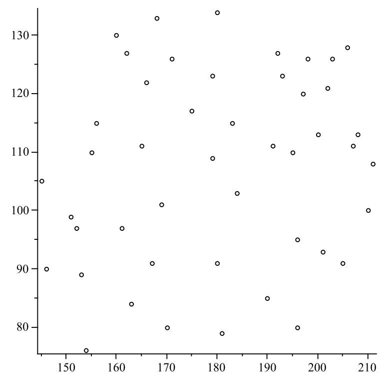
Příklad 4.5. **a)** Příkladem proměnných, které jsou kladně korelovány (= větší hodnotě proměnné x odpovídá větší hodnota proměnné y), je hmotnost (v kg) a výška člověka (v cm):



b) Příkladem proměnných, které jsou záporně korelovány (= větší hodnotě proměnné x odpovídá menší hodnota proměnné y), je průměr golfového skóre (hodnoty 75, 80, 85 atd. na vodorovné ose) a příjem hráče (v tisících dolarů):



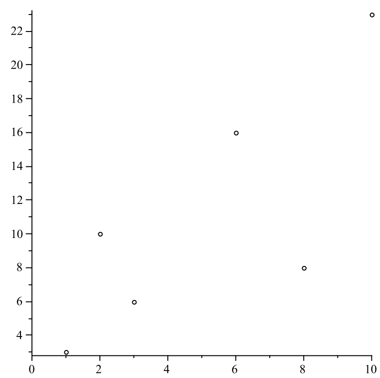
c) Příkladem proměnných, které jsou nekorelované (= nejsou v korelaci), je výška člověka (v cm) a IQ člověka (v íkváččích :-)):



Příklad 4.6. Chceme zjistit, zda existuje korelace mezi počtem piv, které člověk vypije za týden, a počtem fotbalových zápasů, které shlédne za danou sezónu. Z toho důvodu je učiněn malý průzkum: šesti mužům je položena otázka, kolik piva vypijí průměrně za týden a kolik fotbalových zápasů za minulou sezónu sledovali. Získaná data

osoba	1	2	3	4	5	6
počet piv x	3	10	6	1	2	8
počet zápasů y	6	23	16	3	10	8

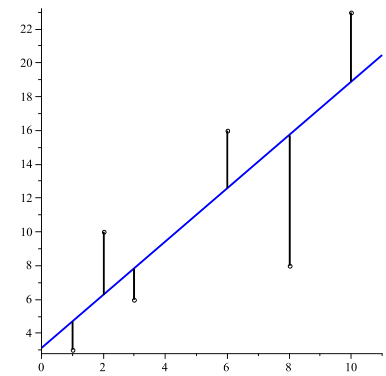
lze graficky znázornit:



Z hodnot naměřených a znázorněných na obrázku příkladu 4.6 chceme získat tzv. **regresní přímku** = nejvhodnější přímku, která zachycuje lineární vztah mezi uvedenými dvěma proměnnými. Otázkou je, v jakém smyslu (vzhledem k jakému kritériu) má být uvedená přímka nejvhodnější či nejlepší, a jak ji najít (= jak určit koeficienty a, b lineárního vztahu $y = a + bx$). Označme proto zadané body $[x_i, y_i]$ pro $i = 1, 2, \ldots, 6$ a $[x_i, y'_i]$ mající stejnou souřadnici x, ale ležící na hledané přímce (platí tedy $y'_i = a + bx_i$).

a) Kritérium $\sum(y_i - y'_i)$: Prvním možným kandidátem pro kritérium vhodnosti regresní přímky se nabízí součet odchylek naměřené hodnoty y_i a predikované hodnoty y'_i – ale tento přístup má nevýhodu, že některé ze členů $(y_i - y'_i)$ jsou kladné, jiné záporné. A protože záporné odchylky vyruší ty kladné odchylky, součet $\sum(y_i - y'_i)$ není dobrým měřítkem celkové chyby pro danou přímku (a navíc – viz [7], str. 442 – tímto kritériem není hledaná přímka určena jednoznačně).

b) Kritérium $\sum |y_i - y'_i|$: Součet absolutních hodnot odchylek je už lepším kritériem vhodnosti přímky – jedná se vlastně o součet vzdáleností naměřeného bodu a predikovaného bodu pro každé x_i, viz obr.:



(délky jednotlivých svislých úseček jsou právě rovny hodnotám $|y_i - y_i'|$)

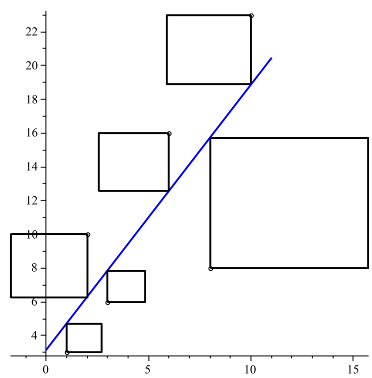
Ale protože výraz obsahující absolutní hodnoty se obtížně matematicky zpracovává, a navíc hledaná přímka opět nemusí být určena jednoznačně (příklad viz [7], str.442), nejvíce se ujalo následující kritérium:

c) Kritérium $\sum(y_i - y_i')^2$: Hledejme takové koeficienty a, b lineárního vztahu $y = a+bx$, aby součet čtverců odchylek $S = \sum_1^6 (y_i - y_i')^2$ byl minimální možný. Každý čtenář těchto slov by si měl nyní uvědomit, že tohle přece už zná – vždyť se jedná o metodu nejmenších čtereců (least squares method) probranou v předmětu BMA3, kapitola 6, oddíl 6.3. Ano, vážení přátelé, regresní přímka není nic jiného než přímka získaná při aproximaci zadaných bodů v rovině metodou nejmenších čtverců. A vlastně stejné kritérium – součet čtverců – bylo používáno v kapitole 3 při analýze rozptylu.

Jen pro úplnost rychle zopakujme, o co se jedná: Chceme najít koeficienty a, b tak, abychom minimalizovali funkci

$$S = \sum_1^n (y_i - (a + bx_i))^2,$$

kde n je počet zadaných bodů $[x_i, y_i]$. Název **metoda nejmenších čtverců** se vžil díky grafickému názoru, že totiž opravdu hledáme takovou přímku, pro kterou je součet ploch čterců odchylek minimální – viz obr.:



(uvedené obrazce jsou skutečně čtverce – o obdélníky se jedná jen zdánlivě, díky různým měřítkům velikosti jednotky na obou osách). Z diferenciálního počtu je známo, že extrém funkce S nastane v tzv. stacionárních bodech, pro které platí $S'_a = 0$, $S'_b = 0$ (na levých stranách rovnic jsou derivace funkce S podle neznámých proměnných a, b), tj. dostáváme:

$$\begin{aligned} \sum 2(y_i - a - bx_i)(-1) &= 0; \\ \sum 2(y_i - a - bx_i)(-x_i) &= 0. \end{aligned}$$

Úpravou dostaneme systém rovnic ve tvaru

$$\begin{aligned} na + b\sum x_i &= \sum y_i; \\ a\sum x_i + b\sum x_i^2 &= \sum (x_i y_i). \end{aligned}$$

Jedná se o systém dvou rovnic o dvou neznámých a, b. Řešíme jej například dosazovací metodou: z první rovnice vyjádříme

$$a = \frac{1}{n}\sum y_i - \frac{b}{n}\sum x_i; \tag{4.1}$$

a dosadíme do druhé rovnice, odkud spočteme

$$b = \frac{n\sum(x_i y_i) - \left(\sum x_i\right)\left(\sum y_i\right)}{n\sum x_i^2 - \left(\sum x_i\right)^2}. \tag{4.2}$$

Při konkrétním výpočtu tedy nejprve určíme b podle 4.2, a pak dosadíme do 4.1 a vypočteme a.

V našem příkladu $b = \frac{606}{384} \doteq 1{,}578$, *a tedy* $a \doteq 3{,}11$, *čili regresní přímka je tvaru*

$$y = 3{,}11 + 1{,}578x.$$

Označíme-li nyní průměr x-ových souřadnic $\overline{x}_x$ ($\overline{x}_x = \frac{1}{n}\sum x_i$) a průměr y-ových souřadnic $\overline{x}_y$ ($\overline{x}_y = \frac{1}{n}\sum y_i$), pak bod $[\overline{x}_x, \overline{x}_y]$ vždy leží na regresní přímce. Skutečně, vzorec 4.1 lze přepsat

$$a = \overline{x}_y - b \cdot \overline{x}_x, \text{ čili } \overline{x}_y = a + b \cdot \overline{x}_x,$$

tj. bod $[\overline{x}_x, \overline{x}_y]$ splňuje rovnici regresní přímky.

Na regresní přímku lze pohlížet jako na nejlepší způsob **lineární predikce** hodnot proměnné Y pomocí hodnot proměnné X. Pokud by například (viz náš příklad 4.6) bylo známo, že jistý člověk vypil za sezónu průměrně devět piv týdně, dosazením za $x = 9$ do rovnice regresní přímky dostaneme, že nejlepší lineární odhad počtu sledovaných fotbalových zápasů za sezónu je u něj

$$y = 3{,}11 + 1{,}578 \cdot 9 = 17{,}312 \doteq 17 \text{ zápasů}.$$

4.4 Korelační koeficient

Jak vhodná je predikce s využitím regresní přímky? Jedním z kritérií posouzení vhodnosti je tzv. **Pearsonův koeficient** $r_2 \in \langle 0; 1\rangle$, kde

$$r^2 = \frac{[n\sum(x_iy_i) - (\sum x_i)(\sum y_i)]^2}{[n\sum x_i^2 - (\sum x_i)^2] \cdot [n\sum y_i^2 - (\sum y_i)^2]}. \tag{4.3}$$

Vzorec má odpudivý tvar, ale existuje v něm jistá symetrie, která usnadňuje jeho zapamatování. Posuďte sami: V čitateli i jmenovateli zlomku je součin dvou závorek (respektive v čitateli je to tatáž závorka umocněna na druhou). V každé závorce je n-násobek součtu součinů MINUS součin součtů, přičemž ta čísla v součinech, která se sčítají, jsou zvlášť sečtena v sumách, které se mezi sebou násobí. V čitateli se vždy jedná o součiny $x_i \cdot y_i$, ve jmenovateli jsou v jedné závorce součiny $x_i \cdot x_i$, ve druhé $y_i \cdot y_i$. Vidíte tu symetrii?

V čitateli zlomku pro výpočet r^2 je druhá mocnina n^2-násobku tzv. kovarinace mezi proměnnými X, Y, což je míra vztahu korelace mezi proměnnými X, Y. Jedná se o číslo z intervalu $\langle -1; 1\rangle$, které je kladné při klané korelovanosti, záporné při záporné korelovanosti a blízké nule při nekorelovanosti obou veličin.

Jmenovatel zlomku pro výpočet r^2 je jakýmsi normalizačním faktorem, který upravuje měřítko pro popis hodnot veličin X, Y. Právě díky tomuto normalizačnímu faktoru se koeficient r^2 nachází v intervalu $\langle 0; 1\rangle$.

Nyní se ještě přesněji vraťme k pojmu **kovarinace**: vydělíme-li čitatele i jmenovatele zlomku 4.3 konstantou n^4, dostaneme

$$r^2 = \frac{[\frac{1}{n}\sum(x_iy_i) - \frac{1}{n^2}(\sum x_i)(\sum y_i)]^2}{[\frac{1}{n}\sum x_i^2 - \frac{1}{n^2}(\sum x_i)^2] \cdot [\frac{1}{n}\sum y_i^2 - \frac{1}{n^2}(\sum y_i)^2]}, \tag{4.4}$$

kde v čitateli zlomku je právě druhá mocnina kovariance veličin X, Y – označme ji s_{XY}^2. A ve jmenovateli posledního uvedeného zlomku je součin rozptylů jednotlivých souborů měření obou proměnných s_X^2, s_Y^2. Čili platí

$$r^2 = \frac{\left[s_{XY}^2\right]^2}{s_X^2 \cdot s_Y^2}. \tag{4.5}$$

Pokud koeficient r^2 je blízký jedné, daná regresní přímka je vhodným modelem korelace. Pokud r^2 je blízký nule, vhodnost lineární regrese je malá. Na našem příkladu 4.6 si ukážeme ještě další význam koeficientu r^2 (respektive jiný pohled na tento koeficient).

Příklad 4.7. V příkladu 4.6 můžeme spočítat $r^2 = 0{,}59$. Vypočtěme dále celkový rozptyl hodnot y_i (počty fotbalových zápasů):

$$s_Y^2 = \frac{1}{6}\sum y_i^2 - \left(\frac{\sum y_i}{6}\right)^2 = 44{,}7.$$

Připomeňme si, že číslo s_Y^2 vyjadřuje rozptyl hodnot y_i samotných, nikoli odhad skutečného roptylu. Teď uvažujme hodnoty y_i' získané pomocí lineární regrese, tj. hodnoty $y_i' = 3{,}11 + 1{,}578 \cdot x_i$, a vypočtěme rozptyl souboru těchto predikovaných hodnot:

$$s_{Y'}^2 = \frac{1}{6}\sum y_i'^2 - \left(\frac{1}{6}\sum y_i'\right)^2 = 26{,}6.$$

Je zajímavé si všimnout, že platí

$$\frac{s_{Y'}^2}{s_Y^2} = \frac{26{,}6}{44{,}7} = 0{,}59 = r^2,$$

čili **koeficient r^2 vyjadřuje, kolik procent rozptylu hodnot y_i je obsaženo v modelu lineární regrese s hodnotami y_i'.**

Spočtěme další zajímavý rozptyl – rozptyl chybových členů $(y_i - y_i')$:

$$s_{Y-Y'}^2 = \frac{1}{6}\sum (y_i - y_i')^2 - \left(\frac{1}{6}\sum (y_i - y_i')\right)^2 = 18{,}1.$$

V příkladu 4.7 platí

$$s_Y^2 = 44{,}7 = 26{,}6 + 18{,}1 = s_{Y'}^2 + s_{Y-Y'}^2,$$

čili opět zde máme situaci dělení koláče rozptylu:

Celkový rozptyl s_Y^2 naměřených hodnot lze rozdělit do dvou složek – tou první je rozptyl $s_{Y'}^2$ obsažený v modelu lineární regrese, tou druhou složkou je rozptyl $s_{Y-Y'}^2$ způsobený chybou či nevhodností modelu (= rozptyl hodnot $y_i - y_i'$).

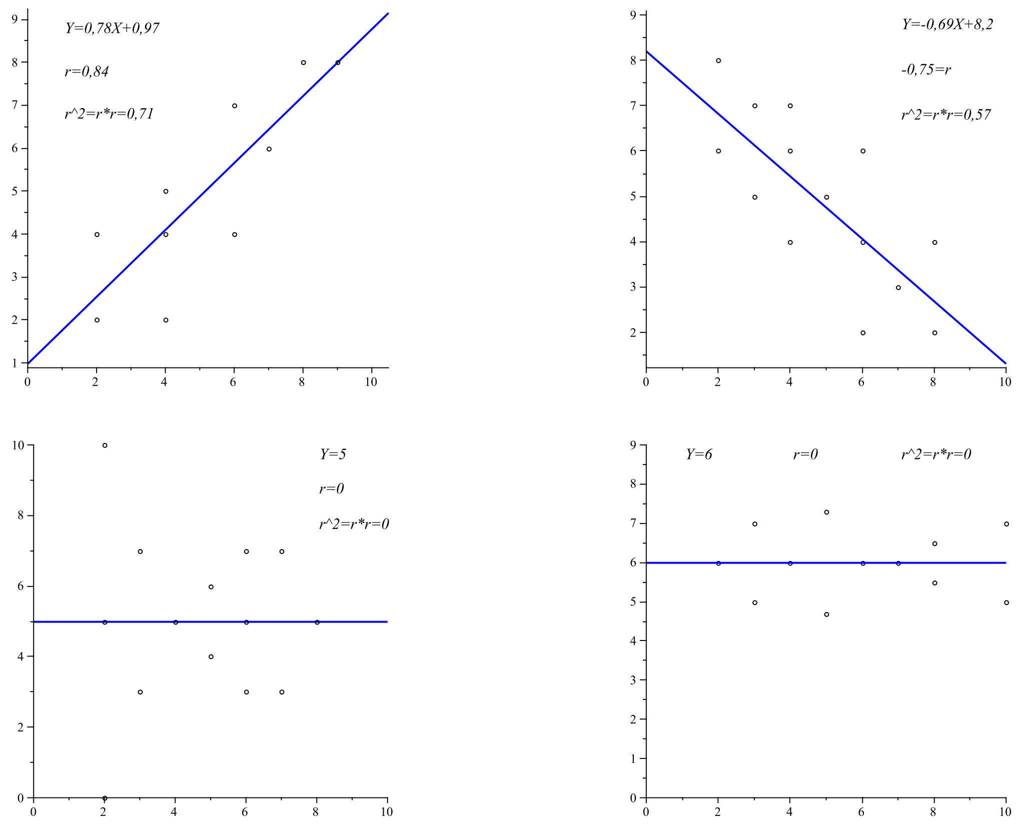
Protože r^2 vyjadřuje procentuelní podíl rozptylu zachyceného lineárním modelem, číslo $(1 - r^2)$ vyjadřuje procentuelní míru rozptylu nezachycenou lineárním modelem regresní přímky, čili míru rozptylu chybových členů $(y_i - y_i')$. Výraz $s_Y \cdot \sqrt{1 - r^2}$ se pak nazývá **střední chyba odhadu** – je to vlastně polovina intervalu spolehlivosti se středem v y_i' (67%-ní interval spolehlivosti). Konkrétně pokud pro danou hodnotu x_i chceme odhadnout y_i, tak hodnota $y_i' \pm s_Y \cdot \sqrt{1 - r^2}$ určuje interval, ve kterém se y_i při opakovaných výskytech bodů s hodnotou x_i nachází s pravděpodobností 67%.

Pearsonův koeficient r^2 určuje, do jaké míry model lineární regrese vysvětluje rozptyl naměřených hodnot. Dále se také často používá koeficient

$$r = \frac{n\sum(x_i y_i) - (\sum x_i)(\sum y_i)}{\sqrt{[n\sum x_i^2 - (\sum x_i)^2] \cdot [n\sum y_i^2 - (\sum y_i)^2]}}, \tag{4.6}$$

který navíc zachycuje, zda korelace mezi proměnnými X, Y je kladná, nebo záporná; $r \in \langle -1; 1\rangle$ a následující příklady uvádějí různé regresní modely společně s výpočtem míry r^2 vhodnosti modelu a korelačního koeficientu r (platí tedy, že koeficeint r^2 je druhou mocninou koeficientu r; naopak ovšem POZOR – koeficient r není odmocninou koeficientu r^2 v tom případě, když je r záporný):

Příklad 4.8. Různé příklady hodnot koeficientů r, r^2 a regresní přímky vidíte na obrázcích:



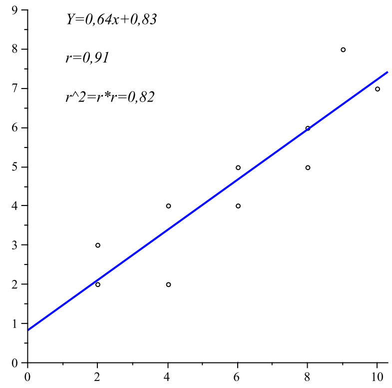
4.5 Test významnosti korelace

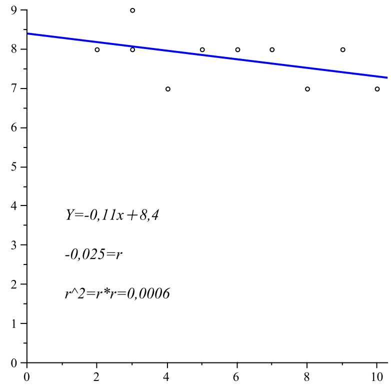
Zatím jsme mluvili o tom, co to korelace je a jak ji měřit. Nyní pár slov o statistické významnosti korelace. Vraťme se k našemu příkladu o fotbalu a pivu.

(ad příklad 4.7) Proveďme statistický test významnosti korelace: jedná se o oboustranný test, kde $t_k = t_k(\alpha = 2q)$*:*

K1: $H_0:\ \ r = 0$*;*
H_1 *: Otázkou je, zda užít* $r \neq 0$ *nebo* $r > 0$ *nebo* $r < 0$*. (jednostranný test* $r > 0$ *bychom užili tehdy, pokud bychom měli k dispozici teoretické (logické) opodstatnění, že jakákoli korelace musí být určitě kladná; alternativně* $r < 0$ *bychom užili, pokud bychom měli k dispozici opodstatnění, že jakákoli korelace, pokud existuje, je určitě záporná; větší diskuse těchto situací viz skripta BMA3, oddíl 13.5 (U-test)); pokud bychom neměli vůbec žádnou představu, jakým způsobem mohou nebo nemohou být studované veličiny korelovány, alternativní hypotéza by měla oboustranný tvar* $r \neq 0$*.*

Dejme tomu, že v našem příkladu se celkem zdá logické, že při větším počtu sledovaných fotbalových zápasů asi nebude tendence vypít méně piva, tj. pokud nějaká korelace existuje, je určitě kladná. Volme tedy hypotézu H_1 *:* $r > 0$ *(jednostranný test).*





K2: *Testovým kritériem bude veličina* $\frac{r\cdot\sqrt{n-2}}{\sqrt{1-r^2}}$.

K3: *Za předpokladu platnosti* H_0 *má testové kritérium rozdělení* t *o* $(n-2)$ *stupních volnosti (tuto skutečnost zde nebudeme dokazovat).*

K4: *Pro* $\alpha = 0{,}05$ *máme* $t_k(q = 0{,}05; \nu = 4) = 2{,}132$ *u jednostranného testu.*

K5: *Hodnota kritéria se rovná*

$$\frac{0{,}77\cdot\sqrt{4}}{\sqrt{1-0{,}59}} = 2{,}41 > 2{,}132,$$

a tedy zamítáme H_0*, korelace je statisticky významná.*

(pokud bychom chtěli být tak konzervativní, že bychom nechtěli připustit jako dostatečně logický předpoklad, že s počtem zápasů chuť přinejmenším neklesá, užili bychom oboustranný test, tj. hypotézu $H_1 : r \neq 0$*, kritickou hodnotu* $t_k(2q = 0{,}05, \nu = 4) = 2,776$*, a tedy výsledkem by bylo „*H_0 *nezamítáme“, což znamená, ne že korelace neexistuje, ale že náš test pro ni nenašel dostatek důkazů).*

Buď jak buď, ze tvaru testového kritéria je vidět, že pro rostoucí počet n *pozorování roste i hodnota tohoto kritéria, tj. i síla testu (pro rostoucí* n *roste pravděpodobnost odhalení korelovanosti obou proměnných),*

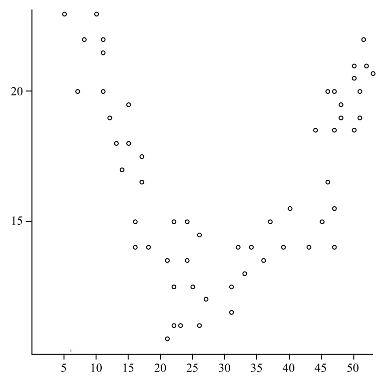
Nyní ještě pár slov o tzv. **zpětné (= inverzní) predikci**. Zatím jsme v modelu lineární regrese mluvili o situaci, že pomocí hodnoty x a regresní přímky se odhadovala hodnota y. Pokud bychom chtěli naopak pro známou hodnotu y odhadnout (predikovat) hodnotu x, museli bychom použít jinou regresní přímku než v dosud uvažovaném případě. Až dosud jsme uvažovali regresi $x \to y$, kde modelem je přímka minimlizující součet čtverců odchylek $(y_i - y_i')$ – pokud nyní chceme použít regresi $y \to x$, kde modelem je přímka minimalizující součet čtverců $(x_i - x_i')$, čili použít jiné vzorce pro nalezení optimálních koeficientů a, b:

$$b = \frac{n\sum(x_i y_i) - \left(\sum x_i\right)\left(\sum y_i\right)}{n\sum y_i^2 - \left(\sum y_i\right)^2}, \quad a = \frac{1}{n}\sum x_i - \frac{b}{n}\sum y_i, \tag{4.7}$$

kde regresní model je tvaru $x = a + by$. Pearsonův koeficient r^2 je u obou regresních přímek stejný. Tyto přímky se však protínají pouze v jediném bodě, a to je bod $[\overline{x}_x, \overline{x}_y]$.

4.6 Další typy regresních modelů

Samozřejmě že vztah mezi dvěma veličinami nemusí být lineární. Například pokud hledáme model (= funkci) popisující závislost věku člověka na době, za jakou uběhne sto metrů, je pravděpodobné, že grafické znázornění prováděného měření bude mít tvar podobný jako na následujícím obrázku:



Z obrázku je zřejmé, že parabola uvedenou závislost popíše lépe než přímka, tj. hledáme koeficienty a, b, c tak, aby funkce $y = a + bx + cx^2$ byla nejvhodnějším modelem ze všech těchto parabolických funkcí (lze najít opět metodou nejmenších čtverců – viz BMA3, kapitola 6). Pak mluvíme o tzv. **kvadratické regresi**.

Jiným typem regresního modelu je tzv. **multilineární model**, který popisuje lineární vztah mezi více než dvěma proměnnými.

Příklad 4.9. Komise výběrového řízení přijímá posluchače medicíny. Má k dispozici některé výsledky dříve přijatých studentů: jejich průměr známek na střední škole, bodové ohodnocení přijímacího testu, a také známkový průměr těchto studentů během studia medicíny:

student	X_1 = průměr VŠ	X_2 = přijímačky	Y =průměr SŠ
1	1,25	620	1,12
2	1,95	630	1,43
3	1,05	790	1,32
4	1,5	710	1,87
5	1,57	690	1,05
⋮	⋮	⋮	⋮

Dvě proměnné byly získány od studentů před započetím studia, třetí (= průměr na VŠ) až po absolvování studia, ke kterému se studenti hlásí. Vytvořme model multilineární regrese

$$Y = a + b_1 X_1 + b_2 X2.$$

Opět určitým postupem lze najít koeficienty a, b_1, b_2, které dobře popisují získaná data. A jaké bude kritérium komise u nových přijímacích zkoušek? Pomocí známých hodnot x_1, x_2 komise odhadne (= predikuje) hodnotu proměnné Y, která pak určuje pořadí v žebříčku přijímaných studentů.

4.7 Regrese směrem k průměru

Příklad 4.10. Střední hodnota IQ celé populace je 100 íkvacích bodů, směrodatná odchylka je 15 bodů. Uvažujme všechny matky, které mají jednu dceru. Kdybychom změřili IQ u některých z těchto matek a IQ jejich dcer, mohli bychom spočítat lineární regresní model a koeficient korelace r mezi proměnnými IQ matek a IQ dcer.

Uvažujme všechny matky, jejichž IQ je 130 íkvacích bodů. Pokud bychom pomocí našeho regresního modelu odhadovali IQ jejich dcer, odhadovaná hodnota by ležela někde mezi 130 a průměrem 100 – byla by zkrátka pravděpodobně (s vysokou pravděpodobností) blíže průměru než IQ jejich matek. Tomuto faktu se říká **regrese směrem k průměru**.

a) Pokud by korelační koeficient byl roven $r = 0$, jaký by byl odhad IQ dcer? $r = 0 \Rightarrow$ IQ matek 130 nelze užít k odhadu IQ dcer, tj. nejlepší odhad, který máme k dispozici pro IQ dcer, je průměr 100 celé populace.

b) $r = 1 \Rightarrow$ IQ matek má velký vliv na IQ dcer, a tedy odhad IQ dcer je 130.

c) $r \in (0; 1) \Rightarrow$ odhadovaná hodnota IQ dcer (= proměnná Y) záleží jak na IQ matek (= proměnné X), tak na střední hodnotě celé populace, a vypočteme ji jako vážený průměr

$$Y = r \cdot X + (1 - r) \cdot \mu,$$

protože oba zdroje informace $x = 130$, $\mu = 100$ mají určitou váhu.

Situaci jsme ještě trochu zjednodušili předpokladem, že proměnné X, Y mají stejný rozptyl. Obecně platí vztah mezi normovanými hodnotami

$$u_y = r \cdot u_x, \quad \text{neboli} \quad \frac{Y - \mu_y}{\sigma_y} = \frac{X - \mu_x}{\sigma_x}$$

(v celém modelu navíc platí $\mu_x = \mu_y$).

Příklad 4.11. Regrese směrem k průměru ve sportu.

Basketbalový tým Seattle Seahawks vyhrál v sezóně 1984 dvanáct zápasů ze šestnácti, pouze čtyři zápasy prohrál. V sezóně 1985, kdy se navíc po roční pauze způsobené zraněním vracel do hry obránce Curt Warner, se očekávala výhra v soutěži. Ale tým získal vítězství jen v osmi zápasech, zbývajících osm byly porážky. Jak je to možné?

Kromě jiných vlivů v týmu bylo možné matematicky zhruba tento výsledek odhadnout i pomocí regrese směrem k průměru. Označme X počet výher týmu v roce 1984, Y počet výher týmu v roce 1985. Korelační koeficient $r = 0{,}4$ vyjadřuje, že korelace mezi oběma ročníky určitě není dokonalá. průměrnbý počet výher týmu za sezónu je $\mu = 8$. Očekávaná (= predikovaná) hodnota s využitím regrese směrem k průměru je

$$y = 0{,}4 \cdot 12 + 0{,}6 \cdot 8 = 9{,}6.$$

A skutečně počet výher za sezónu 1985 se příliš neliší od odhadu 9,6.

Podobný způsob uvažování lze vztáhnout na další sporty. Například hráč basebalu s nejlepším ročním průměrem odpálení má statisticky velkou šanci, že příští rok bude jeho průměr horší. Ale existuje i šťastnější strana mince: hráč s nejnižším průměrem odpálení má statisticky velkou šanci, že příští rok se zlepší.

Příklad 4.12. Regrese směrem k průměru ve škole.

Milan Štencl v průběhu svého studia 1.ročníku FEI VUT (léta letí, kdysi ještě místo FEKT A FIT byla jedna fakulta; taky je vidět dlouhá doba mezi překladem tohoto příkladu z angličtiny a jeho zápisem do počítače) dosáhl průměru 1,0. Ovšem ve druhém ročníku průměr jeho známek klesl na 1,44. Zklamaní Milanovi rodiče připisují tento prospěchový skluz drogám, studentským akcím a holkám, nebo některé zákeřné kombinaci těchto tří zel (prosím jedná se o překlad příkladu z anglické učebnice, nikoli stanovisko autora).

Naštěstí Milan dostal jedničku ze statistiky a dobře zná matematickou zákonitost zvanou regrese směrem k průměru. Proto rodičům vysvětlil, že z čistě statistického hlediska student, který dosáhl v 1.ročníku výsledků vysoko nad průměrem školy, pochopitelně ve druhém ročníku svůj prospěch zhorší (nezapomněl také zdůraznit, že korelace mezi oběma ročníky není nijak velká).

Ovšem to, že Milanovi přátelé, jejichž výsledky v 1.ročníku byly zcela podprůměrné, se ve druhém ročníku zlepšili, jeho rodiče neuklidnilo.

Pojmy k zapamatování

- Ať už vztah mezi dvěma proměnnými je přímý (= změnou hodnot jedné proměnné se změní naměřené hodnoty druhé proměnné) nebo nepřímý (= tzv. **korelační vztah** … obě měřené veličiny vykazují určitý závislostní poměr, ale přímo jejich hodnoty jsou ovlivněny nějakou třetí veličinou), nejjednodušším modelem k popisu vztahu mezi dvěma veličinami je přímka. Konstrukce této přímky vychází z naměřeného vztahu mezi danými dvěma proměnnými. Široce přijímané kritérium pro nalezení této tzv. **regresní přímky** vede k nalezení stejné přímky jako je lineární funkce získaná metodou nejmenších čtverců (také viz BMA3, kapitola 6).
- Je také pochopitelné, že mezi některými dvojicemi veličin existuje jiný než přímkový vztah (např. vazba pomocí kvadratické funkce – nebo mezi danými dvěma veličinami žádný rozumný popsatelný vztah neexistuje), proto vstupuje na scénu i otázka, **do jaké míry je lineární regrese pro popis vztahu mezi dvěma veličinami vhodná**.
- Tuto vhodnost nebo nevhodnost měříme jednak Pearsonovým koeficientem $r^2 \in \langle 0; 1\rangle$, který vysvětluje procentuelní vhodnost popisu rozptylu měřených hodnot přímkovým vztahem; dále korelačním koeficientem $r \in \langle -1; 1\rangle$, který navíc zachycuje, zda je daná kolerace kladná nebo záporná; a konečně lze provést statistický test významnosti přímkové korelace.

Kontrolní otázky

1. U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.
   a) Predikce znamená, že pomocí změny hodnoty první veličiny ovlivníme velikost hodnoty druhé veličiny.
   b) Korelační přístup používáme například v situaci, kdy nemůžeme předem ovlivnit hodnotu žádné z obou studovaných veličin.
   c) Korelační vztah mezi veličinami je možné vysvětlit minimálně třemi různými způsoby.
   d) Regresní přímka je jiná přímka než přímka získaná metodou nejmenších čtverců (BMA3, kap.6).
   e) Regresní přímka $x \to y$ (pomocí hodnot x predikujeme hodnoty y) se liší od regresní přímky $y \to x$ (pomocí hodnot y predikujeme hodnoty x), i když pro nalezení obou přímek používáme stále tytéž naměřené body.
   f) Pearsonův koeficient r^2 je druhou mocninou korelačního koeficientu r.
   g) Pearsonův koeficient procentuelně vyjadřuje vhodnost modelu přímkové regrese pro zadané body v rovině.
   h) Pokud $r = 0$, znamená to, že jedna veličina mění své hodnoty a druhá veličina je neustále konstatní.

Odpovědi na otázky
1a) – N, 1b) – A, 1c) – A, 1d) – N, 1e) – A, 1f) – A, 1g) – A, 1h) – N (pokud jedna veličina je neustále konstantní, pak $r = 0$, ale naopak pokud $r = 0$, tak žádná z veličin konstantní být nemusí).

Cvičení

1. Trenéra basebalu zajímá vztah mezi věkem sportovce a jeho úspěšností odpálení. Náhodně vybraných dvanáct hráčů dosáhlo těchto výsledků (v prvním řádku věk, ve druhém průměrná úspěšnost odpálení):

18	17	31	25	22	24	28	21	21	18	35	41
0,225	0,35	0,15	0,275	0,269	0,2	0,32	0,315	0,195	0,2	0,31	0,275

   a) Určete regresní přímku predikující průměr úspěšnosti na základě věku sportovce.

   b) Jaké procento rozptylu hodnot úspěšnosti je popsáno regresní přímkou?

   c) Proveďte test významnosti korelace.

2. Sportovního psychologa zajímá vztah mezi hmotností diskaře a jeho diskařskou kvalitou. Vybere náhodný vzorek patnácti lidí a změří jejich hmotnost a vzdálenost jejich hodu diskem. Získává data (hmotnost v librách, 1 libra = 0,4536 kg; hod diskem ve stopách, 1 stopa = 12 palců (inches) = 30,48 cm):

osoba	hmotnost	hod diskem
Marie	120	125
Honza	165	215
Eva	105	145
Zuzana	128	129
Robert	220	175
Jiří	170	209
Amálie	115	141
Dušan	156	223
Linda	125	130
Tomáš	190	200
Božena	160	132
Martin	130	250
Petr	200	180
Andrea	100	150
Sylva	130	135

a) Znázorněte tato data v rovině.

b) Určete regresní přímku pro predikci vzdálenosti na základě hmotnosti.

c) Vypočtěte Pearsonův koeficient r^2 – liší se tento statisticky významně od nuly?

d) Zopakujte b),c) pro muže a ženy zvlášť.

3. Vývojový psycholog nalezl dvanáct párů dvojčat, z nichž vždy jedno žije ve vyšším společensko-ekonomickém prostředí, a to druhé v nižším. Všech 24 dětí se podrobilo testu IQ, vysledky (v íkvacích bodech) jsou v tabulce:

číslo páru	dvojče ve „vyšším“ prostředí	dvojče v „nižším“ prostředí
1	98	100
2	115	95
3	101	80
4	125	98
5	120	120
6	102	98
7	92	80
8	110	103
9	105	104
10	75	68
11	112	111
12	90	110

a) Najděte regresní přímku pro predikci IQ „vyššího“ dvojčete na základě IQ „nižšího“ dvojčete.

b) Najděte regresní přímku pro predikci IQ „nižšího“ dvojčete na základě IQ „vyššího“ dvojčete.

c) Jaký je koeficient r^2 mezi IQ obou skupin?

d) Je korelace obou skupin významně odlišná od nuly?

4. Sledováním nákladů X a ceny Y stejného typu výrobku u deseti výrobců byl získán dvourozměrný statistický soubor s koeficientem korelace $r = 0{,}82482$. Na hladině významnosti 0,01 testujte hypotézu, že veličiny X, Y jsou nekorelované.

5. V telefonní společnosti zjišťují, zda volba pracovní pozice (operátor nebo vedoucí) záleží na pohlaví. Deset zaměstnanců společnosti je náhodně vybráno a je jim přiřazena 1 (= žena) nebo 0 (= muž), a ve vztahu ke druhé proměnné 1 (= operátor) nebo 0 (= vedoucí).Byla získána následující data:

osoba	pohlaví x	typ práce y
1	1	1
2	0	1
3	1	0
4	1	0
5	0	0
6	1	1
7	0	0
8	1	1
9	1	1
10	0	0

a) Určete r^2 pro tato data.

b) Proveďte test významnosti hodnoty r.

Výsledky

ad 1. **ad a)** Regresní přímka má tvar $y = 0{,}245348 + 0{,}00046453 \cdot x$.

ad b) $r^2 = 0{,}00307873$, a tedy pouze 0,3% roptylu hodnot je popsáno regresní přímkou.

ad c) Oboustranný test významnosti korelace: $H_0 : r = 0$, $H_1 : r \neq 0$, pak pro $\alpha = 0{,}05 = 2q$ máme $t_k(10) = 2{,}228$. hodnota kritéria testu se rovná

$$\frac{0{,}0554863 \cdot \sqrt{10}}{\sqrt{1 - 0{,}00307873}} = 0{,}17 < 2{,}228;$$

tj. H_0 nezamítáme, korelace mezi veličinami není významná.

ad 2. **ad b)** regresní přímka: $v = -0{,}489 \cdot h + 97{,}152$;

ad c) $r^2 = 0{,}184$, což se významně neliší od nuly;

ad d) muži: regrese $v = -0{,}846 \cdot h + 356{,}183$, $r^2 = 0{,}969$ (korelace je významná); ženy: regrese $v = -0{,}292 \cdot h + 171{,}767$, $r^2 = 0{,}386$ (korelace není významná).

ad 3. **ad a)** $v = 0{,}587 \cdot n + 46{,}634$.

ad b) $n = 0{,}659 \cdot v + 28{,}828$.

ad c) $r^2 = 0{,}387$.

ad d) Jedná se o statisticky významnou korelaci.

ad 4. Jedná se o test $H_0 : r = 0$, $H_1 : r \neq 0$ s kritériem

$$T = \frac{r \cdot \sqrt{n-2}}{\sqrt{1-r^2}}.$$

Pro $\alpha = 0{,}01 = 2q$ máme

$$T = 4{,}1262 > t_{0{,}995}(8) = 3{,}355;$$

tj. hypotézu H_0 zamítáme, korelace veličin je statisticky významná.

ad 5. $r^2 = 0{,}167$, což není statisticky významná korelace.

5 Po analýze rozptylu nebo místo ní

Průvodce studiem

V kapitole 3 jsme testovali hypotézu $H_0 : \mu_1 = \mu_2 = \cdots = \mu_J$ oproti alternativní hypotéze, že H_0 neplatí. Ale hypotéza $H_1 : (H_0$ neplatí) nám dává jen malou informaci o rozdílech mezi jednotlivými skupinami měření, respektive mezi středními hodnotami jednotlivých veličin – neříká nám, která z veličin má rozdílnou střední hodnotu a o kolik.

Podobně při dvourozměrné ANOVA typu 3×4 (viz oddíly 3.1.2, 3.1.3) jsme testovali, zda existuje vliv každého z faktorů a vliv interakce faktorů na hodnotu pozorované proměnné, ale nezabývali jsme se otázkou, co způsobuje rozdíly mezi třemi podmínkami faktoru 1, čtyřmi podmínkami faktoru 2 a která z podmínek má vliv na interakci (a do jaké míry).

Na tyto otázky dává jakousi počáteční odpověď průměr hodnot měření a interval spolehlivosti se středem v průměru každé ze situací experimentu. Ale v této kapitole se budeme zabývat dalšími metodami nebo ukazateli rozdílů mezi středními hodnotami jednotlivých veličin.

Cíle

Po prostudování této kapitoly budete schopni:

- Pracovat s Pearsonovým r^2 koeficientem.
- Vědět, co dělat, když F-test prokáže, že hypotéza H_0 neplatí. Jednou z možných cest je porovnávat každo dvojici skupin měření ve zvláštním statistickém testu, což je sice možné (oddíl 5.0.1 – metoda minimálního významného rozdílu, Scheffeova metoda). Další alternativní možností je tzv. plánované srovnání
- Použít získané znalosti pro řešení konkrétních problémů.

5.0.1 Testování „post-hoc“

Pokud nemáme žádnou předem známou hypotézu o tom, která z testovaných veličin má větší průměr než ty ostatní, použijeme právě testování typu „post-hoc“ (= „po tom“, tedy o hodnotě veličin usuzujeme až po provedení testu, předem není k dispozici žádná teorie).

Příklad 5.1. Chceme zjistit vliv různých metod studia na výsledky zkoušky – padesát náhodně vybraných studentů rozdělíme do pěti skupin po deseti studentech a požádáme, aby se připravovali danou metodou:

- skupina 1: čtení … čtou skripta, ale nechodí na přednášky
- skupina 2: přednáška … chodí na přednášky, ale vůbec nečtou skripta
- skupina 3: svoboda … mají svobodu učit se, jak chtějí
- skupina 4: konzultace I … mají jednu hodinu osobní konzultace týdně
- skupina 5: konzultace II … mají dvě hodiny osobní konzultace týdně

Na konci roku jsme získali průměr výsledků zkoušky v každé ze skupin (v každé skupině $N = 10$):

(čtení)	(přednáška)	(svoboda)	(konzultace I)	(konzultace II)
$\mu_1 \in 82 \pm 6{,}36$	$\mu_2 \in 95 \pm 6{,}36$	$\mu_3 \in 71 \pm 6{,}36$	$\mu_4 \in 83 \pm 6{,}36$	$\mu_5 \in 85 \pm 6{,}36$
$T_1 = 820$	$T_2 = 950$	$T_3 = 710$	$T_4 = 830$	$T_5 = 850$

pro $n = 50$, $T = 4160$. Intervaly spolehlivosti jsou 95%-ní: $\overline{x} \pm t_k \cdot \sqrt{\frac{estRUT}{10}}$.

Provedeme-li klasickou jednorozměrnou ANOVA (oddíl 3.1), obdržíme následující výsledky:

typ rozptylu	V	SS	$est\ R$	F-hodnota	F_k
celkový rozptyl	49	7428	–	–	–
MT	4	2928	732	7,32	2,61
UT	45	4500	100	–	–

Protože $7{,}32 > 2{,}61$, tak zamítáme $H_0 : \mu_1 = \mu_2 = \cdots = \mu_5$. Dokázali jsme tedy, že střední hodnoty veličin se statisticky významně liší, ale nevíme, které. Proto použijeme testování „post-hoc“.

a) **Metoda minimálního významného rozdílu.** Proveďme jeden z deseti dílčích testů, např.

K1: $H_0 : \mu_1 = \mu_2;\quad H_1 : \mu_1 \neq \mu_2.$

K2: Testovým kritériem je veličina $\frac{\overline{X}_1 - \overline{X}_2}{est\ \sigma_{\mu_1-\mu_2}}$.

K3: Jedná se o obyčejný t-test s tím rozdílem, že pro výpočet $est\ RUT$ využijeme i měření ve skupinách 3,4 a 5 (odhad rozptylu pak bude přesnější): $VUT = 45$, $est\ RUT = 100$. Tedy za předpokladu platnosti H_0 má kriterijní funkce t-rozdělení se 45 stupni volnosti. Pak

$$est\ \sigma_{\mu_1-\mu_2} = \sqrt{\frac{est\ RUT}{10} + \frac{est\ RUT}{10}} = \sqrt{10+10} = 4{,}47.$$

K4: $t_k(V = 45, \alpha = 0{,}05) = \pm 2{,}02.$

K5: Po dosazení

$$\frac{\overline{x}_1 - \overline{x}_2}{est\ \sigma_{\mu_1 - \mu_2}} = \frac{82 - 95}{4{,}47} = -2{,}91 \notin (-2{,}02; +2{,}02),$$

a tedy rozdíl $\mu_1 - \mu_2$ je statisticky významný.

A nyní: pro každé dva z devíti zbývajících testů platí následující podmínka statstické významnosti:

$$\frac{|\overline{x}_i - \overline{x}_j|}{4{,}47} \geq 2{,}02, \text{ tj. } \quad |\overline{x}_i - \overline{x}_j| \geq 2{,}02 \cdot 4{,}47 \doteq 9{,}03.$$

Toto číslo 9,03 je tzv. **minimální významný rozdíl**, podle něhož lze rychle rozhodnout, o statistické významnosti jednotlivých rozdílů mezi dvěma středními hodnotami.

V našem příkladu deseti testů tedy můžeme psát odpověď: Statisticky významné jsou rozdíly vždy mezi dvojicí středních hodnot veličin 1 a 2, 1 a 3, 2 a 3, 2 a 4, 2 a 5, 3 a 4, 3 a 5.

Nevýhoda metody:

Potřebujeme otestovat, které z možných $\binom{5}{2}$ dvojic veličin (obecně $\binom{J}{2}$ dvojic) mají statisticky významně odlišné střední hodnoty. Musíme tedy provést $\binom{5}{2}$ testů. Zde dochází k jistým komplikacím pro hladinu významnosti α celkové výpovědi:

Předpokládáme-li nezávislost deseti testů dvojic (deseti t-testů, každý z nich je prováděn na hladině významnosti α), pak pravděpodobnost, že aspoň v jednom z nich provedeme nesprávné rozhodnutí 1.druhu, je rovna

$$1 - p(\text{všech 10 rozhodnutí bude správných}) = 1 - 0{,}95^{10} \doteq 0{,}4,$$

což je docela vysoká pravděpodobnost výskytu chyby 1.druhu.

Nezávislost jednotlivých testů ale téměř nikdy nemůžeme předpokládat (pokud by se například v našem příkladu stalo, že do skupiny 3 by byli vybráni nepříliš inteligentní studenti, tak $\overline{x}_3$ by byl nízký, a to by ovlivnilo ne jeden, ale hned čtyři z testů: test $\mu_3 - \mu_1$, test $\mu_3 - \mu_2$, test $\mu_3 - \mu_4$, test $\mu_3 - \mu_5$; čili testy v našem příkladu nejsou nezávislé).

Tedy celková výpověď deseti prováděných testů nejen že má velkou celkovou chybu prvního druhu ($\geq 0{,}4$), ale v případě závislosti jednotlivých testů, což je spíše pravidlem, nemůžeme pravděpodobnost výskytu chyby prvního druhu vůbec určit.

b) Schefféova metoda. Tato metoda nalezne takové kritérium testu, že celková chyba prvního druhu po provedení všech dílčích testů bude $\leq 0{,}05$ (to je zvolené rozumné α pro celkovou výpověď experimentu).

K1: $H_0 : \mu_1 = \mu_2; \quad H_1 : \mu_1 \neq \mu_2.$

K2: Testovým kritériem je veličina

$$\left(\frac{\overline{X}_1-\overline{X}_2}{est\ \sigma_{\mu_1-\mu_2}}\right)^2 .$$

K3: Za předpokladu platnosti H_0 má kritérium rozdělení $4\cdot F(4,45)$ (obecně je to rozdělení $(J-1)\cdot F(J-1,n-J)$). Tento fakt nyní dokazovat nebudeme (až někdy příště).

K4: Kritická hodnota $4\cdot F_k(4,45)=4\cdot 2{,}61=10{,}44$.

K5: Po dosazení

$$\left(\frac{\overline{x}_1-\overline{x}_2}{est\ \sigma_{\mu_1-\mu_2}})^2=(-2{,}91)^2=8{,}47<10{,}44\right),$$

a tedy hypotézu H_0 nezamítáme, rozdíl středních hodnot veličin není statisticky významný.

Lze určit i minimální významný rozdíl u Scheffého testu: v kritickém případě $t^2=$ $=10{,}44$, a tedy $t=\sqrt{10{,}44}=\pm 3{,}23$. Pak podmínka významnosti má tvar

$$\frac{|\overline{x}_i-\overline{x}_j|}{4{,}47}\geq 3{,}23,\ \text{ tj. }\quad |\overline{x}_i-\overline{x}_j|\geq 3{,}23\cdot 4{,}47\doteq 14{,}44.$$

Podmínka pro rozdíl průměrů je u Scheffého testu přísnější – podle Scheffého testu existuje při vzájemném porovnání uvedených pěti veličin jen jediný významný rozdíl, a to mezi veličinami 2 a 3, protože jen zde přesáhne rozdíl průměrů hodnotu 14,44.

Důležitost Scheffého testu: zvyšujeme minimální významný rozdíl, aby celková pravděpodobnost výskytu chyby prvního druhu po provedení všech deseti (obecně $\binom{J}{2}$) dílčích testů byla menší nebo rovna rozumně zvolené hodnotě 0,05.

c) Další „post-hoc" přístupy: Můžeme shrnout, že
metoda **a)** udržuje malou pravděpodobnost β za cenu velké hodnoty α,
metoda **b)** udržuje malou pravděpodobnost α za cenu velké hodnoty β.

V praxi se někdy používá dalšího postupu, jehož popis je mimo rámec tohoto textu. Pokus o naznačení postupu: Průměry $\overline{x}_1,\ldots,\overline{x}_J$ se srovnají od nejmenšího k největšímu. Pak se

1. srovná nejmenší a největší průměr s využitím jistého kritéria k_1 (při **ne**zamítnutí H_0 se celý proces zastaví).
2. srovná – minimální průměr s druhým největším průměrem;
   – maximální průměr s druhým nejmenším průměrem
   s využitím jistého kritéria k_2 (při **ne**zamítnutí H_0 se celý proces zastaví).
3. celý proces pokračuje pro stále se snižující hodnotu kritéria, dokud některý test neselže (tímto způsobem je nalezen jakýsi statisticky významný rozdíl).

5.0.2 Plánované srovnání v experimentu typu „více vzorků jednou“

Příklad 5.2. Představte si, že jste sociologem zkoumajícím volební preference a chcete ověřit hypotézu, že starší lidé jsou konzervativnější (= nemění rychle své volby). Proto bylo náhodně vybráno šest osob v každé z daných věkových kategorií a všem byl předložen dotazník, který má zjistit, do jaké míry jsou konzervativní (dotazníkem se nyní nebudeme zabývat – stačí nám vědět, že větší počet dosažených bodů v dotazníku znamená větší míru konzervativnosti). Byly získány tyto výsledky (ve všech kategoriích $N = 6$; dále celkem $n = 30$; $SSUT = 300$, $VUT = 25$):

kategorie 1 (20-29 let)	kategorie 2 (30-39 let)	kategorie 3 (40-49 let)	kategorie 4 (50-59 let)	kategorie 5 (60-69 let)
$\mu_1 \in 2 \pm 2{,}91$	$\mu_2 \in 4 \pm 2{,}91$	$\mu_3 \in 4 \pm 2{,}91$	$\mu_4 \in 8 \pm 2{,}91$	$\mu_5 \in 12 \pm 2{,}91$
$T_1 = 12$	$T_2 = 24$	$T_3 = 24$	$T_4 = 48$	$T_5 = 72$

Průměry vykazují rostoucí charakter se zvyšováním věku. Potřebujeme ale zjistit, do jaké míry je tento růst statisticky významný (zda není způsoben pouze náhodnými vlivy). provedeme tedy tzv. **plánované srovnání**:

a) Prvním krokem je volba tzv. vah – to jsou čísla, která budou reprezentovat naši hypotézu o rostoucí konzervativnosti. Váhy musí splňovat tyto podmínky:

1. Každé kategorii je přiřazena jedna váha.
2. Součet všech vah je roven nule.

V našem případě nárůst míry konzervativnosti popisují například váhy $w_1 = 1$, $w_2 = 2$, $w_3 = 3$, $w_4 = 4$, $w_5 = 5$; ovšem aby byla splněna podmínka nulového součtu, musíme volit například $w_1 = -2$, $w_2 = -1$, $w_3 = 0$, $w_4 = 1$, $w_5 = 2$.

b) Nyní nastupuje otázka, do jaké míry jsou odhady $\overline{x}_j$ středních hodnot μ_j v korelaci s našimi hypotézovými vahami w_j. Spočtěme Pearsonův koeficient r^2 (úprava v následujícím vzorci užívá faktu, že $\sum w_j = 0$):

$$r^2 = \frac{\left[J\sum_1^J w_j\overline{x}_j - \left(\sum_1^J w_j\right)\cdot\left(\sum_1^J \overline{x}_j\right)\right]^2}{\left[J\sum_1^J w_j^2 - \left(\sum_1^J w_j\right)^2\right]\cdot\left[J\sum_1^J \overline{x}_j^2 - \left(\sum_1^J \overline{x}_j\right)^2\right]} =$$

$$= \frac{J^2\cdot\left(\sum_1^J w_j\overline{x}_j\right)^2}{\left(J\cdot\sum_1^J w_j^2\right)\cdot\left(J\cdot\sum_1^J \overline{x}_j^2 - \left(\sum_1^J \overline{x}_j\right)^2\right)}.$$

Vynásobíme-li čitatele i jmenovatele v posledním zlomku konstantou $\frac{N}{J^2}$ a nahradíme-li $\overline{x}_j = \frac{T_j}{N}$, dostaneme (rovněž využíváme toho, že $JN = n$, a také $\sum T_j = T$)

$$r^2 = \frac{N \cdot \left(\sum_1^J w_j \overline{x}_j\right)^2}{\left(\sum_1^J w_j^2\right) \cdot \left(\sum_1^J \frac{T_j^2}{N} - \frac{1}{JN}\left(\sum_1^J T_j\right)^2\right)} = \frac{N \cdot \left(\sum_1^J w_j \overline{x}_j\right)^2}{\left(\sum_1^J w_j^2\right) \cdot SSMT}.$$

Nyní je důležité si zpomenout, že koeficient r^2 vyjadřuje, kolik procent rozptylu mezi hodnotami y_i je popsáno uvažovaným modelem, respektive při současném označení, kolik procent různosti mezi průměry $\overline{x}_j$ je popsáno hypotézou lineárního růstu $\vec{w} = (-2, -1, 0, 1, 2)$. Různost mezi průměry je vyjádřena pomocí $SSMT$, zbylý činitel ve zlomku pro výpočet r^2 vyjadřuje tedy míru různosti průměru popsanou naší hypotézou – označme tento člen jako SSH (součet čtverců hypotézy):

$$r^2 = \frac{\frac{N \cdot \left(\sum_1^J w_j \overline{x}_j\right)^2}{\sum_1^J w_j^2}}{SSMT} = \frac{SSH}{SSMT}, \quad \text{tj. } SSH = r^2 \cdot SSMT = \frac{N \cdot \left(\sum_1^J w_j \overline{x}_j\right)^2}{\sum_1^J w_j^2}.$$

V našem příkladu $SSH = 345{,}6$, $SSMT = 384$, tj. $r^2 = 0{,}9$ – 90/rozptylu mezi vzorky je popsáno naší hypotézou.

c) Testujeme hypotézu o významnosti korelace:

Krok 1: H_0: w_j, $\overline{x}_j$ nejsou korelovány;
H_1: jsou.

Krok 2: Kritériem bude $\frac{est\ RH}{est\ RUT}$, kde $est\ RH = \frac{SSH}{VH}$ a víme, že $VH = 1$, protože SSH bylo konstruováno jako představitel jednoho konkrétního průběhu růstu průměrů (RH = rozptyl hodnot průměrů zachycený naší hypotézou; vždy platí $est\ RH = \frac{SSh}{VH} = \frac{SSH}{1} = SSH$).

Krok 3: Za předpokladu platnosti hypotézy H_0 má podíl $\frac{est\ RH}{est\ RUT}$ rozdělení $F(1, 25)$, protože $est\ RH$, $est\ RUT$ jsou odhady téhož rozptylu. Ze zadání víme, že $SSUT = 300$, což bylo vypočteno z jednotlivých měření (dále $VUT = n - J = 30 - 5 = 25$).

Krok 4: Pro $\alpha = 0{,}05$ je $F_k(1, 25) = 4{,}24$.

Krok 5: Hodnota kritéria

$$\frac{est\ RH}{est\ RUT} = \frac{345{,}6}{12} = 28{,}8 > 4{,}24 = F_k,$$

zamítáme tedy H_0, korelace průměrů $\overline{x}_j$ a navržených vah w_j je statisticky významná.

Uvedený test lze shrnout do tabulky podobné tabulce ANOVA:

typ rozptylu	V	SS	$est\ R$	F-hodnota	F_k
celkový	29	684	–	–	–
MT	4	384	–	–	–
– H	1	345,6	345,6	28,8	4,24
UT	25	300	12	–	–

d) Kromě testu c) o nulové korelaci můžeme testovat také následující hypotézu dokonalé korelace. Pro tento test budeme potřebovat informace o zbývající části RMT – o tzv. zbytkovém rozptylu. Doplňme tabulku c) o tyto hodnoty:

typ rozptylu	V	SS	$est\ R$	F-hodnota	F_k
celkový	29	684	–	–	–
MT	4	384	–	–	–
– H	1	345,6	345,6	28,8	4,24
– Z	3	38,4	12,8	1,07	$2{,}99 = F_k(3, 25)$
UT	25	300	12	–	–

Obecně $VZ = J - 2 = VMT - VH$, $SSZ = SSMT - SSH$, $est\ RZ = \frac{SSZ}{VZ}$, kde RZ je rozptyl zbytku (zbytkový rozptyl).

Pokud průměry $\overline{x}_j$ jsou dokonale popsány vahami w_j, bude RZ roven nule? Ne, nějaké odchylky od w_j zde budou – způsobené rozptylem RUT. Čili v případě, že $\overline{x}_j$ jsou v dokonalé korelaci s vahami w_j, platí $RZ = RUT$. Provedeme tedy F-test o shodnosti rozptylů:

Krok 1: H_0: $\overline{x}_j$, w_j jsou perfektně korelované, tedy $r^2 = 1$.
H_1: neplatí H_0.

Krok 2: Kritériem bude podíl odhadů rozptylů, které porovnáváme, tedy $\frac{est\ RZ}{est\ RUT}$.

Krok 3: Za předpokladu platnosti H_0 má podíl $\frac{est\ RZ}{est\ RUT}$ rozdělení $F(3; 25)$.

Krok 4: Pro $\alpha = 0{,}05$ je $F_k(3; 25) = 2{,}99$.

Krok 5: $\frac{est\ RZ}{est\ RUT} = 1{,}07$, což je menší než kritická hodnota, tj. H_0 nezamítáme, korelace je perfektní.

Všimněme si, že test c) je zcela odlišného charakteru z hlediska výpovědi než F-test při analýze rozptylu: místo obecného testu, zda existují rozdíly mezi skupinami, jsme nyní v testu c) testovali, do jaké míry odpovídají naměřené průměry $\overline{x}_j$ naší konkrétní hypotéze o jejich rozdílnosti (tato konkrétní rozdílnost je vyjádřená vahami w_j).

Příklad 5.3. Chceme ověřit hypotézu, že marihuana prodlužuje časový odhad, tj. po požití marihuany se daný časový interval (např. pět minut) jeví člověku mnohem delší. Vybrali jsme náhodně čtyřicet pravidelných kuřáků marihuany a rozdělili do pěti skupin po osmi lidech. Každý člověk v dané skupině vypotřeboval následující dávku cigaret:

- skupina 1 – jedna cigareta s marihuanou

- skupina 2 – dvě cigarety s marihuanou
- skupina 3 – nic
- skupina 4 – jedna placebo cigareta (nevěděli, že v ní není přítomna marihuana)
- skupina 5 – dvě placebo cigarety (nevěděli, že v nich není přítoma marihuana)

Pak byl proveden odhad časového limitu pěti minut a získala se data

(skupina 1)	(skupina 2)	(skupina 3)	(skupina 4)	(skupina 5)
$\mu_1 \in 7 \pm 1{,}05$	$\mu_2 \in 10 \pm 1{,}05$	$\mu_3 \in 5 \pm 1{,}05$	$\mu_4 \in 5 \pm 1{,}05$	$\mu_5 \in 6 \pm 1{,}05$
$T_1 = 56$	$T_2 = 80$	$T_3 = 40$	$T_4 = 40$	$T_5 = 48$

Ve všech skupinách $N = 8$. Spočetl se $SSUT = 75$, a protože $VUT = 35$, mohli být stanoveny intervaly spolehlivosti (95%-ní) pro jednotlivé střední hodnoty: $\overline{x}_j \pm 1{,}05$.

V tomto experimentu budeme chtít testovat dvě hypotézy:

H_1: Marihuana prodlužuje odhad času na rozdíl od těch skupin, kde ji nekouřili. Vhodné váhy jsou například $\vec{w}_1 = (3, 3, -2, -, 2-2)$.

H_2: Čím více marihuany se vykouří, tím delší je časový odhad. Vhodné váhy jsou např. $\vec{w}_2 = (-1, 1, 0, 0, 0)$.

Můžeme nyní vypočítat součty čtverců příslušející jednotlivým hypotézám podle vzorce

$$SSH_i = \frac{N\left(\sum_{j=1}^{J} w_{ij}\overline{x}_j\right)^2}{\sum_{j=1}^{J} w_{ij}^2}, \tag{5.1}$$

tj. $SSH_1 = 96{,}27$, $SSH_2 = 36$. Dále $VH_1 = VH_2 = 1$, protože danému tvaru vah odpovídá vždy jeden stupeň volnosti – tyto váhy představují jeden možný průběh korelace. Nyní lze H_1, H_2 testovat testem analogickým testu c) předchozího příkladu 5.2. Napišme zde pouze tabulku shrnující daná data, ze které bude patrný výsledek testů:

typ rozptylu	V	SS	procento MT	$est\ R$	F-hodnota	F_k
celkový	39	212,6	–	–	–	–
MT	4	137,6	–	–	–	–
– H_1	1	96,3	70	96,3	44,9	$4{,}17 = F_k(1;35)$
– H_2	1	36	26	36	16,8	$4{,}17 = F_k(1;35)$
– Z	2	5,3	4	2,65	1,24	$3{,}32 = F_k(2;35)$
UT	35	75	–	2,14	–	–

Oba testy typu c) podporují platnost hypotéz H_1, H_2 a test d) naznačuje, že zbytkový rozptyl je už roven rozptylu RUT. Je vidět, že platí

$$SSMT = SSH_1 + SSH_2 + SSZ = 137{,}6$$

$$VMT \quad = \quad VH_1 + VH_2 + VZ = 4.$$

Pokud bychom hypotézu H_2 změnili a chtěli místo ní ověřovat hypotézu

H_2': pro významné prodloužení délky časového odhadu je potřeba vykouření dvou a více cigaret s marihuanou. Vhodnými vahami zde jsou např. $\vec{w}_2' = (-1; 4; -1; -; 1-; 1)$,

dostaneme $SSH_2' = 115{,}6$, a tedy $SSH_1 + SSH_2' = 96{,}3 + 115{,}6 = 211{,}9 > 137{,}6 = SSMT$. To nastalo díky tomu, že hypotézy H_1, H_2' nejsou nezávislé.

H_1, H_2 jsou nezávislé: $H_1 : \vec{w}_1 = (3, 3, -2, -2, -2)$, $H_2 : \vec{w}_2 = (-1, 1, 0, 0, 0)$. H_1 pouze říká, že skupiny 1 a 2 vykazují větší časovou prodlevu než ty ostatní. H_2 říká něco zcela odlišného a nezávislého – skupina 2 vykazuje větší prodlevu než skupina 1. Intuitivně je vidět nezávislost H_1, H_2: pokud známe výsledek o vztahu (skupina 1,2) versus (skupiny 3,4,5), neříká nám to nic o vztahu skupina 1 versus skupina 2.

H_1, H_2' jsou závislé: $H_1 : \vec{w}_1 = (3, 3, -2, -2, -2)$, $H_2 : \vec{w}_2' = (-1, 4, -1, -1, -1)$. H_1 říká, že skupiny 1,2 vykazují delší časový odhad než skupiny 3,4,5; H_2' říká, že skupina 2 vykazuje delší časový odhad než skupiny 1,3,4,5. Intuitivně je vidět závislost H_1, H_2': pokud platí H_2', věřili bychom, že platí H_1. Tedy platnost H_2' má vliv na platnost hypotézy H_1.

Obecně budeme ověřovat nezávislost hypotéz matematicky pomocí vah: H_1, H_2 jsou nezávislé, pokud (a jen tehdy, když) $\sum_{j=1}^{J} w_{1j} \cdot w_{2j} = 0$, tj. skalární součin příslušných váhových vektorů je roven nule. V našem příkladu skutečně $\vec{w}_1 \cdot \vec{w}_2 = 0$, tj. H_1, H_2 jsou nezávislé. Otázka nyní zní: je možné najít hypotézu H_3 nezávislou na každé z hypotéz H_1, H_2? Uvažujme např.

H_3 : I u placebo cigaret platí, že s větší spotřebou těchto cigaret se prodlužuje časový odhad. Vhodnými vahami jsou např. $\vec{w}_3 = (0, 0, 0, -1, 1)$.

Výpočtem se lze přesvědčit, že $\vec{w}_1 \cdot \vec{w}_3 = 0 = \vec{w}_2 \cdot \vec{w}_3$, tj. H_3 je nezávislá na H_1 i na H_2. Dále $SSH_3 = 4$, $VH_3 = 1$. Kolik navzájem nezávislých hypotéz je možné formuovat celkem? Tolik, kolik je hodnota VMT – v našem příkladu tedy čtyři. A skutečně je možné zformulovat hypotézu H_4 nezávislou na H_1, H_2 a H_3:

H_4 : Vykouření placebo cigarety prodlužuje časový odhad ve srovnání se skupinou tři, která nekouří nic. Vhodnými vahami jsou např. $\vec{w}_4 = (0, 0, -2, 1, 1)$.

Výpočtem skalárních součinů se lze přesvědčit o nezávislosti H_4, H_1; nezávislosti H_4, H_2; a nezávislosti H_4, H_3 – čili H_1, H_2, H_3, H_4 jsou navzájem nezávislé. Dále lze spočítat $SSH_4 = 1{,}3$, čili skutečně platí, že všechny navzájem nezávislé hypotézy vyčerpají celý $SSMT$:

$$SSH_1 + SSH_2 + SSH_3 + SSH_4 = SSMT = 137{,}6.$$

V praxi obyčejně nehledáme všechny navzájem nezávislé hypotézy, ale jen ty, které tvoří procentuelně významnou část součtu $SSMT$ – ty totiž dostatečně vysvětlují rozptyl mezi jednotlivými skupinami. V našem příkladu hypotézy H_1, H_2 popsaly situaci dostatečně, a hypotézy H_3, H_4 jsme tedy nemuseli hledat, protože vysvětluí celkem jen 5,3% variability $SSMT$, což už není mnoho (testem d) jsme ověřili, že tento zbytkový rozptyl není statisticky významný).

5.0.3 Metody vytváření vah

V příkladech z oddílu 5.0.2 byla volba vah celkem jasná. Podívejme se nyní na některé otázky, které se při volbě vah mohou vynořit. Základním postupem zůstává: zvolit nejprve taková w_j, která modelují žádaný průběh, a pak teprve vhodnou úpravou zaručit, aby platilo $\sum w_j = 0$.

lineární trend – lichý počet skupin

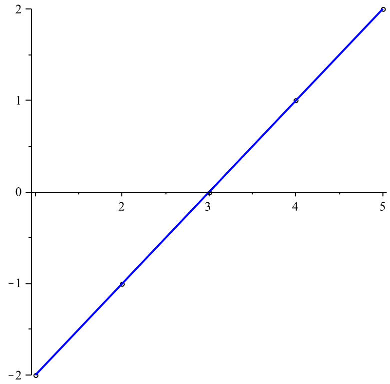
Vraťme se k příkladu 3.7 v oddílu 3.1.2 (tzv. Sternbergův experiment), kde jednou proměnnou byl počet význačných zapamatovaných slov (z jisté vybrané množiny slov) a druhou proměnnou průměrná doba rekce na otázku, zda určité slovo pochází z dané množiny nebo ne. Naše hypotéza říkala, že s rostoucím počtem slov ve význačné množině roste i doba reakce. Pokud bychom měli pět experimentálních skupin (skupina 1 … 1 slovo, skupina 2 … 2 slova, skupina 3 … 3 slova, skupina 4 … 4 slova, skupina 5 … 5 slov ve význačné množině). Vhodnými vahami by byly např. už v minulém oddílu užité $\vec{w} = (-2, -1, 0, 1, 2)$ (všimněte si, že rozdíl sousedních dvou vah je stále stejný – je roven jedné).

