Pravdepodobnosť a štatistika

(zápisky z prednášok doc. Lamoša prednášaných v roku 2000)

Spísali:

Július Loman (*lomo@lomo.sk*)

Peter Kubíni (*8kubini@st.fmph.uniba.sk*)

Peter Hudec (*8hudecp@st.fmph.uniba.sk*)

6. decembra 2000

Obsah

1.2 Axiomatická definícia pravdepodobnosti

Veta 1.2.1 *Nech Ω je neprázdna množina elementárnych udalostí. Nech S je pole náhodných udalostí definované na Ω. Potom pravdepodobnosťou nazveme ľubovoľnú reálnu funkciu P definovanú na S, ktorá spĺňa nasledovné axiómy:*

- *Každej udalosti $A \in S$ je priradené nezávislé číslo $P(A) \geq 0$*
- *Istej udalosti je priradené číslo 1, teda $P(\Omega) = 1$*
- *Pre ľubovoľnú postupnosť udalostí $A_1, A_2, ..., A_n$, ktoré sa navzájom vylučujú $A_i \cap A_j$ je udalosť nemožná pre $i \neq j$*

Potom platí $P(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$

1.3 Geometrická definícia pravdepodobnosti

Definícia 1.3.1 *Nech G je nejaká rovinná množina, nech $g \subset G$. Pričom vieme vypočítať obsah oboch množín. Nech výsledok pokusu je určený náhodným zvolením bodu množiny G, potom pravdepodobnosť toho, že náhodne zvolený bod (nastatie udalosti) padne do množiny g definujeme ako podiel plošných obsahov. $P = \frac{\mu(g)}{\mu(G)}$, pričom $\mu(g) \geq 0$ a $\mu(G) > 0$*

1.4 Klasická definícia pravdepodobnosti

Definícia 1.4.1 *Majme množinu Ω pozostávajúcu z elementárnych vzájomne sa vylučujúcich rovnako možných udalostí. Ak m $(m \leq n)$ z týchto elementárnych udalostí má za následok nastatie udalosti A $(n - m$ elemetárnych udalostí nemá za následok nastatie udalosti $A)$ potom pravdepodobnosť udalosti A definujme ako podiel $P(A) = \frac{m}{n}$.*

1.5 Základné vlastnosti pravdepodobnosti

Veta 1.5.1 (Pravdep. je monotónna, subtrakt. funkcia) *Nech (Ω, S, P) je pravdepodobnostný priestor. $A, B \in S$, $A \subset B$. Potom platí $P(A) \leq P(B)$, $P(B - A) = P(B) - P(A)$.*

Dôkaz: $B = A \cup (B - A), A \cap (B - A) = \emptyset$
$P(B) = P(A \cup (B - A)) = P(A) + P(B - A) \geq 0$
Z predchádzajúceho riadku nám vyplýva: $P(B) \geq P(A)$, pretože $P(B - A)$ je nezáporné. $\square$

Veta 1.5.2 *Nech (Ω, S, P) je pravdepodobnostný priestor. Nech $A \in S$. Potom platí, $P(A^C) = 1 - P(A)$.*

Dôkaz: $A \in S$, $A^C \in S$,
$A \cup A^C \in S$, $\Omega \in S$
$A \cap A^C = \emptyset$
$A \cup A^C = \Omega$
Dostávame: $P(A) + P(A^C) = P(\Omega) = 1$ $\square$

Veta 1.5.3 *Nech (Ω, S, P) je pravdepodobnostný priestor. Nech $\emptyset \in S$. Potom $P(\emptyset) = 0$.*

Veta 1.6.3 *Nech $B \in S, P(B) > 0$. Potom pre ľubovoľnú udalosť A platí*

- $P(A/B) \geq 0$
- $P(\Omega/B) = 1$
- $P(\bigcup_{n=1}^{\infty} A_n/B) = \sum_{n=1}^{\infty} P(A_n/B)$
- $P(A/B) = \frac{P(A \cap B)}{P(B)}$
- $P(A_2/A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)} \Rightarrow P(A_1 \cap A_2) = P(A_1).P(A_2/A_1)$

Veta 1.6.4 (veta o súčine pravdepodobností) *Nech (Ω, S, P), nech A_1, A_2, ..., $A_n \in S$, pričom platí $P(A_1 \cap A_2 \cap ... \cap A_{n-1}) > 0$. Potom platí $P(A_1 \cap A_2 \cap ... \cap A_{n-1} \cap A_n) = P(A_1)P(A_2/A_1)P(A_3/A_1 \cap A_2)...P(A_n/A_1 \cap A_2 \cap ... \cap A_n)$*

Veta 1.6.5 (veta o úmernej pravdepodobnosti) *Nech (Ω, S, P), nech A_1, A_2, ..., $A_n \in S$ je postupnosť disjunktných udalostí s nenulovými pravdepodobnosťami také, že $\bigcup_{i=1}^{\infty} A_i = \Omega$, nech $B \in S$, pričom predpokladáme, že poznáme $P(B/A_i)$. Potom platí $P(B) = \sum_{i=1}^{n} P(A_i)P(B/A_i)$.*

Dôkaz: $B = B \cap \Omega = B \cap (\bigcup_{i=1}^{n} A_i) = \bigcup_{i=1}^{n} (A_i \cap B)$
$P(B) = \sum_{i=1}^{n} P(A_i \cap B) = \sum_{i=1}^{n} P(A_i)P(B/A_i)$ $\square$

Veta 1.6.6 (Prvá Bayesova veta) *Nech (Ω, S, P), nech A_1, A_2, ..., $A_n \in S$ je postupnosť disjunktných udalostí s nenulovými pravdepodobnosťami také, že $\bigcup_{i=1}^{\infty} A_i = \Omega$, nech $B \in S$, pričom $P(B) > 0$, poznáme $P(A_i/B)$. Potom platí*

$$P(B) = \frac{P(A_i)P(B/A_i)}{\sum_{j=1}^{n} P(A_j)P(B/A_j)}; \forall i = 1, ..., n$$

Veta 1.6.7 (Druhá Bayesova veta) *Nech (Ω, S, P), nech $A_1, A_2, ..., A_n \in S$ je postupnosť disjunktných udalostí s nenulovými pravdepodobnosťami také, že $\bigcup_{i=1}^{\infty} A_i = \Omega$, nech $B, C \in S$, pričom $P(B) > 0$, poznáme $P(A_i/B)$, $P(C/A_i \cap B)$, pre $i = 1...n$. Potom platí*

$$P(C/B) = \frac{\sum_{i=1}^{n} P(A_i)P(B/A_i)P(C/A_i \cap B)}{\sum_{j=1}^{n} P(A_j)P(B/A_j)}$$

