M3121 Pravděpodobnost a statistika I

M4122 Pravděpodobnost a statistika II

(prednášky)

1. Pravdepodobnostný priestor

M3121 je základný kurz teórie pravdepodobnosti, na ktorý nadväzuje M4122, v ktorom sú základy matematickej štatistiky.

Skúšať sa bude látka obidvoch semestrov naraz v lete. Cvičenia sú veľmi dôležité.

Podmienky získania kreditov v ZS:

– maximálne 3 neospravedlnené neúčasti na cvičeniach a súčasne

– zisk minimálne 10 bodov z písomky (ohodnotená je maximálne 20 bodmi).

Podmienky získania kreditov v LS: upresní prednášajúci.

História (stručne) teórie pravdepodobnosti sa nájde na
www.math.muni.cz/∼budikova/prf/historie.pdf

Literatúra:

Dupač, V., Hušková, M., Pravděpodobnost a matematická statistika, Karolinum, Praha, 2001.

Zvára, K., Štěpán, J., Pravděpodobnost a matematická statistika, Matfyzpress, Praha, 2001.

Teória pravdepodobnosti je matematická disciplína, ktorá modeluje a popisuje náhodný pokus – pokus, ktorého výsledok dopredu nepoznáme. Teda výsledok pokusu nie je jednoznačne určený podmienkami, za ktorých je realizovaný. Napr. hod kockou. Pokusy, ktorých výsledok je jednoznačne daný podmienkami sa volajú *deterministické*. My budeme popisovať tzv. *stochastické* pokusy. Pritom presnejšie nás zaujímajú také náhodné pokusy, pri ktorých je náhoda akási "regulárna". Konkrétne ak A je ľubovoľný sledovaný náhodný jav, tak požadujeme, aby vykazoval pri opakovanej nezávislej realizácii náhodného pokusu tzv. štatistickú stabilitu, t.j. aby relatívna početnosť

$$f_n(A) = \frac{n_A}{n}$$

výskytu javu A v postupnosti n nezávislých pokusov (pričom n_A je počet nastatí javu A) sa príliš nemenil a s rastúcim n mal "tendenciu" držať sa nejakej konštanty. Obrazne (nepresne) zapísané $\lim_{n\to\infty} f_n(A) = p(A)$. Toto neplatí (úplne) napr. o futbalovom zápase, o predpovedi počasia, atď.

Budeme teda budovať matematickú teóriu – model náhodného (štatisticky stabilného) pokusu.

Značenie:

Ω ... Množina všetkých možných "najjemnejších" výsledkov náhodného pokusu, ktoré ešte treba rozlišovať. Predpokladáme, že vždy Ω je neprázdna. Ω voláme **priestor elementárnych javov**.

ω ... Elementárny jav, prvok Ω; Ω môže byť konečná, spočítateľná aj nespočítateľná; ω je "najjemnejší" výsledok náhodného pokusu.

$A, B, A_1, A_2, ..., A_n$ ... náhodné javy (udalosti)

$\emptyset$ ... nemožný jav

Ω ... istý jav

$A \cup B$ ... jav, ktorý nastane ak nastane alebo A alebo B

$A \cap B$ ... jav, ktorý nastane ak nastane aj A aj B

$A - B$ ... jav, ktorý nastane ak nastane A a nenastane B

$\overline{A} = A^c = \Omega - A$ ... opačný jav k javu A

$\bigcup_{i=1}^{n} A_i$ ... nastane, ak nastane aspoň jeden z javov $A_1, ..., A_n$

$\bigcup_{i=1}^{\infty} A_i$ ... nastane, ak nastane aspoň jeden z javov $A_1, A_2, ...$

$\bigcap_{i=1}^{n} A_i$ ... nastane, ak nastanú všetky javy $A_1, ..., A_n$

$\bigcap_{i=1}^{\infty} A_i$ ... nastane, ak nastanú všetky javy $A_1, A_2, ...$

$\exp \Omega = 2^{\Omega}$ ... systém všetkých podmnožín Ω

$A \subset B$ ... ak nastal jav A, tak nastane jav B (A implikuje B)

Pri náhodnom pokuse okrem priestoru elementárnych javov Ω musíme mať zadaný (popísaný) aj systém náhodných javov.

Definícia 1.1. Nech Ω je ľubovoľná neprázdna množina. Neprázdny systém $\mathcal{A}$ podmnožín množiny Ω sa nazýva $\sigma-$algebra, ak platí

$$A \in \mathcal{A} \Longrightarrow A^c \in \mathcal{A} \tag{1.1 (i)}$$

$$A_1, A_2, ... \in \mathcal{A} \Longrightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A} \qquad (\sigma-\text{aditivita}). \tag{1.2 (ii)}$$

Ukazuje sa rozumná požiadavka odrážajúca naše skúsenosti, aby systém náhodných javov $\mathcal{A}$ v popisovanom náhodnom pokuse bol $\sigma-$algebrou podmnožín množiny elementárnych javov Ω.

Dvojicu $(\Omega, \mathcal{A})$ nazývame **javové pole** a ľubovoľný prvok $A \in \mathcal{A}$ nazývame **náhodný jav** (vzhľadom k $(\Omega, \mathcal{A})$).

Poznámka $(\Omega, \mathcal{A})$ s volá aj **merateľný priestor**.

Poznámka $\omega-$ elementárny jav nie je náhodným javom, ale $\{\omega\}$ ako podmnožina Ω je náhodným javom ak patrí do $\mathcal{A}$.

Povieme, že náhodný jav A nastal, ak (elementárny) výsledok pokusu bol ω, pričom $\omega \in A$. S náhodnými javmi narábame preto ako s množinami. Platia tu de Morganove vzorce

$$\overline{\bigcup_{i\geq 1} A_i} = \bigcap_{i\geq 1} \overline{A_i} \tag{1.3}$$

$$\overline{\bigcap_{i\geq 1} A_i} = \bigcup_{i\geq 1} \overline{A_i}. \tag{1.4}$$

(Dôkaz (1.3): $\omega \in \overline{\bigcup_{i\geq 1} A_i} \Longleftrightarrow \omega \notin \bigcup_{i\geq 1} A_i \Longleftrightarrow \forall\ i \geq 1$ platí $\omega \notin A_i \Longleftrightarrow \forall\ i \geq 1$ platí $\omega \in \overline{A_i} \Longleftrightarrow \omega \in \bigcap_{i\geq 1} \overline{A_i}$.

Dôkaz (1.4) si urobte sami.)

Veta 1.1. Nech $(\Omega, \mathcal{A})$ je javové pole. Potom platí

$$\Omega \in \mathcal{A},\ \emptyset \in \mathcal{A}, \tag{1.5}$$

a pre ľubovoľné prirodzené n a $A_1, ..., A_n, A, B \in \mathcal{A}$

$$\bigcup_{i=1}^{n} A_i \in \mathcal{A},\ \bigcap_{i=1}^{n} A_i \in \mathcal{A},\ A - B \in \mathcal{A}, \tag{1.6}$$

a tiež

$$A_1, A_2, ... \in \mathcal{A} \Longrightarrow \bigcap_{n=1}^{\infty} A_n \in \mathcal{A}. \tag{1.7}$$

Dôkaz: $\mathcal{A}$ je neprázdny systém, teda $\exists\ A \in \mathcal{A}$. Z (1.1) vyplýva, že $\overline{A} \in \mathcal{A}$. Z (1.2) vyplýva, že ak $A = A_1 \in \mathcal{A}, \overline{A} = A_2 \in \mathcal{A}, A_3 = A_1, A_4 = A_1, ...,$ tak $A_1 \cup A_2 \cup A_3 \cup ... = \Omega \in \mathcal{A}$. Z (1.1) tiež $\overline{\Omega} = \emptyset \in \mathcal{A}$.

Ďalej ak $A_1, A_2, ..., A_n \in \mathcal{A}$ a tiež $\emptyset = A_{n+1}, \emptyset = A_{n+2}, ... \in \mathcal{A}$, tak z (1.2) $A_1 \cup A_2 \cup ... A_n \cup \emptyset \cup ... = \bigcup_{i=1}^{n} A_i \in \mathcal{A}$.

Ak $A_1, A_2, ..., A_n \in \mathcal{A}$, teda $\overline{A_1}, \overline{A_2}, ..., \overline{A_n} \in \mathcal{A}$, a $\bigcup_{i=1}^{n} \overline{A_i} \in \mathcal{A}$, ale podľa de Morganovho pravidla (1.4) je $\bigcup_{i=1}^{n} \overline{A_i} = \overline{\bigcap_{i=1}^{n} A_i}$ a preto $\overline{\bigcap_{i=1}^{n} A_i} \in \mathcal{A}$, ale podľa (1.1) $\overline{\overline{\bigcap_{i=1}^{n} A_i}} \in \mathcal{A}$, pričom $\overline{\overline{\bigcap_{i=1}^{n} A_i}} = \bigcap_{i=1}^{n} A_i$.

Teraz nech $A, B \in \mathcal{A}$. Z (1.1) $\overline{B} \in \mathcal{A}$ a z množinovej rovnosti $A - B = A \cap \overline{B}$ dostávame, že $A - B \in \mathcal{A}$.

Nech $A_1, A_2, ... \in \mathcal{A}$. Preto aj $\overline{A}_1, \overline{A}_2, ... \in \mathcal{A}$, teda $\bigcup_{n=1}^{\infty} \overline{A_n} \in \mathcal{A}$ a pomocou (1.1) a de Morganových pravidiel aj $\overline{\bigcup_{n=1}^{\infty} \overline{A_n}} = \bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$. ♣

Definícia 1.2. Majme postupnosť náhodných javov $\{A_n\}_{n=1}^{\infty}$. Hornou limitou postupnosti javov $\{A_n\}_{n=1}^{\infty}$ nazývame množinu všetkých $\omega \in \Omega$, ktoré patria do nekonečne veľa javov A_n. Označujeme ju $\limsup_{n\to\infty} A_n$. (Inak povedané $\limsup_{n\to\infty} A_n$ nastane práve vtedy ak nastane nekonečne veľa javov A_n.)