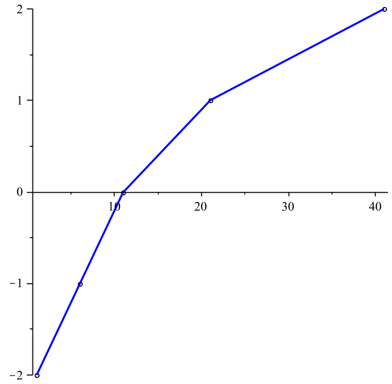
lineární trend – sudý počet skupin

Pokud bychom měli stejný experiment jako v 5.0.3 s jediným rozdílem, že počet skupin by byl sudý (skupina 1 … 1 slovo, skupina 2 … 2 slova, skupina 3 … 3 slova, skupina 4 … 4 slova ve význačné množině), nabízí se váhy $(-2, -1, 1, 2)$ – zde ovšem rozdíl dvou sousedních vah není vždy stejný, někdy je roven jedné, jindy dvěma (těmito různými sousedními rozdíly se nemodeluje lineární nárůst). Proto je lepší vzít váhy např. $-1{,}5; -0{,}5; 0{,}5; 1{,}5$. Ovšem protože práce s desetinnými čísly je poněkud nepohodlná, vynásobíme poslední vahový vektor dvěma: $\vec{w} = (-3, -1, 1, 3)$.

lineární trend – nerovnoměrný nárůst

Uvažujme stejnou hypotézu jako v 5.0.3, 5.0.3 s pěti experimentálními skupinami, ve kterých se ovšem počet význačných slov zvyšuje nerovnoměrně (skupina 1 … 1 slovo, skupina 2 … 6 slov, skupina 3 … 11 slov, skupina 4 … 21 slov, skupina 5 … 41 slov ve význačné množině). Z obrázků je patrné, že váhy $(-2, -1, 0, 1, 2)$ jsou vhodným modelem lineárního trendu v případě 5.0.3, ale ne v případě 5.0.3 (na svislou osu jsou vynášeny váhy, na vodorovnou osu jsou vynášeny počty slov ve významné množině):





Váhy $(-2,-1,0,1,2)$ nejsou v případě pravého obrázku vhodným reprezentantem pro test lineárního trendu. Nerovnoměrný nárůst dobře zachycují samotné počty slov $(1,6,11,21,41)$, zde ovšem ještě neplatí podmínka $\sum w_j = 0$. Tu zajistíme, pokud od každého z čísel odečteme jejich průměr 16: v úvahu přicházejí váhy

$$(1-16, 6-16, 11-16, 21-16, 41-16) = (-15,-10,-5,5,25).$$

Protože vektor vah je dělitelný pěti, ideálními vahami v tomto případě jsou $(-3,-2,-$ $-1,1,5)$ Pro tyto váhy už bude v případě počtu slov $(1,6,11,21,41)$ modelem přímka.

rostoucí trend – neurčené měřítko

Psychologové chtějí zjistit, jak se liší „prestižní hodnota“ některých značek aut. Náhodně vybraná skupina lidí ohodnotí každou ze značek

Ford, Volkswagen, Cadillac, Mercedes, Rolls-Royce

číslem ze stupnice 1(= tuto značku nemohu vystát), 2, 3, 4, 5, 6, 7(= tuto značku bych chtěl mít nejvíc). U každé značky se pak vypočte průměr ohodnocení. Jaké váhy lze využít pro test hypotézy

H: $\mu_F < \mu_V < \mu_C < \mu_M < \mu_R$?

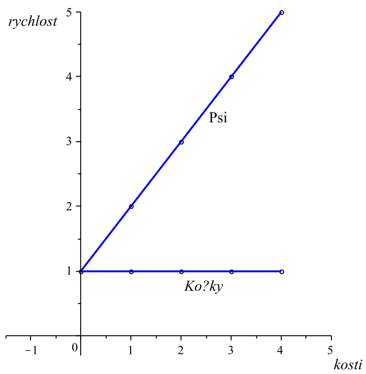
Pokud chceme testovat pouze rozdílnost (v daném pořadí) prestiže jednotlivých značek, nikoliv povahu této rozdílnosti (zda je lineární nebo jiná) – jsou lepší váhy $(-2,-1,0,1,2)$ nebo váhy $(-3,-2,-1,2,4)$?? Jak rozhodnout o vhodnosti vah?

Teorie (Green, B.F.; Tukey,J.W.: Complex analysis of variance: General problems. Psychometrika, 1960, 25, 127-152) naznačuje, že ideálními vahami jsou lineární váhy s dvojnásobnými konci $(-4,-1,0,1,4)$ (dvojnásobnost se projevuje v tom, že místo lineárního -2 a 2 je na okrajích -4 a 4).

váhy při dvourozměrné analýze rozptylu

Zajímá nás, jak se mění rychlost běhu kočky a běhu psa v závislosti na počtu kostí, které je čekají v cíli. Tento experiment lze podrobit dvourozměrné ANOVA, kde faktor 1 = druh zvířete, faktor 2 = počet kostí (0,1,2,3,4).

Máme následující hypotézu: pokud počet kostí = 0, pes i kočka poběží přibližně stejně rychle. Ovšem s rostoucím počtem kostí se rychlost psa zvyšuje, kdežto rychlost kočky zůstává stejná. Grafické znázornění naší hypotézy:



Tuto hypotézu lze testovat dvourozměrnou ANOVA typu $2x5$. Ovšem lze ji také ověřit metodou plánovaného srovnání – jak zkonstruovat v tomto případě váhy? Přiřaďme jedno číslo každé z deseti situací experimentu: Tabulka

	0 kostí	1 kost	2 kosti	3 kosti	4 kosti
psi	1	2	3	4	5
kočky	1	1	1	1	1

dobře modeluje naši hypotézu, čísla ovšem nesplňují podmínku $\sum w_j = 0$. Tuto podmínku zajistíme, když od všech čísel odečteme jejich průměr 2: hledané váhy jsou

	0 kostí	1 kost	2 kosti	3 kosti	4 kosti
psi	-1	0	1	2	3
kočky	-1	-1	-1	-1	-1

.

Nyní bychom mohli provést testování plánovaného srovnání. Předpokládejme, že v každé z deseti experimentálních skupin je pět zvířat, tj. $N = 5$. A můžeme provést vše, co se provádí v jednorozměrném plánovaném srovnání: test významnosti hypotézy na základě

$$SSH = \frac{N \cdot \left(\sum_{\text{podmínky}=1}^{10} w_{kl}\overline{x}_{kl}\right)^2}{\sum_{\text{podmínky}=1}^{10} w_{kl}^2}$$

($VH = 1$) i test kvality korelace na základě odhadu zbytkového rozptylu. V ideálním případě bude test hypotézy významný a test zbytkového rozptylu nevýznamný. Naznačme jen testovou tabulku:

zdroj rozptylu	V	$est\ R$
celkový	49	–
MT	9	–
– H	1	$est\ RH$
– Z	8	$est\ RZ$
UT	40	$est\ RUT$

Čili místo dvourozměrné ANOVA typu $2x5$ lze použít plánované srovnání pro deset podmínek.

5.0.4 Plánované srovnání v experimentech opakovaného měření

Test plánovaného srovnání se v tomto případě neliší od testů v 5.0.2. Proto jen stručně: připomeňme si tabulku testu z kapitoly 3.1.3:

zdroj rozptylu	V
MT	$JN - 1$
$- - R$	$N - 1$
$- - S$	$J - 1$
$- - I$	$(J - 1)(N - 1)$

(kritériem testu byl $\frac{est\ RMS}{est\ RI}$).

Při plánovaném srovnání nejprve určíme váhy způsobem popsaným v oddílu 5.0.3, pak vypočteme SSH analogicky oddílu 5.0.2 ($N =$ počet subjektů) a $SSZ = SSMS - SSH$. Nakonec testujeme obě hypotézy – podobně jako v kapitole 3.1.3 se ve jmenovatelích obou kritérií vyskytuje odhad rozptylu interakce RI:

zdroj rozptylu	V
MT	$JN - 1$
$- - R$	$N - 1$
$- - S$	$J - 1$
$- - - - H$	1
$- - - - Z$	$J - 2$
$- - I$	$(J - 1)(N - 1)$

.

5.0.5 Procentuální podíl celkového rozptylu

Mluvili jsme (v 5.0.2) o procentuálním podílu roztylu RMT popsaném hypotézou. Nyní několik slov o procentuálním podílu celkového rozptylu RC.

Příklad 5.4. Zajímá nás, jak moc je výsledek zkoušky ovlivněn návštěvou přednášky. Náhodně vybereme tři tisíce studentů a rozdělíme do tří skupin po tisíci studentů. Skupinu 1 požádáme, aby chodili na přednášku, ale po jedné hodině odešli domů. Skupinu

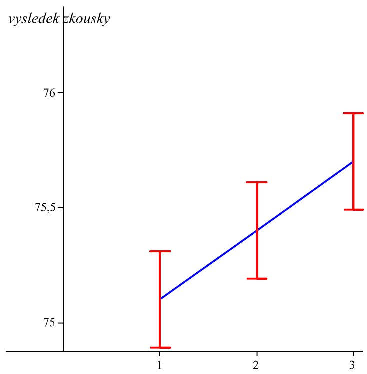
2 požádáme, aby po dvou hodinách odešli domů. A konečně skupinu 3 požádáme, aby vydržela přednášku každý týden celé tři hodiny.

Po zkoušce byly vypočteny průměry bodového hodnocení v každé ze tří skupin: $\overline{x}_1 = 75{,}1$, $\overline{x}_2 = 75{,}4$, $\overline{x}_3 = 75{,}7$. Dále $SSUT = 35964$. Proveďme jednorozměrnou ANOVA pro tyto výsledky (z průměrů lze určit, že $T_1 = 75100$, $T_2 = 75400$, $T_3 = 75700$):

zdroj rozptylu	V	SS	$est\ R$	F-hodnota	F_k
C	2999	36144	–	–	–
MT	2	180	90	7,5	$3 = F_k(2; 2997)$
UT	2997	35964	12	–	–

U testu MT vidíme, že $7{,}5 > 3$, tj. rozdíly jsou významné (i když malé). Protože počet hodnot ve skupinách měření byl velký, test měl velkou sílu a dohalil i malé rozdíly středních hodnot jednotlivých veličin. To je vidět i z grafického znázornění intervalů spolehlivosti:

$$est\ \sigma_{\overline{x}} = \sqrt{\frac{est\ RUT}{1000}} = \frac{\sqrt{12}}{\sqrt{1000}} = 0{,}11 \Rightarrow \mu_j \in \overline{x}_j \pm 1{,}96 \cdot 0{,}11 = \overline{x}_j \pm 0{,}21.$$



Z obrázku je vidět, že intervaly spolehlivosti se nepřekrývají významně, tj. rozdíly jsou významné (velká síla testu zaručila malé intervaly spolehlivosti).

Tedy vliv návštěvy přednášky na výsledek zkoušky je statisticky významný – co z toho plyne? Měli bychom seznámit studenty s výsledkem našeho experimentu, aby to podpořilo návštěvu přednášky? Asi ne, protože skupina 3 je oproti skupině 1 lepší jen o půl bodu. Vliv sice existuje, ale na celkovém výsledku zkoušky se projeví jen nepatrně (významnější vliv zde hrají jiné faktory: doba domácího studia, motivace, inteligence, atd.).

Má tedy smysl zavést jakési měřítko důležitosti vlivu nezávislé proměnné na závislou proměnnou – označme je ω^2: Veličinu ω^2 vypočteme jako podíl **rozptylu závislé proměnné popsaného (vysvětleného) nezávislou proměnnou** a **celkového rozptylu závislé proměnné**. Nepřesně řečeno, $\omega^2 = \frac{SSMT}{SSC}$, přesný vzorec je totiž

$$\omega^2 = \frac{SSMT - (J-1) \cdot est\ RUT}{SSC + est\ RUT}.$$

Důvod, proč čitatel není jen $SSMT$**, ale** $SSMT - (J-1)\cdot est\ RUT$: $SSMT$ se neskládá jen z rozptylu způsobeného různými podmínkami, ale také z přirozeného rozptylu populace (i kdyby rozptyl způsobený podmínkami byl nulový, tak $SSMT$ není roven nule). **Důvod, proč jemnovat kromě** SSC **obsahuje ještě** $est\ RUT$**, je záhadnější a nebudeme jej zde vysvětlovat**.

Vysvětleme nyní význam míry ω^2: pokud $\omega^2 = 1$, tak vztah mezi oběma proměnnými je dokonale významný:

$$\begin{aligned}
SSMT - (J-1)\cdot est\ RUT &= SSC + est\ RUT \\
SSMT - (J-1)\cdot \frac{SSUT}{N-J} &= SSMT + SSUT + \frac{SSUT}{N-J} \\
SSUT(1 + \frac{1}{N-J} + \frac{J-1}{N-J}) &= \\
SSUT &= 0
\end{aligned}$$

(celý rozptyl závislé proměnné je popsán rozptylem nezávislé proměnné)

Naopak pokud $\omega^2 = 0$, tak vztah mezi oběma proměnnými je naprosto bezvýznamný:

$$\begin{aligned}
SSMT &= (J-1)\cdot est\ RUT \\
\frac{SSMT}{J-1} &= est\ RUT \\
est\ RMT &= est\ RUT
\end{aligned}$$

(rozdíly mezi podmínkami jsou způsobeny pouze populačním rozptylem).

V našem příkladu $\omega^2 = 0{,}004$, což je hodně málo. Vztah závislosti mezi chozením na přednášky a výsledkem zkoušky sice existuje, ale jeho význam je malý. Méně než $\frac{2}{10}\%$ celkového rozptylu výsledné zkoušky je způsobeno návštěvností přednášek.

Vzorce pro ω^2 lze odvodit i u dvourozměrné ANOVA: (faktor 1 … J úrovní (řádků), faktor 2 … K úrovní (sloupců)):

$$\begin{aligned}
\omega^2(faktor1) &= \frac{SSMR - (J-1)\cdot est\ RUT}{SSC + est\ RUT}, \\
\omega^2(faktor2) &= \frac{SSMS - (K-1)\cdot est\ RUT}{SSC + est\ RUT}, \\
\omega^2(faktor3) &= \frac{SSI - (J-1)\cdot(K-1)\cdot est\ RUT}{SSC + est\ RUT},
\end{aligned}$$

ω^2 vždy vyjadřuje, kolik procent celkového rozptylu celého experimentu je obsaženo ve faktoru 1, faktoru 2 nebo v interakci těchto faktorů. Podobně i u experimentu opakovaného měření lze vypočíst ω^2 určující, kolik procent celkového podílu experimentu je způsobeno vlivem různých podmínek.

ω^2 Podobně jako Pearsonův koeficient r^2 vyjadřuje jistou míru vhodnosti. V korelační analýze r^2 představuje, kolik procent rozptylu proměnné Y je popsáno lineárním vztahem s jinou proměnnou X. Podobně ω^2 udává, kolik procent celkového rozptylu závislé proměnné je popsáno jakýmkoliv vztahem této závislé proměnné s některou nezávislou proměnnou.

5.0.6 Tři míry závažnosti experimentu

Tou první je hladina významnosti (tou jsme se v této přednášce zejména zabývali), která vyjadřuje naši jistotu, že existuje vztah mezi nezávislou a závislou proměnnou. Druhou mírou je ω^2, která vyjadřuje závažnost vztahu (jak velký je vliv nezávislé proměnné na závislou). A tou třetí mírou je procentuální podíl rozptylu RMT popsaný konkrétní hypotézou v experimentu plánovaného srovnání, který vyjadřuje jak platnost, tak závažnost sledované hypotézy.

Zjistit testem, že mezi dvěma proměnnými existuje vztah, ještě nic neříká o závažnosti tohoto vztahu, jak jsme právě viděli z příkladu 5.4. Čili vysvětlení rozptylu je mnohem důležitější než test významnosti – zde se naskýtá otázka, zda je důležitější ω^2 nebo prcentuální podíl při plánovaném srovnání. Odpověď zní: jak kdy. Pokud se zajímáme zejména o praktický dopad experimentu, nejdůležitější je ω^2 (např. v příkladu 5.4 je $\omega^2 = 0{,}004$, a tedy studenti udělají lépe, když věnují čas přednášky raději osobnímu studiu). Pokud náš experiment má prokázat platnost určité teorie, výsledky plánovaného srovnání jsou docela důležité (plánované srovnání studující vliv návštěvnosti přednášek na výsledky studenta by mělo teoretický význam).

Pojmy k zapamatování

- Tato důležitá kapitola je volným pokračováním kapitoly 3, jen využívá některé drobnosti probrané v kapitole 4 (tím se nechce říci, že kapitola 4 je jen přípravnou kapitolou – naopak, Pearsonův r^2 koeficient je důležitým pojmem v dosavadním toku textu).
- Důležitou otázkou této kapitoly je, co dělat, když F-test prokáže, že hypotéza H_0 neplatí. Jednou z možných cest je porovnávat každo dvojici skupin měření ve zvláštním statistickém testu, což je sice možné (oddíl 5.0.1 – metoda minimálního významného rozdílu, Scheffé-ova metoda), ale trochu těžkopádné. Další alternativní možností je tzv. plánované srovnání (z anglického „planned comparison“ – počínaje oddílem 5.0.2), které skýtá mnohem lepší aparát – nejen že test plánovaného srovnání potvrzuje, že střední hodnoty veličin v jednotlivých skupinách měření jsou různé, ale také přímo testuje hypotézu, jakým konkrétním způsobem jsou tyto střední hodnoty různé. Ba co víc, také přímo při provádění testu vidíme z veličin typu SS (veličiny udávající součty čtverců hodnot), jaký procentuelní podíl rozptýlenosti daných veličin v různých skupinách měření je testovanou hypotézou popsán = vysvětlen.
- Tyto konkrétní rozdíly mezi testovanými středními hodnotami veličin vyjadřujeme v hypotéze tzv. vahami. Celý metodický přístup plánovaného srovnání lze užít nejen v experimentech typu „více skupin jednou“ (oddíly 5.0.2, 5.0.3), ale také ve dvourozměrné ANOVA (oddíl 5.0.3) a v experimentu opakovaného měření (oddíl 5.0.4).
- Důležitost posledních dvou oddílů 5.0.5, 5.0.6 také nelze přecenit, protože zde ilustrují

výpovědní sílu testů plánovaného srovnání nebo význam procentuálního podílu celkového rozptylu ve srovnání s hladinou významnosti statistických testů, která byla užívána ve většině kapitol tohoto textu. Řekl bych že oddíly 5.0.5, 5.0.6 jsou nejdůležitějším vyvrcholením celé knihy.

– Budiž ještě na připomenutí řečeno, že t-testy v oddílu 5.0.1 nebo testy plánovaného srovnání ve zbytku této kapitoly vycházejí všechny z předpokladu normality veličin v každé ze skupin měření – viz též oddíl 2.5.6. V následujících dvou kapitolách tohoto materiálu totiž nastává zlom a budeme se právě zabývat těmi statistickými testy, které předpoklad normality veličin nevyžadují, respektive nesplňují.

Kontrolní otázky

1. U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.
   a) Testování post-hoc znamená, že stanovíme nejprve hypotézy H_0, H_1, a „potom“ provedeme řešení
   b) Když jsou veličiny u Schefféova testu navzájem závislé, celková hladina významnosti u provedených $\binom{J}{2}$ testů se určí jako součet jednotlivých hladin významnosti.
   c) Schefféova metoda je vlastně neco podobného jako metoda minimálního významného rozdílu s tím, že sníží hladinu významnosti celku všech prováděných t-testů pod určitou hodnotu (např. pod hodnotu 0,05).
   d) Plánované srovnání lze užít pouze v experimentech typu „více skupin jednou“, nikoli však v experimentech opakovaného měření.
   e) U testu plánovaného srovnání jsou hypotézy H_0, H_1 vyjádřeny pomocí korelace skupin měření s vektorem tzv. vah.
   f) Pokud vektory vah jsou navzájem nezávislé, odpovídají hypotézám vysvětlujícím různé části celkového rozptylu.
   g) Při testu plánovaného srovnání existuje tolik ruzných navzájem nezávislých vektorů vah, kolik je skupin měření.
   h) Výpovědní hodnota ω^2 je důležitější než hladina význmanosti testu α.

Odpovědi na otázky
1a) – N (charakteristickým znalem post-hoc testů je to, že vycházejí pouze z měření, tj. jsou provedeny "post-hoc" = "po tom" = "po měření"), 1b) – N (u závislých veličin celkovou hladinu významnosti určit vlastně ani neumíme – ale určitě se nezíská jako součet jednotlivých hladin t-testů), 1c) – A, 1d) – N, 1e) – A, 1f) – A, 1g) – N (správná odpověď je: počet skupin měření snížený o jedničku), 1h) – A.

Cvičení

1. Experiment má ukázat, že hnojivo zvyšuje vzrůst trifidů. Každá ze čtyř skupin trifidů po deseti kusech je zasazena do půdy obsahující jiné množství hnojiva. Po roce je změřena výška jednotlivých trifidů (v cm) a vypočten průměr každé skupiny ($N_i = 10$ pro $i = 1, 2, 3, 4$):

(skup. 1: 0 gramů)	(skup. 2: 4 gramy)	(skup. 3: 8 gramů)	(skup. 4: 12 gramů)
$\overline{x}_1 = 2$	$\overline{x}_2 = 5$	$\overline{x}_3 = 5$	$\overline{x}_4 = 8$

Bylo spočteno $SSUT = 360$. Testujte hypotézu, že

a) Průměry vykazují lineární nárůst s nárůstem hnojiva. Kolik procent celkového rozptylu je popsáno touto hypotézou?

b) Testujte hypotézu zbytkového rozptylu.

2. Uvažujme data z příkladu 7.

a) Testujte hypotézu, že tréma člověka roste lineárně s růstem velikosti publika. Jak velká část součtu čtverců $SSMT$ je zachycena touto hypotézou?

b) Proveďte test zbytkového rozptylu příslušející hypotéze a).

3. Je prováděn experiment se čtyřmi skupinami po deseti jedincích. Jsou získány průměry $\overline{x}_1 = 3$, $\overline{x}_2 = 2$, $\overline{x}_3 = 5$, $\overline{x}_4 = 10$ a spočten $SSUT = 120$.

a) Nakreslete průměry s 95%-ními intervaly spolehlivosti.

b) Jaký je minimální významný rozdíl dvou průměrů pro hladinu významnosti 0,01?

c) Jaký je významný rozdíl dvou průměrů při Scheffeově metodě pro hladinu význmanosti 0,01?

d) Testujte hypotézu, že nárůst středních hodnot veličin je lineární. Jaké procento $SSMT$ je vysvětleno touto hypotézou? Je zbytkový rozptyl významný?

4. Provádíme experiment typu „vÍce skupin jednou“ se třemi podmínkami (třemi hodnotami nezávislé proměnné), v každé ze skupin jsou čtyři pozorování zapsaná do tabulky:

podmínka 1	podmínka 2	podmínka 3
2	4	1
8	10	8
14	6	15
10	12	9

a) Nakreslete body, průměry i 95%-ní intervaly spolehlivosti. Existuje zřetelný rozdíl mezi skupinami měření?

b) Proveďte jedorozměrnou ANOVA. Prokázal se rozdíl mezi skupinami?

c) Zvolte váhy pro následující hypotézu: podmínka 2 je lepší než obě další podmínky.

d) Určete SSH, SSZ (vzhledem k c)).

e) Jaké procento $SSMT$ je popsáno hypotézou c)?

f) Je vzhledem k c) podíl zbytkového rozptylu statisticky významný?

5. Uvažujme situaci z příkladu 8.

a) Testujte hypotézu H_1, že obě uvedené metody likvidace moskytů jsou významné ve srovnání s tím, když se neděje nic.

b) Testujte hypotézu H_2 (ortogonální k H_1), že nová metoda je účinnější než DDT.

6. Sociolog zkoumá hypotézu, že průměrná výška mužů v České Republice je větší než průměrná výška mužů ve Slovenské Republice. Získá náhodný výběr 10000 čechů a 9500 slováků a spočte následujíc data (aby se mu dobře počítalo a výška byla mezinárodně srozumitelná, měří výšku ve stopách – při zaokrouhlení na čtyři desetinná místa je jedna stopa dlouhá 0,3048 m):

   ČR ... $\overline{x}_1 = 5{,}75$; $\sum(x_i - \overline{x}_1)^2 = 599$;

   SR ... $\overline{x}_2 = 5{,}74$; $\sum(x_i - \overline{x}_2)^2 = 632$;

   Testujte hypotézu, že tyto průměrné výšky se liší – jaká je hodnota veličiny ω^2 (omega na druhou) pro tento test?

Výsledky

ad 1. **ad a)** H_0: Nárůst vlivu hnojiva není lineární, tj. není v korelaci s vahami $w = (-3; -1, 1, 3)$.

H_1: Nárůst vlivu hnojiva je v korelaci s těmito vahami.

Kritériem je $\frac{est\ RH}{est\ RUT}$. Při platnosti H_0 má kritérium rozdělení $F(1; 36)$. Pro $\alpha = 0{,}05$ máme kritickou hodnotu $F_k(1; 36) \doteq 4{,}1$. Pak

$$\frac{est\ RH}{est\ RUT} = \frac{\frac{SSH}{VH}}{\frac{SSUT}{VUT}} = \frac{162}{10} = 16{,}2 > F_k(1; 36).$$

H_0 zamítáme, korelace s hypotézou je významná.

Procento $SSMT$ popsané hypotézou: $SSMT = 180$, tj. $\frac{SSH}{SSMT} = \frac{162}{180} = 0{,}9$, tj. 90% celkového rozptylu je popsáno hypotézou.

ad b) Test dokonalé korelace:

H_0**:** $\overline{x}_j$, w_j jsou dokonale korelovány, tj. $r^2 = 1$.

H_1**:** neplatí H_0.

kritérium: $\frac{est\ RZ}{est\ RUT}$ – tato funkce má za předpokladu platnosti H_0 rozdělení $F(2; 36)$. Pro $\alpha = 0{,}05$ je kritickou hodnotou $F_k(2; 36) \doteq 3{,}3$.

Zpracování měření:

$$\frac{est\ RZ}{est\ RUT} = \frac{\frac{18}{2}}{10} = 0{,}9 < 3{,}3 = F_k,$$

a tedy H_0 nezamítáme, zbytek rozptylu není významný.

ad 2. ad a)

H_0**:** Průměry $(2; 5; 5{,}25)$ nejsou v korelaci se zlinearizavanými vahami $(-6; -2; 8)$.

H_1**:** Jsou.

Kritérium: $\frac{est\ RH}{est\ RI}$; toto kritérium má za předpokladu platnosti H_0 rozdělení $F(1; 6)$, čili pro $\alpha = 0{,}05$ je kritická hodnota $F_k(1; 6) = 5{,}99$.

Zpracování měření:

$$est\ RH = \frac{SSH}{VH} = \frac{N \cdot (\sum w_j \cdot \overline{x}_j)^2}{1 \cdot \sum w_j^2} = 22{,}15.$$

$$\frac{est\ RH}{est\ RI} = \frac{22{,}15}{0{,}63889} > F_k;$$

H_0 zamítáme, korelace je významná.

Dále pro procento popsaného $SSMT$ platí:

$$\frac{SSH}{SSMT} = \frac{22{,}15}{44{,}25} \doteq 0{,}5;$$

tj. asi 50% součtu čtverců je popsáno hypotézou H_1.

ad b)

H_0**:** Korelace v části a) je perfektní ($r^2 = 1$): $RZ = RI$.

H_1**:** Neplatí H_0, tj. $RZ > RI$

Kritérium: $\frac{est\ RZ}{est\ RI}$ má při předpokladu platnosti H_0 rozdělení $F(1;6)$, a pro $\alpha = 0{,}05$ je $F_k = 5{,}99$.

Zpracování meření:

$$\frac{est\ RZ}{est\ RI} = \frac{44{,}25 - 22{,}15}{0{,}63889} \doteq 3{,}45 < F_k;$$

tj. H_0 nezamítáme, ikdyž zbývá ještě padepát procent nepopsaného rozptylu, tenst neprokázal významnost zbytku (díky „velké volnosti“, respektive přesněji řečeno, nízkému počtu stupňů volnosti)

ad 3. Ad a) $\overline{x}_j \pm 1{,}172$; Ad b) 2,245; Ad c) 3,002; Ad d) $FH(1;36) = 86{,}4$ je statisticky významné, 75,8% $SSMT$ je vysvětleno hypotézou, $FZ(2;36) = 13{,}8$ je významný zbytek.

ad 4. Ad a) $\overline{x}_j \pm 5{,}95$; Ad b) $F = 0{,}22$, což není statisticky významné; Ad c) $(-1;2;-1)$ Ad d) $SSH = 12{,}04$, $SSZ = 0{,}13$ Ad e) je popsáno 98,9% součtu $SSMT$ Ad f) $F(1;9) = 0{,}436$, což není významné.

ad 5. Ad a) $F(1;6) = 7{,}89$ … významná; Ad b) $F(1;6) = 8{,}3$ … významná.

ad 6. Řešení:

typ rozptylu	V	SS	$est\ R$	F-hodnota	F_k ($\alpha = 0{,}05$)
celkový	19499	1231,487179	–	–	–
MT	1	0,487179	0,487179	7,72	$3{,}84 = F_k(1;\infty)$
UT	19498	1231	0,06313428	–	–

Z tabulky plyne, že $F(1;19498) = 7{,}72$, což je hodnota statisticky významná – ale hodnota $\omega^2 = 0{,}000344$ je velmi malá. Závislost mezi veličinami existuje, ale procentuelně nepůsobí velké rozdíly.

6 Rozdělení „chí kvadrát“

Průvodce studiem

Cílem této kapitoly je představit další užitečné rozdělení pravděpodobnosti, a sice rozdělení χ^2. Toto rozdělení, o n stupních volnosti, se definuje jako součet čtverců n normovaných normálních rozdělení U^2. Přitom mnohé veličiny v praxi jsou rozděleny jako χ^2, a tedy toto rozdělení se vyskytuje v mnoha statistických testech.

Cíle

Po prostudování této kapitoly budete schopni:

- Pracovat s rozdělením χ^2, při daných stupních volnosti.
- Použít rozdělení χ^2 ve vhodných statistických testech. lze užít v testu typu $\sigma^2 = konst.$ Pokud měření jsme získali
- Pracovat s tabulkami kritických hodnot rozdělení χ^2.

6.1 Vlastnosti rozdělení χ^2

Uvažujme veličinu X s normálním rozdělením pravděpodobnosti $No(\mu, \sigma^2)$. Víme, že veličina $U = \frac{X-\mu}{\sigma}$ má normované normální rozdělení $No(0; 1)$. Jaké rozdělení má veličina

$$U^2 = \frac{(X-\mu)^2}{\sigma^2}?$$

Především můžeme říci, že hodnoty veličiny U^2 jsou nezáporné. Dále víme, že asi 68% veličiny U leží v intervalu $\langle -1; 1\rangle$, tj. asi 68% veličiny U^2 leží v intervalu $\langle 0; 1\rangle$. Zkrátka a dobře, vlastnosti veličiny U^2 lze odvodit z vlastností veličiny U. A protože se veličina U^2 ve statistice hojně používá, dostala i své jméno: $\chi^2(1)$ … čti: chí kvadrát o jednom stupni volnosti.

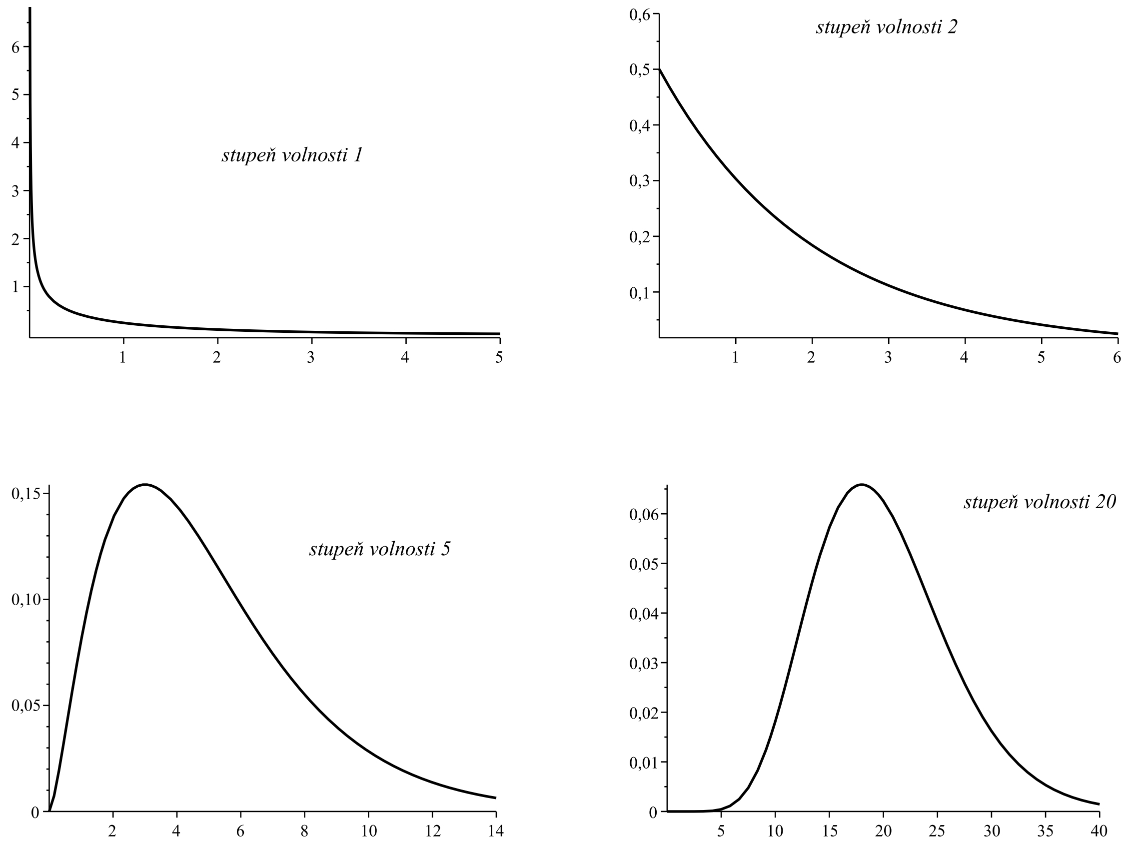
Označení $\chi^2(1)$ naznačuje, že může nastat i více stupňů volnosti než jeden. A skutečně, pokud dvě hodnoty U_1, U_2 veličiny U umocníme na druhou mocninu a sečteme, dostaneme obecně větší hodnotu než $\chi^2(1)$, označujeme ji $\chi^2(2)$ (veličina „chí kvadrát“ se dvěma stupni volnosti):

$$\chi^2(2) = U_1^2 + U_2^2.$$

Čili střední hodnota veličiny $\chi^2(2)$ je větší než střední hodnota veličiny $\chi^2(1)$. Obecně lze pak vyjádřit veličinu o n stupních volnosti:

$$\chi^2(n) = U_1^2 + U_2^2 + \cdots + U_n^2.$$

Příklady χ^2 o různých stupních volnosti jsou na obrázku:



Obecný vzorec hustoty rozdělení $\chi^2(n)$ zní

$$f(x) = \begin{cases} 0 & \ldots \quad x \leq 0; \\ \frac{1}{2^{\frac{n}{2}} \cdot \Gamma\left(\frac{n}{2}\right)} \cdot x^{\left(\frac{n}{2}-1\right)} \cdot e^{-\frac{x}{2}} & \ldots \quad x > 0, \end{cases}$$

kde Γ je tzv. gama-funkce. Co se týká střední hodnoty a rozptylu rozdělení χ^2, vypočteme jen střední hodnotu, rozptyl uvedeme bez důkazu:

$$E\left\{\chi^2(n)\right\} = E(U_1^2 + U_2^2 + \cdots + U_n^2) = 1 + 1 + \cdots + 1 = n.$$

Opravdu platí $E(U_i^2) = 1$:

$$E\left(\frac{X-\mu}{\sigma}\right)^2 = \frac{1}{\sigma^2}\cdot E(X-\mu)^2 = \frac{1}{\sigma^2}(EX^2 - 2\mu EX + \mu^2) = $$
$$= \frac{1}{\sigma^2}(EX^2 - 2\mu^2 + \mu^2) = \frac{1}{\sigma^2}(\sigma^2 + \mu^2 - \mu^2) = 1.$$

Pro výpočet EX^2 jsme využili faktu, že $\sigma^2 = DX = EX^2 - E^2X = EX^2 - \mu^2$, tj. $EX^2 = \sigma^2 + \mu^2$.

$$D\left\{\chi^2(n)\right\} = 2n.$$

Další důležitou vlastností je tzv. aditivita rozdělení χ^2, že totiž

$$\chi^2(n_1) + \chi^2(n_2) = \chi^2(n_1 + n_2),$$

která v podstatě plyne z toho, že $\chi^2(n) = \sum_1^n U_i^2$.

Podobně jako u jiných rozdělení, i zde byly kritické hodnoty rozdělení spočteny jednou provždy a seřazeny do tabulky, takže při konkrétním užívání kritických hodnot nemusíme počítat žádné nechutné integrály (kritické hodnoty viz tabulky 6.1, 6.2).

Například pokud chceme určit kritickou hodnotu h_k tak, že

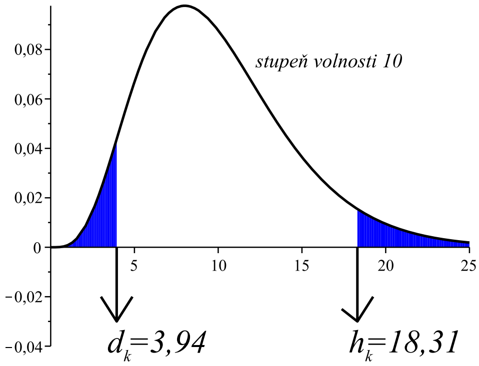
$$P(\chi^2(10) \geq h_k) = 0{,}05,$$

tak ve druhé části tabulky najdeme hodnotu 18,307 na průsečíku řádku 10 a sloupce pro 0,05.

Pokud hledáme d_k tak, aby

$$P(\chi^2(10) \leq d_k) = 0{,}05,$$

tak v první části tabulky na průsečíku řádku 10 a sloupce 0,95 (protože v tabulce jsou kritické hodnoty uvedené pro pravou část podgrafu) najdeme hodnotu 3,94 (viz též obrázek):



Tab. 6.1: Kritické hodnoty jednostranného testu χ^2 – část 1.

$q \rightarrow$ $\nu \downarrow$	0,995	0,990	0,975	0,950	0,900	0,750	0,500
1	$392{,}704 \cdot 10^{-10}$	$157{,}088 \cdot 10^{-9}$	$982{,}069 \cdot 10^{-9}$	$393{,}214 \cdot 10^{-8}$	0,0157908	0,1015308	0,454937
2	0,0100251	0,0201007	0,0506356	0,102587	0,210720	0,575364	1,38629
3	0,0717212	0,114832	0,215795	0,351846	0,584375	1,212534	2,36597
4	0,206990	0,297110	0,484419	0,710721	1,063623	1,92255	3,35670
5	0,411740	0,554300	0,831211	1,145476	1,61031	2,67460	4.35146
6	0,675727	0,872085	1,237347	1,63539	2,20413	3,45460	5,34812
7	0,989265	1,239043	1,68987	2,16735	2,83311	4,25485	6,34581
8	1,344419	1,646482	2,17973	2,73264	3,48954	5,07064	7,34412
9	1,734926	2,087912	2,70039	3,32511	4,16816	5,89883	8,34283
10	2,15585	2,55821	3,24697	3,94030	4,86518	6,73720	9,34182
11	2,60321	3,05347	3,81575	4,57481	5,57779	7,58412	10,3410
12	3,07382	3,57056	4,40379	5,22603	6,30380	8,43842	11,3403
13	3,56503	4,10691	5,00874	5,89186	7,04150	9,29906	12,3398
14	4,07468	4,66043	5,62872	6,57063	7,78953	10,1653	13,3393
15	4,60094	5,22935	6,26214	7,26094	8,54675	11,0365	14,3389
16	5,14224	5,81221	6,90766	7,96164	9,31223	11,9122	15,3385
17	5,69724	6,40776	7,56418	8,67176	10,0852	12,7919	16,3381
18	6,26481	7,01491	8,23075	9,39046	10,8649	13,6753	17,3379
19	6,84398	7,63273	8,90655	10,1170	11,6509	14,5620	18,3376
20	7,43386	8,26040	9,59083	10,8508	12,4426	15,4518	19,3374
21	8,03366	8,89720	10,28293	11,5913	13,2396	16,3444	20,3372
22	8,64272	9,54249	10,9823	12,3380	14,0415	17,2396	21,3370
23	9,26042	10,19567	11,6885	13,0905	14,8479	18,1373	22,3369
24	9,88623	10,8564	12,4011	13,8484	15,6587	19,0372	23,3367
25	10,5197	11,5240	13,1197	14,6114	16,4734	19,9393	24,3366
26	11,1603	12,1981	13,8439	15,3791	17,2919	20,8434	25,3364
27	11,8076	12,8786	14,5733	16,1513	18,1138	21,7494	26,3363
28	12,4613	13,5648	15,3079	16,9279	18,9392	22,6572	27,3363
29	13,1211	14,2565	16,0471	17,7083	19,7677	23,5666	28,3362
30	13,7867	14,9535	16,7908	18,4926	20,5992	24,4776	29,3360
40	20,7065	22,1643	24,4331	26,5093	29,0505	33,6603	39,3354
50	27,9907	29,7067	32,3574	34,7642	37,6886	42,9421	49,3349
60	35,5346	37,4848	40,4817	43,1879	46,4589	52,2938	59,3347
70	43,2752	45,4418	48,7576	51,7393	55,3290	61,6983	69,3344
80	51,1720	53,5400	57,1532	60,3915	64,2778	71,1445	79,3343
90	59,1963	61,7541	65,6466	69,1260	73,2912	80,6247	89,3342
100	67,3276	70,0648	74,2219	77,9295	82,3581	90,1332	99,3341

Tab. 6.2: Kritické hodnoty jednostranného testu χ^2 – část 2.

$q \rightarrow$ $\nu \downarrow$	0,250	0,100	0,050	0,025	0,010	0,005	0,001
1	1,32330	2,70554	3,84146	5,02389	6,63490	7,87944	10,828
2	2,77259	4,60517	5,99147	7,37776	9,21034	10,5966	13,816
3	4,10835	6,25139	7,81473	9,34840	11,3449	12,8381	16,266
4	5,38527	7,77944	9,48773	11,1433	13,2767	14,8602	18,467
5	6,62568	9,23635	11,0705	12,8325	15,0863	16,7496	20,515
6	7,84080	10,6446	12,5916	14,4494	16,8119	18,5476	22,458
7	9,03715	12,0170	14,0671	16,0128	18,4753	20,2777	24,322
8	10,2188	13,3616	15,5073	17,5346	20,0902	21,9550	26,125
9	11,3887	14,6837	16,9190	19,0228	21,6660	23,5893	27,877
10	12,5489	15,9871	18,3070	20,4831	23,2093	25,1882	29,588
11	13,7007	17,2750	19,6751	21,9200	24,7250	26,7569	31,264
12	14,8454	18,5494	21,0261	23,3367	26,2170	28,2995	32,909
13	15,9839	19,8119	22,3621	24,7356	27,6883	29,8194	34,528
14	17,1170	21,0642	23,6848	26,1190	29,1413	31,3193	36,123
15	18,2451	22,3072	24,9958	27,4884	30,5779	32,8013	37,697
16	19,3688	23,5418	26,2962	28,8454	31,9999	34,2672	39,252
17	20,4887	24,7690	27,5871	30,1910	33,4087	35,7185	40,790
18	21,6049	25,9894	28,8693	31,5264	34,8053	37,1564	42,312
19	22,7178	27,2036	30,1435	32,8523	36,1908	38,5822	43,820
20	23,8277	28,4128	31,4104	34,1696	37,5662	39,9968	45,315
21	24,9348	29,6151	32,6705	35,4789	38,9321	41,4010	46,797
22	26,0393	30,8133	33,9244	36,7807	40,2894	42,7956	48,268
23	27,1413	32,0069	35,1725	38,0757	41,6384	44,1813	49,728
24	28,2412	33,1963	36,4151	39,3641	42,9798	45,5585	51,179
25	29,3389	34,3816	37,6525	40,6465	44,3141	46,9278	52,620
26	30,4345	35,5631	38,8852	41,9232	45,6417	48,2899	54,052
27	31,5284	36,7412	40,1133	43,1944	46,9630	49,6449	55,476
28	32,6205	37,9159	41,3372	44,4607	48,2782	50,9933	56,892
29	33,7109	39,0875	42,5569	45,7222	49,5879	52,3356	58,302
30	34,7998	40,2560	43,7729	46,9792	50,8922	53,6720	59,703
40	45,6160	51,8050	55,7585	59,3417	63,6907	66,7659	73,402
50	56,3336	63,1671	67,5048	71,4202	76,1539	79,4900	86,661
60	66,9814	74,3970	79,0819	83,2976	88,3794	91,9517	99,607
70	77,5766	85,5271	90,5312	95,0231	100,425	104,215	112,317
80	88,1303	96,5782	101,879	106,629	112,329	116,321	124,839
90	98,6499	107,565	113,145	118,136	124,116	128,299	137,208
100	109,141	118,498	124,342	129,561	135,807	140,169	149,449

6.2 Využití rozdělení χ^2

6.2.1 Testování hypotézy $\sigma^2 = konst$

Příklad 6.1. Zajímá nás, zda jistý výukový program je schopen naučit jisté partie aritmetiky při výuce na ZŠ. Může se totiž stát, že nadaní studenti budou chápat instrukce počítače, kdežto průměrní žáci budou mít potíže s programem komunikovat a naučí se méně, než by pochopili z výkladu živého učitele.

Je známo, že rozptyl výsledků testu z aritmetiky při klasické výuce je $\sigma_0^2 = 25$. Pokud by naše obava byla oprávněná, rozptyl výsledků znalostí by byl při použití programu větší (nadaní žáci by byli lepší, průměrní žáci horší než obvykle).

Byl proveden experiment, kdy deset žáků se podrobilo programové výuce. Výsledky znalostí (bodové ohodnocení) byly: 68, 90, 70, 91, 72, 80, 85, 82, 91, 95. Odtud

$$\sum_1^{10} (x_i - \overline{x})^2 = 846{,}4$$

a odhad $est\ \sigma^2 = 94{,}04$ populačního roptylu je mnohem větší než $\sigma_0^2 = 25$. Proveďte statistický test, který by potvrdil, že ke zvýšení rozptylu nedošlo náhodou.

K1: H_0: rozptyl výsledků znalostí σ^2 nabytých počítačovu výukou je stejný jako rozptyl $\sigma_0^2 = 25$ při klasické výuce (tj. $\sigma^2 = 25$).

H_1: $\sigma^2 > 25$.

K2: Ukazuje se (uvidíme v bodě K3), že vhodným kritériem testu je $\frac{1}{\sigma_0^2} \cdot \sum_1^{10} (x_i - \overline{x})^2$.

K3: Za předpokladu platnosti H_0 má výraz $\frac{1}{\sigma_0^2} \cdot \sum_1^{10} (x_i - \overline{x})^2$ rozdělení $\chi^2(9)$.

Skutečně, dokažme tento fakt.

Víme, že $\frac{1}{\sigma_0^2} \cdot \sum_1^{n} (x_i - \mu)^2$ má rozdělení $\chi^2(n)$, protože se jedná o součet n čtverců rozdělení U. Ovšem μ neznáme – jaký je tedy vztah mezi $\frac{1}{\sigma_0^2} \cdot \sum_1^{n} (x_i - \mu)^2$ a $\frac{1}{\sigma_0^2} \cdot \sum_1^{10} (x_i - \overline{x})^2$?

Platí

$$\sum (x_i - \mu)^2 = \sum (x_i - \overline{x} + \overline{x} - \mu) = \sum (x_i - \mu)^2 + 2 \cdot \sum (x_i - \overline{x}) \cdot (\overline{x} - \mu) + \sum (\overline{x} - \mu)^2.$$

A protože

$$2 \cdot \sum (x_i - \overline{x}) \cdot (\overline{x} - \mu) = 2(\overline{x} - \mu) \cdot \underbrace{\sum (x_i - \overline{x})}_{=0} = 0,$$

dostaneme

$$\sum (x_i - \overline{x})^2 = \sum (x_i - \overline{x})^2 + \sum (\overline{x} - \mu)^2.$$

Máme tedy rovnici

$$\sum (x_i - \overline{x})^2 = \sum (x_i - \overline{x})^2 + n \cdot (\overline{x} - \mu)^2,$$

kterou vynásobením $\frac{1}{\sigma_0^2}$ lze převést na tvar

$$\underbrace{\frac{\sum (x_i - \mu)^2}{\sigma_0^2}}_{=\chi^2(n)} = \frac{\sum (x_i - \overline{x})}{\sigma_0^2} + \underbrace{\frac{(\overline{x} - \mu)^2}{\frac{\sigma_0^2}{n}}}_{=\chi^2(1)}.$$

To, že člen na levé straně má rozdělení $\chi^2(n)$, už bylo řečeno, člen na pravé straně má rozdělení $\chi^2(1)$, protože se jedná o druhou mocninu normovaného rozdělení průměru $\overline{x}$ se střední hodnotou μ a rozptylem $\frac{\sigma_0^2}{n}$. Dohromady tedy dostáváme to, co jsme chtěli spočítat a dokázat:

$$\frac{x_i - \overline{x}}{\sigma_0^2} = \chi^2(n) - \chi^2(1) = \chi^2(n-1).$$

K4: Pro $\alpha = 0{,}05$ je $\chi_k^2(9) = 16{,}919$ (na průsečíku řádku 9 a sloupce 0,05).

K5: Po dosazení naměřených hodnot

$$\frac{x_i - \overline{x}}{\sigma_0^2} = \frac{846{,}4}{25} = 33{,}86 > 16{,}919 \Rightarrow \text{ zamítáme } H_0,$$

rozptyl je (bohužel) významně větší než 25. A to je tragédie počítačové výuky :-)

V případě oboustranného testu – kdybychom neměli teoretický podklad, že rozptyl poroste, ale bylo by stejně možné, že třeba i klesne – bychom našli dvě kritické hodnoty na řádku 9, a sice $\chi_m = 2{,}7$ (ve sloupci 0,975) a $\chi_v = 19{,}02$ (ve sloupci 0,025) a H_0 bychom zamítli na hladině významnosti $\alpha = 0{,}05$, pokud by hodnota kritéria ležela mimo interval $(2{,}7; 19{,}02)$.

6.2.2 Test druhu rozdělení

Příklad 6.2. Rodičům se narodily už čtyři dcery, a přesto by si přáli i syna. Chystají se mít páté dítě s nadějí, že pravděpodobnost narození chlapce je 0,5 (tj. že rození dětí má podobný charakter – co se týká pohlaví dítěte – jako házení korunou). Ale přece jen si vyhledali data o 1024 rodinách s pěti dětmi a zjistili, v kolika rodinách se narodilo kolik chlapců:

0 chlapců	...	40 rodin
1 kluk	...	184 rodin
2 kluci	...	300 rodin
3 kluci	...	268 rodin
4 kluci	...	196 rodin
5 kluků	...	36 rodin

Pokud tato data o pohlaví při rození dětí mají stejný charakter jako házení korunou, pak je lze dobře popsat binomickým rozdělením s parametry $N = 1024$, $p = 0{,}5$. Teoretické rozdělení pravděpodobnosti a rozdělení četnosti mají následující průběh:

počet chlapců x_i	p_i	četnost f_i
0	1/32	32
1	5/32	160
2	10/32	320
3	10/32	320
4	5/32	160
5	1/32	32

Otázka zní: do jaké míry se shoduje empirické rozdělení četnosti a teoretické rozdělení četnosti, čili: lze nashromážděná data dobře popsat binomickým rozdělením s uvedenými parametry?

Provedeme tzv. **test dobré shody** (v angličtině: good fit test … GFT).

K1: H_0: počet chlapců z pěti narozených dětí lze dobře popsat binomickým rozdělením pro $p = 0{,}5$.

H_1: Nelze.

K2: Jaké kritérium zvolit? Vezmeme součet čtverců rozdílů normalizovaných odchylek naměřené a teoretické četnosti $\sum_1^k \frac{(f_t - f_m)^2}{f_t}$:

počet chlapců	naměřená četnost	teoretická četnost	$(f_t - f_m)^2$	$\frac{(f_t-f_m)^2}{f_t}$
0	40	32	64	2
1	184	160	576	3,6
2	300	320	400	1,25
3	268	320	2704	8,45
4	196	160	1296	8,1
5	36	32	16	0,5

K3: Za předpokladu platnosti H_0 má kritérium z bodu K2 rozdělení $\chi^2(k-1)$. Tento fakt nebudeme dokazovat.

K4: Kritérium je vhodným reprezentantem míry platnosti H_0: pokud H_0 platí, očekáváme, že hodnota kritéria bude malá, pokud neplatí, bude velká. Jedná se tedy o jednostranný test, kde k je počet skupin četnosti, tj. v našem případě $k = 6$. Tedy pro $\alpha = 0{,}05$ je $\chi_k(5) = 11{,}0705$.

K5:

$$\sum_1^6 \frac{(f_m - f_t)^2}{f_t} = 23{,}9 > 11{,}07 \Rightarrow \text{ zamítáme } H_0,$$

binomické rozdělení není příliš dobré pro popis našich dat (tj. pohlaví chlapců při narození se nechová jako počet líců při hodu korunou).

6.2.3 Testování nezávislosti v kontingenční tabulce

Příklad 6.3. Zajímá nás, zda existuje vztah mezi názory na jadernou energii a politickou příslušností. Proto jsme se dotázali nezávisle vybraných 200 lidí, jaký mají názor na jadernou elektrárnu Temelín (neměla by být v provozu – nezajímá mě to – měla by být v provozu), a dále které ze stran ODS, ČSSD dávají větší přednost. Výsledky průzkumu byly sestaveny do kontingenční tabulky (čísla v závorkách vyjadřují empirické pravděpodobnosti = četnosti vydělené číslem 200):

	elektr. by neměla být	nevím	elektr. by měla být	
ČSSD	40 (0,2)	70 (0,35)	40 (0,2)	150 (0,75)
ODS	35 (0,175)	5 (0,025)	10 (0,05)	50 (0,25)
	75 (0,375)	75 (0,375)	50 (0,25)	200 (1,000)

Při testování, zda existuje závislost mezi oběma proměnnými, můžeme použít test χ^2:

K1: H_0: Obě proměnné se chovají nezávisle, čili empirické rozdělení je hodně blízké následujícímu teoretickému rozdělení, kde poslední řádek a poslední sloupec jsou stejné jako v předchozí tabulce (udávající výsledky průzkumu), ovšem ostatní pravděpodobnosti jsou získány vynásobením příslušných pravděpodobností v posledním řádku a sloupci; označme $A_1 \ldots$ ČSSD ($P(A_1) = 0{,}75$); $A_2 \ldots$ ODS ($P(A_2) = 0{,}25$); $B_1 \ldots$ elektrárna NE ($P(B_1) = 0{,}375$); $B_2 \ldots$ nevím ($P(B_2) = 0{,}375$); $B_3 \ldots$ elektrárna ANO ($P(B_3) = 0{,}25$). Nyní pokud A_i, B_j jsou nezávislé, tak

$$P(A_i \cap B_j) = P(A_i) \cdot P(B_j)$$

(například $P(A_1 \cap B_1) = P(A_1) \cdot P(B_1) = 0{,}281$, atd.). Dostáváme tedy tabulku teoretických pravděpodobností (uvedených v závorce), četnosti jsou získány vynásobením příslušné pravděpodobnosti číslem 200 (četnosti tedy nemusí být celočíselné):

	elektr. by neměla být	nevím	elektr. by měla být	
ČSSD	56,2 (0,281)	56,2 (0,281)	37, 6 (0,188)	150 (0,75)
ODS	18,8 (0,094)	18,8 (0,094)	12,4 (0,062)	50 (0,25)
	75 (0,375)	75 (0,375)	50 (0,25)	200 (1,000)

H_1: Obě proměnné jsou závislé, čili nelze je dost dobře popsat příslušným teoretickým rozdělením.

K2: Kritériem bude $\sum_1^k \frac{(f_t - f_m)^2}{f_t}$, kde k je počet různých tříd četností (v našem případě počet různých oken tabulky kromě posledního řádku a posledního sloupce: $k = 2 \cdot 3 = 6$), f_t jsou příslušné teoretické četnosti (=četnosti z poslední tabulky), f_m příslušné naměřené četnosti (z předposlední tabulky).

K3: Za předpokladu platnosti H_0 má kriteriální funkce rozdělení $\chi^2((J-1)(K-1))$, kde J je počet podmínek veličiny A, K je počet podmínek veličiny B (tento fakt nebudeme dokazovat). V našem případě $\chi^2(1 \cdot 2) = \chi^2(2)$.

K4: Pro $\alpha = 0{,}05$ kritická hodnota $\chi_k(2) \doteq 5{,}99$.

K5:

$$\sum_1^6 \frac{(f_t - f_m)^2}{f_t} = \frac{(56{,}2-40)^2}{56{,}2} + \frac{(56{,}2-70)^2}{56{,}2} + \cdots + \frac{(12{,}4-10)^2}{12{,}4} = 32{,}76;$$

toto číslo zdaleka přesahuje kritickou hodnotu testu 5,99, tedy H_0 zamítáme, prokázala se závislost obou veličin.

6.3 Několik poznámek o vztazích mezi různými rozděleními

a) Jak už bylo řečeno, pro rostoucí n se hustota Studentova rozdělení $t(n)$ blíží hustotě rozdělení $U = No(0;1)$, čili

$$t(\infty) = U.$$

b) Platí

$$t^2(n) = F(1,n).$$

c) S využitím a),b) a faktu $U^2 = \chi^(1)$ můžeme psát

$$t^2(\infty) = U^2 = F(1,\infty) = \chi^2(1).$$

d) Víme, že platí $F(n_1-1, n_2-1) = \frac{est_1\,\sigma^2}{est_2\,\sigma^2}$, kde pro odhady téhož σ^2 platí $est_1\,\sigma^2 = \frac{\sum(x_{1i}-\overline{x}_1)^2}{n_1-1}$, $est_2\,\sigma^2 = \frac{\sum(x_{2i}-\overline{x}_2)^2}{n_2-1}$. Protože $\frac{\sum(x_{1i}-\overline{x}_1)}{\sigma^2}$ má rozdělení $\chi^2(n_1-1)$ a $\frac{\sum(x_{2i}-\overline{x}_2)}{\sigma^2}$ má rozdělení $\chi^2(n_2-1)$, můžeme psát

$$F(n_1-1, n_2-1) = \frac{\frac{1}{n_1-1}\cdot\chi^2(n_1-1)}{\frac{1}{n_2-1}\cdot\chi^2(n_2-1)}.$$

e) V příkladu 6.1 jsme testovali hypotézu $\sigma^2 = konst$ pomocí kritéria $\frac{\sum(x_i-\overline{x})^2}{\sigma_0^2}$. Pokud σ_0^2 chápeme jako odhad ropztylu σ_0^2 o nekonečně mnoha stupních volnosti, mohli jsme také provést F-test s kritériem $\frac{est\,\sigma^2}{\sigma_0^2}$, které za předpokladu platnosti H_0 má s využitím d) rozdělení $F(n-1,\infty)$ (kde samozřejmě $est\,\sigma^2 = \frac{\sum(x_i-\overline{x})^2}{n-1}$).

Pojmy k zapamatování

– Cílem této kapitoly bylo představit další užitečné rozdělení pravděpodobnosti, a sice rozdělení χ^2. rozdělení $\chi^2(n)$ (o n stupních volnosti) se definuje jako součet čtverců n normovaných normálních rozdělení U^2. Ukazuje se, že mnohé veličiny v praxi jsou rozděleny jako χ^2, a tedy toto rozdělení se vyskytuje v mnoha statistických testech:

1. V příkladu 6.1 jsme viděli, že rozdělení χ^2 lze užít v testu typu $\sigma^2 = konst$. Pokud měření jsme získali z populace, jejíž rozptyl je σ_0^2, pak kritérium
$$\frac{\sum(x_i - \overline{x})^2}{\sigma_0^2}$$
má rozdělení χ^2 o počtu stupňů volnosti $(n-1)$.
2. V příkladu 6.2 jsme viděli, že χ^2 lze užít v testu, zda naměřená data odpovídají jistému teoretickému rozdělení pravděpodobnosti (toto je asi nejčastější a nejznámější využití rozdělení χ^2, kterému se říká **test dobré shody**). Pokud máme konkrétně n naměřených (či pozorovaných) četností f_m, a jim odpovídající teoretické četnosti f_t, pak kritérium
$$\frac{\sum(f_t - f_m)^2}{f_t}$$
má rozdělení χ^2 o počtu stupňů volnosti $(n-1)$.
3. V příkladu 6.3 jsme viděli, že χ^2 lze užít při testu nezávislosti hodnot v kontingenční tabulce o počtu řádků a sloupců $J \times K$ (contingent = závislý, podmíněný … tj. test si klade otázku: do jaké míry jsou data pozorovaná přibližně stejná jako data vypočtená = nezávislá? tj: je mezi dvěma uvedenými veličinami nějaká závislost?). Tento test je specielním příkladem předchozího testu dobré shody, kdy za teoretické pravděpodobnosti vezmeme ty, co jsou dány součinem součtových pozorovaných pravděpodobností, nikoli měřením. Za této situace má kritérium
$$\frac{\sum(f_m - f_t)^2}{f_t}$$
rozdělení χ^2 o $(J-1)\cdot(K-1)$ stupních volnosti.

Kontrolní otázky

1. U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.
   a) Rozdělení χ^2 o n stupních volnosti je součtem n stejně rozdělených veličin U.
   b) Rozptyl veličiny χ^2 o n stupních volnosti je roven hodnotě $2n$.
   c) Protože hustota rozdělení χ^2 není symetrickou funkcí vzhledem k žádné přímce, některé kritické hodnoty mohou být záporné.
   d) Testy dobré shody lze používat jen u diskrétních náhodných veličin.
   e) Při oboustranném χ^2 testu jsou kritické hodnoty χ^2_m, χ^2_v symetrické vzhledem k průměru, tj. $\chi^2_m = E\chi^2 - d$, $\chi^2_v = E\chi^2 + d$ pro jistou hodnotu d.
   f) Pro rostoucí n se hustota rozdělení $t^2(n)$ blíží hustotě rozdělení $\chi^2(n)$.
   g) Při testech v kontingenční tabulce je počet tříd četnosti vždy sudý.
   h) Test dobré shody je jednostranným (konkrétně pravostranným) testem.

Odpovědi na otázky
1a) – N (χ^2 je definováno jako součet veličin U^2), 1b) – A, 1c) – N (ze vztahu $\chi^2(1) = U^2$ je vidět, že veličina s tímto rozdělením může nabývat pouze kladných hodnot (resp. záporných

hodnot nabývá s nulovou pravděpodobností)), 1d) – N (příslušné četnosti lze počítat i u spojitých veličin, více viz příklad 4), 1e) – N (ne, protože hustota není symetrickou funkcí vzhledem k přímce $x = E\chi^2(n) = n$), 1f) – N (pro rostoucí n se $t^2(n)$ blíží rozdělení $\chi^2(1)$ ... stupeň volnosti je pouze jeden), 1g) – N (např. v příkladu 6.3 stačí, abychom uvažovali tři politické strany místo dvou, a počet tříd by byl $J \times K = 3 \cdot 3 = 9$), 1h) – A.

Cvičení

1. Prodej Kola-loky má normální rozdělení se střední hodnotou 82000 lahví denně a směrodatnou odchylkou 1500 lahví. V zájmu výrobce je snížit tuto směrodatnou odchylku, protože by to učinilo obchod pružnějším. Proto se snaží o novou reklamu výrobku. Prvních deset dnů „nového“ prodeje vykazuje tyto výsledky (v počtech prodaných lahví):

$$81752, 83812, 82104, 82529, 82620, 82033, 81925, 81599, 82730, 81885.$$

Ověřte testem, zda nová reklama snížila směrodatnou odchylku počtu prodaných lahví za den.

2. Výška mužů v USA má normální rozdělení se střední hodnotou 70 palců (jeden palec = 2,54 cm) a směrodatnou odchylkou dva palce. Antropologa Františka Neználka zajímá, zda muži kmene Bora-Bora mají tentýž rozptyl hodnot své výšky. Získá náhodně vybraný vzorek sedmi mužů kmene Bora-Bora:

$$69, 68, 68, 67, 70, 71, 69.$$

Může zamítnou nulovou hypotézu o stejných rozptylech?

3. Honza Kovář pracuje v mincovně. Jeho úkolem je zajistit, aby mince byly dobře vyváženy – aby například při hodu desetikorunou padal rub i líc stejně často. Proto hodí stovkou desetikorun a padne mu 61-krát líc. Testujte následující hypotézy:
H_0: pravděpodobnost padnutí líce je 0,5;
H_1: pravděpodobnost padnutí líce není 0,5.

4. Rozdělení IQ v České Republice je normální (v matematickém slova smyslu) se střední hodnotou 100 a rozptylem 225. Náhodně vybraný vzorek obyvatel Brna prokázal následující rozdělení IQ:

IQ	< 55	$55 - 70$	$70 - 85$	$85 - 100$	$100 - 115$	$115 - 130$	$130 - 145$	> 145
n_i	20	17	29	52	63	42	13	14

kde n_i je četnost. Řekli byste, že brňané dostatečně stejně reprezentují Českou Republiku ve vztahu k IQ? Proveďte statistický test dobré shody v tomto případě.

Výsledky

ad 1. $\sigma_0 = 1500$, $n = 10$,

$$s^2 = \frac{1}{n}\left(\sum x_i^2\right) - \overline{x}^2 = \frac{1}{10} \cdot 67734929545 - \left(\frac{1}{10} \cdot 822989\right)^2 = 384013{,}29,$$

pak $\overline{s^2} = \frac{10}{9} \cdot s^2 = 426681{,}433$. Test:
H_0: $\sigma^2 = 1500^2$;
H_1: $\sigma^2 < 1500^2$;

$$\text{kritérium} \ldots \frac{n \cdot S^2}{\sigma_0^2} = \frac{10 \cdot 384013{,}29}{1500^2} \doteq 1{,}7 < \chi_k^2(9) = 3{,}32;$$

pro $\alpha = 0{,}05$; tedy H_0 nezamítáme, snížení odchylky se neprokázalo.

ad 2. Hodnota kritéria je 5, nemůže zamítnout H_0.

ad 3. Hodnota kritéria je 4,84, což je statisticky významné, zamítáme H_0.

ad 4. Protože normální rozdělení představuje spojitou náhodnou veličinu, dovolte mi celý příklad provést podrobně: Potřebujeme vlastně jen znát pravděpodobnosti, s jakými nabývá normálně rozdělená veličina hodnot z uvedených intervalů – četnosti pak získáme vynásobením těchto pravděpodobností číslem 242 (počet vybraných obyvatel):

IQ	< 55	$55-70$	$70-85$	$85-100$	$100-115$	$115-130$
pst	0,0013499	0,0214002	0,1359052	0,3413447	0, 3413447	0,1359052
n_i	0,33	5,18	32,89	82,605	82,605	32,89

IQ	$130-145$	> 145
pst	0,0214002	0,0013499
n_i	5,18	0,33

Například hodnotu 0,0013499 lze vypočíst následovně:

$$P(X < 55) = P(X \in (0;55)) = \Phi\left(\frac{55-100}{15}\right) - \Phi\left(\frac{0-100}{15}\right) \doteq \Phi(-3) - \Phi(-6{,}67) \doteq$$

$$\doteq \Phi(-3) = 0{,}0013499.$$

Četnost pak spočteme jako $242 \cdot 0{,}0013499 \doteq 0{,}33$. Další hodnoty v tabulce získáme analogicky. Nyní dosazením četností z této tabulky a tabulky zadání do vzorce kritéria testu dostaneme:

$$\text{krit.} = \frac{(0{,}33-20)^2}{0{,}33} + \frac{(5{,}18-17)^2}{5{,}18} + \cdots + \frac{(0{,}33-14)^2}{0{,}33},$$

což je vysoké číslo, mnohem větší než kritická hodnota $\chi_k^2(7)$, takže zamítáme hypotézu H_0 o tom, že Brno je vyváženým reprezentantem IQ v České Republice.