Definícia 1.6.8 *Budeme hovoriť, že udalosti sú navzájom nezávislé, ak platí $P(A \cap B) = P(A)P(B)$.*

Definícia 1.6.9 *Nech $A_1, A_2, ..., A_n$ sú postupnosti udalostí. Budeme hovoriť, že sú po dvoch nezávislé, ak platí $P(A_i \cap A_j) = P(A_i)P(A_j)$, pre $i, j = 1, ..., n; i \neq j$.*

Poznámka 1.6.10 *Uvažujme A_1, A_2, A_3.*
$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$ *(1)*
$P(A_1 \cap A_2) = P(A_1)P(A_2)$ *(2)*
$P(A_1 \cap A_3) = P(A_1)P(A_3)$
$P(A_2 \cap A_3) = P(A_2)P(A_3)$
$(1) \not\Rightarrow (2); (2) \not\Rightarrow (1)$

$$x_i = \frac{k_i - np}{\sqrt{npq}}; i = 1, 2.$$

Veta 1.7.5 (Čebyševova nerovnosť - špeciálne znenie) *Nech $A \in S$ taká, že nastane pri každom pokuse s pravdepodobnosťou $P(A) = p \in (0,1)$. Vykonáme n nezávislých pokusov. Potom pravdepodobnosť toho, že relatívna početnosť udalosti A (označujeme $\frac{k_n}{n}$) sa bude od neznámej pravdepodobnosti líšiť o menej ako $\varepsilon > 0$ sa rovná*

$$P\left(\left|\frac{k_n}{n} - p\right| < \varepsilon\right) \geq 1 - \frac{p(1-p)}{n\varepsilon^2} \geq \frac{1}{4n\varepsilon^2}$$

2.4 Parametre náhodných premenných

Definícia 2.4.1 *Nech náhodná premenná X nadobúda hodnoty x_i,$p_i = (X = x_i)$; $i = 1, 2, ..., n$. Potom výraz tvaru $E(X) = \sum_i x_i p_i$ sa nazýva stredná hodnota diskrétnej náhodnej premennej.*
Tento rad musí konvergovať, teda $\sum |x_i| \, p_i < \infty$.

$$E(X) = \sum_{i=1}^{n} i \binom{n}{i} p^i q^{n-i} = \sum_{i=1} i \frac{n(n-1)!pp^{i-1}q^{n-1-(i-1)}}{i(i-1)!\,[n-1-(i-1)]!} =$$

$$= np \sum_{k=0}^{n-1} \frac{(n-1)!}{k!(n-1-k)!} p^k q^{n-1-k} = np$$

Definícia 2.4.2 *Nech je daná náhodná premenná X s hustotou pravdepodobnosti $f(x)$. Potom výraz $E(X) = \int_{-\infty}^{\infty} x f(x) dx$ nazývame strednou hodnotou spojitej náhodnej premennej, pričom musí platiť $\int_{-\infty}^{\infty} x f(x) dx < \infty$ (teda konverguje).*

Poznámka 2.4.3

- $X = C$; $E(C) = C$
- $X = CX$; $E(CX) = C.E(X)$
- $X = CX + B$; $E(CX + B) = C.E(X) + B$

Veta 2.4.4 *Stredná hodnota odchýlky náhodnej premennej od jej strednej hodnoty je rovná* 0. $E(X - E(X)) = 0$.

2.5 Momenty náhodných premenných

Definícia 2.5.1 *Stredná hodnota $g(x) = x^k$ sa nazýva k-tym začiatočným momentom*

$$\nu_k = E(X^k) = \sum_i x_i^k p_i \text{ - diskrétna}$$

$$\nu_k = E(X^k) = \int_{-\infty}^{\infty} x_i^k f(x) \text{ - spojitá}$$

Definícia 2.5.2 *Stredná hodnota z funkcie $g(x) = (X - A)^k$ je k-ty všeobecný moment $A \in R$ (ľub. konštanta). $\mu_k' = E\left[(X - A)^k\right]$*

Definícia 2.5.3 *Stredná hodnota z funkcie $g(x) = (X - E(X))^k$ sa nazýva k-ty centrálny moment.*

$$\mu_k = E\left[(X - E(X))^k\right] = \sum_i (x_i - E(X))^k p_i$$

2.6 Dvojrozmerná náhodná premenná

Definícia 2.6.1 *Dvojicu* $(X(\omega), Y(\omega))$ *definovanú nad* $\Omega, \omega \in \Omega$, *budeme nazývať dvojrozmernou náhodnou premennou vzhľadom na* S, *ak vzor ľubovoľného intervalu reálnych čísiel tvaru* $I_2 = (-\infty, x) \times (-\infty, y)$ *je prvkom* S, *t.j.*

$$x^{-1}(I_2) = \{\omega_i(X(\omega), Y(\omega) \in I_2)\} = \{\omega, X(\omega) < x, Y(\omega) < y\} \in S, (x, y) \in R_2$$

Definícia 2.6.2 *Distribučnou funkciou dvojrozmernej náhodnej premennej je reálna funkcia 2 reálnych premenných* x, y *definovaná takto:*

$$F(x, y) = P\{\omega, X(\omega) < x, Y(\omega) < y\}$$

Iný spôsob zápisu:

$$F(x, y) = P(X < x, Y < y)(x, y) \in R_2$$

Veta 2.6.3 *Nech* $F(x, y)$ *je distribúciou* (x, y)

- $F(-\infty, y) = 0, F(x, -\infty) = 0$
- $F(-\infty, \infty) = 1$
- $x_1, x_2, x_1 \leq x_2 \Longrightarrow F(x_1, y) \leq F(x_2, y)$
- $y_1, y_2, y_1 \leq y_2 \Longrightarrow F(x, y_1) \leq F(x, y_2)$

potom platí :
$P(x_1 < X < x_2, y_1 < Y < y_2) =$
$P(X < x_2, Y < y_2) - P(X < x_2, Y < y_1) - P(X < x_1, Y < y_2)+$
$+P(X < x_1, Y < y_1) = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x, y_1) \geq 0$

Definícia 2.6.4 *Nech* (X, Y) *nadobúda hodnoty* (x_i, y_i) *s pravdepodobnosťou* $p_{ij} = P(X = x_i, Y = y_i)$ $i, j = 1, ..., n$. *Potom* p_{ij} *sa nazýva rozdelenie pravdepodobnosti* (X, Y) *premenných, ak platí* $\sum_i \sum_j P(X = x_i, Y = y_j) = 1$

