

MICHAL FRIEŠL

**Pravděpodobnost  
a statistika**  
H Y P E R T E X T O V Ě

Obsahuje 65 hesel, v nich 327 křížových odkazů,  
7 obrázků.

Práce na tomto projektu  
**nebyla** podporována žádným grantem.

## Pravděpodobnost a statistika hypertextově

verze k tisku, 2002-09-16\*

Michal Friesl

<http://home.zcu.cz/~friesl/>

[friesl@kma.zcu.cz](mailto:friesl@kma.zcu.cz)

© Mgr. [Michal Friesl](#), Ph.D., 2002

I když autor zpřístupňuje toto své dílo prostřednictvím internetu, **neuděluje** tím nikomu **žádná práva** k užití díla, včetně (ale nikoli výlučně) práva na rozmnožování, rozšiřování, a také zpřístupňování prostřednictvím počítačové sítě.

Užití tohoto díla pro osobní potřebu je možné **pouze** v rozsahu daném zákonem (volná užití).

## OBSAH

Obsah	3
Chybějící stránky	4
Úvod	5
Dostupné verze	5
TeXnická stránka	5
1. Pravděpodobnost	7
1.1. Náhodný jev	7
1.2. Sigma-algebra	7
1.3. Pravděpodobnostní míra	7
1.4. Klasická definice pravděpodobnosti	7
1.5. Nezávislost jevů	8
1.6. Podmíněná pravděpodobnost	8
1.7. Věta o celkové pravděpodobnosti	8
1.8. Bayesova věta	9
2. Náhodná veličina	9
2.1. Náhodná veličina	9
2.2. Rozdělení	9
2.3. Distribuční funkce	9
2.4. Diskrétní rozdělení	10
2.5. Pravděpodobnostní funkce	10
2.6. Spojité rozdělení	10
2.7. Hustota	10
2.8. Intenzita poruch	11
2.9. Střední hodnota	11
2.10. Momenty	12
2.11. Rozptyl	12
2.12. Medián	12
2.13. Kvantil	12
2.14. Poissonovo rozdělení	12
2.15. Exponenciální rozdělení	13
2.16. Normální rozdělení	13
3. Náhodný vektor	13
3.1. Náhodný vektor	13
3.2. Sdružené rozdělení	14
3.3. Marginální rozdělení	14
3.4. Diskrétní sdružené rozdělení	14
3.5. Spojité sdružené rozdělení	14
3.6. Střední hodnota vektoru	15
3.7. Kovariance	15
3.8. Varianční matice	15
3.9. Korelační koeficient	15
3.10. Nezávislé náhodné veličiny	16
3.11. Vícerozměrné normální rozdělení	16
4. Teorie odhadu	16
4.1. Náhodný výběr	16
4.2. Statistika (hodnota)	16
4.3. Odhad	16
4.4. Popisné statistiky	17
4.5. Výběrový průměr	17
4.6. Výběrový rozptyl	17
4.7. Výběrová směrodatná odchylka	17
4.8. Neustrannost	17
4.9. Konzistence	17
4.10. Střední čtvercová chyba	17
4.11. Nejlepší odhad	18
4.12. Maximálně věrohodný odhad	18
4.13. Intervalové odhady	18

4.14. Intervaly spolehlivosti v normálním rozdělení	18
5. Testování hypotéz	19
5.1. Hypotéza	19
5.2. Test	19
5.3. Kritický obor	19
5.4. Chyba 1. a 2. druhu	19
5.5. Hladina významnosti	19
5.6. Síla testu	20
5.7. p-hodnota	20
5.8. Testy střední hodnoty normálního rozdělení	20
5.9. Test o střední hodnotě	21
5.10. Test o rozptylu normálního rozdělení	21
5.11. Párový test	21
5.12. Dvouvýběrový test	22
5.13. Chí-kvadrát test dobré shody	22
5.14. Chí-kvadrát test při neznámých parametrech	23
5.15. Test nezávislosti v kontingenční tabulce	23
5.16. Test nezávislosti v normálním rozdělení	24
Seznam hesel	25
Seznam obrázků	27
Index	28

## CHYBĚJÍCÍ STRÁNKY

Ještě je třeba doplnit stránky o několika rozděleních, stránky z oblastí

- Limitní věty,
- Regrese

a několik dalších. Pak možná přidat tabulky, a určitě vše pročíst a zkontrolovat.

## ÚVOD

Snad mohu říci, že na českém internetu není nadbytek materiálů z oblasti pravděpodobnosti a statistiky a mnoho z těch, které existují, nevyužívá dostatečně hypertextových možností formátů, ve kterých jsou prezentovány (pojem *hypertextově* používám, možná trochu nepřesně ve smyslu přítomnosti křížových odkazů). *Pravděpodobnost a statistika hypertextově* je mým pokusem o vytvoření takového „pořádně hypertextového materiálu“.

Text sleduje výklad z úvodních kurzů pravděpodobnosti a statistiky, nemá nicméně nahrazovat klasický učební text. Obsahuje definice pojmů a tvrzení o nich (i když ne explicitně označené), ne už ale podrobnější vysvětlování a komentáře. Chybí také příklady (viz ovšem sbírku pseudořešených příkladů *Posbírané příklady z pravděpodobnosti a statistiky*). Text se svým uspořádáním spíše podobá jakési encyklopedii a může posloužit třeba těm, kteří si potřebují rychle připomenout některé z pojmů, vztahů, poznatků.

Materiál zveřejňuji ve chvíli, kdy práce na PROGRAMOVÁNÍ MATERIÁLU jsou hotovy, zbývá však PŘIDAT několik stránek (zeleně naznačené odkazy NEFUNGUJÍ<sup>?.?</sup>), opravit text, doplnit obrázky apod., třeba REJSTŘÍK je poněkud zmatený.

Hlášení o chybách, připomínky i náměty samozřejmě rád přijmu.

Michal Friesl  
friesl@kma.zcu.cz

Plzeň, září 2002

## DOSTUPNÉ VERZE

Materiál je v současné době přístupný ve třech variantách

- **HTML verze v podobě klikacích obrázků** určená k online prohlížení přes internet  
Stránky s „hesly“ jsou jednotlivé HTML stránky, odkazy mezi hesly směřují na tyto soubory. Prezentace v podobě obrázků s aktivními oblastmi znamená menší kvalitu zobrazení, k prohlížení ale stačí obyčejný internetový prohlížeč.
- **PDF verze pro obrazovku** zaručující kvalitní zobrazení  
Jediný PDF soubor, jehož stránky tvoří jednotlivá hesla. Odkazy mezi hesly směřují do téhož souboru, soubor (spolu s PDF prohlížečem) je tedy vše, co je k procházení textem potřeba.
- **PDF verze pro tisk** s jinou grafickou úpravou  
Jednotlivá hesla nemají v PDF souboru vyhrazeny samostatné stránky. Hypertextové přesuny v tištěné verzi musí čtenář realizovat sám dle vyznačených čísel.

Jednotlivé verze jsou přístupné z internetové adresy

<http://home.zcu.cz/~friesl/hpsb/>

## TEXNICKÁ STRÁNKA

Text jsem psal samozřejmě v  $\text{\TeX}$ . Použil jsem standardní  $\text{\LaTeX}$ ové balíčky (cstex, amslatex, color, hyperref) a trochu jsem si v  $\text{\TeX}$ u také sám zaprogramoval (50 kB).

Hlavním programem pro vznik všech TŘÍ VERZÍ byl pdf $\text{\TeX}$ . Dále jsem použil csindex pro abecední řazení.

Při výrobě HTML verze přišel navíc na řadu GhostScript pro vyrastování stránek (čistě kvůli lepšímu vzhledu, který se mi jím nepodařilo dosáhnout, ve skutečnosti ještě prošly grafickým programem).

Zvažoval jsem i variantu, kde by HTML stránky nebyly jedním velkým obrázkem, ale nerealizoval ji. Nebyl jsem si jist

výsledkem, který by dal latex2html (nehledě na to, že dosáhnout takové podoby stejně jako ostatních na počítači Pentium 133 MHz s už asi zapomenutými WinNT 3.51 z roku 1995 a nevelkým harddiskem by vyžadovalo větší úsilí), a MathML z  $\text{\TeX}$  vyrobit neumím.

Pro úplnost dodávám, že zdrojové texty jsem psal v editoru, pouštěl jsem dávkové soubory, doplňoval *vlňky* a konvertoval češtinu. Nakonec, internetové stránky posílá PHP, ale šlo by to i bez něj.

Je to vlastně jednoduché.

## 1. PRAVDĚPODOBNOST

1.1. *Náhodný jev.* Je-li dána množina  $\Omega$  (všech výsledků **náhodného pokusu**, tj. pokusu, jehož výsledek není jednoznačně určen podmínkami, za kterých je prováděn), pak **náhodným jevem** (v  $\Omega$ ) nazýváme každou podmnožinu množiny  $\Omega$ .

Jev  $\Omega$  nazýváme **jistý**, jev  $\emptyset$  **nemožný**. Jednotlivé výsledky  $\omega \in \Omega$  se nazývají **elementární jevy**.

Jev označuje určitou vlastnost výsledků pokusu. Pro jevy  $A, B \subset \Omega$  interpretujeme

- $\omega \in A$  — jev  $A$  nastal,
- $\omega \in \Omega \setminus A$  — jev  $A$  nenastal,
- $\omega \in A \cup B$  — aspoň jeden z jevů ( $A$  nebo  $B$ ) nastal,
- $\omega \in A \cap B$  — oba dva jevy ( $A$  i  $B$ ) nastaly,
- $A \subset B$  — nastává-li  $A$ , nastává i  $B$ ,
- $A \cap B = \emptyset$  — jevy  $A$  a  $B$  jsou **neslučitelné** (nenastávají současně).

Ne všechny jevy v  $\Omega$  musí být pozorovatelné a ne všem musí být možné rozumně připsat PRAVDĚPODOBNOST<sup>1,3</sup>.

◊ Viz též: SIGMA-ALGEBRA<sup>1,2</sup>, PRAVDĚPODOBNOSTNÍ MÍRA<sup>1,3</sup>, KLASICKÁ DEFINICE PRAVDĚPODOBNOSTI<sup>1,4</sup>, NÁHODNÁ VELIČINA<sup>2,1</sup>

1.2. *Sigma-algebra.* Systém  $\mathcal{A}$  podmnožin množiny  $\Omega$  se nazývá  **$\sigma$ -algebra**, pokud

- $\Omega \in \mathcal{A}$ ,
- $A \in \mathcal{A} \implies \Omega \setminus A \in \mathcal{A}$ ,
- $A_1, A_2, \dots \in \mathcal{A} \implies \bigcup A_i \in \mathcal{A}$ ,

tedy pokud v  $\mathcal{A}$  je zahrnuta množina  $\Omega$ , s každou množinou je v  $\mathcal{A}$  i její doplněk a s každými spočetně mnoha množinami i jejich sjednocení.

Můžeme si představovat, že  $\sigma$ -algebra představuje souhrn informací, které o výsledcích NÁHODNÉHO POKUSU<sup>1,1</sup> máme — ne všechny JEVI<sup>1,1</sup> v  $\Omega$  musí být totiž pozorovatelné.

1.3. *Pravděpodobnostní míra.* Reálná funkce  $P$  definovaná na  $\sigma$ -ALGEBŘE<sup>1,2</sup> (dostatečně bohatém systému)  $\mathcal{A}$  JEVI<sup>1,1</sup> v  $\Omega$  se nazývá **pravděpodobnostní míra**, pokud

- $P$  má hodnoty v  $\langle 0, 1 \rangle$ ,
- $P(\Omega) = 1$ ,
- Pro disjunkttní množiny (neslučitelné jevy)  $A_i \in \mathcal{A}$ ,  $P(\bigcup A_i) = \sum P(A_i)$ .

Trojice  $(\Omega, \mathcal{A}, P)$  se nazývá **pravděpodobnostní prostor**.

Z těchto vlastností lze odvodit, že  $P(\Omega \setminus A) = 1 - P(A)$ , dále  $P(\emptyset) = 0$ . Pro (nedisjunkttní) množiny lze psát

$$P(A \cup B) = P(A) + P(B) - P(A \cap B),$$

pro disjunkttní poslední člen vypadává. Když  $B \subset A$ , máme  $P(A \setminus B) = P(A) - P(B)$ .

◊ Viz též: KLASICKÁ DEFINICE PRAVDĚPODOBNOSTI<sup>1,4</sup>, ROZDĚLENÍ<sup>2,2</sup>

1.4. *Klasická definice pravděpodobnosti.* Je-li množina výsledků pokusu  $\Omega = \{\omega, \dots, \omega_n\}$  konečná (a neprázdná) a všechny elementární jevy  $\omega_1, \dots, \omega_n$  mají stejnou pravděpodobnost, že nastanou (rovnou  $1/n$ ), můžeme pro všechny JEVI<sup>1,1</sup>  $A \subset \Omega$  použít

**klasickou definici pravděpodobnosti**

$$P(A) = \frac{|A|}{|\Omega|}.$$

Ta splňuje požadavky kladené na PRAVDĚPODOBNOSTNÍ MÍRU<sup>1.3</sup>, za  $\sigma$ -ALGEBRU<sup>1.2</sup> jevů v tomto případě můžeme vzít  $\sigma$ -algebru všech podmnožin  $\Omega$ .

S klasickou definicí ovšem dlouho nevystačíme, protože se zajímáme i o pravděpodobnosti jevů, které vyplývají z pokusů se spočetně (či dokonce nespočetně) mnoha možnými výsledky.

- 1.5. *Nezávislost jevů.* Řekneme, že systém JEVŮ<sup>1.1</sup>  $\mathcal{C}$  je **nezávislý**, jestliže pro každý konečný počet  $k$  jevů  $A_1, \dots, A_k \in \mathcal{C}$  je

$$P\left(\bigcap_i A_i\right) = \prod_i P(A_i).$$

V opačném případě jevy nazýváme závislé.

Dva jevy  $A$  a  $B$  jsou tedy nezávislé, pokud  $P(A \cap B) = P(A)P(B)$ .

Nahradíme-li v  $\mathcal{C}$  některé jevy jejich doplňky, nezávislost (závislost) zůstává zachována.

Jev  $A$  s  $P(A) = 0$  nebo  $P(A) = 1$  je nezávislý s kterýmkoli jiným jevem.