Definícia 1.3. Majme postupnosť náhodných javov $\{A_n\}_{n=1}^{\infty}$. Dolnou limitou postupnosti javov $\{A_n\}_{n=1}^{\infty}$ nazývame množinu všetkých $\omega \in \Omega$, ktoré patria do všetkých A_n s výnimkou konečného počtu týchto javov. Označujeme ju $\liminf_{n\to\infty} A_n$. (Inak povedané $\liminf_{n\to\infty} A_n$ nastane práve vtedy ak nastanú všetky javy A_n s výnimkou konečne veľa týchto javov.)

Lema 1.1. Ak $\{A_n\}_{n=1}^{\infty}$ je postupnosť náhodných javov na $(\Omega, \mathcal{A})$, tak $\liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} A_n$.

Dôkaz: Zrejmý z definícií 1.2 a 1.3.

Veta 1.2. Ak $\{A_n\}_{n=1}^{\infty}$ je postupnosť náhodných javov na $(\Omega, \mathcal{A})$, tak platí

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k, \tag{1.8}$$

$$\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k, \tag{1.9}$$

$$\overline{\limsup_{n\to\infty} A_n} = \liminf_{n\to\infty} \overline{A_n}. \tag{1.10}$$

Dôkaz: $\omega \in \limsup_{n\to\infty} A_n \Longleftrightarrow \omega$ patrí do nekonečne veľa javov $A_n \Longleftrightarrow$ pre $\forall\ n \geq 1\ \exists\ k \geq n$, že $\omega \in A_k \Longleftrightarrow \forall\ n \geq 1$ je $\omega \in \bigcup_{k=n}^{\infty} A_k \Longleftrightarrow \omega \in \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$.

$\omega \in \liminf_{n\to\infty} A_n \Longleftrightarrow \omega$ patrí do každej A_n s výnimkou konečného počtu $A_i \Longleftrightarrow \exists\ n \geq 1$, že $\omega \in \bigcap_{k=n}^{\infty} A_k \Longleftrightarrow \omega \in \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$.

$\omega \in \overline{\limsup_{n\to\infty} A_n} \Longleftrightarrow$ neplatí, že ω patrí do nekonečne veľa javov $A_n \Longleftrightarrow$ neplatí, že $\forall\ n \geq 1\ \exists\ k \geq n$, že $\omega \in A_k \Longleftrightarrow \exists\ n \geq 1\ \forall\ k \geq n\ \omega \notin A_k \Longleftrightarrow \exists\ n \geq 1\ \forall\ k \geq n\ \omega \in \overline{A_k} \Longleftrightarrow \exists\ n \geq 1$, že $\omega \in \bigcap_{k=n}^{\infty} \overline{A_k} \Longleftrightarrow \omega \in \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \overline{A_k} = \liminf_{n\to\infty} \overline{A_n}$. ♣

Definícia 1.3a. Majme postupnosť náhodných javov $\{A_n\}_{n=1}^{\infty}$. Povieme, že postupnosť $\{A_n\}_{n=1}^{\infty}$ má limitu A, ak $A = \limsup_{n\to\infty} A_n = \liminf_{n\to\infty} A_n$. Označujeme $A = \lim_{n\to\infty} A_n$.

Poznámka. Ak $A_1, A_2, ... \in \mathcal{A}$ a $\exists\ A = \limsup_{n\to\infty} A_n = \liminf_{n\to\infty} A_n =$ $\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k \in \mathcal{A}$ (teda limita je z $\mathcal{A}$). Samozrejme z Vety 1.2 vyplýva, že aj $\limsup_{n\to\infty} A_n \in \mathcal{A}$ a $\liminf_{n\to\infty} A_n \in \mathcal{A}$.

Veta 1.3. Nech $\{A_n\}_{n=1}^{\infty}$ je postupnosť náhodných javov na $(\Omega, \mathcal{A})$. Ak $A_1 \subseteq A_2 \subseteq ....$ Potom $\exists\ \lim_{n\to\infty} A_n$ a platí $\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$.

Dôkaz: $\bigcap_{k=n}^{\infty} A_k = A_n$, preto $\omega \in \liminf_{n\to\infty} A_n \iff \omega \in \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k = \bigcup_{n=1}^{\infty} A_n$, čiže $\liminf_{n\to\infty} A_n =$ $\bigcup_{n=1}^{\infty} A_n$.

Aj $\limsup_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$, lebo z Lemy 1.1 $\bigcup_{n=1}^{\infty} A_n = \liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} A_n$ a naopak ak $\omega \in \limsup_{n\to\infty} A_n \iff \omega$ patrí do nekonečne veľa $A_n \iff \omega \in \bigcup_{n=1}^{\infty} A_n$, čiže $\limsup_{n\to\infty} A_n \subseteq \bigcup_{n=1}^{\infty} A_n$.

Preto $\exists\ \lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$. ♣

Veta 1.4. Nech $\{A_n\}_{n=1}^{\infty}$ je postupnosť náhodných javov na $(\Omega, \mathcal{A})$. Ak $A_1 \supseteq A_2 \supseteq ....$ Potom $\exists\ \lim_{n\to\infty} A_n$ a platí $\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n$.

Dôkaz: $\bigcup_{k=n}^{\infty} A_k = A_n$, preto $\omega \in \limsup_{n\to\infty} A_n \iff \omega \in \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k = \bigcap_{n=1}^{\infty} A_n$, čiže $\limsup_{n\to\infty} A_n =$ $\bigcap_{n=1}^{\infty} A_n$.

Aj $\liminf_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n$, lebo z Lemy 1.1 $\liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} A_n$ $= \bigcap_{n=1}^{\infty} A_n$ a naopak ak $\omega \in \bigcap_{n=1}^{\infty} A_n \iff \omega$ patrí do všetkých $A_n \iff \omega \in \liminf_{n\to\infty} A_n$, čiže $\bigcap_{n=1}^{\infty} A_n \subseteq$ $\liminf_{n\to\infty} A_n$.

Preto $\exists\ \lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n$. ♣

Veta 1.5. Nech $\mathcal{S}$ je neprázdny systém podmnožín Ω. potom existuje množinová $\sigma-$algebra $\sigma(\mathcal{S})$ taká, že platí

(i) $\mathcal{S} \subseteq \sigma(\mathcal{S})$,

(ii) ak je $\mathcal{A}^*$ množinová $\sigma-$algebra taká, že $\mathcal{S} \subseteq \mathcal{A}^*$, tak $\sigma(\mathcal{S}) \subseteq \mathcal{A}^*$.

Dôkaz: Položme $\sigma(\mathcal{S})$ prienik množinových $\sigma-$algebier obsahujúcich $\mathcal{S}$. Potom $\mathcal{S} \subseteq \sigma(\mathcal{S})$ a zrejme je $\sigma(\mathcal{S})$ aj $\sigma-$algebrou. ♣

Definícia 1.4. Nech $\mathcal{S}$ je neprázdny systém podmnožín Ω, $\sigma(\mathcal{S})$ je prienik množinových $\sigma-$algebier obsahujúcich $\mathcal{S}$. $\sigma(\mathcal{S})$ sa nazýva *minimálna množinová $\sigma-$algebra generovaná systémom $\mathcal{S}$.*

Borelovské množiny.

Položme

$\Omega = (-\infty, \infty) = \mathtt{R}$,

$\mathcal{S} = \{(-\infty, x >,\quad x \in \mathtt{R}\} \subseteq \exp\Omega = \exp \mathtt{R}$

Minimálna množinová $\sigma-$algebra $\sigma(\mathcal{S}) = \mathcal{B}$ generovaná systémom $\mathcal{S}$ sa volá *borelovská* (množinová) $\sigma-$algebra v $\mathtt{R}$. Jej prvky sa nazývajú *borelovské množiny*.

Poznamenávame len, že borelovská $\sigma-$algebra v $\mathtt{R}$ je totožná aj s minimálnou množinovou $\sigma-$algebrou generovanou systémom množín S všetkých intervalov tvaru $< a, b)$, kde $a < b$ (pozri napr. Riečan, B., O pravdepodobnosti a miere, Alfa, Bratislava, 1972, str. 46).

Analogicky definujeme $\mathcal{B}^n$.

$\Omega = \mathtt{R}^n$,

$\mathcal{S} = \{(-\infty, x_1 > \times ... \times (-\infty, x_n >, \quad x_1, ..., x_n \in \mathtt{R}\}$,

$\sigma(\mathcal{S}) = \mathcal{B}^n$ je *borelovská* (množinová) $\sigma-$algebra v $\mathtt{R}^n$.

Definícia 1.5. (Axiomatická definícia pravdepodobnosti.) Nech $(\Omega, \mathcal{A})$ je javové pole a P reálna množinová funkcia definovaná na $\mathcal{A}$ s vlastnosťami

(i) $P(\Omega) = 1$ (normovaná)

(ii) $\forall\ A \in \mathcal{A} \quad P(A) \geq 0$ (nezáporná)

(iii) ak $\{A_n\}_{n=1}^{\infty}$ je postupnosť po dvoch disjunktných (nezlučiteľných) náhodných javov (t.j. $\forall\ n\ A_n \in \mathcal{A}: \quad A_n \cap A_m = \emptyset$ pre $n \neq m$), tak $P\left(\bigcup_{n=1}^{\infty}\right) = \sum_{n=1}^{\infty} P(A_n)$ ($\sigma-$aditívna).

Potom funkciu P nazývame *pravdepodobnosťou* (na $\mathcal{A}$) a trojicu $(\Omega, \mathcal{A}, P)$ *pravdepodobnostným priestorom.*

Poznámka. Axiomatickú definíciu pravdepodobnosti a pravdepodobnostný priestor zaviedol N.A.Kolmogorov v roku 1933.

Poznámka. Pravdepodobnostný priestor je matematickým modelom (regulárneho) náhodného pokusu.