Maplety

V následujících mapletech si můžete některé studované pojmy přiblížit, případně si sestavit vlastní zadání příkladů.

1. Rozdělení χ^2

7 Neparametrické testy

Průvodce studiem

Jak už bylo řečeno v kapitole 2.5 (oddíl 2.5.6), většina dosud probraných testů vychází z platnosti tří následujících předpokladů:

1. *Naměřené hodnoty x_i jsou navzájem nezávislé.*
2. *Měřená veličina má normální rozdělení.*
3. *Rozptyl uvnitř jednotlivých skupin experimentu je stejný.*

Platnost předpokladu 1 lze obvykle zaručit vhodným navržením experimentu. Pokud jsou však vážně porušeny předpoklady 2 nebo 3, musíme místo parametrického testu vzít vhodný tzv. **neparametrický test**. *Přehled neparametrických testů bude tématem této kapitoly. Je možné, že v této kapitole budou častější numerické chyby, prosím na jejich upozornění, pokud na ně narazíte. Jednotlivé neparametrické testy budou vysvětleny, jak už je zvykem v tomto textu, přímo na příkladech.*

Cíle

Po prostudování této kapitoly budete schopni:

- Pracovat s tzv. neparametrické statistické testy.
- Chápat rozdíly mezi parametrickými a neparametrickými testy.
- Určovat použitelnost paremetrických i neparametrických testů v závislosti na typu řešeného problému.

7.1 Mannův-Whitneyův test podle pořadí

Příklad 7.1. Vraťme se k příkladu 6.1, kde jsme chtěli zjistit, jaký vliv na studenty bude mít počítačová výuka matematiky. Proveďme nyní experiment jiného rázu: náhodně vybraných 24 studentů rozdělíme na dvě skupiny po dvanácti lidech. Jedna skupina se účastnila výuky pod dohledem učitele, druhá příslušné počítačové výuky. Potom se žáci podrobili písemce, kde se měřil čas potřebný na vyřešení zadaných úloh. Získala se data:

$$1\ldots\ \text{bežná výuka: } 43, 12, 21, 41, 39, 23, 27, 37, 35, 31, 33, 29;$$

2 ... počítačová výuka: $3, 13, 1, 5, 11, 10, 9, 8, 6, 2, 4, 7$;

Pomocí všech naměřených hodnot odhadneme rozptyl:

$$est\ \sigma^2 = \frac{SS_1 + SS_2}{V_1 + V_2} = \frac{908{,}92 + 154{,}92}{11 + 11} = 48{,}36$$

A nyní můžeme provést t-test:

K1: H_0: $\mu_1 = \mu_2$ (doba potřebná na vyřešení písemky z aritmetiky má stejnou střední hodnotu u obou skupin);
H_1: $\mu_1 \neq \mu_2$.

K2: Kritériem je $\frac{\overline{x}_1 - \overline{x}_2 - \mu_{1-2}}{est\ \sigma_{1-2}}$, kde $\mu_{1-2} = 0$,

$$est\ \sigma_{1-2} = \sqrt{est\ \sigma_1^2 + est\ \sigma_2^2} = \sqrt{\frac{48{,}36}{12} + \frac{48{,}36}{12}} = 2{,}84.$$

K3: Za předpokladu platnosti H_0 má naše kritérium rozdělení $t(22)$, protože $V_1 + V_2 = 22$.

K4: Pro $\alpha = 0{,}05$ je $t_k(22) = \pm 2{,}09$.

K5: Hodnota kritéria: $\frac{30{,}91 - 6{,}58}{2{,}84} = 8{,}57$, a tak H_0 zamítáme.

Ovšem právě provedený t-test nesplňoval předpoklad rovnosti rozptylů v obou skupinách: $est_1\ \sigma^2 = \frac{SS_1}{V_1} = 82{,}63$, kdežto $est_2\ \sigma^2 = \frac{SS_2}{V_2} = 14{,}08$. První odhad je přibližně šestinásobkem druhého, což už přesahuje rozumnou (= čtyřnásobnou) míru. Toto hrubé porušení předpokladu t-testu je důvodem k detailnějšímu prozkoumání dat pomocí neparametrického testu. Vhodným kandidátem nyní je Mannův-Whitneyův test.

Vraťme se datům z právě opuštěného příkladu 7.1 a podrobme jej Mannovu-Whitneyovu testu:

K1: Uspořádejme všechna měření (z obou souborů dohromady) podle velikosti a zaznamenejme přitom

a) pořadí dané hodnoty;

b) příslušnost dané hodnoty k experimentální skupině (B je běžná výuka, P je počítačová výuka).

hodnota	1	2	3	4	5	6	7	8	9	10	11	12
pořadí	1	2	3	4	5	6	7	8	9	10	11	12
skupina	P	P	P	P	P	P	P	P	P	P	P	B

13	21	23	27	29	31	33	35	37	39	41	43
13	14	15	16	17	18	19	20	21	22	23	24
P	B	B	B	B	B	B	B	B	B	B	B

K2: Nyní vypočteme jakési míry U_1, U_2:

$$U_1 = \sum_{\text{děti ze skupiny P}} (\text{počet dětí z B, které mají horší skóre než dítě } i).$$

Tedy pro dítě s hodnotou 1 na prvním místě pořadí má všech dvanáct dětí z B horší skóre. Pro dítě s hodnotou 2 na druhém místě pořadí má opět všech dvanáct dětí z B horší skóre, atd. až pro dítě s hodnotou 11 na jedenáctém místě má stále všech dvanáct dětí z B horší skóre. A konečně změna, pro dítě s hodnotou 13 na třináctém místě má už jen jedenáct dětí z B horší skóre. Dohromady

$$U_1 = \underbrace{12 + 12 + \cdots + 12}_{\text{jedenáctkrát}} + 11 = 143.$$

Podobně

$$U_2 = \sum_{\text{děti ze skupiny B}} (\text{počet dětí z P, které mají horší skóre než dítě } i).$$

Zde pouze pro dítě s hodnotou 12 na dvanáctém místě má jedno dítě z P horší skóre, tj.

$$U_1 = 1 + \underbrace{0 + 0 + \cdots + 0}_{\text{jedenáctkrát}} = 1.$$

Z konstrukce U_1, U_2 je vidět, že tyto míry zachycují závažnost, s jakou má jedna skupina lepší pořadí než druhá. Samozřejmě existují poněkud pohodlnější vzorce pro jejich výpočet:

$$U_1 = n_1 \cdot n_2 + \frac{n_2 \cdot (n_2 + 1)}{2} - \sum (\text{pořadí hodnot ze skupiny 2}),$$

$$U_2 = n_1 \cdot n_2 + \frac{n_1 \cdot (n_1 + 1)}{2} - \sum (\text{pořadí hodnot ze skupiny 1})$$

(mimochodem, platí také $U_1 + U_2 = n_1 \cdot n_2$ – v našem příkladu tedy $143 + 1 = 12 \cdot 12$). Kritériem testu bude nyní menší z obou vypočtených hodnot: $U = min(U_1, U_2)$. V našem příkladu $U = min(143, 1) = 1$.

K3: Kritická hodnota testu:

a) pro $n_1, n_2 \leq 20$**:** Byly sestaveny tabulky kritických hodnot našeho oboustranného testu, a to pro $\alpha = 0{,}05$ tabulka 7.1, pro $\alpha = 0{,}01$ tabulka 7.2.

b) pro $n_1, n_2 > 20$**:** Rozdělení kritéria U je normální se střední hodnotou $\mu = \frac{n_1 \cdot n_2}{2}$ a směrodatnou odchylkou

$$\sigma = \sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 + 1)}{12}},$$

čili normovanou hodnotu
$$\frac{U-\frac{n_1\cdot n_2}{2}}{\sqrt{\frac{n_1\cdot n_2\cdot(n_1+n_2+1)}{12}}}$$
testujeme s běžnými kritickými hodnotami $\pm 1{,}96$ normálního rozdělení pro $\alpha = 0{,}05$.

K4+K5: V našem příkladu tedy stanovíme kritickou hodnotu

a) z tabulky pro $\alpha = 0{,}05$, $n_1 = n_2 = 12$: $U_k = 37 > U = 1$, tj. H_0 zamítáme ve prospěch alternativní hypotézy H_1.

Všimněte si, že na rozdíl od většiny testů v této přednášce zde zamítnutí H_0 nastane tehdy, když hodnota kritéria kritickou hodnotu NEPŘEKROČÍ. Je to dáno konstrukcí kriterijní funkce – pokud převáží malé hodnoty pořadí v jednom souboru, tak hodnota kritéria je malá, ale současně to znamená jasný a zřetelný rozdíl mezi skupinami.

b) pomocí aproximace normálním rozdělením (nyní méně přesné, ale výpočetně jednodušší):
$$\frac{U-\frac{n_1\cdot n_2}{2}}{\sqrt{\frac{n_1\cdot n_2\cdot(n_1+n_2+1)}{12}}} = -4{,}1 < -1{,}96 \Rightarrow H_0 \text{ zamítáme}$$
ve prospěch alternativní hypotézy H_1.

Vidíme, že výsledky neparametrického testu jsou stejné jako výsledky příslušného parametrického testu 7.1 s porušenými předpoklady. Z toho je vidět, že navzdory porušeným předpokladům t-testu je i při silném rozdílu obou souborů výsledek testu správný, To ovšem nemusí být pravdou, pokud rozdíly mezi oběma skupinami nebudou tak jednoznačné, jak to dokumentují další situace v této kapitole.

Tab. 7.1: Kritické hodnoty Mannova-Whitneyova testu pro $\alpha = 0{,}05$.

	n_2																	
n_1	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
2	–	–	–	–	–	0	0	0	0	1	1	1	1	1	2	2	2	2
3	–	–	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4		0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	13
5			2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6				5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7					8	10	12	14	16	18	20	22	24	26	28	30	32	34
8						13	15	17	19	22	24	26	29	31	34	36	38	41
9							17	20	23	26	28	31	34	37	39	42	45	48
10								23	26	29	33	36	39	42	45	48	52	55
11									30	33	37	40	44	47	51	55	58	62
12										37	41	45	49	53	57	61	65	69
13											45	50	54	59	63	67	72	76
14												55	59	64	67	74	78	83
15													64	70	75	80	85	90
16														75	81	86	92	98
17															87	93	99	105
18																99	106	112
19																	113	119
20																		127

Tab. 7.2: Kritické hodnoty Mannova-Whitneyova testu pro $\alpha = 0{,}01$.

	n_2																	
n_1	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0	0
3	–	–	–	–	–	–	0	0	0	1	1	1	2	2	2	2	3	3
4		–	–	0	0	1	1	2	2	3	4	4	5	5	6	6	7	8
5			0	1	2	3	3	4	5	6	7	7	8	9	10	11	12	13
6				2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7					4	6	7	9	10	12	13	15	16	18	19	21	22	24
8						8	10	12	14	16	18	19	21	23	25	27	29	31
9							11	13	16	18	20	22	24	27	29	31	33	36
10								16	18	21	24	26	29	31	34	37	39	42
11									21	24	27	30	33	36	39	42	45	48
12										27	31	34	37	41	44	47	51	54
13											34	38	42	45	49	53	56	60
14												42	46	50	54	58	63	67
15													51	55	60	64	69	73
16														60	65	70	74	79
17															70	75	81	86
18																81	87	92
19																	93	99
20																		105

7.2 Kruskalův-Wallisův test

Příklad 7.2. Dva psychologové, psycholog P a psycholog N, se doslechli o zajímavém experimentu, kdy byl několika lidem promítnut videozáznam autonehody. Pak byli rozděleni do dvou skupin a první skupině byla položena otázka: Jak rychle jelo bílé auto, když projíždělo kolem jedné stodoly na té venkovské silnici? Zatímco druhé skupině byla položena otázka: Jak rychle jelo bílé auto po té venkovské silnici? Všichni museli na tuto otázku odpovědět.

Potom byla oběma skupinám položena otázka, zda viděli na videozáznamu stodolu. I když na videozáznamu žádná stodola nebyla, na tuto otázku odpovědělo kladně 17% první skupiny, zatímco jen 3% druhé skupiny. Závěrem experimentu bylo, že pokud první otázka obsahovala lživou informaci, tak jaksi mimoděk ji vtiskla do přemýšlení lidí, a ti ji pak považují za pravdivou.

Psychologové P a N si položili otázku, zda někdy lživé nebo zkreslené informace o autonehodě, které jsou zveřejněny v jistých nevinách, neovlivní způsob, jakým si svědek nehody tuto zapamatuje. A také je napadlo: do jaké míry různé prestižní noviny ovlivní očitého svědka autonehody svými zkreslenými informacemi?

Proto se rozhodli provést další experiment. Náhodně vybrali dvanáct lidí, kterým promítli videozáznam jisté autonehody. Potom tito lidi byli rozděleni do tří skupin po čtyřech. Lidé z první skupiny pak dostali přečíst článek o této autonehodě, který byl neznámého původu a byl naprosto pravdivý. Lidé ze druhé skupiny dostali k přečtení článek o téže autonehodě uveřejněný v deníku Rovnost a lidé ze třetí skupiny dostali k přečtení uveřejněný v deníku Dnes. Články v obou denících obsahovaly větší počet chybných informací.

Poté se všechny skupiny podrobily dotazníku, který měl zjistit, do jaké míry si autonehodu pamatují pravdivě. Následující data uvádějí počet chyb jednotlivých vyplněných dotazníků:

noname	Rovnost	MF dnes
4	4	9
7	9	13
1	10	15
1	5	7

Naměřená data byla k dispozici ke statistickému testování. Psycholog P se rozhodl provést F-test jednorozměrné analýzy rozptylu:

K1: H_0: $\mu_1 = \mu_2 = \mu_3$,
H_1: neplatí H_0.

K2: Kritériem bude $\frac{est\ RMT}{est\ RUT}$

K3: Za předpokladu platnosti H_0 má kritérium z bodu (K2) rozdělení $F(2; 9)$.

K4: Pro $\alpha = 0{,}05$ je $F_k(2;9) = 4{,}26$.

K5:

$$\frac{est\ RMT}{est\ RUT} = \frac{\frac{120{,}17}{2}}{\frac{90{,}75}{9}} = 5{,}96 > 4{,}26 \Rightarrow H_0 \text{ zamítáme.}$$

Psycholog N se rozhodl provést neparametrický **Kruskalův - Wallisův jednorozměrný test rozptylu podle pořadí**, který je přirozeným rozšířením předchozího testu 7.1 na více než dvě skupiny měření:

K1: H_0: $\mu_1 = \mu_2 = \mu_3$,
H_1: neplatí H_0.

K2: Provedeme totéž, co v 7.1: všech dvanáct hodnot seřadíme podle velikosti, a každé hodnotě přiřadíme její pořadí.

Vsuvka: V příkladu 7.1 bylo pořadí jednoznačné, protože všechny hodnoty byly navzájem různé. Někdy se ale situace komplikuje. Například hodnotám $(2, 8, 10, 10, 14)$ je potřeba přiřadit pořadí $(1; 2; 3{,}5; 3{,}5; 5)$, tj. dvěma hodnotám 10 se přiřadí průměr jejich pořadí $\frac{3+4}{2} = 3{,}5$. Nebo hodnotám $(2, 8, 10, 10, 10, 14)$ odpovídá pořadí $(1, 2, 4, 4, 4, 6)$, tj. opět se třem hodnotám 10 přiřadí průměr jejich pořadí $\frac{3+4+5}{3} = 4$. Další příklady hodnot a jejich pořadí:

hodnota	2	2	8	10	10	14
pořadí	1,5	1,5	3	4,5	4,5	6

hodnota	12	14	14	14	15	16	16	16	16	21
pořadí	1	3	3	3	5	7,5	7,5	7,5	7,5	10

V našem příkladu tedy:

hodnota	1	1	4	4	5	7	7	9	9	10	13	15
pořadí	1,5	1,5	3,5	3,5	5	6,5	6,5	8,5	8,5	10	11	12

Uvedená pořadí znovu roztřídíme do příslušných tří skupin a v každé skupině je sečteme:

skupina 1	skupina 2	skupina 3
3,5	3,5	8,5
6,5	8,5	11
1,5	10	12
1,5	5	6,5
$R_1 = 13$	$R_2 = 27$	$R_3 = 38$

Čtenář už se jistě těší na kritérium, které se ve druhém kroku testu vždy objeví – kupodivu i v testech neparametrických. Tak tedy: kritériem bude funkce

$$H = \frac{12}{n \cdot (n+1)} \cdot \sum_{skupiny} \frac{R_j^2}{N_j} - 3 \cdot (n+1)$$

kde $n = 12$ je celkový počet osob,
N_j je počet osob ve skupině j
(Nutno dodat, že konstanty 12 a 3 vyskytující se ve vzorci pro H jsou neustále stejné a nezávisí na n nebo na N_j).

K3: —

K4+K5: Kritická hodnota testu:

a) pro malé hodnoty N_j: uvedeno v tabulkách pro tři vzorky – 7.3, 7.4.

b) pro $N_j > 5$: H má rozdělení $\chi^2(J-1)$, čili kritické hodnoty určíme z tabulky pro χ^2 (tabulky 6.1, 6.2).

V našem příkladu tedy stanovíme kritickou hodnotu

a) Na základě tabulky pro $N_j = 4$ (a $\alpha \leq 0{,}05$ hledáme co možná největší) $H_k = 5{,}69$, kdežto $H = 6{,}03$, tedy H_0 zamítáme.

b) Na základě rozdělení $\chi^2(2)$ pro $\alpha = 0{,}05$ máme $\chi_k^2(2) = 5{,}99 < H = 6{,}03$, tedy H_0 zamítáme.

Z porovnání přístupů psychologů N a P je vidět, že kritická hodnota testu N (= neparametrického testu) se vyskytovala blízko hodnoty kritéria, kdežto test psychologa P (= parametrický test) zamítl H_0 přesvědčivěji a jednoznačněji.

Obecně lze říci, že testy parametrické mají větší sílu, dokáží tedy lépe odhalit závislost mezi proměnnými, protože maximálně využívají informace dostupné v datech. V neparametrických testech je část informace z experimentu ztracena, protože využíváme pouze pořadí hodnot, nikoli skutečné hodnoty měření.

Pokud tedy psycholog P zajistil předpoklady parametrického textu, jeho výsledek má větší sílu, protože rozdíl souborů je statsiticky prokázán – může napsat článek do odborného psychologického časopisu. Pokud ovšem předpoklady parametrického testu jsou vážně porušeny, nemáme k dispozici žádný lepší výsledek než výsledek psychologa N (tj. výsledek neparametrického testu).

Tabulka 7.5 udává přehled nejdůležitějších neparametrických testů, které lze použít, pokud nejsou splněny předpoklady příslušných parametrických testů. Z tabulky je patrno několik věcí: za prvé, že pro daná data je někdy možné provést více různých testů. Za druhé, neexistuje neparametrický test, který by vyhovoval situaci dvourozměrné analýzy rozptylu, což je dost nešťastné.

Neparametrické testy představené v této kapitole nejsou vyčerpávající – existují mnohé další. Ale ty zde uvedené jsou nejčastěji používány. Ve zbytku této kapitoly uvedeme na příkladech ty neparametrické testy z tabulky 7.5, které dosud nebyly probrány.

Tab. 7.3: Kritické hodnoty Kruskalova - Wallisova testu pro tři skupiny a $N_j \leq 4$ – část 1.

N_1	N_2	N_3	H_k	α	N_1	N_2	N_3	H_k	α	N_1	N_2	N_3	H_k	α
2	1	1	2,7000	0,500	4	1	1	3,5714	0,200	4	4	3	7,1439	0,010
					4	2	1	4,8214	0,057				7,1364	0,011
2	2	1	3,6000	0,200				4,5000	0,076				5,5985	0,049
								4,0179	0,114				5,5758	0,051
2	2	2	4,5714	0,067	4	2	2	6,0000	0,014				4,5455	0,099
			3,7143	0,200				5,3333	0,033				4,4773	0,102
								5,1250	0,052	4	4	4	7,6538	0,008
3	1	1	3,2000	0,300				4,4583	0,100				7,5385	0,011
								4,1667	0,105				5,6923	0,049
3	2	1	4,2857	0,100	4	3	1	5,8333	0,021				5,6538	0,054
			3,8751	0,133				5,2083	0,050				4,6539	0,097
								5,0000	0,057				4,5001	0,104
3	2	2	5,3572	0,029				4,0556	0,093	5	1	1	3,8571	0,143
			4,7143	0,048				3,8889	0,129	5	2	1	5,2500	0,036
			4,5000	0,067	4	3	2	6,4444	0,008				5,0000	0,048
			4,4643	0,105				6,3000	0,011				4,4500	0,071
								5,4444	0,046				4,2000	0,095
3	3	1	5,1429	0,043				5,4000	0,051				4,1487	0,099
			4,5714	0,100				4,5111	0,098				4,1231	0,103
			4,0000	0,129				4,4444	0,102	5	2	2	6,5333	0,008
					4	3	3	6,7455	0,010				6,1333	0,013
3	2	2	5,3572	0,029				6,7091	0,013				5,1600	0,034
			4,7143	0,048				5,7909	0,046				5,0400	0,056
			4,5000	0,067				5,7273	0,050				4,3733	0,090
			4,4643	0,105				4,7091	0,092				4,2933	0,122
								4,7000	0,101	5	3	1	6,4000	0,012
3	3	2	6,2500	0,011	4	4	1	6,6667	0,010				4,9600	0,048
			5,3611	0,032				6,1667	0,022				4,8711	0,052
			5,1389	0,061				4,9667	0,048				4,0178	0,095
			4,5556	0,100				4,8667	0,054				3,8400	0,123
			4,2500	0,121				4,1667	0,082	5	3	2	6,9091	0,009
								4,0667	0,102				6,8218	0,010
3	3	3	7,2000	0,004	4	4	2	7,0364	0,006				5,2509	0,049
			6,4889	0,011				6,8727	0,011				5,1055	0,052
			5,6889	0,029				5,4545	0,046				4,6509	0,091
			5,6000	0,050				5,2364	0,052				4,4945	0,101
			5,0667	0,086				4,5545	0,098					
			4,6222	0,100				4,4455	0,103					

Tab. 7.4: Kritické hodnoty Kruskalova-Wallisova testu pro 3 skupiny a $N_j \leq 4$ – část 2.

N_1	N_2	N_3	H_k	α	N_1	N_2	N_3	H_k	α
5	3	3	7,0788	0,009	5	5	2	7,3385	0,010
			6,9818	0,011				7,2692	0,010
			5,6485	0,049				5,3385	0,047
			5,5152	0,051				5,2462	0,051
			4,5333	0,097				4,6231	0,097
			4,4121	0,109				4,5077	0,100
5	4	1	6,9545	0,008	5	5	3	7,5780	0,010
			6,8400	0,011				7,5429	0,010
			4,9855	0,044				5,7055	0,046
			4,8600	0,056				5,6264	0,051
			3,9873	0,098				4,5451	0,100
			3,9600	0,102				4,5363	0,102
5	4	2	7,2045	0,009	5	5	4	7,8229	0,010
			7,1182	0,010				7,7914	0,010
			5,2727	0,049				5,6657	0,049
			5,2682	0,050				5,6429	0,050
			4,5409	0,098				4,5229	0,099
			4,5182	0,101				7,5200	0,101
5	4	3	7,4449	0,010	5	5	5	8,0000	0,009
			7,3949	0,011				7,9800	0,010
			5,6564	0,049				5,7800	0,049
			5,6308	0,050				5,6600	0,051
5	4	4	7,7604	0,009				4,5600	0,100
			7,7440	0,011				4,5000	0,102
			5,6571	0,049					
			5,6176	0,050					
			4,6187	0,100					
			4,5527	0,102					
5	5	1	7,3091	0,009					
			6,8364	0,011					
			5,1273	0,046					
			4,9091	0,053					
			4,1091	0,086					
			4,0346	0,105					

Tab. 7.5: Přehled použití parametrických i neparametrických testů

data	**účel testu**	**parametrický test**	**neparametrický test**
jeden vzorek	zjistit, zda střední hodnota populace, ze které byl vzorek vybrán, se liší od jisté hodnoty μ_0	U-test či t-test	znaménkový test
dva vzorky jednou	zjistit, zda střední hodnoty populací, ze kterých byly vzorky vybrány, se rovnají	U-test či t-test pro nezávislé skupiny	Mannův–Whitneyův test
jeden vzorek dvakrát	zjistit, zda střední hodnoty populací, ze kterých byly vzorky vybrány, se rovnají	U-test či t-test typu „jeden vzorek dvakrát“	znaménkový test nebo Wilcoxonův test
více než dva vzorky jednou	zjistit, zda populace, ze kterých byly vzorky vybrány, mají tutéž střední hodnotu	jednorozměrná ANOVA (F-test)	Kruskalův–Wallisův test
jeden vzorek více než dvakrát dvakrát	zjistit, zda populace, ze kterých byly vzorky vybrány, mají tutéž střední hodnotu	jednorozměrná ANOVA opakovaného měření	Friedmanův test
množina položek, z nichž každá má dva parametry	zjistit, zda tyto parametry či proměnné jsou korelovány	Pearsonův test korelace	Spearmanův test korelace
jeden vzorek	zjistit, zda populace, ze které byl vzorek vybrán, má jisté teoretické rozdělení		test χ^2 typu 6.2.2 nebo 6.2.3 nebo Kolmogorovův–Smirnovův test

7.3 Kolmogorovův – Smirnovův test

Příklad 7.3. Chceme zjistit, zda na lidi více působí usměvavé tváře, nebo tváře bez úsměvu. Proto náhodně vybereme deset lidí a zeptáme se jich, která z daných pěti fotografií se jim nejvíce líbí. Fotografie se přitom liší pouze silou úsměvu (1 = žádný úsměv, ..., 5 = úsměv naplno). Získala se data

$x \dots$ foto	1	2	3	4	5
$f_e(x) \dots$ četnost	0	1	0	4	5

Prokažme statistickým testem jejich významnost:

K1: H_0: úsměv na lidi nemá vliv;
H_1: má.

K2: Kritériem bude maximální rozdíl teoretické a empirické hodnoty v histogramu kumulativních četností

$$D = \frac{1}{n} \cdot max|F_t(x) - F_e(x)|$$

pro kumulativní četnost platí $F(x) = \sum_{k \le x} f(k)$.

K3: Při platnosti H_0 má teoretické rozdělení četností tvar

$x \dots$ foto	1	2	3	4	5
$f_t(x) \dots$ četnost	2	2	2	2	2

K4: Stanovení kritické hodnoty: využijeme tabulky 7.6 – v našem příkladu například pro $\alpha = 0{,}01$ a $n = 10$ máme $D_k = 0{,}49$.

K5: Vypočtěmě příslušné kumulativní četnosti a hodnotu kritéria D:

$x \dots$ foto	1	2	3	4	5
$F_e(x)$	0	1	1	5	10
$F_t(x)$	2	4	6	8	10
$\lvert F_e(x) - F_t(x)\rvert$	2	3	5	3	0

$max|F_t(x) - F_e(x)| = 5$, tj. $D = \frac{5}{10} = 0{,}5$, máme $D > D_k$, a tedy zamítáme H_0.

7.4 Wilcoxonův test

Příklad 7.4. Chceme otestovat, zda určitý kurs čtení zlepšuje techniku čtení. Pro sedm náhodně vybraných lidí byla provedena zkouška techniky čtení před a po kursu. Získaly se dva soubory dat (jedná se o experiment opakovaného měření pro tutéž skupinu, tj. experiment typu „jeden vzorek dvakrát“). Pro tato data bychom mohli provést t-test (viz kapitola 2.5), ale užijeme si parametrický **Wilcoxonův test** pro soubor párových hodnot:

jedinec	před	po	rozdíl	absol. rozdíl	pořadí abs. rozdílu	pořadí se znaménkem
A	74	76	−2	2	3	−3
B	81	80	1	1	1,5	1,5
C	85	89	−4	4	6	−6
D	79	88	−9	9	7	−7
E	92	95	−3	3	4,5	−4,5
F	83	80	3	3	4,5	4,5
G	87	86	1	1	1,5	1,5

Tab. 7.6: Kritické hodnoty Kolmogorovova–Smirnovova testu

			$\alpha =$		
(n)	0,20	0,15	0,10	0,05	0,01
1	0,900	0,925	0,950	0,975	0,995
2	0,684	0,726	0,776	0,842	0,929
3	0,565	0,597	0,642	0,708	0,828
4	0,494	0,525	0,564	0,624	0,733
5	0,446	0,474	0,510	0,565	0,669
6	0,410	0,436	0,470	0,521	0,618
7	0,381	0,405	0,438	0,486	0,577
8	0,358	0,381	0,411	0,457	0,543
9	0,339	0,360	0,388	0,432	0,514
10	0,322	0,342	0,368	0,410	0,490
11	0,307	0,326	0,352	0,391	0,468
12	0,295	0,313	0,338	0,375	0,450
13	0,284	0,302	0,325	0,361	0,433
14	0,274	0,292	0,314	0,349	0,418
15	0,266	0,283	0,304	0,338	0,404
16	0,258	0,274	0,295	0,328	0,392
17	0,250	0,266	0,286	0,318	0,381
18	0,244	0,259	0,278	0,309	0,371
19	0,237	0,252	0,272	0,301	0,363
20	0,231	0,246	0,264	0,294	0,356
25	0,21	0,22	0,24	0,27	0,32
30	0,19	0,20	0,22	0,24	0,29
35	0,18	0,19	0,21	0,23	0,27
≥ 35	$\frac{1{,}07}{\sqrt{n}}$	$\frac{1{,}14}{\sqrt{n}}$	$\frac{1{,}22}{\sqrt{n}}$	$\frac{1{,}36}{\sqrt{n}}$	$\frac{1{,}63}{\sqrt{n}}$

Z tabulky je viděl, jak se hodnoty v posledních čtyřech sloupcích konstruují – sloupec rozdílů se zbaví znamének a máme sloupec absolutních hodnot, které se seřadí podle velikosti. Další sloupec obsahuje odpovídající pořadí podle sloupce absolutních hodnot, a poslední sloupec jen eventuelně přidá některým hodnotám záporné znaménko, pokud toto záporné znaménko bylo ve sloupci rozdílů. Proveďme nyní **Wilcoxonův test**:

K1: H_0: Kurs významně nezlepšil techniku čtení ($\mu_1 = \mu_2$)
H_1: Kurs změnil kvalitu techniky čtení $\mu_1 \neq \mu_2$

K2: Kritériem bude minimální ze součtů T_+, T_-, kde $T_+ =$ součet kladných pořadí (v posledním sloupci tabulky); $T_- =$ absolutní hodnota součtu záporných pořadí $W = min(T_+, |T_-|)$.

K4: Kritickou hodnotu W_k určíme z tabulky 7.7 pro Wilcoxonův test (v prvním svislém sloupci této tabulky jsou hodnoty N). V našem případě pro $\alpha = 2q = 0{,}05$ a $N = 7$ oboustranného testu (tj pro určení sloupečku vezmeme záhlaví $\alpha = 2q$) najdeme $W_k = 2$.

Pro jednostranný test bychom použili v tabulce 7.7 záhlaví $\alpha = q$. Pokud N je dostatečně velké ($N \geq 25$), lze W aproximovat normálním rozdělením a příslušnou U-hodnotu určit ze vzorce $U = \frac{W - \frac{N(N+1)}{4}}{\frac{1}{24} \cdot \sqrt{N(N+1)(2N+1)}}$ (pak pro $\alpha = 0{,}05$ můžeme použít známé kritické hodnoty $\pm 1{,}96$).

K5: $W = 7{,}5 > W_k = 2$, tedy H_0 NEzamítáme. POZOR, u tohoto testu je rozhodnutí vedeno v opačném smyslu než u většiny testů v této knize – pokud T_+ reprezentuje součet kladných odchylek a T_- součet záporných, tak zamítnutí H_0 by nastalo tehdy, když by některá z těchto hodnot byla MENŠÍ než kritická hodnota.

Použijeme-li aproximaci normálním rozdělením, tak $U = -5{,}38 \notin \langle -1{,}96; 1{,}96 \rangle$, čili H_0 zamítáme – zde už postupujeme klasicky, tj. pokud vypočtená hodnota kritéria leží mimo interval určený kritickými hodnotami, tak H_0 jsme zamítli.

7.5 Friedmanův test

Příklad 7.5. Chceme porovnat schopnost vidění v různých fázích dne. Náhodně bylo vybráno 11 lidí a provedeny zkoušky zrakových schopností ráno, v poledne, odpoledne a večer. Záskala se následující data (čísla v závorkách udávají pořadí každého jednotlivce v dané ze čtyř kategorií měření (ráno, poledne, odpoledne, večer)):

člověk	ráno	poledne	odpoledne	večer
1	1 (2)	4 (3)	8 (4)	0 (1)
2	3 (2)	2 (1)	4 (3)	13 (4)
3	14 (4)	4 (2)	7 (3)	2 (1)
4	10 (4)	4 (2)	9 (3)	3 (1)
5	10 (4)	4 (2)	5 (3)	3 (1)
6	4 (1)	12 (4)	10 (2)	11 (3)
7	10 (3)	3 (1)	11 (4)	9 (2)
8	1 (2)	3 (3)	10 (4)	0 (1)
9	12 (3)	11 (2)	13 (4)	10 (1)
10	10 (3)	0 (1)	11 (4)	3 (2)
11	2 (2)	3 (3)	13 (4)	1 (1)
	$R_1 = 30$	$R_2 = 24$	$R_3 = 38$	$R_4 = 18$

R_j udává součet pořadí v daném sloupci.

Tab. 7.7: Kritické hodnoty Wilcoxonova testu.

$\alpha = q$	0,05	0,025	0,01	0,005
$\alpha = 2q$	0,10	0,05	0,02	0,01
5	0	–	–	–
6	2	0	–	–
7	3	2	0	–
8	5	3	1	0
9	8	5	3	1
10	10	8	5	3
11	13	10	7	5
12	17	13	9	7
13	21	17	12	9
14	25	21	15	12
15	30	25	19	15
16	35	29	23	19
17	41	34	27	23
18	47	40	32	27
19	53	46	37	32
20	60	42	43	37
21	67	58	49	42
22	75	65	55	48
23	83	73	62	54
24	91	81	69	61
25	100	89	76	68
26	110	98	84	75
27	119	107	92	83
28	130	116	101	91
29	140	126	110	100
30	151	137	120	109
31	163	147	130	118
32	175	159	140	128
33	187	170	151	138
34	200	182	162	148
35	213	195	173	159
36	227	208	185	171
37	241	221	198	182
38	256	235	211	194
39	271	249	224	207
40	286	264	238	220
41	302	279	252	233
42	319	294	266	247
43	336	310	281	261
44	353	327	296	276
45	371	343	312	291
46	389	361	328	307
47	407	378	345	322
48	426	396	362	339
49	446	415	379	355
50	466	434	397	373

Provedeme nyní **Friedmanův test**. Podobně jako když u experimentů typu „více vzorků jednou" používáme při dvou vzorcích Mannův–Whitneův test a při více než dvou Kruskalůlv–Wallisův test, u experimentu opakovaného měření („jeden vzorek vícekrát") použijeme při dvou podmínkách Wilcoxonův test a při více podmínkách Friedmanův test.

K1: H_0: Vidění za různých podmínek se významně neliší;
H_1: liší.

a) Pro malé N, J:

K2: Kritériem je $S = \sum_1^J R_j^2 - \frac{1}{4}N^2 \cdot J(J+1)^2$.

K4: Pro určení kritické hodnoty S_k užijeme tabulku 7.8 při $\alpha = 0{,}05$.

K5: Pokud $S \geq S_k$, zamítáme H_0.

b) Pro větší N, J:

K2: Kritériem je $\chi_r^2 = \frac{12}{N \cdot J \cdot (J+1)} \cdot \left(\sum_1^J R_j^2\right) - 3N(J+1)$.

K3: Při platnosti H_0 má naše kritérium rozdělení $\chi^2(J-1)$.

K4: Pro $\alpha = 0{,}01$ je kritická hodnota $\chi_k^2(3) = 11{,}34$.

K5: $\chi_r^2 = 11{,}95 > 11{,}34$, tedy H_0 zamítáme.

7.6 Spearmanův koeficient korelace mezi pořadími

Příklad 7.6. Zajímá nás, zda existuje vztah mezi návštěvami vězení a kvalitou vězeňského života. Tušíme, že tento vztah by mohl být přímou úměrností, tj. čím větší počet návštěv je ve vězení povolen, tím větší kvalitou vězni oceňují vězeňský život. Získali jsme data z deseti věznic a chceme tato data zpracovat:

vězení	pořadí počtu návštěv	pořadí kvality	d_i	d_i^2
1	1	1,5	0,5	0,25
2	2,5	3	0,5	0,25
3	2,5	1,5	−1	1
4	4	5	1	1
5	5	5	0	0
6	6	8,5	2,5	6,25
7	7	5	−2	4
8	8	7	−1	1
9	9	8,5	−0,5	0,25
10	10	10	0	0

Nyní bychom mohli použít Pearsonův korelační koeficient (kapitola 4), ale v případě, že výstupem měření (= druhý sloupec tabulky dat) jsou nikoli absolutní hodnoty, ale pořadí, vhodnější je ρ = Spearmanův koeficient korelace mezi pořadími:

$$\rho = 1 - \frac{6 \cdot \sum d_i^2}{N(N^2-1)}.$$

V našem příkladu je $\rho = 0{,}915$, což je dost vysoká korelace – při jejím statistickém posouzení je dále ještě možné použít test například s využitím speciální tabulky kritických hodnot, kterou poskytl pan Spearman, tabulky 7.9. Podle této tabulky pro $N = 10$ a $\alpha = 0{,}05$ dostaneme $\rho_k = 0{,}65$, tj. protože $\rho = 0{,}915 \geq \rho_k = 0{,}65$, tak korelace je významná i statisticky.

Tab. 7.8: Kritické hodnoty Friedmanova testu.

N	$J = 3$ 7 N	$J = 4$	N	$J = 5$	N	$J = 6$	
3	18	2	20	2	38	2	64
4	26	3	37	3	64	3	103,5
5	32	4	52	4	88		
6	42	5	65	5	112		
7	50	6	76	6	136		
8	50	7	91				
9	56	8	102				
10	62	9	115				
11	72	10	128				
12	78						
13	86						
14	86						
15	96						
16	104						
17	104						
18	114						
19	122						
20	126						
21	128						
22	134						
23	136						
24	150						
25	152						
	$R_1 = 30$		$R_2 = 24$		$R_3 = 38$		$R_4 = 18$

Tab. 7.9: Kritické hodnoty pro Spearmanův test významnosti korelačního koeficientu ρ.

	$\alpha =$				
N	0,20	0,10	0,05	0,02	0,01
4		1,00			
5	0,80	0,90		1,00	
6	0,66	0,83	0,89	0,94	1,00
7	0,57	0,71	0,79	0,89	0,93
8	0,52	0,64	0,74	0,83	0,88
9	0,48	0,60	0,68	0,78	0,83
10	0,45	0,56	0,65	0,73	0,79
11	0,41	0,52	0,61	0,71	0,77
12	0,39	0,50	0,59	0,68	0,75
13	0,37	0,47	0,56	0,65	0,71
14	0,36	0,46	0,54	0,63	0,69
15	0,34	0,44	0,52	0,60	0,66
16	0,33	0,42	0,51	0,58	0,64
17	0,32	0,41	0,49	0,57	0,62
18	0,31	0,40	0,48	0,55	0,61
19	0,30	0,39	0,46	0,54	0,60
20	0,29	0,38	0,45	0,53	0,58
21	0,29	0,37	0,44	0,51	0,56
22	0,28	0,36	0,43	0,50	0,55
23	0,27	0,35	0,42	0,49	0,54
24	0,27	0,34	0,41	0,48	0,53
25	0,26	0,34	0,40	0,47	0,52
26	0,26	0,33	0,39	0,46	0,51
27	0,25	0,32	0,38	0,45	0,50
28	0,25	0,32	0,38	0,44	0,49
29	0,24	0,31	0,47	0,44	0,48
30	0,24	0,31	0,36	0,43	0,47

Pojmy k zapamatování

- Tato kapitola oproti většině předchozích kapitol představuje tzv. neparametrické statistické testy. U parametrických testů byl obyčejně nějaký parametr rozdělení neznámý, a ten byl podroben danému testu (např. μ, σ). Nyní u neparametrických testů žádný takový parametr není u daného testu k dispozici – odtud název „neparametrické“.
- Parametrické testy většinou lépe a rychleji vyvrátí hypotézu H_0 (řečeno obratem z kapitol 13 a 14 textu BMA3, mají větší sílu) – což je jejich cílem, pokud skutečně H_0 má být zamítnuta. Když ovšem nejsou splněny tři důležité předpoklady POUŽITELNOSTI těchto testů (viz úvodní odstavec této kapitoly 7), máme k dispozici pouze testy neparametrické.
- **Tabulka 7.5 uvádí přehled statistických testů neparametrických i parametrických použitých v tomto textu, společně s přehledem situací a vhodností jejich použitelnosti**.

Kontrolní otázky

1. U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.
   a) Pokud u t-testu není splněn předpoklad homogenního (= stejného) rozptylu, máme k dispozici neparametrický Kruskalův–Wallisův test.
   b) Mannův–Whitneyův test a Kruskalův–Wallisův test používají podobnou filosofii – místo jednotlivých hodnot měření zpracovávají jen pořadí těchto měření v souboru uspořádaném podle velikosti hodnot.
   c) Pokud $W < W_k$ u Wilcoxonova testu, tak H_0 nezamítáme.
   d) Neparametrické testy mají obecně větší sílu (= větší schopnost správně zamítnout H_0, když neplatí).
   e) Kolmogorovův–Smirnovův test užívá ve svém kritériu distribuční funkce – jak teoretického, tak empirického rozdělení.
   f) U Friedmanova testu nás zajímá pořadí nikoli vzhledem k různým podmínkám, ale vzhledem k různým lidem v každé podmínce.
   g) Wilcoxonův test přiřazuje pořadí absolutním hodnotám rozdílů.
   h) Test plánovaného srovnání (viz oddíl 5.0.2) je neparametrický.

Odpovědi na otázky
1a) – N (správně je k dispozici Mannův–Whitneův test), 1b) – A, 1c) – N (Wilcoxonův test je výjimkou z většiny testů v tomto textu), 1d) – N (právě naopak, parametrické testy mají větší sílu, protože neparametrické testy užívají většinou spíše jen pořadí měřených hodnot něž přímo tyto hodnoty), 1e) – A, 1f) – N (je to naopak, právě u Friedmanova testu používáme pořadí „NAPŘÍČ“ – vzhledem k různým podmínkám u daného člověka), 1g) – A, 1h) – N (Nejsem si stoprocentně jistý, protože jsem nečetl všechny důkazy a odvození, ale plánované srovnání podle mne vychází z předpokladu normálního rozdělení původní populace; a i kdyby tato myšlenka nebyla správná, z povahy $H_0: \ r^2 = 1$ je vidět, že se testuje hodnota parametru – tedy parametrický test).

Cvičení

1. Je prováděn experiment, který má zjistit, že očekávání u člověka ovlivňuje výsledek experimentu. Každému z devíti náhodně vybraných učitelů Ústavu matematiky je přidělen jeden náhodně vybraný student. Čtyřek učitelům je neformálně řečeno, že jejich student je hloupý, pěti ostatním je řečeno, že jejich student je celkem inteligentní. Pak každý vyučující podrobil svého studenta testu ze statistiky, a pečlivě přitom pokládal otázky a zaznamenával počet chyb. Získala se data

   a) U studentů, kteří byli označeni za celkem inteligentní: 6, 8, 5, 12.

   b) U studentů, kteří byli označeni za hloupé: 10, 2, 14, 9, 4.

   Ověřte neparametrickým testem, zda tato data potvrzují hypotézu, že vyučující, který studenta předem (= a priori) považuje za chytrého, mu napočítá méně chyb.

2. Dělníky na stavbě zajímá otázka, zda mají více knih politici nebo psychologové. Navštíví kancelář několika z nich a spočítají všechny knihy na regálech. Určete pomocí neparametrického testu, zda existuje významný rozdíl v počtu knih u těchto dvou profesí.

   Politikové (9 lidí): 87, 72, 65, 54, 67, 76, 73, 82, 104.
   Psychologové (12 lidí): 131, 94, 77, 88, 116, 90, 87, 76, 95, 164, 127, 77.

3. Dva politologové se zabývají otázkou, jak se mění nálada příznivců ČSSD a ODS. Mají hypotézu, že voliči ODS jsou stále více odrazováni výstřelky své strany a jejich zapojení v regionálních i národních volbách klesá. Proto se zeptali tří členů ČSSD a čtyř členů ODS, jak často se účastnili voleb v posledních několika letech. Zjistili toto: Členové ČSSD ... 4, 3, 1; členové ODS ... 2, 1, 0, 0. Ověřte neparametrickým testem, že členové ODS volí méně než členové ČSSD.

4. Lékovědec Jan Zelený zkoumal vliv nového léku na pacienty trpící nespavostí. Náhodně vybraných patnáct pacientů rozdělil na tři skupiny po pěti, z nichž skupina 1 užívala lék A, skupina 2 lék B a skupina 3 neužívala žádný lék. Získal data vyjadřující průměrný počet hodin spánku každého z pacientů. Pak tato data seřadil podle pořadí od nejlepšího k nejhoršímu ve smyslu jejich schopnosti spát. Proveďte pro tato data neparametrický test:

lék A	lék B	žádný lék
3	1	2
6	5	4
11	9	7
12	10	8
15	14	13

Výsledky

ad 1. Použijeme Mannův–Whitneyův test:

hodnota	2	4	5	6	8	9	10	12	14
pořadí	1	2	3	4	5	6	7	8	9
skupina	H	H	I	I	I	H	H	I	H

Pak $U_1 = U_I = 20 + 15 - 25 = 10$, $U_2 = 20 + 10 - 20 = 10$, odtud $10 > U_k = 1$, tj. H_0 nezamítáme (inverzní vlastnosti pro zamítnutí na rozdíl od většiny testů v tomto textu), oba soubory se statisticky významně neliší (tzv. „haló efekt“) se tedy mezi vyučujícími UMAT nerozšířil statisticky významně.

ad 2. Použijeme opět Mannův–Whitneyův test: při určování pořadí např. dvě hodnoty 76 mají pořadí 6,5, dvě hodnoty 77 mají pořadí 8,5, apod.

$$U_1 = 9 \cdot 12 + \frac{12 \cdot 13}{2} - 171 = 15, \quad U_2 = 9 \cdot 12 + \frac{9 \cdot 10}{2} - 60 = 93,$$

tedy $U = 15 < 26 = U_k$ a ZAMÍTÁME H_0 (inverzní vlastnosti pro zamítnutí na rozdíl od většiny testů v tomto textu), mezi profesemi je významný rozdíl.

ad 3. Použijeme Mannův–Whitneův test:

hodnota	0	0	1	1	2	3	4
pořadí	1,5	1,5	3,5	3,5	5	6	7
skupina	ODS	ODS	ODS	ČSSD	ODS	ČSSD	ČSSD

Pak $U_1 = U_{ODS} 10{,}5$, $U_2 = U_{\text{ČSSD}} = 1{,}5$, tedy $U = 1{,}5$. Tabulková hodnota není k dispozici, proto použijeme aproximaci normálním rozdělením, dostaneme

$$\frac{1{,}5 - \frac{3 \cdot 4}{2}}{\sqrt{\frac{3 \cdot 4(3+4+1)}{12}}} = -1{,}591 \in (-1{,}96; +1{,}96),$$

tedy H_0 nezamítáme, rozdíly mezi stranami nebyly testem prokázány. Je vidět, že menší počet měření má malou výpovědní sílu.

ad 4. $H = 0{,}86$, což není statisticky významné, tj. jednotlivé skupiny se neliší.

8 Operační výzkum

Průvodce studiem

V této kapitole si nejdříve jsme nejdříve na různých příkladech z praxe ukážeme, že všechny se dají popsat pomocí lineárních funkcí a tedy že se dají řešit metodami lineárního programování.

Dále se budeme nejdříve zabývat grafickým řešením a následnou analýzou citlivosti, tj. tím, jak se změní řešení při změně některého ze vstupních parametrů úlohy.

Potom přejdeme k algebraickému řešení. Ukážeme si převedení úlohy na kanonický tvar pomocí pomocných proměnných a případné použití umělých proměnných. Seznámíte se se simplexovou metodou řešení úloh lineárního programování.

V závěru kapitoly se budeme věnovat některým úskalím simplexové metody, zejména degeneraci, neexistenci nebo nejednoznačnosti řešení.

Cíle

Po prostudování této kapitoly budete schopni:

- Slovně zadanou úlohu vyjádřit matematickými vztahy.
- Převést úlohu na kanonický tvar.
- Umět používat pomocné a umělé proměnné.
- Řešit úlohy lineárního programování.
- Provádět analýzu citlivosti.
- Chápat omezení a úskalí řešení úlok lineárního programování.

Co je to operační výzkum a kdy vznikl? Jak už název napovídá, jedná se o výzkum operací. Systematický útok na tuto oblast vědění byl zahájen v průběhu druhé světové války. Ano, a operace, o kterých se zde mluví, byly původně operace vojenského charakteru: taktické, organizační i zásobovací.

Metody operačního výzkumu tedy spojuje zejména oblast praxe, ve které vznikaly – spadá sem optimální rozdělování surovin, výrobků a pracovních sil, plánování projektů, úlohy zásobování, řešení problémů opotřebení a obnovy zařízení, otázky spojené s čekáním na obsluhu, výběr nejlepší strategie, stanovení harmonogramu činností, atd. Po válce se výzkum těchto otázek nezastavil, protože dané poznatky nacházely své uplatnění zejména v ekonomice, kde se setkáváme prakticky se stejnými situacemi (konec konců, ekonomika je také trochu válka).

Z teoretického hlediska jsou však problémy vzniklé na jednom bitevním poli řešeny v mnoha oblastech matematiky: problém optimálního rozdělení zdrojů a pracovních sil se převedl na hledání maxima nebo minima lineární funkce (disciplína: lineární programování) nebo nelineární funkce (disciplíny: matematické programování nebo nelineární programování). Plánování projektu našlo své řešení v úlohách síťové analýzy (teorie grafů). Volba optimální strategie podnítila rozvoj celého vědního oboru – teorie strategických her. Teorie pravděpodobnosti přinesla svou trochu do mlýna v otázkách front (nemyslí se fronty válečné, ale fronty v opravně, menze, bance, na maso nebo na mobil). Sekvenční, opakující se procesy vedly k rozvoji dynamického programování, řešícího úlohy užitím jakýchsi rekurzívních optimalizačních algoritmů.

Zkrátka a dobře, dnešnímu studentovi se předmět zvaný operační výzkum může jevit jako exkurze do mnoha různých oblastí matematiky. A skutečně, snad každá z matematických disciplín má k některému z uvedených problémů co říct. A přece existuje něco, co všechny tyto metody spojuje: snaha nalézt optimum – podle daných kritérií ty nejlepší hodnoty uvažovaných proměnných, nejlepší řešení studované situace.

Latinské „optimus“ znamená „nejlepší“. My ve svém vyjadřování nepoužíváme pouze slovo „optimální“ (= nejlepší), ale i „optimálnější“ (= více nejlepší) a „nejoptimálnější“ (= nejvíce nejlepší). A tato humorná situace, kterou nám připravila naše drahá čeština, dobře vystihuje, jací vlastně jsme: nestačí nám mít to nejlepší, ale hledáme ještě něco více. Přeji vám, abyste to nejlepší z nejlepších ve svém životě objevili.

8.1 Lineární programování

Předmětem lineárního programování je řešení úlohy, jejíž matematický tvar zní:

nalezněte minimum (nebo maximum) funkce $z = \sum_{j=1}^{n} c_j x_j$ za omezujících podmínek

$$x_j \geq 0 \text{ pro } j = 1, 2, \ldots, n \text{ (tzv. triviální podmínky)},$$

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i \text{ (nebo } = b_i \text{ nebo } \geq b_i) \text{ pro } i = 1, 2, \ldots, m.$$

Na úlohách lineárního programování není nejzajímavější matematická formulace nebo řešení (i když algoritmus řešení má krásný geometrický význam), ale to, kolik různých úloh z praxe lze na tento matematický model převést. Uveďme nyní několik typických představitelů z oblasti ekonomické praxe.

Příklad 8.1. Společnost RED vlastní závod na výrobu vnitřních a venkovních nátěrů domů. K výrobě se používá dvou základních surovin A, B. Maximální denní dostupnost suroviny A je 6 tun, suroviny B 8 tun. Na výrobu jedné tuny vnějšího nátěru je potřeba 1 tuna A a 2 tuny B, na výrobu jedné tuny vnitřního nátěru 2 tuny A a 1 tuna B. Průzkum trhu ukázal, že

a) denní výroba vnitřního nátěru nesmí překročit denní výrobu vnějšího nátěru o více než 1 tunu;

b) denní výroba vnitřního nátěru nesmí překročit 2 tuny.

Prodejní cena 1 tuny vnějšího nátěru je 3000 dolarů, vnitřního 2000 dolarů. Jaké množství obou nátěrů musí společnost vyrábět, aby byl její obrat maximální?

Matematická formulace úlohy:

1. označíme proměnné:
   x ... denní výroba vnějšího nátěru (v tunách)
   y ... denní výroba vnitřního nátěru
2. budeme hledat maximum funkce $z = 3x + 2y$, která vyjadřuje denní obrat (v tisících dolarů).
3. omezující podmínky:
   - omezení na spotřebu suroviny A:
     na výrobu vnějšího nátěru se spotřebuje x tun suroviny A denně
     na výrobu vnitřního nátěru se spotřebuje $2y$ tun suroviny A denně
     maximální dostupnost suroviny A je 6 tun denně
     celkem tedy dostáváme omezení: **1.** $x + 2y \leq 6$
   - omezení na spotřebu suroviny B: **2.** $2x + y \leq 8$
   - ad průzkum trhu a): **3.** $y - x \leq 1$
   - ad průzkum trhu b): **4.** $y \leq 2$
   - triviální omezení, chceme vyrobit kladné množství obou nátěrů: **5.** $x \geq 0$
     6. $y \geq 0$.

Příklad 8.2. Politika bankovních půjček. Banka přemýšlí o různých typech klientů. Z dřívější zkušenosti je známa následující tabulka:

typ půjčky	úroky	pst zadlužení
osobní	0,14	0,1
automobilová	0,13	0,07
domácí	0,12	0,03
zemědělská	0,125	0,05
obchodní	0,1	0,02

Zadlužení nevykazuje žádný úrokový zisk. Konkurence jiných peněžních institucí v regionu vyžaduje, aby banka přidělila alespoň 40% z vybraných 12 miliónů na zemědělské a obchodní půjčky. Dále domácí půjčky musí nabýt objemu aspoň jako automobilové a osobní půjčky dohromady. Celková pravděpodobnost zadlužení nesmí přesáhnout 0,04. Jaké je optimální rozdělení daných 12 miliónů?

Matematická formulace úlohy:

1. označíme proměnné:
   x_1 ... osobní půjčky (v miliónech)
   x_2 ... automobilové půjčky
   x_3 ... půjčky domácnostem
   x_4 ... zemědělské půjčky
   x_5 ... obchodní půjčky
2. budeme maximalizovat rozdíl „zisk z úroků“ minus „ztracené fondy“, tj.

$$\begin{aligned} z = (&0,14 \cdot 0,9\, x_1 + 0,13 \cdot 0,93\, x_2 + 0,12 \cdot 0,97\, x_3 + 0,125 \cdot 0,95\, x_4 \\ &+ 0,1 \cdot 0,98\, x_5) - 0,1\, x_1 - 0,07\, x_2 - 0,03\, x_3 - 0,05\, x_4 - 0,02\, x_5. \end{aligned}$$

Po úpravě dostaneme

$$z = 0,026\,x_1 + 0,0509\,x_2 + 0,0864\,x_3 + 0,06875\,x_4 + 0,078\,x_5.$$

3. omezující podmínky:
   - banka na půjčky vyhradila 12 miliónů: $x_1 + x_2 + x_3 + x_4 + x_5 \leq 12$
   - aspoň 40% půjček musí být zemědělské nebo obchodní: $x_4 + x_5 \geq 0,4 \cdot 12$
   - omezení na půjčky domácnostem: $x_3 \geq x_1 + x_2$
   - omezení celkového zadlužení:
     $$\frac{0,1\,x_1 + 0,07\,x_2 + 0,03\,x_3 + 0,05\,x_4 + 0,02\,x_5}{x_1 + x_2 + x_3 + x_4 + x_5} \leq 0,04$$
     (aby toto omezení bylo lineárním, musíme je zbavit zlomku)
   - triviální omezení: $x_i \geq 0$ pro $i = 1, \ldots, 5$.

Matematická formulace problému je opět úlohou nalezení maxima z lineární funkce za lineárních omezení, tj. úlohou lineárního programování.

Příklad 8.3. Pozemkové hospodaření. Společnost BIRD vlastní 800 hektarů půdy u jezera Ozark. V dané oblasti jsou vydána následující bezpečnostní opatření, aby se zamezilo znečištění vody:

1) Mohou zde být stavěny jen domy pro jednu až tři rodiny, přičemž domů pro jednu rodinu musí být aspoň 50%.

2) Aby byl omezen počet odpadních nádrží, je vyžadována minimální rozloha

   2 hektary na dům pro 1 rodinu
   3 hektary na dům pro 2 rodiny
   4 hektary na dům pro 3 rodiny

3) Na 200 rodin musí být vyhrazena rekreační plocha 1 hektaru

4) Podzemní voda nesmí být získávána za účelem použití v rodinách nebo na zahradě.

Odhaduje se, že 15% plochy z daných 800 hektarů bude využito k budování ulic a přístupových cest. Zisk u jednotlivých typů domů je:

dům pro 1 rodinu ... 10 000 \$
dům pro 2 rodiny ... 15 000 \$
dům pro 3 rodiny ... 20 000 \$

Cena přívodu vody je přímo úměrná počtu postavených domů, ale minimálně se jedná o 100 000 \$. Dále spotřeba vody přiváděné novým přívodem je omezena na 200 000 galonů denně. Následující tabulka shrnuje údaje ceny a spotřeby vody:

typ domu	1 rodina	2 rodiny	3 rodiny	rekreační plocha
cena přívodu ($)	1 000	1 200	1 400	800
spotřeba vody (galony/den)	400	600	840	450

Jak využít pozemek, aby zisk byl maximální?

Matematická formulace úlohy:

1. označíme proměnné:
   x_1 ... počet postavených domů pro 1 rodinu
   x_2 ... počet postavených domů pro 2 rodiny
   x_3 ... počet postavených domů pro 3 rodiny
   x_4 ... počet rekreačních ploch

2. budeme hledat maximum funkce $z = 10\,000\,x_1 + 15\,000\,x_2 + 20\,000\,x_3$.

3. omezující podmínky:

   – omezení počtu domů: $2\,x_1 + 3\,x_2 + 4\,x_3 + x_4 \leq 680$

   – domů pro 1 rodinu je aspoň 50%: $x_1 \geq x_2 + x_3$

   – minimální počet rekreačních zařízení: $x_4 \geq \frac{x_1+2\,x_2+3\,x_3}{200}$

   – přípojky by měly pokrýt pořizující cenu přívodu vody:

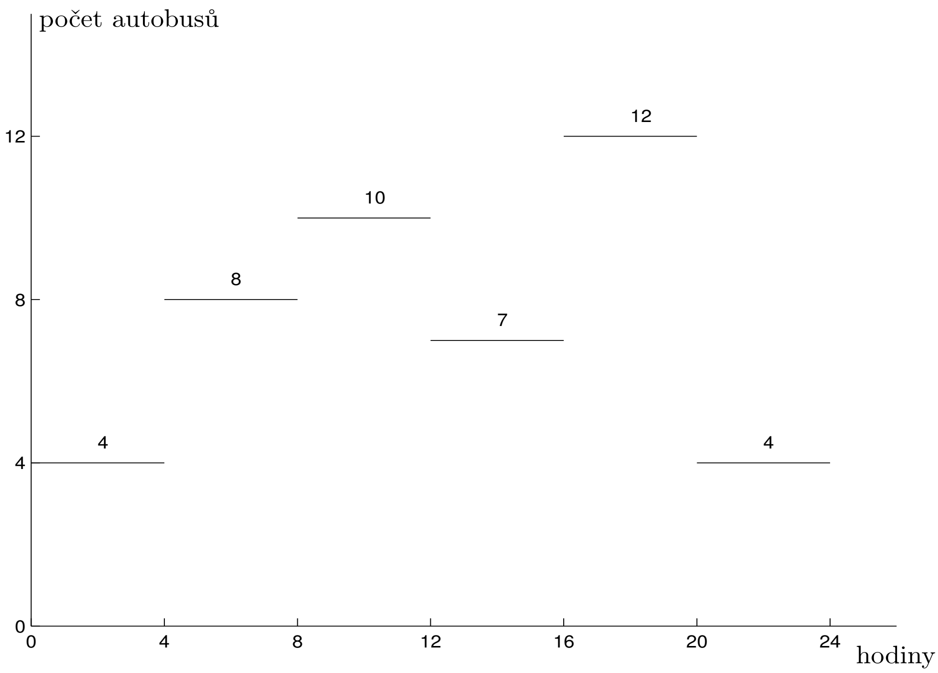
   $$1\,000\,x_1 + 1\,200\,x_2 + 1\,400\,x_3 + 800\,x_4 \leq 100\,000$$

   – omezení na spotřebu vody:

   $$400\,x_1 + 600\,x_2 + 840\,x_3 + 450\,x_4 \leq 200\,000$$

   – triviální omezení: $x_i \geq 0$ pro $i = 1, \ldots, 4$.

Příklad 8.4. Jízdní řád autobusů. Progresívní dopravní podnik studuje autobusovou síť ve městě. Bylo zjištěno, kolik autobusů během dne je třeba (počet autobusů je stejný vždy ve čtyřhodinovém časovém úseku):



Vytvořte rozvrh využití autobusů tak, aby každý autobus pracoval 8 hodin po sobě, a pak měl 16 hodin volno.

Matematická formulace úlohy:

1. proměnné x_i ... počet autobusů pracujících během různých směn

2. budeme hledat minimum funkce $z = \sum x_i$ vyjadřující celkový počet autobusů v prostoru.

3. zadání je nejasné: nevíme, kolik směn by mělo být a kdy začínají.
   Při třísměnném provozu:

$$\left.\begin{array}{llcl} \text{1. směna} & 8^{00} - 16^{00} & \dots & x_1 \geq 10 \\ \text{2. směna} & 16^{00} - 24^{00} & \dots & x_2 \geq 12 \\ \text{3. směna} & 24^{00} - \ \ 8^{00} & \dots & x_3 \geq \ \ 8 \end{array}\right\} \Rightarrow z = x_1 + x_2 + x_3 \geq 30.$$

   Lepšího výsledku lze dosáhnout v šestisměnném provozu:

$$\begin{array}{llcl} \text{1. směna} & 0^{00} - \ \ 8^{00} & \dots & x_1 \\ \text{2. směna} & 4^{00} - 12^{00} & \dots & x_2 \\ \text{3. směna} & 8^{00} - 16^{00} & \dots & x_3 \\ \text{4. směna} & 12^{00} - 20^{00} & \dots & x_4 \\ \text{5. směna} & 16^{00} - 24^{00} & \dots & x_5 \\ \text{6. směna} & 20^{00} - \ \ 4^{00} & \dots & x_6 \end{array}$$

tj. minimalizujeme funkce $z = \sum_{i=1}^{6} x_i$ za omezení

$$\begin{aligned} x_1 + x_6 &\geq 4 \\ x_1 + x_2 &\geq 8 \\ x_2 + x_3 &\geq 10 \\ x_3 + x_4 &\geq 7 \\ x_4 + x_5 &\geq 12 \\ x_5 + x_6 &\geq 4 \\ x_i &\geq 0 \quad , \; i = 1, \dots, 6. \end{aligned}$$

Tato úloha má řešení $(0, 10, 0, 12, 0, 4)$, tj. minimální počet autobusů je 26.

8.2 Grafické řešení úlohy lineárního programování

Zabývejme se nyní nejprve grafickým řešením našeho matematického modelu. Z geometrického významu řešení úlohy lze totiž odvodit obecný algebraický postup řešení. Vše bude vysvětleno na prvním zformulovaném příkladu výroby dvou nátěrů.

Ad Příklad 8.1. Nalezněte maximum funkce $z = 3\,x + 2\,y$ za omezení

$$\begin{array}{rr} \mathbf{1}. & x + 2\,y \leq 6 \\ \mathbf{2}. & 2\,x + y \leq 8 \\ \mathbf{3}. & -x + y \leq 1 \\ \mathbf{4}. & y \leq 2 \\ \mathbf{5}. & x \geq 0 \\ \mathbf{6}. & y \geq 0 \end{array}$$

Řešení. Jedná se o úlohu nalezení globálního extrému funkce z na množině přípustných hodnot (označme ji M) zadané šesti omezeními. Každé omezení jednoznačně určuje jednu polorovinu. Všechna omezení musí platit současně, tj. množina M je průnikem šesti polorovin (viz obr.8.1):

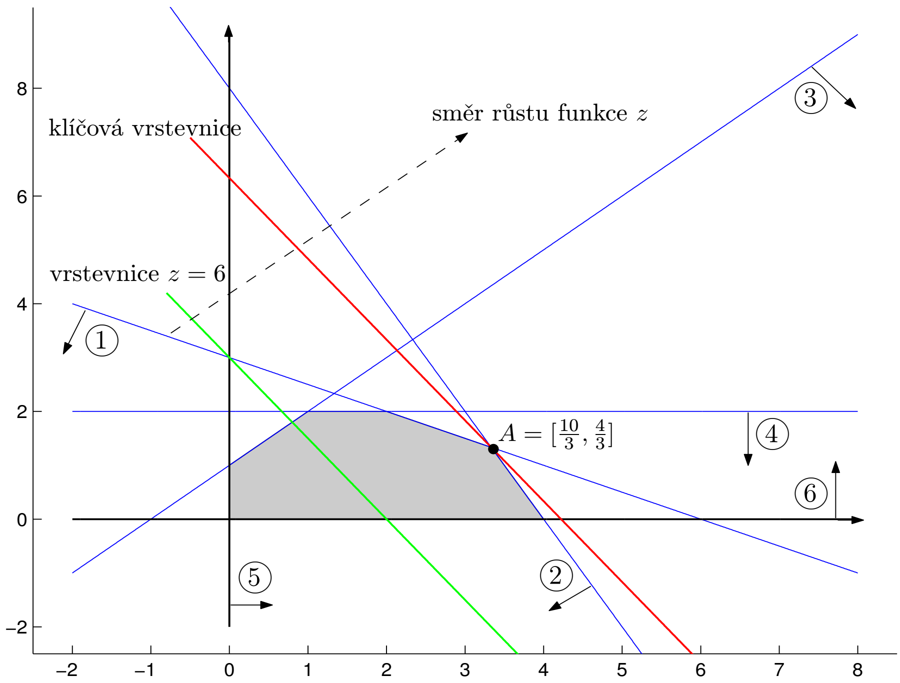
Studujme nyní vrstevnice funkce z na množině M:
Vrstevnicemi jsou křivky tvaru $3\,x + 2\,y = d$, kde d je hodnota vrstevnice; tj. rovnoběžné přímky $y = -\frac{3}{2} + \frac{d}{2}$. Pokud libovolnou z těchto vrstevnic posunujeme ve směru růstu funkce z (= směru růstu konstanty d) kolmém na všechny vrstevnice, dojdeme k optimálnímu řešení, kterým je poslední neprázdný průnik vrstevnice s maximální hodnotou na M a množiny M.
V našem případě je optimum v bodě A, který získáme jako průsečík přímek **1** a **2** . Optimální hodnota funkce z v bodě A je

$$z(A) = 3 \cdot \frac{10}{3} + 2 \cdot \frac{4}{3} = \frac{38}{3}.$$

Poznámka 8.5. Protože grafem funkce z je rovina, optimum musí ležet v některém vrcholu množiny přípustných hodnot (nebo ve více vrcholech, pokud klíčová vrstevnice prochází více vrcholy).