- ◊ Viz též: NEZÁVISLÉ NÁHODNÉ VELIČINY<sup>3.10</sup>, PODMÍNĚNÁ PRAVDĚPODOBNOST<sup>1.6</sup>

- 1.6. *Podmíněná pravděpodobnost.* **Podmíněná pravděpodobnost** jevu  $A$  za podmínky, že nastal jev  $B$  (s  $P(B) > 0$ ) je definována jako

$$P_B(A) = P(A | B) = \frac{P(A \cap B)}{P(B)}.$$

Podmíněná pravděpodobnost  $P_B$  je opět PRAVDĚPODOBNOSTNÍ MÍROU<sup>1.3</sup>.

Pro jevy  $A, B$  s kladnými pravděpodobnostmi je ekvivalentní

- $A, B$  jsou NEZÁVISLÉ<sup>1.5</sup>
- $P_B(A) = P(A)$
- $P_A(B) = P(B)$ .

- ◊ Viz též: BAYESOVA VĚTA<sup>1.8</sup>, VĚTA O CELKOVÉ PRAVDĚPODOBNOSTI<sup>1.7</sup>

- 1.7. *Věta o celkové pravděpodobnosti.* Tvoří-li  $H_1, H_2, \dots$  rozklad pravděpodobnostního prostoru (navzájem disjunktní,  $\sum_i P(H_i) = 1$ ,  $P(H_i) > 0$ ), pak pro každý jev  $A$

$$P(A) = \sum_i P(A | H_i) P(H_i).$$

Známe-li (PODMÍNĚNÉ<sup>1.6</sup>) pravděpodobnosti jevu  $A$  za hypotéz  $H_i$  a známe-li pravděpodobnosti, s jakými tyto hypotézy nastávají, umíme spočítat celkovou pravděpodobnost jevu  $A$ .

- ◊ Viz též: BAYESOVA VĚTA<sup>1.8</sup>, PODMÍNĚNÁ PRAVDĚPODOBNOST<sup>1.6</sup>

- 1.8. *Bayesova věta.* Tvoří-li  $H_1, H_2, \dots$  rozklad pravděpodobnostního prostoru (navzájem disjunktní,  $\sum_i P(H_i) = 1$ ,  $P(H_i) > 0$ ), a  $A$  je jev s  $P(A) > 0$ , pak pro každé  $k$

$$P(H_k | A) = \frac{P(A | H_k) P(H_k)}{\sum_i P(A | H_i) P(H_i)}.$$

Nastal-li jev  $A$ , můžeme na základě této informace opravit (aktualizovat) apriorní pravděpodobnosti  $P(H_k)$  na aposteriorní  $P(H_k | A)$ .

Ve jmenovateli zlomku je CELKOVÁ PRAVDĚPODOBNOST<sup>1.7</sup> jevu  $A$ .

- ◊ Viz též: PODMÍNĚNÁ PRAVDĚPODOBNOST<sup>1.6</sup>, VĚTA O CELKOVÉ PRAVDĚPODOBNOSTI<sup>1.7</sup>

## 2. NÁHODNÁ VELIČINA

- 2.1. *Náhodná veličina.* Funkce  $X : \Omega \mapsto \mathbf{R}$  definovaná na prostoru  $\Omega$  se  $\sigma$ -ALGEBROU<sup>1.2</sup>  $\mathcal{A}$  taková, že  $\{\omega; X(\omega) \in (a, b)\} \in \mathcal{A}$  pro libovolné  $-\infty \leq a < b \leq \infty$ , se nazývá (reálná) **náhodná veličina**.

Požadavek, aby vzory intervalů při zobrazení  $X$  ležely v  $\mathcal{A}$  (jde o měřitelnost funkce  $X$ ), je jen technický požadavek zajišťující, že vzorům rozumných množin z  $\mathbf{R}$  bude možné přiřadit PRAVDĚPODOBNOST<sup>1.3</sup>.

Používáme označení

$$[X \in B] = \{\omega \in \Omega; X(\omega) \in B\} = X^{-1}(B)$$

pro vzor množiny  $B \subset \mathbf{R}$  při zobrazení  $X$ .

- 2.2. *Rozdělení. Rozdělením* NÁHODNÉ VELIČINY<sup>2.1</sup>  $X$  definované na pravděpodobnostním prostoru  $(\Omega, \mathcal{A}, P)$  nazýváme PRAVDĚPODOBNOSTNÍ MÍRU<sup>1.3</sup>  $Q$

$$Q(B) = P[X \in B], \quad B \in \mathcal{B}(\mathbf{R}),$$

působící na vhodném systému podmnožin  $\mathbf{R}$ , tzv. Borelovské  $\sigma$ -ALGEBŘE<sup>1.2</sup>  $\mathcal{B}(\mathbf{R})$ . Jde o nejmenší  $\sigma$ -algebru v  $\mathbf{R}$  obsahující všechny intervaly.

Náhodná veličina  $X$  přenáší pravděpodobnost  $P$  z  $\Omega$  na  $\mathbf{R}$ ,  $Q$  je „obrazem“ míry  $P$ . Ke každé pravděpodobnostní míře  $Q$  na  $\mathbf{R}$  existuje náhodná veličina, jejímž je rozdělením.

Rozdělení popisuje pravděpodobnostní chování náhodné veličiny (při dané pravděpodobnostní míře  $P$ ), nikoliv veličinu samotnou. Můžeme tak mít dvě různé veličiny, na stejném nebo na různých pravděpodobnostních prostorech, které mají stejné rozdělení. Změníme-li  $P$ , může se rozdělení veličiny  $X$  změnit.

- ◊ Viz též: SPOJITÉ ROZDĚLENÍ<sup>2.6</sup>, DISKRÉTNÍ ROZDĚLENÍ<sup>2.4</sup>

- 2.3. *Distribuční funkce. Distribuční funkcí* ROZDĚLENÍ<sup>2.2</sup>  $Q$  (resp. VELIČINY<sup>2.1</sup>  $X$  s tímto rozdělením) rozumíme funkci

$$F(x) = Q((-\infty, x]) = P[X \leq x], \quad x \in \mathbf{R}.$$

Pomocí distribuční funkce můžeme zapisovat PRAVDĚPODOBNOST<sup>1.3</sup>, že hodnoty veličiny  $X$  padnou do intervalu  $(a, b) \subset \mathbf{R}$ , resp. že  $X$  nabude právě hodnoty  $a \in \mathbf{R}$ , jako

$$P[X \in (a, b)] = F(b) - F(a),$$

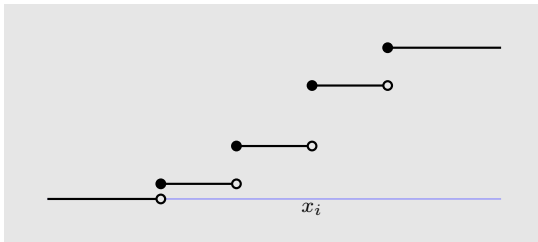
$$\text{resp. } P[X = a] = F(a) - F(a-),$$

kde  $F(a-)$  značí limitu  $F$  v bodě  $a$  zleva.

Distribuční funkce má hodnoty v  $(0, 1)$ , je neklesající, zprava spojitá s limitami  $F(-\infty) = 0$  a  $F(\infty) = 1$ . Každé funkci na

$\mathbf{R}$  s těmito vlastnostmi odpovídá nějaké rozdělení na  $\mathbf{R}$  (a existuje tedy i náhodná veličina), jehož je distribuční funkce.

◊ Viz též: PRAVDĚPODOBNOSTNÍ FUNKCE<sup>2.5</sup>, HUSTOTA<sup>2.7</sup>, INTENZITA PORUCH<sup>2.8</sup>



Distribuční funkce diskrétního rozdělení

4. **Diskrétní rozdělení.** NÁHODNÁ VELIČINA<sup>2.1</sup>  $X$  má **diskrétní rozdělení**, existuje-li spočetná množina  $M = \{x_1, x_2, \dots\} \subset \mathbf{R}$  (nosič rozdělení) taková, že

$$P[X \in M] = 1.$$

DISTRIBUČNÍ FUNKCE<sup>2.3</sup> takového rozdělení je schodovitá, se skoky v bodech  $x_i, i = 1, 2, \dots$

◊ Viz též: PRAVDĚPODOBNOSTNÍ FUNKCE<sup>2.5</sup>

2.5. **Pravděpodobnostní funkce.** Pravděpodobnostní funkcí rozumíme nezápornou funkci nulovou všude s výjimkou spočetně bodů  $x_1, x_2, \dots \in \mathbf{R}$ , v nichž nabývá postupně hodnot  $p_1, p_2, \dots \in (0, 1)$ , přičemž  $\sum p_i = 1$ .

Je-li dáno DISKRÉTNÍHO ROZDĚLENÍ<sup>2.4</sup>  $Q$ , jeho pravděpodobnostní funkcí rozumíme funkci

$$P(x) = Q(\{x\}) = P[F(x) - F(x-)] \\ = \begin{cases} p_i & x = x_i, \quad i = 1, 2, \dots, \\ 0 & x \neq x_i, \end{cases}$$

kde  $F$  je DISTRIBUČNÍ FUNKCE<sup>2.3</sup>  $Q$ ,  $x_i$  body skoků  $F$  a  $p_i$  velikosti skoků v těchto bodech.

Pro NÁHODNOU VELIČINU<sup>2.1</sup>  $X$  s takovým rozdělením je

$$P[X = x_i] = p_i, \quad i = 1, 2, \dots$$

$$P[X \in B] = \sum_{i, x_i \in B} p_i, \quad B \subset \mathbf{R}.$$

Každou pravděpodobnostní funkcí (tj. dvojicemi  $(x_i, p_i), i = 1, 2, \dots$ , kde  $x_i \in \mathbf{R}, p_i \in (0, 1), \sum p_i = 1$ ) je popsáno nějaké diskrétní rozdělení. Pro jeho distribuční funkci platí

$$F(x) = \sum_{i, x_i \leq x} p_i, \quad x \in \mathbf{R}.$$

◊ Viz též: HUSTOTA<sup>2.7</sup>

2.6. **Spojité rozdělení.** NÁHODNÁ VELIČINA<sup>2.1</sup>  $X$  s DISTRIBUČNÍ FUNKCÍ<sup>2.3</sup>  $F$  má **spojité rozdělení**, existuje-li funkce  $f(x)$  taková, že

$$F(x) = P[X \leq x] = \int_{-\infty}^x f(t) dt, \quad x \in \mathbf{R}.$$

Distribuční funkce takového rozdělení je (v řeči analýzy) absolutně spojitá.

◊ Viz též: HUSTOTA<sup>2.7</sup>, DISKRÉTNÍ ROZDĚLENÍ<sup>2.4</sup>, SPOJITÉ SDRUŽENÉ ROZDĚLENÍ<sup>3.5</sup>

2.7. **Hustota.** Jako **hustotou** označujeme každou funkci  $f(x)$  s vlastnostmi

$$f(x) \geq 0, \quad x \in \mathbf{R}, \quad \int_{\mathbf{R}} f(x) dx = 1.$$

Funkci  $f$  nazýváme hustotou SPOJITÉHO ROZDĚLENÍ<sup>2.6</sup> s DISTRIBUČNÍ FUNKCÍ<sup>2.3</sup>  $F$ , pokud

$$F(x) = \int_{-\infty}^x f(t) dt, \quad x \in \mathbf{R}.$$

Taková  $f$  má vlastnosti hustoty a je  $f(x) = F'(x)$  pro „skoro všechna“  $x \in \mathbf{R}$  (např. v bodech, kde  $F'$  neexistuje, můžeme  $f$  volit libovolně).

Pro VELIČINU<sup>2.1</sup>  $X$  s hustotou  $f$  platí

$$P[X = x] = 0, \quad x \in \mathbf{R},$$

a je-li  $I$  interval s krajními body  $a < b$  (ať už uzavřený, polouzavřený, nebo otevřený), pak

$$P[X \in I] = F(b) - F(a) = \int_a^b f(x) dx.$$

Každá hustota  $f$  je hustotou nějakého ROZDĚLENÍ<sup>2.2</sup>, příslušná distribuční funkce má tvar

$$F(x) = \int_{-\infty}^x f(t) dt, \quad x \in \mathbf{R}.$$

◊ Viz též: PRAVDĚPODOBNOSTNÍ FUNKCE<sup>2.5</sup>, SPOJITÉ SDRUŽENÉ ROZDĚLENÍ<sup>3.5</sup>

2.8. *Intenzita poruch.* Pro nezápornou NÁHODNOU VELIČINU<sup>2.1</sup>  $X$  se SPOJITÝM ROZDĚLENÍM<sup>2.6</sup> definujeme **intenzitu poruch**

$$\lambda(x) = \lim_{h \rightarrow 0^+} \frac{P[X \in (x, x+h) \mid X > x]}{h} = \frac{f(x)}{1 - F(x)}$$

tam, kde  $F(x) < 1$ , přičemž  $f$  a  $F$  jsou HUSTOTA<sup>2.7</sup> a DISTRIBUČNÍ FUNKCE<sup>2.3</sup>  $X$ .

Představuje-li  $X$  dobu do poruchy nějakého zařízení, intenzita  $\lambda(x)$  vyjadřuje, že pokud do času  $x$  nedošlo k poruše, tak PRAVDĚPODOBNOST<sup>1.3</sup>, že k ní dojde v následujícím okamžiku malé délky  $h$ , je přibližně  $\lambda(x)h$ .

Intenzita charakterizuje ROZDĚLENÍ<sup>2.2</sup> nezáporné náhodné veličiny, lze psát

$$F(x) = 1 - \exp\left(-\int_0^x \lambda(t) dt\right), \quad x > 0.$$

2.9. *Střední hodnota.* **Střední hodnota** NÁHODNÉ VELIČINY<sup>2.1</sup>  $X$ , resp.  $g(X)$ , kde  $g: \mathbf{R} \rightarrow \mathbf{R}$  je (měřitelná) funkce, je pro  $X$  DISKRÉTNĚ ROZDĚLENOU<sup>2.4</sup> s PRAVDĚPODOBNOSTNÍ FUNKCÍ<sup>2.5</sup>  $P$  definována jako

$$E X = \sum_i x_i P(x_i),$$

$$\text{resp. } E g(X) = \sum_i g(x_i) P(x_i),$$

a pro  $X$  SPOJITĚ ROZDĚLENOU<sup>2.6</sup> s HUSTOTOU<sup>2.7</sup>  $f$  jako

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

$$\text{resp. } E g(X) = \int_{-\infty}^{\infty} g(x) f(x) dx,$$

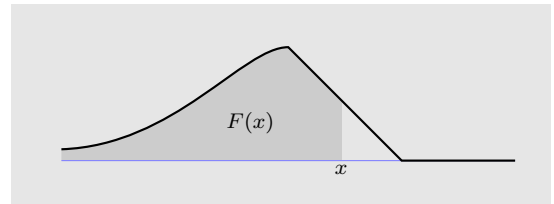
mají-li výrazy napravo smysl.