Definícia 2.6.5 *Distribučná funkcia diskrétnej 2-rozmernej náhodnej premennej je definovaná takto:*

$$F(x, y) = P(X < x, Y < y) = \sum_{\{i, x_i < x\}} \sum_{\{j, y_j < y\}} p_{ij}$$

Definícia 2.6.6 *Distribučná funkcia spojitej 2-rozmernej náhodnej premennej je definovaná takto :*

$$F(x, y) = P(X < x, Y < y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y) dx dy$$

Definícia 2.6.7 *Marginálnymi pravdepodobnosťami zložiek diskrétnej premennej* (X, Y) *nazývame výrazy tvaru*

$$\sum_{j}^{n} p_{ij} = p_{i\circ}, \sum_{i}^{m} p_{ij} = p_{\circ j}$$

- $cov(x,x) = D(x), cov(y,y) = D(y)$

Veta 2.7.5 *Nech X, Y sú nezávislé náhodné premenné, potom $cov(x,y) = 0$.*

Dôkaz: Smola bude neskôr. TODO □

Definícia 2.7.6 *Nech (X,Y) je dvojrozmerná náh. premenná. Nech existujú $D(X), D(Y)$, μ_{11}.Potom výraz tvaru:*

$$\rho(x,y) = \frac{E[(X - E(X))(Y - E(Y))]}{\sqrt{E[(X - E(X))^2]}\sqrt{E[(Y - E(Y))^2]}} = \frac{cov(x,y)}{\sigma(x)\sigma(y)}$$

sa nazýva koeficient korelácie. ρ nadobúda hodnoty z intervalu od -1 po 1.

Veta 2.7.7 *Nech náhodné premenné sú zviazané lineárnym vzťahom $y = ax+b$, kde $a \neq 0$,b je ľubovoľná konštanta, potom $\rho_{x,y} = \pm 1$*

Dôkaz: Priamo z definície koeficientu korelácie □

2.8 Lineárna regresia

$y = ax + b$

$S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 = \min$

Chceme aby, $\frac{\delta S(a,b)}{\delta a} = 0, \frac{\delta S(a,b)}{\delta b} = 0,$

$$0 = \frac{\delta S(a,b)}{\delta a} = 2\sum_{i=1}^{n}(y_i - ax_i - b)(-x_i)$$

z toho nám po úprave plynie:

$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i x_i$$

z druhej rovnice dostávame:

$$0 = \frac{\delta S(a,b)}{\delta b} = 2\sum_{i=1}^{n}(y_i - ax_i - b)(-1)$$

$$a\sum_{i=1}^{n} x_i + bn = \sum_{i=1}^{n} y_i$$

Takže máme sústavu:

$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i x_i$$

$$a\sum_{i=1}^{n} x_i + bn = \sum_{i=1}^{n} y_i$$

2.9 Lineárna regresia z hľadiska pravdepodobnosti

Definícia 2.9.1 *Nech (X, Y) sú náhodne premenné. Regresnou priamkou náhodne premennej Y vzhľadom na náhodne premennú X budeme nazývať priamku tvaru $y = ax + b$, ak platí*

$$E[(Y - aX - b)^2] \leq E[(Y - a'X - b')^2], \text{ pre ľubovoľné } a', b'$$

Definícia 2.9.2 *Priamka $x = cy + d$ sa nazýva regresnou priamkou X vzhľadom na Y, ak platí*

$$E[(X - cY - d)^2] \leq E[(X - c'Y - d')^2], \text{ pre ľubovoľné } c', d'$$

Veta 2.9.3 *Nech (X, Y) je 2-rozmerná náhodne premenná. Nech*

$$E(X) = \mu_x, D(X) = \sigma_x^2, E(Y) = \mu_y, D(Y) = \sigma_y^2$$

$$\mu_{11} = cov(X, Y) = E(XY) - E(X)E(Y) = E[(X - \mu_x)(Y - \mu_y)]$$

Potom priamka tvaru:

$$Y - E(X) = \frac{\mu_{11}}{\sigma_x^2}(X - E(X))$$

je regresnou priamkou náhodne premennej Y vzhľadom na náhodne premennú X.

Dôkaz:

$$S(a', b') = E[(Y - a'X - b')^2] = [y^2 + a'^2x^2 + b'^2 - 2a'xy - 2b'y + 2a'b'x] =$$

$$= E(Y^2) + a'^2E(X^2) - 2a'E(XY) - 2b'E(Y) + 2a'b'E(X) + b'^2 =$$

$$E(Y^2) - E^2(Y) + E^2(Y) + a'^2E(X^2) - a'^2E^2(X) + a'^2E^2(X) - 2a'(E(XY) -$$

$$-E(X)E(Y) + E(X)E(Y)) - 2b'E(Y) + 2a'b'E(X) + b'^2$$

ďalej

$$\sigma_y^2 + E^2(Y) + a'^2(\sigma_x^2 + E^2(X)) - 2a'(\mu_{11} + E(X)E(Y)) + b'^2$$

keďže:

$$\frac{\delta S(a', b')}{\delta a} = 0 \qquad \frac{\delta S(a', b')}{\delta b} = 0$$

z toho vyplýva:

$$\frac{\delta S(a', b')}{\delta a} = 2a'(\sigma_x^2 + E^2(X)) - 2(\mu_{11} + E(X)E(Y)) + 2b'E(X) = 0$$

3 Zákony rozdelenia pravdepodobností

3.1 Binomické rozdelenie

Definícia 3.1.1 *náhodná premenná X má binomické rozdelenie, ak nadobúda hodnoty j s pravdepodobnosťou*

$$p_j = P(X = j) = \binom{n}{j} p^j q^{n-j};\ j = 0, 1, \ldots, n$$

Príklad 3.1.2 *Teraz si ukážeme, že je to rozdelenie pravdepodobností, teda $\sum_{j=0}^{n} p_j = 1$*

$$\sum_{j=0}^{n} p_j = \sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j} = (p+q)^n = (p + (1-p))^n = 1$$

distribučná funkcia

$$F(x) = P(X < x) = \sum_{\{j; x_j < x\}} p_j = \sum_{\{j; j < x\}} p_j$$

stredná hodnota

$$E(X) = \sum_j x_j p_j = \sum_{j=0}^{n} j p_j = \sum_{j=0}^{n} j \binom{n}{j} p^j q^{n-j} = np$$

disperzná hodnota

$$D(X) = E\left(X - E\left(X\right)\right) = \underbrace{E(X^2)}_{*} - E^2(X) = n(n-1)p^2 + np - n^2p^2$$

$$\begin{aligned} * \quad = \quad & \sum_{j=0}^{n} j^2 \binom{n}{j} p^j q^{n-j} = \sum_{j=0}^{n} (j(j-1) + j) \binom{n}{j} p^j q^{n-j} = \\ & \underbrace{\sum_{j=0}^{n} j(j-1) \binom{n}{j} p^j q^{n-j}}_{**} + \underbrace{\sum_{j=0}^{n} j \binom{n}{j} p^j q^{n-j}}_{=np} \end{aligned}$$

$$\begin{aligned} ** = \sum_{j=2}^{n} j(j-1) \tfrac{n(n-1)(n-2)!}{j(j-1)(j-2)!} p^2 p^{j-2} q^{n-2-(j-2)} \stackrel{k=j-2}{=} \\ n(n-1)p^2 \underbrace{\sum_{k=0}^{n-2} \binom{n-2}{k} p^k q^{n-2-k}}_{=1} = n(n-1)p^2 \end{aligned}$$