Príklad 1.1. Nech Ω je konečná množina elementárnych javov, t.j. $\Omega = \{\omega_1, ..., \omega_n\}, \mathcal{A} = \exp \Omega$. Pre $A = \{\omega_{i_1}, ..., \omega_{i_k}\} \subseteq \Omega$ nech $P(A) = \sum_{j=1}^{k} P(\{\omega_{i_j}\})$, pričom $\forall\ i\ P(\{\omega_i\}) \geq 0$, $\sum_{i=1}^{n} P(\{\omega_i\}) = 1$. Potom $(\Omega, \mathcal{A}, P)$ je pravdepodobnostný priestor.

Špeciálne: Ak v Príklade 1.1 je $P(\{\omega_i\}) = \frac{1}{n}$ pre $i = 1, 2, ..., n$, tak hovoríme o klasickom pravdepodobnostnom pokuse (klasickej definícii pravdepodobnosti, klasickom pravdepodobnostnom priestore), pričom

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

($|A|$ je počet elementárnych javov v A).

Váhová definícia pravdepodobnosti: Nech Ω je nanajvýš spočítateľná množina, teda $\Omega = \{\omega_1, ..., \omega_n, ...\}$, $\mathcal{A} = \exp \Omega$, $P(A) = \sum_{\omega_{i_j} \in A} P(\{\omega_{i_j}\})$, pričom $\forall\ n\ P(\{\omega_n\}) = p_n \geq 0$ a $\sum_{n=1}^{\infty} P(\{\omega_n\}) = 1$.

Geometrická definícia pravdepodobnosti: Nech $\Omega \in \mathcal{B}^n$ je borelovská množina, ktorej Lebesgueova miera $\mu(\Omega)$ je konečná a kladná, $\mathcal{A} = \mathcal{B}^n(\Omega)$ (systém všetkých borelovských podmnožín Ω), pravdepodobnosť $P(A) = \frac{\mu(A)}{\mu(\Omega)}$ pre $A \in \mathcal{A}$.

2. Vlastnosti pravdepodobnosti

Veta 2.1. Nech $(\Omega, \mathcal{A}, P)$ je pravdepodobnostný priestor. Potom pravdepodobnosť P má nasledujúce vlastnosti:

(i) $P(\emptyset) = 0$

(ii) $A, B \in \mathcal{A},\ A \cap B = \emptyset \implies P(A \cup B) = P(A) + P(B)$

(iii) $A, B \in \mathcal{A},\ A \subset B \implies P(B - A) = P(B) - P(A)$ (subtraktívnosť)

(iv) $A, B \in \mathcal{A},\ A \subset B \implies P(A) \leq P(B)$ (monotónnosť)

(v) $A \in \mathcal{A} \implies 0 \leq P(A) \leq 1$

(vi) $A \in \mathcal{A} \implies P(\overline{A}) = 1 - P(A)$

(vii) $A, B \in \mathcal{A} \implies P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(viii) $A_1, ..., A_n \in \mathcal{A} \implies P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) -$
$- \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} P(A_i \cap A_j) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} P(A_i \cap A_j \cap A_k) + ... +$
$+(-1)^{n+1} P(A_1 \cap A_2 \cap ... \cap A_n)$

(ix) $A_1, ..., A_n \in \mathcal{A} \implies P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i)$

Dôkaz:

(i) $P(\Omega) = P(\Omega \cup \emptyset \cup ...) = P(\Omega) + P(\emptyset) + ... = 1 \implies P(\emptyset) = 0$;

(ii) $A, B \in \mathcal{A},\ A \cap B = \emptyset \implies P(A \cup B \cup \emptyset \cup ...) = P(A) + P(B) + P(\emptyset) + ... = P(A) + P(B)$;

(iii), (iv) $A, B \in \mathcal{A},\ A \subset B \implies B = A \cup (B - A)$ (nezlučiteľné). Teda $P(B) = P(A) + P(B - A)$ a preto $P(B - A) = P(B) - P(A)$, ale aj $P(A) = P(B) - P(B - A)$. Keďže $P(B - A) \geq 0$, je $P(B) \geq P(A)$;

(v) $A \in \mathcal{A}, \emptyset \in \mathcal{A}, \emptyset \subset A \subset \Omega \implies$ (z (i),(iv)) $0 = P(\emptyset) \leq P(A) \leq P(\Omega) = 1$;

(vi) $A \in \mathcal{A}, A \cup \overline{A} = \Omega \implies$ (z (ii)) $1 = P(\Omega) = P(A \cup \overline{A}) = P(A) + P(\overline{A})$, čiže $P(\overline{A}) = 1 - P(A)$;

(vii) $A, B \in \mathcal{A}$, teda sa dá písať $A \cup B = [A - (A \cap B)] \cup (A \cap B) \cup [B - (A \cap B)]$ (disjunktné) $\implies P(A \cup B) = P(A - (A \cap B)) + P(A \cap B) + P(B - (A \cap B)) =$ (z (iii)) $P(A) - P(A \cap B) + P(A \cap B) + P(B) - P(A \cap B) = P(A) + P(B) - P(A \cap B)$;

(viii) indukciou pomocou (vii) (pozri napr. Riečan, B., O pravdepodobnosti a miere, Alfa, Bratislava, 1972)

(ix) $P\left(\bigcup_{i=1}^n A_i\right) = P\left(\bigcup_{i=1}^{n-1} A_i\right) + P(A_n) - P\left(A_n \cap \bigcup_{i=1}^{n-1} A_i\right) \leq P\left(\bigcup_{i=1}^{n-1} A_i\right) + P(A_n)$

$P\left(\bigcup_{i=1}^{n-1} A_i\right) = P\left(\bigcup_{i=1}^{n-2} A_i\right) + P(A_{n-1}) - P\left(A_{n-1} \cap \bigcup_{i=1}^{n-2} A_i\right) \leq P\left(\bigcup_{i=1}^{n-2} A_i\right) + P(A_{n-1})$

$\vdots$

$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \leq P(A_1) + P(A_2)$

a sčítaním máme $P\left(\bigcup_{i=1}^n A_i\right) \leq P(A_1) + P(A_2) + ... + P(A_n)$. ♣

Veta 2.2. Nech $(\Omega, \mathcal{A})$ je javové pole, P reálna množinová funkcia definovaná na $\mathcal{A}$ s vlastnosťami

(i) $P(\Omega) = 1$

(ii) $\forall\, A \in \mathcal{A} \quad P(A) \geq 0$

(iii) $A, B \in \mathcal{A},\ A \cap B = \emptyset \implies P(A \cup B) = P(A) + P(B)$ (aditivita, nie $\sigma-$aditivita)

Potom nasledujúce vlastnosti sú ekvivalentné

(1) P je pravdepodobnosť na $(\Omega, \mathcal{A})$

(2) $A_1, A_2, ... \in \mathcal{A}, A_n \subseteq A_{n+1} \implies \lim_{n\to\infty} P(A_n) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = P(\lim_{n\to\infty} A_n)$ (spojitosť zdola)

(3) $A_1, A_2, ... \in \mathcal{A}, A_n \supseteq A_{n+1} \implies \lim_{n\to\infty} P(A_n) = P\left(\bigcap_{i=1}^{\infty} A_i\right) = P(\lim_{n\to\infty} A_n)$ (spojitosť zhora)

(4) $A_1, A_2, ... \in \mathcal{A}, A_n \supseteq A_{n+1},\ \bigcap_{n=1}^{\infty} A_n = \emptyset \implies \lim_{n\to\infty} P(A_n) = 0$ (spojitosť zhora v $\emptyset$).

Dôkaz:

(1) $\implies$ (2)

P je $\sigma-$aditívna, teda ak $B_1, B_2, ... \in \mathcal{A}, B_i \cap B_j = \emptyset$ pre $i \neq j \implies P\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} P(B_i)$. Položme $B_1 = A_1, B_2 = A_2 - A_1, B_3 = A_3 - A_2, \ldots$. Platí $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$ a $B_i \cap B_j = \emptyset$ pre $i \neq j$. Dostávame

$P\left(\bigcup_{i=1}^{\infty} A_i\right) = P\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{n=1}^{\infty} P(B_n) = \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) =$
$\lim_{n\to\infty}[P(B_1) + P(B_2) + ... + P(B_n)] = \lim_{n\to\infty}[P(A_1) + P(A_2) - P(A_1) + ... + P(A_n) - P(A_{n-1})] = \lim_{n\to\infty} P(A_n)$.

(2) $\implies$ (3)

$A_n \supseteq A_{n+1}$, preto $\overline{A_n} \subseteq \overline{A_{n+1}}$ a podľa (2) $\lim_{n\to\infty} P(\overline{A_n}) = P\left(\bigcup_{i=1}^{\infty} \overline{A_i}\right) = P\left(\overline{\bigcap_{i=1}^{\infty} A_i}\right)$ (de Morgan) $= 1 - P\left(\bigcap_{n=1}^{\infty} A_n\right)$. Teda $\lim_{n\to\infty} P(A_n) = \lim_{n\to\infty}[1 - P(\overline{A_n})] = 1 - [1 - P\left(\bigcap_{n=1}^{\infty} A_n\right)] = P\left(\bigcap_{n=1}^{\infty} A_n\right)$.

(3) $\implies$ (4)

Ak $A_1, A_2, ... \in \mathcal{A}, A_n \supseteq A_{n+1},\ \bigcap_{n=1}^{\infty} A_n = \emptyset$, tak $\lim_{n\to\infty} P(A_n) = P\left(\bigcap_{i=1}^{\infty} A_i\right) = P(\emptyset) = 0$.

(4) $\implies$ (1)

Nech $B_1, B_2, ... \in \mathcal{A}$. Platí $B_i \cap B_j = \emptyset$ pre $i \neq j$. ďalej platí $P\left(\bigcup_{i=1}^{\infty} B_i\right) = P\left(B_1 \cup ... \cup B_{n-1} \cup \bigcup_{i=n}^{\infty} B_i\right)$. Ak označíme $\bigcup_{i=n}^{\infty} B_i = C_n$, potom $C_n \supseteq C_{n+1}$ a $\bigcap_{n=1}^{\infty} C_n = \emptyset$ (lebo $\bigcap_{n=1}^{\infty} C_n = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} B_i = \limsup_{n\to\infty} B_n = \{\omega \in \Omega :\ \omega$ patrí do nekonečne veľa $B_i\} = \emptyset$, lebo B_i sú po dvoch disjunktné). Teda podľa (4) $\lim_{n\to\infty} P(C_n) = 0$.