Střední hodnota  $E X$  představuje průměrnou hodnotu (ve smyslu váženého průměru), které veličina  $X$  nabývá.

Střední hodnota je lineární operátor. Jsou-li  $X, Y$  náhodné veličiny definované na stejném PRAVDĚPODOBNOSTNÍM PROSTORU<sup>1.3</sup>  $\Omega$  a  $a, b \in \mathbf{R}$ , platí

$$E a = a, \quad E(aX + bY) = a E X + b E Y.$$

Operátor  $E$  také zachovává monotonii. Je-li  $X \leq Y$  (tj.  $X(\omega) \leq Y(\omega) \forall \omega \in \Omega$ ), pak  $E X \leq E Y$ .



Hustota rozdělení

◊ Viz též: MOMENTY<sup>2.10</sup>, ROZPTYL<sup>2.11</sup>, STŘEDNÍ HODNOTA VEKTORU<sup>3.6</sup>

2.10. *Momenty.* **Momenty** ROZDĚLENÍ<sup>2.2</sup> jsou STŘEDNÍ HODNOTY<sup>2.9</sup> mocnin příslušné NÁHODNÉ VELIČINY<sup>2.1</sup>.

$E X^k$  je  $k$ -tý (necentrální) moment rozdělení veličiny  $X$ ,  $E(X - E X)^k$  je  $k$ -tý centrální moment ( $k \in \mathbf{N}$ ).

◊ Viz též: ROZPTYL<sup>2.11</sup>

2.11. *Rozptyl.* **Rozptyl** (ROZDĚLENÍ<sup>2.2</sup>) VELIČINY<sup>2.1</sup>  $X$  je definován jako druhý CENTRÁLNÍ MOMENT<sup>2.10</sup> rozdělení  $X$ ,

$$\text{var } X = E(X - E X)^2 = E X^2 - (E X)^2,$$

pokud je  $E X^2$  konečná.

Pro  $c \in \mathbf{R}$  platí

$$\text{var } c = 0, \quad \text{var } cX = c^2 \text{var } X.$$

Rozptyl vyjadřuje variabilitu rozdělení kolem jeho STŘEDNÍ HODNOTY<sup>2.9</sup>. Je-li  $E X^2 < \infty$ , Čebyševova nerovnost odhaduje (velice hrubě) PRAVDĚPODOBNOST<sup>1.3</sup> odchylky  $X$  od její střední hodnoty

$$P[|X - E X| \geq \varepsilon] \leq \frac{\text{var } X}{\varepsilon^2}, \quad \varepsilon > 0.$$

◊ Viz též: VARIANČNÍ MATICE<sup>3.8</sup>

2.12. *Medián.* **Medián** ROZDĚLENÍ<sup>2.2</sup> veličiny  $X$  je libovolné číslo  $\tilde{x} \in \mathbf{R}$  splňující

$$P[X < \tilde{x}] \leq 1/2 \leq P[X \leq \tilde{x}].$$

Jde o jakousi „prostřední“ hodnotu rozdělení, podobně jako STŘEDNÍ HODNOTA<sup>2.9</sup> je měřítkem polohy rozdělení.

Pro SPOJITĚ ROZDĚLENOU<sup>2.6</sup> veličinu s DISTRIBUTUČNÍ FUNKCÍ<sup>2.3</sup>  $F$  je mediánem  $\tilde{x}$  číslo splňující  $F(\tilde{x}) = 1/2$ .

◊ Viz též: KVANTIL<sup>2.13</sup>

2.13. *Kvantil.* Pro  $p \in (0, 1)$  je  $p$ -**kvantil**  $x_p$  ROZDĚLENÍ<sup>2.2</sup> s DISTRIBUTUČNÍ FUNKCÍ<sup>2.3</sup>  $F$  definován jako číslo splňující

$$F(x_{p-}) \leq p \leq F(x_p).$$

Je-li  $F$  spojitá,  $p$ -kvantilem jsou hodnoty  $F^{-1}(p)$ .

Pro  $p = 1/2$ , resp.  $p = 1/4, 3/4$  hovoříme o MEDIÁNU<sup>2.12</sup>, resp. dolním a horním kvantilu.

2.14. *Poissonovo rozdělení.* **Poissonovo rozdělení** s parametrem  $\lambda > 0$ , značení např.  $Po(\lambda)$ , je DISKRÉTNÍ ROZDĚLENÍ<sup>2.4</sup> na množině celých nezáporných čísel s PRAVDĚPODOBNOSTNÍ FUNKCÍ<sup>2.5</sup>

$$P(k) = \begin{cases} e^{-\lambda} \lambda^k / k!, & k = 0, 1, 2, \dots, \\ 0 & \text{jinak.} \end{cases}$$

NÁHODNÁ VELIČINA<sup>2.1</sup> s Poissonovým rozdělením,  $X \sim Po(\lambda)$ , má STŘEDNÍ HODNOTU<sup>2.9</sup> a ROZPTYL<sup>2.11</sup> rovny parametru rozdělení,

$$E X = \text{var } X = \lambda,$$

hodnoty DISTRIBUTUČNÍ FUNKCE<sup>2.3</sup> bývají tabelovány.

◊ Viz též: <CLVPO><sup>?.?</sup>

- 2.15. *Exponenciální rozdělení.* **Exponenciální rozdělení** s parametrem  $\lambda > 0$ , značení např.  $\text{Exp}(\lambda)$ , je **SPOJITÉ ROZDĚLENÍ**<sup>2.6</sup> na množině kladných čísel s **DISTRIBUČNÍ FUNKCÍ**<sup>2.3</sup>

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

tj. s **HUSTOTOU**<sup>2.7</sup>

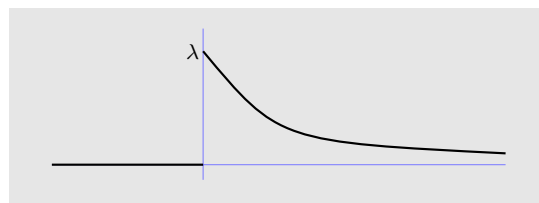
$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0 & \text{jinak.} \end{cases}$$

Toto rozdělení má konstantní **INTENZITU PORUCH**<sup>2.8</sup>,  $\lambda(x) \equiv \lambda$ ,  $x > 0$ .

**NÁHODNÁ VELIČINA**<sup>2.1</sup> s exponenciálním rozdělením,  $X \sim \text{Exp}(\lambda)$ , má **STŘEDNÍ HODNOTU**<sup>2.9</sup> a **ROZPTYL**<sup>2.11</sup>

$$E X = 1/\lambda, \quad \text{var } X = 1/\lambda^2.$$

Někdy se jako parametr exponenciálního rozdělení udává nikoli intenzita poruch  $\lambda$ , ale její převrácená hodnota, tj. střední hodnota rozdělení  $\theta = 1/\lambda$ .



Hustota exponenciálního rozdělení

- 2.16. *Normální rozdělení.* **Normální (Gaussovo) rozdělení** se **STŘEDNÍ HODNOTOU**<sup>2.9</sup>  $\mu \in \mathbf{R}$  a **ROZPTYLEM**<sup>2.11</sup>  $\sigma^2$  ( $\sigma > 0$  je **SMĚRODATNÁ ODCHYLKA**<sup>2.2</sup>), značení  $N(\mu, \sigma^2)$ , je **SPOJITÉ ROZDĚLENÍ**<sup>2.6</sup> s **HUSTOTOU**<sup>2.7</sup>

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}, \quad x \in \mathbf{R}.$$

Normální rozdělení s nulovou střední hodnotou a jednotkovým rozptylem,  $N(0, 1)$ , se nazývá **standardní (normované) normální**, jeho hustota se značí

$$\varphi(x) = (1/\sqrt{2\pi})e^{-x^2/2}, \quad x \in \mathbf{R}.$$

**DISTRIBUČNÍ**<sup>2.3</sup> a **KVANTILOVÁ**<sup>2.13</sup> funkce rozdělení  $N(0, 1)$ ,

$$\Phi(x) = \int_{-\infty}^x \varphi(t) dt, \quad u_p = \Phi_{-1}(p), \quad p \in (0, 1),$$

bývají tabelovány, vzhledem k symetrii  $N(0, 1)$  platí

$$\Phi(-x) = 1 - \Phi(x), \quad u_{1-p} = -u_p.$$

Při lineární transformaci zůstává normalita rozdělení zachována. Je-li  $X \sim N(\mu, \sigma^2)$  a  $a, b \in \mathbf{R}$ , pak  $aX + b \sim N(a\mu + b, b^2\sigma^2)$ . Pro distribuční funkci  $F_X$  **VELIČINY**<sup>2.1</sup>  $X \sim N(\mu, \sigma^2)$  můžeme tedy psát

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right), \quad x \in \mathbf{R}.$$

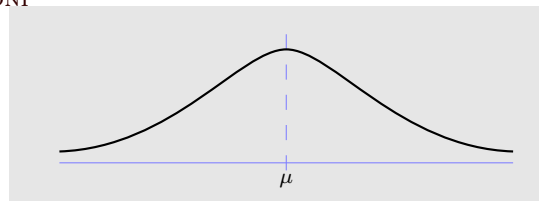
◊ Viz též: **VÍCEROZMĚRNÉ NORMÁLNÍ ROZDĚLENÍ**<sup>3.11</sup>

### 3. NÁHODNÝ VEKTOR

- 3.1. *Náhodný vektor.* **Náhodným vektorem** rozumíme vektor složený z **NÁHODNÝCH VELIČIN**<sup>2.1</sup> definovaných na stejném **PRAVDĚPODOBNOTNÍM PROSTORU**<sup>1.3</sup>,  $\mathbf{X} = (X_1, \dots, X_n)$ .

Náhodný vektor  $\mathbf{X}$  můžeme chápat také jako náhodnou veličinu s hodnotami v  $\mathbf{R}^n$ .

◊ Viz též: **SDRUŽENÉ ROZDĚLENÍ**<sup>3.2</sup>



Hustota normálního rozdělení

3.2. *Sdružené rozdělení. Sdružené rozdělení* NÁHODNÉHO VEKTORU<sup>3.1</sup>  $\mathbf{X} = (X_1, \dots, X_n)$  s hodnotami v  $\mathbf{R}^n$  je charakterizováno sdruženou distribuční funkcí

$$F(\mathbf{x}) = P[X_1 \leq x_1, \dots, X_n \leq x_n],$$

$$\mathbf{x} = (x_1, \dots, x_n) \in \mathbf{R}^n.$$

◊ Viz též: MARGINÁLNÍ ROZDĚLENÍ<sup>3.3</sup>, DISKRÉTNÍ SDRUŽENÉ ROZDĚLENÍ<sup>3.4</sup>, SPOJITÉ SDRUŽENÉ ROZDĚLENÍ<sup>3.5</sup>

3.3. *Marginální rozdělení. Marginálním rozdělením* NÁHODNÉHO VEKTORU<sup>3.1</sup>  $\mathbf{X} = (X_1, \dots, X_n)$  rozumíme ROZDĚLENÍ<sup>2.2</sup> jeho složek.

Má-li vektor  $\mathbf{X}$  SDRUŽENOU DISTRIBUTUČNÍ FUNKCÍ<sup>3.2</sup>  $F(x_1, \dots, x_n)$ , pak marginální distribuční funkce  $j$ -té složky  $X_j$  je

$$F_j(x_j) = P[X_j \leq x_j]$$

$$= \lim_{x_i \rightarrow \infty, i \neq j} F(x_1, \dots, x_n), \quad x_j \in \mathbf{R},$$

tj. v  $F(x_1, \dots, x_n)$  všechny ostatní složky kromě  $j$ -té posíláme k  $\infty$ .

Obecněji, marginální rozdělení podvektoru  $(X_{i_1}, \dots, X_{i_k})$ ,  $i_1, \dots, i_k \in \{1, \dots, n\}$ ,  $k < n$ , má distribuční funkci

$$F_{i_1, \dots, i_k}(x_{i_1}, \dots, x_{i_k}) = P[X_{i_1} \leq x_{i_1}, \dots, X_{i_k} \leq x_{i_k}]$$

$$= \lim_{x_i \rightarrow \infty, i \neq i_1, \dots, i_k} F(x_1, \dots, x_n), \quad x_{i_1}, \dots, x_{i_k} \in \mathbf{R}.$$

◊ Viz též: NEZÁVISLÉ NÁHODNÉ VELIČINY<sup>3.10</sup>

3.4. *Diskrétní sdružené rozdělení.* NÁHODNÝ VEKTOR<sup>3.1</sup>  $\mathbf{X}$  má **diskrétní rozdělení**, pokud existuje spočetná množina bodů  $M = \{\mathbf{x}_1, \mathbf{x}_2, \dots\}$  taková, že  $P[\mathbf{X} \in M] = 1$ .

Jako pravděpodobnostní funkci takového vektoru označujeme funkci

$$P(\mathbf{x}) = \begin{cases} P[\mathbf{X} = \mathbf{x}], & \mathbf{x} \in M \\ 0, & \mathbf{x} \notin M. \end{cases}$$

MARGINÁLNÍ ROZDĚLENÍ<sup>3.3</sup> diskrétně rozděleného vektoru  $\mathbf{X}$  jsou DISKRÉTNÍ<sup>2.4</sup>, PRAVDĚPODOBNOSTNÍ FUNKCE<sup>2.5</sup>  $j$ -té složky je

$$P_j(x) = P[X_j = x] = \sum_{\mathbf{x} \in M; x_j = x} \dots \sum P(\mathbf{x}), \quad x \in \mathbf{R},$$

tj. při daném  $x$  sčítáme sdruženou pravděpodobnostní funkci přes všechny hodnoty ostatních složek.