3.3 Hypergeometrické rozdelenie

Skúsme najprv malý príklad:

Príklad 3.3.1 *Majme N prvkov a nech M má určitú vlastnosť ($M \leq N$). Vykonajme náhodný výber rozsahu n. Aká je pravdepodobnosť, že medzi n prvkami bude práve x prvkov s vlastnosťou M?*

Riešenie 3.3.2

- $\binom{N}{n}$ - *počet všetkých výberov*
- $\binom{M}{x}$ - *počet výberov, ktoré majú tú vlastnosť*
- $\binom{N-M}{n-x}$ - *počet výberov, ktoré nemajú tú vlastnosť*
- $\binom{M}{x}\binom{N-M}{n-x}$ - *počet priaznivých výberov*

  teda dostávame $P(X=x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$ - *hypergeometrická schéma*

Definícia 3.3.3 *Náhodná premenná X má hypergeometrické rozdelenie, ak nadobúda hodnoty x s pravdepodobnosťou*

$$P(X=x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}};\ x = 0, 1, \ldots, \min(M, n)$$

Príklad 3.3.4 *Teraz si ukážeme, že je to rozdelenie pravdepodobností, teda $\sum_{x=0} P(X=x) = 1$ (viď kombinatoricka analýza = identita)*

stredná hodnota

$$E(X) = \frac{1}{\binom{N}{n}} \sum_x x\binom{M}{x}\binom{N-M}{n-x} = \frac{1}{N\binom{N-1}{n-1}} M \sum_x \binom{M-1}{x-1}\binom{N-M}{n-x} =$$
$$\frac{M}{\frac{N}{n}} = n\frac{M}{N}$$

v úlohách sa používa $p = \frac{M}{N}$

Poznámka 3.3.5 *Limitou hypergeometrického rozdelenia je binomické rozdelenie, pre $n \to \infty$.*

Príklad 3.3.6 *Za smenu sa vyrobí $1000ks$ výrobkov. Z nich $50ks$ býva chybných. Aká je pravdepodobnosť, že v zásielke $n = 100ks$ bude $x = 10ks$ chybných?*

Riešenie 3.3.7 *Ako teda vyzerajú čísla n, p ?*

- $n = 100$
- $p = \frac{50}{1000} = 0,05$
- $P(X=10) = \frac{\binom{50}{10}\binom{1000-50}{100-10}}{\binom{1000}{100}} = \frac{\binom{50}{10}\binom{950}{90}}{\binom{1000}{100}} \doteq \binom{100}{10}.0,05^{10}.0,95^{90} \doteq 0,00166$

Definícia 3.4.4 *Hovoríme, že náhodná premenná ϕ má normované normálne rozdelenie, ak hustota pravdepodobnosti je tvaru:*

$$\phi(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}, y \in (-\infty, \infty)$$

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

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

po substitúcii dostaneme:

$$Y = \frac{x-\mu}{\sigma}$$

$\Phi(y) = P(Y < y) = P(\frac{x-\mu}{\sigma} < y) = P(X < \mu + \sigma y) = F(\mu + \sigma y)$
z toho:
$\Phi(y) = F(\mu + \sigma y)$
$Y(y) = \frac{d}{dy}\Phi(y) = \frac{d}{dy}F(\mu+\sigma y) = \sigma f(\mu+\sigma y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{\mu+\sigma y-\mu}{\sigma})^2} = \frac{1}{\sqrt{2\pi}} e^{\frac{1}{2}y^2}$

Príklad 3.4.5 *Aká je pravdepodobnosť, že náhodná premenná X nadobúda hodnoty väčšie ako x_0?*

Riešenie 3.4.6 $P(X > x_0) = 1 - P(X \leq x_0) = 1 - P(\frac{X-\mu}{\sigma} \leq \frac{x_0-\mu}{\sigma}) = 1 - P(U < \frac{x_0-\mu}{\sigma}) =$
$= 1 - \Phi(\frac{x_0-\mu}{\sigma})$

Príklad 3.4.7 $X \sim N(20, 4)$.*Aká je pravdepodobnosť, že náhodne premenná X nadobúda hodnoty väčšie ako 23?*

Riešenie 3.4.8 $P(X > 23) = 1 - P(X \leq 23) = 1 - P(\frac{x-23}{2} \leq \frac{23-20}{2}) =$
$= 1 - P(U \leq 1,5) = 1 - \Phi(1,5) = 1 - 0,933 = 0,067$

Príklad 3.4.9 $X \sim N(28,38; 0,16)$. *Aká je pravdepodobnosť toho, že $27 < X < 30$?*

Riešenie 3.4.10 $P(27 < X < 30) = P(\frac{27-28,38}{0,4} < \frac{x-28,38}{0,4} < \frac{30-28,38}{0,4}) = P(-3,45 < U < 4,05) = \Phi(4,05) - \Phi(-3,45) = 0,99997 - 0,00028 = 0,99969$

Príklad 3.4.11 $X \sim N(\mu, \sigma^2)$. *Aká je pravdepodobnosť, že náhodná premenná nadobúda hodnoty z intervalu*

- $(\mu - \sigma, \mu + \sigma)$
- $(\mu - 2\sigma, \mu + 2\sigma)$
- $(\mu - 3\sigma, \mu + 3\sigma)$