Optimální řešení neexistuje, pokud

- množina M je prázdná
- množina M je neohraničená ve směru růstu funkce z.



Obr. 8.1: Grafické znázornění řešení Příkladu 8.1.

Netriviální omezení se nazývají

– **klíčová** ... pokud prochází bodem optima

– **neklíčová** ... pokud neprochází bodem optima.

8.3 Analýza citlivosti na základě grafického náhledu

Zabývejme se nyní chvíli analýzou citlivosti, tj. tím, jak se změní řešení při změně některého ze vstupních parametrů úlohy. Podle toho, jaký parametr se mění, odpovězme u Příkladu 8.1 na následující otázky:

a) Jak moc má smysl zvyšovat pravou stranu klíčových omezení, aby se zlepšovala optimální hodnota funkce z?

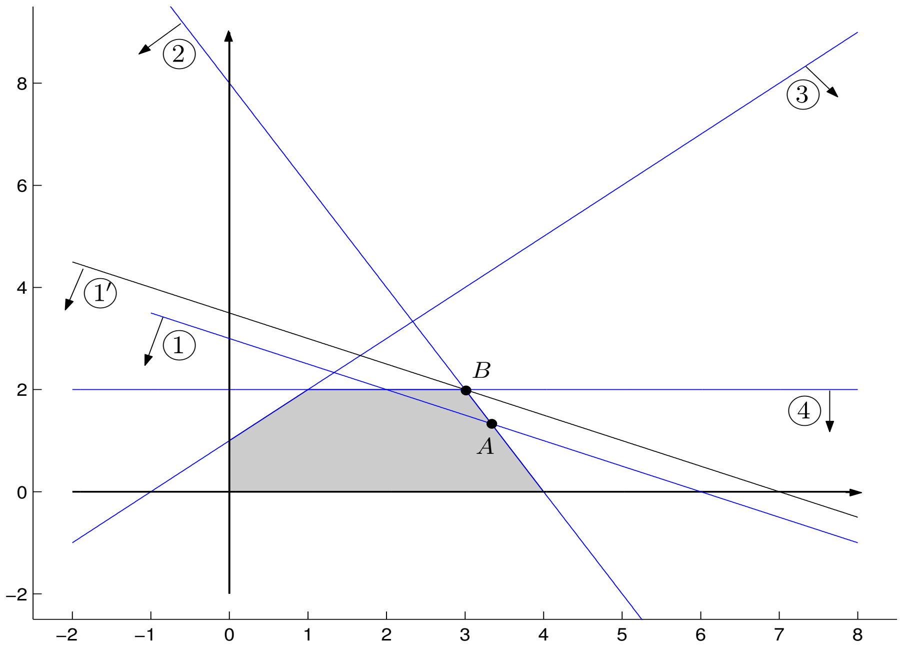
b) Jak moc má smysl snížit pravou stranu neklíčových omezení, aby hodnota optima zůstala zachována?

c) Pokud bychom chtěli zajistit zvýšení pravé strany některého z klíčových omezení, které z nich má největší prioritu?

d) Jak moc můžeme měnit koeficienty funkce z, aby bod optima zůstal zachován? (funkční hodnota v bodě optima se samozřejmě změní)

Ad Příklad 8.1.

ad a) Klíčová omezení jsou **1** a **2** .
Omezení **1** : $x + 2\,y \le 6$
Graficky je tato situace znázorněna na obr.8.2. Pokud posunujeme přímku **1** ve směru růstu funkce z, při posunu až do **1'** se toto omezení stává nadbytečným, neboť množina přípustných hodnot se nemění přidáním nebo odebráním omezení **1'** . Posunovat **1** dále než do **1'** tedy nemá smysl.



Obr. 8.2: Grafické znázornění posunutí přímky klíčového omezení.

Nahradíme-li tedy v našem případě omezení **1** omezením **1'** , optimum bude nyní v bodě B, což je průsečík omezení **2** a **4** :

$$B = [3; 2] \Rightarrow z(B) = 3 \cdot 3 + 2 \cdot 2 = 13$$

$$B \in \mathbf{1}' \Rightarrow \mathbf{1}' : x + 2\,y \le 7$$

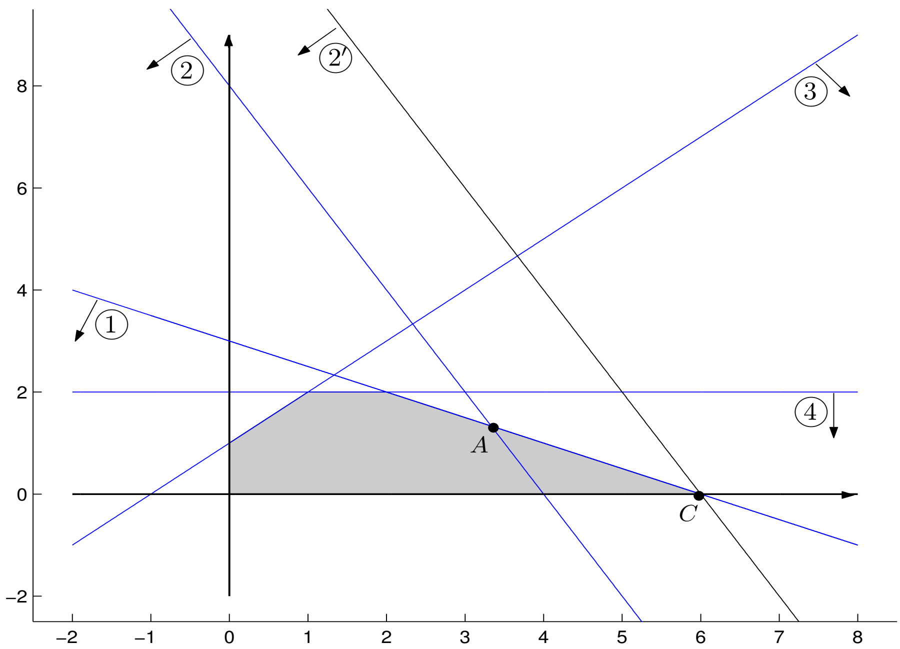
Omezení **2** : $2\,x + y \le 8$
Graficky je tato situace znázorněna na obr.8.3. Přímku **2** má smysl posunovat až do **2'** , kdy se omezení **2'** stává nadbytečným (jeho přidáním nebo odebráním se nemění množina přípustných hodnot).
Nahradíme-li v našem případě omezení **2** omezením **2'** , optimum nastane v bodě C, což je průsečík omezení **1** a **6** :

$$C = [6; 0] \Rightarrow z(C) = 3 \cdot 6 + 2 \cdot 0 = 18$$

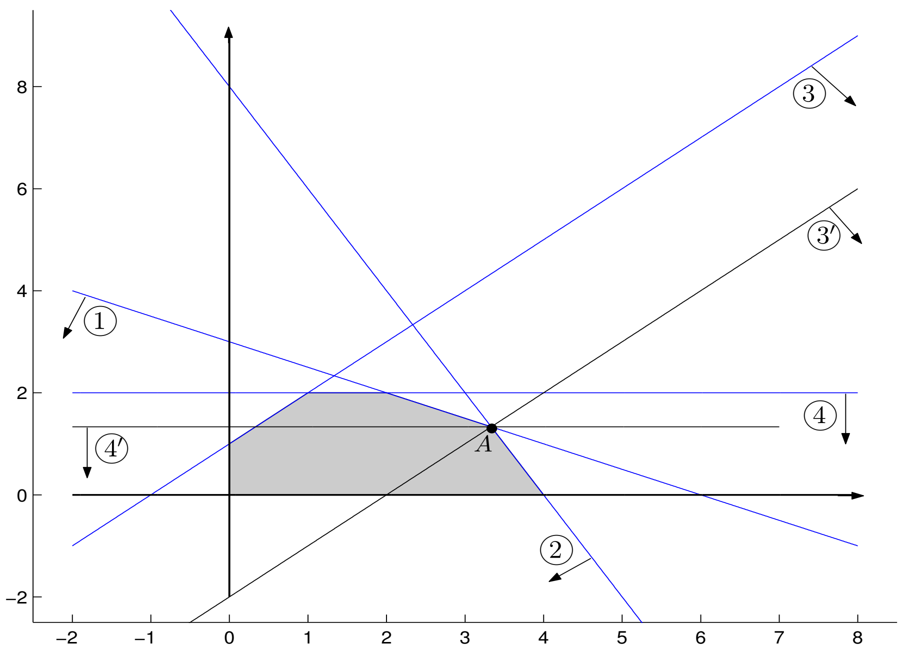
$$C \in \mathbf{2}' \Rightarrow \mathbf{2}' : 2\,x + y \le 12$$

ad b) Neklíčová netriviální omezení jsou **3** a **4** . Protože neprochází bodem optima A původní úlohy, můžeme jejich pravou stranu snižovat, aniž by se bod optima změnil. Tato situace je znázorněna na obr.8.4.
Omezení **3** lze měnit až na **3'** , tj.

$$A \in \mathbf{3}' \Rightarrow \mathbf{3}' : -x + y \le -2.$$



Obr. 8.3: Grafické znázornění posunutí přímky klíčového omezení.



Obr. 8.4: Grafické znázornění posunutí přímek neklíčových omezení.

Omezení **4** lze měnit až na **4’** , tj.

$$A \in \mathbf{4'} \Rightarrow \mathbf{4'} : y \leq \frac{4}{3}.$$

V obou případech změny zůstává optimum v bodě $A = [\frac{10}{3}, \frac{4}{3}]$, tj. $z(A) = \frac{38}{3}$ je optimální funkční hodnota.

ad c) Shrňme modifikace netriviálních omezení do tabulky:

omezení	pravá strana se změní o ...	změna funkce z	jednotková změna z
1	$7-6=1$ (bod optima ... B)	$z(B)-z(A)=\frac{1}{3}$	$\frac{1/3}{1}=\frac{1}{3}$
2	$12-8=4$ (bod optima ... C)	$z(C)-z(A)=\frac{16}{3}$	$\frac{16/3}{4}=\frac{4}{3}$
3	$-2-1=-3$ (bod optima ... A)	$z(A)-z(A)=0$	$\frac{0}{-3}=0$
4	$\frac{4}{3}-2=-\frac{2}{3}$ (bod optima ... A)	$z(A)-z(A)=0$	$\frac{0}{-2/3}=0$

Údaje uvedené v posledním sloupci tabulky jsou tzv. *stínové ceny*, tj. stínová cena příslušná danému omezení = změna funkce z při jednotkovém zvýšení pravé strany klíčového omezení a nebo jednotkovém snížení pravé strany neklíčového omezení.

Stínové ceny neklíčových omezení jsou vždy nulové. Maximální stínová cena určuje omezení s největší prioritou změny pravé strany, tj. omezení **2** má největší prioritu změny pravé strany.

ad d) Odpovězme např. na otázku, jaká změna koeficientu a funkce $z = a\,x + 2\,y$ ještě zachová bod optima A?

Odpověď lze opět odvodit z geometrického názoru: Aby bod A zůstal bodem optima, vrstevnice funkce z musí mít sklon mezi dvěma mezními hodnotami určenými sklony přímek **1** a **2** :

$$\left.\begin{array}{rl} \text{přímka } \mathbf{1}: & y = -\frac{1}{2}x + 3 \\ \text{vrstevnice funkce } z: & y = -\frac{a}{2}x + \frac{d}{2} \\ \text{přímka } \mathbf{2}: & y = -\frac{4}{2}x + 8 \end{array}\right\} \Rightarrow \text{porovnáním koeficientů u } x \text{ máme:}$$

$$-\frac{a}{2} \in \left\langle -\frac{1}{2}; -\frac{4}{2} \right\rangle \Rightarrow a \in \langle 1; 4 \rangle \,.$$

Tj. bodem optima zůstane A, pokud cena venkovního nátěru se pohybuje v rozmezí od 1 do 4 tisíc dolarů.

Z příkladu je vidět, že analýza citlivosti je stejně důležitá jako řešení původní úlohy.

8.4 Algebraické řešení úlohy lineárního programování – simplexová metoda

Mezi všemi formulacemi úlohy lineárního programování existuje jeden přesný, tzv. *kanonický tvar*, který je jediný vhodný pro použití algebraického řešení. Pokud úloha není v kanonickém tvaru, musíme ji na tento tvar převést.

Tak tedy, kanonický tvar je charakteristický tím, že

- všechna omezení jsou rovnicemi
- všechna omezení mají nezápornou pravou stranu

– pro všechny proměnné x_j platí: $x_j \geq 0$

Příklad 8.6. Následující úlohu převeďte na kanonický tvar:

nalezněte minimum funkce $z = 2\,x_1 + 3\,x_2$ za omezení

$$
\begin{aligned}
x_1 + x_2 &= 10 \\
-2\,x_1 + 3\,x_2 &\leq -5 \\
7\,x_1 - 4\,x_2 &\leq \ 6 \\
x_1 &\in \mathbb{R} \\
x_2 &\geq \ 0
\end{aligned}
$$

Řešení. Zápornou pravou stranu omezení jednoduše odstraníme vynásobením nerovnosti číslem (-1).

Nerovnost převedeme na rovnici tak, že v případě $\leq$ k levé straně přičteme, v případě $\geq$ od levé strany odečteme, novou nezápornou proměnnou p.

Posledním problémem je nahradit neohraničenou proměnnou $x_1 \in \mathbb{R}$ nějakými nezápornými proměnnými. Toho lze docílit např. substitucí $x_1 = x_1' - x_1''$, kde $x_1' \geq 0,\ x_1'' \geq 0$. Libovolné reálné číslo lze skutečně vyjádřit jako rozdíl dvou nezáporných čísel, např.:

$$
\begin{aligned}
x_1 = 2 \quad &\Rightarrow \quad x_1' = 2,\ x_1'' = 0 \\
x_1 = -3 \quad &\Rightarrow \quad x_1' = 0,\ x_1'' = 3.
\end{aligned}
$$

Kanonický tvar úlohy tedy bude:

nalezněte minimum funkce $z = 2\,(x_1' - x_1'') + 3\,x_2$ za omezení

$$
\begin{aligned}
x_1' - x_1'' + x_2 &= 10 \\
2\,x_1' - 2\,x_1'' - 3\,x_2 - p_1 &= \ 5 \\
7\,x_1' - 7\,x_1'' - 4\,x_2 + p_2 &= \ 6
\end{aligned}
$$

$$
x_1' \geq 0,\ x_1'' \geq 0,\ x_2 \geq 0,\ p_1 \geq 0,\ p_2 \geq 0.
$$

Převod na kanonický tvar je sympatický v tom, že zachovává řešení původní úlohy. Tedy řešení úlohy v kanonickém tvaru je stejné jako řešení původní úlohy (omezíme-li se na původní proměnné).

Při algebraickém řešení se využívá toho faktu, že optimum funkce z musí nastat v některém z vrcholů množiny přípustných hodnot. V algoritmu řešení tedy najdeme nějaký libovolný vrchol množiny přípustných hodnot. Potom vybereme ten z jeho sousedních vrcholů, ve kterém nastane největší „zlepšení“ funkce z („zlepšení“ = zvýšení při maximalizaci a snížení při minimalizaci) a přesuneme se do něj. Proces výběru sousedního vrcholu, který „zlepšuje“ funkci z, opakujeme tak dlouho, dokud to jde. Když všechno dobře dopadne, jsme na konci procesu v optimálním vrcholu.

Ještě poznámka k názvu metody: pokud množina přípustných hodnot je ohraničená a počet jejích vrcholů je o 1 větší než její dimenze, nazýváme ji simplex. Odtud i název simplexová metoda, neboť se v daném kroku pohybujeme mezi vrcholy určitého simplexu.

Ad Příklad 8.1 Celou metodu vysvětlíme na našem příkladu výroby nátěrů, jejíž kanonický tvar je:

nalezněte maximum funkce $z = 3x + 2y$ za omezení

$$\begin{array}{rcl} x + 2y + p_1 & & = 6 \\ 2x + \ \ y \qquad + p_2 & & = 8 \\ -x + \ \ y \qquad\qquad + p_3 & & = 1 \\ y \qquad\qquad\qquad + p_4 & & = 2 \end{array}$$

$$x,\ y,\ p_1,\ p_2,\ p_3,\ p_4 \geq 0.$$

Řešení. Kanonický tvar úlohy přepíšeme do tzv. *simplexové tabulky*:

	x	y	p_1	p_2	p_3	p_4	řešení
z	-3	-2	0	0	0	0	$= 0$
p_1	1	2	1	0	0	0	6
p_2	2	1	0	1	0	0	8
p_3	-1	1	0	0	1	0	1
p_4	0	1	0	0	0	1	2

Vysvětlení k tabulce:
V prvním řádku tabulky jsou označeny sloupce příslušné jednotlivým proměnným. Ve druhém řádku jsou zapsány koeficienty rovnice $z - 3x - 2y = 0$, je třeba si jen uvědomit tvar funkce z se všemi proměnnými

$$z = 3x + 2y + 0p_1 + 0p_2 + 0p_3 + 0p_4.$$

V dalších řádcích jsou zapsány koeficienty jednotlivých omezujících rovnic. V našem případě omezení tvoří systém 4 rovnic o 6 neznámých. Najdeme jisté jeho řešení následovně:
Vybereme 4 proměnné, tzv. *bázické* proměnné … v našem případě p_1, p_2, p_3, p_4.
Ostatní, tzv. *nebázické* proměnné, položíme rovny 0, tj. $x = 0$, $y = 0$. Tím vznikne systém 4 rovnic o 4 neznámých p_1, p_2, p_3, p_4, který má jediné řešení. Protože příslušná matice systému je jednotková, vektor pravých stran je přímo vektorem řešení, tj.

$$p_1 = 6,\ p_2 = 8,\ p_3 = 1,\ p_4 = 2.$$

Tímto způsobem jsme našli souřadnice výchozího vrcholu simplexu

$$\boldsymbol{x}^0 = (0, 0, 6, 8, 1, 2).$$

Přechod na sousední vrchol simplexu lze algebraicky zařídit tak, že jednu z nebázických proměnných zaměníme za jednu z bázických a pak postup opakujeme, nebázické proměnné položíme rovny 0, zbylý systém s bázickými proměnnými vyřešíme.
Poznámka k terminologii:

- *přípustné řešení* … libovolný bod množiny přípustných hodnot (nemusí být vrchol), jehož všechny souřadnice jsou nezáporné
- *přípustné bázické řešení* … takové přípustné řešení, které odpovídá některému vrcholu množiny přípustných hodnot.

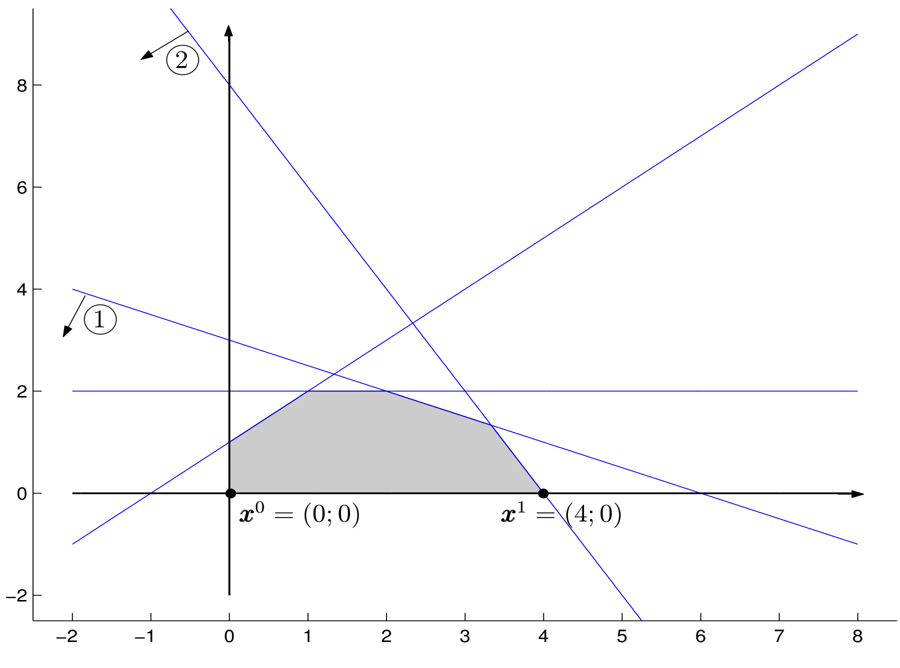
Realizujme nyní přechod do sousedního vrcholu, který nejvíce „zlepší“ hodnotu funkce z:

a) Za vstupní proměnnou vybereme x, protože příslušný koeficient v řádku funkce z je nejvíce záporný, tj. při převodu na druhou stranu rovnice nejvíce zvýší hodnotu funkce z. Pokud už žádné číslo v řádku z není záporné, dané přípustné řešení je už optimální.

b) Proměnná x určuje sloupec, ve kterém vybereme kladné hodnoty. K těmto hodnotám vypočteme tzv. kladné změny neboli podíly (příslušná pravá strana)/(hodnota): $\frac{6}{1}$ a $\frac{8}{2}$. Minimální z těchto kladných změn určuje řádek, jehož proměnná je výstupní: v našem případě minimální kladná změna je $\frac{8}{2}$, tj. p_2 je výstupní proměnná. Prvek na průsečíku vstupního sloupce a výstupního řádku se nazývá *pivotový prvek*.

	x ↓	y	p_1	p_2	p_3	p_4	řešení
z	-3	-2	0	0	0	0	0
p_1	**1**	2	1	0	0	0	**6**
← p_2	**2**	1	0	1	0	0	**8**
p_3	-1	1	0	0	1	0	1
p_4	0	1	0	0	0	1	2

Důvod výběru řádku s minimální kladnou změnou:



Obr. 8.5: Grafické znázornění přechodu mezi dvěma vrcholy.

Náš výchozí vrchol je $\boldsymbol{x}^0 = (0;0)$. Vstupní sloupec určuje, že největší nárůst funkční hodnoty funkce z nastane pro proměnnou x, tj. budeme se pohybovat ve směru růstu proměnné x (= v kladném směru osy x). Následující vrchol množiny přípustných hodnot v tomto směru $\boldsymbol{x}^1 = (4;0)$ je dán prvním nejbližším omezením, které protne osu x, což je omezení **2** s kladnou změnou 4, nikoli omezení **1** s kladnou změnou 6 (bod $(6;0)$ není vrchol M). Minimální kladná změna (= 4) vyjadřuje změnu souřadnice x při přechodu z vrcholu $\boldsymbol{x}^0$ do $\boldsymbol{x}^1$.

Nakonec ještě poznamenejme, že kladná změna nemusí obecně znamenat vzrůst x-ové souřadnice; kladný směr = směr růstu funkce z při maximalizaci, resp. směr klesání z při minimalizaci.

Nyní eliminačními úpravami zajistíme, aby se na místě pivotového prvku objevila hodnota 1 a zbylé hodnoty vstupního sloupce byly 0 (včetně koeficientu v řádku z).
POZOR! K danému řádku lze přičítat násobek pouze pivotového (= výstupního) řádku, žádného dalšího! Tímto dodatečným omezením se totiž docílí důležité vlastnosti, že sloupce odpovídající novým bázickým proměnným budou opět vytvářet jednotkovou matici (i když jejich pořadí bude přeházené) a navíc bude splněna důležitá podmínka (která musí být zaručena v každém kroku simplexové tabulky), že koeficienty v řádku z příslušné bázickým proměnným jsou rovny 0:

	x	y	p_1	p_2	p_3	p_4	řešení
z	0	$-\frac{1}{2}$	0	$\frac{3}{2}$	0	0	12
p_1	0	$\boldsymbol{\frac{3}{2}}$	1	$-\frac{1}{2}$	0	0	**2**
x	1	$\boldsymbol{\frac{1}{2}}$	0	$\frac{1}{2}$	0	0	**4**
p_3	0	$\boldsymbol{\frac{3}{2}}$	0	$\frac{1}{2}$	1	0	**5**
p_4	0	**1**	0	0	0	1	**2**

V posledním sloupci řádku pro funkci z je uvedena její nová hodnota (ve vrcholu $\boldsymbol{x}^1$). Ta se zvýšila o násobek příslušného koeficientu v řádku funkce z (= 3) a kladné změny proměnné x (= 4), tj. o $3 \cdot 4 = 12$.
Nebázické proměnné položíme rovny 0, tj.

$$\begin{aligned} y &= 0 \\ p_2 &= 0. \end{aligned}$$

Dostáváme nové bázické řešení:

$$\begin{aligned} p_1 &= 2 \\ x &= 4 \\ p_3 &= 5 \\ p_4 &= 2. \end{aligned}$$

Nový vrchol, do kterého jsme se dostali, má tedy souřadnice

$$\boldsymbol{x}^1 = (4, 0, 2, 0, 5, 2).$$

Algoritmus pokračuje dalším krokem: záporná hodnota v řádku z určuje vstupní proměnnou y, výstupní proměnná je určena minimální kladnou změnou

$$\min\left\{\frac{2}{3/2}, \frac{4}{1/2}, \frac{5}{3/2}, \frac{2}{1}\right\} = \frac{4}{3},$$

která nastane v 1.řádku (= p_1 řádek).

Po normalizaci pivotového prvku a vynulování vstupního sloupce pomocí přičtení násobku pivotového řádku k nepivotovým řádkům dostáváme tabulku

	x	y	p_1	p_2	p_3	p_4	řešení
z	0	0	$\frac{1}{3}$	$\frac{4}{3}$	0	0	$12\frac{2}{3}$
y	0	1	$\frac{2}{3}$	$-\frac{1}{3}$	0	0	$\frac{4}{3}$
x	1	0	$-\frac{1}{3}$	$\frac{2}{3}$	0	0	$\frac{10}{3}$
p_3	0	0	-1	1	1	0	3
p_4	0	0	$-\frac{2}{3}$	$\frac{1}{3}$	0	1	$\frac{2}{3}$

Nová hodnota funkce z se zvýšila o $\frac{1}{2} \cdot \frac{4}{3}$, tj. o $\frac{2}{3}$.
Nebázické proměnné položíme rovny 0, tj.

$$p_1 = 0$$
$$p_2 = 0.$$

Dostáváme nové bázické řešení:

$$y = \tfrac{4}{3}$$
$$x = \tfrac{10}{3}$$
$$p_3 = 3$$
$$p_4 = \tfrac{2}{3}.$$

Nový vrchol, do kterého jsme se dostali, má tedy souřadnice

$$\boldsymbol{x}^2 = (\frac{10}{3}, \frac{4}{3}, 0, 0, 3, \frac{2}{3}).$$

Protože v řádku funkce z už nejsou záporné hodnoty, nalezené řešení je optimální, tj.

$$x = \frac{10}{3}, \; y = \frac{4}{3}$$

je optimální množství denní výroby obou nátěrů.

Poznámka k výběru vstupní proměnné:

1) Výběr sloupce odpovídajícího záporné, ale ne nejvíce záporné hodnotě v řádku z by také vedl ke zvýšení hodnoty funkce z, ale to zvýšení by nebylo největší možné.
2) Při minimalizaci je v algoritmu jediný rozdíl: vstupní sloupec je ten, který odpovídá maximální kladné hodnotě v řádku z (ta při převodu na druhou stranu rovnice nejvíce zmenší funkční hodnotu).

8.5 Analýza citlivosti pomocí výstupní simplexové tabulky

Celou analýzu citlivosti lze provést pomocí simplexové tabulky, a to i v úlohách pro vyšší dimenzi, kde už grafické řešení není možné.

Ad Příklad 8.1 Odpovědi na jednotlivé otázky analýzy citlivosti lze vyčíst ze závěrečné (výstupní) simplexové tabulky.

a) Jaký je význam optimálních hodnot $p_1 = p_2 = 0,\ p_3 = 3,\ p_4 = \frac{2}{3}$?

$p_1 = p_2 = 0$ ... omezení $\mathbf{1}, \mathbf{2}$ jsou klíčová, obě suroviny jsou maximálně využity (jedná se o omezení dostupnosti zdrojů)

$p_3 = 3$ ... pravou stranu omezení **3** lze snížit při zachování optima o hodnotu 3, tj.

$\mathbf{3}' : -x + y \leq -2$ ... denní výroba vnějšího nátěru může překročit denní výrobu vnitřního nátěru až o 2 tuny (jedná se o omezení poptávky)

$p_4 = \dfrac{2}{3}$ ... pravou stranu **4** lze snížit o $\dfrac{2}{3}$, tj.

$\mathbf{4}' : y \leq \dfrac{4}{3}$ ... poptávka může být ještě o $\frac{2}{3}$ snížena.

b) Řádek funkce z v optimální tabulce udává stínové ceny příslušné jednotlivým omezením $\rightarrow$ každé omezení je neoddělitelně spjato s jednou přídavnou proměnnou

stínová cena omezení **1**je rovna $\dfrac{1}{3}$ $\rightarrow$ snížení pravé strany **1** o 1 povede ke snížení obratu o $\dfrac{1}{3}$ tisíce dolarů

2 $\frac{4}{3}$

3 0

4 0 $\rightarrow$ kdyby stínová cena **4** byla místo 0 kladná, znamenalo by to, že má smysl zvyšovat poptávku po vnitřním nátěru, čehož lze docílit zvýšením podílu společnosti na trhu.

c) Informace o změně pravých stran klíčových omezení.
Jak moc má smysl zvyšovat pravou stranu omezení **1** ? Kdybychom toto omezení uvažovali ve tvaru $x + 2\,y \leq 6 + \Delta_1$, výsledná simplexová tabulka by byla tatáž až na sloupec pravých stran:

	x	y	p_1	p_2	p_3	p_4	řešení
z			$\frac{1}{3}$				$12\frac{2}{3} + \frac{1}{3}\Delta_1 \geq 0$
y			$\frac{2}{3}$				$\frac{4}{3} + \frac{2}{3}\Delta_1 \geq 0$
x			$-\frac{1}{3}$				$\frac{10}{3} - \frac{1}{3}\Delta_1 \geq 0$
p_3			-1				$3 - 1\Delta_1 \geq 0$
p_4			$-\frac{2}{3}$				$\frac{2}{3} - \frac{2}{3}\Delta_1 \geq 0$

Modifikace pravých stran (= koeficienty u Δ_1) je určena koeficienty ve sloupci výstupní tabulky odpovídajícím přídavné proměnné p_1 v omezení **1** .
Nerovnosti uvedené v pravém sloupci tabulky musí platit, aby řešení bylo přípustné. Vyřešením máme $\Delta_1 \in \langle -2; 1 \rangle$, tj. dostupnost suroviny A má smysl zvýšit o 1 tunu.

Analogicky lze rozebrat i další klíčové omezení **2** . Přitom výsledek získaný pro dané klíčové omezení platí pro tu situaci, že pravé strany ostatních omezení neměníme (vždy sledujeme změnu pouze jednoho omezení).

d) Informace o změně koeficientu funkce z.
Sledujme vliv změny koeficientu u proměnné x: $z = (3 + \delta_1)\,x + 2\,y$.
Optimální tabulka:

	x	y	p_1	p_2	p_3	p_4	řešení
z	0	0	$\left(\frac{1}{3}-\frac{1}{3}\delta_1\right)$	$\left(\frac{4}{3}+\frac{2}{3}\delta_1\right)$	0	0	$12\frac{2}{3}+\frac{10}{3}\delta_1$
y							
x	1	0	$-\frac{1}{3}$	$\frac{2}{3}$	0	0	$\frac{10}{3}$
p_3							
p_4							

Koeficienty u δ_1 jsou dány hodnotami v řádku x v mimobázických sloupcích.
Aby se jednalo o maximum, musí platit

$$\frac{1}{3}-\frac{1}{3}\delta_1 \geq 0,\ \frac{4}{3}+\frac{2}{3}\delta_1 \geq 0\ \Rightarrow \delta_1 \in \langle -2; 1\rangle,$$

tj. koeficient u x ve funkci z musí být v intervalu $\langle 3-2; 3+1\rangle = \langle 1; 4\rangle$.
Situace změny nebázického koeficientu je ještě jednodušší: $z = 3\,x + 2\,y + (0+\delta_3)\,p_1$ způsobí optimální tabulku s řádkem z ve tvaru

	x	y	p_1	p_2	p_3	p_4	řešení
z	0	0	$\frac{1}{3}-\delta_3$	$\frac{4}{3}$	0	0	$12\frac{2}{3}$

U δ_3 je při maximalizaci vždy koeficient (-1), při minimalizaci vždy koeficient $+1$.

e) Při analýze citlivosti jsme se dosud nezabývali otázkou, co způsobí změna koeficientu na levé straně některého omezení. K této otázce se vrátíme v kapitole 2 a zodpovíme ji s využitím teorie duality.

8.6 Obecný tvar simplexové metody s využitím umělých proměnných

Pokud sloupce odpovídající počátečním bázickým proměnným ve vstupní tabulce simplexové metody nevytváří jednotkovou matici, nenulové souřadnice příslušného bázického řešení nejsou přímo rovny pravým stranám simplexové tabulky – tato matice je získána až vyřešením příslušného systému rovnic. A zde se může stát, že některá souřadnice řešení bude záporná, tedy výchozí bázické řešení nebude přípustné.

Příklad 8.7. Minimalizujte funkci $z = 4\,x_1 + x_2$ za podmínek

$$\begin{aligned} 3\,x_1 + x_2 &= 3 \\ 4\,x_1 + 3\,x_2 &\geq 6 \\ x_1 + 2\,x_2 &\leq 4 \end{aligned}$$

$$x_1,\ x_2 \geq 0.$$

Řešení. Kanonický tvar úlohy bude:

najděte minimum funkce $z = 4x_1 + x_2$ za podmínek

$$\begin{aligned} 3x_1 + x_2 \phantom{{}- p_1 + p_2} &= 3 \\ 4x_1 + 3x_2 - p_1 \phantom{{}+ p_2} &= 6 \\ x_1 + 2x_2 \phantom{{}- p_1} + p_2 &= 4 \end{aligned}$$

$$x_1,\ x_2,\ p_1,\ p_2 \geq 0.$$

Vstupní simplexová tabulka má tvar

	x_1	x_2	p_1	p_2	řešení
z	-4	-1	0	0	0
x_2	3	1	0	0	3
p_1	4	3	-1	0	6
p_2	1	2	0	1	4

Za vstupní bázické proměnné jsme zvolili x_2, p_1, p_2, tj. položili jsme $x_1 = 0$. Vyřešíme příslušný systém rovnic

$$\left.\begin{aligned} x_2 \phantom{{}- p_1 + p_2} &= 3 \\ 3x_2 - p_1 \phantom{{}+ p_2} &= 6 \\ 2x_2 \phantom{{}- p_1} + p_2 &= 4 \end{aligned}\right\} \Rightarrow \quad \begin{aligned} x_2 &= 3 \\ p_1 &= 3 \\ p_2 &= -2 \end{aligned}$$

Dané řešení $(0, 3, 3, -2)$ není přípustné, protože jeho čtvrtá souřadnice je záporná. Nejedná se tedy o vrchol množiny přípustných hodnot.

Abychom zaručili, že nalezené vstupní bázické řešení bude přípustné, použijeme jedné ze dvou následujících metod využívajících tzv. umělé proměnné.

a) *Dvoufázová metoda*

1.fáze: Přidáme do levých stran systému nerovností další proměnné, díky nimž zde vznikne jednotková podmatice, a řešíme úlohu lineárního programování s novou funkcí r (ta se vždy minimalizuje, i kdyby původní úloha byla maximalizace):

Najděte minimum funkce $r = u_1 + u_2$ za podmínek

$$\begin{aligned} 3x_1 + x_2 + u_1 \phantom{{}- p_1 + u_2 + p_2} &= 3 \\ 4x_1 + 3x_2 - p_1 + u_2 \phantom{{}+ p_2} &= 6 \\ x_1 + 2x_2 \phantom{{}- p_1 + u_2} + p_2 &= 4 \end{aligned}$$

$$x_1,\ x_2,\ p_1,\ p_2,\ u_1,\ u_2 \geq 0.$$

Vstupní simplexová tabulka má pak tvar

		x_1	x_2	p_1	u_1	u_2	p_2	řešení
min	r	0	0	0	-1	-1	0	0
	u_1	3	1	0	1	0	0	3
	u_2	4	3	-1	0	1	0	6
	p_2	1	2	0	0	0	1	4

Zvolili jsme bázické proměnné u_1, u_2, p_2 a položili $x_1 = x_2 = p_1 = 0$. Ovšem nemůžeme provést optimalizační krok, protože není splněna podmínka, o které už byla řeč: koeficienty v řádku funkce r v bázických sloupcích musí být rovny 0. Abychom vynulovali koeficienty -1 ve sloupcích u_1, u_2, a přitom zachovali nuly ve sloupci p_2, přičteme k řádku r řádky u_1, u_2:

		$\downarrow$						
		x_1	x_2	p_1	u_1	u_2	p_2	řešení
min	r	7	4	-1	0	0	0	9
$\leftarrow$	u_1	**3**	1	0	1	0	0	**3**
	u_2	**4**	3	-1	0	1	0	**6**
	p_2	**1**	2	0	0	0	1	**4**

Nyní můžeme provést optimalizaci:

$$\boldsymbol{x}^0 = (0, 0, 0, 3, 6, 4).$$

Vstupní sloupec určíme podle maximálního kladného prvku v řádku r, výstupní řádek podle minimální kladné změny $\min\left\{\frac{3}{3}, \frac{6}{4}, \frac{4}{1}\right\} = \frac{3}{3}$.

			$\downarrow$					
		x_1	x_2	p_1	u_1	u_2	p_2	řešení
min	r	0	$\frac{5}{3}$	-1	$-\frac{7}{3}$	0	0	2
	x_1	1	$\boldsymbol{\frac{1}{3}}$	0	$\frac{1}{3}$	0	0	**1**
$\leftarrow$	u_2	0	$\boldsymbol{\frac{5}{3}}$	-1	$-\frac{4}{3}$	1	0	**2**
	p_2	0	$\boldsymbol{\frac{5}{3}}$	0	$-\frac{1}{3}$	0	1	**3**

Optimalizace:

$$\boldsymbol{x}^1 = (1, 0, 0, 0, 2, 3)$$

$\max r = \frac{5}{3} \Rightarrow$ vstupní souřadnice je x_2
$\min\left\{\frac{1}{1/3}, \frac{2}{5/3}, \frac{3}{5/3}\right\} = \frac{2}{5/3}$

	x_1	x_2	p_1	u_1	u_2	p_2	řešení
r	0	0	0	-1	-1	0	0
x_1	1	0	$\frac{1}{5}$	$\frac{3}{5}$	$-\frac{1}{5}$	0	$\frac{3}{5}$
x_2	0	1	$-\frac{3}{5}$	$-\frac{4}{5}$	$\frac{3}{5}$	0	$\frac{6}{5}$
p_2	0	0	1	1	-1	1	1

Optimum úlohy je

$$\boldsymbol{x}^2 = \left(\frac{3}{5}, \frac{6}{5}, 0, 0, 0, 1\right),$$

protože v řádku funkce r (mimo pravou stranu, která neslouží k určování vstupní proměnné) už není kladný prvek.
Všimněme si také, že $r(\boldsymbol{x}^2) = 0$. Kdyby tomu tak nebylo a optimální funkční hodnota by byla nenulová, znamenalo by to, že původní úloha (s funkcí z) nemá řešení, tj. 2.fází metody bychom už nepokračovali.

2.fáze: Nyní se vrátíme k minimalizaci funkce z původní úlohy, přičemž omezení opíšeme z výstupní tabulky 1.fáze. Přitom sloupce u_1, u_2 vyloučíme, protože nejsou bázické. Kdyby některá z proměnných u_1, u_2 byla bázická v optimální tabulce 1.fáze, museli bychom ji uvažovat i ve 2.fázi a zajistit, aby vypadla z báze (Dautzig 1963).
Řešíme tedy úlohu:

najděte minimum funkce $z = 4\,x_1 + x_2$ za podmínek

$$\begin{aligned} x_1 \quad + \tfrac{1}{5}\,p_1 \quad &= \tfrac{3}{5} \\ x_2 - \tfrac{3}{5}\,p_1 \quad &= \tfrac{6}{5} \\ p_1 + p_2 &= 1 \end{aligned}$$

$$x_1,\, x_2,\, p_1,\, p_2 \geq 0.$$

Systém podmínek nyní obsahuje sloupce, ze kterých lze složit jednotkovou matici, čili volbou bázických proměnných x_1, x_2, p_2 a položením $p_1 = 0$ dostaneme už přípustné bázické řešení:

		x_1	x_2	p_1	p_2	řešení
min	z	-4	-1	0	0	0
	x_1	1	0	$\mathbf{\frac{1}{5}}$	0	$\mathbf{\frac{3}{5}}$
	x_2	0	1	$-\frac{3}{5}$	0	$\frac{6}{5}$
$\longleftarrow$	p_2	0	0	$\mathbf{1}$	1	$\mathbf{1}$
	z	0	0	$\frac{1}{5}$	0	$\frac{18}{5}$
				$\uparrow$		

Protože v bázických sloupcích řádku z nejsou nuly, musíme tento řádek upravit před prováděním optimalizačního kroku.
Nyní můžeme provést optimalizaci:

$$\boldsymbol{x}^0 = \left(\frac{3}{5}, \frac{6}{5}, 0, 1\right); \; z(\boldsymbol{x}^0) = \frac{18}{5}.$$

max $z = \frac{1}{5} \Rightarrow p_1$ je vstupní proměnná
min $\left\{\frac{3/5}{1/5}, \frac{1}{1}\right\} = 1 \Rightarrow p_2$ je výstupní proměnná.

	x_1	x_2	p_1	p_2	řešení
z	0	0	0	$-\frac{1}{5}$	$\frac{17}{5}$
x_1	1	0	0	$-\frac{1}{5}$	$\frac{2}{5}$
x_2	0	1	0	$\frac{3}{5}$	$\frac{9}{5}$
p_1	0	0	1	1	1

Optimalizace:

$$\boldsymbol{x}^1 = \left(\frac{2}{5}, \frac{9}{5}, 1, 0\right); \; z(\boldsymbol{x}^1) = \frac{17}{5}.$$

Dostáváme tedy optimální řešení úlohy:

$$x_1 = \frac{2}{5};\ x_2 = \frac{9}{5}.$$

b) *Penalizační metoda*
Tato metoda je ekvivalentní dvoufázové metodě – rozdíl je však v tom, že obě fáze jsou zabudovány najednou v jedné úloze lineárního programování.

Ad Příklad 8.7 Najděte minimum funkce $z = 4x_1 + x_2 + M(u_1 + u_2)$, kde M je libovolně velká kladná konstanta (v případě maximalizace funkce z by u M bylo znaménko –) za podmínek

$$\begin{aligned} 3x_1 + x_2 + u_1 &= 3 \\ 4x_1 + 3x_2 - p_1 + u_2 &= 6 \\ x_1 + 2x_2 + p_2 &= 4 \end{aligned}$$

$$x_1,\ x_2,\ p_1,\ p_2,\ u_1,\ u_2 \geq 0.$$

Protože M může nabývat vysokých hodnot, pro optimum úlohy musí platit $u_1 = u_2 = 0$. Pokud to nenastane, původní úloha nemá řešení.
Systém omezení je tedy stejný jako v 1.fázi dvoufázové metody, pouze došlo ke změně funkce z.

		x_1	x_2	p_1	u_1	u_2	p_2	řešení
min	z	-4	-1	0	$-M$	$-M$	0	0
$\longleftarrow$	u_1	**3**	1	0	1	0	0	**3**
	u_2	**4**	3	-1	0	1	0	**6**
	p_2	**1**	2	0	0	0	1	**4**
	z	$-4+7M$	$-1+4M$	$-M$	0	0	0	$9M$
		$\uparrow$						

Před použitím optimalizačního kroku musíme opět vynulovat bázické pozice v řádku z.
Optimalizace:

$$\boldsymbol{x}^0 = (0,0,0,3,6,4);\ z(\boldsymbol{x}^0) = 9M$$

max $z = -4 + 7M \Rightarrow x_1$ je vstupní proměnná
min $\left\{\frac{3}{3}, \frac{6}{4}, \frac{4}{1}\right\} = \frac{3}{3} \Rightarrow u_1$ je výstupní proměnná.

$$\downarrow$$

		x_1	x_2	p_1	u_1	u_2	p_2	řešení
min	z	0	$\frac{1+5M}{3}$	$-M$	$\frac{4-7M}{3}$	0	0	$4+2M$
$\longleftarrow$	x_1	1	$\mathbf{\frac{1}{3}}$	0	$\frac{1}{3}$	0	0	**1**
	u_2	0	$\mathbf{\frac{5}{3}}$	-1	$\frac{4}{3}$	1	0	**2**
	p_2	0	$\mathbf{\frac{5}{3}}$	0	$-\frac{1}{3}$	0	1	**3**

Optimalizace:

$$\boldsymbol{x}^1 = (1,0,0,0,2,3);\ z(\boldsymbol{x}^1) = 4 + 2M$$

max $z = \frac{1+5M}{3} \Rightarrow x_2$ je vstupní proměnná
min $\left\{\frac{1}{1/3}, \frac{2}{5/3}, \frac{3}{5/3}\right\} = \frac{6}{5} \Rightarrow u_2$ je výstupní proměnná.

$\downarrow$

		x_1	x_2	p_1	u_1	u_2	p_2	řešení
min	z	0	0	$\frac{1}{5}$	$(\frac{8}{5} - M)$	$(-\frac{1}{5} - M)$	0	$\frac{18}{5}$
	x_1	1	0	$\mathbf{\frac{1}{5}}$	$\frac{3}{5}$	$-\frac{1}{5}$	0	$\mathbf{\frac{3}{5}}$
	x_2	0	1	$-\frac{3}{5}$	$-\frac{4}{5}$	$\frac{3}{5}$	0	$\frac{6}{5}$
$\leftarrow$	p_2	0	0	**1**	1	-1	1	**1**

Optimalizace:

$$\boldsymbol{x}^2 = (\frac{3}{5}, \frac{6}{5}, 0, 0, 0, 1);\ z(\boldsymbol{x}^2) = \frac{18}{5}$$

max $z = \frac{1}{5} \Rightarrow p_1$ je vstupní proměnná
min $\left\{\frac{3/5}{1/5}, \frac{1}{1}\right\} = 1 \Rightarrow p_2$ je výstupní proměnná.

	x_1	x_2	p_1	u_1	u_2	p_2	řešení
z	0	0	0	$(\frac{7}{5} - M)$	$-M$	$-\frac{1}{5}$	$\frac{17}{5}$
x_1	1	0	0	$\frac{2}{5}$	0	$-\frac{1}{5}$	$\frac{2}{5}$
x_2	0	1	0	$-\frac{1}{5}$	0	$\frac{3}{5}$	$\frac{9}{5}$
p_1	0	0	1	1	-1	1	1

Optimalizace:

$$\boldsymbol{x}^3 = (\frac{2}{5}, \frac{9}{5}, 1, 0, 0, 0);\ z(\boldsymbol{x}^3) = \frac{17}{5}$$

je optimum, protože v řádku funkce z (mimo pravou stranu) nejsou už kladné hodnoty. Dostáváme tedy optimální řešení úlohy:

$$x_1 = \frac{2}{5};\ x_2 = \frac{9}{5}.$$

V úloze jsme při řešení mohli místo abstraktního M použít dostatečně velkou hodnotu, např. $M = 10\,000$. Ale (zejména v počítačovém zpracování algoritmu) při použití konkrétního M dochází k zaokrouhlovací chybě. Proto je výhodnější používat abstraktní M (i v programu, lze totiž dodefinovat algebraické operace sčítání, násobení a porovnávání hodnot obecně i pro abstraktní M s využitím pomocné proměnné, do které je možné ukládat koeficient u M).

Příklad 8.8. Je nutné v následujících úlohách použít umělé proměnné?

a) Najděte maximum funkce $z = x_1 + x_2$ za omezení

$$2\,x_1 + 3\,x_2 = 5$$
$$7\,x_1 + 2\,x_2 \leq 6$$

$$x_1,\ x_2 \geq 0.$$

b) Najděte minimum funkce $z = x_1 + x_2 + x_3 + x_4$ za omezení

$$\begin{aligned} 2\,x_1 + \,x_2 + x_3 &= 7 \\ 4\,x_1 + 3\,x_2 + x_4 &\leq 8 \end{aligned}$$

$$x_1,\, x_2,\, x_3,\, x_4 \geq 0.$$

Řešení.

ad a) Ano, úloha přeformulovaná pro užití simplexové metody zní:

maximalizujte funkci $z = x_1 + x_2 - Mu$ za podmínek

$$\begin{aligned} 2\,x_1 + 3\,x_2 + u &= 5 \\ 7\,x_1 + 2\,x_2 + p &= 6 \end{aligned}$$

$$x_1,\, x_2,\, u,\, p \geq 0.$$

Příslušná simplexová tabulka:

		x_1	x_2	u	p	řešení
max	z	-1	-1	M	0	0
	u	2	3	1	0	5
	p	7	2	0	1	6
	z	$-1-2M$	$-1-3M$	0	0	$-5M$

Nesmíme zapomenout vynulovat bázické koeficienty v řádku z, aby byla tabulka připravená pro optimalizační krok!

ad b) Ne, samotná omezení obsahují už jednotkovou podmatici

		x_1	x_2	x_3	x_4	řešení
min	z	-1	-1	-1	-1	0
	x_3	2	1	1	0	7
	x_4	4	3	0	1	8
	z	5	3	0	0	15

Musíme opět vynulovat bázické pozice v řádku z.

8.7 Úskalí simplexové metody

a) *Degenerace*

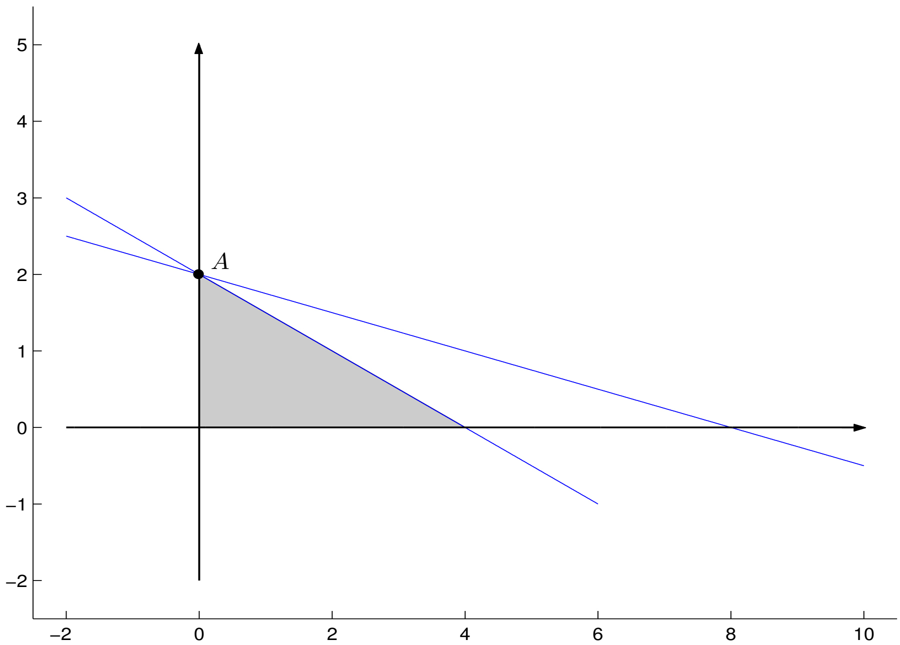
- **matematicky:** bázická souřadnice řešení je rovna 0, tj. může se stát, že jednomu vrcholu množiny přípustných hodnot odpovídá více kroků simplexové tabulky; speciálně je zde nebezpečí tzv. zacyklení, tj. procházení několika vrcholů množiny přípustných hodnot stále dokola, aniž bychom šli k optimálnímu vrcholu.

- **prakticky:** některé omezení je nadbytečné, ale většinou nepoznáme, které to je.
- **řešení:** zkusíme v algoritmu pokračovat, zacyklení většinou nenastane, nebo lze zacyklení většinou vyloučit tím, že pro volbu výstupního řádku na chvíli uvažujeme místo nul na pravé straně systému malá $\varepsilon > 0\ldots$ tzv. ε-modifikace.

Příklad 8.9. Maximalizujte funkci $z = 3\,x_1 + 9\,x_2$ za podmínek

$$\begin{aligned} x_1 + 4\,x_2 &\le 8 \\ x_1 + 2\,x_2 &\le 4 \\[1ex] x_1,\, x_2 &\ge 0. \end{aligned}$$

Řešení. Grafické řešení je znázorněno na obr.8.6.



Obr. 8.6: Grafické řešení příkladu 8.9.

První z obou netriviálních omezení je nadbytečné; optimum je v bodě $A = [0; 2]$.

b) *Více optimálních řešení*

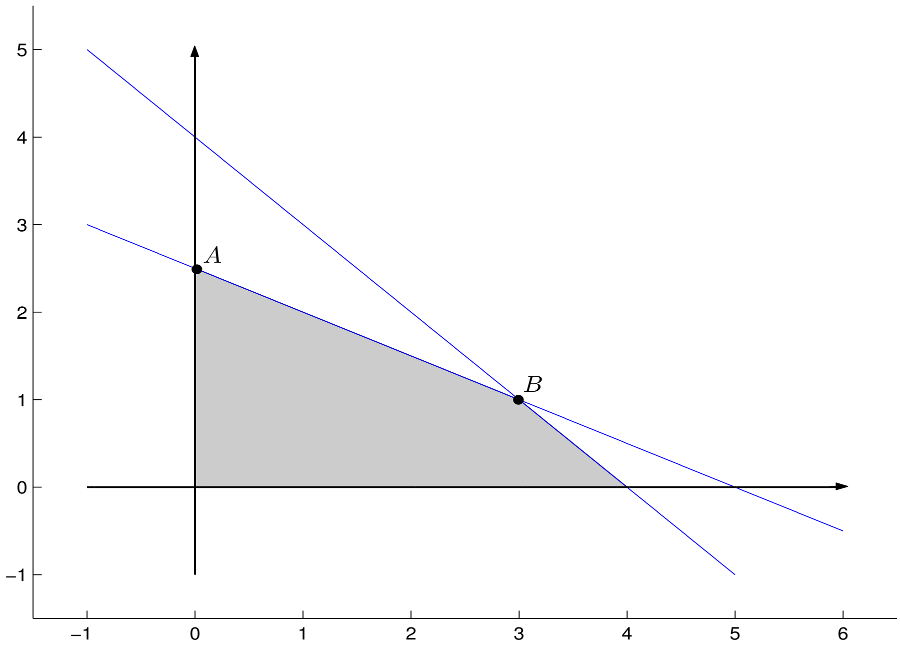
- **matematicky:** nebázický koeficient v řádku z je roven 0 ve výstupní simplexové tabulce; pak pokud s proměnnou příslušnou tomuto koeficientu vstoupíme do báze, příslušné řešení bude rovněž optimální.
- **prakticky:** optimum není pouze v obou vrcholech $\boldsymbol{x}^1, \boldsymbol{x}^2$, ale v každém bodě hrany mezi nimi $\alpha\,\boldsymbol{x}^1 + (1-\alpha)\,\boldsymbol{x}^2$, kde $\alpha \in \langle 0; 1\rangle$.
  Pokud optimum nastane ve 3 vrcholech $\boldsymbol{x}^1, \boldsymbol{x}^2, \boldsymbol{x}^3$, nastane rovněž v každém bodě trojúhelníka těmito vrcholy určeného $\boldsymbol{x} = \alpha_1\boldsymbol{x}^1 + \alpha_2\boldsymbol{x}^2 + \alpha_3\boldsymbol{x}^3$, kde $\alpha_i \in \langle 0; 1\rangle$ a platí $\alpha_1 + \alpha_2 + \alpha_3 = 1$ (tzv. *konvexní kombinace* vrcholů).

Příklad 8.10. Maximalizujte funkci $z = 2\,x_1 + 4\,x_2$ za omezení

$$\begin{aligned} x_1 + 2\,x_2 &\le 5 \\ x_1 + x_2 &\le 4 \end{aligned}$$

$$x_1,\, x_2 \ge 0.$$

Řešení. Grafické řešení je znázorněno na obr.8.7.



Obr. 8.7: Grafické řešení příkladu 8.10.

Libovolný bod úsečky AB je řešením, tj.

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \alpha \begin{pmatrix} 0 \\ 2,5 \end{pmatrix} + (1-\alpha) \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \quad \alpha \in \langle 0; 1 \rangle.$$

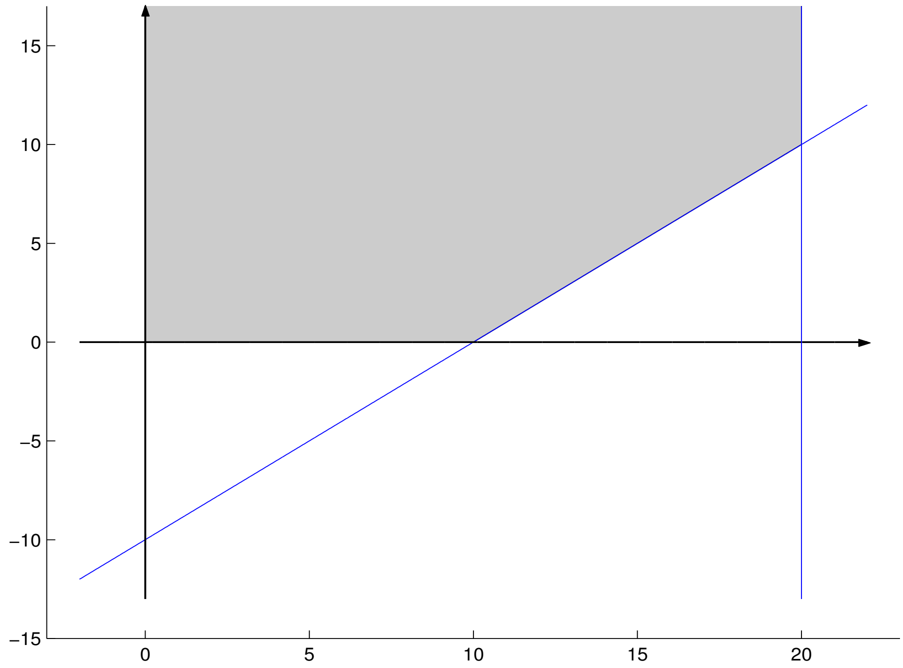
c) *Neomezené řešení*

- **matematicky:** všechny hodnoty ve vstupním sloupci v některém optimalizačním kroku jsou ≤ 0, tj. neexistuje žádná kladná změna pro daný sloupec.
- **prakticky:** úloha není dobře formulována (chybí určité omezení, popřípadě některý parametr není dobře odhadnut).
- **řešení:** řekneme, že funkce z je ve směru svého „zlepšování“ neomezená, popřípadě dodáme další omezení nebo přehodnotíme koeficienty úlohy.

Příklad 8.11. Maximalizujte funkci $z = 2\,x_1 + 4\,x_2$ za omezení

$$\begin{aligned} x_1 + 2\,x_2 &\le 5 \\ x_1 + x_2 &\le 4 \end{aligned}$$

$$x_1,\, x_2 \ge 0.$$



Obr. 8.8: Grafické řešení příkladu 8.11.

Řešení. Grafické řešení je znázorněno na obr.8.8.
Množina přípustných hodnot je neohraničená ve směru růstu funkce z.

d) *Množina přípustných hodnot je prázdná*

- **matematicky:** v 1.fázi dvoufázové metody je optimální hodnota kladná.
- **prakticky:** úloha není dobře formulována, omezení jsou v rozporu.
- **řešení:** některá omezení odstraníme.

Pojmy k zapamatování

- V této kapitole jsme nejdříve na různých příkladech z praxe ukázali matematickou formulaci úlohy lineárního programování, jejíž tvar obecně zní:

nalezněte minimum (maximum) funkce $z = \sum_{j=1}^{n} c_j x_j$ za omezujících podmínek

$$x_j \geq 0 \text{ pro } j = 1, 2, \ldots, n \text{ (tzv. triviální podmínky)},$$
$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i \text{ (nebo } = b_i \text{ nebo } \geq b_i) \text{ pro } i = 1, 2, \ldots, m.$$

- Dále jsme se zabývali grafickým řešením této úlohy a následnou analýzou citlivosti, tj. tím, jak se změní řešení při změně některého ze vstupních parametrů úlohy. Sledovali jsme především změny netriviálních omezení, které mohou být dvojího druhu:

  - klíčová ... pokud prochází bodem optima
  - neklíčová ... pokud neprochází bodem optima.

– Kromě toho jsme se také věnovali algebraickému řešení – tzv. *simplexové metodě*. Pro její nastolení jsme definovali tzv. *kanonický tvar* úlohy lineárního programování, který je charakteristický tím, že

  – všechna omezení jsou rovnicemi

  – všechna omezení mají nezápornou pravou stranu

  – pro všechny proměnné x_j platí: $x_j \geq 0$.

– Úlohu lineárního programování lze řešit algebraicky pomocí *simplexové tabulky* následujícím způsobem:

  1. převedeme na kanonický tvar (přidáním pomocných proměnných, vynásobením (-1), substitucí $x_i = x_i' - x_i''$ pro neomezenou proměnnou x_i)
  2. dodáme umělé proměnné u_i do některých rovnic, abychom zaručili existenci jednotkové matice (eventuelně může mít přeházené sloupce)
  3. vyloučíme z funkce z proměnné určující jednotkovou matici substitucí za tyto proměnné z některé z podmínek omezení
  4. sestavíme vstupní simplexovou tabulku, do prvního řádku zapíšeme rovnici $z - \sum$(kombinace nebázických proměnných) = absolutní člen, do ostatních řádků opíšeme omezení.

– V závěru kapitoly se věnujeme některým úskalím simplexové metody, zejména degeneraci, neexistenci nebo nejednoznačnosti řešení.

Kontrolní otázky

1. U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.
   Úlohou lineárního programování je nalezení globálního extrému libovolné funkce na množině M zadané soustavou lineárních nerovnic.
   a) Každá úloha, která má přípustné řešení, má i optimální řešení.
   b) Každá úloha, která má optimální řešení, má i přípustné řešení.
   c) Pokud má úloha optimální řešení, je určeno jednoznačně.
   d) Změna některého z neklíčových omezení může ovlivnit hodnotu optimálního řešení.
   e) Existuje úloha lineárního programování, která nelze převést do kanonického tvaru?
   f) Pivotový prvek v simplexové tabulce může být záporné číslo.
   g) Při simplexové metodě může dojít k zacyklení, tj. procházení několika vrcholů množiny přípustných hodnot stále dokola, aniž bychom šli k optimálnímu vrcholu.

Odpovědi na otázky
1 – N, 1a) – N, 1b) – A, 1c) – N, 1d) – A, 1e) – N, 1f) – N, 1g) – A.

Cvičení

1. Společnost chce investovat 1 000 dolarů měsíčně na reklamu svého výrobku. Minuta vysílání v rozhlasových pořadech ji stojí 5 dolarů, minuta v televizi 100 dolarů. Společnost by ráda využila rozhlasu časově aspoň dvakrát více než TV. Minulá zkušenost naznačuje, že minuta vysílání v TV způsobí asi 25-krát větší ohlas než minuta rozhlasového vysílání. Formulujte problém nalezení optimálního rozdělení investic jako úlohu lineárního programování (neřešte ji).

2. Navrhněte výrobu krmné směsi, máte-li k dispozici produkty A,B,C,D, které stojí $200, 260, 180, 340$ Kč a obsahují komponenty a,b,c, jejichž obsah je dán následující tabulkou:

	A	B	C	D	požadavek *
a	10	8	12	6	92
b	6	10	4	14	88
c	4	6	2	12	72

* minimální množství komponenty ve směsi

Vytvořte matematický model, úlohu neřešte.

3. Firma má dva provozy. V prvním provozu vyrábí výrobek A, který je částečně finálním výrobkem a částečně polotovarem pro druhý provoz. Ve druhém provozu vyrábí výrobek B. Denní výroba je omezena 3 000 kg suroviny S. Na jednu jednotku A je třeba 5 kg suroviny S, na jednotku výrobku B je třeba jedna jednotka A a 2 kg suroviny S. Výrobku A je nutno vyrobit aspoň 250 jednotek (finálně, ne jako součást výrobku B). Cena výroby jednotky A je 5 Kč, jednotky B 10 Kč. Vytvořte program výroby, který zabezpečuje maximální odbyt. Neřešte, pouze uveďte matematickou formulaci problému. Nezapomeňte, že ty jednotky A, které se spotřebují na výrobu B, už do ceny odbytu nezapočítáváme (ty jsou totiž už zakalkulovány v ceně jednotky B).

4. Tyč 12 m dlouhá se má řezat na 3 různé délky, $A = 5,65$; $B = 3,25$; $C = 2,40$ a to takto: A nejméně 26 kusů, B nejméně 48 kusů, C nejméně 124 kusů. Protože výroba pokračuje i v dalších dekádách, je přípustné řezat do zásoby. Vytvořte řezné plány a matematický model programu, pro který odpad bude minimální.

5. Řešte graficky následující úlohu:

$$\text{nalezněte minimum funkce } z = x_1 - x_2 \text{ za podmínek} \qquad \begin{aligned} 2\,x_1 + x_2 &\geq 2 \\ -3\,x_1 + 2\,x_2 &\leq 6 \\ x_1 + x_2 &\leq 4 \\ x_1 \geq 0,\ x_2 &\geq 0. \end{aligned}$$

6. Řešte graficky následující úlohu:

$$\text{nalezněte minimum funkce } z = x_1 - x_2 \text{ za podmínek} \quad \begin{aligned} 2\,x_1 + x_2 &\geq 2 \\ -3\,x_1 + 2\,x_2 &\leq 6 \\ x_1 \geq 0,\ x_2 &\geq 0. \end{aligned}$$

7. Formulujte takovou úlohu lineárního programování v dimenzi 2 (pro proměnné x, y), aby body $(1, 3)$, $(2, 2)$ byly jejím optimálním řešením a bod $(3, 1)$ už ne. Využijte grafické názornosti úlohy.

8. Uvažujte následující úlohu lineárního programování:

$$\text{nalezněte maximum funkce } z = 5\,x + 3\,y \text{ za omezení} \quad \begin{aligned} x + y &\leq 4 \\ 5\,x + 2\,y &\leq 10 \\ x \geq 0,\ y &\geq 0. \end{aligned}$$

   Určete

   a) optimální řešení úlohy (graficky);
   b) zvýšení pravé strany v první nerovnosti, které maximálně zlepší hodnotu optima (a současně učiní tuto nerovnost nadbytečnou) a určete tuto změnu optima;
   c) stínové ceny příslušné optimálnímu řešení a);
   d) o kolik můžeme zvýšit koeficient proměnné y ve funkci z, aby vrchol optima (= optimální hodnota proměnných) zůstal zachován (hodnota optima se samozřejmě změní)?