◊ Viz též: SPOJITÉ SDRUŽENÉ ROZDĚLENÍ<sup>3.5</sup>

3.5. *Spojité sdružené rozdělení.* NÁHODNÝ VEKTOR<sup>3.1</sup>  $\mathbf{X} = (X_1, \dots, X_n)$  se SDRUŽENOU<sup>3.2</sup> distribuční funkcí  $F(x_1, \dots, x_n)$  má **spojité rozdělení**, pokud existuje nezáporná funkce  $n$  proměnných  $f$ , že

$$F(x_1, \dots, x_n)$$

$$= \int_{(-\infty, x_1) \times \dots \times (-\infty, x_n)} \dots \int f(t_1, \dots, t_n) dt_1 \dots dt_n$$

pro všechna  $(x_1, \dots, x_n) \in \mathbf{R}$ . Tato funkce se nazývá sdružená hustota vektoru  $\mathbf{X}$ .

MARGINÁLNÍ ROZDĚLENÍ<sup>3.3</sup> spojitě rozděleného vektoru  $\mathbf{X}$  jsou SPOJITÁ<sup>2.6</sup>, složka  $X_j$  má HUSTOTU<sup>2.7</sup>

$$f_j(x_j) = \int_{\mathbf{R}^{n-1}} \cdots \int f(x_1, \dots, x_n) \times dx_1 \dots dx_{j-1} dx_{j+1} \dots dx_n,$$

tj. pro dané  $x_j$  integrujeme hustotu sdruženého rozdělení přes všechny hodnoty ostatních složek.

◊ Viz též: DISKRÉTNÍ SDRUŽENÉ ROZDĚLENÍ<sup>3.4</sup>

3.6. **Střední hodnota vektoru.** **Střední hodnotou** NÁHODNÉHO VEKTORU<sup>3.1</sup>  $\mathbf{X} = (X_1, \dots, X_n)$  rozumíme vektor STŘEDNÍCH HODNOT<sup>2.9</sup> jeho složek

$$E \mathbf{X} = (E X_1, \dots, E X_n).$$

Střední hodnotu veličiny  $g(\mathbf{X})$ , kde  $g : \mathbf{R}^n \mapsto \mathbf{R}$ , lze pro VEKTOR S DISKRÉTNÍM ROZDĚLENÍM<sup>3.4</sup> počítat jako

$$E g(\mathbf{X}) = \sum_{\mathbf{x} \in M} g(\mathbf{x}) P(\mathbf{x})$$

a pro vektor SE SPOJITÝM ROZDĚLENÍM<sup>3.5</sup> jako

$$E g(\mathbf{X}) = \int_{\mathbf{R}^n} \cdots \int g(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}.$$

3.7. **Kovariance.** **Kovariance** mezi složkami  $X_i$  a  $X_j$  NÁHODNÉHO VEKTORU<sup>3.1</sup>  $\mathbf{X}$  je

$$\begin{aligned} \text{cov}(X_i, X_j) &= E(X_i - E X_i)(X_j - E X_j) \\ &= E X_i X_j - E X_i E X_j. \end{aligned}$$

Je-li  $i = j$ , dostáváme  $\text{cov}(X_i, X_i) = \text{var } X_i$ , tedy ROZPTYL<sup>2.11</sup>.

S pomocí kovariance můžeme psát pro rozptyl lineární kombinace VELIČIN<sup>2.1</sup>  $X$  a  $Y$  s konstantami  $a, b \in \mathbf{R}$

$$\begin{aligned} \text{var}(aX + bY) \\ = a^2 \text{var } X + 2ab \text{cov}(X, Y) + b^2 \text{var } Y. \end{aligned}$$

◊ Viz též: VARIANČNÍ MATICE<sup>3.8</sup>

3.8. **Varianční matice.** **Varianční matice** NÁHODNÉHO VEKTORU<sup>3.1</sup>  $\mathbf{X} = (X_1, \dots, X_n)$  je matice složená z KOVARIANCÍ<sup>3.7</sup> mezi jeho složkami,

$$\text{var } \mathbf{X} = (\text{cov}(X_i, X_j))_{i,j=1,\dots,n}.$$

Varianční matice je pozitivně semidefinitní.

◊ Viz též: ROZPTYL<sup>2.11</sup>

3.9. **Korelační koeficient.** **Korelační koeficient** mezi složkami  $X_i$  a  $X_j$  NÁHODNÉHO VEKTORU<sup>3.1</sup>  $\mathbf{X}$  je definován jako

$$\rho_{ij} = \frac{\text{cov}(X_i, X_j)}{\sqrt{\text{var } X_i \text{var } X_j}}.$$

Tato hodnota z intervalu  $\langle -1, 1 \rangle$  je mírou lineární závislosti mezi  $X_i$  a  $X_j$ .

Při  $\rho = 0$  mluvíme o nekorelovaných veličinách, krajní hodnoty  $\rho_{ij} = \pm 1$  znamenají lineární závislost mezi  $X_i$  a  $X_j$ . Při  $\rho_{ij} = \pm 1$  totiž existují konstanty  $a, b \in \mathbf{R}$  (znaménko  $b$  je přitom stejné jako u  $\rho_{ij}$ ), že s pravděpodobností 1 je  $X_j = a + bX_i$ .

- 3.10. *Nezávislé náhodné veličiny.* NÁHODNÉ VELIČINY<sup>2.1</sup>  $X_1, X_2, \dots$ , definované na stejném PRAVDĚPODOBNOSTNÍM PROSTORU<sup>1.3</sup>, se nazývají **nezávislé**, jestliže pro každé  $n \in \mathbf{N}$  je

$$F_{(X_1, \dots, X_n)}(x_1, \dots, x_n) = F_1(x_1) \cdots F_n(x_n)$$

pro všechna  $(x_1, \dots, x_n) \in \mathbf{R}^n$ , tedy pokud SDRUŽENÁ<sup>3.2</sup> distribuční funkce vektoru  $(X_1, \dots, X_n)$  je součinem distribučních funkcí jednotlivých veličin. Obdobně definujeme nezávislost NÁHODNÝCH VEKTORŮ<sup>3.1</sup>.

Pro SPOJITĚ ROZDĚLENÉ<sup>3.5</sup>  $(X_1, \dots, X_n)$  je tato shoda ekvivalentní shodě HUSTOT<sup>2.7</sup>

$$f_{(X_1, \dots, X_n)}(x_1, \dots, x_n) = f_1(x_1) \cdots f_n(x_n)$$

pro všechna  $x_1, \dots, x_n \in \mathbf{R}$ .

Jsou-li veličiny  $X_i$  a  $X_j$  nezávislé, pak

$$E X_i X_j = E X_i E X_j, \quad \text{cov}(X_i, X_j) = \rho_{ij} = 0.$$

Nezávislé veličiny jsou tedy NEKORELOVANÉ<sup>3.9</sup>. Opačné tvrzení obecně neplatí.

Jsou-li NÁHODNÉ VELIČINY<sup>2.1</sup>  $X$  a  $Y$  nezávislé, pak vztah pro ROZPTYL<sup>2.11</sup> součtu, resp. rozdílu neobsahuje KOVARIANCI<sup>3.7</sup>,

$$\text{var}(X \pm Y) = \text{var } X + \text{var } Y.$$

- 3.11. *Vícerozměrné normální rozdělení.*  $n$ -**rozměrné normální** rozdělení  $N_n(\boldsymbol{\mu}, \mathbf{V})$ , kde  $\boldsymbol{\mu}$  je sloupcový vektor z  $\mathbf{R}^n$  a  $\mathbf{V}$  pozitivně definitní matice řádu  $n$ , je rozdělení s HUSTOTOU<sup>3.5</sup>

$$f(x_1, \dots, x_n) = \frac{1}{\sqrt{2\pi \det \mathbf{V}}} \exp(-(\mathbf{x} - \boldsymbol{\mu})' \mathbf{V}^{-1} (\mathbf{x} - \boldsymbol{\mu})/2),$$

$\mathbf{x} = (x_1, \dots, x_n)' \in \mathbf{R}^n$ .

Parametr  $\boldsymbol{\mu}$  je STŘEDNÍ HODNOTOU<sup>3.6</sup> a  $\mathbf{V}$  VARIANČNÍ MATICÍ<sup>3.8</sup> tohoto rozdělení.

Má-li NÁHODNÝ VEKTOR<sup>3.1</sup>  $\mathbf{X}$  rozdělení  $N_n(\boldsymbol{\mu}, \mathbf{V})$ , jeho složky mají také NORMÁLNÍ ROZDĚLENÍ<sup>2.16</sup>,  $X_i \sim N(\mu_i, \sigma_i^2)$ , kde  $\sigma_i^2$  je  $i$ -tý diagonální prvek ve  $\mathbf{V}$ . Normálně rozdělené jsou i lineární kombinace složek.

Jsou-li složky  $X_i$  a  $X_j$  náhodného vektoru se sdruženým normálním rozdělení NEKORELOVANÉ<sup>3.9</sup>, pak  $X_i$  a  $X_j$  jsou NEZÁVISLÉ<sup>3.10</sup>. Ve sdruženém normálním rozdělení je tedy nekorrelovanost ekvivalentní nezávislosti (což obecně neplatí).

## 4. TEORIE ODHADU

- 4.1. *Náhodný výběr.* Posloupnost NÁHODNÝCH VELIČIN<sup>2.1</sup> (příp. NÁHODNÝCH VEKTORŮ<sup>3.1</sup>)  $X_1, \dots, X_n$  se nazývá **náhodný výběr** (rozsahu  $n$  z ROZDĚLENÍ<sup>2.2</sup> s DISTRIBUČNÍ FUNKCÍ<sup>2.3</sup>  $F$ ), jsou-li tyto veličiny NEZÁVISLÉ<sup>3.10</sup> a všechny mají stejné rozdělení ( $F$ ).

Taková posloupnost veličin odpovídá výběru ze základního souboru s vrácením.

- 4.2. *Statistika (hodnota).* **Statistikou** rozumíme funkci NÁHODNÉHO VÝBĚRU<sup>4.1</sup>.

- 4.3. *Odhad.* **Odhadem** parametrické funkce  $g(\theta)$  nazýváme libovolnou STATISTIKU<sup>4.2</sup>  $S(X_1, \dots, X_n)$ , která nezávisí na  $\theta$ .

- 4.4. *Popisné statistiky.* **Popisné statistiky** popisují některé charakteristiky NÁHODNÉHO VÝBĚRU<sup>4.1</sup>. Jedná se o výběrové protějšky obdobných charakteristik ROZDĚLENÍ<sup>2.2</sup>, jako např. VÝBĚROVÝ PRŮMĚR<sup>4.5</sup>, VÝBĚROVÝ ROZPTYL<sup>4.6</sup>, VÝBĚROVÝ SMĚRODATNÁ ODCHYLKA<sup>4.7</sup>, VÝBĚROVÁ KOVARIANCE<sup>2.7</sup>, VÝBĚROVÝ KORELAČNÍ KOEFICIENT<sup>2.7</sup>, VÝBĚROVÝ MEDIÁN<sup>2.7</sup>.
- 4.5. *Výběrový průměr.* **Výběrovým průměrem** z NÁHODNÉHO VÝBĚRU<sup>4.1</sup>  $X_1, \dots, X_n$  rozumíme veličinu

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

Jde o výběrový protějšek STŘEDNÍ HODNOTY<sup>2.9</sup>.

Výběrový průměr  $\bar{X}$  je NESTRANNÝM<sup>4.8</sup> a KONZISTENTNÍM<sup>4.9</sup> odhadem střední hodnoty  $E X_i$  (existuje-li tato). Je také NEJLEPŠÍM<sup>4.11</sup> odhadem mezi lineárními (tj. tvaru  $\sum_i a_i X_i$ ) nestrannými odhady  $E X_i$ , v BINOMICKÉM<sup>2.7</sup>, POISSONOVĚ<sup>2.14</sup>, NORMÁLNÍM<sup>2.16</sup>, či EXPONENCIÁLNÍM<sup>2.15</sup> rozdělení dokonce mezi všemi odhady.

- 4.6. *Výběrový rozptyl.* **Výběrovým rozptylem** z NÁHODNÉHO VÝBĚRU<sup>4.1</sup>  $X_1, \dots, X_n$ ,  $n > 1$ , rozumíme veličinu

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

Jde o výběrový protějšek ROZPTYLU<sup>2.11</sup>.

Výběrový rozptyl  $s^2$  je NESTRANNÝM<sup>4.8</sup> a KONZISTENTNÍM<sup>4.9</sup> odhadem rozptylu  $\text{var } X_i$  (existuje-li tento).

Pro odhad rozptylu lze použít také  $\frac{1}{n} \sum_i (X_i - \bar{X})^2 = \frac{n-1}{n} s^2$ , který má menší rozptyl než  $\sigma^2$ , na rozdíl od  $s^2$  je však vychýlený.

- 4.7. *Výběrová směrodatná odchylka.* **Výběrová směrodatná odchylka**  $s$  je odmocninou z VÝBĚROVÉHO ROZPTYLU<sup>4.6</sup>  $s^2$ .
- 4.8. *Nestrannost.* ODHAD<sup>4.3</sup>  $S_n(X_1, \dots, X_n)$  parametrické funkce  $g(\theta)$  se nazývá **nestranný** (nevychýlený), jestliže

$$E_\theta S_n(X_1, \dots, X_n) = g(\theta) \quad \text{pro všechna } \theta.$$

VÝBĚROVÝ PRŮMĚR<sup>4.5</sup> a VÝBĚROVÝ ROZPTYL<sup>4.6</sup> coby odhady STŘEDNÍ HODNOTY<sup>2.9</sup> a ROZPTYLU<sup>2.11</sup> jsou nestranné (existují-li příslušné momenty).

- 4.9. *Konzistence.* ODHAD<sup>4.3</sup>  $S_n(X_1, \dots, X_n)$  parametrické funkce  $g(\theta)$  se nazývá (silně) **konzistentní**, jestliže

$$P_\theta \left[ \lim_{n \rightarrow \infty} S_n(X_1, \dots, X_n) = g(\theta) \right] = 1 \quad \text{pro všechna } \theta.$$

VÝBĚROVÝ PRŮMĚR<sup>4.5</sup> a VÝBĚROVÝ ROZPTYL<sup>4.6</sup> coby odhady STŘEDNÍ HODNOTY<sup>2.9</sup> a ROZPTYLU<sup>2.11</sup> jsou konzistentní (existují-li příslušné momenty).