Riešenie 3.4.12 *Vypočítame pravdepodobnosť pre interval $(\mu - \sigma, \mu + \sigma)$. Ostatné prípady riešime analogicky.*

$$\begin{aligned} P(\mu - \sigma y < X < \mu + \sigma y) &= P(\tfrac{\mu-\sigma-\mu}{\sigma} < \tfrac{x-\mu}{\sigma} < \tfrac{m+\sigma-\mu}{\sigma}) = \\ &= P(-1 < U < 1) = \Phi(1) - \Phi(-1) = \\ &= 0,84134 - 0,15666 = \\ &= 0,68268 = 68,2\% \end{aligned}$$

alebo ak máme disperziu tvaru:

$$S_*^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

platí:

$$\begin{aligned} E(S_*^2) &= \frac{n-1}{n}\sigma^2 \\ D(S_*^2) &= \frac{(n-1)^2}{n^3}(\mu_4 - \frac{n-3}{n-1}\sigma^4) \end{aligned}$$

Veta 3.5.5 (4) *Nech* $(X_1, X_2, ..., X_n) \sim N(\mu, \sigma^2)$, *nech* σ^2 *poznáme. Potom náhodná premenná*[4]

$$U = \frac{\overline{x} - \mu}{\sigma}\sqrt{n} \sim N(0,1)$$

Definícia 3.5.6 *Číslo* $u_{1-\frac{\alpha}{2}}$ *nazývame* $(1 - \frac{\alpha}{2})$*-tým kvantilom normálneho rozdelenia, ak platí*

$$\Phi(u_{1-\frac{\alpha}{2}}) = P(U < u_{1-\frac{\alpha}{2}}) = \int_{-\infty}^{u_{1-\frac{\alpha}{2}}} f(x)dx = 1 - \frac{\alpha}{2}$$

Definícia 3.5.7 [5]*Náhodná premenná* X *má* χ^2 *rozdelenie s n stupňami voľnosti, ak jej funkcia hustoty (pravdepodobnosti) je (definované len pre* $X > 0$*)*

$$f_n(X) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}X^{\frac{n}{2}-1}e^{-\frac{x}{2}}$$

$\nu_k = E(X^k) = 2^k\frac{\Gamma(k+\frac{n}{2})}{\Gamma(\frac{n}{2})}$ *a ďalej* $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$ *potom* $E(X) = n$ *a disperzia* $D(X) = 2n$

Definujeme náhodnú premennú $U = \frac{X-n}{\sqrt{2n}} \sim N(0,1) \quad n > 30$

Definícia 3.5.8 *Súčet stupňov voľnosti rozumieme najmenší počet náhodných premenných TODO súčet štvorcov odchýliek pomocou lineárnej transformácie. Ich počet určíme tak, že od počtu všetkých sčítancov v súčte štvorcov odchýliek odpočítame počet tých lineárnych väzieb, s ktorými sú sčítance zviazané. Napríklad:*

$Q_1 = \Sigma_{i=1}^{n}(\frac{X_i - \overline{X}}{\sigma})^2 \quad \overline{X} = \frac{1}{n}\Sigma x_i$ *(n-1) stupňov voľnosti*

$Q_2 = \Sigma_{i=1}^{n}(\frac{X_{ij} - \overline{X}_{i\cdot}}{\sigma})^2 \quad \overline{X}_{1\cdot} = \frac{1}{p}\Sigma_{j=1}^{p}X_{ij}$ *má (np-n) stupňovvoľnosti*

Definícia 3.5.9 *Číslo* $\chi^2_\alpha(n-1)$ *nazývame kritickou hodnotou* χ^2 *rozdelenia ak platí*

$P(X > \chi^2_\alpha(n-1)) = \int_{\chi^2_\alpha(n-1)}^{\infty} f_n(x)dx = \alpha$

[4]Náhodná premenná U sa používa na testovanie hypotézy o neznámom parametri μ alebo na odhad intervalu spoľahlivosti pre neznámu premennú μ.

[5]vraj netreba vedieť na skúšku

Veta 3.5.16 (8) *Nech* $(X_{11}, X_{12}, ..., X_{1n_1}) \sim N(\mu_1, \sigma^2)$, *nech* $(X_{21}, X_{22}, ..., X_{2n_2}) \sim N(\mu_2, \sigma^2)$, *a* σ^2 *je neznámy parameter. Potom náhodná premenná*

$$t = \frac{\overline{X_1} - \overline{X_2} - (\mu_1 - \mu_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \; kde \; s^2 = \frac{(s_1^2(n_1 - 1) + s_2^2(n_2 - 1))}{n_1 - n_2 - 2}$$

$$s_1^2 = \tfrac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_{1i} - \overline{X_1})^2 \qquad s_2^2 = \tfrac{1}{n_2 - 1} \sum_{i=1}^{n_2} (X_{2i} - \overline{X_2})^2$$

má Studentovo rozdelenie o $(n_1 + n_2 - 2)$ *st. voľnosti.*

3.6 Fisherovo-Snedecorovo rozdelenie (F-rozdelenie)

Definícia 3.6.1 [8] *Hovoríme, že náhodná premenná* X *má F-S rozdelenie s* (n_1, n_2) *st. voľnosti, ak jej funkcia hustoty je tvaru*

$$f_{n_1,n_2}(X) = \frac{\Gamma(\frac{n_1+n_2}{2})}{\Gamma(\frac{n_1}{2})\Gamma(\frac{n_1}{2})} (\tfrac{n_2}{n_2})^{\frac{n_1}{n_2}} X^{\frac{n_1}{2}-1} (1 + \tfrac{n_1}{n_2} X)^{\frac{-n_1+n_2}{2}}$$
$$x > 0, n_1 \to \infty, n_2 \to \infty$$

Definícia 3.6.2 *Číslo* X_α *nazývame kritickou hodnotou F-S rozdelenia, ak náhodná premenná* X *nadobúda hodnoty väčšie ako* X_α.

$$P(X > X_\alpha) = \int_{X_\alpha}^{\infty} f_n(X) dx = \alpha$$

Veta 3.6.3 (9) [9]*Nech* $(X_{11}, X_{12}, ..., X_{1n_1}) \sim N(\mu_1, \sigma_1^2)$, $(X_{21}, X_{22}, ..., X_{2n_2}) \sim N(\mu_2, \sigma_2^2)$ $\mu_1, \mu_2, \sigma_1, \sigma_2$ *sú neznáme a nech* S_1^2, S_2^2 *sú výberové rozptyly, kde* $S_1^2 > S_2^2$. *Potom náhodná premenná* $F = \frac{S_1^2}{S_2^2}$ *má* $F-S$ *rozdelení o* $(n_1 - 1, n_2 - 1)$ *stupňoch voľnosti. Pričom*

$$S_1^2 = \tfrac{1}{n_1 - 1} \sum_{i=1}^{n_1} (x_{1_i} - \overline{x_1})^2 \qquad S_1^2 = \tfrac{1}{n_2 - 1} \sum_{i=1}^{n_1} (x_{2_i} - \overline{x_2})^2$$