Počítajme pre ľubovoľné $n \geq 2$:

$P\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{n-1} P(B_i) + P(C_n)$ (aditivita P).

Preto platí:

$P\left(\bigcup_{i=1}^{\infty} B_i\right) = \lim_{n\to\infty} P\left(\bigcup_{i=1}^{\infty} B_i\right) = \lim_{n\to\infty}[\sum_{i=1}^{n-1} P(B_i) + P(C_n)] =$
$\lim_{n\to\infty} \sum_{i=1}^{n-1} P(B_i) + \lim_{n\to\infty} P(C_n) = \sum_{i=1}^{\infty} P(B_i)$. ♣

Veta 2.3. Nech $(\Omega, \mathcal{A}, P)$ je pravdepodobnostný priestor, $A_n \in \mathcal{A},\ n = 1, 2, ...$ a existuje $\lim_{n\to\infty} A_n = A$. Potom $P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n)$.

Dôkaz:

Pre reálnu číselnú postupnosť $\{a_n\}_{n=1}^{\infty}$ platí: a je hromadným bodom $\{a_n\}_{n=1}^{\infty}$ ak a je limitou nejakej vybranej podpostupnosti z postupnosti $\{a_n\}_{n=1}^{\infty}$. Množina hromadných bodov každej reálnej postupnosti má najväčší a najmenší prvok. Najväčší prvok je $\limsup_{n\to\infty} a_n$ a najmenší prvok je $\liminf_{n\to\infty} a_n$. Postupnosť $\{a_n\}_{n=1}^{\infty}$ má limitu práve vtedy ak $\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = \lim_{n\to\infty} a_n$ (Jarník, V., Diferenciální počet II, Academia, Praha, 1976).

Ďalej označme $\bigcap_{i=n}^{\infty} A_i = B_n,\ \bigcup_{i=n}^{\infty} A_i = C_n,\ P(B_n) = b_n, P(C_n) = c_n$. Zrejme $B_n \subseteq B_{n+1},\ C_n \supseteq C_{n+1},\ n = 1, 2, ...$

Podľa Vety 2.2 $\exists\ \lim_{n\to\infty} b_n = \lim_{n\to\infty} P(B_n) = P(\bigcup_{n=1}^{\infty} B_n)$ a podľa Vety 2.2 $\exists\ \lim_{n\to\infty} c_n = \lim_{n\to\infty} P(C_n) = P(\bigcap_{n=1}^{\infty} C_n)$.

Z predpokladov vety tiež $A = \lim_{n\to\infty} A_n = \liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n$.

Počítajme:

$P(\lim_{n\to\infty} A_n) = P(\liminf_{n\to\infty} A_n) = P\left(\bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} A_i\right) = P\left(\bigcup_{n=1}^{\infty} B_n\right) =$
$\lim_{n\to\infty} P(B_n) = \lim_{n\to\infty} b_n = \liminf_{n\to\infty} b_n = \liminf_{n\to\infty} P\left(\bigcap_{i=n}^{\infty} A_i\right) \leq$
$\liminf_{n\to\infty} P(A_n) \leq \limsup_{n\to\infty} P(A_n) \leq \limsup_{n\to\infty} P\left(\bigcup_{i=n}^{\infty} A_i\right) =$
$\limsup_{n\to\infty} P(C_n) = \limsup_{n\to\infty} c_n = \lim_{n\to\infty} c_n = \lim_{n\to\infty} P(C_n) = P\left(\bigcap_{n=1}^{\infty} C_n\right) = P\left(\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i\right) =$
$P(\limsup_{n\to\infty} A_n) = P(\lim_{n\to\infty} A_n)$.

Preto všade platí rovnosť a $P(\lim_{n\to\infty} A_n) = \liminf_{n\to\infty} P(A_n) =$
$\limsup_{n\to\infty} P(A_n) =$ (Jarník) $\lim_{n\to\infty} P(A_n)$. ♣

Veta 2.4. (Borelova-Cantelliho lema) Nech $A_n,\ n = 1, 2, ...$ je postupnosť náhodných javov na $(\Omega, \mathcal{A}, P)$ a $\sum_{n=1}^{\infty} P(A_n) < \infty$. Potom $P(\limsup_{n\to\infty} A_n) = 0$.

Dôkaz:

$0 \leq P(\limsup_{n\to\infty} A_n) = P\left(\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i\right) = \lim_{n\to\infty} P\left(\bigcup_{i=n}^{\infty} A_i\right)$ (Veta 2.2, lebo $\{\bigcup_{i=n}^{\infty} A_i\}_{n=1}^{\infty}$ je klesajúca postupnosť.

Platí tiež: $\bigcup_{i=n}^{\infty} A_i = A_n \cup A_{n+1} \cup ... = A_n \cup (A_{n+1} - A_n) \cup (A_{n+2} - \bigcup_{i=n}^{n+1} A_i) \cup (A_{n+3} - \bigcup_{i=n}^{n+2} A_i) \cup ...$, pričom $A_n, A_{n+1} - A_n, A_{n+2} - \bigcup_{i=n}^{n+1} A_i, ...$ sú disjunktné. Preto

$\lim_{n\to\infty} P(\bigcup_{i=n}^{\infty} A_i) = \lim_{n\to\infty} P((A_n \cup (A_{n+1} - A_n) \cup (A_{n+2} - \bigcup_{i=n}^{n+1} A_i) \cup ...)) =$
$\lim_{n\to\infty}(P(A_n) + P(A_{n+1} - A_n) + P(A_{n+2} - \bigcup_{i=n}^{n+1} A_i) + ...) \leq \lim_{n\to\infty} \sum_{i=n}^{\infty} P(A_i) =$
$= \lim_{n\to\infty}(\sum_{i=1}^{\infty} P(A_i) - \sum_{i=1}^{n-1} P(A_i)) = \sum_{i=1}^{\infty} P(A_i) - \sum_{i=1}^{\infty} P(A_i) = 0.$

Teda $P(\limsup_{n\to\infty} A_n) = 0$. ♣

3. Podmienená pravdepodobnosť

Príklad 3.1. Majme urnu s a čiernymi a b bielymi guľkami. Guľku po vytiahnutí *nevrátime* späť. Označme náhodný jav

B_1 – v prvom ťahu vytiahneme bielu guľku

B_2 – v druhom ťahu vytiahneme bielu guľku

Zaujíma nás pravdepodobnosť, s akou v druhom ťahu vytiahneme bielu guľku, ak vieme, že v prvom ťahu sme vytiahli bielu guľku.

Riešenie:

$$P(B_1) = \frac{b}{a+b}.$$

Podobne

$$P(B_2|B_1) = \frac{b-1}{a+b-1}.$$

Označenie $P(B_2|B_1)$ znamená podmienená pravdepodobnosť nahodného javu B_2 ak nastal náhodný jav B_1. Platí tiež

$$P(B_1 \cap B_2) = \frac{b(b-1)}{(a+b)(a+b-1)},$$

lebo všetkých možností (výsledkov) dvoch ťahov je $b(b-1+a) + a(b+a-1) = b^2 - b + ab + ab + a^2 - a = (a+b)(a+b-1)$ a "priaznivých" $b(b-1)$.

$$\begin{array}{llllll}
(bi_1, bi_2) & (bi_1, bi_3) & \dots & (bi_1, bi_b) & (bi_1, \check{c}_1) & \dots (bi_1, \check{c}_a) \\
(bi_2, bi_1) & (bi_2, bi_3) & \dots & (bi_2, bi_b) & (bi_2, \check{c}_1) & \dots (bi_2, \check{c}_a) \\
\vdots & & & & & \\
(bi_b, bi_1) & (bi_b, bi_2) & \dots & (bi_b, bi_{b-1}) & (bi_b, \check{c}_1) & \dots (bi_b, \check{c}_a)
\end{array}$$

$$\begin{array}{llllll}
(\check{c}_1, bi_1) & (\check{c}_1, bi_2) & \dots & (\check{c}_1, bi_b) & (\check{c}_1, \check{c}_2) & \dots (\check{c}_1, \check{c}_a) \\
\vdots & & & & & \\
(\check{c}_a, bi_1) & (\check{c}_a, bi_2) & \dots & (\check{c}_a, bi_b) & (\check{c}_a, \check{c}_1) & \dots (\check{c}_1, \check{c}_{a-1})
\end{array}$$

Môžeme ale písať

$$P(B_2|B_1) = \frac{P(B_2 \cap B_1)}{P(B_1)} = \frac{\frac{b(b-1)}{(a+b)(a+b-1)}}{\frac{b}{a+b}} = \frac{b-1}{a+b-1}.$$

Teda ohraničili sme sa namiesto Ω na B_1

$$\begin{array}{llllll}
(bi_1, bi_2) & (bi_1, bi_3) & \dots & (bi_1, bi_b) & (bi_1, \check{c}_1) & \dots (bi_1, \check{c}_a) \\
(bi_2, bi_1) & (bi_2, bi_3) & \dots & (bi_2, bi_b) & (bi_2, \check{c}_1) & \dots (bi_2, \check{c}_a) \\
\vdots & & & & & \\
(bi_b, bi_1) & (bi_b, bi_2) & \dots & (bi_b, bi_{b-1}) & (bi_b, \check{c}_1) & \dots (bi_b, \check{c}_a)
\end{array}$$

a z náhodného javu B_2 berieme "len tú časť, ktorá je v B_1".