- 4.10. *Střední čtvercová chyba.* **Střední čtvercová chyba** (MSE, mean squared error) ODHADU<sup>4.3</sup>  $S(X_1, \dots, X_n)$  reálné parametrické funkce  $g(\theta)$  je definována jako

$$\text{MSE}_\theta S = E_\theta (S(X_1, \dots, X_n) - g(\theta))^2,$$

pokud tato STŘEDNÍ HODNOTA<sup>2.9</sup> existuje.

Udává, jak se v průměru liší odhad od odhadované hodnoty, měříme-li odlišnost pomocí druhé mocniny jeho odchylky od správné hodnoty. Na základě MSE můžeme odhady porovnávat, určovat, který je LEPŠÍ<sup>4.11</sup>.

- 4.11. *Nejlepší odhad.* Řekneme, že  $\text{ODHAD}^{4.3} S_1$  je (dle STŘEDNÍ ČTVERCOVÉ CHYBY<sup>4.10</sup>, MSE) **lepší** než  $S_2$ , pokud

$$\text{MSE}_\theta S_1 \leq \text{MSE}_\theta S_2 \quad \text{pro všechna } \theta.$$

Odhad  $S$  nazveme **nejlepším** odhadem  $g(\theta)$  (ve smyslu MSE), pokud pro libovolný jiný odhad  $S'$  je

$$\text{MSE}_\theta S \leq \text{MSE}_\theta S' \quad \text{pro všechna } \theta$$

- 4.12. *Maximálně věrohodný odhad.* Máme-li NÁHODNÝ VÝBĚR<sup>4.1</sup>  $\mathbf{X} = (X_1, \dots, X_n)$  s HUSTOTOU<sup>2.7</sup>  $f_\theta(\mathbf{x})$ , odhadem metodou **maximální věrohodnosti** rozumíme tu hodnotu  $\theta$ , pro niž (při daných  $\mathbf{X}$ ) funkce  $L(\theta; \mathbf{X}) = f_\theta(\mathbf{X})$  nabývá maxima,

$$\hat{\theta}(\mathbf{X}) = \arg \max_{\theta} (L(\theta; \mathbf{X})).$$

Maximálně věrohodné odhady jsou (za předpokladu regularity) asymptoticky optimální (asymptoticky nestranné a asymptoticky eficientní)

- 4.13. *Intervalové odhady.* Je-li počet pozorování  $n$  malý, může se často bodový odhad dost lišit od odhadovaného parametru. Zavádějí se proto intervalové odhady.

Interval  $(D(X_1, \dots, X_n), H(X_1, \dots, X_n))$  nazveme **intervalovým odhadem**  $\theta$  o spolehlivosti  $1-\alpha$  nebo  $(1-\alpha)$ -**intervalem spolehlivosti** ( $\alpha \in (0, 1)$ ), jestliže

$$P_\theta[D(X_1, \dots, X_n) < \theta < H(X_1, \dots, X_n)] = 1 - \alpha.$$

Je-li  $H(X_1, \dots, X_n) = \infty$ , resp.  $D(X_1, \dots, X_n) = -\infty$ , mluvíme o dolním, resp. horním, intervalovém odhadu.

Krajní body intervalu spolehlivosti jsou funkcemi NÁHODNÉHO VÝBĚRU<sup>4.1</sup>, jsou to tedy NÁHODNÉ VELIČINY<sup>2.1</sup>. Definiční rovnost vyjadřuje, že intervalový odhad „pokrývá“ skutečnou hodnotu  $\theta$  s pravděpodobností  $1 - \alpha$ . Tím není tvar odhadu definován jednoznačně, obvykle hledáme nejkratší interval s takovou vlastností.

- 4.14. *Intervaly spolehlivosti v normálním rozdělení.* Je-li  $X_1, \dots, X_n$  NÁHODNÝ VÝBĚR<sup>2.1</sup> z  $N(\mu, \sigma^2)$ <sup>2.16</sup>, pak

$$\left( \bar{X} - u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{X} + u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right),$$

$$\left( \bar{X} - t_{1-\alpha/2}(n-1) \frac{s}{\sqrt{n}}, \bar{X} + t_{1-\alpha/2}(n-1) \frac{s}{\sqrt{n}} \right),$$

resp.

$$\left( \frac{(n-1)\sigma^2}{\chi_{n-1, 1-\alpha/2}^2}, \frac{(n-1)\sigma^2}{\chi_{n-1, \alpha/2}^2} \right),$$

kde  $\bar{X}$  je VÝBĚROVÝ PRŮMĚR<sup>4.5</sup> a  $s^2$  VÝBĚROVÝ ROZPTYL<sup>4.6</sup>, jsou  $1 - \alpha$  INTERVALY SPOLEHLIVOSTI<sup>4.13</sup> pro  $\mu$  (první použijeme, když  $\sigma^2$  známe, druhý, když neznáme), resp. pro  $\sigma^2$ .

Jednostranné intervaly dostaneme, když jeden krajní bod nahradíme  $\pm\infty$  a u druhého použijeme  $\alpha$  místo  $\alpha/2$ .

Díky CENTRÁLNÍ LIMITNÍ VĚTĚ<sup>7.7</sup> a KONZISTENCÍ<sup>4.9</sup>  $s^2$  můžeme stejný tvar odhadu použít i při výběrech z jiných rozdělení s konečným rozptylem,

$$\left( \bar{X} - u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{X} + u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right)$$

pak představuje přibližně  $1 - \alpha$  intervalový odhad střední hodnoty  $E X_1$ .

## 5. TESTOVÁNÍ HYPOTÉZ

5.1. *Hypotéza.* **Hypotézou** rozumíme tvrzení o ROZDĚLENÍ<sup>2.2</sup> pozorované NÁHODNÉ VELIČINY<sup>2.1</sup> (např. o rozdělení NÁHODNÉHO VÝBĚRU<sup>4.1</sup>).

Jsou-li rozdělení, která přicházejí v úvahu, parametrizována pomocí parametru  $\theta \in \Theta$ , považujeme za hypotézu také tvrzení o hodnotě tohoto parametru.

**Jednoduchou hypotézu** představuje hypotéza  $H: \theta = \theta_0$ , že parametr  $\theta$  je roven danému číslu  $\theta_0 \in \Theta$ . Hypotéza  $H': \theta \in \Theta_0$ , že parametr  $\theta$  pochází z dané (víceprvkové) množiny  $\Theta_0 \subset \Theta$ , je **hypotézou složenou**.

◊ Viz též: TEST<sup>5.2</sup>

5.2. *Test.* **Test hypotézy** je pravidlo, které pozorované hodnotě NÁHODNÉ VELIČINY<sup>2.1</sup> (obvykle realizaci NÁHODNÉHO VÝBĚRU<sup>4.1</sup>) přiřadí rozhodnutí HYPOTÉZU<sup>5.1</sup> zamítnout nebo nezamítnout.

Obvykle proti **nulové hypotéze**  $H_0: \theta \in \Theta_0$ , že hodnota parametru  $\theta$  pochází z dané množiny  $\Theta_0 \subset \Theta$ , kde  $\Theta$  je množina možných hodnot parametru, stavíme **alternativní hypotézu**  $H_1: \theta \in \Theta_1$ , kde  $\Theta_0 \cup \Theta_1 = \Theta$ , že tomu tak není.

Zamítnutí  $H_0$  pak fakticky znamená rozhodnout se pro  $H_1$ .

◊ Viz též: KRITICKÝ OBOR<sup>5.3</sup>, CHYBA 1. A 2. DRUH<sup>5.4</sup>

5.3. *Kritický obor.* Je-li TEST<sup>5.2</sup> založen na NÁHODNÉ VELIČINĚ<sup>2.1</sup> (pozorování)  $X$  s hodnotami v  $\mathbf{R}^n$ , pak jeho **kritický obor** je množina  $W \subset \mathbf{R}^n$  těch hodnot  $X$ , pro něž test HYPOTÉZU<sup>5.1</sup> zamítá.

Kritický obor lze často popsat prostřednictvím kritického oboru pro vhodnou (jednorozměrnou) STATISTIKU<sup>4.2</sup> pozorovaných hodnot.

◊ Viz též: CHYBA 1. A 2. DRUH<sup>5.4</sup>

5.4. *Chyba 1. a 2. druhu.* Rozhodnutí TESTU<sup>5.2</sup> o HYPOTÉZE<sup>5.1</sup> nemusí být vždy správné. **Chyba 1. druhu** nastává, když je hypotéza zamítnuta, přestože platí, **chyba 2. druhu**, když hypotéza zamítnuta není, přestože neplatí.

Kvalita testu je dána pravděpodobnostmi, s jakými tyto chyby nastávají. Je-li test hypotézy  $H_0: \theta \in \Theta_0$  proti ALTERNATIVĚ<sup>5.2</sup>  $H_1: \theta \in \Theta_1$  o parametru  $\theta$  založený na hodnotách NÁHODNÉ VELIČINY<sup>2.1</sup>  $X$  s KRITICKÝM OBOREM<sup>5.3</sup>  $W$ , pak pravděpodobnost chyby 1. druhu je

$$P_\theta[X \in W], \quad \theta \in \Theta_0$$

a chyby 2. druhu

$$P_\theta[X \notin W], \quad \theta \in \Theta_1,$$

kde  $P_\theta$  označuje PRAVDĚPODOBNOSTNÍ MÍRU<sup>1.3</sup>, při níž má  $X$  ROZDĚLENÍ<sup>2.2</sup> s parametrem  $\theta$ .

Při daném rozsahu VÝBĚRU<sup>4.1</sup> obvykle nelze současně minimalizovat pravděpodobnosti obou druhů chyb.

◊ Viz též: HLADINA VÝZNAMNOSTI<sup>5.5</sup>, SÍLA TESTU<sup>5.6</sup>

5.5. *Hladina významnosti.* TESTEM<sup>5.2</sup> na **hladině významnosti**  $\alpha$  ( $\alpha$ -testem),  $\alpha \in (0, 1)$ , rozumíme test, u něž pravděpodobnost CHYBY 1. DRUH<sup>5.4</sup> nepřekračuje hodnotu  $\alpha$ .

Je žádoucí, aby test měl nízkou hladinou významnosti, rozhodnutí o zamítnutí HYPOTÉZY<sup>5.1</sup> pak interpretujeme ve smyslu, že hypotéza neplatí. Obvykle volené hodnoty pro hladinu  $\alpha$  bývají 0,05, 0,01 nebo 0,001.

◇ Viz též: SÍLA TESTU<sup>5.6</sup>, P-HODNOTA<sup>5.7</sup>

- 5.6. *Síla testu.* Má-li TEST<sup>5.2</sup> (s KRITICKÝM OBOREM<sup>5.3</sup>  $W$ ) pravděpodobnost CHYBY 2. DRUHU<sup>5.4</sup>  $P_\theta[X \notin W]$ , pak hodnota

$$P_\theta[X \in W] = 1 - P_\theta[X \notin W]$$

se nazývá **síla testu** proti ALTERNATIVĚ<sup>5.2</sup>  $\theta \in \Theta_1$ .

Při hledání optimálního testu stanovíme HLADINU VÝZNAMNOSTI<sup>5.5</sup>  $\alpha$  (omezíme pravděpodobnost CHYBY 1. DRUHU<sup>5.4</sup>) a mezi  $\alpha$ -testy hledáme ten s největší silou (nejmenší chybou 2. druhu).

- 5.7. *p-hodnota.* **P-hodnota** TESTU<sup>5.2</sup> je u testů, kde má tato definice smysl, pravděpodobnost, s jakou testovací STATISTIKA<sup>4.2</sup> nabývá hodnot „horších“ (více svědčících proti testované HYPOTÉZE<sup>5.1</sup>), než je pozorovaná hodnota statistiky.

*P-hodnota* je obvyklým výstupem počítačových programů na testování hypotéz, udává mezní HLADINU VÝZNAMNOSTI<sup>5.5</sup>, při které bychom hypotézu ještě zamítali.

Hypotézu  $H_0$  zamítáme na hladině  $\alpha$ , právě když *p-hodnota* je menší než  $\alpha$ .

- 5.8. *Testy střední hodnoty normálního rozdělení.* Pro testování HYPOTÉZY<sup>5.1</sup> o STŘEDNÍ HODNOTĚ<sup>2.9</sup>  $\mu$  veličin  $X_1, \dots, X_n$  NÁHODNÉHO VÝBĚRU<sup>4.1</sup> z NORMÁLNÍHO ROZDĚLENÍ<sup>2.16</sup>  $N(\mu, \sigma^2)$

$$H_0: \mu = \mu_0 \quad \text{proti} \quad H_1: \mu \neq \mu_0,$$

kde  $\mu_0 \in \mathbf{R}$  je dané číslo, používáme TESTY<sup>5.2</sup> založené na STATISTIKÁCH<sup>4.2</sup>

$$U = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}, \quad \text{resp.} \quad T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}},$$

kde  $\bar{X}$  je VÝBĚROVÝ PRŮMĚR<sup>4.5</sup> a  $s^2$  VÝBĚROVÝ ROZPTYL<sup>4.6</sup>. Statistiku volíme podle toho, zda ROZPTYL<sup>2.11</sup>  $\sigma^2$  je známý, nebo ne, druhý z testů se nazývá **t-test**.

Hypotézu  $H_0$  zamítáme na HLADINĚ VÝZNAMNOSTI<sup>5.5</sup>  $\alpha$ , pokud

$$|U| > u_{1-\alpha/2}, \quad \text{resp.} \quad |T| > t_{1-\alpha/2}(n-1),$$

kde  $u_{1-\alpha/2}$ , resp.  $t_{1-\alpha/2}(n-1)$  jsou  $(1-\alpha/2)$  KVANTILY<sup>2.13</sup> rozdělení  $N(0, 1)$ , resp.  $t_{n-1}$ ??.

Za platnosti  $H_0$  mají  $T$ , resp.  $U$ , ROZDĚLENÍ<sup>2.2</sup>  $N(0, 1)$ , resp.  $t$ -rozdělení o  $n-1$  stupních volnosti, KRITICKÝ OBOR<sup>5.3</sup> tvoří hodnoty  $U$ , resp.  $T$ , blízké  $\pm\infty$ .