9. Simplexovou metodou vyřešte následující úlohu lineárního programování:

$$\text{najděte maximum funkce } z = 2\,x_1 + x_2 - 5\,x_3 \text{ za podmínek} \quad \begin{aligned} x_1 + x_2 + x_3 &= 7 \\ 2\,x_1 - 5\,x_2 + x_3 &\geq 10 \\ x_1, x_2, x_3 &\geq 0. \end{aligned}$$

10. Simplexovou metodou vyřešte následující úlohu lineárního programování:

$$\text{najděte maximum funkce } z = x_1 + 2\,x_2 + 3\,x_3 - x_4 \text{ za podmínek} \quad \begin{aligned} x_1 + 2\,x_2 + 3\,x_3 &= 15 \\ 2\,x_1 + x_2 + 5\,x_3 &= 20 \\ x_1 + 2\,x_2 + x_3 + x_4 &= 10 \\ x_1, x_2, x_3, x_4 &\geq 0. \end{aligned}$$

Výsledky

ad 1. matematická formulace úlohy: maximalizujte funkci $z = x + 25\,y$ za podmínek $5\,x + 100\,y = 1\,000$, $x \geq 2\,y$, $x, y \geq 0$.

ad 2. matematická formulace úlohy: minimalizujte funkci $z = 200\,x_1 + 260\,x_2 + 180\,x_3 + 340\,x_4$ za podmínek $10\,x_1 + 8\,x_2 + 12\,x_3 + 6\,x_4 \geq 92,\ 6\,x_1 + 10\,x_2 + 4\,x_3 + 14\,x_4 \geq 88,\ 4\,x_1 + 6\,x_2 + 2\,x_3 + 12\,x_4 \geq 72,\ x_1, x_2, x_3, x_4 \geq 0$.

ad 3. matematická formulace úlohy: maximalizujte funkci $z = 5\,(x - y) + 10\,y$ za podmínek $x - y \geq 250,\ 5\,x + 2\,y \leq 3\,000,\ x, y \geq 0$.

ad 4. matematická formulace úlohy: minimalizujte funkci $z = 0,7\,x_1 + 0,7\,x_2 + 0,7\,x_3 + 1,55\,x_4 + 1,55\,x_5 + 0\,x_6 + 2,25\,x_7$ za podmínek $x_1 + 2\,x_2 + x_5 \geq 26,\ x_1 + 2\,x_3 + x_4 + 3\,x_7 \geq 48,\ x_1 + 2\,x_3 + 3\,x_4 + 2\,x_5 + 5\,x_6 \geq 124,\ x_1, x_2, x_3, x_4, x_5, x_6, x_7 \geq 0$, (proměnné x_i odpovídají jednotlivým možnostem řezu tyče, tj. např. $x_1 = ABC,\ x_2 = AA$ atd.).

ad 5. $A = \left[\frac{2}{5}, \frac{18}{5}\right],\ z(A) = -\frac{16}{5}$.

ad 6. Minimum neexistuje.

ad 7. Jedna z možných matematických formulací úlohy: maximalizujte funkci $z = x + y$ za podmínek $x + y \leq 4,\ x - y \leq 0,\ x, y \geq 0$.

ad 8. ad a) $V = \left[\frac{2}{3}, \frac{10}{3}\right],\ z(V) = \frac{40}{3}$; ad b) pravou stranu první nerovnosti můžeme zvýšit na 5, optimum pak bude $z(0; 5) = 15$; ad c) stínová cena pro 1.omezení je $\frac{5}{3}$, stínová cena pro 2.omezení je $\frac{2}{3}$; ad d) koeficient u y lze zvýšit o 2.

ad 9. $\boldsymbol{x} = (7, 0, 0);\ z(\boldsymbol{x}) = 14$.

ad 10. $\boldsymbol{x} = \left(\frac{5}{2}, \frac{5}{2}, \frac{5}{2}, 0\right);\ z(\boldsymbol{x}) = 15$.

9 Dualita v úlohách lineárního programování

Průvodce studiem

Tato kapitola je uvedením do teorie duality. Co to je dualita?

Jde o dvě struktury. První je primární (s latinského primus = první) a od něj se odvozuje duální (z latinského dualis = dvojitý, dvojný), která je sice odvozena z první, ale je s ní zcela rovnocenná. Z druhé strany jestliže vezmeme jako výchozí druhou strukturu,tak k ní můžeme přiřadit první strukturu jako duální.

Například v klasické Eukleidovské geometrii je řada tvrzení o bodech a přímkách. Všichni znáte větu:
„Každé dva různé body určují právě jednu přímku".
V tomto případě máme primární body a duální jsou přímky. Pokud ale otočíme pozici bodů a přímek, dostaneme opět pravdivé tvrzení:
„Každé dvě různoběžné přímky určují právě jeden bod = jejich průsečík."
Právě tak i další tvrzení, kde výchozí jsou body a z nich vyplývají vlastnosti přímek, můžeme prohodit mezi sebou body a přímky a znovu dostaneme pravdivé tvrzení, které bude duální a kde budeme vycházet z přímek a dostaneme tvrzení o bodech.

Při studiu vztahu mezi primární a duální úlohou uvidíme, že obě úlohy někdy „stojí proti sobě" (v něčem se liší), jindy „jsou v souladu" (v něčem jsou stejné nebo se doplňují), ale vždy jsou navzájem rovnocenné, žádná není nadřazena té druhé.

Cíle

Po prostudování této kapitoly budete schopni:

- K dané primární úloze přiřadit duální úlohu.
- Chápat vztahy mezi řešením primární a duální úlohy.
- Používat duální simplexovou metodu.
- Provádět analýzu citlivosti.
- Chápat ekonomickou interpretaci duality.

9.1 Formulace duální úlohy lineárního programování

Původní úlohu lineárního programování označujeme jako *primární:*

> nalezněte minimum (nebo maximum) funkce $z = \sum\limits_{j=1}^{n} c_j x_j$ za omezujících podmínek
>
> $$\sum_{j=1}^{n} a_{ij} x_j = b_i \text{ pro } i = 1, 2, \dots, m,$$
>
> $$x_j \geq 0 \text{ pro } j = 1, 2, \dots, n.$$

Povšimněte si, že primární úloha je v kanonickém tvaru.
K této primární úloze konstruujeme tzv. *duální* úlohu podle následujících pravidel:

primární úloha	duální úloha
minimalizace	maximalizace, všechna omezení typu $\leq$, proměnné neohraničené
maximalizace	minimalizace, všechna omezení typu $\geq$, proměnné neohraničené

Vztah koeficientů primární a duální úlohy je možné znázornit v tabulce:

primární proměnné

$$
\begin{array}{l|cccccc|c|cl}
 & x_1 & x_2 & \dots & x_j & \dots & x_n & & & \\
\text{pravé strany duálních omezení} \rightarrow & c_1 & c_2 & \dots & c_j & \dots & c_n & & & \\
\hline
 & a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} & b_1 & y_1 & \\
 & a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} & b_2 & y_2 & \\
\text{koeficienty levých stran duálních omezení} & \vdots & & & \vdots & & \vdots & \vdots & \vdots & \text{duální proměnné} \\
 & a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} & b_m & y_m & \\
\hline
 & & & & \uparrow & & & \uparrow & & \\
 & & & & \text{j-té duální omezení} & & & \text{koeficienty duální funkce } w & &
\end{array}
$$

Příklad 9.1. Formulujte duální úlohu k úloze:

maximalizujte funkci $z = 5\,x_1 + 12\,x_2 + 4\,x_3$ za podmínek

$$
\begin{aligned}
x_1 + 2\,x_2 + \,x_3 &\leq 10 \\
2\,x_1 - \,x_2 + 3\,x_3 &= 8 \\
x_1,\, x_2,\, x_3 &\geq 0.
\end{aligned}
$$

Řešení. Nejprve si musíme vytvořit kanonický tvar primární úlohy:

maximalizujte funkci $z = 5\,x_1 + 12\,x_2 + 4\,x_3 + 0\,p$ za podmínek

$$\begin{aligned} x_1 + 2\,x_2 + \,x_3 + p &= 10 \\ 2\,x_1 - \,x_2 + 3\,x_3 \phantom{{}+p} &= 8 \\ x_1,\, x_2,\, x_3,\, p &\geq 0. \end{aligned}$$

Duální úloha je tedy tvaru:

minimalizujte funkci $w = 10\,y_1 + 8\,y_2$ za podmínek

$$\begin{aligned} y_1 + 2\,y_2 &\geq 5 \\ 2\,y_1 - \,y_2 &\geq 12 \\ y_1 + 3\,y_2 &\geq 4 \\ y_1 \geq 0;\ y_2 &\in \mathbb{R}, \end{aligned}$$

tj. na y_2 není kladeno žádné omezení.

Příklad 9.2. Formulujte duální problém k následujícímu:

minimalizujte funkci $z = 5\,x_1 - 2\,x_2$ za omezení

$$\begin{aligned} -x_1 + \,x_2 &\geq -3 \\ 2\,x_1 + 3\,x_2 &\leq 5 \\ x_1,\, x_2 &\geq 0. \end{aligned}$$

Řešení. Nejprve kanonický tvar primární úlohy:

minimalizujte funkci $z = 5\,x_1 - 2\,x_2 + 0\,p_1 + 0\,p_2$ za podmínek

$$\begin{aligned} x_1 - \,x_2 + p_1 \phantom{{}+p_2} &= 3 \\ 2\,x_1 + 3\,x_2 \phantom{{}+p_1} + p_2 &= 5 \\ x_1,\, x_2,\, p_1,\, p_2 &\geq 0. \end{aligned}$$

Duální úloha je tvaru:

maximalizujte funkci $w = 3\,y_1 + 5\,y_2$ za podmínek

$$\begin{aligned} y_1 + 2\,y_2 &\leq 5 \\ -y_1 + 3\,y_2 &\leq -2 \\ y_1,\, y_2 &\leq 0. \end{aligned}$$

Příklad 9.3. Formulujte duální problém k následujícímu:

maximalizujte funkci $z = 5\,x_1 + 6\,x_2$ za omezení

$$\begin{aligned} x_1 + 2\,x_2 &= 5 \\ -x_1 + 5\,x_2 &\geq 3 \\ 4\,x_1 + 7\,x_2 &\leq 8 \end{aligned}$$

$$x_1 \in \mathbb{R},\ x_2 \geq 0.$$

Řešení. Kanonický tvar primární úlohy:

maximalizujte funkci $z = 5\,x_1' - 5\,x_1'' + 6\,x_2 + 0\,p_1 + 0\,p_2$ za podmínek

$$\begin{array}{rcrcrcrcrcl} x_1' & - & x_1'' & + & 2\,x_2 & & & & & = & 5 \\ -x_1' & + & x_1'' & + & 5\,x_2 & - & p_1 & & & = & 3 \\ 4\,x_1 & - & 4\,x_1'' & + & 7\,x_2 & & & + & p_2 & = & 8 \end{array}$$

$$x_1',\ x_1'',\ x_2,\ p_1,\ p_2 \geq 0.$$

Duální úloha je tvaru:

minimalizujte funkci $w = 5\,y_1 + 3\,y_2 + 8\,y_3$ za podmínek

$$\begin{array}{rcrcrcr} y_1 & - & y_2 & + & 4\,y_3 & \geq & 5 \\ -y_1 & + & y_2 & - & 4\,y_3 & \geq & -5 \\ 2\,y_1 & + & 5\,y_2 & + & 7\,y_3 & \geq & 6 \\ & & -y_2 & & & \geq & 0 \\ & & & & y_3 & \geq & 0 \end{array}$$

$$y_1 \in \mathbb{R}.$$

Všimněme si, že sloučením prvních dvou podmínek vznikne rovnost a že na proměnnou y_1 není vztaženo žádné omezení.

Jenom pro pořádek si připomeňme, že před řešením duální úlohy je opět třeba tuto úlohu převést do kanonického tvaru.

9.2 Vztah mezi řešením primární a duální úlohy

Vyřešme nyní primární úlohu a úlohu k ní duální a všimněme si souvislostí mezi jednotlivými simplexovými tabulkami.

Ad Příklad 9.1. Vyřešme nejprve primární úlohu. Kanonický tvar doplníme umělou proměnnou u s cílem vytvoření jednotkové podmatice:

maximalizujte funkci $z = 5\,x_1 + 12\,x_2 + 4\,x_3 + 0\,p - M u$ za podmínek

$$\begin{aligned} x_1 + 2x_2 + x_3 + p &= 10 \\ 2x_1 - x_2 + 3x_3 + u &= 8 \end{aligned}$$

$$x_1,\ x_2,\ x_3,\ p,\ u \geq 0.$$

Nyní je vše připraveno pro provedení simplexové metody pro primární úlohu.

	max $z =$	5	12	4	**0**	-M	
		x_1	x_2	x_3	p	u	řešení
krok 0:	z	-5	-12	-4	0	M	$=0$
	p	1	2	1	1	0	10
$x_3 \leftrightarrow$	u	2	-1	**3**	0	1	8
	z	$-5-2M$	$-12+M$	$-4-3M$	0	0	$-8M$

Nejprve jsme museli vynulovat bázické pozice v řádku z. Duální omezení odpovídající počáteční bázi získáme z tabulky ze sloupců p a u. Tj. jsou to omezení

$$\begin{aligned} y_1 &\geq 0 \\ y_2 &\geq -M \end{aligned}$$

(do pravých stran bereme původní nevynulované koeficienty s nezměněným znaménkem – zakroužkované hodnoty v tabulce).

		x_1	x_2	x_3	p	u	řešení
krok 1:	z	$-\frac{7}{3}$	$-\frac{40}{3}$	0	0	$\frac{4}{3}+M$	$\frac{32}{3}$
$x_2 \leftrightarrow$	p	$\frac{1}{3}$	$\boldsymbol{\frac{7}{3}}$	0	1	$-\frac{1}{3}$	$\frac{22}{3}$
	x_3	$\frac{2}{3}$	$-\frac{1}{3}$	1	0	$\frac{1}{3}$	$\frac{8}{3}$

		x_1	x_2	x_3	p	u	řešení
krok 2:	z	$-\frac{3}{7}$	0	0	$\frac{40}{7}$	$-\frac{4}{7}+M$	$\frac{368}{7}$
	x_2	$\frac{1}{7}$	1	0	$\frac{3}{7}$	$-\frac{1}{7}$	$\frac{22}{7}$
$x_1 \leftrightarrow$	x_3	$\boldsymbol{\frac{5}{7}}$	0	1	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{26}{7}$

		x_1	x_2	x_3	p	u	řešení
krok 3:	z	0	0	$\frac{3}{5}$	$\frac{29}{5}$	$-\frac{2}{5}+M$	$54\frac{4}{5}$
	x_2	0	1	$-\frac{1}{5}$	$\frac{2}{5}$	$-\frac{1}{5}$	$\frac{12}{5}$
	x_1	1	0	$\frac{7}{5}$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{26}{5}$

Dostáváme optimální řešení:

$$x_1 = \frac{26}{5},\ x_2 = \frac{12}{5},\ x_3 = 0.$$

Z tabulky můžeme také vyčíst optimální koeficienty funkce z příslušné primární počáteční bázi – hodnoty označené čárkovaně.
Nyní vyřešme duální úlohu převedenou na kanonický tvar doplněný o umělé proměnné:

$$\text{minimalizujte funkci } w = 10\,y_1 + 8\,y_2' - 8\,y_2'' + M\,(u_1 + u_2 + u_3) \text{ za podmínek}$$

$$\begin{array}{rcl} y_1 + 2\,y_2' - 2\,y_2'' - p_1 \qquad\qquad + u_1 \qquad\qquad & = & 5 \\ 2\,y_1 - y_2' + y_2'' \qquad - p_2 \qquad\qquad + u_2 \qquad & = & 12 \\ y_1 + 3\,y_2' - 3\,y_2'' \qquad\qquad - p_3 \qquad\qquad + u_3 & = & 4 \end{array}$$

$$y_1,\, y_2',\, y_2'',\, p_1,\, p_2,\, p_3,\, u_1,\, u_2,\, u_3 \geq 0.$$

krok 0:

min $w =$	10	8	-8	0	0	0	**M**	M	**M**	
	y_1	y_2'	y_2''	p_1	p_2	p_3	u_1	u_2	u_3	řešení
w	-10	-8	8	0	0	0	$-M$	$-M$	$-M$	$= 0$
u_1	1	2	-2	-1	0	0	1	0	0	5
u_2	2	-1	1	0	-1	0	0	1	0	12
u_3	1	3	-3	0	0	-1	0	0	1	4
w	$-10+4M$	$-8+4M$	$8-4M$	$-M$	$-M$	$-M$	0	0	0	$21M$

Primární omezení odpovídající počáteční duální bázi jsou

$$\begin{aligned} x_1 &\leq M \\ x_2 &\leq M \\ x_3 &\leq M \end{aligned}$$

(do pravých stran opět bereme původní nevynulované koeficienty s nezměněným znaménkem – zakroužkované hodnoty v tabulce).

⋮

krok 4:

	y_1	y_2'	y_2''	p_1	p_2	p_3	u_1	u_2	u_3	řešení
w	0	0	0	$-\frac{26}{5}$	$-\frac{12}{5}$	0	$\frac{26}{5} - M$	$\frac{12}{5} - M$	$-M$	$54\frac{4}{5}$
p_3	0	0	0	$-\frac{7}{5}$	$\frac{1}{5}$	1	$\frac{7}{5}$	$-\frac{1}{5}$	-1	$\frac{3}{5}$
y_2''	0	-1	1	$\frac{2}{5}$	$-\frac{1}{5}$	0	$-\frac{2}{5}$	$\frac{1}{5}$	0	$\frac{2}{5}$
y_1	1	1	0	$-\frac{1}{5}$	$-\frac{2}{5}$	0	$\frac{1}{5}$	$\frac{2}{5}$	0	$\frac{29}{5}$

Dostáváme optimální řešení:

$$y_1 = \frac{29}{5},\; y_2 = y_2' - y_2'' = 0 - \frac{2}{5} = -\frac{2}{5}.$$

Z tabulky můžeme také vyčíst optimální koeficienty funkce w odpovídající počáteční duální bázi – hodnoty označené čárkovaně.

Vztah mezi řešením primární a duální úlohy:

1) optimální hodnota funkce z = optimální hodnota funkce w

2)

$$\begin{pmatrix} \text{vektor optimálních koeficientů funkce } z \\ \text{příslušných počáteční primární bázi} \end{pmatrix} = \begin{pmatrix} \text{vektor rozdílů levé minus pravé strany} \\ \text{duálních omezení příslušných počáteční} \\ \text{primární bázi} \end{pmatrix}$$

Ad Příklad 9.1 Příslušná vektorová rovnice z bodu 2) vztahu mezi primární a duální úlohou zde má tvar

$$\begin{pmatrix} \frac{29}{5} \\ -\frac{2}{5} + M \end{pmatrix} = \begin{pmatrix} y_1 - 0 \\ y_2 - (-M) \end{pmatrix} \Rightarrow y_1 = \frac{29}{5},\ y_2 = -\frac{2}{5},\ w = 54\frac{4}{5}.$$

Čili pomocí řešení primární úlohy jsme z tohoto vztahu získali řešení duální úlohy. Naopak uvážíme-li, že duální úloha k duální úloze je původní primární úloha, lze pomocí řešení duální úlohy určit řešení primární úlohy:

$$\begin{pmatrix} \frac{26}{5} - M \\ \frac{12}{5} - M \\ -M \end{pmatrix} = \begin{pmatrix} x_1 - M \\ x_2 - M \\ x_3 - M \end{pmatrix} \Rightarrow x_1 = \frac{26}{5},\ x_2 = \frac{12}{5},\ x_3 = 0,\ z = 54\frac{4}{5}.$$

Z porovnání primární a duální úlohy vidíme, že

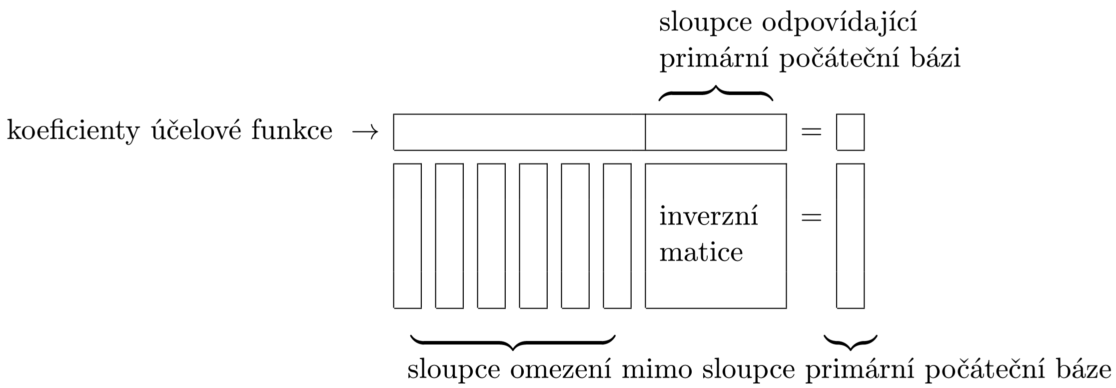
$$\begin{pmatrix} \text{libovolná funkční hodnota přípustného} \\ \text{bázického řešení maximalizace} \end{pmatrix} \leq \begin{pmatrix} \text{libovolná funkční hodnota přípustného} \\ \text{bázického řešení minimalizace} \end{pmatrix}$$

bez ohledu na to, která úloha je primární a která duální. Když tedy známe „dobré“ řešení minimalizace a „dobré“ řešení maximalizace (dobré v tom smyslu, že funkční hodnoty v obou případech se od sebe moc neliší), můžeme s jistotou vědět, že skutečné optimum obou úloh má funkční hodnotu v intervalu určeném těmito dvěma „dobrými“ funkčními hodnotami.

9.3 Pojem inverzní matice

V k-tém kroku simplexové tabulky lze všechny hodnoty této tabulky určit na základě tzv. inverzní matice (a zadání primární a duální úlohy):

tabulka k-tého kroku:



a) určení sloupců omezení mimo sloupce počáteční báze:

$$\begin{pmatrix} \text{sloupec} \\ \text{v } k\text{-tém} \\ \text{kroku} \end{pmatrix} = \begin{pmatrix} \text{inverzní ma-} \\ \text{tice v } k\text{-tém} \\ \text{kroku} \end{pmatrix} \cdot \begin{pmatrix} \text{sloupec} \\ \text{v 0-tém} \\ \text{kroku} \end{pmatrix}$$

Ad Příklad 9.1. *krok 1*: určení sloupce proměnné x_1:

$$\begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{3} \\ 0 & \frac{1}{3} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

krok 2: určení sloupce pravých stran:

$$\begin{pmatrix} \frac{22}{7} \\ \frac{26}{7} \end{pmatrix} = \begin{pmatrix} \frac{3}{7} & -\frac{1}{7} \\ \frac{1}{7} & \frac{2}{7} \end{pmatrix} \cdot \begin{pmatrix} 10 \\ 8 \end{pmatrix}$$

b) určení koeficientů účelové funkce:
Nejprve určíme hodnoty duálních proměnných y_i podle vztahu

$$\begin{pmatrix} \text{vektor pro-} \\ \text{měnných} \\ y_i \text{ v } k\text{-tém} \\ \text{kroku} \end{pmatrix} = \begin{pmatrix} \text{vektor koeficientů} \\ \text{nad inverzní ma-} \\ \text{ticí v } k\text{-tém kroku} \\ \text{primární tabulky} \end{pmatrix} \cdot \begin{pmatrix} \text{inverzní} \\ \text{matice} \\ \text{v } k\text{-tém} \\ \text{kroku} \end{pmatrix},$$

a pak lze vypočíst primární řádek účelové funkce z vektorové rovnice

$$\begin{pmatrix} \text{vektor koefici-} \\ \text{entů u } x_j \text{ úče-} \\ \text{lové funkce v } k\text{-} \\ \text{tém kroku} \end{pmatrix} = \begin{pmatrix} \text{vektor rozdílů levé minus} \\ \text{pravé strany odpovídají-} \\ \text{cího duálního omezení s} \\ \text{hodnotami proměnných } y_i \\ \text{z } k\text{-tého kroku} \end{pmatrix}.$$

Ad Příklad 9.1. V našem příkladu máme vždy v daném kroku nad inverzní maticí následující koeficienty primární účelové funkce:

krok	báze	koeficienty účelové funkce nad inverzní maticí
0	(p, u)	$(0, -M)$
1	(p, x_3)	$(0, 4)$
2	(x_2, x_3)	$(12, 4)$
3	(x_2, x_1)	$(12, 5)$

(všimněte si, že v 0-tém kroku bereme původní koeficienty s nezměněným znaménkem).
Tj. například hodnoty v řádku funkce z ve 3.(= optimálním) kroku simplexové tabulky určíme následovně:
Nejprve najdeme hodnoty y_i ve 3.kroku

$$(y_1, y_2) = (12, 5) \cdot \begin{pmatrix} \frac{2}{5} & -\frac{1}{5} \\ \frac{1}{5} & \frac{2}{5} \end{pmatrix} = \left(\frac{29}{5}, -\frac{2}{5}\right)$$

a pak příslušné koeficienty funkce z ve 3.kroku:

$$\begin{pmatrix} \text{koeficient u } x_1 \\ \text{koeficient u } x_2 \\ \text{koeficient u } x_3 \\ \text{koeficient u } p \\ \text{koeficient u } u \end{pmatrix} = \begin{pmatrix} y_1 + 2\,y_2 - 5 \\ 2\,y_1 - y_2 - 12 \\ y_1 + 3\,y_2 - 4 \\ y_1 - 0 \\ y_2 - (-M) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \frac{3}{5} \\ \frac{29}{5} \\ -\frac{2}{5} + M \end{pmatrix}.$$

c) Určení hodnoty účelové funkce v k-tém kroku:

dosazením sloupce řešení z k-tého kroku do účelové funkce vypočteme její hodnotu v k-tém kroku.

Pomocí sloupců příslušných inverzní matici lze tedy získat hodnoty ve všech ostatních sloupcích. Toto lze využít při programování algoritmu simplexové metody – více o tom v kapitole 3.

9.4 Ekonomická interpretace duality

Uvažujme následující primární a duální úlohu:

P:

$$\boxed{\begin{array}{c}\text{maximalizujte funkci } z = \sum\limits_{j=1}^{n} c_j x_j \text{ za omezujících podmínek} \\ \sum\limits_{j=1}^{n} a_{ij} x_j = b_i \text{ pro } i = 1, 2, \ldots, m, \\ x_j \geq 0 \text{ pro } j = 1, 2, \ldots, n.\end{array}}$$

D:

$$\boxed{\begin{array}{c}\text{minimalizujte funkci } w = \sum\limits_{i=1}^{m} b_i x_i \text{ za omezujících podmínek} \\ \sum\limits_{i=1}^{m} a_{ij} y_i \geq c_j \text{ pro } j = 1, 2, \ldots, n, \\ y_i \geq 0 \text{ pro } i = 1, 2, \ldots, m.\end{array}}$$

Jednotlivé koeficienty a funkce mají následující ekonomický význam:

- c_j ... zisk jednotkového výstupu činnosti j (obyčejně je diktován trhem)
- b_i ... dostupné množství zdroje i
- a_{ij} ... množství zdroje i potřebné pro jednotkový výstup činnosti j
- z ... zisk
- y_i ... cena jednotkového množství zdroje i (= tzv. stínová cena – udává, jak moc jednotkové zvýšení dostupnosti zdroje i zvýší zisk z)
- w ... využití zdrojů (odčerpání zdrojů)

Ad Příklad 9.1 Význam optimálních hodnot $y_1 = \frac{29}{5}$, $y_2 = -\frac{2}{5}$.
Abychom zvýšili optimum funkce z, musíme

– zvýšit dostupnost zdroje 1
– snížit dostupnost zdroje 2.

Vždy platí $z \leq w$.
Řešení není optimální, pokud zisk $z <$ využití zdrojů w
$\max z$ ... maximalizujeme zisk
$\min w$ ... minimalizujeme využití zdrojů pro daný zisk

Podmínka optimality v k-tém kroku maximalizace ... koeficient funkce z u proměnné x_j je roven $\sum\limits_{i=1}^{m} a_{ij} y_i - c_j \geq 0$.

$\uparrow$

z_j ... tzv. připsaná cena

Pokud $z_j - c_j \geq 0$, tj. $z_j \geq c_j$, proměnná x_j nemůže být vstupní proměnnou optimalizačního kroku

$\uparrow$

z_j ... tzv. redukovaná cena

(zvyšování proměnné x_j nepřinese větší zisk).

ad primární úloha: položky b_i lze někdy zvýšit (stínové ceny určují prioritu) novými investicemi

ad duální úloha: aby se zvýšila ziskovost činnosti j, snažíme se snížit její připsanou cenu $z_j = \sum\limits_{i=1}^{m} a_{ij} y_i$, což se obyčejně dosahuje snížením koeficientu spotřeby a_{kj} odpovídajícího největší duální souřadnici y_k

Příklad 9.4. Uvažujme výrobní halu, kde tři různé typy výrobků prochází každý třemi různými linkami. Limity doby přístupu ke každé lince jsou po řadě 430, 460 a 420 minut denně a jednotkový zisk výrobků je 3,2 a 5. Tabulka udává dobu (v minutách) průchodu výrobků jednotlivými linkami:

	výrobek 1	výrobek 2	výrobek 3
linka 1	1	2	1
linka 2	3	0	2
linka 3	1	4	0

Řešení. Primární úloha bude tvaru:

maximalizujte denní zisk $z = 3\,x_1 + 2\,x_2 + 5\,x_3$ za podmínek

$$\begin{aligned} x_1 + 2\,x_2 + x_3 &\le 430 \\ 3\,x_1 \phantom{{}+ 2\,x_2} + 2\,x_3 &\le 460 \\ x_1 + 4\,x_2 \phantom{{}+ 2\,x_3} &\le 420 \end{aligned}$$

$$x_1,\, x_2,\, x_3 \ge 0.$$

Výstupní simplexová tabulka má tvar:

	x_1	x_2	x_3	p_1	p_2	p_3	řešení
z	4	0	0	1	2	0	1350
x_2	$-\frac{1}{4}$	1	0	$\frac{1}{2}$	$-\frac{1}{4}$	0	100
x_3	$\frac{3}{2}$	0	1	0	$\frac{1}{2}$	0	230
p_3	2	0	0	-2	1	1	20

Analýza: optimální řešení neobsahuje výrobek 1 ($x_1 = 0$), tj. tento výrobek není ziskový ($z_1 > c_1 = 3$), ale můžeme jej ziskovým učinit snížením z_1 ($z_1 = y_1 + 3\,y_2 + y_3$).

Protože z duální úlohy získáváme duální řešení
$$y_1 = 1$$
$$y_2 = 2$$
$$y_3 = 0,$$
má smysl něco dělat jen s první a druhou linkou, větší priorita je dávána druhé lince. Zkoumejme tedy druhou nerovnost primární úlohy; její pravou stranu zvyšovat zatím nechceme, zabývejme se tedy snížením koeficientů na levé straně:

	výrobek 1	výrobek 2	výrobek 3
linka 1	1	2	1
linka 2	$3 - r$	0	2
linka 3	1	4	0

Jak velké musí být r, aby se stal výrobek 1 ziskovým? Tak, aby $z_1 \le c_1$, tj. $1 + (3 - r)\,2 \le 3 \Rightarrow r \ge 2$.

9.5 Duální simplexová metoda

Jestliže v případě maximalizační úlohy není řešení optimální, aspoň pro jeden koeficient j přepočítávané funkce z platí:

$$\sum_{i=1}^{m} a_{ij} y_i - c_j < 0, \text{ tj.}$$

$$(*) \quad \sum_{i=1}^{m} a_{ij} y_i < c_j.$$

Všimneme-li si blíže omezení $(*)$, toto omezení se vyskytuje v duální úloze, ale s opačnou nerovností $(\geq)$. Tedy primární řešení není optimální $\Leftrightarrow$ příslušné duální řešení je nepřípustné. Odtud plyne hlavní myšlenka duální simplexové metody:

> Pokud primární bázické řešení je nepřípustné (některá jeho souřadnice je < 0), ale platí podmínka optimality ($z_j - c_j \geq 0$ pro všechna j), snažíme se duální simplexovou metodou přejít do vrcholu, který stále splňuje podmínku optimality, a navíc už je přípustný (jeho souřadnice ≥ 0). První takový vrchol, do kterého touto metodou dorazíme, je optimum primární úlohy.

Postup: na rozdíl od regulární simplexové metody probíhá optimalizační krok v opačném pořadí – nejprve vybereme výstupní řádek (a sice podle nejvíce záporné souřadnice ve sloupci pravých stran), a potom vstupní sloupec (podle minimální absolutní hodnoty podílu „koeficient v řádku funkce / koeficient ve výstupním řádku“, přičemž kladné a nulové jmenovatele vypouštíme; pokud takové jsou všechny, úloha nemá přípustné řešení).

Příklad 9.5. Minimalizujte funkci $z = 2\,x_1 + x_2$ za podmínek

$$\begin{aligned} 3\,x_1 + \,x_2 &\geq 3 \\ 4\,x_1 + 3\,x_2 &\geq 6 \\ x_1 + 2\,x_2 &\leq 3 \\ x_1,\, x_2 &\geq 0. \end{aligned}$$

Řešení. Pokud první dvě nerovnosti vynásobíme (-1), vyhneme se použití umělých proměnných, ale za cenu toho, že nalezené bázické řešení není přípustné:

	x_1	x_2	p_1	p_2	p_3	řešení
z	$\mathbf{-2}$	$\mathbf{-1}$	0	0	0	0
p_1	-3	-1	1	0	0	-3
p_2	$\mathbf{-4}$	$\boxed{\mathbf{-3}}$	0	1	0	-6
p_3	1	2	0	0	1	3

$\Rightarrow$ podmínka optimality je splněna (koeficienty jsou ≤ 0)

Z hodnot v posledním sloupci je vidět, že řešení je nepřípustné; pokusíme se nalézt přípustný vrchol pomocí duální simplexové metody:
Výstupní řádek určíme pomocí maximálně záporné souřadnice řešení – to je v našem případě $p_2 = -6$. V tomto řádku p_2 jsou dva záporné koeficienty – k nim vytvoříme podíly $|z - hodnota/p_2 - hodnota|$: $\left|\frac{-2}{-4}\right|$, $\left|\frac{-1}{-3}\right|$. Minimální z nich je ten druhý, vstupní sloupec tedy bude x_2.

Zbytek algoritmu je stejný jako u původní simplexové metody. Na místo pivotového prvku (-3) se snažíme dostat v dalším kroku hodnotu $(+1)$, ostatní hodnoty v pivotovém sloupci chceme vynulovat, a to skrze přičtení jistého násobku pivotového řádku k danému řádku. Dostaneme tabulku

	x_1	x_2	p_1	p_2	p_3	řešení
z	$-\frac{2}{3}$	0	0	$-\frac{1}{3}$	0	2
p_1	$-\frac{5}{3}$	0	1	$-\frac{1}{3}$	0	-1
x_2	$\frac{4}{3}$	1	0	$-\frac{1}{3}$	0	2
p_3	$-\mathbf{\frac{5}{3}}$	0	0	$\frac{2}{3}$	1	-1

$\Rightarrow$ podmínka optimality je stále splněna

Z hodnot v posledním sloupci je vidět, že řešení opět není přípustné; zopakujeme tedy optimalizační krok duální simplexové metody:
Výstupní řádek si můžeme vybrat z řádků p_1 a p_3 díky minimální hodnotě (-1). Vyberme tedy například řádek p_3. Jediná záporná hodnota určuje vstupní sloupec x_1.

	x_1	x_2	p_1	p_2	p_3	řešení
z	0	0	0	$-\frac{3}{5}$	$-\frac{2}{5}$	$\frac{12}{5}$
p_1	0	0	1	-1	-1	0
x_2	0	1	0	$\frac{1}{5}$	$\frac{4}{5}$	$\frac{6}{5}$
p_3	1	0	0	$-\frac{2}{5}$	$-\frac{3}{5}$	$\frac{3}{5}$

Z hodnot v posledním sloupci je vidět, že řešení je přípustné, a tedy i optimální.

9.6 Analýza citlivosti v celé své kráse

a) Změna pravé strany některého omezení: Může vést jen k tomu, že řešení nebude přípustné (některá souřadnice bude < 0), podmínka optimality zůstane zachována $\Rightarrow$ můžeme pokračovat použitím duální simplexové metody.

Ad Příklad 8.1. Změníme-li v zadání úlohy pravou stranu omezení **1** na 7 a pravou stranu omezení **2** na 4, výstupní tabulka simplexové metody změněné úlohy se bude od výstupní tabulky původní úlohy lišit pouze ve sloupci pravých stran – řešení bude ale nepřípustné (některá jeho souřadnice bude < 0). Po jednom kroku duální simplexové metody dospějeme k optimu (jehož hodnota bude horší). Podrobněji viz samostatné cvičení.

b) Přidání nového omezení: Výsledek pozměněné úlohy se nemění, pokud bod optima toto omezení splňuje. V opačném případě musíme doplnit omezení na rovnost (pomocnou proměnnou), vyloučit z této rovnosti optimální bázické proměnné původní úlohy (do nové báze se navíc přidá nová pomocná proměnná, čili původní simplexová tabulka je doplněna o řádek i sloupec) a případně použít duální simplexovou metodu, pokud je to potřeba.

Ad Příklad 8.1. Chceme-li k omezením původní úlohy přidat podmínku $x \leq 3$, doplněním na rovnost máme $x + p_5 = 3$, výstupní simplexová tabulka původní úlohy se doplní o řádek a sloupec. Protože proměnná x je bázická, musíme přidaný řádek upravit tím, že od něj odečteme řádek **x** . Tak se poruší nezápornost pravé strany tabulky a provedením jednoho kroku duální simplexové dospějeme k novému optimálnímu řešení (funkční hodnota v něm bude nižší – to se ale dalo čekat, že přidáním dalšího omezení se nezlepší funkční hodnota optima). Podrobněji samostatně.

c) Změna koeficientů účelové funkce: Jeden způsob přepočtu byl popsán na str.230. Uveďme zde ještě jeden způsob pro přepočet změny koeficientů, které stojí u bázických proměnných výstupní tabulky původní úlohy. Tento způsob užívá duálních proměnných a duálních omezení. Vysvětlíme jej na příkladu.

Ad Příklad 8.1. Pokud funkce v původní úloze bude změněna na $z = 5\,x + 4\,y$, pořadí koeficientů vzhledem k bázi výstupní tabulky je $\begin{matrix} y & x & p_3 & p_4 \\ (4, & 5, & 0, & 0) \end{matrix}$, a tedy

$$\begin{pmatrix} y_1, & y_2, & y_3, & y_4 \end{pmatrix} = \begin{pmatrix} 4, & 5, & 0, & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & 0 & 0 \\ -\frac{1}{3} & \frac{2}{3} & 0 & 0 \\ -1 & 1 & 1 & 0 \\ -\frac{2}{3} & \frac{1}{3} & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1, & 2, & 0, & 0 \end{pmatrix}.$$

Pomocí rozdílů levých a pravých stran duálních omezení určíme nový z-řádek ve výstupní tabulce:

$$\begin{aligned} \text{nový koeficient u} \quad x:&\quad y_1 + 2\,y_2 - y_3 - 5 = 0 \\ y:&\quad 2\,y_1 + y_2 + y_3 + y_4 - 4 = 0 \\ p_1:&\quad y_1 = 1 \\ p_2:&\quad y_2 = 2 \\ p_3:&\quad y_3 = 0 \\ p_4:&\quad y_4 = 0 \end{aligned}$$

Vlastně stačilo přepočítat jen nebázické koeficienty, protože bázické jsou rovny nule. Všechny nové z-koeficienty jsou ≥ 0, tj. bod optima se nezmění (i když funkční hodnota v něm ano). Pokud by některý koeficient byl záporný, museli bychom najít simplexovou metodou zlepšení.

Kdyby funkce v původní úloze byla změněna na $z = 4\,x + y$, po přepočtení z-koeficientů by bylo nutné provést jeden krok simplexové metody, abychom našli nový bod optima.

d) Změna levých stran omezení: Má smysl analyzovat jen změnu nebázického sloupce levé strany (při bázické změně je lepší vyřešit celou úlohu znovu); z příslušné duální nerovnosti lze hned zjistit, zda se neporušila podmínka optimality.

Ad Příklad 8.1. Pro $z = 4\,x + y$ a změnu druhého sloupce levých stran $a_{12} = 4$, $a_{22} = 3$ má příslušné duální omezení tvar $4\,y_1 + 3\,y_2 + y_3 + y_4 \geq 1$. Po výpočtu duálních proměnných vidíme, že podmínka platí.

Porušení podmínky optimality řeší klasická simplexová metoda.

e) Přidání nové činnosti (nového sloupce): – tj. změna funkce z i matice (a_{ij}) současně (pokud se na situaci díváme tak, že sloupec tam už dříve byl, ale všechny jeho koeficienty byly nulové).

Přidání nové činnosti má smysl jen tehdy, pokud zlepší hodnotu optima. Vysvětlíme přepočet na příkladu.

Ad Příklad 8.1. Přidáním nového výrobku do našich úvah vznikne úloha

$$\text{maximalizujte funkci } z = 3\,x + 2\,y + \tfrac{3}{4}\,n \text{ za podmínek}$$

$$
\begin{aligned}
x + 2y + \tfrac{3}{4}n &\le 6 \\
2x + \ \ y + \tfrac{3}{4}n &\le 8 \\
-x + \ \ y - \ \ n &\le 1 \\
y \qquad &\le 2 \\
x, y, n &\ge 0.
\end{aligned}
$$

Protože proměnnou n považujeme za nebázickou, duální řešení zůstává stejné: $y_1 = \frac{1}{3}$, $y_2 = \frac{4}{3}$, $y_3 = 0$, $y_4 = 0$. Nové duální omezení $\frac{3}{4}y_1 + \frac{3}{4}y_2 - y_3 \ge \frac{3}{2}$ není splněno, příslušný z-koeficient je roven $\frac{3}{4}y_1 + \frac{3}{4}y_2 - y_3 - \frac{3}{2} = -\frac{1}{4}$. Zbytek sloupce proměnné n vypočteme pomocí inverzní matice:

$$
\begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & 0 & 0 \\ -\frac{1}{3} & \frac{2}{3} & 0 & 0 \\ -1 & 1 & 1 & 0 \\ -\frac{2}{3} & \frac{1}{3} & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{3}{4} \\ \frac{3}{4} \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{4} \\ -1 \\ -\frac{1}{4} \end{pmatrix}.
$$

Nyní protože z-koeficient v novém sloupci je záporný, přidáme jej k optimální tabulce původní úlohy a provedeme jeden krok klasické simplexové metody. Dostaneme nové řešení, které zlepší hodnotu optima původní úlohy.

f) Při změně pravých a levých stran omezení současně: Dochází zde ke složitějším změnám a je výhodnější celou úlohu vyřešit znovu, analýza citlivosti už není tak pomocná.

Duální simplexová metoda nachází své využití nejen při analýze citlivosti, ale i v některých dalších algoritmech, např. v metodě řezů celočíselného lineárního programování.

Pojmy k zapamatování

– V této kapitole byl podán stručný úvod do teorie duality. Nejprve jsme původní úlohu lineárního programování (její kanonický tvar) označili jako *primární*. K této úloze konstruujeme tzv. *duální* úlohu podle následujících pravidel:

primární úloha	duální úloha
minimalizace	maximalizace, všechna omezení typu $\le$, proměnné neohraničené
maximalizace	minimalizace, všechna omezení typu $\ge$, proměnné neohraničené

– Na příkladu jsme pak pomocí simplexové metody zkoumali vztah mezi řešením primární a duální úlohy. Ten se dá stručně charakterizovat takto:

1) optimální hodnota funkce z = optimální hodnota funkce w.

2) (vektor optimálních koeficientů funkce z příslušných počáteční primární bázi) = (vektor rozdílů levé minus pravé strany duálních omezení příslušných počáteční primární bázi)

– Dále jsme zjistili, že všechny hodnoty simplexové tabulky lze určit pomocí tzv. inverzní matice. Toto lze využít zejména při programování algoritmu simplexové metody a také při analýze citlivosti.

– Zajímavá je také ekonomická interpretace duality, která se dá velice stručně charakterizovat jako maximalizace zisku z při současné minimalizaci využití zdrojů w pro daný zisk.

– V souvislosti s dualitou jsme se zabývali duální simplexovou metodou, jejíž hlavní myšlenka zní:

  Pokud primární bázické řešení je nepřípustné (některá jeho souřadnice je < 0), ale platí podmínka optimality ($z_j - c_j \geq 0$ pro všechna j), snažíme se duální simplexovou metodou přejít do vrcholu, který stále splňuje podmínku optimality, a navíc už je přípustný (jeho souřadnice ≥ 0). První takový vrchol, do kterého touto metodou dorazíme, je optimum primární úlohy.

– Na závěr kapitoly jsme se věnovali využití poznatků z teorie duality k získání efektivnějších postupů v analýze citlivosti.

Kontrolní otázky

1. U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.
   a) Abychom mohli formulovat duální úlohu, primární úloha musí být v kanonickém tvaru.
   b) Duální úloha existuje jen tehdy, pokud primární úloha je úlohou minimalizace.
   c) Koeficienty duální funkce w zapsané v simplexové tabulce jsou totéž co pravé strany primárních omezení.
   d) Optimální hodnota funkce z je menší než optimální hodnota funkce w.
   e) I když primární řešení není optimální, příslušné duální řešení je přípustné.
   f) Pokud primární řešení je přípustné, příslušné duální řešení je optimální.
   g) Pokud příslušné duální řešení je přípustné, primární řešení je optimální.
   h) Pokud je alespoň jedna hodnota v posledním sloupci simplexové tabulky kladná, řešení primární úlohy je přípustné.
   i) Přidáním nového omezení se optimální hodnota funkce z může zlepšit.

Odpovědi na otázky
1a) – A, 1b) – N, 1c) – A, 1d) – N, 1e) – N, 1f) – N, 1g) – A, 1h) – N, 1i) – N.

Cvičení

1. Uvažujte následující úlohu lineárního programování:

$$\text{najděte maximum funkce } z = 2x_1 + 4x_2 + 4x_3 - 3x_4 \text{ za podmínek} \quad \begin{aligned} x_1 + x_2 + x_3 &= 4 \\ x_1 + 4x_2 + x_4 &= 8 \\ x_1, x_2, x_3, x_4 &\geq 0. \end{aligned}$$

   a) Formulujte k této úloze úlohu duální.

b) Najděte řešení duální úlohy pomocí optimální tabulky primární úlohy.

2. Uvažujte následující úlohu lineárního programování:

$$\text{najděte maximum funkce } z = 5\,x_1 + 2\,x_2 + 3\,x_3 \text{ za podmínek } \begin{aligned} x_1 + 5\,x_2 + 2\,x_3 &= 30 \\ x_1 - 5\,x_2 - 6\,x_3 &\le 40 \\ x_1, x_2, x_3 &\ge 0. \end{aligned}$$

a) Formulujte k této úloze úlohu duální.

b) Najděte řešení duální úlohy pomocí optimální tabulky primární úlohy.

3. Vyřešte duální simplexovou metodou úlohu:

$$\text{najděte minimum funkce } z = 2\,x_1 + 3\,x_2 \text{ za podmínek } \begin{aligned} 2\,x_1 + 3\,x_2 &\le 30 \\ x_1 + 2\,x_2 &\ge 10 \\ x_1, x_2 &\ge 0. \end{aligned}$$

4. Uvažujte zadání z příkladu 2. Proveďte následující analýzu citlivosti:

a) jak se změní řešení při změně pravé strany na $\begin{pmatrix} 35 \\ 15 \end{pmatrix}$?

b) jak se změní řešení při přidání nového omezení $x_1 + x_3 \le 5$?

c) jak se změní řešení při změně funkce z na $z = x_1 + x_2 - 2\,x_3$?

d) jak se změní řešení při změně 2.sloupce omezení na $\begin{pmatrix} 4 \\ 1 \end{pmatrix}$?

Výsledky příkladů viz **??**.

Výsledky

ad 1. ad a) formulace duální úlohy: minimalizujte funkci $w = 4\,y_1 + 8\,y_2$ za podmínek $y_1 + y_2 \ge$ $\ge 2,\ y_1 + 4\,y_2 \ge 4,\ y_1 \ge 4, y_2 \ge -3$; ad b) $\boldsymbol{y} = (4, 0);\ w(\boldsymbol{y}) = 16$.

ad 2. ad a) formulace duální úlohy: minimalizujte funkci $w = 30\,y_1 + 40\,y_2$ za podmínek $y_1 +$ $+ y_2 \ge 5,\ 5\,y_1 - 5\,y_2 \ge 2,\ 2\,y_1 - 6\,y_2 \ge 3,\ y_1 \in \mathbb{R}, y_2 \ge 0$; ad b) $\boldsymbol{y} = (5, 0);\ w(\boldsymbol{y}) = 150$.

ad 3. $\boldsymbol{x} = (0, 5);\ z(\boldsymbol{x}) = 15$.

ad 4. ad a) $\boldsymbol{x} = \left(30, 0, \frac{5}{2}\right),\ z(\boldsymbol{x}) = \frac{315}{2}$; ad b) $\boldsymbol{x} = (5, 5, 0),\ z(\boldsymbol{x}) = 35$; ad c) hodnota optima se nemění $\boldsymbol{x} = (30, 0, 0)$, změní se pouze hodnota účelové funkce $z(\boldsymbol{x}) = 30$; ad d) nic se nezmění.

Maplety

V následujících mapletech si můžete některé studované pojmy přiblížit, případně si sestavit vlastní zadání příkladů.

1. Výpočet inverzní matice

10 Dopravní úloha

Průvodce studiem

V této kapitole se budeme věnovat řešení dopravní úlohy. Jde sice také o úlohu lineárního programování, ale protože má určité specifické vlastnosti, ukážeme si jiné metody řešení.

Nejdříve si ukážeme, jak převést úlohu na vyvážený tavr, potom si projdeme tři způsoby získání přípustného řešení a nakonec si ukážeme jakým způsobem budeme provádět optimalizaci.

V závěru si ukážeme řešení přiřazovací úlohy.

Cíle

Po prostudování této kapitoly budete schopni:

- Formulovat zadání dopravní úlohy, upravit zadání na vyvážený tvar, najít počáteční přípustné řešení a provádět optimalizační kroky, které Vás dovedou k řešení.
- Řešit modifikací dopravního problému, tzv. přiřazovací úlohu.
- Řešit další modifikací dopravní úlohy tzv. *problém překladu materiálu.*

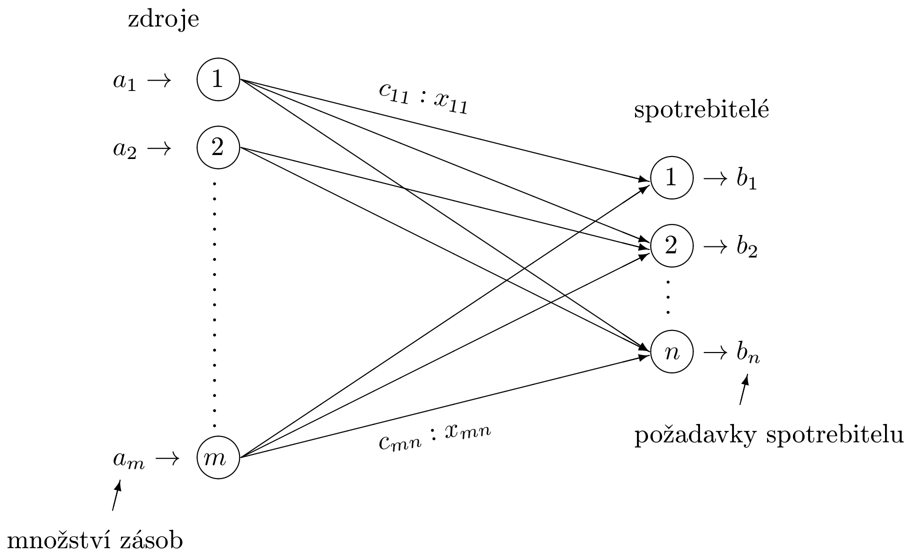
10.1 Úvod

Klasická verze dopravní úlohy má následující zadání: rozhodněte o způsobu rozvozu jednoho typu výrobku (nebo suroviny) z m závodů do n spotřebitelských skladů tak, aby se minimalizovala celková cena dopravy, je-li dáno

$$
\begin{aligned}
c_{ij} &\ \ldots \text{ jednotková cena dopravy ze zdroje } i \text{ na místo určení } j,\\
a_i &\ \ldots \text{ množství zásob se zdroji } i,\ i = 1, 2, \ldots, m,\\
b_j &\ \ldots \text{ požadavek spotřebitele } j,\ j = 1, 2, \ldots, n.
\end{aligned}
$$

Výstupem úlohy jsou optimální hodnoty proměnných x_{ij} (= množství jednotek dopravovaných ze zdroje i ke spotřebiteli j) a celková cena dopravy $\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n} c_{ij}x_{ij}$.

Tuto konkrétní úlohu praxe lze matematicky reprezentovat užitím teorie grafů:



Ale užitím grafových algoritmů ji řešit nebudeme. Dopravní úlohu lze také formulovat jako úlohu lineárního programování:

$$\text{minimalizujte } z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} \text{ za podmínek}$$
$$\sum_{j=1}^{n} x_{ij} \leq a_i \text{ pro } i = 1, 2, \ldots, m$$
$$\sum_{i=1}^{m} x_{ij} \geq b_j \text{ pro } j = 1, 2, \ldots, n$$
$$x_{ij} \geq 0.$$

Ale protože dopravní úloha má jakýsi speciální tvar, nebudeme ji řešit přímo simplexovou metodou; použijeme algoritmus, ve kterém je simplexová metoda skryta, ale který je rychlejší a přehlednější.

Právě popsaný model dopravní úlohy budeme vždy řešit až po eventuálním převedení na tzv. vyvážený tvar (vyvážený dopravní model), kdy $\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$, tj. nabídka = poptávka. Tento vyvážený tvar je analogií kanonického tvaru úlohy lineárního programování.

Příklad 10.1. Automobilová společnost MG má závody v Los Angeles, Detroitu a New Orleans a distribuční střediska v Denveru a Miami. Kapacity výrobních závodů pro plánované období jsou po řadě 1000, 1500 a 1200 ks aut, poptávka středisek je 2300 a 1400 ks. Cena dopravy 1 auta na jednu míli je 8 centů (cent je setina dolaru). Určete optimální rozdělení dopravy od výrobců ke spotřebitelským místům. Vzdálenosti mezi místy (v mílích) jsou uvedeny v tabulce:

	Denver	Miami
Los Angeles	1000	2690
Detroit	1250	1350
New Orleans	1275	850

Řešení. a) Cenu dopravy c_{ij} jednoho auta z místa výroby do spotřebitelského střediska lze určit vynásobením vzdálenosti míst číslem 0,08, což je cena za 1 kus. Nyní všechny informace sestavíme do přehledné tabulky, ze které budeme při řešení vycházet (ceny c_{ij} jsou uváděny v pravém horním rohu jednotlivých polí):

	Denver	Miami	
Los Angeles	80	215	1000
Detroit	100	108	1500
New Orleans	102	68	1200
	2300	1400	

$1000 + 1500 + 1200 = 2300 + 1400 \Rightarrow$ jedná se už o vyváženou úlohu, protože kapacity výrobců jsou stejné jako požadavky odběratelů. U této úlohy můžeme tedy spustit řešící algoritmus.

b) Pokud by kapacita výroby v Detroitu nebyla 1500, ale 1300, museli bychom dopravní úlohu nejprve vyvážit; protože kapacita výroby by nebyla dostačující, zavedli bychom fiktivního výrobce:

	Denver	Miami	
Los Angeles	80	215	1000
Detroit	100	108	1500
New Orleans	102	68	1200
Fiction	0	0	200
	2300	1400	

Ceny dopravy od fiktivního výrobce zvolíme rovny 0, protože fiktivní výrobce je nenáročný. Nyní už „celková kapacita výroby = celková poptávka“, čili úloha je vyvážena a připravena k tomu, aby na ni byl vypuštěn řešící algoritmus.

c) Naopak, pokud by kapacita výroby byla větší než poptávka, úlohu bychom vyvážili zavedením fiktivního spotřebitele. Například při poptávce Denveru 1900 místo 2300 ks bychom zavedli fiktivního spotřebitele s poptávkou 400 ks:

	Denver	Miami	Al Capone	
Los Angeles	80	215	0	1000
Detroit	100	108	0	1500
New Orleans	102	68	0	1200
	1900	1400	400	

Kdybychom chtěli, aby např. Detroit navzdory menší poptávce vyvezl všechna vyrobená auta, místo 0 bychom cenu c_{23} dopravy z Detroitu fiktivnímu odběrateli zvolili kladnou a dostatečně velkou, a tím by se zaručilo, že v optimálním řešení $x_{23} = 0$.

□

Řešení úloh a), b), c) zde nebudeme uvádět. Místo toho si řekneme další dva příklady formulace dopravního modelu, a až poté uvedeme algoritmus řešení dopravních úloh.

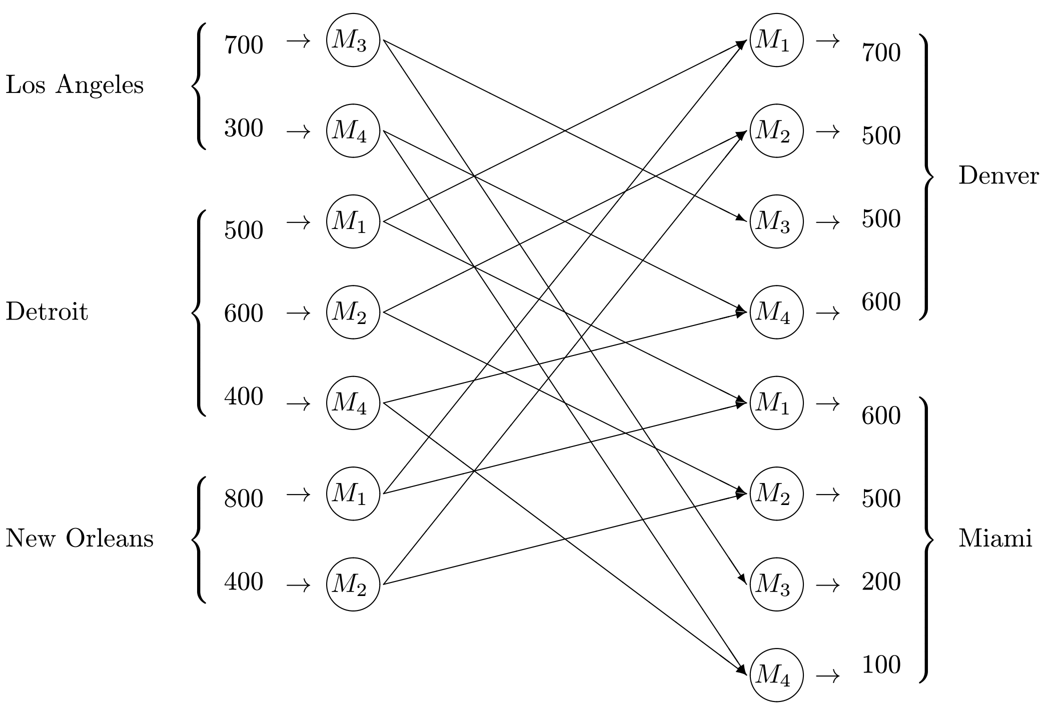
Příklad 10.2. Model dopravy většího počtu typů výrobků: Společnost MG vyrábí čtyři různé typy aut (M_1, M_2, M_3, M_4). Kapacity jednotlivých závodů a požadavky distribučních center jsou uvedeny v tabulkách:

závod vyrábí	M_1	M_2	M_3	M_4	celkem
Los Angeles	–	–	700	300	1000
Detroit	500	600	–	400	1500
New Orleans	800	400	–	–	1200

požadavek distribučního střediska	M_1	M_2	M_3	M_4	celkem
Denver	700	500	500	600	2300
Miami	600	500	200	100	1400

Vzdálenosti mezi městy jsou uvedeny v př.10.1, jednotkové ceny dopravy jsou stejné pro všechny typy aut. Určete optimální rozdělení dopravy.

Matematická formulace úlohy:
Každý ze závodů rozdělíme na tolik podzávodů, kolik typů auta daný závod vyrábí. A podobně každé distribuční centrum rozdělíme na čtyři podcentra. Odpovídající grafová reprezentace úlohy:



Přípravou pro řešení je tabulka dopravní úlohy. Ceny c_{ij} těch tras, které nemá smysl realizovat, položíme rovny velkému číslu M. Dopravní tabulka má pak tvar

		Denver				Miami				
		M_1	M_2	M_3	M_4	M_1	M_2	M_3	M_4	
Los Angeles	M_3	M	M	80	M	M	M	215	M	700
	M_4	M	M	M	80	M	M	M	215	300
Detroit	M_1	100	M	M	M	108	M	M	M	500
	M_2	M	100	M	M	M	108	M	M	600
	M_4	M	M	M	100	M	M	M	108	400
New Orleans	M_1	102	M	M	M	68	M	M	M	800
	M_2	M	102	M	M	M	68	M	M	400
		700	500	500	600	600	500	200	100	

Úloha je vyvážená, dopravní tabulka je připravena pro vypuštění řešícího algoritmu.

Příklad 10.3. Plánování výroby: Společnost potřebuje vyrobit (= má poptávku)

během 1.měsíce 100 ks výrobku
2.měsíce 200
3.měsíce 180
4.měsíce 300

Poptávka v aktuálním měsíci může být pokryta

a) nadbytkem výroby v předchozím měsíci

b) výrobou v aktuálním měsíci

c) předem objednanou výrobou v následujícím měsíci.

Výrobní cena je 4 dolary za kus, cena skladování 0,5 dolaru za kus a měsíc, penalizační poplatek za výrobek objednaný předem je 2 dolary za kus a měsíc. Výrobní kapacita je omezena jinými vyráběnými výrobky a je 50 ks v 1.měsíci
180 ks ve 2.měsíci
280 ks ve 3.měsíci
270 ks ve 4.měsíci.

Vytvořte plán výroby na následující 4 měsíce, který minimalizuje celkovou cenu nákladů.

Matematická formulace úlohy:
Matematická formulace úlohy je stejná jako u dopravního problému. Analogie mezi oběma úlohami jsou:

dopravní problém	plán výroby
i ... zdroj	i ... výrobní období
j ... spotřebitel	j ... poptávkové období
a_i ... kapacita zdroje i	a_i ... výrobní kapacita období i
b_j ... poptávka spotřebitele j	b_j ... poptávka v období j

$$c_{ij} = \begin{cases} \text{výrobní cena v období } i, & \text{při } i = j \\ \text{výrobní cena v } i + \text{cena skladování od } i \text{ do } j, & \text{při } i < j \\ \text{výrobní cena v } i + \text{penalizační cena od } j \text{ do } i, & \text{při } i > j \end{cases}$$

Při těchto analogiích lze problém sestavit do dopravní tabulky:

období	1	2	3	4	
1	4	4,5	5	5,5	50
2	6	4	4,5	5	180
3	8	6	4	4,5	280
4	10	8	6	4	270
	100	200	180	300	

Tedy i tuto a podobné úlohy, které přímo nesouvisí s dopravní tematikou, lze řešit dopravním algoritmem.

10.2 Řešení dopravního problému

A. Počáteční krok

V počátečním kroku nalezneme nějaké přípustné rozdělení dopravy, které nemusí být nutně optimálním (v optimalizačním kroku je pak případně vylepšíme). Existuje více metod pro nalezení počátečního rozdělení, ukážeme si tři z nich.

a) Metoda severozápadního rohu (SZR)

1) Proměnné x_{11} (která se nachází v severozápadním rohu tabulky) přiřadíme maximální možnou hodnotu dopravovaných jednotek. Protože tedy x_{11} je maximální možné, vyčerpá se tím kapacita výroby prvního výrobce (nebo poptávka prvního spotřebitele), takže ve zbylých polích prvního řádku (resp. prvního sloupce) umístíme pomlčky naznačující, že se zde žádné jednotky dopravovat nebudou.

2) Pokračujeme přiřazením maximální možné hodnoty do pole x_{21} při vyčerpání řádku 1 v kroku 1 (resp. do pole x_{12} při vyčerpání prvního sloupce v kroku 1). Zkrátka a dobře – další vyplňované pole je vždy směrem na východ nebo na jih podle toho, zda byl v předchozím kroku vyčerpán sloupec nebo řádek.

3) Krok 2 opakujeme tak dlouho, dokud se nedostaneme do pravého dolního (= jihovýchodního) pole tabulky. U vyvážené úlohy tím docílíme toho, že rozdělení jednotek dopravy bude splňovat požadavky na kapacity ve všech řádcích i sloupcích.

Příklad 10.4. Tuto i další metody objasníme na následující tabulkové formulaci dopravní úlohy:

	S_1	S_2	S_3	S_4	
V_1	10	0	20	11	15
V_2	12	7	9	20	25
V_3	0	14	16	18	5
	5	15	15	10	

Řešení. Pro počáteční rozdělení úlohy se třemi výrobci a čtyřmi spotřebiteli užijeme metodu severozápadního rohu: $x_{11} = 5$, což je poptávka v celém prvním sloupci, čili do polí s indexy $(2, 1)$ a $(3, 1)$ umístíme pomlčku a posuneme se směrem na východ, do pole $(1, 2)$. Kapacita ve 2. sloupci je 15, ovšem hodnotu 15 do pole $(1, 2)$ přiřadit nemůžeme, protože by se překročil součet 15 (= kapacita 1. výrobce) v 1. řádku: $x_{12} = 10$, a tím se vyčerpala kapacita výrobce V_1, protože $x_{11} + x_{12} = 5 + 10 = 15$. Čili do polí (1,3) a (1,4) můžeme umístit pomlčku. A posuneme se směrem na jih – do pole (2,2), atd. až dostaneme výsledné vstupní rozdělení získané metodou SZR, uvedené v následující tabulce:

	S_1	S_2	S_3	S_4	
V_1	10 5	0 10	20 –	11 –	15
V_2	12 –	7 5	9 15	20 5	25
V_3	0 –	14 –	16 –	18 5	5
	5	15	15	10	

Celková cena dopravy při tomto rozdělení je

$$5 \cdot 10 + 10 \cdot 0 + 5 \cdot 7 + 15 \cdot 9 + 5 \cdot 20 + 5 \cdot 18 = 410 \text{ peněžních jednotek.}$$

□

Všimněme si důležité skutečnosti, že při m řádcích a n sloupcích dopravní tabulky má přiřazení dopravy tu vlastnost, že pomlčka není umístěna v $(m + n - 1)$ polích. Tento fakt musí platit pro počáteční rozdělení, i pro každý optimalizační krok. Při metodě SZR je platnost tohoto faktu ohrožena v situaci, kdy při jistém opakování kroku 2 proměnná x_{ij} vyčerpá kapacitu v řádku i i sloupci j současně – v takové situaci nemůžeme umístit pomlčky do zbylých polí v řádku i i sloupci j, protože nepomlčkových polí by bylo málo (méně než $m + n - 1$). Proto např. umístíme pomlčky jen do řádku i. Dále přiřadíme $x_{i+1,j} = 0$, a do zbylých polí sloupce j už můžeme umístit pomlčku. Chtěl bych upozornit na to, že je velký rozdíl v tom, zda je v určitém poli hodnota 0, nebo pomlčka. Tento rozdíl bude patrný v optimalizačním kroku, ale už nyní můžeme prozradit, že pomlčkové pole označuje nebázickou proměnnou, pole s hodnotou 0 označuje bázickou proměnnou s hodnotou 0 (čili jedná se o tzv. degenerované řešení, co se týká analogie mezi tímto algoritmem a simplexovou metodou). A protože při optimalizačním kroku musíme vědět, které proměnné jsou bázické a které ne, je třeba dobře rozlišovat mezi nulou a pomlčkou.

Příklad 10.5. Příklad degenerovaného počátečního rozdělení v dopravní úloze při metodě SZR:

	S_1	S_2	S_3	
V_1	5 (0)	– (2)	– (1)	5
V_2	0 (2)	5 (1)	5 (5)	10
V_3	– (2)	– (4)	5 (3)	5
	5	5	10	

Nepomlčková pole označují hodnoty bázických proměnných.

b) Indexová metoda (IM)
Tato metoda začíná přiřazením maximální možné hodnoty proměnné s ohledem na kapacity v poli s minimální cenou c_{ij} jednotkové dopravy (čísla v polích vpravo nahoře). Pak umístíme pomlčky v polích vyčerpaného řádku (nebo sloupce) a pokračujeme s přiřazováním v poli s cenou minimální ze všech polí zbývajících. Problém degenerace, kdy vyplnění proměnné x_{ij} vyčerpá řádek i sloupec současně, musíme opět řešit citlivě – umístíme pomlčky jen ve zbývajících polích řádku i a do některého dalšího pole ve sloupci j přiřadíme 0. Teprve pak sloupec j doplníme pomlčkami.

Ad Příklad 10.4. Připomeňme si tabulkovou formulaci dopravní úlohy

	S_1	S_2	S_3	S_4	
V_1	10	0	20	11	15
V_2	12	7	9	20	25
V_3	0	14	16	18	5
	5	15	15	10	

Minimální jednotkové ceny jsou $c_{12} = 0 = c_{31}$. Volíme např. pole (1,2) a přiřadíme maximální možný počet jednotek s ohledem na kapacity v řádku i sloupci: $x_{12} = 15$. Tím se vyčerpá řádek i sloupec, tj. bude se jednat o degenerované řešení. Do prvního řádku ve zbylých polích umístíme pomlčky, ve druhém sloupci musíme vyplnit jednu nulu, např. $x_{22} = 0$, abychom zaručili dostatečný počet bázických (= nepomlčkových) polí, a do zbylého pole (3,2) můžeme umístit pomlčku.

	S_1	S_2	S_3	S_4	
V_1	– (10)	15 (0)	– (20)	– (11)	15
V_2	(12)	0 (7)	(9)	(20)	25
V_3	(0)	– (14)	(16)	(18)	5
	5	15	15	10	

Další minimální zatím nevyplněné pole je určeno cenou $c_{31} = 0$. Vyplníme maximálně, tj. $x_{31} = 5$, což opět vyčerpá kapacitu současně třetího řádku i prvního sloupce, tj. třetí řádek doplníme pomlčkami a v prvním sloupci musíme přidat jednu nulu: $x_{21} = 0$.