Definícia 3.1. Majme pravdepodobnostný priestor $(\Omega, \mathcal{A}, P)$ a $B \in \mathcal{A}$ je vybraný náhodný jav taký, že $P(B) > 0$. *Podmienená pravdepodobnosť* náhodného javu $A \in \mathcal{A}$ za podmienky nastatia náhodného javu B je

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{3.1}$$

Poznámka. Z (3.1) vyplýva

$$P(A \cap B) = P(A|B)P(B) \tag{3.2}$$

pričom sa predpokladá, že $P(B) > 0$. Pretože $A \cap B \subseteq B$, teda $P(B) = 0 \Longrightarrow P(A \cap B) = 0$, vzťah (3.2) má význam aj pre $P(B) = 0$. Vzťah (3.2) je "symetrický" aj pre A, teda

$$P(A \cap B) = P(B|A)P(A). \tag{3.3}$$

Z (3.2) a (3.3) máme

$$P(A|B)P(B) = P(B|A)P(A). \tag{3.4}$$

Označenie:
Nech jav $B \in \mathcal{A}$ je pevne daný, pričom $P(B) > 0$. Definujme

$$P_B : \quad \mathcal{A} \rightarrow < 0, 1 > \qquad P_B(A) = P(A|B).$$

Veta 3.1 P_B je pravdepodobnosť na $(\Omega, \mathcal{A})$ (pre každý jav B, pre ktorý je $P(B) > 0$).

Dôkaz:
$P_B(\Omega) = \dfrac{P(B \cap \Omega)}{P(B)} = \dfrac{P(B)}{P(B)} = 1,$
$P_B(A) = \dfrac{P(A \cap B)}{P(B)} \geq 0$ pre $\forall\ A \in \mathcal{A}$
$A_n \in \mathcal{A},\ n = 1, 2, ...,\ A_i \cap A_j = \emptyset$ pre $i \neq j$, potom

$$P_B\left(\bigcup_{i=1}^{\infty} A_i\right) = \frac{P\left(\bigcup_{i=1}^{\infty} A_i \cap B\right)}{P(B)} = \frac{P\left(\bigcup_{i=1}^{\infty}(A_i \cap B)\right)}{P(B)} =$$

$$= \frac{\sum_{i=1}^{\infty} P(A_i \cap B)}{P(B)} = \sum_{i=1}^{\infty} P(A_i|B) = \sum_{i=1}^{\infty} P_B(A_i). \quad \clubsuit$$

Veta 3.2 Platí

(i) $P(A|\Omega) = P(A)$ pre $\forall\ A \in \mathcal{A}$

(ii) $P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)P(A_2|A_1)P(A_3|A_2 \cap A_1)...P(A_n|A_1 \cap A_2 \cap ... \cap A_{n-1})$ ak $P\left(\bigcap_{i=1}^{n-1} A_i\right) > 0$ (veta o násobení pravdepodobnosti).

Dôkaz:

(i) $P(A|\Omega) = \dfrac{P(A \cap \Omega)}{P(\Omega)} = \dfrac{P(A)}{P(\Omega)} = P(A);$

(ii) Z (3.2) je $P\left(\bigcap_{i=1}^{n} A_i\right) =$
$P\left(A_n \cap \bigcap_{i=1}^{n-1} A_i\right) =$
$P\left(A_n | \bigcap_{i=1}^{n-1} A_i\right) P\left(\bigcap_{i=1}^{n-1} A_i\right) =$
$P\left(A_n | \bigcap_{i=1}^{n-1} A_i\right) P\left(A_{n-1} \cap \bigcap_{i=1}^{n-2} A_i\right) =$

$P\left(A_n|\bigcap_{i=1}^{n-1}A_i\right)P\left(A_{n-1}|\bigcap_{i=1}^{n-2}A_i\right)P\left(\bigcap_{i=1}^{n-2}A_i\right)=...=$
$P(A_1)P(A_2|A_1)P(A_3|A_2\cap A_1)...P(A_n|A_1\cap A_2\cap...\cap A_{n-1}).$ ♣

Definícia 3.2. Majme pravdepodobnostný priestor $(\Omega,\mathcal{A},P)$. Náhodné javy $A_1,A_2,...\in\mathcal{A}$ tvoria *úplný systém javov*, ak platí

$$A_i\cap A_j=\emptyset,\quad i\neq j,\quad \text{a}\quad \bigcup_{i=1}^{\infty}A_i=\Omega. \tag{3.5}$$

Poznámka. úplný systém javov môže byť aj konečný.

Veta 3.3. (Vzorec pre úplnú pravdepodobnosť) Nech $A_1,A_2,...$ je úplný systém javov v pravdepodobnostnom priestore $(\Omega,\mathcal{A},P)$ taký, že

$$P(A_i)>0,\quad i=1,2,...\ . \tag{3.6}$$

Potom platí

$$P(B)=\sum_{i=1}^{\infty}P(B|A_i)P(A_i). \tag{3.7}$$

Dôkaz:

$P(B)=P(B\cap\Omega)=P\left(B\cap\bigcup_{i=1}^{\infty}A_i\right)=P\left(\bigcup_{i=1}^{\infty}(B\cap A_i)\right)=\sum_{i=1}^{\infty}P(B\cap A_i)=\sum_{i=1}^{\infty}P(B|A_i)P(A_i)$ (podľa (3.2)) ♣

Veta 3.4. (1. Bayesov vzorec) Nech $A_1,A_2,...$ je úplný systém javov v pravdepodobnostnom priestore $(\Omega,\mathcal{A},P)$ taký, že

$$P(A_i)>0,\quad i=1,2,...\ .$$

Ak $P(B)>0$, tak platí

$$P(A_j|B)=\frac{P(B|A_j)P(A_j)}{\sum_{i=1}^{\infty}P(B|A_i)P(A_i)},\quad j=1,2,...\ . \tag{3.8}$$

Dôkaz:

Pre ľubovoľné j je pomocou (3.2) a (3.7)

$$P(A_j|B)=\frac{P(B\cap A_j)}{P(B)}=\frac{P(B|A_j)P(A_j)}{\sum_{i=1}^{\infty}P(B|A_i)P(A_i)}.\quad ♣$$

Veta 3.5. (2. Bayesov vzorec) Nech $A_1,A_2,...$ je úplný systém javov v pravdepodobnostnom priestore $(\Omega,\mathcal{A},P)$ taký, že $P(A_i)>0,\quad i=1,2,...$. ďalej $A\in\mathcal{A}$, že $P(A)>0$ a $B\in\mathcal{A}$. Platí

$$P(B|A)=\frac{\sum_{\{i:\ P(A\cap A_i)>0\}}P(A_i)P(A|A_i)P(B|A\cap A_i)}{\sum_{i=1}^{\infty}P(A|A_i)P(A_i)}. \tag{3.9}$$

Dôkaz: spravte si sami.

Poznámka. Vety 3.3, 3.4 a 3.5 platia aj v prípade, že úplny systém javov je konečný.

Poznámka. $P(A_j)$ v Bayesových vzorcoch sú tzv. *apriorné pravdepodobnosti* a $P(A_j|B)$ *aposteriórne pravdepodobnosti* (po vykonaní pokusu s výsledkom B).

Poznámka. V prípade 1. Bayesovho vzorca ide o riešenie situácie, keď máme hypotézy $A_1,...,$ ktoré sa navzájom vylučujú, ale vyčerpávajú všetky možnosti. Poznáme ich (apriorné) pravdepodobnosti $P(A_i)$. Nastal jav A a poznáme pravdepodobnosti $P(A|A_i)$. Pýtame sa na (aposteriórne; nové, ktoré berú do úvahy skutočnosť, že mastal A) pravdepodobnosti $P(A_i|A)$

V prípade 2. Bayesovho vzorca ak nastal jav A, pýtame sa na pravdepodobnosť javu B.

Poznámka. Nie je vždy jednoduché voliť správny pravdepodobnostný model pre výpočet podmienených pravdepodobností.

Príklad 3.2. (Lekárska diagnostika) Vieme, že určitou (konkrétnou) chorobou *Ch* trpí 1% populácie. Choroba je diagnostikovaná na základe vyšetrenia, ktorého spoľahlivosť je

(i) 95% ak vyšetrovaná osoba trpí chorobou *Ch*

(ii) 70 % ak vyšetrovaná osoba netrpí chorobou *Ch*.

Vyšetrujeme náhodne zvolenú osobu. Určte pravdepodobnosť správnej diagnózy, ak výsledok vyšetrenia je

(a) pozitívny (podľa výsledku vyšetrenia je osoba chorá)

(b) negatívny (podľa výsledku vyšetrenia je osoba zdravá).

Riešenie:

Označme jav

A – vyšetrovaná osoba trpí chorobou *Ch* (je chorá)

B – výsledok vyšetrovania je pozitívny

Zo zadania vieme

$P(A) = 0.01$ (pravdepodobnosť, že vybraná osoba je chorá) Táto pravdepodobnosť sa volá <u>prevalencia</u> alebo tiež apriorná pravdepodobnosť choroby

Vyšetrenie (spoľahlivosť vyšetrenia) sa charakterizuje dvomi charakteristikami, a síce

pravdepodobnosťou $P(B|A) = 0.95$ tzv. <u>citlivosť</u> testu alebo aj <u>senzitivita</u> testu

pravdepodobnosťou $P(\overline{B}|\overline{A}) = 0.7$ tzv. <u>špecificita</u> testu.

(a) Máme určiť vlastne $P(A|B)$ (lebo v tomto prípade výsledok testu bol pozitívny, teda test hovorí, že vyšetrovaná osoba je chorá (diagnóza je, že pacient je chorý) a my máme určiť pravdepodobnosť správnej dignózy).

Zo zadania vieme, že $P(A) = 0.01$, $P(\overline{A}) = 0.99$, $P(B|A) = 0.95$ a $P(B|\overline{A}) = 1 - P(\overline{B}|\overline{A}) = 1 - 0.7 = 0.3$. Podľa Bayesovho vzorca ($A, \overline{A}$ sú hypotézy)

$$P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(\overline{A})P(B|\overline{A})} = \frac{0.01 \cdot 0.95}{0.01 \cdot 0.95 + 0.99 \cdot 0.3} = 0.030995.$$

Je to aj aposteriórna pravdepodobnosť, že pacient je chorý, ak výsledok testu bol pozitívny. Je to prekvapivý výsledok. čakali by sme "omnoho lepší" výsledok.