---

[8] tak takúto haluz určite nebude treba vedieť na skúšku :-)

[9] Vetu používame na testovanie hypotézy o rovnosti rozptylov základných súborov. $H_0 : (\sigma_1^2 = \sigma_2^2)$

Príklad 4.3.3 $f(x,\lambda) = \frac{e^{-\lambda}\lambda^j}{j!}$ *- funkcia Poissonovho rozdelenia*
Metódou vierohodnosti odhadujeme λ *Potrebujeme vedieť* $(k_1, k_2, ..., k_n)$.
$L(k,\lambda) = \frac{e^{-\lambda}\lambda^{k_1}}{k_1!}\frac{e^{-\lambda}\lambda^{k_2}}{k_2!}...\frac{e^{-\lambda}\lambda^{k_n}}{k_n!}$
Urobíme $\ln L(k,\lambda) = -\lambda + k_1 \ln\lambda - \ln k_1! - \lambda + k_2 \ln\lambda - \ln k_2! - ... - \lambda + k_n \ln k_n! = -n\lambda + \sum_{i=1}^n k_i \ln\lambda - \sum_{i=1}^n \ln k_i!$
$\frac{d}{d\lambda}\ln L(k,\lambda) = -n + \sum_{i=1}^n k_i \frac{1}{\lambda} = 0 \Rightarrow \sum_{i=1}^n k_i = n\lambda \Rightarrow \hat{\lambda} = \frac{1}{n}\sum_{i=1}^n k_i = \overline{k}$

Definícia 4.3.4 *Nech* $(X_1, X_2, ..., X_n)$ *je náhodný výber zo základného súboru s náhodným rozdelením* $f(x,\Theta)$ *kde* $\Theta = (\Theta_1, \Theta_2, ..., \Theta_n)$. *Potom funkciu náhodného rozdelenia pri realizovaných hodnotách* $(x_1, x_2, ..., x_n)$ *tvaru* $L(x,\Theta) = f(x_1,\Theta), ..., f(x_n,\Theta) = \prod_{i=1}^n f(x_i,\Theta)$ *nazývame funkciou vierohodnosti. Jej prirodzený logaritmus je tvaru* $\ln L(x,\Theta) = \sum_{i=1}^n \ln f(x_i,\Theta)$

Definícia 4.3.5 *Maximálnym vierohodným odhadom nazveme hodnotu vektoru parametrov* $\hat{\Theta} = (\hat{\Theta}_1, \hat{\Theta}_2, ..., \hat{\Theta}_n)$, *ktorý maxim. funkciou vierohodnosti alebo jej prirodzený logaritmus. Odhady* $\hat{\Theta}_i$ *získame riešením systému rovníc*

$$\frac{\partial \ln L(\mathbf{x}, \Theta_1, \Theta_2, ..., \Theta_r)}{\partial \Theta_i} = 0 \textit{ pre } i = 1, 2, ..., r$$

Príklad 4.3.6 $(X_1, X_2, ..., X_n) \sim N(\mu, \sigma^2)$

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

$$\begin{aligned} L(x,\mu,\sigma^2) &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x_1-\mu}{\sigma})^2} . \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x_2-\mu}{\sigma})^2} ... \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x_n-\mu}{\sigma})^2} = \\ &= \frac{1}{\sqrt{\sigma^2}\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x_1-\mu}{\sigma})^2} ... = \\ &= \frac{1}{(\sigma^2)^{\frac{1}{2}}(2\pi)^{\frac{1}{2}}} e^{-\frac{1}{2}(\frac{x_1-\mu}{\sigma})^2} ... \frac{1}{(\sigma^2)^{\frac{1}{2}}(2\pi)^{\frac{1}{2}}} e^{-\frac{1}{2}(\frac{x_n-\mu}{\sigma})^2} = \\ &= \frac{1}{(\sigma^2)^{\frac{1}{2}}(2\pi)^{\frac{1}{2}}} e^{-\frac{1}{2}\sum_{i=1}^n (\frac{x_i-\mu}{\sigma})^2} \end{aligned}$$

$$\ln L(\mathbf{x}, 1, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma) - \frac{1}{2}\sum_{i=1}^n (\frac{x_i-\mu}{\sigma})^2$$

$$\frac{\partial \ln L(x,\mu,\sigma^2)}{\partial\mu} = -\frac{1}{2}L[TODO]\sum_{i=1}^n (\frac{x_i-\mu}{\sigma})(-1) = 0$$
$$\sum_{i=1}^n x_i = \mu n \Rightarrow \hat{\mu} = \frac{1}{n}\sum_{i=1}^n x_i$$

$$\frac{\partial \ln L(x,\hat{\mu},\sigma^2)}{\partial\sigma^2} = -\frac{n}{2}\frac{1}{\sigma^2} + \sum_{i=1}^n (x_i - \hat{\mu})^2 \frac{1}{2\sigma^4} = 0$$
$$= n\sigma^2 + \sum_{i=1}^n (x_i - \hat{\mu})^2 = 0 \Rightarrow \sum_{i=1}^n (x_i - \hat{\mu})^2 = n\sigma^2 \Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \hat{\mu})^2$$

4.4 Intervalové odhady

Veta 4.4.1 (10) *Nech* $(X_1, X_2, ..., X_n) \sim N(\mu, \sigma^2)$, *kde* μ *je neznáme číslo. Nech* $\alpha \in (0,1)$ *a* $n_{1-\frac{\alpha}{2}}$ *je* $(1-\frac{\alpha}{2})$*-kvantilom normálneho rozdelenia.*
$\Phi(u_1 - \frac{\alpha}{2}) = \int_{-\infty}^{u_{1-\frac{\alpha}{2}}} f(x)dx = 1 - \frac{\alpha}{2}$
Potom obojstranný interval spoľahlivosti pre neznámy parameter je
$P(\overline{X} - \frac{\sigma}{\sqrt{u}} u_{1-\frac{\alpha}{2}} \leq \mu \leq \overline{X} + \frac{\sigma}{\sqrt{u}} u_{1-\frac{\alpha}{2}}) = 1 - \alpha$

Príklad 4.4.2 *Uvažujme nasl. pravdep. rovnicu*
$P(|U| > u) = \alpha$
Podľa vety č. (4) vieme, že $U - \frac{\overline{X}-\mu}{\sigma}\sqrt{u} \sim N(0,1)$. *Použijeme ju k odvodeniu obojstranného intervalu spoľahlivosti pre* μ.
$P(|U| \leq u) = 1 - \alpha$

- $t_{1-\frac{\alpha}{2}} \geq \frac{\overline{X}-\mu}{S}\sqrt{n} \Longrightarrow \mu \geq \overline{X} - \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}}$

□

Príklad 4.4.6 *Máme dané realizácie daného výpočtu* $\sim N\left(\mu, \sigma^2\right)$ $(2, 12, 8, 11, 22, -4, 10, 25, 17, -5)$. *Máme vypočítať* 95%-*OIS pre* μ.