	S_1	S_2	S_3	S_4	
V_1	$-^{10}$	15^{0}	$-^{20}$	$-^{11}$	15
V_2	0^{12}	0^{7}	^{9}	^{20}	25
V_3	5^{0}	$-^{14}$	$-^{16}$	$-^{18}$	5
	5	15	15	10	

Zbývá dokončit vyplnění posledního řádku, kde jsou hodnoty proměnných jednoznačně určeny (protože se jedná o vyváženou úlohu): $x_{23} = 15$, $x_{24} = 10$. Výsledné počáteční rozdělení dopravy má tedy tvar:

	S_1	S_2	S_3	S_4	
V_1	$-^{10}$	15^{0}	$-^{20}$	$-^{11}$	15
V_2	0^{12}	0^{7}	15^{9}	10^{20}	25
V_3	5^{0}	$-^{14}$	$-^{16}$	$-^{18}$	5
	5	15	15	10	

Celková cena dopravy v tomto případě je

$$10 \cdot 20 + 15 \cdot 9 = 335 \text{ finančních jednotek.}$$

Je pochopitelné, že počáteční rozdělení dopravy provedené indexovou metodou bude mít zpravidla vždy nižší celkovou cenu dopravy, protože metoda SZR vůbec nebrala ohled na ceny c_{ij}. A následující metoda by obecně měla být ještě lepší než indexová.

c) Vogelova aproximační metoda (VAM)
Tato metoda nejenže začíná vyplňovat pole s minimální cenou, ale navíc přidává kritérium, že ze všech polí s malými cenami začne tím polem, které splňuje navíc tzv. penalizační podmínku (a je jistým způsobem nejvhodnější ze vhodných).
Postup:

1) Vypočteme pro každý řádek i sloupec tzv. penalizaci, která je rovna rozdílu „(druhá nejmenší cena) minus (nejmenší cena)“.

2) Zvolíme řádek nebo sloupec s největší penalizací, vybereme zde pole s minimální cenou c_{ij} a vyplníme maximální možnou hodnotu x_{ij}. Dále doplníme pomlčky do vyčerpaného řádku nebo sloupce. Pokud vyplnění x_{ij} vyčerpá kapacitu řádku i sloupce současně, musíme vyplnit zbylá pole pomlčkami jen například v daném řádku, ve sloupci musíme umístit do jednoho pole 0 a až do zbylých pomlčky; zkrátka a dobře, pokud vyplnění jisté proměnné vyčerpá řádek i sloupec současně a hodnoty některých polí v tabulce v příslušném řádku nebo sloupci ještě nejsou vyplněny, vždy se musí do některého pole v tom řádku nebo v tom sloupci doplnit jedna 0.

3) Kroky 1 a 2 opakujeme tak dlouho, až v tabulce zbude pouze jeden řádek nebo sloupec, kterého vyplnění je jednoznačně dáno. Přitom při výpočtu penalizací už nebereme v úvahu ceny v polích, kde už je vyplněna hodnota nebo pomlčka.

Ad Příklad 10.4. Penalizace určíme jako rozdíl dvou nejnižších cen v daném řádku nebo sloupci:

	S_1	S_2	S_3	S_4	
V_1	$-\;^{10}$	^{0}	^{20}	^{11}	10
V_2	$0\;^{12}$	^{7}	^{9}	^{20}	2
V_3	$5\;^{0}$	$-\;^{14}$	$-\;^{16}$	$-\;^{18}$	$\boxed{14}$
	10	7	7	7	

Maximální penalizace 14 určuje třetí řádek, ve kterém vyplníme co největší hodnotu s ohledem na kapacity (ty zde teď nejsou napsány, ale jsou uvedeny na předchozí straně): $x_{31} = 5$, což vyčerpá kapacitu řádku i sloupce současně, čili např. přiřadíme $x_{21} = 0$ a zbylá pole v 1.sloupci a 3.řádku vyplníme pomlčkami.

Pokračujeme dalším krokem, čili znovu přepočteme penalizace řádků a sloupců, přičemž ve třetím řádku a prvním sloupci to už nemá smysl, a ani pro ostatní řádky a sloupce ceny z těchto vyplněných polí už nebudeme uvažovat:

	S_1	S_2	S_3	S_4	
V_1	$-\;^{10}$	^{0}	$-\;^{20}$	^{11}	11
V_2	$0\;^{12}$	^{7}	$15\;^{9}$	^{20}	2
V_3	$5\;^{0}$	$-\;^{14}$	$-\;^{16}$	$-\;^{18}$	
		7	$\boxed{11}$	9	

Dvě z penalizací jsou rovny 11, zvolme jednu z nich, například tu která určuje třetí sloupec. Ve třetím sloupci vybereme minimální $c_{23} = 9$ a přiřadíme maximálně: $x_{23} = 15$. Tím se vyčerpá kapacita sloupce, do posledního pole ve 3.sloupci, které zbývá vyplnit, píšeme pomlčku.

V další fázi jsou už penalizace celkem jednoznačné (vždy rozdíly dvou cen nevyplněných polí v daném řádku či sloupci):

	S_1	S_2	S_3	S_4	
V_1	$-\;^{10}$	^{0}	$-\;^{20}$	^{11}	11
V_2	$0\;^{12}$	$10\;^{7}$	$15\;^{9}$	$-\;^{20}$	$\boxed{13}$
V_3	$5\;^{0}$	$-\;^{14}$	$-\;^{16}$	$-\;^{18}$	
		7		9	

Maximální penalizaci určuje druhý řádek, vybereme v něm dosud nevyplněné pole s minimální c_{ij}: $x_{22} = 10$, tím je vyčerpána kapacita druhého řádku.

Nyní zbývá k vyplnění už jen poslední řádek, kterým je pozičně první řádek – doplníme vzhledem k požadavkům kapacit jednoznačně určené hodnoty $x_{12} = 5$, $x_{14} = 10$.

	S_1	S_2	S_3	S_4
V_1	$-^{10}$	5^{0}	$-^{20}$	10^{11}
V_2	0^{12}	10^{7}	15^{9}	$-^{20}$
V_3	5^{0}	$-^{14}$	$-^{16}$	$-^{18}$

Celková cena dopravy je

$$5 \cdot 0 + 10 \cdot 11 + 0 \cdot 12 + 10 \cdot 7 + 15 \cdot 9 + 5 \cdot 0 = 315 \text{ finančních jednotek,}$$

což je ještě lepší výsledek než počáteční rozdělení získané metodou indexovou. Počáteční rozdělení nalezené pomocí VAM je už dost dobré, ale stále to nemusí být řešení nejlepší. Takové algoritmy, které sice nenaleznou optimální, ale jakési dost dobré řešení, nazýváme *heuristikami*.

Ukázali jsme si tedy tři metody, které naleznou počáteční rozdělení dopravy v dopravní úloze. Čtenář tohoto textu by si možná mohl položit otázku, zda je nutné tolik metod a zda nestačí hodnoty do tabulky pro začátek jakýmkoliv způsobem, který vyhovuje kapacitním požadavkům řádků a sloupců. Odpovím na tuto otázku, i když si ji čtenář zrovna nepokládá:

Ne každé rozdělení dopravy je povoleným počátečním rozdělením – zakázána jsou rozdělení, která obsahují tzv. uzavřený okruh, což je pravoúhelník s hranami vodorovnými či svislými a s vrcholy v nepomlčkových polích.

Ad Příklad 10.4. Např. rozdělení

–	5	10	–
–	10	5	10
5	–	–	–

uzavřeným okruhem je zde čtverec (1,2)–(1,3)–(2,3)–(2,2)

nebo rozdělení

0	–	15	–
–	15	–	10
5	–	0	–

uzavřeným okruhem je zde čtverec (1,1)–(1,3)–(3,3)–(3,1)

není povoleným počátečním rozdělením. Tedy počáteční rozdělení dopravy lze provést i jiným způsobem než uvedenými metodami, ale takovému vyplnění nepomlčkových polí, které vytváří vrcholy uzavřeného okruhu, se musíme vyvarovat (na vrcholy uzavřeného okruhu nemusí být využita všechna nepomlčková pole). Metody a), b), c) jsou konstruovány tak, aby při jejich použití uzavřený okruh pomocí nepomlčkových polí nevznikl.

B. Optimalizační krok

Tento krok vysvětlíme na příkladě 10.4 – vezmeme počáteční rozdělení dopravy získané metodou SZR a budeme je dále zlepšovat. Protože se jedná o úlohu lineárního programování, optimalizační krok sestává ze tří částí:

a) určení vstupní proměnné (= vstupního pole) ze všech nebázických (= pomlčkových polí)

b) určení výstupní proměnné (= výstupního pole, které se stane v dalším optimalizačním kroku pomlčkovým)

c) provedení optimalizačního kroku (= přepočítání nového bázického řešení)

Optimalizační krok č.1:

a) určení vstupní proměnné: Nejprve pro dané rozdělení dopravy přiřadíme každému řádku i každému sloupci tzv. multiplikátory u_i, v_j, které jsou řešením systému rovnic

$$u_i + v_j = c_{ij}$$

pro nepomlčková pole:

	$v_1 =$	$v_2 =$	$v_3 =$	$v_4 =$
$u_1 = 0$	5^{10}	10^{0}	$-^{20}$	$-^{11}$
$u_2 =$	$-^{12}$	5^{7}	15^{9}	5^{20}
$u_3 =$	$-^{0}$	$-^{14}$	$-^{16}$	5^{18}

Důležité upozornění:

1) pravé strany rovnic jsou tvořeny cenami c_{ij} v pravém horním rohu, nikoli počtem jednotek uprostřed pole
2) neznámých je o jednu více než rovnic. Protože nám stačí najít jedno z těch nekonečně mnoha řešení, jednu neznámou můžeme zvolit. Zpravidla volíme $u_1 = 0$.

Nyní pro pole $(1,1)$ platí $u_1 + v_1 = 10 \Rightarrow v_1 = 10$ (protože $u_1 = 0$).
Pokračujeme dalším nepomlčkovým polem $(1,2)$: $u_1 + v_2 = 0 \Rightarrow v_2 = 0$.
Dále pole $(2,2)$: $u_2 + v_2 = 7 \Rightarrow u_2 = 7$ (protože $v_2 = 0$).
Dále $(2,3)$: $u_2 + v_3 = 9 \Rightarrow v_3 = 2$ (protože $u_2 = 7$).
Pole $(2,4)$: $u_2 + v_4 = 20 \Rightarrow v_4 = 14$.
A konečně pole $(3,4)$: $u_3 + v_4 = 18 \Rightarrow u_3 = 4$.
A tím jsou všechna nepomlčková pole vyčerpána. Z popsaného řešení je vidět, že rovnice řešíme v takovém pořadí, že při řešení dané rovnice už hodnotu jedné neznámé známe – pak druhou neznámou můžeme lehce dopočítat. Tohle jsme prováděli pouze pro pole, ve kterých se nevyskytuje pomlčka!

Nyní po určení multiplikátorů pokročíme k pomlčkovým polím a určíme pro ně jakési přepočtené ceny $\overline{c}_{pq}$ ze vztahu

$$\overline{c}_{pq} = u_p + v_q - c_{pq} \qquad \text{(pouze pro pomlčková pole!).}$$

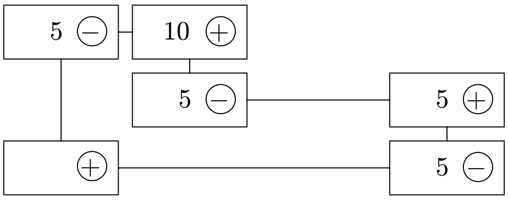
Pomocné ceny $\overline{c}_{pq}$ v pomlčkových polích budeme psát například vlevo dole (výpočet je přirozený: u_i v příslušném řádku plus v_j v příslušném sloupci minus cena c_{ij}), tedy $\overline{c}_{13} = u_1 + v_3 - c_{13} = 0 + + 2 - 20 = -18$, atd.

	$v_1 = 10$	$v_2 = 0$	$v_3 = 2$	$v_4 = 14$
$u_1 = 0$	5^{10}	10^{0}	$-^{20}$ -18	$-^{11}$ 3
$u_2 = 7$	$-^{12}$ 5	5^{7}	15^{9}	5^{20}
$u_3 = 4$	$-^{0}$ 14	$-^{14}$ -10	$-^{16}$ 0	5^{18}

A nyní tedy konečně určení vstupního pole: za vstupní proměnnou vezmeme tu na pozici s maximální kladnou cenou $\overline{c}_{pq}$ (podobně jako při minimalizaci v jakékoli jiné úloze LP). Pokud jsou všechny $\overline{c}_{pq} \leq 0$, dané rozdělení dopravy je optimální.

V našem případě $\max \overline{c}_{pq} = 14$ na pozici $(3,1)$ ... máme určeno vstupní pole.

b) určení výstupního pole: Uvažujme teď bázická (=nepomlčková) pole společně s dodávaným polem $(3,1)$. Bázická pole jsou maximální množinou polí, která nevytváří (neobsahují) uzavřený okruh. Dodáním vstupního pole do této množiny se v této množině uzavřený okruh už vytvoří – a sice vždy právě jeden!
V našem případě dodáním pozice $(3,1)$ do množiny $\{(1,1),(1,2),(2,2),(2,3),(2,4),(3,4)\}$ vznikne uzavřený okruh $(3,1)-(1,1)-(1,2)-(2,2)-(2,4)-(3,4)$ (pozici $(2,3)$ neuvažujeme, protože pole $(2,3)$ není vrcholem daného okruhu).



Vstupní pole nyní vždy označíme znakem ⊕ , sousední vrcholy vstupního pole označme ⊖ , sousední pole polí ⊖ označme ⊕ . Zkrátka a dobře, při postupném procházení vrcholů uzavřeného okruhu se střídají znaménka. Jaký význam udávají znaménka $+,-$? Přičteme-li do polí označených ⊕ určitý počet jednotek dopravy a odečteme-li z polí ⊖ tutéž hodnotu, příslušné kapacity výrobců a spotřebitelů ve všech řádcích i sloupcích zůstanou v platnosti (protože v každém řádku nebo sloupci je vždy jedno pole ⊕ a jedno pole ⊖).

Nyní tedy určení výstupního pole: za výstupní pole vybereme to z polí ⊖ , které obsahuje minimální počet jednotek dopravy. Pokud je těchto minimálních hodnot více (jako v našem příkladu, kdy všechna pole ⊖ obsahují stejnou minimální hodnotu 5), vybereme jedno z nich. Zvolme tedy za výstupní např. pole (3,4).

c) provedení optimalizačního kroku: Byla-li záměna bázické proměnné v případě o-becné úlohy LP početně nejsložitější, u dopravní úlohy se pro změnu jedná o nejjednodušší část optimalizačního kroku: do polí ⊕ přičteme hodnotu z výstupního pole, z polí ⊖ ji odečteme. Ostatní pole zůstanou nezměněna. Přitom ve výstupním poli se objeví pomlčka.

Pokud by se stalo (jako je tomu v našem příkladě), že výstupní pole nebylo určeno jednoznačně, i tak napíšeme pomlčku pouze na pozici $(3,4)$, do pozic $(1,1)$ a $(2,2)$ musíme napsat 0. Zkrátka a dobře, nepomlčkových (=bázických) polí musí být vždy $m+n-1$ (tedy v našem případě 6). Už víme, proč to je důležité: nepomlčková hodnota jednotek dopravy (tedy i 0) označuje pole sloužící k určení multiplikátorů u_i, v_j, kdežto pomlčka označuje pole sloužící k výpočtu pomocné ceny $\overline{c}_{pq}$.

Dostáváme tedy rozdělení dopravy

0^{10}	15^{0}	$-^{20}$	$-^{11}$
$-^{12}$	0^{7}	15^{9}	10^{20}
5^{0}	$-^{14}$	$-^{16}$	$-^{18}$

Optimalizační krok č.2:

a) určení vstupního pole: Musíme znovu přepočítat multiplikátory u_i, v_j, protože některé z nich budou jiné díky změně nepomlčkových pozic. Volíme $u_1 = 0$ a další neznámé dopisujeme do tabulky ve vhodném pořadí pomocí nepomlčkových polí:

	$v_1 = 10$	$v_2 = 0$	$v_3 = 2$	$v_4 = 13$
$u_1 = 0$	0^{10}	15^{0}	$-18\;{-}^{20}$	$2\;{-}^{11}$
$u_2 = 7$	$5\;{-}^{12}$	0^{7}	15^{9}	10^{20}
$u_3 = -10$	5^{0}	$-24\;{-}^{14}$	$-24\;{-}^{16}$	$-15\;{-}^{18}$

Nyní do téže tabulky (v rámci urychlení času) vpisujeme do pomlčkových polí pomocné ceny $\overline{c}_{pq}$ (vlevo dole). Protože $\max \overline{c}_{pq} = 5 \Rightarrow$ vstupní pole je $(2, 1)$.

b) určení výstupního pole: Uzavřený okruh vzniklý po dodání vstupního pole k bázickým je

$0 \ominus$	$15 \oplus$
$\oplus$	$0 \ominus$

Obě pole $\ominus$ obsahují stejnou minimální hodnotu $\min\limits_{\text{pole } \ominus} x_{ij} = 0 \Rightarrow$ výstupní pole opět není určeno jednoznačně, volíme např. $(1, 1)$.

c) provedení optimalizačního kroku: Jedinou změnou v tomto kroku je, že pomlčka a nula v polích $(1, 1)$ a $(2, 1)$ se vyměnily – ale to, jak víme, je změna podstatná.

${-}^{10}$	15^{0}	${-}^{20}$	${-}^{11}$
0^{12}	0^{7}	15^{9}	10^{20}
5^{0}	${-}^{14}$	${-}^{16}$	${-}^{18}$

Optimalizační krok č.3:

a) určení vstupního pole: Pomocí nepomlčkových polí určíme u_i, v_j. Pak v pomlčkových polích určíme pomocné ceny $\overline{c}_{pq}$ (vlevo dole).

	$v_1 = 5$	$v_2 = 0$	$v_3 = 2$	$v_4 = 13$
$u_1 = 0$	$-5\;{-}^{10}$	15^{0}	$-18\;{-}^{20}$	$2\;{-}^{11}$
$u_2 = 7$	0^{12}	0^{7}	15^{9}	10^{20}
$u_3 = -5$	5^{0}	$-19\;{-}^{14}$	$-19\;{-}^{16}$	$-10\;{-}^{18}$

$\max \overline{c}_{pq} = 2 \Rightarrow$ vstupní pole je $(1, 4)$.

b) určení výstupního pole: Vzniklý uzavřený okruh je

15 $\ominus$	$\oplus$
0 $\oplus$	10 $\ominus$

tedy $\min\limits_{\text{pole } \ominus} x_{ij} = 10 \Rightarrow (2, 4)$ je výstupní pole.

c) provedení optimalizačního kroku: Hodnotu 10 odečteme z polí $\ominus$ a přičteme do polí $\oplus$:

$-^{10}$	5^{0}	$-^{20}$	10^{11}
0^{12}	10^{7}	15^{9}	$-^{20}$
5^{0}	$-^{14}$	$-^{16}$	$-^{18}$

Optimalizační krok č.4:
Pro vzniklé rozdělení dopravy vypočteme u_i, v_j a ceny $\overline{c}_{pq}$:

	$v_1 = 5$	$v_2 = 0$	$v_3 = 2$	$v_4 = 11$
$u_1 = 0$	-5 $-^{10}$	5^{0}	-18 $-^{20}$	10^{11}
$u_2 = 7$	0^{12}	10^{7}	15^{9}	-2 $-^{20}$
$u_3 = -5$	5^{0}	-19 $-^{14}$	-19 $-^{16}$	-12 $-^{18}$

Protože všechny $\overline{c}_{pq} < 0$, dané rozdělení dopravy je optimální.

Co k tomu dodat? Snad jakési vysvětlení celého algoritmu: u_i, v_j jsou vlastně duální proměnné k dopravní úloze, $u_i + v_j \leq c_{ij}$ jsou duální omezení, ve kterých pro bázické proměnné musí nastat rovnost ($u_i + v_j = c_{ij}$). A $\overline{c}_{pq}$ jsou koeficienty účelové funkce nebázických proměnných v k-tém kroku (proto $\overline{c}_{pq} = u_p + v_q - c_{pq}$).

10.3 Přiřazovací úloha

Do logického sledu výkladu patří následující úloha, kteráž slove přiřazovací:
Jak přiřadit m úkolů (dělníků, pracovníků) k n strojům, aby celková cena nákladů byla minimální (resp. celkový užitek byl maximální)?

Tato úloha je úlohou LP, protože ji lze formulovat následovně:
Nalezněte minimum funkce $z = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} c_{ij} x_{ij}$, kde c_{ij} je cena přiřazení úlohy i ke stroji j za omezení

$$\left.\begin{array}{l}\sum\limits_i x_{ij} = 1\\ \sum\limits_j x_{ij} = 1\end{array}\right\}\quad \text{pro } i, j = 1, 2, \ldots, n,$$

přičemž x_{ij} může nabývat hodnoty 0 (i-tý úkol není přiřazen j-tému stroji) nebo 1 (i-tý úkol je přiřazen j-tému stroji).
Je to vlastně úloha binárního programování. Ba co víc, jedná se i o dopravní problém, kde výrobce představují úkoly a spotřebitele představují stroje:

$$\begin{array}{ccc}
 & & \text{stroje} \\
 & & \begin{array}{cccc} 1 & 2 & \ldots & n \end{array} \\
\text{úkoly} & \begin{array}{c} 1 \\ 2 \\ \vdots \\ m \end{array} & \begin{array}{|cccc|} \hline c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \vdots & & & \\ c_{m1} & c_{m2} & \ldots & c_{mn} \\ \hline \end{array} \begin{array}{c} 1 \\ 1 \\ \vdots \\ 1 \end{array} \\
 & & \begin{array}{cccc} 1 & 1 & \ldots & 1 \end{array}
\end{array}$$

Kapacity úkolů a požadavky strojů jsou rovny 1. Dopravní úlohu lze řešit po vyvážení (zaručení toho, že platí $m = n$).

Příklad 10.6. Následující přiřazení pro $m = n = 3$ bylo získáno metodou SZR:

stroje

úkoly	1	2	3	
1	$1\ ^{5}$	$0\ ^{7}$	$–\ ^{9}$	1
2	$–\ ^{14}$	$1\ ^{10}$	$0\ ^{12}$	1
3	$–\ ^{15}$	$–\ ^{13}$	$1\ ^{16}$	1
	1	1	1	

(řešení přiřazení je degenerované v každém optimalizačním kroku).

Zvláštní struktura přiřazovacího modelu ovšem dovoluje vyvinout rychlejší metodu, než je algoritmus dopravní úlohy. Využijeme následujícího faktu:

Věta 10.7. *Pokud k řádku nebo sloupci matice* (c_{ij}) *je přičtena konstanta, bod optima se nemění, jen se o konstantu změní hodnota účelové funkce.*

Důkaz. pokud od i-tého řádku odečteme konstantu p_i
j-tého sloupce odečteme konstantu q_j,

dostaneme nový cenový koeficient $c'_{ij} = c_{ij} - p_i - q_j$, a pak

$$\begin{aligned}
z' &= \sum_{i=1}^{n}\sum_{j=1}^{n} c'_{ij}x_{ij} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}(c_{ij} - p_i - q_j)x_{ij} \\
&= z - \sum_{i=1}^{n} p_i \underbrace{\sum_{j=1}^{n} x_{ij}}_{1} - \sum_{j=1}^{n} q_j \underbrace{\sum_{i=1}^{n} x_{ij}}_{1} \\
&= z - \text{konst.}]
\end{aligned}$$

□

Vytvoříme tedy algoritmus pro případ minimalizace nákladů přiřazení. Hlavní myšlenkou algoritmu je:

Pokud můžeme v matici (c'_{ij}) vyrobit tolik nul, že z nich lze vybrat bázi, je tato báze optimální bod, protože cena z' je nulová (a záporná být nemůže).

Ad Příklad 10.6. Připomeňme si tabulkovou formulaci úlohy

5	7	9
14	10	12
15	13	16

Od každého řádku odečteme minimální cenu: $p_1 = 5$
$p_2 = 10$
$p_3 = 13.$

Získáme matici:

0	2	4
4	0	2
2	0	3

Odečteme ještě od každého sloupce minimální cenu v tomto sloupci: $q_1 = 0$
$q_2 = 0$
$q_3 = 2.$

Vznikla matice (c'_{ij}):

$_{1}$ 0	2	2
4	0	$_{1}$ 0
2	$_{1}$ 0	1

Ze čtyř polí oceněných nulou nyní lze vybrat bázické pozice – je to možné jediným způsobem. Tento úkol je analogický úkolu najít na šachovnici 3×3 z jistých pozic takovou kombinaci rozestavění věží tak, aby jedna druhou navzájem neohrožovaly. Tedy $x_{11} = x_{23} = x_{32} = 1$, ostatní $x_{ij} = 0$. Všimněme si také, že $z = p_1 + p_2 + p_3 + q_3 = 30$.

Někdy ovšem pouhé odečítání konstanty od řádku nebo sloupce k nalezení optima nevede a je nutno provést optimalizační krok. To vysvětlíme následujícím příkladem.

Příklad 10.8. Úloha je dána následující tabulkou:

		stroje				
		1	2	3	4	
	1	1	4	6	3	$p_1 = 1$
úkoly	2	9	7	10	9	$p_2 = 7$
	3	4	5	11	7	$p_3 = 4$
	4	8	7	8	5	$p_4 = 5$

Po odečtení konstant od řádků dostáváme tabulku:

0	3	5	2
2	0	3	2
0	1	7	3
3	2	3	0
		$q_3 = 3$	

Po odečtení konstant i od sloupců nastává situace, kdy nelze z polí ohodnocených nulou vybrat bázi:

0	3	2	2
2	0	0	2
0	1	4	3
3	2	0	0

Musíme provést **optimalizační krok:** Co nejmenším počtem vodorovných a svislých přímek vyškrtáme všechna pole ohodnocená nulou. Pak ze všech nepřeškrtnutých polí vybereme nejmenší cenu, kterou odečteme ode všech nepřeškrtnutých polí a přičteme k ceně v průsečících „škrtacích“ přímek.

V našem případě je minimální cena v nepřeškrtnutých polích rovna 1; provedením změn optimalizačního kroku získáváme cenové ohodnocení, ve kterém už lze nalézt optimum $x_{11} = x_{23} = x_{32} = x_{44} = 1$. V opačném případě bychom museli pokračovat dalším optimalizačním krokem.

${}_{1}\,0$	2	1	1
3	0	${}_{1}\,0$	2
0	${}_{1}\,0$	3	2
4	2	0	${}_{1}\,0$

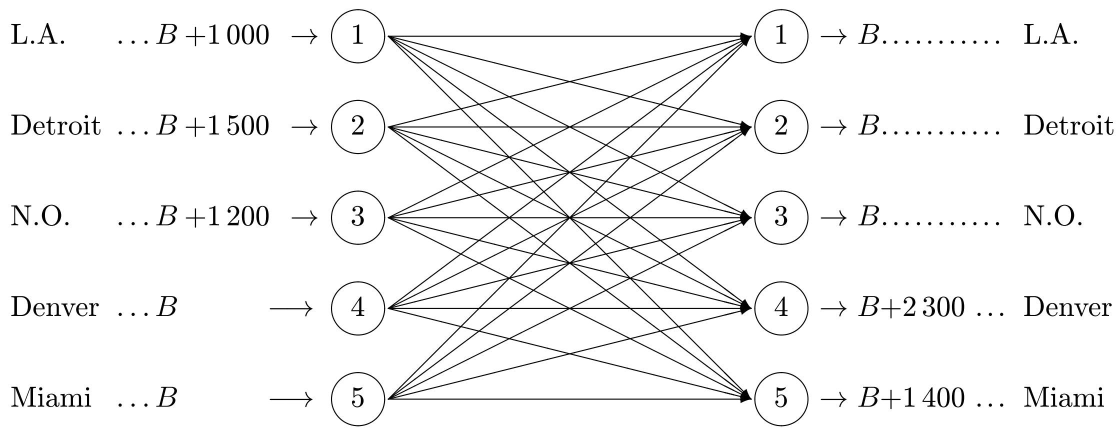
Pokud by cenové ohodnocení v přiřazovací úloze vyjadřovalo zisk nebo obrat a naším úkolem by bylo najít maximum přiřazení, odečtením všech cen od maximální ceny v tabulce bychom získali jakési ohodnocení popsané cenovými zbytky (rozdíly) n_{ij}, na které pak už lze nasadit právě popsaný algoritmus minimalizace. Převod původních cen na cenové zbytky:

$$n_{ij} = \max_{i,j} \{c_{ij}\} - c_{ij}.$$

10.4 Problém překladu materiálu

Jedná se o modifikaci dopravní úlohy, kdy se některým výrobkům v modelu povoluje procházet jiné zdroje a spotřebitele, než dorazí k tomu svému – každý uzel sítě může být současně zdrojem i spotřebitelem.

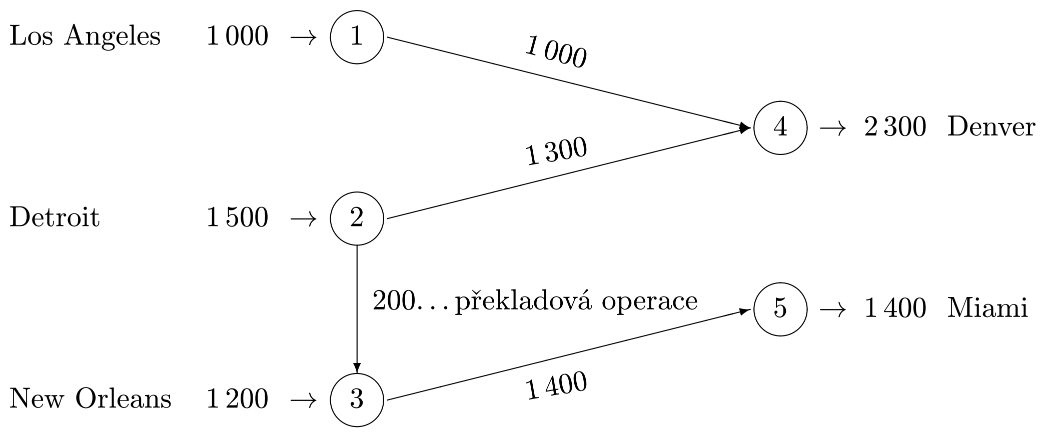
Ad Příklad 10.1.a) Uvažujme zadání z příkladu 10.1.a); doplňme ještě požadavek, že každý zdroj i spotřebitel může být navíc ještě dopravním uzlem, tj. dostáváme překladový model pro 5 zdrojů a 5 spotřebitelů. Příslušná síťová formulace, kde $B = \sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j = 3700$ označuje povolené množství překládaného materiálu, je tvaru:



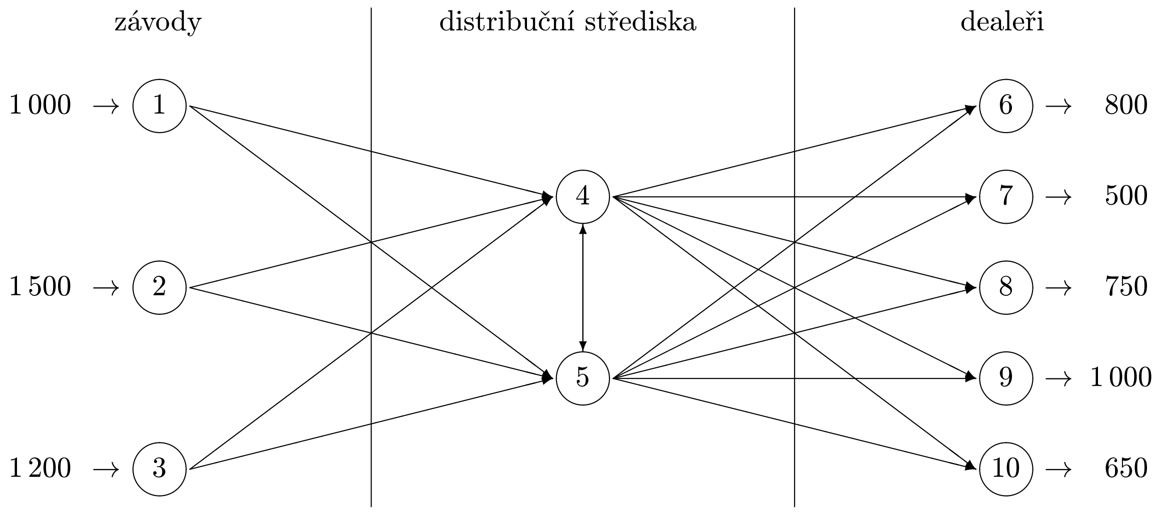
Některé jednotkové ceny jsou stejné jako v původní úloze; jednotková cena dopravy z místa do sebe sama je rovna nule. Cena z X do Y může být jiná než cena dopravy z Y do X (v opačném směru je použit jiný druh dopravy).
Pro $B = 3\,700$ je optimální řešení úlohy v tabulce:

	L.A.	Detroit	N.O.	Denver	Miami	
L.A.	3 700 ^{0}	– ^{130}	– ^{90}	1 000 ^{80}	– ^{215}	4 700
Detroit	– ^{135}	3 700 ^{0}	200 ^{101}	1 300 ^{100}	– ^{108}	5 200
N.O.	– ^{95}	– ^{105}	3 500 ^{0}	^{102}	1 400 ^{68}	4 900
Denver	– ^{79}	– ^{99}	– ^{110}	3 700 ^{0}	^{205}	3 700
Miami	– ^{200}	– ^{107}	– ^{72}	– ^{205}	3 700 ^{0}	3 700
	3 700	3 700	3 700	6 000	5 100	

Toto řešení lze mimo hodnoty na hlavní diagonále znázornit i síťově:



Rozšiřme nyní příklad 10.1.a) o fakt, že spotřebitelé (4), (5) ještě zboží rozváží pěti dealerům, tj. uzly (4) a (5) jsou jediné překladové uzly:



Není dovolena doprava ze závodů přímo k dealerům – vše se musí přepravit přes uzly 4 a 5 ($B = 3\,700$ je přidáno pouze zdrojům a spotřebitelům 4 a 5).
Tabulková reprezentace dané úlohy:

	4	5	6	7	8	9	10	
1			M	M	M	M	M	1 000
2			M	M	M	M	M	1 500
3			M	M	M	M	M	1 200
4								3 700
5								3 700
	3 700	3 700	800	500	750	1 000	650	

Dostatečně velká konstanta M v příslušné tabulkové reprezentaci síťové úlohy zaručuje, že daná souřadnice nebude vybrána pro optimální rozdělení dopravy. Poněvadž je povolen překlad pouze v uzlech 4 a 5 , cesty zpět z uzlů 4 , 5 do 1 , 2 , 3 v tabulce neuvažujeme. Pokud bychom povolili překlad v bodech 1 až 5 , příslušná tabulková reprezentace už tyto cesty zpět musí zapracovat:

	1	2	3	4	5	6	7	8	9	10	
1						M	M	M	M	M	4 700
2						M	M	M	M	M	5 200
3						M	M	M	M	M	4 900
4											3 700
5											3 700
	3 700	3 700	3 700	3 700	3 700	800	500	750	1 000	650	

Pojmy k zapamatování

– Úvodem kapitoly jsme formulovali zadání dopravní úlohy: rozhodněte o způsobu rozvozu jednoho typu výrobku (nebo suroviny) z m závodů do n spotřebitelských skladů tak, aby se minimalizovala celková cena dopravy, je-li dáno

c_{ij} ... jednotková cena dopravy ze zdroje i na místo určení j,
a_i ... množství zásob se zdroji i, $i = 1, 2, \ldots, m$,
b_j ... požadavek spotřebitele j, $j = 1, 2, \ldots, n$.

Výstupem úlohy jsou optimální hodnoty proměnných x_{ij} (= množství jednotek dopravovaných ze zdroje i ke spotřebiteli j) a celková cena dopravy $\sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$.

– Model dopravní úlohy lze řešit až po eventuálním převedení na tzv. *vyvážený tvar*, kdy $\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$, tj. nabídka = poptávka. Tohoto tvaru lze dosáhnout zavedením fiktivního výrobce nebo fiktivního spotřebitele. Cenu dopravy u fiktivních členů našeho systému položíme rovnu nule, případně volíme dostatečně velké kladné číslo (záleží na požadavcích úlohy).

– Dále jsme se zabývali již samotným řešením dopravního problému, které lze rozdělit do následujících fází:

A. Počáteční krok: V počátečním kroku nalezneme nějaké přípustné rozdělení dopravy, které nemusí být nutně optimálním. Ukázali jsme si 3 metody pro nalezení počátečního rozdělení

a) **Metoda severozápadního rohu (SZR)**
Metoda přiděluje maximální možnou hodnotu dopravovaných jednotek do jednotlivých polí tak, že začne na pozici $(1, 1)$ (severozápadní roh tabulky) a po-

kračuje jihovýchodním směrem. Do ostatních polí, kde je již hodnota přepravovaných jednotek vyčerpána, přiřadíme pomlčku.

b) **Indexová metoda (IM)**
Tato metoda začíná přiřazením maximální možné hodnoty proměnné s ohledem na kapacity v poli s minimální cenou c_{ij} jednotkové dopravy (čísla v polích vpravo nahoře). Pak umístíme pomlčky v polích vyčerpaného řádku (nebo sloupce) a pokračujeme s přiřazováním v poli s cenou minimální ze všech polí zbývajících.

c) **Vogelova aproximační metoda (VAM)**
Tato metoda nejenže začíná vyplňovat pole s minimální cenou, ale navíc přidává kritérium, že ze všech polí s malými cenami začne tím polem, které splňuje navíc tzv. penalizační podmínku.

B. Optimalizační krok Protože se jedná o úlohu lineárního programování, optimalizační krok sestává ze tří částí:

a) určení vstupní proměnné (= vstupního pole) ze všech nebázických (= pomlčkových polí)
b) určení výstupní proměnné (= výstupního pole, které se stane v dalším optimalizačním kroku pomlčkovým)
c) provedení optimalizačního kroku (= přepočítání nového bázického řešení)

– Modifikací dopravního problému je tzv. *přiřazovací úloha*, jejíž formulace zní:
Jak přiřadit m úkolů (dělníků, pracovníků) k n strojům, aby celková cena nákladů byla minimální (resp. celkový užitek byl maximální)? Tato úloha je úlohou LP, protože ji lze formulovat následovně:
Nalezněte minimum funkce $z = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij}$, kde c_{ij} je cena přiřazení úlohy i ke stroji j za omezení

$$\left.\begin{array}{l} \sum_i x_{ij} = 1 \\ \sum_j x_{ij} = 1 \end{array}\right\} \quad \text{pro } i, j = 1, 2, \ldots, n,$$

přičemž x_{ij} může nabývat hodnoty 0 (i-tý úkol není přiřazen j-tému stroji) nebo 1 (i-tý úkol je přiřazen j-tému stroji).

– Řešení tohoto problému je velmi podobné řešení dopravní úlohy a je demonstrováno na příkladě. Využívá se zde toho, že proměnná x_{ij} je binární.

– Další modifikací dopravní úlohy je tzv. *problém překladu materiálu*, kdy se některým výrobkům v modelu povoluje procházet jiné zdroje a spotřebitele, než dorazí k tomu svému – každý uzel sítě může být současně zdrojem i spotřebitelem. Při řešení je třeba uvědomit si správnou tabulkovou reprezentaci zadané úlohy. Algoritmus je pak stejný jako u dopravního problému.

Kontrolní otázky

1. U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.

a) Při řešení dopravního problému musíme úlohu upravit tak, aby byl stejný počet „zdrojů“ i „spotřebitelů“, tj. $m = n$.
b) Při řešení dopravního problému můžeme přidávat fiktivní „zdroje“ nebo fiktivní „spotřebitele“.
c) Při použití některé z metod SZR, IM nebo VAM nalezneme vždy přípustné řešení.
d) Metoda SZR může skončit v jakémkoliv rohu dopravní tabulky.
e) Metoda IM je zpravidla lepší než SZR, protože bere také ohled na ceny c_{ij}.
f) Při použití metody VAM nemůže nastat problém degenerovaného řešení.
g) Při použití metody SZR může vzniknout řešení, které obsahuje uzavřený okruh pomocí nepomlčkových polí.
h) Bázická pole mohou tvořit uzavřený okruh.
i) Bázických polí musí být vždy $m + n$.
j) Při řešení přiřazovacího problému musíme úlohu upravit tak, aby byl stejný počet „úkolů“ i „strojů“, tj. $m = n$.
k) Dopravním algoritmem lze řešit pouze úlohy, které přímo souvisí s dopravní tematikou.

Odpovědi na otázky
1a) – N, 1b) – A, 1c) – A, 1d) – N, 1e) – A, 1f) – N, 1g) – N, 1h) – N, 1i) – N, 1j) – A, 1k) – N

Cvičení

1. Najděte počáteční řešení v následující dopravní úloze

   a) metodou SZR
   b) indexovou metodou
   c) metodou VAM.

   Užitím nejlepšího počátečního řešení vypočtěte optimální řešení.

	S_1	S_2	S_3	S_4	
V_1	10	20	5	7	10
V_2	13	9	12	8	20
V_3	4	15	7	9	30
V_4	14	7	1	0	40
V_5	3	12	5	19	50
	60	60	20	10	

2. Najděte počáteční řešení v následující dopravní úloze

   a) metodou SZR

   b) indexovou metodou

   c) metodou VAM.

   Užitím nejlepšího počátečního řešení vypočtěte optimální řešení.

	S_1	S_2	S_3	S_4	
V_1	23	27	16	18	30
V_2	12	17	20	51	40
V_3	22	28	12	32	53
	22	35	25	41	

3. Uvažujme dopravní úlohu pro případ tří výrobců a tří spotřebitelů:

	S_1	S_2	S_3	
V_1	1	0	2	4
V_2	3	5	4	6
V_3	1	2	3	10
	3	5	12	

   Čísla v 5.řádku jsou kapacity spotřebitelů, čísla v 5.sloupci tabulky kapacity výrobců, jinak jsou v tabulce uvedeny ceny dopravy.

   a) Metodou severozápadního rohu určete přípustnou vstupní verzi.

   b) Najděte optimální rozdělení dopravy dostatečným opakováním optimalizačního kroku.

4. Vyřešte následující přiřazovací úlohu, kde se mají optimálně přiřadit 4 operátoři ke 4 strojům. Přitom se požaduje, aby 1.operátor nebyl přidělen ke 3.stroji a 3.operátor nebyl přidělen ke 4.stroji. Tabulka udává ceny jednotlivých přiřazení, chceme minimalizovat celkovou cenu.

	Stroj 1	Stroj 2	Stroj 3	Stroj 4
Operátor 1	5	5	–	2
Operátor 2	7	4	2	3
Operátor 3	9	3	5	–
Operátor 4	7	2	6	7

Výsledky příkladů viz **??**.

Výsledky

ad 1. Metoda VAM dávala nejlepší počáteční řešení. Byly třeba 3 iterační kroky k nalezení optimálního řešení $x_{13} = 10$, $x_{22} = 20$, $x_{31} = 30$, $x_{42} = 30$, $x_{44} = 10$, $x_{51} = 30$, $x_{52} = x_{53} = 10$ a $z = 820$.

ad 2. Metoda VAM našla přímo optimální řešení $x_{14} = 30$, $x_{21} = 5$, $x_{22} = 35$, $x_{31} = 17$, $x_{33} = 25$, $x_{34} = 11$ a $z = 2221$.

ad 3. Optimální řešení je $x_{12} = 4$, $x_{23} = 6$, $x_{31} = 3$, $x_{32} = 1$, $x_{33} = 6$ a $z = 47$.

ad 4. Optimální řešení je $x_{14} = x_{23} = x_{32} = x_{41} = 1$, $x_{ij} = 0$ pro ostatní i, j, celková cena je 14.

11 Dynamické programování

Průvodce studiem

V této kapitole se budeme zabývat další možnost hledání extrému funkce za daných omezení, tzv. dynamickým programováním. Přitom si ukážeme, že při hledání optimální možnosti je někdy nejrychlejší rozdělit daný problém na několik podproblémů (fází) a najít optimum v dané fázi se zřetelem na kombinace možností z předchozí fáze.

Nejtěžším úkolem u úloh dynamického programování je vhodně formulovat zadaný problém. Je třeba správně definovat fáze, stavy a přípustné alternativy.

Ukážeme si dva možné typy algoritmu: přímý postup a zpětný postup. Použití přímého nebo zpětného postupu algoritmu v praxi závisí na tom, který typ rekurentních rovnic je jednodušší.

Při řešení úloh dynamického programování nemusí být funkce, jejíž optimum hledáme, nutně lineární.

Cíle

Po prostudování této kapitoly budete schopni:

- Formulovat úlohu dynamického programování.
- Vhodně definovat jednotlivé fáze, stavy a možné alternativy.
- Hledat optimální řešení problémů dynamického programování.

Hlavní myšlenka dynamického programování: Při hledání optimální možnosti je někdy nejrychlejší rozdělit daný problém na několik podproblémů (fází) a najít optimum v dané fázi se zřetelem na kombinace možností z předchozí fáze – algoritmus je tedy rekurzivní. Podrobněji jej vysvětlíme na následujícím příkladu.

Příklad 11.1. Společnost chce investovat 5 miliónů dolarů do rozvoje svých tří závodů. Každý ze závodů předložil několik návrhů, jak by peníze mohl využít a s jakým výnosem. Jak rozdělit peníze, aby byl celkový výnos maximální?

	závod 1		závod 2		závod 3	
návrh	c_1	R_1	c_2	R_2	c_3	R_3
1	0	0	0	0	0	0
2	1	5	2	8	1	3
3	2	6	3	9	–	–
4	–	–	4	12	–	–

Jednotkou je jeden milión dolarů; nulový návrh u každého ze závodů naznačuje, že bereme v úvahu i možnost, že určitá část investic nemusí být vyčerpána.

Řešení. Nejprve musíme definovat a) fáze j

b) stavy x_j ve fázi j

c) přípustné alternativy k_j ve fázi j.

Existují dva možné typy algoritmu: přímý postup a zpětný postup.

1) Přímý postup:

ad a) fáze = závod

ad b) x_1 ... množství investic přidělené závodu 1
x_2 ... množství investic přidělené závodům 1 a 2
x_3 ... množství investic přidělené závodům 1, 2 a 3

ad c) označme

k_j ... ne konstanta, ale proměnná podle počtu přípustných alternativ ve fázi j

k_j^* ... optimální hodnota alternativy ve fázi j

$R_j(k_j)$ ... výnos alternativy k_j ve fázi j

$f_j(x_j)$ ... optimální výnos fází $1, 2, \ldots, j$ za daného stavu proměnné x_j

$c_j(k_j)$ ... cena alternativy k_j ve fázi j

Za daného označení nyní hledáme $f_3(x_3 = 5)$; klíčem je sestavení rekurzivních rovnic

$$f_1(x_1) = \max_{c_1(k_1) \leq x_1} (R_1(k_1))$$

$$f_j(x_j) = \max_{c_j(k_j) \leq x_j} \Big(R_j(k_j) + \underbrace{f_{j-1}\overbrace{(x_j - c_j(k_j))}^{x_{j-1}}}_{\text{maximální výnos celkem za fáze } 1,2,\ldots,j-1}\Big)$$

Nyní už přistoupíme k přímému průchodu jednotlivých fází:

fáze 1:

$$f_1(x_1) = \max_{c_1(k_1) \le x_1} (R_1(k_1)) \qquad \text{pro } k_1 = 1,2,3$$

(v závodu 1 existují tři alternativy)

x_1	$R_1(k_1)$			optim. řešení fáze 1	
	$k_1 = 1$	$k_1 = 2$	$k_1 = 3$	$f_1(x_1)$	k_1^*
0	$\boxed{0}$	–	–	0	1
1	0	$\boxed{5}$	–	5	2
2	0	5	$\boxed{6}$	6	3
3	0	5	$\boxed{6}$	6	3
4	0	5	$\boxed{6}$	6	3
5	0	5	$\boxed{6}$	6	3

Sloupec $k_1 = 1$ odpovídá té variantě (č.1), že závod 1 nedostane peníze; sloupec $k_1 = 2$ odpovídá variantě, že závod 1 dostane peníze na návrh 2, sloupec $k_1 = 3$ odpovídá variantě, že závod 1 dostane peníze na návrh 3 (žádný ze závodů nemůže získat peníze na více návrhů současně, tj. více variant ve fázi 1 neexistuje); tyto sloupce udávají výnosy jednotlivých variant pro možné hodnoty proměnné x_1.

fáze 2:

$$f_2(x_2) = \max_{c_2(k_2) \le x_2} (R_2(k_2) + f_1(x_2 - c_2(k_2)))$$

Tabulka má následující tvar (k_2^* je číslo optimálního návrhu v řádku fáze 2, ovšem už v závislosti na fázi 1):

x_2	$R_2(k_2) + f_1(x_2 - c_2(k_2))$					
	$k_2 = 1$	$k_2 = 2$	$k_2 = 3$	$k_2 = 4$	$f_2(x_2)$	k_2^*
0	$0 + 0 = \boxed{0}$	–	–	–	0	1
1	$0 + 5 = \boxed{5}$	–	–	–	5	1
2	$0 + 6 = 6$	$8 + 0 = \boxed{8}$	–	–	8	2
3	$0 + 6 = 6$	$8 + 5 = \boxed{13}$	$9 + 0 = 9$	–	13	2
4	$0 + 6 = 6$	$8 + 6 = \boxed{14}$	$9 + 5 = \boxed{14}$	$12 + 0 = 12$	14	$2 \vee 3$
5	$0 + 6 = 6$	$8 + 6 = 14$	$9 + 6 = 15$	$12 + 5 = \boxed{17}$	17	4

fáze 3:

$$f_3(x_3) = \max_{\substack{c_3(k_3) \le x_3 \\ k_3 = 1,2}} (R_3(k_3) + f_2(x_3 - c_3(k_3)))$$

V poslední fázi potřebujeme už jen jeden řádek: $x_3 = 5$ (i když i např. řádek $x_3 = 4$ by byl zajímavý tím, že by ukazoval optimum pro 4 milióny celkových investic apod., čili jednalo by se o jistou citlivostní analýzu – tento pojem viz kapitola 8.1).

x_3	$R_3(k_3) + f_2(x_3 - c_3(k_3))$			
	$k_3 = 1$	$k_3 = 2$	$f_3(x_3)$	k_3^*
5	$0 + 17 = \boxed{17}$	$3 + 14 = \boxed{17}$	17	$1 \vee 2$

výsledek: Zpětně projdeme k_j^*, abychom viděli možné kombinace řešení:

x_3	k_3^*	x_2	k_2^*	x_1	k_1^*	(k_1^*, k_2^*, k_3^*)
5	1	$(5 - 0 = 5)$	4	$(5 - 4 = 1)$	2	$(2, 4, 1)$
			2	$(4 - 2 = 2)$	3	$(3, 2, 2)$
	2	$(5 - 1 = 4)$	3	$(4 - 3 = 1)$	2	$(2, 3, 2)$

Z posledního sloupce je vidět, že existují tři optimální rozdělení investic pro jednotlivé návrhy.

2) Zpětný postup: Je to druhá možnost řešení; ovšem kromě opačného postupu od závodu 3 k závodu 1 (jiné rekurentní rovnice) budou jinak definovány stavy:

ad a) fáze jsou stejné

ad b) y_1 ... množství investic přidělené závodům 1, 2 a 3
y_2 ... množství investic přidělené závodům 2 a 3
y_3 ... množství investic přidělené závodu 3

ad c) $f_j(y_j)$ ... optimální výnos fází $j, j+1, \ldots, n$ za daného stavu proměnné y_j.

Pro toto označení můžeme psát rekurzivní rovnice:

$$f_3(y_3) = \max_{\substack{k_3 \\ c_3(k_3) \leq x_3}} (R_3(k_3))$$

$$f_j(y_j) = \max_{\substack{k_j \\ c_j(k_j) \leq x_j}} (R_j(k_j) + f_{j+1}(y_j - c_j(k_j)))$$

fáze 3:

	$R_3(k_3)$		optimální řešení	
y_3	$k_3 = 1$	$k_3 = 2$	$f_3(y_3)$	k_3^*
0	$\boxed{0}$	–	0	1
1	0	$\boxed{3}$	3	2
2	0	$\boxed{3}$	3	2
3	0	$\boxed{3}$	3	2
4	0	$\boxed{3}$	3	2
5	0	$\boxed{3}$	3	2

fáze 2:

	$R_2(k_2) + f_3(y_2 - c_2(k_2))$					
y_2	$k_2 = 1$	$k_2 = 2$	$k_2 = 3$	$k_2 = 4$	$f_2(y_2)$	k_2^*
0	$0 + 0 = \boxed{0}$	–	–	–	0	1
1	$0 + 3 = \boxed{3}$	–	–	–	3	1
2	$0 + 3 = 3$	$8 + 0 = \boxed{8}$	–	–	8	2
3	$0 + 3 = 3$	$8 + 3 = \boxed{11}$	$9 + 0 = 9$	–	11	2
4	$0 + 3 = 3$	$8 + 3 = 11$	$9 + 3 = \boxed{12}$	$12 + 0 = \boxed{12}$	12	$3 \vee 4$
5	$0 + 3 = 3$	$8 + 3 = 11$	$9 + 3 = 12$	$12 + 3 = \boxed{15}$	15	4

fáze 1: Stačí jen řádek: $y_1 = 5$.

	$R_1(k_1) + f_2(y_1 - c_1(k_1))$				
y_1	$k_1 = 1$	$k_1 = 2$	$k_1 = 3$	$f_1(y_1)$	k_1^*
5	$0 + 15 = 15$	$5 + 12 = \boxed{17}$	$6 + 11 = \boxed{17}$	17	$2 \vee 3$

„Rozpletením“ v přímém směru dostaneme stejná 3 řešení.

Použití přímého nebo zpětného postupu algoritmu v praxi závisí na tom, který typ rekurentních rovnic je jednodušší.

Nejtěžším úkolem v algoritmech dynamického programování je definice stavu (tj. správné vyjádření závislosti fáze na předchozích fázích v rekurentní rovnici).

Procvičme na několika příkladech:

Příklad 11.2. Na konci každého roku se rozhoduje, zda určitý stroj necháme v provozu, nebo jej vyměníme za nový; pokud zůstane v provozu, klesne zisk (díky větší poruchovosti), pokud koupíme nový, bude to něco stát. Určete fáze, alternativy v každé fázi a stavy.

Řešení. fáze j .... rok j

alternativy .... a) nechat daný stroj

b) koupit nový stroj

stav j .... stáří stroje na počátku roku j

Příklad 11.3. Vedoucí chce zjistit počet dělníků potřebných na nejbližších 5 týdnů, když zná počty potřebné pro každý týden.
Jsou známy také ceny, které zaplatí za – přijetí

– propuštění

– nevyužitý čas práce dělníka

(na str.296 je příklad uveden pro konkrétní hodnoty).
Kolik dělníků každý týden přijmout a propustit, aby se minimalizovala celková cena s tím spojená? Určete fáze, alternativy v každé fázi a stavy.

Řešení. fáze j .... týden j

alternativy .... počet dělníků – přijatých

– propuštěných v týdnu j

stav týdne j ...... počet dělníků připravených k práci na počátku týdne j (= na konci týdne $(j-1)$). Počty dělníků přijatých nebo propuštěných v předchozích fázích zde nehraje roli.

Příklad 11.4. Problém nákladu (problém plnění batohu). Na člun chceme naložit některé z položek $1, 2, \ldots, N$ tak, aby náklad nepřekročil určitou hmotnost a měl přitom maximální hodnotu. Označení:

w_i ... hmotnost položky i

v_i ... hodnota položky i

k_i ... počet položek i

w ... maximální hmotnost celého nákladu

Uvažujme následující zadání

i	w_i	v_i
1	2	65
2	3	80
3	1	30

$w = 5.$

Řešení. Jedná se vlastně o úlohu celočíselného lineárního programování:

$$\text{maximalizujte funkci } z = v_1k_1 + \cdots + v_Nk_N \text{ za podmínky}$$
$$w_1k_1 + \cdots + w_Nk_N \leq w,\ k_i \text{ jsou nezáporná celá čísla.}$$

Ale vyřešme tuto úlohu pomocí algoritmu dynamického programování:

fáze j	položka j
stav y_j ve fázi j	celková hmotnost přiřazená fázím $j, j+1, \dots N$
	$(y_1 = w,\ y_i = 0, 1, \dots, w$ pro $i > 1)$
alternativa k_j ve fázi j	počet jednotek j-té položky, kterou bereme s sebou.

Tato úloha je podobná příkladu 11.1 – jedná se opět v jistém smyslu o přidělení zdrojů. Zpětné rekurentní rovnice jsou tvaru:

$$f_N(y_N) = \max_{\substack{k_N \\ y_N}} (v_N\, k_N)$$

$$f_j(y_j) = \max_{\substack{k_j \\ y_j}} (v_j\, k_j + f_{j+1}(y_j - w_j\, k_j)) \quad \text{pro } j = 1, 2, \dots, N-1$$

Zpětným postupem (fáze 3, fáze 2, fáze 1) si bystrý čtenář ověří, že $(2, 0, 1)$ je optimální rozdělení položek.

Příklad 11.5. Problém spolehlivosti. Elektronické zařízení se skládá ze tří hlavních komponent zapojených sériově, tj. selhání komponenty vede k selhání celého zařízení. Paralelním přidáním rezervních komponent lze zvýšit spolehlivost zařízení. Každá komponenta může být až dvakrát zálohovaná (tj. maximálně lze paralelně zapojit 3 kusy komponenty téhož typu). Celková cena zařízení má být maximálně 10 000 dolarů. Tabulka udává, jak roste spolehlivost $R_j(k_j)$ a cena $c_j(k_j)$ pro různý počet komponent téhož typu:

	$j=1$		$j=2$		$j=3$	
k_j	R_1	c_1	R_2	c_2	R_3	c_3
1	0,6	1	0,7	3	0,5	2
2	0,8	2	0,8	5	0,7	4
3	0,9	3	0,9	6	0,9	5

(ceny jsou v tisících dolarů). Jak zapojit náhradní komponenty, aby spolehlivost při dané ceně byla maximální?

Řešení. celková spolehlivost = součin jednotlivých spolehlivostí

úkol: nalezněte maximum funkce $R = \prod_{j=1}^{N} R_j(k_j)$ za podmínky

$$\sum_{j=1}^{N} c_j(k_j) \leq 10 \text{ (tisíc dolarů).}$$

Použijeme zpětný postup dynamického programování:

fáze j .... typ komponenty j

stav y_j ve fázi j .... celková cena přiřazená typům $j, j+1, \dots N$

alternativy k_j ve fázi j .... počet paralelně zapojených jednotek typu j ($k_j \in \{1,2,3\}$)

$f_j(y_j)$ .... celková optimální spolehlivost typů $j, j+1, \dots N$ pro daný kapitál y_j

Rekurentní rovnice:

$$\boxed{\begin{aligned} f_N(y_N) &= \max_{\substack{k_N \\ c_N(k_N) \le y_N}} \left(R_N(k_N)\right) \\ f_j(y_j) &= \max_{\substack{k_j \\ c_j(k_j) \le y_j}} \left(R_j(k_j)\, f_{j+1}(y_j - c_j(k_j))\right) \quad \text{pro } j = 1,2,\dots,N-1 \end{aligned}}$$

$y_j \in \{0,1,2,\dots,10\}$; ovšem počet řádků v jednotlivých fázích můžeme snížit následující úvahou:

fáze 3....... $y_3 \ge 2$ a $y_3 \le 10 -$ (cena 1 komponenty typu 1 + cena 1 komponenty typu 2),
tj. $2 \le y_3 \le 6$

fáze 2....... $y_2 \ge$ cena 1 komponenty typu 2 + cena 1 komponenty typu 3
$\le 10 -$ cena 1 komponenty typu 1
celkem $5 \le y_2 \le 9$

fáze 1....... $y_1 \ge$ cena komponenta 1 + komponenta 2 + komponenta 3
≤ 10
celkem $6 \le y_1 \le 10$

Můžeme se tedy pustit do rekurentního algoritmu (zpětný postup):

fáze 3:

	$R_3(k_3)$			optimální řešení	
y_3	$k_3 = 1$	$k_3 = 2$	$k_3 = 3$	$f_3(y_3)$	k_3^*
2	$\boxed{0,5}$	–	–	$0,5$	1
3	$\boxed{0,5}$	–	–	$0,5$	1
4	$0,5$	$\boxed{0,7}$	–	$0,7$	2
5	$0,5$	$0,7$	$\boxed{0,9}$	$0,9$	3
6	$0,5$	$0,7$	$\boxed{0,9}$	$0,9$	3

fáze 2:

y_2	$R_2(k_2)\cdot f_3(y_2-c_2(k_2))$				
	$k_2=1$	$k_2=2$	$k_2=3$	$f_2(y_2)$	k_2^*
5	$0,7\cdot 0,5=\boxed{0,35}$	–	–	$0,35$	1
6	$0,7\cdot 0,5=\boxed{0,35}$	–	–	$0,35$	1
7	$0,7\cdot 0,7=\boxed{0,49}$	$0,8\cdot 0,5=0,4$	–	$0,49$	1
8	$0,7\cdot 0,9=\boxed{0,63}$	$0,8\cdot 0,5=0,4$	$0,9\cdot 0,5=0,45$	$0,63$	1
9	$0,7\cdot 0,9=\boxed{0,63}$	$0,8\cdot 0,7=0,56$	$0,9\cdot 0,5=0,45$	$0,63$	1

fáze 1:

y_1	$R_1(k_1)\cdot f_2(y_1-c_1(k_1))$				
	$k_1=1$	$k_1=2$	$k_1=3$	$f_1(y_1)$	k_1^*
10	$0,6\cdot 0,63=0,378$	$0,8\cdot 0,63=\boxed{0{,}504}$	$0,9\cdot 0,49=0,441$	$0,504$	2

Optimální varianta je tedy $(k_1^*,k_2^*,k_3^*)=(2,1,3)$.

Ad Příklad 11.3 (pro konkrétní hodnoty)

týden	*1*	*2*	*3*	*4*	*5*
potřebný počet dělníků b_j	*5*	*7*	*8*	*4*	*6*

Pokud počet dělníků aktuálního týdne přesáhne počet dělníků minulého týdne, musíme započítat cenu spojenou se získáním nových pracovních sil. Na druhé straně, pokud je daný týden více dělníků než potřebujeme, stojí to určitou přebytkovou cenu.

y_j .... *počet dělníků na počátku týdne* j

$c_1(y_j-b_j)$ .... *přebytková cena, když* y_j *převýší* b_j

$c_2(y_j-y_{j-1})$ ...... *cena získání nových pracovních sil*

Data vedoucího ukazují, že

$$c_1(y_j-b_j)=3\,(y_j-b_j),\ \ j=1,2,\dots,5$$

$$c_2(y_j-y_{j-1})=\begin{cases}4+2\,(y_j-y_{j-1}) & \dots\ y_j>y_{j-1}\\ 0 & \dots\ y_j\le y_{j-1}\end{cases}$$

(propouštění nic nestojí)

$$y_0=5,\ \ y_5=6.$$

Jaké jsou optimální hodnoty y_1,y_2,y_3,y_4*?*

Řešení. fáze j .... j-tý týden
stav ve fázi j .... y_{j-1} (= počet dělníků na konci fáze $j-1$)
alternativy ve fázi j .... y_j (= počet dělníků v týdnu j)
$f_j(y_{j-1})$ .... optimální cena řízení pracovních sil za týdny $j, j+1, \ldots, 5$ pro dané y_{j-1}

Rekurentní rovnice zpětného postupu:

$$\boxed{\begin{aligned} f_5(y_4) &= \min_{y_5=b_5} \left(c_1(y_5-b_5)+c_2(y_5-y_4)\right) \\ f_j(y_{j-1}) &= \min_{y_j \geq b_j} \left(c_1(y_j-b_j)+c_2(y_j-y_{j-1})+f_{j+1}(y_j)\right) \quad \text{pro } j=1,2,3,4 \end{aligned}}$$

počet řádků v jednotlivých fázích můžeme snížit touto úvahou:

$$4 = b_4 < b_5 = 6 \Rightarrow y_4 \in \{4,5,6\}.$$

Dále $y_3 = 8$ (protože propouštění nic nestojí).
Dále $y_2 \in \{7,8\}$.
Dále $y_1 \in \{5,6,7,8\}$ (musíme uvažovat i $y_1 = 6,7,8$, protože nevíme, zda není levnější nabrat v 1.týdnu 8 lidí a nechat si je až do 3.týdne).

A nyní už se můžeme pustit do zpětného postupu:

fáze 5: $b_5 = 6$

	$c_1(y_5-6)+c_2(y_5-y_4)$	optimální řešení	
y_4	$y_5 = 6$	$f_5(y_4)$	y_5^*
4	$3\cdot 0+4+2\cdot 2=\boxed{8}$	8	6
5	$3\cdot 0+4+2\cdot 1=\boxed{6}$	6	6
6	$3\cdot 0+0=\boxed{0}$	0	6

fáze 4: $b_4 = 4$

	$c_1(y_4-4)+c_2(y_4-y_3)+f_5(y_4)$				
y_3	$y_4 = 4$	$y_4 = 5$	$y_4 = 6$	$f_4(y_3)$	y_4^*
8	$0+0+8=8$	$3\cdot 1+0+6=9$	$3\cdot 2+0+0=\boxed{6}$	6	6

fáze 3: $b_3 = 8$

y_2	$c_1(y_3-8)+c_2(y_3-y_2)+f_4(y_3)$		
	$y_3 = 8$	$f_3(y_2)$	y_3^*
7	$0+4+2\cdot 1+6=\boxed{12}$	12	8
8	$0+0+6=\boxed{6}$	6	8

fáze 2: $b_2 = 7$

y_1	$c_1(y_2-7)+c_2(y_2-y_1)+f_3(y_2)$			
	$y_2 = 7$	$y_2 = 8$	$f_2(y_1)$	y_2^*
5	$0+4+2\cdot 2+12=20$	$3\cdot 1+4+2\cdot 3+6=\boxed{19}$	19	8
6	$0+4+2\cdot 1+12=18$	$3\cdot 1+4+2\cdot 2+6=\boxed{17}$	17	8
7	$0+0+12=\boxed{12}$	$3\cdot 1+4+2\cdot 1+6=15$	12	7
8	$0+0+12=12$	$3\cdot 1+0+6=\boxed{9}$	9	8

fáze 1: $b_1 = 5$

y_0	$c_1(y_1-5)+c_2(y_1-y_0)+f_2(y_1)$					
	$y_1 = 5$	$y_1 = 6$	$y_1 = 7$	$y_1 = 8$	$f_1(y_0)$	y_1^*
5	$0+0+19=\boxed{19}$	$3\cdot 1+\ 4+$ $2\cdot 1+17=26$	$3\cdot 2+\ 4+$ $2\cdot 2+12=26$	$3\cdot 3+4+$ $2\cdot 3+9=28$	19	5

optimum: $y_0 = 5 \Rightarrow$

$$\begin{aligned} y_1 &= 5\\ y_2 &= 8\\ y_3 &= 8\\ y_4 &= 6\\ y_5 &= 6 \end{aligned}$$

Tedy ideální strategií je
1.týden … 0 (máme 5 dělníků z minulého týdne)
2.týden … 3 najmeme
3.týden … 0
4.týden … 2 propustíme
5.týden … 0.

Poslední úloha se lišila od těch předchozích v tom, že v každém řádku tabulky jsme hledali minimální hodnotu (jednalo se o minimalizační úlohu).

Všechny probírané úlohy byly jednodimenzionální; u větší dimenze neúměrně narůstá počet přípustných kombinací proměnných.

Úlohy lineárního programování lze rovněž řešit pomocí algoritmu dynamického programování (viz př.11.4), ale počet omezení určuje počet proměnných v úloze, tj. dynamické programování se zde nepoužívá pro více než 1 omezení (nebo než 2 v případě dvou proměnných).