Pro test  $H_0$  proti jednostranné alternativě

$$H_1': \mu > \mu_0, \quad \text{resp.} \quad H_1'': \mu < \mu_0,$$

na hladině významnosti  $\alpha$  použijeme v případě známého  $\sigma^2$  kritický obor

$$U > u_{1-\alpha}, \quad \text{resp.} \quad U < -u_{1-\alpha},$$

při neznámém  $\sigma^2$  pak

$$T > t_{1-\alpha}(n-1), \quad \text{resp.} \quad T < -t_{1-\alpha}(n-1)$$

(kritický obor představují hodnoty statistik blízké  $+\infty$  u  $H_1'$ , resp.  $-\infty$  u  $H_1''$ ).

◇ Viz též: DVOUVÝBĚROVÝ TEST<sup>5.12</sup>, PÁROVÝ TEST<sup>5.11</sup>, TEST O STŘEDNÍ HODNOTĚ<sup>5.9</sup>, TEST O ROZPTYLU NORMÁLNÍHO ROZDĚLENÍ<sup>5.10</sup>, TEST NEZÁVISLOSTI V NORMÁLNÍM ROZDĚLENÍ<sup>5.16</sup>

- 5.9. *Test o střední hodnotě.* Pro přibližné testování HYPOTÉZY<sup>5.1</sup> o STŘEDNÍ HODNOTĚ<sup>2.9</sup>  $\mu = E X_i$  veličin v NÁHODNÉM VÝBĚRU<sup>4.1</sup>  $X_1, \dots, X_n$

$$H_0: \mu = \mu_0 \quad \text{proti} \quad H_1: \mu \neq \mu_0,$$

kde  $\mu_0 \in \mathbf{R}$  je dané číslo, použijeme, podobně jako u *t*-TESTU<sup>5.8</sup> v NORMÁLNÍM ROZDĚLENÍ<sup>2.16</sup>  $N(\mu, \sigma^2)$ , STATISTIKU<sup>4.2</sup>

$$T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}},$$

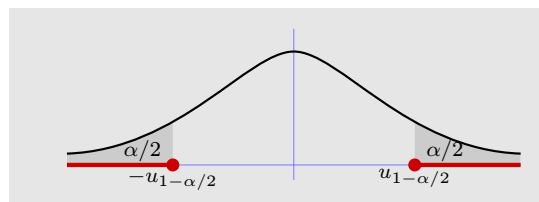
kde  $\bar{X}$  je VÝBĚROVÝ PRŮMĚR<sup>4.5</sup> a  $s^2$  VÝBĚROVÝ ROZPTYL<sup>4.6</sup>.

Hypotézu  $H_0$  zamítneme na HLADINĚ VÝZNAMNOSTI<sup>5.5</sup>  $\alpha$  v případě, že

$$|T| > u_{1-\alpha/2},$$

kde  $u_p$  značí *p*-KVANTIL<sup>2.13</sup> rozdělení  $N(0, 1)$ <sup>2.16</sup>.

Za  $H_0$  má  $T$  díky CENTRÁLNÍ LIMITNÍ VĚTĚ<sup>2.7</sup> a KONZISTENCI<sup>4.9</sup> výběrového rozptylu asymptoticky rozdělení  $N(0, 1)$ , KRITICKÝ OBOR<sup>5.3</sup> představují hodnoty  $T$  blízké  $\pm\infty$ .



Kritický obor oboustranného testu

- 5.10. *Test o rozptylu normálního rozdělení.* Nechť  $X_1, \dots, X_n$  je NÁHODNÝ VÝBĚR<sup>4.1</sup> z NORMÁLNÍHO ROZDĚLENÍ<sup>2.16</sup>  $N(\mu, \sigma^2)$ . Pro testování HYPOTÉZY<sup>5.1</sup> o hodnotě ROZPTYLU<sup>2.11</sup>  $\sigma^2$

$$H_0: \sigma^2 = \sigma_0^2 \quad \text{proti} \quad H_1: \sigma^2 \neq \sigma_0^2,$$

kde  $\sigma_0^2 > 0$  je dané číslo, používáme TEST<sup>5.2</sup> založený na STATISTICE<sup>4.2</sup>

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2},$$

kde  $s^2$  je VÝBĚROVÝ ROZPTYL<sup>4.6</sup>.

Hypotézu  $H_0$  zamítneme na HLADINĚ VÝZNAMNOSTI<sup>5.5</sup>  $\alpha$ , když

$$\chi^2 > \chi_{1-\alpha/2}^2(n-1), \quad \text{nebo} \quad \chi^2 < \chi_{\alpha/2}^2(n-1),$$

kde  $\chi_p^2(n-1)$  značí *p*-KVANTIL<sup>2.13</sup> rozdělení  $\chi_{n-1}^2$ <sup>2.7</sup>.

Za  $H_0$  má statistika  $\chi^2$  rozdělení  $\chi_{n-1}^2$ , KRITICKÝ OBOR<sup>5.3</sup> tvoří hodnoty  $\chi^2$  blízké 0 nebo  $\infty$ .

◊ Viz též: TESTY STŘEDNÍ HODNOTY NORMÁLNÍHO ROZDĚLENÍ<sup>5.8</sup>, TEST NEZÁVISLOSTI V NORMÁLNÍM ROZDĚLENÍ<sup>5.16</sup>

- 5.11. *Párový test. Párový t-test* je testem HYPOTÉZY<sup>5.1</sup> o hodnotě rozdílu mezi STŘEDNÍMI HODNOTAMI<sup>2.9</sup>  $\mu_1$  a  $\mu_2$  složek NÁHODNÝCH VEKTORŮ<sup>3.1</sup> v NÁHODNÉM VÝBĚRU<sup>4.1</sup>  $(X_1, Y_1), \dots, (X_n, Y_n)$  z DVOUROZMĚRNÉHO NORMÁLNÍHO ROZDĚLENÍ<sup>3.11</sup>,

$$H_0: \mu_1 - \mu_2 = d \quad \text{proti} \quad H_1: \mu_1 - \mu_2 \neq d,$$

kde  $d \in \mathbf{R}$  je dané číslo (pro  $d = 0$  to znamená shodu středních hodnot složek  $\mu_1 = \mu_2$ ).

Tento TEST<sup>5.2</sup> zamítá  $H_0$  na HLADINĚ VÝZNAMNOSTI<sup>5.5</sup>  $\alpha$ , pokud

$$\left| \frac{\bar{Z} - d}{s_Z/\sqrt{n}} \right| > t_{1-\alpha/2}(n-1),$$

kde  $\bar{Z}$  a  $s^2$  jsou VÝBĚROVÝ PRŮMĚR<sup>4.5</sup> a VÝBĚROVÝ ROZPTYL<sup>4.6</sup> náhodného výběru složeného z rozdílů

$$Z_i = X_i - Y_i, \quad i = 1, \dots, n,$$

a  $t_p(n)$  značí *p*-KVANTIL<sup>2.13</sup> rozdělení  $t_n$ <sup>2.7</sup>.

Ve skutečnosti provádíme *t*-TEST<sup>5.8</sup> o střední hodnotě  $\mu_1 - \mu_2$  výběru NORMÁLNĚ ROZDĚLENÝCH<sup>2.16</sup> veličin  $Z_1, \dots, Z_n$ .

Párový test můžeme používat pouze pokud je výběr tvořen dvojicemi měření.

◊ Viz též: DVOUVÝBĚROVÝ TEST<sup>5.12</sup>, TEST NEZÁVISLOSTI V NORMÁLNÍM ROZDĚLENÍ<sup>5.16</sup>

5.12. *Dvouvýběrový test. Dvouvýběrový t-test* je testem HYPOTÉZY<sup>5.1</sup> o hodnotě rozdílu STŘEDNÍCH HODNOT<sup>2.9</sup>  $\mu_1$  a  $\mu_2$  ve dvou NEZÁVISLÝCH<sup>3.10</sup> náhodných výběrech  $X_1, \dots, X_n$  z NORMÁLNÍHO ROZDĚLENÍ<sup>2.16</sup>  $N(\mu_1, \sigma^2)$  a  $Y_1, \dots, Y_n$  z  $N(\mu_2, \sigma^2)$  se stejným (i když neznámým) ROZPTYLEM<sup>2.11</sup>  $\sigma^2$ ,

$$H_0: \mu_1 - \mu_2 = d \quad \text{proti} \quad H_1: \mu_1 - \mu_2 \neq d,$$

kde  $d \in \mathbf{R}$  je dané číslo (pro  $d = 0$  to znamená shodu středních hodnot složek  $\mu_1 = \mu_2$ ).

Použijeme-li k testování STATISTIKU<sup>4.2</sup>

$$T = \frac{\bar{X} - \bar{Y} - d}{\sqrt{(n-1)s_X^2 + (m-1)s_Y^2}} \sqrt{\frac{nm(n+m-2)}{n+m}},$$

kde  $s_X^2$  a  $s_Y^2$  jsou příslušné VÝBĚROVÉ ROZPTYLY<sup>4.6</sup>, zamítneme  $H_0$  na HLADINĚ VÝZNAMNOSTI<sup>5.5</sup>  $\alpha$ , pokud

$$|T| > t_{1-\alpha/2}(n+m-2),$$

kde  $t_p(n)$  je označení pro  $p$ -KVANTIL<sup>2.13</sup> rozdělení  $t_n$ ??.

Za platnosti  $H_0$  má statistika  $T$  rozdělení  $t_{n+m-2}$ , KRITICKÝ OBOR<sup>5.3</sup> tvoří hodnoty  $T$  blízké  $\pm\infty$ .

Předpoklad nezávislosti testovaných výběrů (které mohou být různého rozsahu a měly by, alespoň přibližně, mít stejný rozptyl) je podstatný.

◊ Viz též: PÁROVÝ TEST<sup>5.11</sup>

5.13. *Chí-kvadrát test dobré shody.* Pro testování HYPOTÉZY<sup>5.1</sup> o hodnotách PRAVDĚPODOBNOTÍ<sup>1.3</sup>  $p_1, \dots, p_k$ , se kterými nastávají jednotlivé výsledky NÁHODNÉHO POKUSU<sup>1.1</sup> (o  $k$  možných výsledcích), na základě pozorování četností  $n_1, \dots, n_k$  výsledků při  $n$  NEZÁVISLÝCH<sup>1.5</sup> opakováních pokusu

$$H_0: p_j = p_1^0, \dots, p_k = p_k^0$$

proti

$$H_1: p_j \neq p_j^0 \quad \text{alespoň pro jedno } j,$$

kde  $p_1^0, \dots, p_k^0 \in (0, 1)$ ,  $\sum_j p_j^0 = 1$ , jsou daná čísla, se používá TEST<sup>5.2</sup> založený na STATISTICE<sup>4.2</sup>

$$\chi^2 = \sum_{j=1}^k \frac{(n_j - o_j)^2}{o_j} = \sum_{i=1}^k \frac{n_j^2}{o_j} - n,$$

kde  $o_j = np_j^0$ .

Za předpokladu, že  $o_j > 5$  pro všechna  $j = 1, \dots, k$ , zamítáme hypotézu  $H_0$  přibližně na HLADINĚ VÝZNAMNOSTI<sup>5.5</sup>  $\alpha$ , pokud

$$\chi^2 > \chi_{1-\alpha}^2(k-1),$$

kde  $\chi_{1-\alpha}^2(k-1)$  je  $(1-\alpha)$ -KVANTIL<sup>2.13</sup> rozdělení  $\chi_{k-1}^2$ ??.

Není-li podmínka  $o_j > 5 \quad \forall j$  splněna, je možné nejdříve některé výsledky (tj. četnosti a pravděpodobnosti) sloučit.

Za platnosti  $H_0$  jsou  $o_j$  očekávanými četnostmi  $E X_j$  a statistika  $\chi^2$  má asymptoticky pro  $n \rightarrow \infty$  rozdělení  $\chi_{k-1}^2$ . KRITICKÝ OBOR<sup>5.3</sup> je tvořen hodnotami  $\chi^2$  blízkými  $\infty$ .

Pro testování hypotézy, že NÁHODNÝ VÝBĚR<sup>4.1</sup>  $X_1, \dots, X_n$  pochází z daného ROZDĚLENÍ<sup>2.2</sup> (se známými parametry) rozdělíme obor hodnot VELIČIN<sup>2.1</sup>  $X_i$  na  $k$  intervalů  $I_1, \dots, I_k$  a za výsledek  $i$ -tého pokusu považujeme zařazení hodnoty  $X_i$  do příslušného intervalu. Testujeme pak hypotézu o pravděpodobnostech

$$p_j = P[X_i \in I_j], \quad j = 1, \dots, k,$$

na základě počtů pozorování  $n_j = \#\{i; X_i \in I_j\}$ ,  $j = 1, \dots, k$ , v jednotlivých intervalech.

◊ Viz též: CHÍ-KVADRÁT TEST PŘI NEZNÁMÝCH PARAMETRECH<sup>5.14</sup>

5.14. *Chí-kvadrát test při neznámých parametrech.* Pro testování HYPOTÉZY<sup>5.1</sup> o hodnotách PRAVDĚPODOBNOSTÍ<sup>1.3</sup>  $p_1, \dots, p_k$ , se kterými nastávají jednotlivé výsledky NÁHODNÉHO POKUSU<sup>1.1</sup> (o  $k$  možných výsledcích) z pozorování  $n_1, \dots, n_k$  četností výsledků při  $n$  NEZÁVISLÝCH<sup>1.5</sup> opakováních pokusu

$$H_0: p_j = p_j(\theta), \quad j = 1, \dots, k$$

proti

$$H_1: p_j \neq p_j(\theta) \quad \text{alespoň pro jedno } j,$$

kde  $p_j(\theta)$  jsou dané funkce a  $\theta \in \mathbf{R}^m$ ,  $m < k - 1$ , neznámý parametr, se používá TEST<sup>5.2</sup> založený na STATISTICE<sup>4.2</sup> analogické té z testu PŘI ZNÁMÝCH PARAMETRECH<sup>5.13</sup>,

$$\chi^2 = \sum_{j=1}^k \frac{(n_j - o_j)^2}{o_j} = \sum_{i=1}^k \frac{n_j^2}{o_j} - n,$$

kde  $o_j = np_j(\hat{\theta})$  a  $\hat{\theta}$  je řešením soustavy rovnic

$$\sum_j \frac{n_j}{p_j(\hat{\theta})} \frac{\partial p_j(\hat{\theta})}{\partial \hat{\theta}_i} = 0, \quad i = 1, \dots, m,$$

čili  $\hat{\theta}$  je MAXIMÁLNĚ VĚROHODNÝM<sup>4.12</sup> odhadem  $\theta$  z pozorování  $n_1, \dots, n_k$ .