Riešenie 4.4.7

- $n = 10$
- $\overline{X} = 9,5$
- $\alpha = 0,05$
- $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = 103,4 \Longrightarrow S = 10,17$
- $t_{1-\frac{\alpha}{2}} = t_{1-\frac{0,05}{2}} = t_{0,975}(9) = 2,362$ *(v tabuľkách)*

  ďalej si to už hádam každý zráta sám :-)

Teraz sa pozrieme na jednostranné IS

LIS(príp. PIS) :-)	PIS
$P\left(t > t_{1-\alpha}\right) = \alpha$	$P\left(t < t_{1-\alpha}\right) = \alpha$
$P\left(t \leq t_{1-\alpha}\right) = 1-\alpha$	$P\left(t \geq t_{1-\alpha}\right) = 1-\alpha$
$\frac{\overline{X}-\mu}{S}\sqrt{n} \leq t_{1-\alpha} \Longrightarrow \mu \geq \overline{X} - \frac{S}{\sqrt{n}} t_{1-\alpha}$	$\frac{\overline{X}-\mu}{S}\sqrt{n} \geq t_{1-\alpha} \Longrightarrow \mu \leq \overline{X} + \frac{S}{\sqrt{n}} t_{1-\alpha}$

Veta 4.4.8 (12) *Nech* $(X_1, X_2, \ldots, X_n) \sim N\left(\mu, \sigma^2\right)$, *kde* μ, σ^2 *sú neznáme parametre. Nech* $\chi \in (0,1)$,$\chi_{\frac{\alpha}{2}}$, $\chi_{1-\frac{\alpha}{2}}$ *sú kritické hodnoty* χ^2*- rozdelenia. Potom OIS pre neznámy parameter* σ^2 *je tvaru*

$$P\left(\overline{X} - \frac{(n-1)\,S^2}{\chi^2_{\frac{\alpha}{2}}} \leq \sigma^2 \leq \overline{X} + \frac{(n-1)\,S^2}{\chi^2_{1-\frac{\alpha}{2}}} t_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

Dôkaz: Podľa vety(5) platí, že náhodná premenná $\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ stupňa voľnosti

- $P\left(\chi^2_{1-\frac{\alpha}{2}} \leq \chi^2 \leq \chi^2_{\frac{\alpha}{2}}\right) = 1-\alpha$
- $P\left(\chi^2_{1-\frac{\alpha}{2}} \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi^2_{\frac{\alpha}{2}}\right) = 1-\alpha$
- $\chi^2_{1-\frac{\alpha}{2}} \leq \frac{(n-1)S^2}{\sigma^2} \Longrightarrow \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}}$
- $\chi^2_{\frac{\alpha}{2}} \geq \frac{(n-1)S^2}{\sigma^2} \Longrightarrow \sigma^2 \geq \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}}$

5 Testovanie špecifických hypotéz

5.1 Štatistická hypotéza

- predpoklady kladené na rozdelenie pravdepodobností náhodných premenných alebo na ich parametre
- overovanie správnosti - *testovanie*
- hypotéza, ktorej platnosť overujeme = *testovaná (nulová) hypotéza*
- proti testovanej hypotéze kladieme *alternatívnu hypotézu*

5.2 Testovacie kritérium, chyba 1. a 2. druhu

- na overovanie používame štatistiky $g(x)$ - testovacie kritérium[12]
- obor hodnôt, ktoré môže testovacie kritérium nadobúdať je rozdelený na dva disjunktné obory
  - kritický W_α(obor zamietnutia testovanej hypotézy)
  - doplnkový $W_{\alpha'}$ (obor prijatia testovanej hypotézy)
- ak hodnota testovaného kritéria vypočítaná na základe experimentu nadobúda hodnotu z kritického oboru, tak testovanú hypotézu H_0 zamietneme, ak z doplnkového oboru, testovaciu hypotézu prijímame

Definícia 5.2.1 *Chybu, ktorá spočíva v zamietnutí testovanej hypotézy, hoci je pravdivá, nazývame chybou prvého druhu, tzn. platí:*

$$P(G_0(x) \in W_\alpha / H_0) = \alpha$$

H_0 je pravdivá, ak hodnoty sú z doplnkového oboru.
H_1 je pravdivá, ak hodnoty sú z kritického oboru.

Definícia 5.2.2 *Chyba, ktorá spočíva v zamietnutí alternatívnej hypotézy nazývame chybou druhého druhu.*

$$P(G_0(x) \in W'_\alpha / H_1) = P(G_0(x) \notin W_\alpha / H_1) = \beta$$

Prakticky, zvolíme chybu prvého druhu α a použijeme také testy, aby chyba druhého druhu bola čo najmenšia.

[12] niektorá z viet 4-9(10)

Ak platí

$$U_0 = \frac{\overline{X} - \mu_0}{\sigma}\sqrt{n} > u_{1-\alpha}$$

tak testovanú hypotézu $H_0(\mu = \mu_0)$ zamietame a prijímame alternatívnu hypotézu $H_1(\mu \neq \mu_0)$..

Ak platí

$$|U_0| = \frac{|\overline{X} - \mu_0|}{\sigma}\sqrt{n} \leq u_{1-\alpha}$$

tak testovanú hypotézu H_0 prijímame s rizikom α.

b) Majme náhodný výber $N(\mu, \sigma^2)$, kde μ je neznámy parameter a σ^2 je známe. Testujeme hypotézu $H_0 : \mu = \mu_0$, $H_1 : \mu < \mu_0$. Za testovacie kritérium zvolíme náhodnú premennú

$$U = \frac{\overline{X} - \mu}{\sigma}\sqrt{n} \sim N(0,1)$$

podľa vety číslo 4. Kritická oblasť pre jednostrannú alternatívnu hypotézu je daná nerovnosťou

$$P(U < -u_{1-\alpha}) = \alpha$$

Ak platí

$$U_0 = \frac{\overline{X} - \mu_0}{\sigma}\sqrt{n} < -u_{1-\alpha}$$

tak testovanú hypotézu $H_0(\mu = \mu_0)$ zamietame a prijímame alternatívnu hypotézu $H_1(\mu \neq \mu_0)$..