Celkom máme 29 700 + 950 = 30 650 pozitívnych výsledkov, z toho správne pozitívnych je 950, čiže $P(A|B) = \dfrac{950}{30650} = 0.030995$.
(b) Analogicky (zase $A, \overline{A}$ sú hypotézy)

$$P(\overline{A}|\overline{B}) = \frac{P(\overline{A})P(\overline{B}|\overline{A})}{P(\overline{A})P(\overline{B}|\overline{A}) + P(A)P(\overline{B}|A)} = \frac{0.99 \cdot 0.7}{0.99 \cdot 0.7 + 0.01 \cdot 0.05} = 0.99928.$$

Je to aposteriórna pravdepodobnosť, že pacient nie je chorý, ak výsledok testu bol negatívny. Naozaj celkovo máme 69 300 + 50 = 69 350 negatívnych výsledkov, z toho správne negatívnych je 69 300 a teda pravdepodobnosť správnej diagnózy u negatívnych výsledkov testu je $P(\overline{A}|\overline{B}) = \dfrac{69300}{69350} = 0.99928$.

Nezávislosť náhodných javov

Definícia 3.3. Majme pravdepodobnostný priestor $(\Omega, \mathcal{A}, P)$. Náhodné javy $A, B \in \mathcal{A}$ sú *nezávislé* (vzhľadom k pravdepodobnosti P) ak $P(A \cap B) = P(A)P(B)$.

Definícia 3.4. Majme pravdepodobnostný priestor $(\Omega, \mathcal{A}, P)$. Náhodné javy $A_1, A_2, ... \in \mathcal{A}$ sú *skupinovo (združene) nezávislé* (vzhľadom k pravdepodobnosti P) ak pre ľubovoľné $k \in \{1, 2, ...\}$ a ľubovoľnú skupinu indexov $\{i_1, ..., i_k\} \subseteq \{1, 2, ...\}$ platí

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P(A_{i_j}).$$

Náhodné javy $A_1, A_2, ... \in \mathcal{A}$ sú **po dvoch nezávislé**, ak každé dva sú nezávislé.

Poznámka. Zrejme ľubovoľný jav $A \in \mathcal{A}$ a jav istý Ω sú nezávislé. Takisto ľubovoľný jav $A \in \mathcal{A}$ a jav nemožný $\emptyset$ sú nezávislé.

Poznámka. Pozor, je rozdiel medzi disjunktnými (nezlučiteľnými) javmi (nemôžu naraz nastať, $A \cap B = \emptyset$) a nezávislými javmi (tu treba pravdepodobnosť).

Príklad 3.3. V urne sú 4 lístky $\{000, 110, 101, 011\}$. Náhodné javy

$$A_i - \{\text{náhodne vytiahnutý lístok má na } i\text{−tom mieste } 1\}, \quad i = 1, 2, 3,$$

sú po dvoch nezávislé, ale nie sú (združene) nezávislé, lebo

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{2}$$

$$P(A_1 \cap A_2) = P(A_1 \cap A_3) = P(A_2 \cap A_3) = \frac{1}{4}$$

$$P(A_1 \cap A_2 \cap A_3) = 0.$$

Veta 3.6. Nech $A_1, A_2, ..., A_n \in \mathcal{A}$ sú združene nezávislé javy. Platí

(i) ľubovoľná postupnosť $\tilde{A}_1, \tilde{A}_2, ..., \tilde{A}_n$, kde $\tilde{A}_k = A_k$ alebo $\tilde{A}_k = \overline{A_k}$ je postupnosť združene nezávislých javov;

(ii) $P\left(\bigcup_{k=1}^{n} A_k\right) = 1 - \prod_{k=1}^{n} (1 - P(A_k))$.

Dôkaz:

(i) Ak A_1, A_2 sú nezávislé, tak

$$P(A_1 \cap \overline{A_2}) = P(A_1 - (A_1 \cap A_2)) = P(A_1) - P(A_1 \cap A_2) =$$

$$= P(A_1) - P(A_1)P(A_2) = P(A_1)(1 - P(A_2)) = P(A_1)P(\overline{A_2}),$$

teda $A_1, \overline{A_2}$ sú nezávislé. Tak isto

$$P(\overline{A_1} \cap \overline{A_2}) = P(\overline{(A_1 \cup A_2)}) = 1 - P(A_1 \cup A_2) = 1 - P(A_1) - P(A_2) + P(A_1 \cap A_2) =$$

$$= 1 - P(A_1) - P(A_2) + P(A_1)P(A_2) = (1 - P(A_1))(1 - P(A_2)) = P(\overline{A_1})P(\overline{A_2}),$$

čiže aj $\overline{A_1}$ a $\overline{A_2}$ sú nezávislé. Dôkaz dokončíme indukciou (pozri Riečan, B., O pravdepodobnosti a miere, Alfa, Bratislava, 1972, alebo Dupač, Hušková, Pravděpodobnost a matematická statistika).

(ii) Z de Morganových pravidiel a z (i)

$$1 - P\left(\bigcup_{k=1}^{n} A_k\right) = P\left(\overline{\bigcup_{k=1}^{n} A_k}\right) = P\left(\bigcap_{k=1}^{n} \overline{A_k}\right) = \prod_{k=1}^{n} (1 - P(A_k)) . \quad \clubsuit$$

Veta 3.7. (Borelova lema) Nech $A_1, A_2, ... \in \mathcal{A}$ je postupnosť nezávislých javov, t.j. . Potom

$$P(\limsup_{n\to\infty} A_n) = 0 \vee 1$$

podľa toho, či rad $\sum_{n=1}^{\infty} P(A_n)$ konverguje alebo diverguje.

Dôkaz:

Ak $\sum_{n=1}^{\infty} P(A_n) < \infty \Longrightarrow P(\limsup_{n\to\infty} A_n) = 0$ podľa Borelovej-Cantelliho lemy (A_i ani nemusia byť nezávislé).

Ak $\sum_{n=1}^{\infty} P(A_n) = \infty$, tak

$P(\limsup_{n\to\infty} A_n) = P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right) = P\left(\bigcap_{n=1}^{\infty} B_n\right) =$

(kde $B_n = \bigcup_{k=n}^{\infty} A_k \supseteq B_{n+1}$)

$=$ (Veta 2.2. (3)) $\lim_{n\to\infty} P(B_n) =$

$(B_n = A_n \cup (A_n \cup A_{n+1}) \cup (A_n \cup A_{n+1} \cup A_{n+2}) \cup ...)$

$=$ (Veta 2.2. (2)) $\lim_{n\to\infty} \left[\lim_{N\to\infty} P\left(\bigcup_{k=n}^{N} A_k\right)\right] =$

$=$ (de Morgan) $\lim_{n\to\infty} \lim_{N\to\infty} P\left(\overline{\bigcap_{k=n}^{N} \overline{A_k}}\right) =$

$= \lim_{n\to\infty} \lim_{N\to\infty} \left[1 - P\left(\bigcap_{k=n}^{N} \overline{A_k}\right)\right] =$

$=$ (nezávislosť) $1 - \lim_{n\to\infty} \lim_{N\to\infty} \left[\prod_{k=n}^{N} P(\overline{A_k})\right] =$

$= 1 - \lim_{n\to\infty} \lim_{N\to\infty} \prod_{k=n}^{N} (1 - P(A_k)) \geq 1 - \lim_{n\to\infty} \lim_{N\to\infty} e^{-\sum_{k=n}^{N} P(A_k)}$

(lebo $0 \leq P(A_k) = x_k \leq 1$ a $1 - x_k \leq e^{-x_k}$, teda $\prod_{k=n}^{N} (1 - x_k) \leq \prod_{k=n}^{N} e^{-x_k}$, čiže $-\prod_{k=n}^{N} (1 - x_k) \geq -\prod_{k=n}^{N} e^{-x_k} = -e^{-\sum_{k=n}^{N} P(A_k)}$).

Pretože $\sum_{n=1}^{\infty} P(A_n) = \infty$, čiže $\lim_{N\to\infty} \sum_{n=1}^{N} P(A_n) = \infty$ a aj $\lim_{N\to\infty} \sum_{n=k}^{N} P(A_n) = \infty$ pre každé n. Teda $\lim_{n\to\infty} \lim_{N\to\infty} \sum_{n=k}^{N} P(A_n) = \infty$ a $\lim_{n\to\infty} \lim_{N\to\infty} e^{-\sum_{n=k}^{N} P(A_n)} = 0$.

Dostávame $P(\limsup_{n\to\infty} A_n) = 1$. $\clubsuit$

4. Náhodná veličina

Snažíme sa výsledok pokusu vyjadriť číslom (počet padnutých šestiek na 10 kockách; doba po ktorú svieti žiarovka; počet baktérií v jednotkovom objeme vody; atď.).

Snažíme sa "pretransformovať" výsledok pokusu, náhodné javy na číselnú os. Pravdepodobnostný priestor "zobraziť" na číselnú os tak, aby sa dala spočítať pravdepodobnosť všetkým "rozumným" množinám reálnych čísel. Teda chceme nájsť vhodné zobrazenie

$$(\Omega, \mathcal{A}, P) \to \mathtt{R}$$

pričom prepokladáme, že $(\Omega, \mathcal{A}, P)$ máme dané, určené napr. verbálne (slovne). Ukazuje sa rozumné vziať na reálnej osi borelovskú $\sigma-$algebru $\mathcal{B}$ a hľadať vhodné zobrazenie

$$X: \quad (\Omega, \mathcal{A}) \to (\mathtt{R}, \mathcal{B})$$

tak, aby sme mohli spočítať (udať) pravdepodobnosť každej borelovskej množiny $B \in \mathcal{B}$. Zadefinujme si takúto "vhodnú" funkciu.

Definícia 4.1. Majme daný pravdepodobnostný priestor $(\Omega, \mathcal{A}, P)$. Reálnu funkciu X definovanú na Ω pre ktorú platí

$$\forall\ B \in \mathcal{B} \Longrightarrow \{\omega \in \Omega: \quad X(\omega) \in B\} \in \mathcal{A} \tag{4.1}$$

nazývame *náhodná veličina*.