Předpokládá se, že v nějakém okolí skutečné hodnoty  $\theta$  (včetně) jsou funkce  $p_1, \dots, p_k$  kladné,  $\sum_j p_j = 1$ , a zobrazení  $(p_1, \dots, p_k): \mathbf{R}^m \rightarrow \mathbf{R}^k$  je dvakrát spojitě diferencovatelné a regulární. Statistika  $\chi^2$  pak má asymptoticky rozdělení  $\chi_{k-m-1}^2$ .

Hypotézu  $H_0$  zamítáme přibližně na HLADINĚ VÝZNAMNOSTI<sup>5.5</sup>  $\alpha$ , pokud

$$\chi^2 > \chi_{1-\alpha}^2(k - m - 1),$$

$\chi_p^2(n)$  značí  $p$ -KVANTIL<sup>2.13</sup> rozdělení  $\chi_n^{2??}$ .

5.15. *Test nezávislosti v kontingenční tabulce.* Pro testování HYPOTÉZY<sup>5.1</sup> o NEZÁVISLOSTI<sup>3.10</sup> složek  $X, Y$  NÁHODNÉHO VEKTORU<sup>3.1</sup> s DISKRÉTNÍM ROZDĚLENÍM<sup>3.4</sup>, kde složky nabývají hodnot  $x_1, \dots, x_r$  a  $y_1, \dots, y_s$ ,  $r, s \geq 2$ ,

$$H_0: X, Y \text{ nezávislé} \quad \text{proti} \quad H_1: X, Y \text{ závislé},$$

se používá  $\chi^2$ -TEST DOBRÉ SHODY PŘI NEZNÁMÝCH PARAMETRECH<sup>5.14</sup> vycházející z pozorovaných četností  $n_{ij}$ ,  $i = 1, \dots, r$ ,  $j = 1, \dots, s$ , dvojic hodnot  $(x_i, y_j)$  v NÁHODNÉM VÝBĚRU<sup>4.1</sup> rozsahu  $n$  z rozdělení  $(X, Y)$ . (Tyto četnosti zapisujeme do tzv. **kontingenční tabulky**.) TEST<sup>5.2</sup> pracuje se STATISTIKOU<sup>4.2</sup>

	$y_1$	$\dots$	$y_s$	$\sum$
$x_1$	$n_{11}$	$\dots$	$n_{1s}$	$n_{1.}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$
$x_r$	$n_{r1}$	$\dots$	$n_{rs}$	$n_{r.}$
$\sum$	$n_{.1}$	$\dots$	$n_{.s}$	$n$

Kontingenční tabulka

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - o_{ij})^2}{o_{ij}} = n \sum_{i=1}^r \sum_{j=1}^s \frac{n_{ij}^2}{n_i \cdot n_{\cdot j}} - n,$$

kde  $o_{ij} = n_i \cdot n_{\cdot j} / n$ , přičemž

$$n_i = \sum_{j=1}^s n_{ij}, \quad n_{\cdot j} = \sum_{i=1}^r n_{ij}.$$

jsou řádkové a sloupcové součty v tabulce.

Za předpokladu, že  $o_{ij} > 5$  pro všechna  $i, j$ , zamítáme hypotézu  $H_0$  přibližně na HLADINĚ VÝZNAMNOSTI<sup>5.5</sup>  $\alpha$ , když

$$\chi^2 > \chi_{1-\alpha}^2((r-1)(s-1)),$$

$\chi_p^2(n)$  je označení pro  $p$ -KVANTIL<sup>2.13</sup> rozdělení  $\chi_n^{2??.?}$ .

Jsou-li některá  $o_{ij} < 5$ , je možné nejdříve sloučit příslušné řádky nebo sloupec se sousedními, počet řádků i sloupců se však nesmí zredukovat na jeden.

Za  $H_0$  představují  $o_{ij}$  odhady očekávaných četností dvojic  $(x_i, y_j)$  a statistika  $\chi^2$  má asymptoticky (pro  $n \rightarrow \infty$ ) rozdělení  $\chi_{(r-1)(s-1)}^2$ .

◊ Viz též: TEST NEZÁVISLOSTI V NORMÁLNÍM ROZDĚLENÍ<sup>5.16</sup>

- 5.16. *Test nezávislosti v normálním rozdělení.* Pro testování NEZÁVISLOSTI<sup>3.10</sup> složek NÁHODNÝCH VEKTORŮ<sup>3.1</sup>  $(X_1, Y_1), \dots, (X_n, Y_n)$  v NÁHODNÉM VÝBĚRU<sup>4.1</sup> z DVOUROZMĚRNÉHO NORMÁLNÍHO ROZDĚLENÍ<sup>3.11</sup> vektoru  $(X, Y)$ ,

$$H_0: X, Y \text{ nezávislé} \quad \text{proti} \quad H_1: X, Y \text{ závislé},$$

se používá test založený na STATISTICĚ<sup>4.2</sup>

$$T = \frac{r}{\sqrt{1-r^2}} \sqrt{n-2},$$

kde  $r$  je VÝBĚROVÝ KORELAČNÍ KOEFICIENT<sup>?.?</sup>.

HYPOTÉZU<sup>5.1</sup>  $H_0$  zamítáme na HLADINĚ VÝZNAMNOSTI<sup>5.5</sup>  $\alpha$ , pokud

$$|T| > t_{1-\alpha/2}(n-2),$$

kde pravou stranu tvoří  $(1-\alpha)$ -KVANTIL<sup>2.13</sup> rozdělení  $t_{n-2}^{??.?}$ .

Hypotéza  $H_0$  je díky normalitě ekvivalentní nulovosti KORELAČNÍHO KOEFICIENTU<sup>3.9</sup> mezi  $X$  a  $Y$ . Za platnosti  $H_0$  má statistika  $T$  rozdělení  $t_{n-2}^{??.?}$ , KRITICKÝ OBOR<sup>5.3</sup> tvoří hodnoty  $T$  blízké  $\pm\infty$ .

◊ Viz též: TEST NEZÁVISLOSTI V KONTINGENČNÍ TABULCE<sup>5.15</sup>, PÁROVÝ TEST<sup>5.11</sup>

## SEZNAM HESEL

Abecední seznam číslovaných hesel.

BAYESOVA VĚTA<sup>1.8</sup>  
 DISKRÉTNÍ ROZDĚLENÍ<sup>2.4</sup>  
 DISKRÉTNÍ SDRUŽENÉ ROZDĚLENÍ<sup>3.4</sup>  
 DISTRIBUČNÍ FUNKCE<sup>2.3</sup>  
 DVOUVÝBĚROVÝ TEST<sup>5.12</sup>  
 EXPONENCIÁLNÍ ROZDĚLENÍ<sup>2.15</sup>  
 HLADINA VÝZNAMNOSTI<sup>5.5</sup>  
 HUSTOTA<sup>2.7</sup>  
 HYPOTÉZA<sup>5.1</sup>  
 CHÍ-KVADRÁT TEST DOBRÉ SHODY<sup>5.13</sup>  
 CHÍ-KVADRÁT TEST PŘI NEZNÁMÝCH PARAMETRECH<sup>5.14</sup>  
 CHYBA 1. A 2. DRUH<sup>5.4</sup>  
 INTENZITA PORUCH<sup>2.8</sup>  
 INTERVALOVÉ ODHADY<sup>4.13</sup>  
 INTERVALY SPOLEHLIVOSTI V NORMÁLNÍM ROZDĚLENÍ<sup>4.14</sup>  
 KLASICKÁ DEFINICE PRAVDĚPODOBNOTI<sup>1.4</sup>  
 KONZISTENCE<sup>4.9</sup>  
 KORELAČNÍ KOEFICIENT<sup>3.9</sup>  
 KOVARIANCE<sup>3.7</sup>  
 KRITICKÝ OBOR<sup>5.3</sup>  
 KVANTIL<sup>2.13</sup>  
 MARGINÁLNÍ ROZDĚLENÍ<sup>3.3</sup>  
 MAXIMÁLNĚ VĚROHODNÝ ODHAD<sup>4.12</sup>  
 MEDIÁN<sup>2.12</sup>  
 MOMENTY<sup>2.10</sup>  
 NÁHODNÁ VELIČINA<sup>2</sup>  
 NÁHODNÁ VELIČINA<sup>2.1</sup>  
 NÁHODNÝ JEVIŠTĚ<sup>1.1</sup>  
 NÁHODNÝ VEKTOR<sup>3</sup>  
 NÁHODNÝ VEKTOR<sup>3.1</sup>  
 NÁHODNÝ VÝBĚR<sup>4.1</sup>  
 NEJLEPŠÍ ODHAD<sup>4.11</sup>  
 NESTRANNOST<sup>4.8</sup>  
 NEZÁVISLÉ NÁHODNÉ VELIČINY<sup>3.10</sup>  
 NEZÁVISLOST JEVIŠTĚ<sup>1.5</sup>  
 NORMÁLNÍ ROZDĚLENÍ<sup>2.16</sup>  
 ODHAD<sup>4.3</sup>  
 P-HODNOTA<sup>5.7</sup>  
 PÁROVÝ TEST<sup>5.11</sup>  
 PODMÍNĚNÁ PRAVDĚPODOBNOTA<sup>1.6</sup>  
 POISSONOVO ROZDĚLENÍ<sup>2.14</sup>  
 POPISNÉ STATISTIKY<sup>4.4</sup>  
 PRAVDĚPODOBNOTA<sup>1</sup>  
 PRAVDĚPODOBNOTNÍ FUNKCE<sup>2.5</sup>  
 PRAVDĚPODOBNOTNÍ MÍRA<sup>1.3</sup>  
 ROZDĚLENÍ<sup>2.2</sup>  
 ROZPTYL<sup>2.11</sup>  
 SDRUŽENÉ ROZDĚLENÍ<sup>3.2</sup>  
 SIGMA-ALGEBRA<sup>1.2</sup>  
 SÍLA TESTU<sup>5.6</sup>  
 SPOJITÉ ROZDĚLENÍ<sup>2.6</sup>  
 SPOJITÉ SDRUŽENÉ ROZDĚLENÍ<sup>3.5</sup>  
 STATISTIKA (HODNOTA)<sup>4.2</sup>  
 STŘEDNÍ ČTVERCOVÁ CHYBA<sup>4.10</sup>  
 STŘEDNÍ HODNOTA<sup>2.9</sup>  
 STŘEDNÍ HODNOTA VEKTORU<sup>3.6</sup>

TEORIE ODHADU<sup>4</sup>  
TEST<sup>5.2</sup>  
TEST NEZÁVISLOSTI V KONTINGENČNÍ TABULCE<sup>5.15</sup>  
TEST NEZÁVISLOSTI V NORMÁLNÍM ROZDĚLENÍ<sup>5.16</sup>  
TEST O ROZPTYLU NORMÁLNÍHO ROZDĚLENÍ<sup>5.10</sup>  
TEST O STŘEDNÍ HODNOTĚ<sup>5.9</sup>  
TESTOVÁNÍ HYPOTÉZ<sup>5</sup>  
TESTY STŘEDNÍ HODNOTY NORMÁLNÍHO ROZDĚLENÍ<sup>5.8</sup>  
VARIANČNÍ MATICE<sup>3.8</sup>  
VĚTA O CELKOVÉ PRAVDĚPODOBNOSTI<sup>1.7</sup>  
VÍCEROZMĚRNÉ NORMÁLNÍ ROZDĚLENÍ<sup>3.11</sup>  
VÝBĚROVÁ SMĚRODATNÁ ODCHYLKA<sup>4.7</sup>  
VÝBĚROVÝ PRŮMĚR<sup>4.5</sup>  
VÝBĚROVÝ ROZPTYL<sup>4.6</sup>

## SEZNAM OBRÁZKŮ

Abecední seznam obrázků.