Ak platí

$$|U_0| = \frac{|\overline{X} - \mu_0|}{\sigma}\sqrt{n} \geq -u_{1-\alpha}$$

tak testovanú hypotézu H_0 prijímame s rizikom α.

Príklad 5.3.1 *Treba overiť, či pevnosť ocele zodpovedá predpokladanej hodnote 66 jednotiek t.j.* $\mu_0 = 66, \sigma = 3,011$. *Na základe realizácie náhodného výberu budeme na hladine významnosti* $\alpha = 0,01$ *študovať hypotézu* $H_0 : (\mu = 66), H_1 : (\mu \neq 66)$

Riešenie 5.3.2 *Vieme, že* $\mu_0 = 66, \sigma = 3,011, n = 10$, *na základe experimentu sme zistili, že* $\overline{X} = 67,8$.
Musíme zistiť testovaciu hodnotu $U_0 = \frac{\overline{X}-\mu_0}{\sigma}\sqrt{n} = \frac{67,8-66}{3,011}\sqrt{10} = 1,89$
Hľadáme v tabuľkách kvantil $u_{1-\frac{\alpha}{2}} = u_{1-\frac{0,01}{2}} = u_{0,995} = 2,57$
Keďže testovacia hodnota $1,89$ *leží v doplnkovom obore, príjmeme hypotézu* H_0

5.3.3 Poznáme σ_1^2, σ_2^2

Majme $(X_{11}, X_{12}, ..., X_{1n_1}) \sim N(\mu_1, \sigma_1^2)$,$(X_{21}, X_{22}, ..., X_{2n_1}) \sim N(\mu_2, \sigma_2^2)$, kde σ_1^2, σ_2^2 sú dané. Testujme hypotézu $H_0 : (\mu = \mu_0), H_1 : (\mu \neq \mu_0)$. Za testovacie kritérium si zvolíme náhodnú premennú z vety číslo 7 majúcu normálne rozdelenie:

$$U = \frac{\overline{X_1} - \overline{X_2} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

Po zvolení α vieme nájsť kvantily. Kritický obor je určený nasledovnou nerovnosťou

$$P(|U| > u_{1-\frac{\alpha}{2}}) = \alpha$$

Testovacia hodnota je (rozdiel $\mu_1 - \mu_2$ vypadne, pretože podľa hypotézy je to 0)

$$U_0 = \frac{\overline{X_1} - \overline{X_2} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Ak $|U_0| > u_{1-\frac{\alpha}{2}}$, H_0 zamietame, inak H_0 prijímame.

a) jednostranná hypotéza

Majme $(X_{11}, X_{12}, ..., X_{1n_1}) \sim N(\mu_1, \sigma_1^2)$,$(X_{21}, X_{22}, ..., X_{2n_1}) \sim N(\mu_2, \sigma_2^2)$, kde

σ_1^2, σ_2^2 sú dané. Testujme hypotézu $H_0 : (\mu = \mu_0), H_1 : (\mu \neq \mu_0)$. Za testovacie kritérium si zvolíme náhodnú premennú z vety číslo 7 majúcu normálne rozdelenie:

$$U = \frac{\overline{X_1} - \overline{X_2} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

Po zvolení α vieme nájsť kvantily. Kritický obor je určený nasledovnou nerovnosťou

$$P(U > u_{1-\alpha}) = \alpha$$

Testovacia hodnota je (rozdiel $\mu_1 - \mu_2$ vypadne, pretože podľa hypotézy je to 0)

$$U_0 = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Ak $U_0 > u_{1-\alpha}$, H_0 zamietame, inak H_0 prijímame.

b) obdobne pre druhú stranu

Majme $(X_{11}, X_{12}, ..., X_{1n_1}) \sim N(\mu_1, \sigma_1^2)$,$(X_{21}, X_{22}, ..., X_{2n_1}) \sim N(\mu_2, \sigma_2^2)$, kde

σ_1^2, σ_2^2 sú dané. Testujme hypotézu $H_0 : (\mu = \mu_0), H_1 : (\mu \neq \mu_0)$. Za testovacie kritérium si zvolíme náhodnú premennú z vety číslo 7 majúcu normálne rozdelenie:

$$U = \frac{\overline{X_1} - \overline{X_2} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

5.3.5 Nepoznáme μ, σ^2

Nech $(X_{11}, X_{12}, ..., X_{1n_1}) \sim N(\mu, \sigma^2)$, kde μ, σ^2 sú neznáme. Testujeme hypotézu $H_0 : (\sigma^2 = \sigma_0^2), H_1 : (\sigma^2 \neq \sigma_0^2)$. Za testovacie kritérium si zvolíme náhodnú premennú

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$

podľa vety číslo 5 je toto χ^2-rozdelenie s n-1 stupňami voľnosti. Kritický obor je daný nerovnosťami

$$\chi^2 \leq \chi^2_{1-\frac{\alpha}{2}}$$
$$\chi^2 \geq \chi^2_{1-\frac{\alpha}{2}}$$

Vypočítame hodnotu $\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2}$, ak $\chi_0^2 = \frac{(n-1)S_1^2}{\sigma_0^2} \leq \chi^2_{1-\frac{\alpha}{2}}$ alebo $\chi_0^2 = \frac{(n-1)S_1^2}{\sigma_0^2} \geq \chi^2_{\frac{\alpha}{2}}$, tak hypotézu zamietame, inak ju prijímame.

a) Testujeme hypotézu $H_0 : (\sigma^2 = \sigma_0^2), H_1 : (\sigma^2 > \sigma_0^2)$. Kritická oblasť je daná nerovnicou

$$\chi_0^2 > \chi_\alpha^2$$

b) Testujeme hypotézu $H_0 : (\sigma^2 = \sigma_0^2), H_1 : (\sigma^2 < \sigma_0^2)$. Kritická oblasť je daná nerovnicou

$$\chi_0^2 > \chi^2_{1-\alpha}$$

Príklad 5.3.7 *Pri meraní teploty steny sme namerali hodnoty* $0,61,\ 0,62, \ldots$ $,0,59$. *Treba testovať hypotézu* $H_0 : (\sigma^2 = \sigma_0^2), H_1 : (\sigma^2 > \sigma_0^2)$ *s mierou rizika* $\alpha = 0,003$ *a* $\sigma_0^2 = 0,003$, *počet nameraných hodnôt je* 8.

$$\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{7.0,00064}{0,003} = 1,49$$

testovacie kritériá sú

$$\chi^2_{1-\frac{\alpha}{2}} = 1,69$$
$$\chi^2_{\frac{\alpha}{2}} = 16,0$$

H_0 *je zamietnuté, lebo* $1,49$ *leží mimo intervalu* $1,69$ *a* $16,0$.