Náhodnú veličinu niekedy voláme aj náhodná premenná. Funkciu X splňujúcu (4.1) nazývame *merateľná funkcia*, prvky $\sigma-$algebry $\mathcal{B}$ *merateľné množiny*. Množinu $\{\omega \in \Omega: \quad X(\omega) \in B\}$ zapisujeme (skrátene) $\{X \in B\}$ alebo $\{X^{-1}(B)\}$. Je zrejmé, že náhodnou veličinou X zobrazíme elemenárny výsledok pokusu ω na reálne číslo. Keď sa zrealizuje elemenárny jav ω, tak realizácia náhodnej veličiny je (reálne číslo) $x = X(\omega)$. Keď máme zadanú (určenú) náhodnú veličinu X, tak pre každú borelovskú množinu $B \in \mathcal{B}$ vieme určiť $\{\omega \in \Omega: \ X(\omega) \in B\} = \{X^{-1}(B)\}$. špeciálne pre každé reálne číslo x je $\{\omega \in \Omega: \quad X(\omega) \in (-\infty, x)\} = \{X < x\} \in \mathcal{A}$.

Poznámka. Dá sa ukázať, že k tomu, aby X bola náhodná veličina je nutné a stačí, aby $\forall x \in \mathtt{R}\ \ \{X < x\} \in \mathcal{A}$.

Keď máme daný pravdepodobnostný priestor $(\Omega, \mathcal{A}, P)$ (matematicky popísaný náhodný pokus) a náhodnu veličinu X (t.j. reálnu funkciu s vlastnosťou (4.1)), tak každej borelovskej (merateľnej) množine B vieme priradiť (určiť) pravdepodobnosť predpisom

$$P_X(B) = P\{\omega \in \Omega: \quad X(\omega) \in B\} = P(\{X^{-1}(B)\}).$$

Definícia 4.2. Množinová funkcia

$$P_X(B) = P(\{X^{-1}(B)\}), \quad B \in \mathcal{B} \tag{4.2}$$

sa nazýva *rozdelenie pravdepodobnosti* náhodnej veličiny X.

Rozdelenie pravdepodobnosti náhodnej veličiny X je teda pravdepodobnostná miera (pravdepodobnosť) na $\sigma-$algebre $\mathcal{B}$ borelovských množín indukovaná náhodnou veličinou X. Presvedčte sa, že naozaj P_X spĺňa všetky tri vlastnosti z Definície 1.5.

Poznámka. Pri náhodnom pokuse - hode kockou náhodnú veličinu fyzicky zrealizujeme tak, že na jednotlivé steny kocky nakreslíme bodky. Môžeme na jednotlivé steny kocky aj napísať (nejaké konkrétne) čisla. Ak sme každej stene kocky priradili určité čislo, zostrojili sme istú nahodnú veličinu. Iný priklad na náhodnú veličinu je merací prístroj (napríklad voltmeter). Určitému napätiu v sieti priradí číslo - hodnotu napätia. V reálnej elektrickej sieti aj konštantné napätie nie je "pevné", ale fluktuuje (vplyvom náhodných porúch).

Náhodná veličina teda je pevne daná funkcia, ktorá ale svoje hodnoty nadobúda "nahodne". Pravdepodobnostné správanie sa náhodnej veličiny X, teda rozdelenie pravdepodobnosti náhodnej veličiny X je určené systémom pravdepodobností

$$P(\{X < x\}), \quad x \in \mathtt{R}.$$

Pravdepodobnostné správanie sa nahodnej veličiny úplne a jednoznačne popisuje distribučná funkcia.

Definícia 4.3. Nech X je náhodná veličina definovaná na $(\Omega, \mathcal{A}, P)$. Reálna funkcia

$$F_X(x) = P(\{X < x\}) = P(\{\omega \in \Omega : \ X(\omega) < x\})$$

sa nazýva *distribučná funkcia* náhodnej veličiny X. (Budeme značiť aj $F(x)$, ak nedochádza k nedorozumeniu.)

Veta 4.1. Distribučná funkcia je

(i) neklesajúca

(ii) spojitá zľava

(iii) $\lim_{x\to-\infty} F(x) = 0, \quad \lim_{x\to\infty} F(x) = 1.$

Dôkaz:

(i) Zvoľme reálne $a < b$. Zrejme

$$\{X < b\} = \{\omega : \ X(\omega) < b\} = \{X < a\} \cup \{a \leq X < b\}$$

pričom posledné dva javy sú nezlučiteľné. Podľa aditívnej vlastnosti pravdepodobnosti je

$$P(\{X < b\}) = P(\{X < a\}) + P(\{a \leq X < b\}),$$

čiže

$$0 \leq P(\{a \leq X < b\}) = P(\{X < a\}) - P(\{X < b\}) = F_X(b) - F_X(a),$$

teda $F_X(b) \geq F_X(a)$.

(ii) Pre ľubovoľné (ale pevné) x nech $\{x_i\}_{i=1}^{\infty}$ je akákoľvek neklesajúca postupnosť taká, že konverguje zľava k x (teda $x_n \to x^-$). Nech $A_i = \{\omega : \ X(\omega) < x_i\}$ a $A = \{\omega : \ X(\omega) < x\}$. Zrejme $\bigcup_{i=1}^{\infty} A_i = A, \ A_i \subseteq A_{i+1}$. Zo spojitosti zdola pravdepodobnosti dostávame

$$F_X(x) = P(A) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} P(A_i) = \lim_{i\to\infty} F_X(x_i).$$

(iii) Vezmime ľubovoľnú nerastúcu postupnosť $\{x_i\}_{i=1}^{\infty}$ takú, že $x_n \to -\infty$. Pri označení B_n, A_n z (ii) tentokrát $A_i \supseteq A_{i+1}$ a $\bigcap_{i=1}^{\infty} A_i = \emptyset$. Preto

$$\lim_{n\to\infty} F_X(x_n) = \lim_{n\to\infty} P(A_n) = P(\lim_{n\to\infty} A_n) = P(\emptyset) = 0.$$

Ak teraz vezmeme ľubovoľnú neklesajúcu $\{x_i\}_{i=1}^{\infty}$ takú, že $x_n \to \infty$. Pri označení A_n z (ii) je $A_i \subseteq A_{i+1}$ a

$$\lim_{n\to\infty} F_X(x_n) = \lim_{n\to\infty} P(A_n) = P\left(\bigcup_{n=1}^{\infty} A_n\right) = P(\Omega) = 1. \quad \clubsuit$$

Veta 4.2. Pre distribučnú funkciu F_X platí

$$P(\{X = x\}) = F_X(x+0) - F_X(x), \quad x \in \mathtt{R}. \tag{4.3}$$

($F_X(x+0) = \lim_{y\to x^+} F_X(y)$ (limita sprava).)

Dôkaz: Platí $\{X \le x\} = \{X = x\} \cup \{X < x\}$ a ak vezmeme ľubovoľnú nerastúcu $\{x_i\}_{i=1}^{\infty}$ takú, že $x_n \to x^+$, tak $\{X \le x\} = \bigcap_{n=1}^{\infty}\{X < x_n\}$, pričom $\{X < x_n\} \supseteq \{X < x_{n+1}\}$. Zo spojitosti pravdepodobnosti zhora platí

$$P(\{X \le x\}) = P\left(\bigcap_{n=1}^{\infty}\{X < x_n\}\right) = \lim_{n\to\infty} P(\{X < x_n\}) =$$

$$= \lim_{n\to\infty} F(x_n) = F_X(x+0),$$

(pričom zo spojitosti pravdepodobnosti zhora vyplýva, že $\lim_{n\to\infty} P(\{X < x_n\})$ existuje a je jediná pre akúkoľvek nerastúcu postupnosť $\{x_i\}_{i=1}^{\infty}$ takú, že $x_n \to x^+$). Preto

$$P(\{X = x\}) = P(\{X \le x\}) - P(\{X < x\}) = F_X(x+0) - F_X(x). \quad \clubsuit$$

Dôsledok. $F_X(\cdot)$ je spojitá v x práve vtedy ak $P(\{X = x\}) = 0$ (lebo $F_X(\cdot)$ je zľava spojitá vždy a sprava práve ak $P(\{X = x\}) = 0$).

Veta 4.3. Distribučná funkcia má najviac spočitateľne veľa bodov nespojitosti (skokov).

Dôkaz: Označme

$C_n = \{$ množina bodov, v ktorých má $F_X(\cdot)$ skok väčší ako $\frac{1}{n}$ $\}$.

Veľkosť skoku v bode x je vlastne (podľa (4.3)) $F_X(x+0) - F_X(x) = P(\{X = x\})$ a $C_n = \{x \in \mathtt{R} : \ P(\{X = x\}) > \frac{1}{n}\}$. Pretože hodnoty pravdepodobnosti ležia v intervale $< 0, 1 >$, môže mať C_n najviac $(n-1)$ prvkov.

Množina bodov

$$C = \{x : \ F_X(\cdot) \ \text{ má v bode } x \text{ nejaký skok}\} = \bigcup_{n=2}^{\infty} C_n.$$

Pretože C je spočitateľným zjednotením konečných množín, je nanajvýš spočitateľná. $\clubsuit$

Lebesgueova - Stieltjesova miera

Stručne si povieme o Lebesgueovej - Stieltjesovej miere. Majme danú reálnu funkciu F s vlastnosťami

(i) neklesajúca

(ii) spojitá zľava.

Majme systém S všekých intervalov tvaru $< a, b)$, kde $a < b$. Potom je množinová funkcia μ definovaná na S predpisom $\mu(< a, b)) = F(b) - F(a)$ $\sigma-$aditívna. (dôkaz pozri napr. Riečan, B., O pravdepodobnosti a miere, Alfa, Bratislava, 1972, Veta 5.2.1). Z teórie miery potom existuje práve jedna miera μ_F definovaná na systéme $\mathcal{B}$ všetkých borelovských množín taka, že $\mu_F(< a, b)) = F(b) - F(a)$ (dôkaz pozri napr. Riečan, B., O pravdepodobnosti a miere, Alfa, Bratislava, 1972, Veta 5.2.2). Miera μ_F sa nazýva Lebesgueova - Stieltjesova miera indukovaná funkciou F. Dá sa ukázať, že ak navyše platí, že $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to \infty} F(x) = 1$, tak μ_F je pravdepodobnosť na $(\mathtt{R}, \mathcal{B})$.