DISTRIBUČNÍ FUNKCE DISKRÉTNÍHO ROZDĚLENÍ<sup>2.4</sup>

HUSTOTA EXPONENCIÁLNÍHO ROZDĚLENÍ<sup>2.15</sup>

HUSTOTA NORMÁLNÍHO ROZDĚLENÍ<sup>2.16</sup>

HUSTOTA ROZDĚLENÍ<sup>2.7</sup>

KONTINGENČNÍ TABULKA<sup>5.15</sup>

KRITICKÝ OBOR OBOUSTRANNÉHO TESTU<sup>5.9</sup>

PRAVDĚPODOBNOSTNÍ A DISTRIBUČNÍ FUNKCE<sup>2.5</sup>

## INDEX

## Stránky s výskytem položek

- $n$ -rozměrné normální, 16<sup>3.11</sup>, 21<sup>5.11</sup>, 24<sup>5.16</sup>
- $p$ -kvantil, 12<sup>2.13</sup>, 13<sup>2.16</sup>, 20<sup>5.8</sup>, 21<sup>5.9</sup>, 21<sup>5.10</sup>, 21<sup>5.11</sup>, 22<sup>5.12</sup>, 22<sup>5.13</sup>, 23<sup>5.14</sup>, 23<sup>5.15</sup>, 24<sup>5.16</sup>
- $t$ -test, 20<sup>5.8</sup>, 21<sup>5.9</sup>, 21<sup>5.11</sup>
- $\langle$ in.bi $\rangle$ , 17<sup>4.5</sup>
- $\langle$ in.clv $\rangle$ , 18<sup>4.14</sup>, 21<sup>5.9</sup>
- $\langle$ in.chi2 $\rangle$ , 21<sup>5.10</sup>, 22<sup>5.13</sup>, 23<sup>5.14</sup>, 23<sup>5.15</sup>
- $\langle$ in.smodch $\rangle$ , 13<sup>2.16</sup>
- $\langle$ in.t $\rangle$ , 20<sup>5.8</sup>, 21<sup>5.11</sup>, 22<sup>5.12</sup>, 24<sup>5.16</sup>
- $\langle$ in.vybkor $\rangle$ , 17<sup>4.4</sup>, 24<sup>5.16</sup>
- $\langle$ in.vybkov $\rangle$ , 17<sup>4.4</sup>
- $\langle$ in.vybmed $\rangle$ , 17<sup>4.4</sup>
- alternativní hypotéza, 19<sup>5.2</sup>, 19<sup>5.4</sup>, 20<sup>5.6</sup>
- Bayesova věta, 9<sup>1.8</sup>
- diskrétní rozdělení, 10<sup>2.4</sup>, 10<sup>2.5</sup>, 11<sup>2.9</sup>, 12<sup>2.14</sup>, 14<sup>3.4</sup>, 14<sup>3.4</sup>, 15<sup>3.6</sup>, 23<sup>5.15</sup>
- Distribuční funkcí, 9<sup>2.3</sup>, 10<sup>2.4</sup>, 10<sup>2.5</sup>, 10<sup>2.6</sup>, 10<sup>2.7</sup>, 11<sup>2.8</sup>, 12<sup>2.12</sup>, 12<sup>2.13</sup>, 12<sup>2.14</sup>, 13<sup>2.15</sup>, 13<sup>2.16</sup>, 16<sup>4.1</sup>
- hladina významnosti, 19<sup>5.5</sup>, 20<sup>5.6</sup>, 20<sup>5.7</sup>, 20<sup>5.8</sup>, 21<sup>5.9</sup>, 21<sup>5.10</sup>, 21<sup>5.11</sup>, 22<sup>5.12</sup>, 22<sup>5.13</sup>, 23<sup>5.14</sup>, 23<sup>5.15</sup>, 24<sup>5.16</sup>
- hustotou, 10<sup>2.7</sup>, 11<sup>2.8</sup>, 11<sup>2.9</sup>, 13<sup>2.15</sup>, 13<sup>2.16</sup>, 14<sup>3.5</sup>, 16<sup>3.10</sup>, 18<sup>4.12</sup>
- hypotéza, 19<sup>5.1</sup>, 19<sup>5.2</sup>, 19<sup>5.3</sup>, 19<sup>5.4</sup>, 19<sup>5.5</sup>, 20<sup>5.7</sup>, 20<sup>5.8</sup>, 21<sup>5.9</sup>, 21<sup>5.10</sup>, 21<sup>5.11</sup>, 22<sup>5.12</sup>, 22<sup>5.13</sup>, 23<sup>5.14</sup>, 23<sup>5.15</sup>, 24<sup>5.16</sup>
- chyba
1. druhu, 19<sup>5.4</sup>, 19<sup>5.5</sup>, 20<sup>5.6</sup>
  2. druhu, 19<sup>5.4</sup>, 20<sup>5.6</sup>
- intenzitu poruch, 11<sup>2.8</sup>, 13<sup>2.15</sup>
- intervalem spolehlivosti, 18<sup>4.13</sup>, 18<sup>4.14</sup>
- intervalovým odhadem, 18<sup>4.13</sup>
- jednoduchá hypotéza, 19<sup>5.1</sup>
- jev
- elementární, 7<sup>1.1</sup>
  - jistý, 7<sup>1.1</sup>
  - náhodný, 7<sup>1.1</sup>, 7<sup>1.2</sup>, 7<sup>1.3</sup>, 7<sup>1.4</sup>, 8<sup>1.5</sup>
  - nemožný, 7<sup>1.1</sup>
- kontingenční tabulka, 23<sup>5.15</sup>
- konzistentní, 17<sup>4.5</sup>, 17<sup>4.6</sup>, 17<sup>4.9</sup>, 18<sup>4.14</sup>, 21<sup>5.9</sup>
- Korelační koeficient, 15<sup>3.9</sup>, 16<sup>3.10</sup>, 16<sup>3.11</sup>, 24<sup>5.16</sup>
- Kovariance, 15<sup>3.7</sup>, 15<sup>3.8</sup>, 16<sup>3.10</sup>
- kritický obor, 19<sup>5.3</sup>, 19<sup>5.4</sup>, 20<sup>5.6</sup>, 20<sup>5.8</sup>, 21<sup>5.9</sup>, 21<sup>5.10</sup>, 22<sup>5.12</sup>, 22<sup>5.13</sup>, 24<sup>5.16</sup>
- lepší, 18<sup>4.11</sup>
- Marginálním rozdělením, 14<sup>3.3</sup>, 14<sup>3.4</sup>, 14<sup>3.5</sup>
- maximální věrohodnosti, 18<sup>4.12</sup>, 23<sup>5.14</sup>
- Medián, 12<sup>2.12</sup>, 12<sup>2.13</sup>
- Momenty, 12<sup>2.10</sup>, 12<sup>2.11</sup>
- náhodná veličina, 9<sup>2.1</sup>, 9<sup>2.2</sup>, 9<sup>2.3</sup>, 10<sup>2.4</sup>, 10<sup>2.5</sup>, 10<sup>2.6</sup>, 10<sup>2.7</sup>, 11<sup>2.8</sup>, 11<sup>2.9</sup>, 12<sup>2.10</sup>, 12<sup>2.11</sup>, 12<sup>2.14</sup>, 13<sup>2.15</sup>, 13<sup>2.16</sup>, 13<sup>3.1</sup>, 15<sup>3.7</sup>, 16<sup>3.10</sup>, 16<sup>4.1</sup>, 18<sup>4.13</sup>, 18<sup>4.14</sup>, 19<sup>5.1</sup>, 19<sup>5.2</sup>, 19<sup>5.3</sup>, 19<sup>5.4</sup>, 22<sup>5.13</sup>
- náhodný pokus, 7<sup>1.1</sup>, 7<sup>1.2</sup>, 22<sup>5.13</sup>, 23<sup>5.14</sup>
- náhodný výběr, 16<sup>4.1</sup>, 16<sup>4.2</sup>, 17<sup>4.4</sup>, 17<sup>4.5</sup>, 17<sup>4.6</sup>, 18<sup>4.12</sup>, 18<sup>4.13</sup>, 19<sup>5.1</sup>, 19<sup>5.2</sup>, 19<sup>5.4</sup>, 20<sup>5.8</sup>, 21<sup>5.9</sup>, 21<sup>5.10</sup>, 21<sup>5.11</sup>, 22<sup>5.13</sup>, 23<sup>5.15</sup>, 24<sup>5.16</sup>
- Náhodným vektorem, 13<sup>3.1</sup>, 14<sup>3.2</sup>, 14<sup>3.3</sup>, 14<sup>3.4</sup>, 14<sup>3.5</sup>, 15<sup>3.6</sup>, 15<sup>3.7</sup>, 15<sup>3.8</sup>, 15<sup>3.9</sup>, 16<sup>3.10</sup>, 16<sup>3.11</sup>, 16<sup>4.1</sup>, 21<sup>5.11</sup>, 23<sup>5.15</sup>, 24<sup>5.16</sup>
- nejlepším, 17<sup>4.5</sup>, 17<sup>4.10</sup>, 18<sup>4.11</sup>
- neslučitelné jevy, 7<sup>1.1</sup>

- nestranný, 17<sup>4.5</sup>, 17<sup>4.6</sup>, 17<sup>4.8</sup>  
 nezávislé, 16<sup>3.10</sup>, 16<sup>3.11</sup>, 16<sup>4.1</sup>, 22<sup>5.12</sup>, 23<sup>5.15</sup>, 24<sup>5.16</sup>  
 nezávislost  
   jevů, 8<sup>1.5</sup>, 8<sup>1.6</sup>, 22<sup>5.13</sup>, 23<sup>5.14</sup>  
 nulová hypotéza, 19<sup>5.2</sup>  
  
 Odhadem, 16<sup>4.3</sup>, 17<sup>4.8</sup>, 17<sup>4.9</sup>, 17<sup>4.10</sup>, 18<sup>4.11</sup>  
  
 p-hodnota, 20<sup>5.7</sup>  
 Popisné statistiky, 17<sup>4.4</sup>  
 pravděpodobnost  
   celková, 8<sup>1.7</sup>, 9<sup>1.8</sup>  
   klasická, 7<sup>1.4</sup>  
   podmíněná, 8<sup>1.6</sup>, 8<sup>1.7</sup>  
 Pravděpodobnostní funkcí,  
   10<sup>2.5</sup>, 11<sup>2.9</sup>, 12<sup>2.14</sup>, 14<sup>3.4</sup>  
 pravděpodobnostní míra,  
   7<sup>1.1</sup>, 7<sup>1.3</sup>, 7<sup>1.4</sup>, 8<sup>1.6</sup>, 9<sup>2.1</sup>, 9<sup>2.2</sup>, 9<sup>2.3</sup>, 11<sup>2.8</sup>, 12<sup>2.11</sup>, 19<sup>5.4</sup>, 22<sup>5.13</sup>, 23<sup>5.14</sup>  
 pravděpodobnostní prostor,  
   7<sup>1.3</sup>, 11<sup>2.9</sup>, 13<sup>3.1</sup>, 16<sup>3.10</sup>  
  
 rozdělení  
   exponenciální, 13<sup>2.15</sup>, 17<sup>4.5</sup>  
   normální, 13<sup>2.16</sup>, 16<sup>3.11</sup>, 17<sup>4.5</sup>, 18<sup>4.14</sup>, 20<sup>5.8</sup>, 21<sup>5.9</sup>, 21<sup>5.10</sup>, 21<sup>5.11</sup>, 22<sup>5.12</sup>  
   Poissonovo, 12<sup>2.14</sup>, 17<sup>4.5</sup>  
 Rozdělením, 9<sup>2.2</sup>, 9<sup>2.3</sup>, 10<sup>2.7</sup>, 11<sup>2.8</sup>, 12<sup>2.10</sup>, 12<sup>2.11</sup>, 12<sup>2.12</sup>, 12<sup>2.13</sup>, 14<sup>3.3</sup>, 16<sup>4.1</sup>, 17<sup>4.4</sup>, 19<sup>5.1</sup>, 19<sup>5.4</sup>, 20<sup>5.8</sup>, 22<sup>5.13</sup>  
 Rozptyl, 12<sup>2.11</sup>, 12<sup>2.14</sup>, 13<sup>2.15</sup>, 13<sup>2.16</sup>, 15<sup>3.7</sup>, 16<sup>3.10</sup>, 17<sup>4.6</sup>, 17<sup>4.8</sup>, 17<sup>4.9</sup>, 20<sup>5.8</sup>, 21<sup>5.10</sup>, 22<sup>5.12</sup>  
  
 Sdružené rozdělení, 14<sup>3.2</sup>, 14<sup>3.3</sup>, 14<sup>3.5</sup>, 16<sup>3.10</sup>  
 sigma algebra, 7<sup>1.2</sup>, 7<sup>1.3</sup>, 7<sup>1.4</sup>, 9<sup>2.1</sup>, 9<sup>2.2</sup>  
 síla testu, 20<sup>5.6</sup>  
  
 složená hypotéza, 19<sup>5.1</sup>  
 spojitě rozdělení, 10<sup>2.6</sup>, 10<sup>2.7</sup>, 11<sup>2.8</sup>, 11<sup>2.9</sup>, 12<sup>2.12</sup>, 13<sup>2.15</sup>, 13<sup>2.16</sup>, 14<sup>3.5</sup>, 14<sup>3.5</sup>, 15<sup>3.6</sup>, 16<sup>3.10</sup>, 16<sup>3.11</sup>  
 Statistikou, 16<sup>4.2</sup>, 16<sup>4.3</sup>, 19<sup>5.3</sup>, 20<sup>5.7</sup>, 20<sup>5.8</sup>, 21<sup>5.9</sup>, 21<sup>5.10</sup>, 22<sup>5.12</sup>, 22<sup>5.13</sup>, 23<sup>5.14</sup>, 23<sup>5.15</sup>, 24<sup>5.16</sup>  
 Střední čtvercová chyba,  
   17<sup>4.10</sup>, 18<sup>4.11</sup>  
 Střední hodnota, 11<sup>2.9</sup>, 12<sup>2.10</sup>, 12<sup>2.11</sup>, 12<sup>2.12</sup>, 12<sup>2.14</sup>, 13<sup>2.15</sup>, 13<sup>2.16</sup>, 15<sup>3.6</sup>, 17<sup>4.5</sup>, 17<sup>4.8</sup>, 17<sup>4.9</sup>, 17<sup>4.10</sup>, 20<sup>5.8</sup>, 21<sup>5.9</sup>, 21<sup>5.11</sup>, 22<sup>5.12</sup>  
 Střední hodnotou, 15<sup>3.6</sup>, 16<sup>3.11</sup>  
  
 test, 19<sup>5.2</sup>, 19<sup>5.3</sup>, 19<sup>5.4</sup>, 19<sup>5.5</sup>, 20<sup>5.6</sup>, 20<sup>5.7</sup>, 20<sup>5.8</sup>, 21<sup>5.10</sup>, 21<sup>5.11</sup>, 22<sup>5.13</sup>, 23<sup>5.14</sup>, 23<sup>5.15</sup>  
   chí-kvadrát, 22<sup>5.13</sup>, 23<sup>5.14</sup>  
   při neznámých parametrech, 23<sup>5.14</sup>, 23<sup>5.15</sup>  
   nezávislosti  
     v normálním rozdělení, 24<sup>5.16</sup>  
   párový, 21<sup>5.11</sup>  
   střední hodnoty  
     dvouvýběrový, 22<sup>5.12</sup>  
     normálního rozdělení, 20<sup>5.8</sup>  
 test o rozptylu, 21<sup>5.10</sup>  
  
 Varianční matice, 15<sup>3.8</sup>, 16<sup>3.11</sup>  
 Výběrová směrodatná odchylka, 17<sup>4.4</sup>, 17<sup>4.7</sup>  
 Výběrovým průměrem,  
   17<sup>4.4</sup>, 17<sup>4.5</sup>, 17<sup>4.6</sup>, 17<sup>4.8</sup>, 17<sup>4.9</sup>, 18<sup>4.14</sup>, 20<sup>5.8</sup>, 21<sup>5.9</sup>, 21<sup>5.11</sup>  
 Výběrovým rozptylem,  
   17<sup>4.4</sup>, 17<sup>4.6</sup>, 17<sup>4.7</sup>, 17<sup>4.8</sup>, 17<sup>4.9</sup>, 18<sup>4.14</sup>, 20<sup>5.8</sup>, 21<sup>5.9</sup>, 21<sup>5.10</sup>, 21<sup>5.11</sup>, 22<sup>5.12</sup>