Pojmy k zapamatování

- V této kapitole jsme ukázali další možnost hledání extrému funkce za daných omezení.
- Hlavní myšlenka dynamického programování zní: při hledání optimální možnosti je někdy nejrychlejší rozdělit daný problém na několik podproblémů (fází) a najít optimum v dané fázi se zřetelem na kombinace možností z předchozí fáze – algoritmus je tedy rekurzivní.
- Nejtěžší úkol u úloh dynamického programování je vhodně formulovat zadaný problém. Je třeba správně definovat a) fáze j

  b) stavy x_j ve fázi j

  c) přípustné alternativy k_j ve fázi j.
- Připomeňme označení, které se používá:

  k_j^* ... optimální hodnota alternativy ve fázi j

  $R_j(k_j)$ ... výnos alternativy k_j ve fázi j

  $f_j(x_j)$ ... optimální výnos fází $1, 2, \ldots, j$ za daného stavu proměnné x_j

  $c_j(k_j)$ ... cena alternativy k_j ve fázi j
- Existují dva možné typy algoritmu: přímý postup a zpětný postup

1) Přímý postup: vede k sestavení a řešení rekurzivních rovnic

$$f_1(x_1) = \max_{c_1(k_1)\le x_1} (R_1(k_1))$$

$$f_j(x_j) = \max_{c_j(k_j)\le x_j} (R_j(k_j) + f_{j-1}(x_j - c_j(k_j))),$$

jejichž výsledky se v jednotlivých fázích zapisují do tabulky.

2) Zpětný postup: Je to druhá možnost řešení; ovšem kromě opačného postupu (jiné rekurentní rovnice) budou jinak definovány stavy a optimální výnos $f_j(y_j)$:

- $f_j(y_j)$ ... optimální výnos fází $j, j+1, \ldots, n$ za daného stavu proměnné y_j.

Pro toto označení sestavíme a řešíme rekurzivní rovnice:

$$f_n(y_n) = \max_{\substack{k_n \\ c_n(k_n) \le x_n}} (R_n(k_n))$$

$$f_j(y_j) = \max_{\substack{k_j \\ c_j(k_j) \le x_j}} (R_j(k_j) + f_{j+1}(y_j - c_j(k_j)))$$

- Použití přímého nebo zpětného postupu algoritmu v praxi závisí na tom, který typ rekurentních rovnic je jednodušší.

– Při řešení úloh dynamického programování nemusí být funkce, jejíž optimum hledáme, nutně lineární (viz př.11.5), což je velmi výhodné. Naopak, úlohy celočíselného lineárního programování lze rovněž řešit pomocí algoritmu dynamického programování (viz př.11.4), ale počet omezení určuje počet proměnných v úloze, tj. dynamické programování se zde nepoužívá pro více než 1 omezení.

Kontrolní otázky

1. U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.
   a) Dynamické programování používá rekurzivní algoritmus.
   b) Algoritmem dynamického programování lze řešit pouze úlohy na hledání maxima nějaké funkce.
   c) Při zmenšení pravé strany omezení musíme přepočítat všechny fáze algoritmu dynamického programování znovu.
   d) Přímý i zpětný postup algoritmu dynamického programování vedou vždy ke stejnému výsledku.
   e) Každou úlohu dynamického programování lze převést na úlohu lineárního programování.
   f) Každou úlohu lineárního programování lze řešit algoritmem dynamického programování.

Odpovědi na otázky
1a) – A, 1b) – N, 1c) – N, 1d) – A, 1e) – N, 1f) – A.

Cvičení

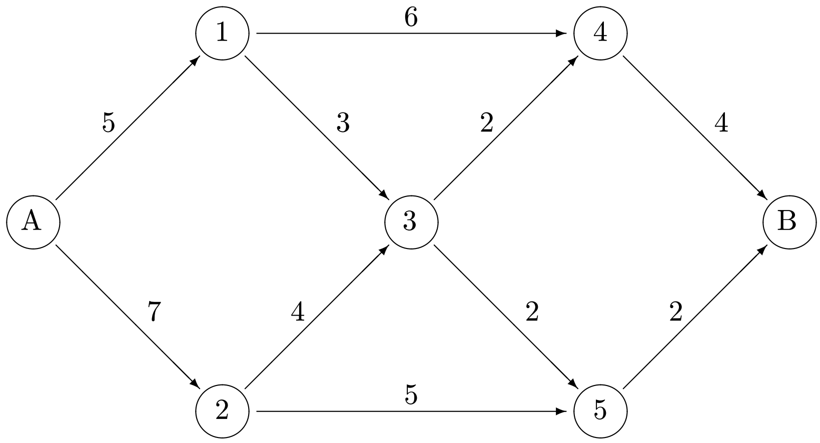
1. Společnost chce investovat 8 miliónů dolarů do rozvoje svých tří závodů. Každý ze závodů předložil několik návrhů, jak by peníze mohl využít a s jakým výnosem. Jak rozdělit peníze, aby byl celkový výnos maximální?

	závod 1		závod 2		závod 3	
návrh	c_1	R_1	c_2	R_2	c_3	R_3
1	3	5	3	4	0	0
2	4	6	4	5	2	3
3	–	–	5	8	3	5
4	–	–	–	–	6	9

   Jednotkou je jeden milión dolarů. Použijte přímý postup algoritmu dynamického programování k získání optimálního řešení.

2. (hledání nejkratší cesty)
   Graf na obrázku znázorňuje možné cesty vedoucí z města A do města B procházející městy

označenými čísly 1 až 5. Hodnoty nad šipkami udávají délky jednotlivých cest. Naším úkolem je najít nejkratší cestu z A do B.



Formulujte tento problém jako úlohu dynamického programování. Definujte fáze, stavy a přípustné alternativy, pak najděte optimální řešení.

3. Formulujte následující úlohu celočíselného lineárního programování jako úlohu dynamického programování

najděte maximum funkce $z = 2\,x_1 + 3\,x_2 + 4\,x_3$ za omezení $\quad 2\,x_1 + 2\,x_2 + 3\,x_3 \leq 4$

$$x_i \in \mathbb{Z},\ x_i \geq 0, \text{ pro } i = 1, 2, 3.$$

Určete optimální řešení přímým postupem algoritmu dynamického programování.

4. Studentka si musí vybrat 10 volitelných předmětů, které budou přednášeny na 4 různých katedrách, přitom na každé katedře musí absolvovat minimálně 1 předmět. Při rozhodování, kolik předmětů si na každé katedře zvolí, se řídí požadavkem, aby maximalizovala svoje „poznání“ ve všech čtyřech oblastech. Dobře ví, že pokud její volba pro danou katedru přesáhne určitý počet předmětů, její vědomosti z daného oboru se již nebudou dále zvyšovat. To proto, že buď bude množství probírané látky příliš velké pro její pochopení, nebo se bude daná tematika v předmětech neustále opakovat. Vytvořila si tedy pro každou katedru 100-bodovou stupnici poznání (100 = maximální poznání) v závislosti na počtu studovaných předmětů, která je zapsána v následující tabulce

	počet předmětů									
katedra	1	2	3	4	5	6	7	8	9	10
I	25	50	60	80	100	100	100	100	100	100
II	20	70	90	100	100	100	100	100	100	100
III	40	60	80	100	100	100	100	100	100	100
IV	10	20	30	40	50	60	70	80	90	100

Formulujte tento problém jako úlohu dynamického programování a určete optimální řešení.

Výsledky

ad 1. Optimální rozdělení investic pro jednotlivé návrhy je $(1, 3, 1)$ se ziskem 13 miliónů dolarů.

ad 2. Optimální cesta je $A \to 1 \to 3 \to 5 \to B$, celková minimální vzdálenost $= 12$.

ad 3. $\boldsymbol{x} = (0, 2, 0)$, $z(\boldsymbol{x}) = 6$.

ad 4. Optimální počty předmětů na jednotlivých katedrách jsou $(2, 3, 4, 1)$, celkový počet bodů je 250.

12 Modely skladových zásob

Průvodce studiem

V této kapitole se budeme věnovat matematickému zpracováním odpovědi na následující otázku z teorie zásob: Kdy a kolik zboží je optimální objednat?

Rozhodnutí o skladové politice je vždy prováděno s ohledem na celkovou cenu, kterou chceme minimalizovat. Tato cena se skládá z ceny zboží, plus ceny objednávky, plus ceny dopravy, plus ceny skladování, plus ceny penalizační (za škody vzniklé při nedostatku zásob).

Podle typu poptávky zboží rozdělujeme matematické modely na deterministické *(je známa velikost poptávky zboží) a* pravděpodobnostní *(poptávka přesně není známa, pouze její pravděpodobná hodnota).*

Dále můžeme modely dělit podle toho, jestli se velikost poptávky s rostoucím časem mění nebo ne, a to na statické *a* dynamické *(v případě deterministických modelů) a na* stacionární *a* nestacionární *(v případě pstních modelů).*

Ukážeme si možnosti řešení jednotlivých typů úloh.

Cíle

Po prostudování této kapitoly budete schopni:

- Odpovědět na otázku: Kdy a kolik zboží je optimální objednat?
- Stanovit celkovou cenu zboží.
- Rozlišit od sebe jednotlivé modely na deterministické a pravděpodobnostní.
- Rozlišit od sebe modely statické, dynamické, stacionární a nestacionární.
- Najít vhodné řešení pro jednotlivé typy úloh.

12.1 Úvod

V této kapitole se budeme zabývat matematickým zpracováním odpovědi na následující otázku z teorie zásob: Kdy a kolik zboží je optimální objednat?

Přitom lze uvažovat různé situace skladového provozu: někdy se zásoby ve skladu kontrolují jednou za čas (za měsíc, za rok ...), u jiného typu zboží se kontrolují nepřetržitě (a k nové objednávce dojde, pokud zásoby klesnou na určitou hodnotu).

Rozhodnutí o skladové politice je prováděno s ohledem na celkovou cenu (tuto celkovou cenu chceme minimalizovat):

celková cena = cena zboží + cena objednávky (dopravy) + cena skladování + cena penalizační za škody vzniknuvší při nedostatku zásob

Podle typu poptávky zboží rozdělujeme matematické modely na:

a) deterministické – je známa velikost poptávky zboží
   1. statické … velikost poptávky je konstantní, nemění se
   2. dynamické … velikost poptávky je v různých obdobích rovna různým konstantám

b) pravděpodobnostní – poptávka přesně není známa, pouze hustota (nebo pravděpodobnostní funkce), která vyjadřuje jistou pravděpodobnou hodnotu poptávky
   1. stacionární … hustota (nebo pstní funkce) poptávky se nemění v čase
   2. nestacionární … hustota (nebo pstní funkce) s časem mění svůj tvar

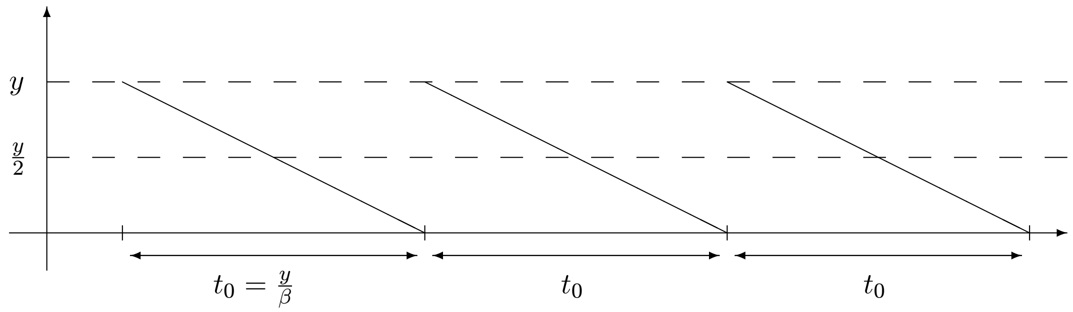
roste matematická složitost modelů

12.2 Deterministické modely

12.2.1 Statický model pro jednu položku

Předpoklady použití modelu:

- poptávka je konstantní … β jednotek zboží za jednotku času
- zásoby jsou jednorázové doplňovány po určité časové periodě t_0
- není uvažován nedostatek zásob (záporné zásoby na skladě znamenají nevyřízenou objednávku zboží)
- velikost objednávky … y jednotek zboží
- dodací lhůta objednávky je známá konstanta … L časových jednotek
- graf množství zboží ve skladu v závislosti na čase lze znázornit:



– cena objednávky (dopravy) … K finančních jednotek

– h … cena skladování jednotky zboží za jednotku času

Řešení modelu: jak zvolit optimální velikost objednávky y, aby celková cena byla minimální?
Zavedeme funkci vyjadřující celkovou cenu za jednotkový čas … $TCU(y)$
(omlouvám se, ale ze slabosti vůči angličtině jsem ponechal anglické značení:
TCU = Total Cost per Unit (time))

$$TCU(y) = \frac{K}{y/\beta} + h \cdot \underbrace{\frac{y}{2}}_{\text{průměrné množství zboží na skladě}}$$

$$\left(\text{čili} \quad \begin{matrix}\text{celková cena}\\ \text{za jednotkový čas}\end{matrix} = \begin{matrix}\text{cena objednávky (dopravy)}\\ \text{za jednotkový čas}\end{matrix} + \begin{matrix}\text{cena skladování}\\ \text{za jednotkový čas}\end{matrix}\right)$$

Nyní řešením rovnice

$$\frac{\partial}{\partial y}(TCU(y)) = 0$$

najdeme minimum

$$y_0 = \sqrt{\frac{2K\beta}{h}}$$

(tzv. Wilsonova ekonomická velikost objednávky).
Často existuje časové prodlení L mezi objednáním a dodáním zboží. Objednávku tedy musíme realizovat vždy L časových jednotek před vyčerpáním zásob.

Příklad 12.1. Denní poptávka zboží je asi 100 jednotek. Cena objednávky je vždy 100 Kč navíc k ceně zboží. Skladovací cena jednotky zboží na 1 den je 0,02 Kč. Dodací lhůta je 12 dní. Určete ekonomickou velikost objednávky a upřesněte, jak dlouho před vyčerpáním zásob se musí objednávka provést.

Řešení. $y_0 = \sqrt{\frac{2\cdot 100\cdot 100}{0{,}02}} = 1\,000$ jednotek, $t_0 = \frac{y_0}{\beta} = \frac{1\,000}{100} = 10$ dní. Každých 10 dní je nutno provést novou objednávku 1 000 jednotek, a sice po ustálení cyklu vždy 2 dny před vyčerpáním zásob.

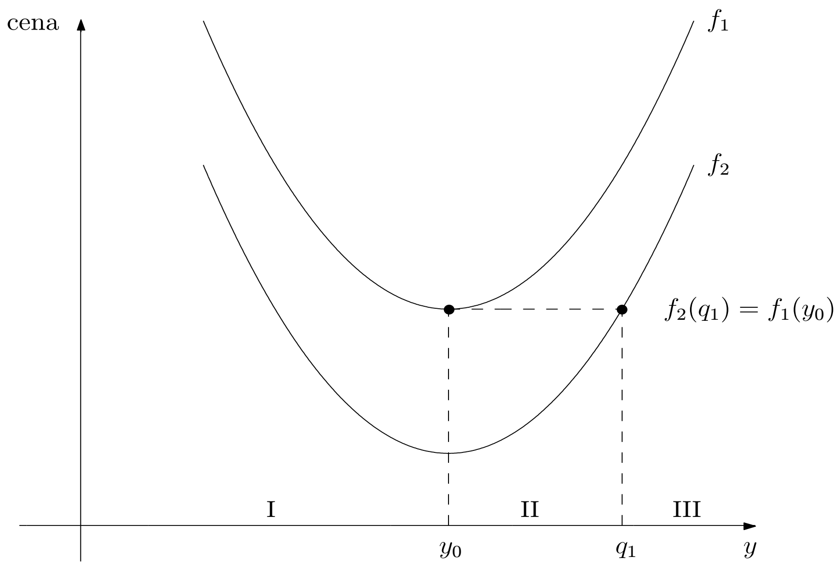
12.2.2 Statický model pro jednu položku s diskontními cenami

Předpoklady použití modelu jsou stejné jako u předchozího modelu až na to, že do celkové ceny se navíc započítává i fakt, že při větším objednaném množství klesá cena jednotky zboží:

$$\begin{array}{c}\text{celková cena}\\ \text{za jednotkový čas}\end{array} = \begin{array}{c}\text{cena zboží}\\ \text{za jednotkový čas}\end{array} + \begin{array}{c}\text{cena objednávky}\\ \text{za jednotkový čas}\end{array} + \begin{array}{c}\text{cena skladování}\\ \text{za jednotkový čas}\end{array}$$

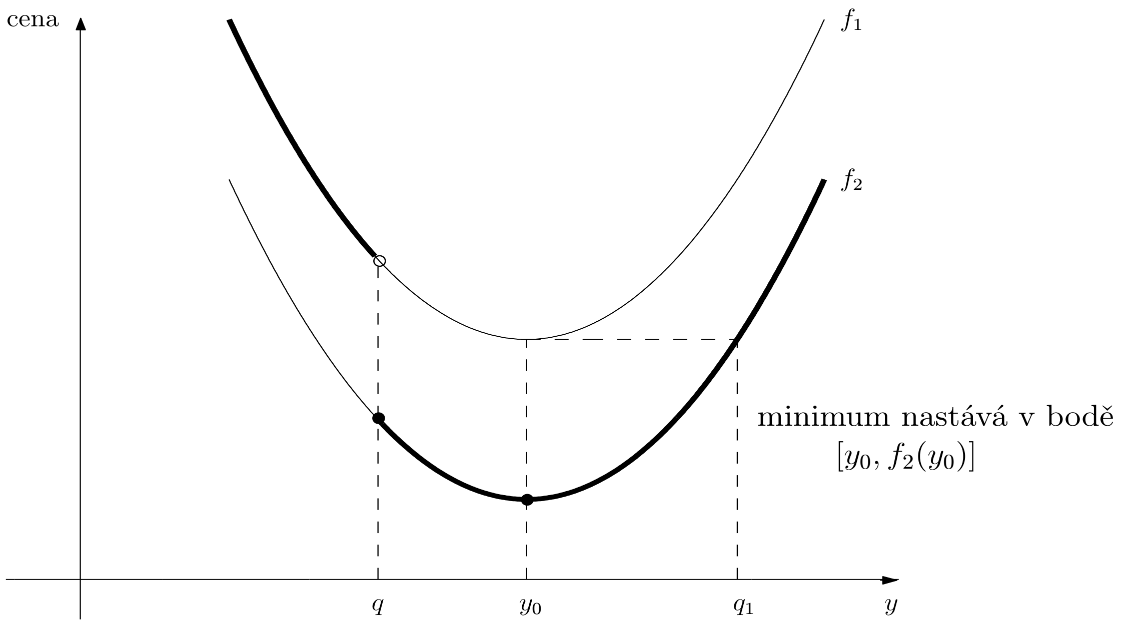
$$TCU(y) = \begin{cases} f_1(y) = \beta \cdot c_1 + \frac{K \cdot \beta}{y} + \frac{h}{2} \cdot y \ \dots \ y < q \\ f_2(y) = \beta \cdot c_2 + \frac{K \cdot \beta}{y} + \frac{h}{2} \cdot y \ \dots \ y \geq q \end{cases} \qquad (c_1 > c_2)$$

Obě z funkcí f_1, f_2 nabývají svého minima ve stejném bodě $y_0 = \sqrt{\frac{2K\beta}{h}}$:

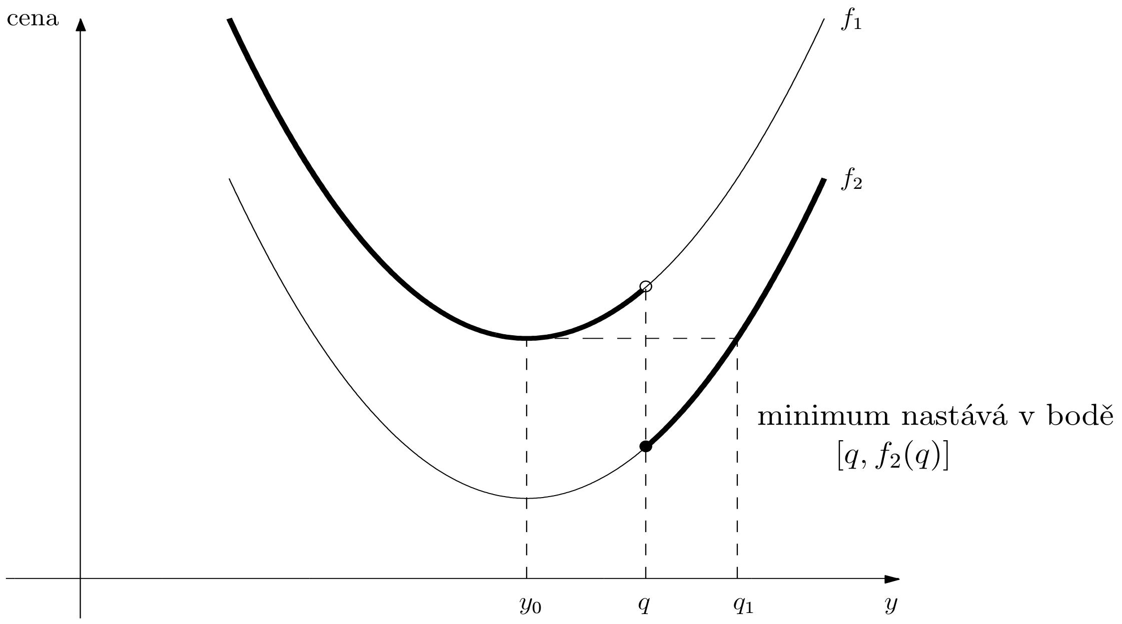


Body y_0, q_1 rozdělí vodorovnou osu na části I, II, III. Řešení modelu závisí na tom, ve které z částí I, II, III se nachází kritické množství q, při jehož dosažení dochází k diskontnímu snížení ceny jednotky zboží z c_1 na c_2.

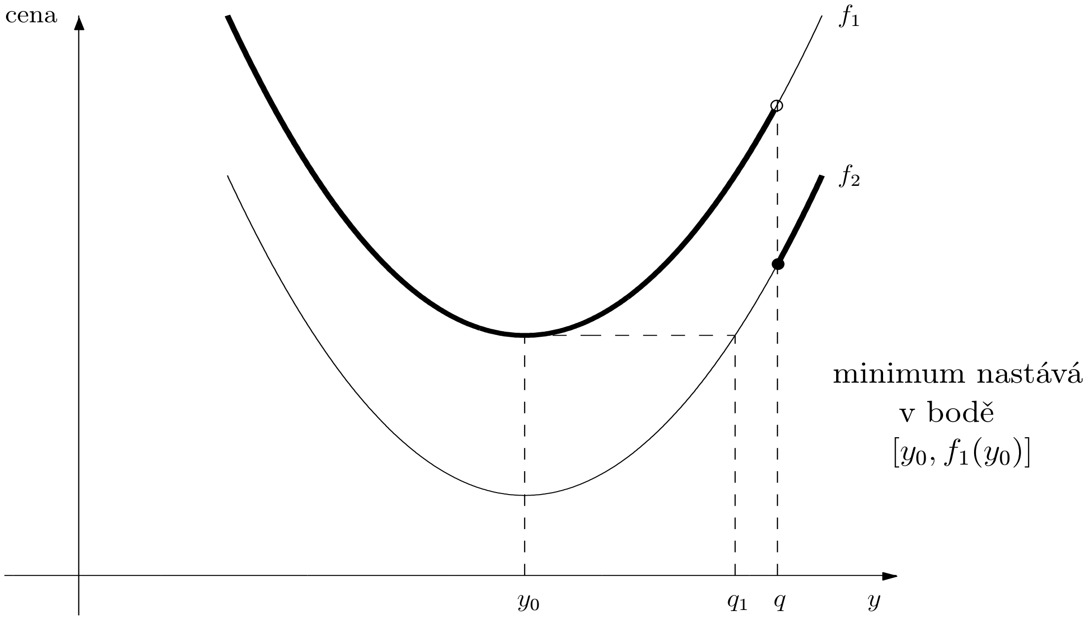
a) $q \in$ I: (silně je zvýrazněn graf funkce TCU)



b) $q \in$ II:



c) $q \in$ III:



Postup při rozhodování: určíme $y_0 = \sqrt{\frac{2K\beta}{h}}$; pokud $y_0 > q$, jedná se o případ a); jinak určíme q_1 tak, aby $f_2(q_1) = f_1(y_0)$; pak pro $q < q_1$ nastane případ b), jinak c).

Příklad 12.2. Rozhodněte o optimální velikosti objednávky při následujících údajích: $K = 10$ Kč, $h = 1$ Kč, $\beta = 5$ jednotek zboží za den, $c_1 = 2$ Kč, $c_2 = 1$ Kč, $q = 15$ jednotek.

Řešení. $y_0 = 10$ jednotek zboží, $y_0 < q \Rightarrow$ řešme rovnici $f_1(10) = f_2(q_1)$:
úpravou dostaneme

$$q_1^2 - 30\,q_1 + 100 = 0 \Rightarrow q_{1_{1,2}} = \begin{cases} 3,82 \\ 26,18 \end{cases}$$

Protože objednáváme na více než 1 den a $\beta = 5$, reálně připadá v úvahu jen $q_1 = 26,18$; $q < q_1 \Rightarrow q \in \mathrm{II}$, tj. optimální je objednat 15 jednotek zboží, celková cena za jednotkový čas je $f_2(q) = f_2(15) = 15,83$ Kč za 1 den.

12.2.3 Statický model pro více druhů zboží s omezením skladového prostoru

Předpoklady použití modelu:

– maximální plocha skladování $\dots A$

– požadavek skladovací plochy na jednotku i-tého druhu zboží $\ldots a_i$

– velikost objednávky zboží druhu $i \ldots y_i$
musí platit

$$(*) \qquad \sum_{i=1}^{n} a_i\, y_i \le A$$

– zboží každého druhu je jednorázově doplňováno s periodou $t_i = \frac{y_i}{\beta_i}$, kde β_i, K_i, h_i mají stejný význam jako v modelu 12.2.1, ale pro zboží druhu i není uvažován nedostatek zásob ani diskontní ceny (tj. model jako 12.2.1, ale ve vyšší dimenzi).

Řešení modelu: Hledáme minimum ceny za jednotku času

$$TCU(y_1, \ldots, y_n) = \sum_{i=1}^{n} \left(\frac{K_i\, \beta_i}{y_i} + \frac{h_i\, y_i}{2} \right)$$

za omezující podmínky

$$\sum_{i=1}^{n} a_i\, y_i \le A; \;\; y_i \ge 0.$$

Pokud $y_i^0 = \sqrt{\frac{2\, K_i\, \beta_i}{h_i}}$ splňují omezení $(*)$, máme optimum.
V opačném případě postupujeme tzv. Lagrangeovou metodou: hledáme minimum funkce

$$L(\lambda, y_1, \ldots, y_n) = \sum_{i=1}^{n} \left(\frac{K_i\, \beta_i}{y_i} + \frac{h_i\, y_i}{2} \right) + \lambda \left(\sum_{i=1}^{n} a_i\, y_i - A \right) \;\; \text{pro } \lambda > 0.$$

Systém $\begin{bmatrix} \frac{\partial L}{\partial y_i} = 0 \\ \frac{\partial L}{\partial \lambda} = 0 \end{bmatrix}$ lze převést na $\begin{bmatrix} y_i^0 = \sqrt{\frac{2\, K_i\, \beta_i}{h_i + 2\, \lambda\, a_i}} \\ \sum\limits_{i=1}^{n} a_i\, y_i^0 - A = 0 \end{bmatrix}$.
Poslední uvedený systém lze řešit numericky: zvyšujeme λ s určitým krokem tak dlouho, až výraz $\sum\limits_{i=1}^{n} a_i\, y_i - A$ je záporný; z posledních dvou hodnot λ pak interpolací najdeme optimální λ, tj. i optimální y_i^0.

Příklad 12.3. Určete optimální objednávané množství při následujících datech ($A = 25\, m^2$):

druh zboží i	K_i (Kč)	β_i $\left(\frac{\text{jednotek}}{\text{za jednotku času}}\right)$	h_i (Kč)	a_i (m^2)
1	10	2	0,3	1
2	5	4	0,1	1
3	10	4	0,2	1

Řešení. Vypočteme $\sqrt{\frac{2\,K_i\,\beta_i}{h_i+2\,\lambda\,a_i}}$ pro různé hodnoty λ:

λ	y_1	y_2	y_3	$\sum\limits_{i=1}^{n} a_i\,y_i - A$
0	$11,5$	$20,0$	$24,5$	$31,0$
$0,05$	$10,0$	$14,1$	$17,3$	$16,4$
$0,1$	$9,0$	$11,5$	$14,9$	$10,4$
$0,15$	$8,2$	$10,0$	$13,4$	$6,6$
$0,2$	$7,6$	$8,9$	$12,2$	$3,7$
$0,25$	$7,1$	$8,2$	$11,3$	$1,6$
$0,3$	$6,7$	$7,6$	$10,6$	$-0,1$

Optimální λ je téměř $\doteq 0,3$, tj. $y_1^0 \doteq 6,7;\ y_2^0 \doteq 7,6;\ y_3^0 \doteq 10,6$

$t_1 = 3,35;\ t_2 = 1,9;\ t_3 = 2,65$ č.j.

(Pro $A \geq 56$ nastane optimum už pro $\lambda = 0$ ... už je dost prostoru pro hodnoty v 1.řádku.)

12.2.4 Dynamický model pro jednu položku a N období

Předpoklady použití modelu:

- poptávka v období $i \ldots D_i$ $\qquad (D_i \in \mathbb{R})$
- množství zboží objednávané na začátku období $i \ldots z_i$
- množství zboží, které je na skladě na začátku období i před přijetím objednávaného z_i $\ldots x_i$
- jednotková cena skladování, jestliže zboží skladované v období i zůstane na skladě i na začátku období $i+1 \ldots h_i$
- cena objednávky v období $i \ldots K_i$
- výrobní cena zboží při objednávaném množství $z_i \ldots c_i(z_i)$
  (může se měnit v různých obdobích, lze uvažovat i diskontní ceny)
- cena nákupu zboží v období $i \ldots C_i(z_i) = \begin{cases} 0 & \ldots\ z_i = 0 \\ K_i + c_i(z_i) & \ldots\ z_i > 0 \end{cases}$

Řešení modelu: Úkolem je určit optimální z_i, pro které je minimální součet

$$\text{cena nákupu} + \text{cena skladování}$$

celkem za všechna období $1, 2, \ldots, N$.

Použijeme přímý algoritmus dynamického programování:
označme $f_i(x_{i+1}) \ldots$ minimální celková cena (nákupu i skladování) za období $1, 2, \ldots, i$, je-li dáno x_{i+1}
rekurzivní rovnice:

$$\boxed{\begin{aligned} f_1(x_2) &= \min_{0 \le z_1 \le D_1 + x_2} (C_1(z_1) + h_1\, x_2) \\ f_i(x_{i+1}) &= \min_{0 \le z_i \le D_i + x_{i+1}} (C_i(z_i) + h_i\, x_{i+1} + f_{i-1}(x_i)) \quad \text{pro } i = 2, \ldots, N \end{aligned}}$$

kde $h_i\, x_{i+1}$ …cena skladování za období i (obecněji by mohla být místo konstanty h_i i funkce $H_i(x_{i+1})$ nebo místo x_{i+1} i třeba $\frac{1}{2}(x_i + x_{i+1})$, apod.)
$x_{i+1} = x_i + z_i - D_i$, tj. $f_{i-1}(x_{i+1} + D_i - z_i)$.

Dále v jakékoliv fázi musí platit $0 \le x_{i+1} \le D_{i+1} + \cdots + D_N$
(na skladě nemůžeme nechat víc, než je poptávka za všechna následující období).

Příklad 12.4. Uvažujme hospodaření jedné položky v průběhu tří období ($N = 3$):

období i	D_i (v jednotkách)	K_i (v Kč)	h_i (v Kč)
1	3	3	1
2	2	7	3
3	4	6	2

Dále $x_1 = 1$ (na začátku prvního období máme na skladě jednu jednotku zboží), nákupní cena je 10 Kč za jednotku u prvních tří jednotek a 20 Kč za každou další jednotku, tj.

$$C_i(z_i) = \begin{cases} 10\, z_i & \ldots\ 0 \le z_i \le 3 \\ 30 + 20\,(z_i - 3) & \ldots\ z_i \ge 4. \end{cases}$$

Stanovte optimální plán objednávek.

Řešení. Použijeme právě popsaný algoritmus rekurentní rovnice:
fáze 1: $D_1 = 3$

$0 \le x_2 \le D_2 + D_3 = 6$

$2 = D_1 - x_1 \le z_1 \le D_1 + D_2 + D_3 - x_1 = 8$

		$f_1(z_1 \vert x_2) = C_1(z_1) + h_1\,x_2$							optimum	
		$z_1 = 2$	3	4	5	6	7	8		
x_2	$h_1\,x_2$	$C_1(z_1) = 23$	33	53	73	93	113	133	$f_1(x_2)$	z_1^*
0	0	23	–	–	–	–	–	–	23	2
1	1	–	34	–	–	–	–	–	34	3
2	2	–	–	55	–	–	–	–	55	4
3	3	–	–	–	76	–	–	–	76	5
4	4	–	–	–	–	97	–	–	97	6
5	5	–	–	–	–	–	118	–	118	7
6	6	–	–	–	–	–	–	139	139	8

fáze 2: $D_2 = 2$

$0 \leq x_3 \leq D_3 = 4$

$0 \leq z_2 \leq D_2 + D_3 - x_2 \leq D_2 + D_3 = 6$ (optimální x_2 neznáme)

		$f_2(z_2 \vert x_3) = C_2(z_2) + h_2\,x_3 + f_1(x_3 + D_2 - z_2)$							optimum	
		$z_2 = 0$	1	2	3	4	5	6		
x_3	$h_2\,x_3$	$C_2(z_2)$=0	7+10=17	7+20=27	7+30=37	7+50=57	7+70=77	7+90=97	$f_2(x_3)$	z_2^*
0	0	0+0+55 =55	17+0+34 =51	27+0+23 =50	–	–	–	–	50	2
1	3	0+3+76 =79	17+3+55 =75	27+3+34 =64	37+3+23 =63	–	–	–	63	3
2	6	0+6+97 =103	17+6+76 =99	27+6+55 =88	37+6+34 =77	57+6+23 =86	–	–	77	3
3	9	0+9+118 =127	17+9+97 =123	27+9+76 =112	37+9+55 =101	57+9+34 =100	77+9+23 =109	–	100	4
4	12	0+12+139 =151	17+12+118 =147	27+12+97 =136	37+12+76 =125	57+12+55 =124	77+12+34 =123	97+12+23 =132	123	5

fáze 3: $D_3 = 4$

$x_4 = 0$

$z_3 = 0, 1, 2, 3, 4$

		$f_3(z_3 \vert x_4) = C_3(z_3) + h_3\,x_4 + f_2(x_4 + D_3 - z_3)$					optimum	
		$z_3 = 0$	1	2	3	4		
x_4	$h_3\,x_4$	$C_3(z_3)$=0	6+10=16	6+20=26	6+30=36	6+50=56	$f_3(x_4)$	z_3^*
0	0	0+0+123 =123	16+0+100 =116	26+0+77 =103	36+0+63 =99	56+0+50 =106	99	3

Nejnižší celkový součet je 99, tj. $z_3 = 3$

$z_2 = 3$

$z_1 = 2$

Poznámka. Zvláštní případ pro konkávní cenové funkce $C_i(z)$ – tento případ většinou nastane, protože např. ceny u diskontní politiky mají konkávní průběh:
pro $x_1 = 0$ (případ $x_1 > 0$ by se převedl na $x_1 = 0$ odečtením x_1 od D_1) a např. $N = 5$, $D_i = 10, 15, 20, 50, 70$:

1) počet řádků lze podstatně redukovat; např. ve 3.fázi lze místo 121 řádků

$$\left.\begin{array}{l} x_4 = 0 \\ x_4 = 1 \\ \vdots \\ x_4 = 50 + 70 = 120 \end{array}\right\} \text{uvažovat pouze řádky 3: } \left\{\begin{array}{l} x_4 = 0 \\ x_4 = 50 \\ x_4 = 50 + 70. \end{array}\right.$$

Neboli – optimum nastane pro některý z případů $x_i = 0$

$x_i = D_i$

$x_i = D_i + D_{i+1}$

$\vdots$

$x_i = D_i + \cdots + D_N$

2) počet sloupců tabulky lze podstatně redukovat; např. ve fázi 3 můžeme místo 141 sloupců

$$\left.\begin{array}{l} z_3 = 0 \\ z_3 = 1 \\ z_3 = 2 \\ \vdots \\ z_3 = 20 + 50 + 70 = 140 \end{array}\right\} \text{uvažovat pouze 4 sloupce: } \left\{\begin{array}{l} z_3 = 0 \\ z_3 = 20 \\ z_3 = 20 + 50 \\ z_3 = 20 + 50 + 70. \end{array}\right.$$

Neboli – optimum nastane pro některý z případů $z_i = 0$

$z_i = D_i$

$z_i = D_i + D_{i+1}$

$\vdots$

$z_i = D_i + \cdots + D_N$

(v případě z_1 nemusíme dokonce uvažovat ani $z_1 = 0$, protože při $x_1 = 0$ musí být z_1 rovno aspoň hodnotě D_1)

Příklad 12.5. Sestavte plán objednávek pro $N = 4$,

$h_i = 1$ pro všechna i,

$C_i(z_i) = 2\, z_i$ pro všechna i,

$x_1 = 15$ jednotek

a je-li dáno

období i	D_i (v jednotkách)	K_i (v Kč)
1	76	98
2	26	114
3	90	185
4	67	70

Řešení. Odečtením $D_1 - x_1 = 76 - 15 = 61$ převedeme na případ $x_1 = 0$. Využijeme předchozí poznámky, díky níž se podstatně zredukuje počet řádků i sloupců:

fáze 1: $D_1 = 61$

		$f_1(z_1 \mid x_2) = C_1(z_1) + h_1\, x_2$				optimum	
		$z_1 = 61$	61+26=87	61+26+90=177	61+26+90+67=244		
x_2	$h_1\, x_2$	C_1=98+2·61=220	98+2·87=272	98+2·177=452	98+2·244=586	$f_1(x_2)$	z_1^*
0	0	220+0=220	–	–	–	220	61
26	26	–	272+26=298	–	–	298	87
116	116	–	–	452+116=568	–	568	177
183	183	–	–	–	586+183=769	769	244
objednáváme na období...		1	$1, 2$	$1, 2, 3$	$1, 2, 3, 4$		

fáze 2: $D_2 = 26$

		$f_2(z_2 \mid x_3) = C_2(z_2) + h_2\, x_3 + f_1(x_3 + D_2 - z_2)$				optimum	
		$z_2 = 0$	26	26+90=116	26+90+67=183		
x_3	$h_2\, x_3$	$C_2(z_2)$=0	114+2·26=166	114+2·116=346	114+2·183=480	$f_2(x_3)$	z_2^*
0	0	0+0+298=298	166+0+220=386	–	–	298	0
90	90	0+90+568 =658	případ $x_2 = 90$ není optimální	346+90+220 =656	–	656	116
157	157	0+157+769 =926	případ $x_2 = 131$ není optimální	případ $x_2 = 41$ není optimální	480+157+220 =857	857	183
objednáváme na období...		–	2	$2, 3$	$2, 3, 4$		

fáze 3: $D_3 = 90$

		$f_3(z_3 \mid x_4) = C_3(z_3) + h_3\,x_4 + f_2(x_4 + D_3 - z_3)$			optimum	
		$z_3 = 0$	90	$90 + 67 = 157$		
x_4	$h_3\,x_4$	$C_3(z_3)$=0	185+2·90=365	185+2·157=499	$f_3(x_4)$	z_3^*
0	0	0+0+656=656	365+0+298=663	–	656	0
67	67	0+67+857 =924	případ není optimální	499+67+298 =864	864	157
objednáváme na období...		–	3	$3, 4$		

fáze 4: $D_4 = 67$

		$f_4(z_4 \mid x_5) = C_4(z_4) + h_4\,x_5 + f_3(x_5 + D_4 - z_4)$		optimum	
		$z_4 = 0$	$z_4 = 67$		
x_5	$h_4\,x_5$	$C_4(z_4) = 0$	$70 + 2 \cdot 67 = 204$	$f_4(x_5)$	z_4^*
0	0	$0 + 0 + 864 = 864$	$204 + 0 + 656 = 860$	860	67
objednáváme na období...		–	4		

optimální plán: $z_4 = 67$

$z_3 = 0$

$z_2 = 116$

$z_1 = 61$

$$f_4(x_5) = 860 \ \ldots \text{celková cena objednávka + skladování}$$

12.2.5 Dynamický model plánování výroby jedné položky na N období

Cílem modelu je naplánovat výrobu tak, aby se minimalizovala její celková cena.
Označení a předpoklady použití modelu:

c_i ... výrobní cena jednotky zboží v pracovní době období i

d_i ... výrobní cena jednotky zboží v přesčasové době období i ($c_i < d_i$)

h_i ... skladovací cena jednotky zboží, které zůstalo na skladu na konci období i

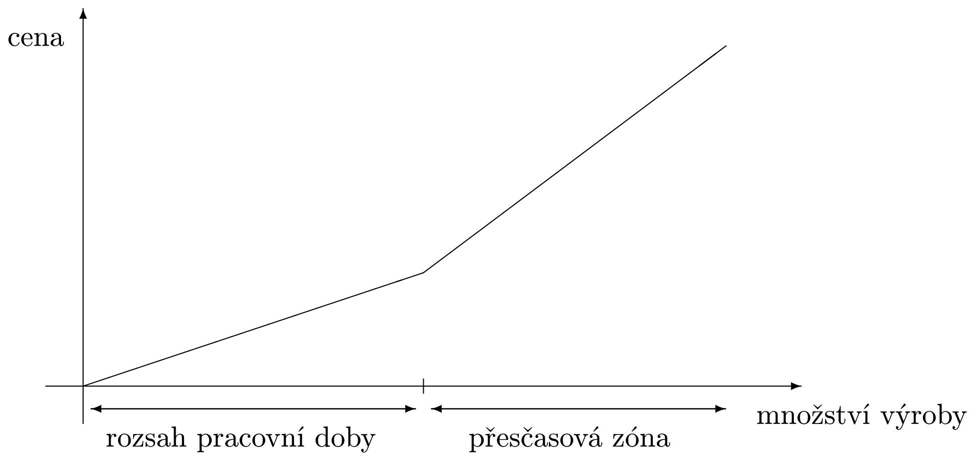
p_i ... penalizační cena jednotky zboží objednaného v období i, ovšem vyrobeného v období $i + 1$

a_{R_i} ... výrobní kapacita (v počtu jednotek) v období i v pracovní době

a_{T_i} ... výrobní kapacita v období i v přesčasové době

b_i ... poptávka (v počtu jednotek) v období i

Graf závislosti ceny na vyrobením množství je konvexní:



Graf může mít i více sklonů $c_i, d_i, e_i, f_i,$ apod. při zachování konvexnosti, což je logický požadavek růstu ceny. Následující algoritmus pak lze rozšířit i na tento případ většího množství cenových skoků.

a) řešení modelu pro $p_i = 0$ (neuvažujeme případ, že výroba v i nepokryje objednávky v i):
Úlohu lze formulovat jako dopravní úlohu, kde

- dodavatelům odpovídají řádné + přesčasové doby v jednotlivých obdobích
- spotřebitelům odpovídají jednotlivá období
- cenám dopravy jednotky zboží od k-tého dodavatele k l-tému spotřebiteli odpovídají součty ceny výrobní a skladovací v daném typu pracovní doby a daném období

Údaje lze pak sestavit do tabulky dopravní úlohy:

	1	2	3	$\dots$	N	navíc	
R_1	c_1	$c_1 + h_1$	$c_1 + h_1 + h_2$	$\dots$	$c_1 + h_1 + \dots + h_{N-1}$	0	a_{R_1}
T_1	d_1	$d_1 + h_1$	$d_1 + h_1 + h_2$	$\dots$	$d_1 + h_1 + \dots + h_{N-1}$	0	a_{T_1}
R_2		c_2	$c_2 + h_2$	$\dots$	$c_2 + h_2 + \dots + h_{N-1}$	0	a_{R_2}
T_2		d_2	$d_2 + h_2$	$\dots$	$d_2 + h_2 + \dots + h_{N-1}$	0	a_{T_2}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
R_N					c_N	0	a_{R_N}
T_N					d_N	0	a_{T_N}
	b_1	b_2	b_3	$\dots$	b_N	s	

$s = \sum a_i - \sum b_j \dots$ případný přidaný sloupec, který vyvažuje kapacitu výroby $\sum a_i$ a kapacitu poptávky $\sum b_j$; většinou je na vyrovnání kapacit potřeba sloupec, protože výroba je zpravidla větší než poptávka.

Díky předpokladu $p_i = 0$ platí

$$(*) \qquad \sum_{i=1}^{k} (a_{R_i} + a_{T_i}) \geq \sum_{j=1}^{k} b_j, \qquad k = 1, 2, \dots, N$$

(výroba v obdobích $1, 2, \dots, k$ musí naplnit poptávku v obdobích $1, 2, \dots, k$),

a proto k optimálnímu rozdělení výroby dospějeme takto:

– v prvním sloupci vybereme pole s nejnižší cenou a vložíme do něj hodnotu co možná největší vzhledem k řádkové i sloupcové kapacitě; pak pokračujeme druhou nejnižší cenou v prvním sloupci, atd.

– ve druhém sloupci vybereme pole s nejnižší cenou a vyplníme maximálně vzhledem k chybějícím kapacitám; atd.

– po projítí všech sloupců je výsledek už optimální plán.

Příklad 12.6. Je dáno:

	kapacity výroby		poptávka
období i	a_{R_i}	a_{T_i}	b_i
1	100	50	120
2	150	80	200
3	100	100	250
4	200	50	200
	550	280	770

$c_i = 2, d_i = 3, h_i = 0,1$ pro všechna i.

Řešení. Kapacita výroby je o 60 jednotek větší než poptávka, úlohu tedy vyvážíme přidáním sloupce s kapacitou 60 jednotek

	1	2	3	4	navíc	
R_1	2	2,1	2,2	2,3	M	100
T_1	3	3,1	3,2	3,3	M	50
R_2		2	2,1	2,2	M	150
T_2		3	3,1	3,2	M	80
R_3			2	2,1	M	100
T_3			3	3,1	M	100
R_4				2	M	200
T_4				3	M	50
	120	200	250	200	60	

V pravém horním rohu polí jsou vepsány jednotkové ceny výroby + skladování. Konstanta M označuje dostatečně velké číslo (větší než jakákoli jiná cena v tabulce).

Algoritmus vyplnění tabulky:

1) V prvním sloupci je minimální cena 2 v poli $(1, 1)$ – zde vložíme maximální možnou hodnotu 100 (tím se vyčerpá kapacita v 1.řádku a všechna ostatní pole v 1. řádku můžeme proškrtnout pomlčkou).

   Do pole $(2, 1)$ vložíme zbývající hodnotu 20, která vyčerpá kapacitu 1.sloupce.

2) Ve 2.sloupci do pole $(3, 3)$ s minimální cenou 2 vložíme hodnotu 150 omezenou kapacitou 3.řádku (tím je tato vyčerpána a ostatní pole ve 3.řádku proškrtneme pomlčkou).

   Další neproškrtnuté pole s minimální cenou ve 2.sloupci je $(4, 2)$, kam vložíme zbylou kapacitu 50 druhého sloupce. Poslední zbylé pole $(2, 2)$ proškrtneme pomlčkou.

3) Atd., výsledné vyplnění je dáno tabulkou

	1	2	3	4	navíc	
R_1	100 (2)	– (2,1)	– (2,2)	– (2,3)	– (M)	100
T_1	20 (3)	– (3,1)	20 (3,2)	– (3,3)	10 (M)	50
R_2		150 (2)	– (2,1)	– (2,2)	– (M)	150
T_2		50 (3)	30 (3,1)	– (3,2)	– (M)	80
R_3			100 (2)	– (2,1)	– (M)	100
T_3			100 (3)	– (3,1)	– (M)	100
R_4				200 (2)	– (M)	200
T_4				– (3)	50 (M)	50
	120	200	250	200	60	

b) řešení modelu pro $p_i \neq 0$: Už nemusí být splněna (∗), tj. po vyplnění tabulky postupem z a) musíme ještě provést optimalizační krok, kterým se eventuelně sníží celková cena výroby + skladování. Optimalizační krok dopravní úlohy byl podrobně popsán v kapitole 10, proto čtenáře odkazuji na příslušný přednáškový text.

Příklad 12.7. Optimalizujte plán na tři období:

i	a_{R_i}	a_{T_i}
1	15	10
2	15	0
3	20	15

Poptávky a jednotkové ceny výroby jsou uvedeny v tabulce:

	1	2	3	navíc	
R_1	5	6	7	M	15
T_1	10	11	12	M	10
R_2	7	5	6	M	15
R_3	9	7	5	M	20
T_3	14	12	10	M	15
	20	35	15	5	

Řešení. Po vyplnění způsobem popsaným v a) dospějeme k tabulce:

	1	2	3	navíc	
R_1	15 (5)	– (6)	– (7)	– (M)	15
T_1	– (10)	5 (11)	– (12)	5 (M)	10
R_2	5 (7)	10 (5)	– (6)	– (M)	15
R_3	– (9)	20 (7)	– (5)	– (M)	20
T_3	– (14)	– (12)	15 (10)	– (M)	15
	20	35	15	5	

Celková cena $= 15 \cdot 5 + 5 \cdot 11 + 5 \cdot 7 + \cdots + 15 \cdot 10 = 505$ $f.j.$ (položku $5 \cdot M$ na pozici $(2,4)$ jsme nebrali v úvahu, protože se jedná o fiktivní období).

Toto rozdělení lze ještě zlepšit optimalizačním krokem (viz kapitola 10), dospějeme k tabulce

	1	2	3	navíc	
R_1	15 (5)	– (6)	– (7)	– (M)	15
T_1	5 (10)	5 (11)	– (12)	– (M)	10
R_2	– (7)	15 (5)	– (6)	– (M)	15
R_3	– (9)	5 (7)	15 (5)	– (M)	20
T_3	– (14)	10 (12)	– (10)	5 (M)	15
	20	35	15	5	

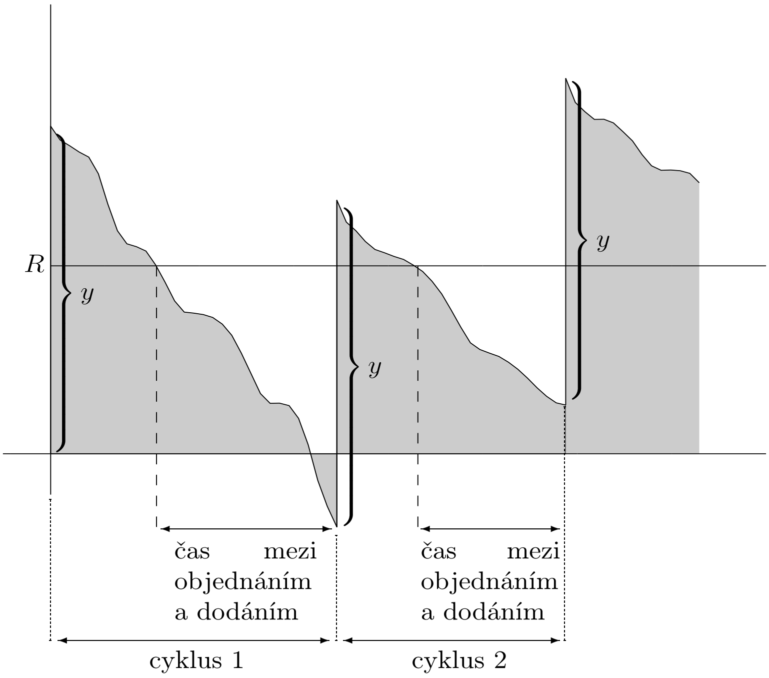
s celkovou cenou 485 $f.j.$ (opět jsme nepočítali fiktivní cenu $5 \cdot M$ v posledním sloupci).

12.3 Pravděpodobnostní modely

12.3.1 Model nepřetržité kontroly pro jednu položku

Označení a předpoklady použití modelu:

– zboží na skladě je neustále kontrolováno a jestliže jeho množství klesne na jistou hodnotu R, přikročí se k objednávce o velikosti y:



– náhodná veličina poptávky mezi objednáním a dodáním $\ldots X$

– podmíněná hustota poptávky X během času mezi objednáním a dodáním (nenulová pro $x > 0$) $\ldots r(x|\,t)$

– hustota doby t mezi objednáním a dodáním $\ldots s(t)$

– hustota poptávky x během doby mezi objednáním a dodáním:

$$f(x) = \int_0^\infty r(x|\,t) \cdot s(t)\, dt$$

– velikost objednávky v jednom cyklu $\ldots y$

– očekávaná celková poptávka za jednotku času (zpravidla 1 rok) $\ldots D$

– jednotková skladovací cena za jednotku času $\ldots h$

– jednotková cena penalizaci při nevyřízené objednávce za jednotku času $\ldots p$

– cena objednávky (doprava) $\ldots K$

Řešení modelu: Určíme optimální R a y tak, aby celková cena skladování za jednotku času byla minimální

– očekávané množství na skladě na konci cyklu $\ldots\ E(R-x) = R - Ex$
na začátku cyklu $\ldots\ y + R - Ex$

- průměrná velikost zásob za jeden cyklus ... $\overline{H} = \frac{(y+R-Ex)+(R-Ex)}{2} = \frac{y}{2} + R - Ex$ (v tomto členu $\overline{H}$ je zanedbána situace, kdy $R - Ex$ je záporná; předpokládá se, že se tím dopustíme malé chyby)
- množství zboží, kterého je v daném cyklu nedostatek (a tudíž prodej nelze vyřídit)

$$S(x) = \begin{cases} 0 & \dots x \le R \\ x - R & \dots x > R \end{cases}$$

- průměrná velikost nevyřízené poptávky ... $\overline{S} = \int\limits_0^\infty S(x) \cdot f(x)\,dx = \int\limits_R^\infty (x-R) f(x)\,dx$
- celková cena skladování za jednotku času =

= průměrná cena objednávky za jednotku času + očekávaná cena skladování + očekávaná penalizace v důsledku později vyřízených objednávek

$$TAC(y,R) = \frac{DK}{y} + h\left(\frac{y}{2} + R - Ex\right) + p \cdot \frac{D\overline{S}}{y}$$

Zde došlo k dalšímu zjednodušení, že totiž penalizační člen uvažujeme nezávislý na době zpoždění dodávky, i když závislý je.

Stacionární body funkce $TAC(y,R)$ najdeme řešením systému

$$\left.\begin{aligned} \frac{\partial TAC(y,R)}{\partial y} &= 0 \\ \frac{\partial TAC(y,R)}{\partial R} &= 0 \end{aligned}\right\} \Rightarrow \quad \begin{aligned} -\frac{DK}{y^2} + \frac{h}{2} - \frac{pD\overline{S}}{y^2} &= 0 \\ h - \frac{pD}{y}\int\limits_R^\infty f(x)\,dx &= 0 \end{aligned}$$

Systém lze přepsat ve tvaru

$$y = \sqrt{\frac{2D(K + p\overline{S})}{h}} \tag{12.1}$$

$$\int\limits_R^\infty f(x)\,dx = \frac{hy}{pD} \tag{12.2}$$

a řešit numericky iterační metodou (pokud tato konverguje k přesnému řešení). Zavádíme následující postup:

a) y_1 určíme ze zjednodušené verze rovnice (12.1):

$$y_1 = \sqrt{\frac{2DK}{h}}$$

b) dosazením y_1 do (12.2) určíme R_1

c) dosazením R_1 do (12.1) určíme y_2

d) dosazením y_2 do (12.2) určíme R_2 atd.

Za jistých předpokladů (pro $R = 0$ máme $\widehat{y} = \sqrt{\frac{2D(K+pEx)}{h}}$ z rov. (12.1)

$\widetilde{y} = \frac{pD}{h}$ z rov. (12.2),

pro $\widetilde{y} \geq \widehat{y}$ řešení existuje)

platí $y_n \to y^*,\ R_n \to R^*$.

Příklad 12.8. Určete řešení pro $K = 100$ Kč, $D = 1\,000$ jednotek zboží, $p = 10$ Kč, $h = 2$ Kč a rovnoměrné rozdělení poptávky X na intervalu $[0; 100]$.

Řešení. $\widetilde{y} = \frac{pD}{h} = 5\,000 > \widehat{y} = 774,5 \Rightarrow$ řešení existuje.
Nejprve určíme $\overline{S}$:

$$\overline{S} = \int_R^{\infty} (x-R)f(x)\,dx = \int_R^{100} (x-R)\cdot\frac{1}{100}\,dx = \frac{R^2}{200} - R + 50$$

Rovnice (12.1) a (12.2) jsou tvaru:

(12.1) $$y = \sqrt{\frac{2\cdot 1\,000\left(100+10\left(\frac{R^2}{200}-R+50\right)\right)}{2}}$$

(12.2) $$\int_R^{100} \frac{1}{100}\,dx = \frac{2y}{10\cdot 1\,000} \Rightarrow R = 100 - \frac{y}{50}$$

Nyní užitím iteračního procesu:

1. iterace: pro $\overline{S} = 0$ určíme z (12.1): $y_1 = \sqrt{\frac{2DK}{h}} = 316$
z (12.2): $R_1 = 100 - \frac{316}{50} = 93,68$

2. iterace: $\overline{S} = \frac{R_1^2}{200} - R_1 + 50 = 0,19971 \Rightarrow$ $y_2 = 319,37$ z rovnice (12.1)
$R_2 = 93,612$ z rovnice (12.2)

3. iterace: $y_3 = 319,43$ dosazením R_2 do (12.1)
$R_3 = 93,612$ dosazením y_3 do (12.2)

Hodnoty y_3, R_3 se už moc neliší od y_2, R_2, uzavíráme tedy

$$y^* = 319,4;\ \ R^* = 93,61.$$

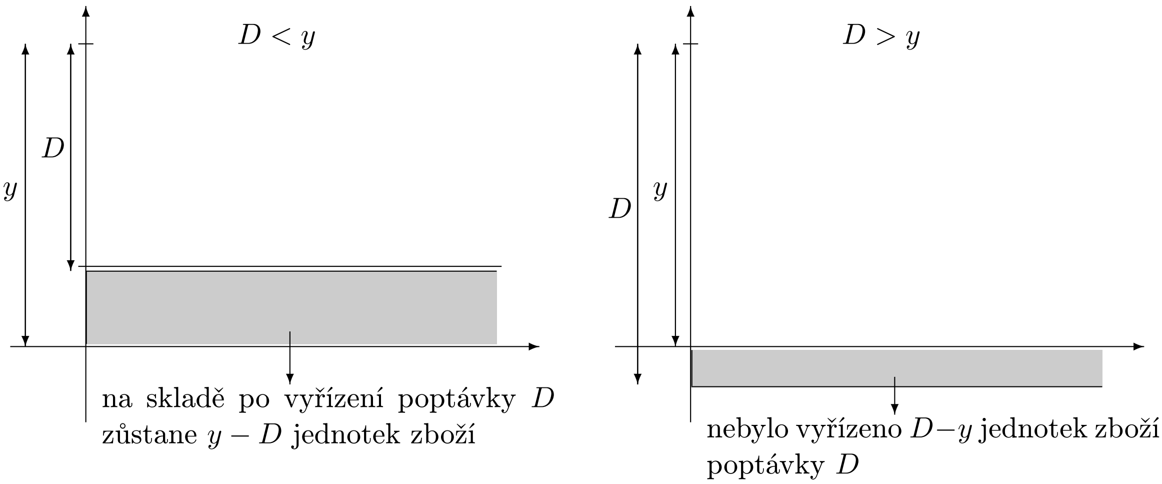
12.3.2 Model pro jednu položku a jedno období s jednorázovou objednávkou na začátku období a jednorázovou poptávkou

Předpoklady užití modelu a označení:

– na začátku období (před provedením objednávky) máme na skladě x jednotek zboží

– neuvažujeme cenu objednávky K (objednává se pouze jednou)

– po obdržení objednávky velikosti $(y - x)$ se odčerpá poptávka D
a na skladě zůstane $H(y) = \begin{cases} 0 & \dots D \geq y \\ y - D & \dots D < y, \end{cases}$
přičemž se nedostává $G(y) = \begin{cases} 0 & \dots D < y \\ D - y & \dots D \geq y \end{cases}$ jednotek zboží



– penalizační cena (při nevyřízení prodeje zboží) jednotky zboží za jednotku času $\dots p$

– cena skladování jednotky zboží za jednotku času (= za dané období) $\dots h$

– výrobní cena jednotky zboží $\dots c$

Vzhledem k typu poptávky budeme uvažovat dvě verze modelu:

a) poptávka je spojitá náhodná veličina s hustotou $f(D)$

$$\begin{array}{c} \text{očekávaná celková} \\ \text{cena za období} \end{array} = \begin{array}{c} \text{výrobní cena} \\ \text{zboží} \end{array} + \begin{array}{c} \text{očekávaná cena} \\ \text{skladování} \end{array} + \begin{array}{c} \text{očekávaná cena} \\ \text{penalizace} \end{array}$$

$$\begin{aligned} EC(y) &= c\,(y - x) + h \int\limits_0^\infty H(y)\, f(D)\, dD + p \int\limits_0^\infty G(y)\, f(D)\, dD \\ &= c\,(y - x) + h \int\limits_0^y (y - D)\, f(D)\, dD + p \int\limits_y^\infty (D - y)\, f(D)\, dD. \end{aligned}$$

Optimální y^* určíme nalezením stacionárního bodu, tj. řešením rovnice $\frac{\partial EC(y)}{\partial y} = 0$:

$$c + h \int_0^y f(D)\,dD - p \int_y^\infty f(D)\,dD = 0$$

$$c + h \int_0^y f(D)\,dD - p \left(1 - \int_0^y f(D)\,dD\right) = 0$$

Odtud y^* získáme z rovnice

$$\int_0^{y^*} f(D)\,dD = \frac{p-c}{p+h}$$

Model dává rozumný výsledek jen pro $p > c$, což je požadováno proto, že integrál na levé straně rovnosti je kladný.
Optimálním rozhodnutím pak tedy je:

$$\begin{array}{l} y^* > x \ldots \text{ objednat } y^* - x \\ y^* \le x \ldots \text{ neobjednávat další zboží} \end{array}$$

Příklad 12.9. Určete řešení pro $h = 0,5$ Kč, $p = 4,5$, $c = 0,5$ a hustotu

$$f(D) = \begin{cases} \frac{1}{10} & \ldots\ D \in \langle 0; 10\rangle \\ 0 & \ldots \text{ jinak} \end{cases} .$$

Řešení. Máme $\frac{p-c}{p+h} = 0,8$; tj.

$$P(D \le y^*) = \int_0^{y^*} f(D)\,dD = 0,8$$

$$\int_0^{y^*} \frac{1}{10}\,dD = 0,8$$

$$\frac{y^*}{10} = 0,8 \Rightarrow y^* = 8$$

b) poptávka je diskrétní náhodná veličina s pstní funkcí $f(D_i)$

$$EC(y) = c\,(y - x) + h \sum_{D_i=0}^{y} (y - D_i)\,f(D_i) + p \sum_{D_i=y+1}^{\infty} (D_i - y)\,f(D_i)$$

Nutnou podmínkou pro minimum je

$$EC(y^* - 1) \ge EC(y^*) \text{ a současně } EC(y^* + 1) \ge EC(y^*).$$

Dosazením EC máme (po jisté úpravě)

$$F(y^* - 1) = P(D \leq y^* - 1) \leq \frac{p - c}{p + h} \leq P(D \leq y^*) = F(y^*)$$

(jak všichni ví, F označuje příslušnou distribuční funkci).

Příklad 12.10. Určete řešení pro $h = 1$, $p = 1$, $c = 2$ a je-li poptávka zadána pstní funkcí $f(D_i)$:

D_i	0	1	2	3	4	5
$f(D_i)$	0,1	0,2	0,25	0,2	0,15	0,1

Řešení. Sestrojíme distribuční funkci F (jedná se o kumulativní psti získané sečtením pstí, že $D_i \leq y_i$):

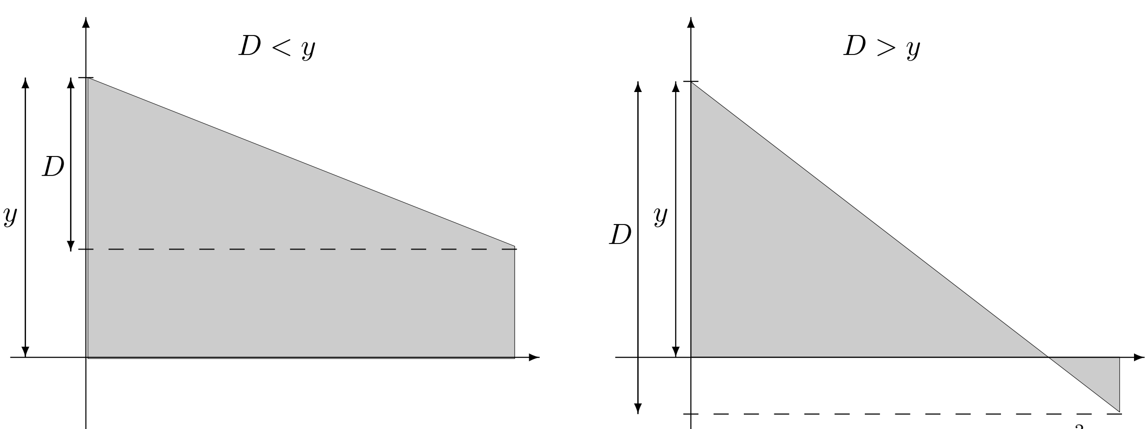
y_i	0	1	$\boxed{2}$	3	4	5
$F(y_i)$	0,1	0,3	0,55	0,75	0,9	1

$\frac{p-c}{p+h} = 0,4$, tj. $y^* = 2$.

12.3.3 Model pro jednu položku a jedno období se stejnoměrnou poptávkou v průběhu celého období

Předpoklady užití modelu a označení:

- na skladě máme x jednotek, jednorázově objednáme $(y - x)$, tj. máme k dispozici celkem y jednotek zboží
- poptávka za celé období je D
- neuvažujeme cenu K objednávky
- průběh zásob v závislosti na čase:



průměrně je na skladě ... $y - \frac{D}{2}$ průměrně je na skladě ... $\frac{y^2}{2D}$

průměrně se nedostává ... 0 průměrně se nedostává ... $\frac{(D-y)^2}{2D}$

Při řešení modelu si vybereme např. jen tu situaci, že y je spojitá veličina:

$$EC(y) = c\,(y-x) + h\left\{\int_0^y \left(y - \frac{D}{2}\right) f(D)\,dD + \int_y^\infty \frac{y^2}{2D}\, f(D)\,dD\right\} + p\int_y^\infty \frac{(D-y)^2}{2D}\, f(D)\,dD$$

Najdeme stacionární bod řešením rovnice $\frac{\partial EC(y)}{\partial y} = 0$:

$$c + h\left(\int_0^y f(D)\,dD + \int_y^\infty \frac{y}{D}\, f(D)\,dD\right) - p\int_y^\infty \frac{D-y}{D}\, f(D)\,dD = 0.$$

Úpravou dostaneme vztah pro optimální hodnotu y^*:

$$\int_0^{y^*} f(D)\,dD + y^*\int_{y^*}^\infty \frac{f(D)}{D}\,dD = \frac{p-c}{p+h}$$

Ad Příklad 12.9. Uvažujme situaci z příkladu 12.9 jen s tím rozdílem, že poptávka není jednorázová, ale konstantní (rovnoměrně rozložená) v průběhu celého období. Pak vztah pro optimální hodnotu y^* lze psát ve tvaru

$$\int_0^y \tfrac{1}{10}\,dD + y\int_y^{10} \tfrac{1}{10\,D}\,dD = 0,8$$

$$\tfrac{1}{10}\,(y - y\,\ln y + 2,3\,y) = 0,8$$

$$3,3\,y - y\,\ln y - 8 = 0$$

Řešením dané nelineární rovnice najdeme numericky:

$$y^* = 4,5.$$

12.3.4 Model pro jednu položku a jedno období s jednorázovou poptávkou na začátku období, přičemž uvažujeme cenu K objednávky

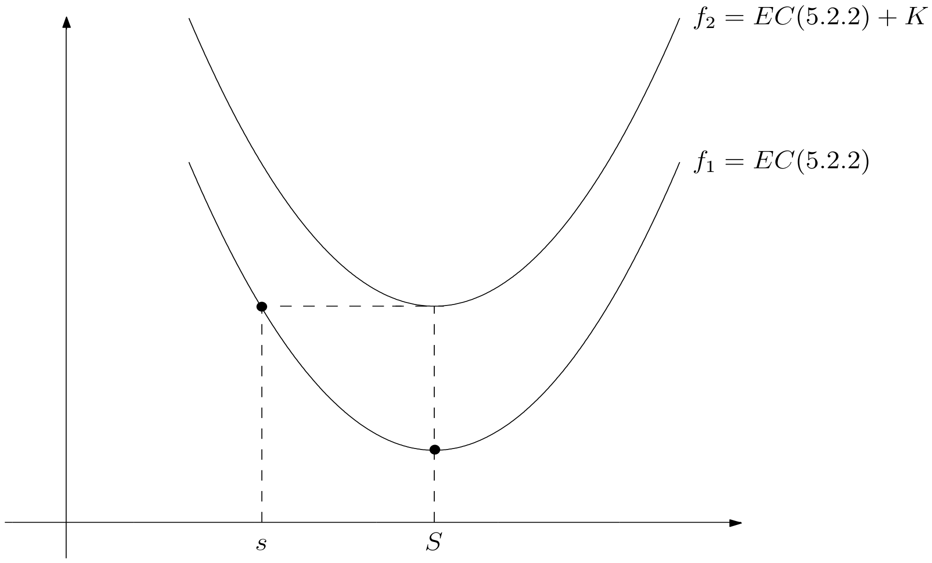
Předpoklady modelu jsou stejné jako u 12.3.2 s tím rozdílem, že uvažujeme cenu K objednávky.

Řešení modelu je analogické řešení modelu 12.2.2 (s diskontní sazbou). Pokud k objednávce nedojde, je výsledek stejný jako u modelu 12.3.2; jinak musíme při realizaci objednávky k celkové ceně modelu 12.3.2 přičíst konstantu K.

Uvažujme tedy větve $f_1 = EC(12.3.2)$

$$f_2 = EC(12.3.2) + K.$$

Analogicky modelu 12.2.2 (viz obrázek):



Označme s ... takový bod, že $f_1(s) = f_2(S)$

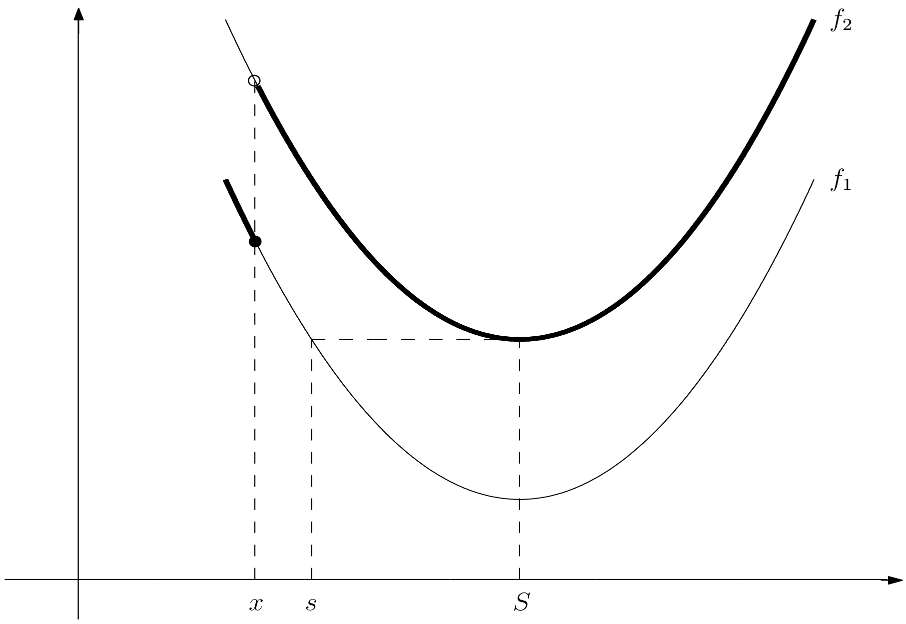
S ... minimum funkcí f_1, f_2

Na rozdíl od modelu 12.2.2 je bod s nalevo od bodu S.
S určíme na základě 12.3.2:

$$\int_0^S f(D)\,dD = \frac{p-c}{p+h}$$

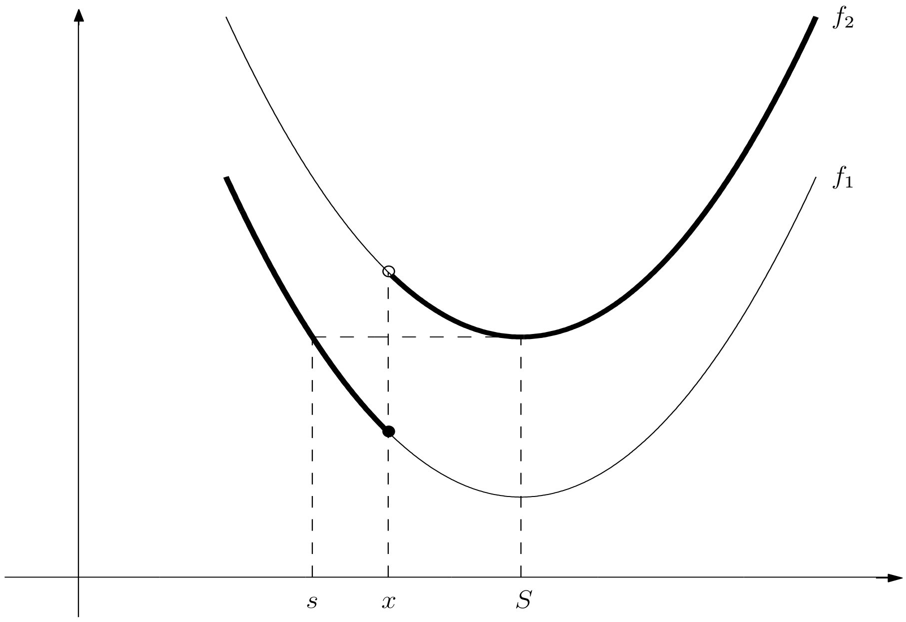
Body s, S rozdělí reálnou osu na tři části; podle toho, ve které z nich se nachází hodnota x počtu jednotek zboží na skladě před provedením objednávky. Mluvíme zde o tzv. $s - S$ politice.

a) $x < s$:



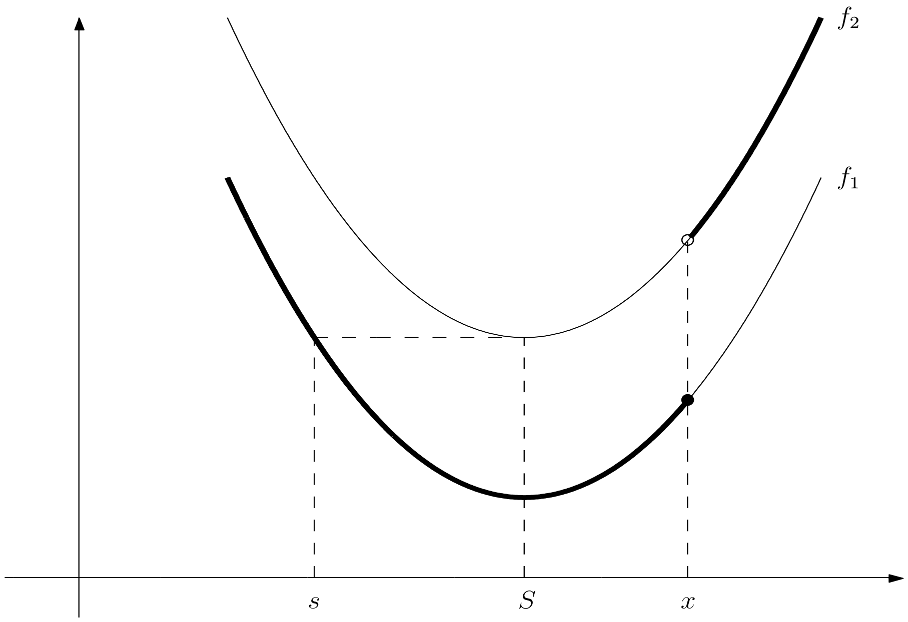
Protože x jednotek už neobjednáváme, jejich cena je $f_1(x) = EC(x, 12.3.2)$; $(y - x)$ jednotek objednáváme s cenou $f_2(x) = K + EC(x, 12.3.2)$. Optimum nastane pro S …tj. objednat $(S - x)$.

b) $s \leq x \leq S$:



Minimum ceny je v bodě x, tj. rozhodnutí: neobjednávat více.

c) $S < x$:



Minimum je v S, opět je nejlepší neobjednávat nic víc.

Ad Příklad 12.9. Uvažujme situaci v příkladu 12.9 z modelu 12.3.2, navíc ovšem cenu objednávky $K = 25$ Kč; $x = 0$. Protože $y^* = 8$, máme $S = 8$. Najdeme s:

$$
\begin{aligned}
EC(y, 12.3.2) &= 0,5\,(y-x) + 0,5 \int_0^y \frac{1}{10}(y-D)\,dD + 4,5 \int_y^{10} \frac{1}{10}(D-y)\,dD \\
&= 0,5\,(y-x) + 0,5 \left[y\,D - \frac{D^2}{2}\right]_0^y + 0,45 \left[\frac{D^2}{2} - y\,D\right]_y^{10} \\
&= 0,25\,y^2 - 4\,y + 22,5 - 0,5\,x \\
&= 0,25\,y^2 - 4\,y + 22,5
\end{aligned}
$$

$$
\begin{aligned}
EC(s) &= K + EC(S) \\
0,25\,s^2 - 4\,s + 22,5 &= 25 + 0,25 \cdot 64 - 4 \cdot 8 + 22,5 \\
s^2 - 16\,s - 36 &= 0 \\
s_{1,2} &= \begin{cases} -2 \\ 18 \end{cases}
\end{aligned}
$$

Volíme to s, jež je menší než $S = 8$, tj. $s = -2$. To znamená, že případ $x < s$ nemůže nastat, tj. rozhodnutí je v každém případě neobjednávat.

Pojmy k zapamatování

– V této kapitole jsme se zabývali matematickým zpracováním odpovědi na následující otázku z teorie zásob: Kdy a kolik zboží je optimální objednat?

– Rozhodnutí o skladové politice je prováděno s ohledem na celkovou cenu (tuto celkovou cenu chceme minimalizovat):

celková cena = cena zboží + cena objednávky (dopravy) + cena skladování + cena penalizační za škody vzniknuvší při nedostatku zásob

– Podle typu poptávky zboží rozdělujeme matematické modely na *deterministické* (je známa velikost poptávky zboží) a *pravděpodobnostní* (poptávka přesně není známa, pouze její pravděpodobná hodnota). Dále můžeme modely dělit podle toho, jestli se velikost poptávky s rostoucím časem mění nebo ne, a to na *statické* a *dynamické* (v případě deterministických modelů) a na *stacionární* a *nestacionární* (v případě pstních modelů).

– A. Deterministické modely **A.1. Statický model pro jednu položku**
Předpoklady použití modelu:

  – poptávka je konstantní ... β jednotek zboží za jednotku času
  – zásoby jsou jednorázové doplňovány po určité časové periodě t_0
  – není uvažován nedostatek zásob (záporné zásoby na skladě znamenají nevyřízenou objednávku zboží)
  – velikost objednávky ... y jednotek zboží
  – dodací lhůta objednávky je známá konstanta ... L časových jednotek
  – cena objednávky (dopravy) ... K finančních jednotek
  – h ... cena skladování jednotky zboží za jednotku času

Řešení modelu: řešením je optimální velikost objednávky y_0 taková, aby celková cena byla minimální

$$y_0 = \sqrt{\frac{2K\beta}{h}}$$

Objednávku musíme realizovat vždy L časových jednotek před vyčerpáním zásob.

A.2. Statický model pro jednu položku s diskontními cenami
Předpoklady použití modelu jsou stejné jako u předchozího modelu až na to, že do celkové ceny se navíc započítává i fakt, že pokud objednané množství překročí nějakou hranici q, klesá cena jednotky zboží. Tj. celková cena za jednotku času je dána vztahem:

$$TCU(y) = \begin{cases} f_1(y) = \beta \cdot c_1 + \frac{K\cdot\beta}{y} + \frac{h}{2}\cdot y \ \dots \ y < q \\ f_2(y) = \beta \cdot c_2 + \frac{K\cdot\beta}{y} + \frac{h}{2}\cdot y \ \dots \ y \geq q \end{cases} \qquad (c_1 > c_2)$$

Řešení modelu: určíme $y_0 = \sqrt{\frac{2K\beta}{h}}$; pokud $y_0 > q$, optimální hodnota objednávky je y_0; jinak určíme q_1 tak, aby $f_2(q_1) = f_1(y_0)$; pak pro $q < q_1$ objednáme q jednotek zboží, jinak je optimální hodnota opět y_0.

A.3. Statický model pro více druhů zboží s omezením skladového prostoru
Předpoklady použití modelu:

- maximální plocha skladování $\ldots A$
- požadavek skladovací plochy na jednotku i-tého druhu zboží $\ldots a_i$
- velikost objednávky zboží druhu $i \ldots y_i$
  musí platit
$$(*) \qquad \sum_{i=1}^{n} a_i\, y_i \le A$$
- zboží každého druhu je jednorázově doplňováno s periodou $t_i = \frac{y_i}{\beta_i}$, kde β_i, K_i, h_i mají stejný význam jako v modelu A.1, ale pro zboží druhu i není uvažován nedostatek zásob ani diskontní ceny.

Řešení modelu: Pokud $y_i^0 = \sqrt{\frac{2\,K_i\,\beta_i}{h_i}}$ splňují omezení $(*)$, máme optimum.

V opačném případě řešíme systém $\begin{bmatrix} y_i^0 = \sqrt{\frac{2\,K_i\,\beta_i}{h_i+2\,\lambda\,a_i}} \\ \sum\limits_{i=1}^{n} a_i\, y_i^0 - A = 0 \end{bmatrix}$ numericky: zvyšujeme λ s určitým krokem tak dlouho, až výraz $\sum\limits_{i=1}^{n} a_i\, y_i - A$ je záporný; z posledních dvou hodnot λ pak interpolací najdeme optimální λ, tj. i optimální y_i^0.

A.4. Dynamický model pro jednu položku a N období
Předpoklady použití modelu:

- poptávka v období $i \ldots D_i \qquad (D_i \in \mathbb{R})$
- množství zboží objednávané na začátku období $i \ldots z_i$
- množství zboží, které je na skladě na začátku období i před přijetím objednávaného $z_i \ldots x_i$
- jednotková cena skladování, jestliže zboží skladované v období i zůstane na skladě i na začátku období $i+1 \ldots h_i$
- cena objednávky v období $i \ldots K_i$
- výrobní cena zboží při objednávaném množství $z_i \ldots c_i(z_i)$
  (může se měnit v různých obdobích, lze uvažovat i diskontní ceny)
- cena nákupu zboží v období $i \ldots C_i(z_i) = \begin{cases} 0 & \ldots z_i = 0 \\ K_i + c_i(z_i) & \ldots z_i > 0 \end{cases}$

Řešení modelu: Úkolem je určit optimální z_i. Použijeme přímý algoritmus dynamického programování, rekurzivní rovnice jsou tvaru:

$$f_1(x_2) = \min_{0 \le z_1 \le D_1 + x_2} (C_1(z_1) + h_1\, x_2)$$

$$f_i(x_{i+1}) = \min_{0 \le z_i \le D_i + x_{i+1}} (C_i(z_i) + h_i\, x_{i+1} + f_{i-1}(x_i)) \quad \text{pro } i = 2, \ldots, N$$

V jakékoliv fázi musí platit $0 \le x_{i+1} \le D_{i+1} + \dots + D_N$.

A.5. Dynamický model plánování výroby jedné položky na N období
Cílem modelu je naplánovat výrobu tak, aby se minimalizovala její celková cena.
Označení a předpoklady použití modelu:

- výrobní cena jednotky zboží v pracovní době období $i \dots c_i$
- výrobní cena jednotky zboží v přesčasové době období i ($c_i < d_i$) $\dots d_i$
- skladovací cena jednotky zboží, které zůstalo na skladu na konci období $i \dots h_i$
- penalizační cena jednotky zboží objednaného v období i, ovšem vyrobeného v období $i + 1 \dots p_i$
- výrobní kapacita (v počtu jednotek) v období i v pracovní době $\dots a_{R_i}$
- výrobní kapacita v období i v přesčasové době $\dots a_{T_i}$
- poptávka (v počtu jednotek) v období $i \dots b_i$

Řešení modelu

a) $p_i = 0$**:** Díky tomuto předpokladu zde platí

$$(*) \qquad \sum_{i=1}^{k}(a_{R_i} + a_{T_i}) \ge \sum_{j=1}^{k} b_j, \qquad k = 1, 2, \dots, N$$

Úlohu lze formulovat jako dopravní úlohu a zapsat do tabulky. Protože platí $(*)$, stačí projít všechny sloupce uvedeným algoritmem pouze jednou a již máme optimální řešení.

b) $p_i \neq 0$**:** Už nemusí být splněna $(*)$, tj. po vyplnění tabulky postupem z a) musíme ještě provést optimalizační krok, kterým se eventuelně sníží celková cena výroby + skladování.

- B. Pravděpodobnostní modely **B.1. Model nepřetržité kontroly pro jednu položku**
Označení a předpoklady použití modelu:

- zboží na skladě je neustále kontrolováno a jestliže jeho množství klesne na jistou hodnotu R, přikročí se k objednávce o velikosti y
- náhodná veličina poptávky mezi objednáním a dodáním $\dots X$
- podmíněná hustota poptávky X během času mezi objednáním a dodáním (nenulová pro $x > 0$) $\dots r(x|\,t)$
- hustota doby t mezi objednáním a dodáním $\dots s(t)$
- hustota poptávky x během doby mezi objednáním a dodáním:

$$f(x) = \int_0^\infty r(x|\,t) \cdot s(t)\, dt$$

- velikost objednávky v jednom cyklu $\dots y$
- očekávaná celková poptávka za jednotku času (zpravidla 1 rok) $\dots D$

– jednotková skladovací cena za jednotku času … h

– jednotková cena penalizaci při nevyřízené objednávce za jednotku času … p

– cena objednávky (doprava) … K

Řešení modelu: Následujícím postupem určíme optimální R a y tak, aby celková cena skladování za jednotku času byla minimální:

a) Vypočteme y_1

$$y_1 = \sqrt{\frac{2DK}{h}}$$

b) dosazením y_1 do

$$\int_{R_n}^{\infty} f(x)\,dx = \frac{hy_n}{pD}$$

určíme R_1.

c) dosazením R_1 do

$$y_{n+1} = \sqrt{\frac{2D(K + p\overline{S}_n)}{h}},$$

kde

$$\overline{S}_n = \int_{R_n}^{\infty} (x - R_n) f(x)\,dx$$

určíme y_2.

Za jistých předpokladů platí $y_n \to y^*,\ R_n \to R^*$, tj. kroky b) a c) opakujeme tak dlouho, až jsou hodnoty dostatečně přesné.

B.2. Model pro jednu položku a jedno období s jednorázovou objednávkou na začátku období a jednorázovou poptávkou

Předpoklady užití modelu a označení:

– na začátku období (před provedením objednávky) máme na skladě x jednotek zboží

– neuvažujeme cenu objednávky K (objednává se pouze jednou)

– po obdržení objednávky velikosti $(y - x)$ se odčerpá poptávka D
a na skladě zůstane $H(y) = \begin{cases} 0 & \ldots D \geq y \\ y - D & \ldots D < y, \end{cases}$
přičemž se nedostává $G(y) = \begin{cases} 0 & \ldots D < y \\ D - y & \ldots D \geq y \end{cases}$ jednotek zboží

– penalizační cena (při nevyřízení prodeje zboží) jednotky zboží za jednotku času … p

– cena skladování jednotky zboží za jednotku času (= za dané období) … h

– výrobní cena jednotky zboží … c

Řešení modelu:

a) poptávka je spojitá náhodná veličina s hustotou $f(D)$
Optimální y^* získáme řešením rovnice

$$\int_0^{y^*} f(D)\,dD = \frac{p-c}{p+h}$$

b) poptávka je diskrétní náhodná veličina s pstní funkcí $f(D_i)$
Optimální y^* získáme řešením rovnice

$$F(y^*-1) \leq \frac{p-c}{p+h} \leq F(y^*)$$

Optimálním rozhodnutím pak je:

$$y^* > x \ldots \text{ objednat } y^* - x$$

$$y^* \leq x \ldots \text{ neobjednávat další zboží.}$$

B.3. Model pro jednu položku a jedno období se stejnoměrnou poptávkou v průběhu celého období
Předpoklady užití modelu a označení:

– na skladě máme x jednotek, jednorázově objednáme $(y-x)$, tj. máme k dispozici celkem y jednotek zboží
– poptávka za celé období je D
– neuvažujeme cenu K objednávky

Řešení modelu: Optimální hodnotu y^* (pro spojitou náhodnou veličinu y) získáme řešením rovnice:

$$\int_0^{y^*} f(D)\,dD + y^* \int_{y^*}^{\infty} \frac{f(D)}{D}\,dD = \frac{p-c}{p+h}.$$

B.4. Model pro jednu položku a jedno období s jednorázovou poptávkou na začátku období, přičemž uvažujeme cenu K objednávky
Předpoklady modelu jsou stejné jako u B.2 s tím rozdílem, že uvažujeme cenu K objednávky.
Řešení modelu: Je analogické řešení modelu A.2. Pokud k objednávce nedojde, je výsledek stejný jako u modelu B.2; jinak musíme při realizaci objednávky k celkové ceně modelu B.2 přičíst konstantu K.
Uvažujeme tedy větve $f_1 = EC(B.2)$
$f_2 = EC(B.2) + K.$
Označme $s \ldots$ takový bod, že $f_1(s) = f_2(S)$
$S \ldots$ minimum funkcí f_1, f_2
Body s, S rozdělí reálnou osu na tři části; optimální řešení najdeme podobně jako v modelu A.2.

Kontrolní otázky

1. U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.
   a) Časová perioda t_0, po které jsou zásoby jednorázově doplněny, je ve „statickém modelu pro jednu položku“ konstantní.
   b) Ve „statickém modelu pro jednu položku“ je optimální hodnota y_0 jediným minimem funkce $TCU(y)$.
   c) Ve „statickém modelu pro jednu položku s diskontními cenami“ může být funkce $TCU(y)$ na některém intervalu konkávní.
   d) V případě „statického modelu pro více druhů zboží s omezením skladového prostoru“ je zahrnut i případ nedostatečných zásob zboží libovolného druhu i.
   e) Při hledání optima „statického modelu pro více druhů zboží s omezením skladového prostoru“ se řeší systém n rovnic o n neznámých.
   f) V případě „dynamického modelu pro jednu položku a N období“ musíme v každém období vyčerpat všechny zásoby.
   g) Při řešení „dynamického modelu plánování výroby jedné položky na N období“ využíváme dynamického programování.
   h) V případě „dynamického modelu plánování výroby jedné položky na N období“ je zahrnut i případ nedostatečných zásob zboží v libovolném období i.
   i) Časová perioda t, po které jsou zásoby jednorázově doplněny, je v „modelu nepřetržité kontroly pro jednu položku“ konstantní.
   j) V případě „modelu pro jednu položku a jedno období s jednorázovou objednávkou na začátku období a jednorázovou poptávkou“ je optimální velikost objednávky přímo úměrná rozdílu penalizační a výrobní ceny jednotky zboží.
   k) V „modelu pro jednu položku a jedno období se stejnoměrnou poptávkou v průběhu celého období“ je poptávka za celé období známá konstanta.
   l) V případě „modelu pro jednu položku a jedno období s jednorázovou poptávkou na začátku období, přičemž uvažujeme cenu K objednávky“ je zahrnut i případ nedostatečných zásob zboží.

Odpovědi na otázky
1a) – A, 1b) – A, 1c) – N, 1d) – N, 1e) – N, 1f) – N, 1g) – N, 1h) – A, 1i) – N, 1j) – A, 1k) – N, 1l) – A.

Cvičení

1. Roční poptávka zboží je 1500 jednotek. Cena objednávky je vždy 20 Kč navíc k ceně zboží. Skladovací cena jednotky zboží je 2 Kč na měsíc a není dovolen nedostatek zásob.
   a) Určete optimální velikost objednávky a čas mezi jednotlivými objednávkami.
   b) Určete rozdíl roční celkové ceny optimální strategie a roční celkové ceny, kdyby se objednávalo každý měsíc 12-krát za rok.

2. Cena zboží je 4 Kč za jednotku, ale při odběru 150 kusů a více dostaneme 10% slevu. Firma, která spotřebuje 20 kusů za den, se rozhoduje, jestli se jí vyplatí využít akční nabídky. Cena objednávky je 50 Kč a skladovací cena jednotky zboží je $0,30$ Kč za 1 den. Měla by firma využít slevu?

3. Pro potřeby výrobního procesu jsou skladovány 4 druhy materiálu. Poptávka po všech typech je stále stejná a není povolen nedostatek zásob. D_i označuje potřebné množství i-tého druhu za rok. V souladu se značením v kapitole 12.2.3 jsou dány následující údaje:

materiál i	K_i	β_i	h_i	D_i
1	100	10	$0,1$	10 000
2	50	20	$0,2$	5 000
3	90	5	$0,2$	7 500
4	20	10	$0,1$	5 000

Určete optimální velikosti objednávek pro jednotlivé druhy materiálu, jestliže požadujeme, aby maximální počet všech objednávek nepřekročil 200.

4. Uvažujme hospodaření jedné položky v průběhu čtyř období:

období i	D_i	K_i	h_i
1	5	5	1
2	7	7	1
3	11	9	1
4	3	7	1

Nákupní cena je 1 Kč za jednotku u prvních šesti jednotek a 2 Kč za každou další jednotku. Stanovte optimální plán objednávek.

5. Poptávka po produktu může být v následujících čtyřech obdobích uspokojena buď výrobou v normální pracovní době nebo v přesčasové době. Kapacity výroby a poptávka pro jednotlivá období jsou dány v následující tabulce:

období i	kapacity výroby		poptávka
	a_{R_i}	a_{T_i}	b_i
1	120	50	160
2	70	20	80
3	90	80	150
4	70	50	100

Výrobní cena jednotky zboží je po všechna období stejná, v normální pracovní době je 1 Kč, v přesčasové době je 2 Kč. Skladovací cena jednotky zboží je také po všechna období stejná a je rovna $0,5$ Kč za jedno období. Dále také nesmí nastat případ, že by v některém období výroba nepokryla objednávky. Stanovte optimální plán výroby.

6. Uvažujme model nepřetržité kontroly pro jednu položku. Určete optimální řešení pro $K = 20$ Kč, $D = 10\,000$ jednotek zboží, $p = 4$ Kč, $h = 2$ Kč a normální rozdělení poptávky X se střední hodnotou 100 a rozptylem 4.

7. Jednorázová poptávka zboží pro jedno období má exponenciální rozdělení se střední hodnotou 10 jednotek. Objednáváme jednorázově na začátku období, cena skladování a penalizační cena jednotky zboží za jednotku času jsou 1 Kč, resp. 3 Kč. Výrobní cena jednotky zboží je 2 Kč. Určete optimální velikost objednávky, jestliže před provedením objednávky máme na skladě 2 jednotky zboží. Jaká je optimální velikost objednávky, jestliže před provedením objednávky máme na skladě 5 jednotek zboží?

8. Řešte příklad 7, jestliže jednorázová poptávka zboží pro jedno období má Poissonovo rozdělení se střední hodnotou 10 jednotek.

9. Poptávka zboží, která má rozdělení dáno hustotou
$$f(D) = \begin{cases} 1 - \frac{1}{2}D & \dots\ D \in \langle 0, 2\rangle \\ 0 & \dots \text{ jinak,} \end{cases}$$
je konstantní (stejnoměrně rozdělená) v průběhu celého období. Objednáváme jednorázově na začátku období, cena skladování a penalizační cena jednotky zboží za jednotku času jsou 1 Kč, resp. 4 Kč. Výrobní cena jednotky zboží je 2 Kč. Určete optimální velikost objednávky.

10. Určete optimální strategii objednávání u modelu pro jednu položku a jedno období s jednorázovou poptávkou, je-li dáno
$$f(D) = \begin{cases} \frac{1}{5}D & \dots\ D \in \langle 5, 10\rangle \\ 0 & \dots \text{ jinak,} \end{cases}$$
$h = 1$ Kč, $p = 5$ Kč a $c = 3$ Kč. Uvažujeme také cenu objednávky $K = 5$ Kč a víme, že před provedením objednávky máme na skladě 10 jednotek zboží.

Výsledky příkladů viz **??**.

Výsledky

ad 1. ad a) $y_0 \approx 50$ jednotek, $t_0 \approx 12$ dní; ad b) rozdíl celkových cen za rok je 540, 20 Kč.

ad 2. Firma by měla objednat 150 kusů zboží a využit tak slevu.

ad 3. Úlohu je třeba řešit s omezením $\sum_{i=1}^{4} \frac{D_i}{y_i} \leq 200$, dostaneme pak jiný vztah pro výpočet $\boldsymbol{y}_0$. Optimální řešení je $\boldsymbol{y}_0 = (201,49; 123,09; 110,57; 119,58)$ $(\lambda^* \approx 0,103)$.

ad 4. $(z_1, z_2, z_3, z_4) = (5, 7, 14, 0)$ nebo $(6, 6, 14, 0)$.

ad 5. Plán výroby v normální pracovní době je $\boldsymbol{x}_R = (120, 70, 90, 70)$, v přesčasové době $\boldsymbol{x}_T = (40, 10, 60, 30)$ s celkovou cenou 630 Kč.

ad 6. Poptávku X je třeba transformovat na standardizované normální rozdělení, jehož hodnoty najdeme v tabulce; optimální řešení je $y^* = 447,34$; $R^* = 2,379$.

ad 7. Jestliže $x = 2$, objednat 0,877 jednotek. Pro $x = 5$ neobjednávat další zboží.

ad 8. Jestliže $x = 2$, objednat 6 jednotek. Pro $x = 5$ objednat 3 jednotky.

ad 9. $y^* = 2,71$.

ad 10. Pro $x < 3,78$ objednat $(6,7 - x)$; jinak neobjednávat.

Maplety

V následujících mapletech si můžete některé studované pojmy přiblížit, případně si sestavit vlastní zadání příkladů.

1. Určitý integrál
2. Integrování

13 Pravděpodobnostní dynamické programování

Průvodce studiem

V poslední kapitole se budeme zabývat problematikou pravděpodobnostního dynamického programování. Pro obecnou formulaci úloh pravděpodobnostního dynamického programování budeme využívat Markovské řetězce.

Základní otázky, na které budeme hledat odpověď, jsou: Jakou posloupnost strategií zvolit, aby celkové výnosy za dané období byly maximální? Jaký výnos celkem přinese daná posloupnost strategií v nejbližších obdobích?

Řěšení budeme hledat pomocí rekurzivní rovnice.

Cíle

Po prostudování této kapitoly budete schopni:

- Formulovat úlohy pravděpodobnostního dynamického programování.
- Využívat Markovské řetězce.
- Hledat řešení vybraných typů úloh.

Následující úloha jen slouží jako demonstrace problematiky. Využití uváděných metod spadá i do jiných oblastí, právě i do probírané teorie skladových zásob, dále teorie obnovy, řízení toku peněz, regulace vodní nádrže, apod.

Příklad 13.1. Problém zahradnice.
Zahradnice každý rok testuje kvalitu půdy své zahrady. Výsledky dělí do tří kategorií: výborná (stav 1), dobrá (stav 2), špatná (stav 3).
Všimla si, že může předpokládat, že úrodnost půdy závisí na její kvalitě pouze v předchozím roce, tj. že psti přechodu z jednoho stavu do druhého lze reprezentovat Markovského řetězcem:

$$
(p^1_{ij}) = \boldsymbol{P}^1 = \begin{array}{c|ccc} & 1 & 2 & 3 \to \text{stav půdy v příštím roce} \\ 1 & 0,2 & 0,5 & 0,3 \\ 2 & 0 & 0,5 & 0,5 \\ 3 & 0 & 0 & 1 \\ \downarrow & & & \\ \text{stav půdy letos} & & & \end{array}
$$

Z matice přechodových pstí je vidět, že když se půda nechá ladem, jde její produktivita od deseti k pěti. Samozřejmě, že psti přechodu mohou být ovlivněny přihnojováním půdy:

$$
(p^2_{ij}) = \boldsymbol{P}^2 = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 0,3 & 0,6 & 0,1 \\ 0,1 & 0,6 & 0,3 \\ 0,05 & 0,4 & 0,55 \end{pmatrix} \end{array}
$$

($\boldsymbol{P}^2$ neznamená „pé na druhou“, ale „pé s indexem 2“).
Příslušné roční výnosy půdy (ve stovkách dolarů), když půda přijde ze stavu i do stavu j, jsou:

- když se nehnojí:

$$
(r^1_{ij}) = \boldsymbol{R}^1 = \begin{pmatrix} 7 & 6 & 3 \\ 0 & 5 & 1 \\ 0 & 0 & -1 \end{pmatrix}
$$

- když se hnojí:

$$
(r^2_{ij}) = \boldsymbol{R}^2 = \begin{pmatrix} 6 & 5 & -1 \\ 7 & 4 & 0 \\ 6 & 3 & -2 \end{pmatrix}
$$

Z matic $\boldsymbol{R}^1, \boldsymbol{R}^2$ je například vidět, že proces $1 \to 1$ má nižší výnos při hnojení než bez hnojení, protože ve výnosu při hnojení je započítána cena hnojiva.
Otázka úlohy:

1) Hnojit či nehnojit v daném roce, aby celkové výnosy za `k` let byly maximální?

2) Jaký výnos celkem přinese daná posloupnost rozhodnutí v nejbližších `k` letech? (Např. jednoduchá je odpověď v případě tzv. stacionární politiky

   „hnojit jen za stavu 3“:

   v tomto případě se matice s výhledem na $k = \infty$ období najdou snadno:

$$
\boldsymbol{P} = \begin{pmatrix} 0,2 & 0,5 & 0,3 \\ 0 & 0,5 & 0,5 \\ 0,05 & 0,4 & 0,55 \end{pmatrix}, \quad \boldsymbol{R} = \begin{pmatrix} 7 & 6 & 3 \\ 0 & 5 & 1 \\ 6 & 3 & -2 \end{pmatrix}
$$

   (první dva řádky jsou shodné s maticemi $\boldsymbol{P}^1, \boldsymbol{R}^1$, třetí řádek s maticemi $\boldsymbol{P}^2, \boldsymbol{R}^2$)).

Řešení.

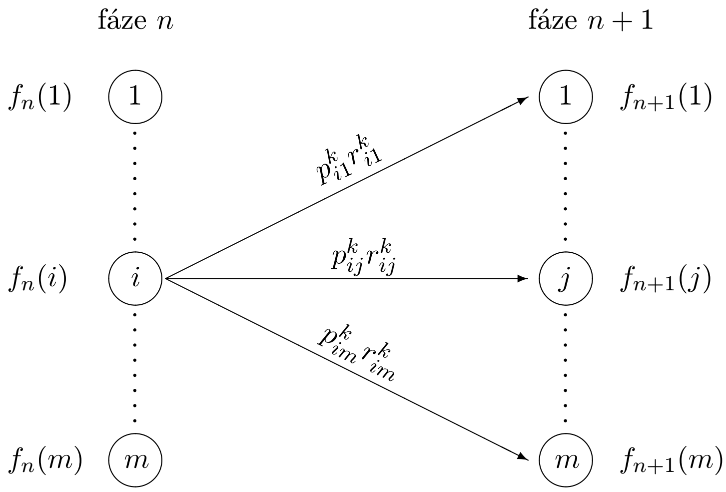
a) řešení úlohy pro konečné $k = N$

Úlohu vyřešíme pomocí dynamického programování:

m ... počet možných stavů (v našem příkladu $m = 3$)

$f_n(i)$ ... optimální (očekávaný) výnos z fází $n, n+1, \ldots, N$ za daného stavu i na počátku období n

Využijeme zpětné rekurzivní rovnice (viz obrázek):



$$\boxed{\begin{array}{l} f_n(i) = \max\limits_k \left\{ \sum\limits_{j=1}^{m} p_{ij}^k \left[r_{ij}^k + f_{n+1}(j) \right] \right\}, \; n = 1, 2, \ldots, N \\[2ex] f_{N+1} = 0 \text{ pro každé } j \end{array}}$$

Označíme-li $v_i^k = \sum\limits_{j=1}^{m} p_{ij}^k r_{ij}^k$, rekurzivní rovnice mají tvar

$$f_N(i) = \max_k \left\{ v_i^k \right\}$$

$$f_n(i) = \max_k \left\{ v_i^k + \sum_{j=1}^{m} p_{ij}^k f_{n+1}(j) \right\} \quad \text{pro } n = 1, 2, \ldots, N-1$$

Např. v naší úloze pro $k = 1$

$$\left.\begin{array}{l} v_1^1 = 0,2 \cdot 7 + 0,5 \cdot 6 + 0,3 \cdot 3 = 5,3 \\ v_2^1 = 0 \cdot 0 + 0,5 \cdot 5 + 0,5 \cdot 1 = 3 \\ v_3^1 = 0 \cdot 0 + 0 \cdot 0 + 1 \cdot (-1) = -1 \end{array}\right\} \begin{array}{l} \text{zisky jednotlivých} \\ \text{stavů při alternativě} \\ \text{„nehnojit“} \end{array}$$

Ad Příklad 13.1. Předpokládejme $N = 3$ ... tříletý horizont plánování

$$\left.\begin{aligned} v_1^2 &= 4,7 \\ v_2^2 &= 3,1 \\ v_3^2 &= 0,4 \end{aligned}\right\} \text{ zisky jednotlivých stavů při alternativě „hnojit“}$$

fáze 3:

	v_i^k		opt. řešení	
i	$k = 1$	$k = 2$	$f_3(i)$	k^*
1	$5,3$	$4,7$	$5,3$	1
2	3	$3,1$	$3,1$	2
3	-1	$0,4$	$0,4$	2

fáze 2:

	$v_i^k + p_{i1}^k f_3(1) + p_{i2}^k f_3(2) + p_{i3}^k f_3(3)$		opt. řešení	
i	$k = 1$	$k = 2$	$f_2(i)$	k^*
1	$5,3 + 0,2 \cdot 5,3 + 0,5 \cdot 3,1$ $+0,3 \cdot 0,4 = 8,03$	$4,7 + 0,3 \cdot 5,3 + 0,6 \cdot 3,1$ $+0,1 \cdot 0,4 = 8,19$	$8,19$	2
2	$3 + 0 + 0,5 \cdot 3,1$ $+0,5 \cdot 0,4 = 4,75$	$3,1 + 0,1 \cdot 5,3 + 0,6 \cdot 3,1$ $+0,3 \cdot 0,4 = 5,61$	$5,61$	2
3	$-1 + 0 + 0 + 1 \cdot 0,4$ $= -0,6$	$0,4 + 0,05 \cdot 5,3 + 0,4 \cdot 3,1$ $+0,55 \cdot 0,4 = 2,13$	$2,13$	2

fáze 1:

	$v_i^k + p_{i1}^k f_2(1) + p_{i2}^k f_2(2) + p_{i3}^k f_2(3)$		opt. řešení	
i	$k = 1$	$k = 2$	$f_1(i)$	k^*
1	$5,3 + 0,2 \cdot 8,19 + 0,5 \cdot 5,61$ $+0,5 \cdot 2,13 = 10,38$	$4,7 + 0,3 \cdot 8,19 + 0,6 \cdot 5,61$ $+0,1 \cdot 2,13 = 10,74$	$10,74$	2
2	$3 + 0 + 0,5 \cdot 5,61$ $+0,5 \cdot 2,13 = 6,87$	$3,1 + 0,1 \cdot 8,19 + 0,6 \cdot 5,61$ $+0,3 \cdot 2,13 = 7,92$	$7,92$	2
3	$-1 + 0 + 0 + 1 \cdot 2,13$ $= 1,13$	$0,4 + 0,05 \cdot 8,19 + 0,4 \cdot 5,61$ $+0,55 \cdot 2,13 = 4,23$	$4,23$	2

Můžeme psát odpověď: zahradnice by měla v prvním a druhém roce hnojit vždy, a ve třetím roce hnojit tehdy, pokud je stav 2 nebo 3.

Optimální zisk záleží na stavu půdy v roce plánování (= před prvním rokem, který jsme v plánování uvažovali).

Při půdě ve stavu	1 bude	$10,74 \cdot 100\,\$$
	2	$7,92 \cdot 100\,\$$
	3	$4,23 \cdot 100\,\$$

Výhodou zpětného chodu metody dynamického programování je to, že pokud chceme zvýšit plánovací horizont o 1 rok (tj. na 4 roky), všechna data zůstanou využita, pouze při fázi 1 přidáme ještě další fázi – fázi 0.

Tato „úloha zahradnice“ může být zobecněna dvěma způsoby:

– přechodové psti mohou být každý rok jiné $\dots p_{ij}^{k,n}$. Pak by byly rovnice ve tvaru

$$
\begin{aligned}
f_N(i) &= \max_k \left\{ v_i^{k,N} \right\} \\
f_n(i) &= \max_k \left\{ v_i^{k,n} + \sum_{j=1}^{m} p_{ij}^{k,n} f_{n+1}(j) \right\} \quad \text{pro } n = 1, 2, \dots, N-1
\end{aligned}
$$

– chceme znát současnou hodnotu očekávaných zisků (tj. výnosy příštích let budou násobeny koeficientem $\alpha < 1$):

D dolarů příští rok $= \alpha \cdot D$ dolarů dnes
($\alpha = \frac{1}{1+t}$, kde t je roční úroková míra (= interest rate)).

Pak bychom mohli užít rovnice ve tvaru

$$
\begin{aligned}
f_N(i) &= \max_k \left\{ v_i^{k} \right\} \\
f_n(i) &= \max_k \left\{ v_i^{k} + \alpha \sum_{j=1}^{m} p_{ij}^{k} f_{n+1}(j) \right\} \quad \text{pro } n = 1, 2, \dots, N-1
\end{aligned}
$$

Při vyhodnocení stacionární politiky (=stále stejné politiky v každém roce) užíváme vztahu

$$
f_n(i) = N_i + \sum_{j=1}^{m} p_{ij}\, f_{n+1}(j)
$$

roční zisk zvolené politiky

příslušný řádek matice zvolené politiky

Například pro stacionární politiku „hnojit jen ve stavu 3“ je

$$\left.\begin{aligned} v_1 &= v_1^1 = 5,3 \\ v_2 &= v_2^1 = 3 \\ v_3 &= v_3^2 = 0,4 \end{aligned}\right\} \quad \text{odhad hodnoty } f_n(i) \text{ pro nižší } n \text{ lze určit a uspořádat do tabulky:}$$

i \ n	3	2	1
1	$5,3$	$5,3+0,2\cdot 5,3+0,5\cdot 3$ $+0,3\cdot 0,4=7,98$	$5,3+0,2\cdot 7,98+0,5\cdot 4,7$ $+0,3\cdot 2,09=\underline{\underline{9,87}}$
2	3	$3+0+0,5\cdot 3+0,5\cdot 0,4$ $=4,7$	$3+0,5\cdot 4,7+0,5\cdot 2,09$ $=\underline{\underline{6,39}}$
3	$0,4$	$0,4+0,05\cdot 5,3+0,4\cdot 3$ $+0,55\cdot 0,4=2,09$	$0,4+0,05\cdot 7,98+0,4\cdot 4,7$ $+0,55\cdot 2,09=\underline{\underline{3,83}}$

zisky vzhledem k počátečnímu stavu

b) řešení úlohy pro k nekonečné (nebo hodně velké)

1) *metoda úplného vyčíslení* ... ohodnotíme všechny stacionární politiky a vybereme tu optimální (lze ji užít jen při nízkém počtu stacionárních politik):

pro S stacionárních politik
$$\begin{array}{c} \boldsymbol{P}^1, \boldsymbol{R}^1 \\ \boldsymbol{P}^2, \boldsymbol{R}^2 \\ \vdots \\ \boldsymbol{P}^S, \boldsymbol{R}^S \end{array}$$

vypočteme

- v_i^s ... výnos jednoho kroku politiky s ve stavu i ($i = 1, 2, \ldots, m$)
- π_i^s ... dlouhodobé stacionární psti pro jednotlivé politiky (slovo „politika“ je zde užito v ženském rodě):
  řešením systému
  $$\boldsymbol{\pi}^s \cdot \boldsymbol{P}^s = \boldsymbol{\pi}^s$$
  $$\sum_{i=1}^{m} \pi_i^s = 1$$
  (je zde maticové zapsán systém $(m+1)$ lineárních rovnic o m neznámých (jedna z prvních m rovnic je nadbytečná))
- očekávaný výnos pro politiku s:
  $$E_s = \sum_{i=1}^{m} \pi_i^s v_i^s$$

Pak optimální politika je ta, pro níž je E_s maximální.

Ad Příklad 13.1. Díky 3 stavům a dvěma možnostem v každém z nich existuje $2^3 = 8$ politik: hnojit v každém stavu
nehnojit nikdy
hnojit jen ve stavu 1
hnojit jen ve stavu 2
hnojit jen ve stavu 3
hnojit jen ve stavech 1 nebo 2
hnojit jen ve stavech 1 nebo 3
hnojit jen ve stavech 2 nebo 3

Kdybychom pro každou z nich určili $\boldsymbol{P}^s, \boldsymbol{R}^s, v_i^s, \pi_i^s$, zjistili bychom, že optimální stacionární politika je „hnojit v každém stavu“.

2) *metoda zlepšení politiky (+ neuvažujeme inflaci)*
Už pro 4 alternativy v každém roce (např. hnojit, nehnojit, hnojit 2× ročně, hnojit 3× ročně) je množství stacionárních politik $4^3 = 256$. Zde je výhodnější užít následující

metodu:

Přepišme rovnici jednoho roku politiky

$$f_n(i) = v_i + \sum_{j=1}^{m} p_{ij} f_{n+1}(j)$$

do tvaru

$$f_\eta(i) = v_i + \sum_{j=1}^{m} p_{ij} f_{\eta-1}(j) \qquad (*)$$

kde η je počet fází, které zbývá vzít v úvahu (abychom mohli studovat asymptotický proces $\eta \to \infty$);
dále

$$(\pi_1, \pi_2, \ldots, \pi_m) \ldots \text{psti rovnovážného stavu asymptotického procesu,}$$

očekávaný zisk za jednu fázi (= jeden rok) je

$$E = \pi_1 v_1 + \cdots + \pi_m v_m.$$

Lze ukázat, že pro velká η platí

$$f_\eta(i) = \eta \cdot E + f(i),$$

kde $f(i)$ je jakási konstantní hodnota pro stav i očekávaná při asymptotickém chování.
Rekurzivní rovnici $(*)$ lze tedy psát ve tvaru

$$\eta \cdot E + f(i) = v_i + \sum_{j=1}^{m} p_{ij} \left\{(\eta - 1) \cdot E + f(j)\right\}$$

a po úpravě

$$E = v_i + \left(\sum_{j=1}^{m} p_{ij}\, f(j) \right) - f(i) \qquad \text{pro } i = 1, 2, \ldots, m$$

To je systém m rovnic o $(m+1)$ neznámých, tj. jednu z nich lze zvolit, např. $f(m) = 0$.

krok 1: zvolíme libovolnou politiku s, tj. $\boldsymbol{P}^s, \boldsymbol{R}^s$;
předpokládejme $f^s(m) = 0$ a vyřešme systém m rovnic o m neznámých

$$E^s = v_i^s + \left(\sum_{j=1}^{m} p_{ij}^s\, f^s(j) \right) - f^s(i), \qquad i = 1, 2, \ldots, m$$

krok 2: (zlepšení politiky)
pro každý stav i určíme alternativu k s maximální hodnotou výrazu

$$v_i^k + \sum_{j=1}^{m} p_{ij}^k\, f^s(j); \qquad \to \quad \text{tím se určí maximální } E \text{ pro každý stav}$$

to bude nová politika t.

Pokračujeme opakováním kroků 1,2 tak dlouho, až $s = t$.

Ad Příklad 13.1.

krok 1: zvolme politiku s … „nehnojit vůbec“ $\Rightarrow$ $\boldsymbol{P}^s$ = původní $\boldsymbol{P}^1$
$\boldsymbol{R}^s$ = původní $\boldsymbol{R}^1$

řešme systém

$$\begin{array}{ll} E + f(1) - 0,2\,f(1) - 0,5\,f(2) - 0,3\,f(3) &= 5,3 \\ E + f(2) \qquad\qquad - 0,5\,f(2) - 0,5\,f(3) &= 3 \\ E + f(3) \qquad\qquad\qquad\qquad - f(3) &= -1 \\ \hline \end{array}$$

$$E = -1 \Rightarrow \begin{array}{l} f(1) \approx 12,88 \\ f(2) = 8 \\ f(3) = 0 \end{array}$$

krok 2:

i	$v_i^k + p_{i1}^k f(1) + p_{i2}^k f(2) + p_{i3}^k f(3)$		opt. řešení	
	$k = 1$	$k = 2$	$f(i)$	k^*
1	$5,3 + 0,2 \cdot 12,88$ $+0,5 \cdot 8 = 11,875$	$4,7 + 0,3 \cdot 12,88$ $+0,6 \cdot 8 = 13,36$	$13,36$	2
2	$3 + 0,5 \cdot 8 = 7$	$3,1 + 0,1 \cdot 12,88$ $+0,6 \cdot 8 = 9,19$	$9,19$	2
3	$-1 + 0 = -1$	$0,4 + 0,05 \cdot 12,88$ $+0,4 \cdot 8 = 4,24$	$4,24$	2

$$\Downarrow$$

nová politika t … „hnojit vždy“; $t \neq s$, tj. pokračujeme.

krok 1’: $\boldsymbol{P}^2, \boldsymbol{R}^2 \Rightarrow$ řešíme systém, získáváme $E = 2,26$
$f(1) = 6,75$
$f(2) = 3,79$
$f(3) = 0$

krok 2’:

i	$v_i^k + p_{i1}^k f(1) + p_{i2}^k f(2) + p_{i3}^k f(3)$		opt. řešení	
	$k = 1$	$k = 2$	$f(i)$	k^*
1	$5,3 + 0,2 \cdot 6,75$ $+0,5 \cdot 3,79 = 8,54$	$4,7 + 0,3 \cdot 6,75$ $+0,6 \cdot 3,79 = 8,99$	$8,99$	2
2	$3 + 0,5 \cdot 3,79 = 4,89$	$3,1 + 0,1 \cdot 6,75$ $+0,6 \cdot 3,79 = 6,05$	$6,05$	2
3	$-1 + 0 = -1$	$0,4 + 0,05 \cdot 6,75$ $+0,4 \cdot 3,79 = 2,25$	$2,25$	2

$$\Downarrow$$

$t = s$, tj. končíme, $E = 2,26$ je zisk optimální politiky

Tedy k optimálnímu řešení stacionární politiky jsme dospěli už po dvou iteracích, nemuseli jsme procházet všech 8 politik.

3) *metoda zlepšení politiky (+ uvažuje inflaci)*

$$f_\eta(i) = \max_k \left\{ v_i^k + \alpha \sum_{j=1}^{m} p_{ij}^k f_{\eta-1}(j) \right\}$$

Pro $\eta \to \infty$ je $f_\eta(i) = f(i)$, tj. $f_\eta(i)$ nezávisí na η narozdíl od metody zlepšení bez uvažování inflace (je to vidět z toho, že $E = 0$ ... budoucí zisk se vlivem inflace blíží k nule). Postup řešení: opět jsou zde kroky 1,2, ovšem v kroku 1 je systém m rovnic o m neznámých ($E = 0$ ... vypadlo ze systému) a ve 2.kroku volíme variantu s maximální hodnotou výrazu

$$v_i^k + \alpha \sum_{j=1}^{m} p_{ij}^k f^s(j)$$

Oba kroky opakujeme tak dlouho, až nová varianta je stejná jako ta stará.

Ad Příklad 13.1.

krok 1: počáteční politika ... $(1, 1, 1)$
řešením systému

$$\begin{aligned} f(1) - 0,6\,(0,2\,f(1) + 0,5\,f(2) + 0,3\,f(3)) &= 5,3 \\ f(2) - 0,6\,(\phantom{0,2\,f(1) +{}} 0,5\,f(2) + 0,5\,f(3)) &= 3 \\ f(3) - 0,6\,(f(3)) &= -1 \end{aligned}$$

dostaneme

$$\begin{aligned} f(1) &\approx 6,6 \\ f(2) &\approx 3,21 \\ f(3) &\approx -2,5. \end{aligned}$$

krok 2:

	$v_i^k + \alpha(p_{i1}^k f(1) + p_{i2}^k f(2) + p_{i3}^k f(3))$		
i	$k = 1$	$k = 2$	k^*
1	$5,3 + 0,6\,(0,2 \cdot 6,6 + 0,5 \cdot 3,21$ $-0,3 \cdot 2,5) = 6,6$	$4,7 + 0,6\,(0,3 \cdot 6,6 + 0,6 \cdot 3,21$ $-0,1 \cdot 2,5) = 6,89$	2
2	$3 + 0,6\,(0,5 \cdot 3,21$ $-0,5 \cdot 2,5) = 3,21$	$3,1 + 0,6\,(0,1 \cdot 6,6 + 0,6 \cdot 3,21$ $-0,3 \cdot 2,5) = 4,2$	2
3	$-1 + 0,6\,(-2,5) = -2,5$	$0,4 + 0,6\,(0,05 \cdot 6,6 + 0,4 \cdot 3,21$ $-0,55 \cdot 2,5) = 0,54$	2

tj. $s_1 = (2, 2, 2)$

krok 1': řešením systému

$$f(1) - 0,6\,(\;0,3\,f(1) + 0,6\,f(2) + \;0,1\,f(3)) = 4,7$$
$$f(2) - 0,6\,(\;0,1\,f(1) + 0,6\,f(2) + \;0,3\,f(3)) = 3,1$$
$$f(3) - 0,6\,(0,05\,f(1) + 0,4\,f(2) + 0,55\,f(3)) = 0,4$$

dostaneme

$$f(1) \approx 8,88$$
$$f(2) \approx 6,62$$
$$f(3) \approx 3,37.$$

krok 2':

i	$k=1$	$k=2$	k^*
1	$8,95$	$8,88$	1
2	$5,99$	$6,62$	2
3	$1,02$	$3,37$	2

tj. $s_2 = (1,2,2)$ … „hnojit ve stavu 2 nebo 3"

$\left.\begin{array}{l}\textbf{krok 1''}\\ \textbf{krok 2''}\end{array}\right\}$ vede k variantě $\underline{\underline{s_3 = (1,2,2) = s_2}}$.

Úloha má tedy jiné řešení než v případě, kdy jsme neuvažovali inflaci.

Díky tomu, že se zde využívá jistý optimalizační krok zlepšení, absolvent předmětu OPV by i věřil, že úlohu lze přeformulovat jako úlohu lineárního programování. A skutečně je tomu tak. Ale pro velká k, m algoritmus lineárního programování není moc rychlý, rychlejší je právě uvedená metoda zlepšení politiky.

Pojmy k zapamatování

- V poslední kapitole jsme přímo na příkladu z praxe demonstrovali problematiku pravděpodobnostního dynamického programování. Obecně lze úlohu pstního dynamického programování formulovat takto:
  Máme m možných stavů nějakého objektu a S politik (strategií), kterými můžeme ovlivňovat pravděpodobnost změn jednotlivých stavů za dané období.
- Pravděpodobnosti přechodu od jednoho stavu k jinému pro s-tou politiku jsou reprezentovány tzv. *Markovským řetězcem*, tj. čtvercovou maticí řádu m, která je stochastická (její prvky jsou nezáporné a součet řádků je 1). Tuto matici značíme $\boldsymbol{P}^s$, $s = 1, \ldots, S$. Dále také známe matice výnosů jednotlivých politik $\boldsymbol{R}^1, \ldots, \boldsymbol{R}^S$.

- Základní otázky, na které chceme odpovědět, jsou:
  1) Jakou posloupnost strategií zvolit, aby celkové výnosy za N období byly maximální?

2) Jaký výnos celkem přinese daná posloupnost strategií v nejbližších N obdobích?

– **Řešení**

a) řešení úlohy pro konečné $k = N$

Na obě otázky najdeme odpověď algoritmem dynamického programování. Jednotlivé fáze jsou v tomto případě období a přípustnými alternativami jsou naše možné strategie. Označíme

$f_n(i)$ ... optimální (očekávaný) výnos z fází $n, n+1, \dots, N$ za daného stavu i na počátku období n

$v_i^k = \sum_{j=1}^{m} p_{ij}^k r_{ij}^k$ ... výnos k-té politiky za daného stavu i.

K řešení využijeme zpětné rekurzivní rovnice

$$f_N(i) = \max_k \{v_i^k\}$$
$$f_n(i) = \max_k \left\{ v_i^k + \sum_{j=1}^{m} p_{ij}^k f_{n+1}(j) \right\} \quad \text{pro } n = 1, 2, \dots, N-1$$

Tato úloha jde ještě zobecnit např. tak, že přechodové psti mohou být každé období jiné nebo pokud chceme znát současnou hodnotu očekávaných zisků. Princip řešení je stejný, jen se částečně pozmění výpočet jednotlivých výnosů. Naopak, při vyhodnocení stacionární politiky se algoritmus zjednoduší.

b) řešení úlohy pro k nekonečné (nebo hodně velké)

1) *metoda úplného vyčíslení...* ohodnotíme všechny stacionární politiky a vybereme tu optimální (lze ji užít jen při nízkém počtu stacionárních politik).
Řešením systému

$$\boldsymbol{\pi}^s \cdot \boldsymbol{P}^s = \boldsymbol{\pi}^s$$
$$\sum_{i=1}^{m} \pi_i^s = 1$$

vypočteme vektor $\boldsymbol{\pi}^s$ dlouhodobých stacionárních pstí pro jednotlivé politiky a určíme očekávaný výnos pro politiku s:

$$E_s = \sum_{i=1}^{m} \pi_i^s v_i^s$$

Pak optimální politika je ta, pro níž je E_s maximální.

2) *metoda zlepšení politiky (+ neuvažujeme inflaci)*

krok 1: zvolíme libovolnou politiku s, tj. $\boldsymbol{P}^s, \boldsymbol{R}^s$;
předpokládejme $f^s(m) = 0$ a vyřešme systém m rovnic o m neznámých

$$E^s = v_i^s + \left(\sum_{j=1}^{m} p_{ij}^s \, f^s(j) \right) - f^s(i), \qquad i = 1, 2, \dots, m$$

krok 2: (zlepšení politiky)
pro každý stav i určíme alternativu k s maximální hodnotou výrazu

$$v_i^k + \sum_{j=1}^{m} p_{ij}^k\, f^s(j); \qquad \to \quad \text{tím se určí maximální } E \text{ pro každý stav}$$

to bude nová politika t.

Pokračujeme opakováním kroků 1,2 tak dlouho, až $s = t$.

3) *metoda zlepšení politiky (+ uvažuje inflaci)*
Předpokládáme zde, že $E = 0$, neboť budoucí zisk se vlivem inflace blíží k nule). Postup řešení: opět jsou zde kroky 1,2, ovšem v kroku 1 je systém m rovnic o m neznámých ($E = 0 \ldots$ vypadlo ze systému) a ve 2.kroku volíme variantu s maximální hodnotou výrazu

$$v_i^k + \alpha \sum_{j=1}^{m} p_{ij}^k\, f^s(j)$$

Oba kroky opakujeme tak dlouho, až nová varianta je stejná jako ta stará.

Kontrolní otázky

1. U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.
   a) Matice pravděpodobností $\boldsymbol{P}$ přechodu z jednoho stavu do druhého nemusí být obecně čtvercová.
   b) Matice zisku $\boldsymbol{R}$ musí mít jen nezáporné prvky.
   c) Při průběhu algoritmu dynamického programování hledáme maximu z celkového výnosu vzhledem ke zvolené strategii.
   d) Při řešení úloh této kapitoly pomocí algoritmu dynamického programování používáme zpětnou rekurzi.
   e) Matice pravděpodobností $\boldsymbol{P}$ přechodu z jednoho stavu do druhého se může každé období měnit.
   f) Pojem „stacionární politika“ znamená, že se všechny strategie každé období pravidelně střídají.
   g) Při hledání optimální strategie na (nekonečně) mnoho období dopředu se omezujeme pouze na stacionární politiky.
   h) Inflace neovlivňuje celkový dlouhodobý zisk.

Odpovědi na otázky
1a) – N, 1b) – N, 1c) – A, 1d) – A, 1e) – A, 1f) – N, 1g) – A, 1h) – N.

Cvičení

1. Firma každoročně kontroluje úspěšnost prodeje svého výrobku na trhu a rozhoduje, jestli je uspokojivá (stav 1) nebo není (stav 2). Na základě těchto poznatků pak rozhoduje, zda

investovat do reklamy na tento výrobek a zvýšit tak jeho prodej. Matice $\boldsymbol{P}_1$ a $\boldsymbol{P}_2$ udávají pravděpodobnosti přechodu mezi stavy s využitím a bez využití reklamy v průběhu roku. Příslušné zisky jsou dány maticemi $\boldsymbol{R}_1$ a $\boldsymbol{R}_2$. Stanovte optimální politiku rozhodování v průběhu 3 let.

$$\boldsymbol{P}_1 = \begin{pmatrix} 0,9 & 0,1 \\ 0,6 & 0,4 \end{pmatrix}, \quad \boldsymbol{R}_1 = \begin{pmatrix} 2 & -1 \\ 1 & -3 \end{pmatrix}$$

$$\boldsymbol{P}_2 = \begin{pmatrix} 0,7 & 0,3 \\ 0,2 & 0,8 \end{pmatrix}, \quad \boldsymbol{R}_2 = \begin{pmatrix} 4 & 1 \\ 2 & -1 \end{pmatrix}.$$

2. Společnost může propagovat svůj výrobek prostřednictvím reklamy ve třech médiích: v rádiu, televizi nebo v novinách. Týdenní ceny reklamy v jednotlivých médiích jsou postupně 200 Kč, 900 Kč a 300 Kč. Může také hodnotit jeho prodejnost v průběhu každého týdne jako (1) průměrnou, (2) dobrou, (3) nejlepší. Následující matice udávají psti přechodu mezi stavy pro jednotlivá média

$$\begin{array}{ccc} \text{Rádio} & \text{Televize} & \text{Noviny} \\ \begin{pmatrix} 0,4 & 0,5 & 0,1 \\ 0,1 & 0,7 & 0,2 \\ 0,1 & 0,2 & 0,7 \end{pmatrix} & \begin{pmatrix} 0,7 & 0,2 & 0,1 \\ 0,3 & 0,6 & 0,1 \\ 0,1 & 0,7 & 0,2 \end{pmatrix} & \begin{pmatrix} 0,2 & 0,5 & 0,3 \\ 0 & 0,7 & 0,3 \\ 0 & 0,2 & 0,8 \end{pmatrix} \end{array}.$$

Příslušné týdenní zisky (v tisících Kč) jsou

$$\begin{array}{ccc} \text{Rádio} & \text{Televize} & \text{Noviny} \\ \begin{pmatrix} 400 & 520 & 600 \\ 300 & 400 & 700 \\ 200 & 250 & 500 \end{pmatrix} & \begin{pmatrix} 1\,000 & 1\,300 & 1\,600 \\ 800 & 1\,000 & 1\,700 \\ 600 & 700 & 1\,100 \end{pmatrix} & \begin{pmatrix} 400 & 530 & 710 \\ 350 & 450 & 800 \\ 250 & 400 & 650 \end{pmatrix} \end{array}.$$

Stanovte optimální politiku rozhodování pro následující 3 týdny.

3. Stanovte optimální rozhodovací politiku u příkladu 1 na nekonečně mnoho let užitím metody úplného vyčíslení.

4. Stanovte optimální rozhodovací politiku u příkladu 1 na nekonečně mnoho let užitím metody zlepšení politiky.

Výsledky

ad 1. V prvním a druhém roce: investovat jen pokud je prodej výrobku neuspokojivý; ve třetím roce: neinvestovat do reklamy.

ad 2. Pokud je prodejnost výrobku průměrná, využít reklamy v rádiu; jinak využít reklamy v novinách.

ad 3. Nikdy neinvestovat do reklamy.

ad 4. Investovat do reklamy jen v případě neuspokojivého prodeje.

Literatura

[1] J. Anděl: *Matematická statistika*, Praha, SNTL, 1978

[2] J. Anděl: *Statistiské metody*, Praha, Matfyzpress, 1993

[3] L. Bican: *Lineární algebra,* SNTL 1979, rozšířené vydání 2001

[4] G. Birkhoff, T.C. Bartee: *Aplikovaná algebra,* Alfa, Bratislava 1981

[5] G. Birkhoff, S. MacLane: *Algebra,* Alfa,Bratislava 1973

[6] M. Budíková, Š. Mikoláš, P. Osecký: *Teorie pravděpodobnosti a matematická statistika*, Brno, MU – Př.fak, 1998

[7] S.C. Chapra, R.P. Canale: *Numerical Methods for Engineers.* McGraw-Hill 2002, New York (4th Edition).

[8] M. Capiński, T. Zastawniak: *Probability Through Problems*, Springer, ISBN 0-387-95063-X

[9] M. Demlová, J. Nagy: *Algebra,* MVšT —III, SNTL 1982

[10] J. Diblík, A. Haluzíková, J. Baštinec: *Numerické metody a matematická statistika.* VUT Brno, 1987 (skriptum)

[11] J. Diblík, J. Baštinec: *Matematika IV.* Nakladatelství VUT v Brně, 1991 (skriptum)

[12] N. Dudorkin: *Operační analýza,* FEL ČVUT, Praha, 1997.

[13] B. Fajmon, J. Koláček: *Pravděpodobnost, statistika a operační výzkum*, FEKT VUT, Brno 2005. (Elektronická studijní opora)

[14] L.E. Garner: *Calculus and analytic geometry*, London, 1988

[15] A. Hald: *A histori of parametric statistical Inference from Bernoulli to Fisher*, 1713 - 1935. Springer, ISBN 0-387-46408-5

[16] A. Haluzíková, V. Kudláček, B. Zástěra: *Numerické metody a matematická statistika.* SNTL Brno, 1983 (skriptum)

[17] V. Havel, J. Holenda: *Linaární algebra,* SNTL 1984

[18] Z. Horský: *Množiny a matematické struktury,* MVšT — I, SNTL 1980

[19] Z. Horský: *Vektorové prostory,* MVšT — II, SNTL 1980

[20] C.W. CHurchman, R.L. Ackoff, E.L. Arnoff: *Úvod do operačného výskumu.* ALFA, Bratislava 1968.

[21] K. Itô: *Stochastic Processes*, Springer, Berlin, New York 2004

[22] O. Kallenberg: *Foundations of Modern Probability*, Springer, ISBN 0-387-95313-2

[23] A. Laščiak a kol.: *Dynamické modely.* ALFA, Bratislava 1985.

[24] J. Likeš, J. Machek: *Počet pravděpodobnosti*, SNTL, Praha 1981

[25] J. Likeš, J. Machek: *Matematická statistika*, SNTL, Praha 1983

[26] J. Loftus, E. Loftus: *Essence of Statistics. Second Edition*, Alfred A. Knopf, New York 1988.

[27] David G. Luenberger: *Linear and Nonlinear Programing, second edition*, Kluwer Academic Publishers, ISBN 1-1020-7593-6

[28] B. Maroš: *Empirické modely I.* Skriptum FSI, nakladatelství CERM 2001, Brno.

[29] D.C. Montgomery, G.C. Runger: *Applied Statistics and Probability for Engineers.* Third Edition. John Wiley & Sons, Inc., New York 2003.

[30] I. Miller, M. Miller: *John E. Freund's Mathematical Statistics. Sixth Edition.* Prentice Hall, Inc., New Jersey 1999. Předchozí vydání publikováno pod názvem Freund, J.E.: Mathematical Statistics, Fifth Edition.

[31] K. Rektorys a kol.: *Přehled užité matematiky.* SNTL Praha

[32] Z. Riečanová a kol. *Numerické metódy a matematická štátistika.* Alfa Bratislava 1987

[33] J. Seger, R. Hindls: *Statistické metody v tržním hospodářství*, Victoria Publishing, Praha 1995, ISBN 80-7187-058-7

[34] T. Šalát: *Metrické priestory,* Alfa, Bratislava 1981

[35] M. Šikulová, Z. Karpíšek: *Matematika IV – Pravděpodobnost a matematická statistika.*VUT Brno, 1987 (skriptum)

[36] H.A. Taha: *Operations research. An Introduction. Eighth Edition*, Pearson Prentice Hall, New Jersey 2007.

[37] O. Tyc: *Operační analýza.* MZLU Brno 2002.

[38] T.H.Wonnacot, R.J.Wonnacot: *Statistika pro obchod a hospodářství.* Victoria Publishing, Praha, ISBN 80-85605-09-0

[39] J. Zapletal: *Operační analýza.* Kunovice 1995.