Poznamenávame len, že ak za funkciu $F(\cdot)$ zvolíme funkciu $F(x) = x$, tak μ_F sa volá Lebesgueova miera.

Poznámka. Lebesgueova-Stieltjesova miera sa zavádza všeobecnejšie pre funkcie F ktoré sú neklesajúce a spojité zľava (nemusia byť len distribučné funkcie). Pre nás je dôležitý prípad keď F je distribučná funkcia.

Poznámka. Ak máme náhodnú veličinu X a jej distribučnú funkciu F_X, tak na systéme S intervalov $< a, b)$, kde $a < b$ je

$$P_X(< a, b)) = P(\{a \leq X < b\}) = F_X(b) - F_X(a) = \mu_{F_X}(< a, b))$$

a preto pravdepodobnostná miera μ_F je totožná s rozdelením pravdepodobnosti P_X (podrobnejšie pozri napr. v Riečan, B., O pravdepodobnosti a miere, Alfa, Bratislava, 1972). Z teórie integrálu pre každú $B \in \mathcal{B}$ je

$$P_X(B) = \mu_{F_X}(B) = \int_B d\mu_{F_X}(x) \ \ \text{(Lebesgueov-Stieltjesov integrál)} \ =$$

$$= \int_B d\mu_{F_X} = \int_B dF_X(x) \ \ \text{(iné značenie)}.$$

Platí aj nasledujúca veta ("opak" Vety 4.1):

Veta 4.4. Nech F je neklesajúca, spojitá zľava a $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to \infty} F(x) = 1$. Potom existuje náhodná veličina X tak, že F je jej distribučná funkcia.

Dôkaz: Povedali sme, že μ_F je pravdepodobnosť na $(\mathtt{R}, \mathcal{B})$ a preto $(\mathtt{R}, \mathcal{B}, \mu_F)$ je pravdepodobnostný priestor. Definujme teraz na $\mathtt{R}$ náhodnú veličinu X vzťahom $X(x) = x$. Je zrejmé, že X je náhodná veličina, lebo ak $B \in \mathcal{B}$, tak $X^{-1}(B) = B$ je borelovská množina. Nech G je distribučná funkcia náhodnej veličiny X, potom

$$G(x) = \mu_F(\{X^{-1}((-\infty, x))) = \mu_F((-\infty, x)) = \mu_F(\bigcup_{n=1}^{\infty} < x - n, x)) =$$

$$= \lim_{n \to \infty} \mu_F(< x - n, x)) = \lim_{n \to \infty} (F(x) - F(x - n)) = F(x) - \lim_{n \to \infty} F(x - n) = F(x).$$

Teda F je distribučná funkcia náhodnej veličiny X. ♣

5. Diskrétne náhodné veličiny

(náhodné veličiny diskrétneho typu)

Náhodným veličinám zodpovedajú určité distribučné funkcie (teda aj určité Lebesgueove-Stieltjesove miery).

Definícia 5.1. Nech $\{p_i\}_{i=1}^{\infty}$ je rad kladných čísel takých, že $\sum_{i=1}^{\infty} p_i = 1$ a $M = \{x_i\}_{i=1}^{\infty}$ je (ľubovoľná) postupnosť rôznych reálnych čísel. Funkcia $(x_i, p_i)_{i=1}^{\infty}$ na $\{x_i\}_{i=1}^{\infty}$ sa nazýva *pravdepodobnostná funkcia*. Poznamenajme len, že postupnosti $\{x_i\}$ a $\{p_i\}$ môžu byť aj konečné.

Poznámka Pravdepodobnostná funkcia môže byť chápaná aj ako $(x_i, p_i)_{i \in J}$, kde J je konečná alebo spočitateľná indexová množina.

Veta 5.1. Nech $(x_i, p_i)_{i=1}^{\infty}$ je pravdepodobnostná funkcia. Položme

$$F(x) = \sum_{x_i < x} p_i.$$

Potom je funkcia F neklesajúca, spojitá zľava a $\lim_{x \to -\infty} F(x) = 0, \lim_{x \to \infty} F(x) = 1$ (teda distribučná funkcia).

Dôkaz: Nech $x < y$. Potom $F(y) - F(x) = \sum_{\{x_i:\ x \le x_i < y\}} p_i \ge 0$, teda $F(y) \ge F(x)$.

Nech je x pevné číslo. Nech ϵ je ľubovoľné kladné číslo. Pretože $\sum_{i=1}^{\infty} p_i = 1$, existuje také n, že $\sum_{i=n}^{\infty} p_i < \epsilon$. Vezmeme $\delta > 0$ tak, aby sa v $(x - \delta, x)$ nenachádzalo žiadne z čísel $x_1, ..., x_{n-1}$. Potom pre $y \in (x - \delta, x >$ je

$$F(x) - F(y) = \sum_{\{x_i:\ y \le x_i < x\}} p_i \le \sum_{i=n}^{\infty} p_i < \epsilon.$$

Nech $\{z_m\}_{m=1}^{\infty}$ je (ľubovoľná) taká klesajúca postupnosť, že $\lim_{m \to \infty} z_m = -\infty$. Nech ϵ je ľubovoľné kladné číslo. Pretože $\sum_{i=1}^{\infty} p_i = 1$, existuje také n_0, že $\sum_{i=n_0}^{\infty} p_i < \epsilon$. Pre toto ϵ ale $\exists m_0$, že $z_{m_0} < \min_{i \in \{1,...,n_0-1\}} \{x_i\}$ (pričom $\forall m > m_0$ je $z_m < z_{m_0}$). Preto $\forall m > m_0$ je $0 \le F(z_m) \le F(z_{m_0}) = \sum_{\{x_i:\ x_i < z_{m_0}\}} p_i \le \sum_{i=n_0}^{\infty} p_i < \epsilon$. Teda pre ľubovoľnú klesajúcu k $-\infty$ postupnosť $\{z_m\}_{m \ge 1}$ konverguje $\{F(z_m)\}_{m \ge 1}$ k nule.

Nech $\{z_n\}_{n=1}^{\infty}$ je (ľubovoľná) taká rastúca postupnosť, že $\lim_{n \to \infty} z_n = \infty$. Označme $\mathrm{card}\{x_i :\ x_i < z_n\} = j_n$. Zrejme $\lim_{n \to \infty} j_n = \infty$. Pretože $\sum_{i=1}^{\infty} p_i = 1$, platí

$$\lim_{n \to \infty} F(z_n) = \lim_{n \to \infty} \sum_{x_i < z_n} p_i = \lim_{n \to \infty} \sum_{i=1}^{j_n} p_i = 1. \quad \clubsuit$$

Pretože F z predchádzajúcej vety je distribučná funkcia, existuje náhodná veličina, ktorá má túto distribučnú funkciu. Takúto náhodnú veličinu nazývame *diskrétna náhodná veličina* alebo *náhodná veličina diskrétneho typu*.

Pre rozdelenie pravdepodobnosti diskrétnej náhodnej veličiny s pravdepodobnostnou funkciou $(x_i, p_i)_{i=1}^{\infty}$ platí

$$P_X(B) = \sum_{x_i \in B} p_i. \tag{5.1}$$

Táto náhodná veličina nadobúda s nenulovými pravdepodobnosťami práve tie reálne hodnoty x_i, pre ktoré je $p_i > 0$, pritom pre tieto hodnoty je $P(\{X = x_i\}) = p_i$ a pre ľubovoľné $x \notin \{x_1, x_2, ...\}$ je $P(\{X = x\}) = 0$. Samozrejme platí pre ľubovoľné $x \in \mathtt{R}$, že $P(\{X = x\}) = F(x + 0) - F(x)$.

Predchádzajúce úvahy majú aj takú interpretáciu, že ak máme reálnu funkciu X, ktorá nadobúda hodnoty z množiny $M = \{x_1, x_2, ...\}$ s pravdepodobnosťami $P(\{X = x_i\}) = p_i$, pričom $p_i > 0,\ i = 1, 2, ...$ a $\sum_{i=1}^{\infty} p_i = 1$, tak X je náhodná premenná s rozdelením pravdepodobnosti (5.1).

Príklady diskrétnych náhodných veličín

Náhodná veličina s alternatívnym rozdelením pravdepodobnosti

(Alternatívne rozdelenie pravdepodobnosti)

Majme $M = \{0, 1\}$, (teda $x_1 = 0, x_2 = 1$) a ďalej $p_1 = 1 - \theta,\ p_2 = \theta$, pričom $\theta \in < 0, 1 >$. Nazveme $0-$ neúspech a $1-$úspech. Potom funkcia X (schválne nehovoríme kde je definovaná), ktorá nadobúda hodnoty 0 a 1 s pravdepodobnosťami $P(X = 0) = 1 - \theta$ a $P(X = 1) = \theta$ je náhodná premenná. Rozdelenie pravdepodobnosti tejto náhodnej premennej sa nazýva *alternatívne rozdelenie pravdepodobnosti* s parametrom θ a píšeme $X \sim A(\theta)$. Modelujeme (matematicky popisujeme) ním situáciu, keď máme pokus s dvomi možnými výsledkami - "úspechom" a "neúspechom". Pravdepodobnosť úspechu je θ a neúspechu $1-\theta$. Jej distribučná funkcia je

Ľahko skonštruujeme v tomto prípade priestor elementárnych javov $\Omega = \{\omega_1, \omega_2\}$ a $\sigma-$algebru náhodných javov $\mathcal{A} = \{\{\emptyset\}, \{\omega_1\}, \{\omega_2\}, \Omega\}$. Pravdepodobnosť $P(\{\emptyset\})$
$= 0, P(\{\omega_1\}) = 1 - \theta, P(\{\omega_2\}) = \theta, P(\Omega\}) = 1$. Náhodná veličina X je definovaná nasledovne:

$$X(\omega_1) = 0, \quad X(\omega_2) = 1.$$

Pravda, toto všetko už "nepotrebujeme". Stačí nám poznať pravdepodobnostnú funkciu náhodnej veličiny X.

Binomické rozdelenie pravdepodobnosti

Majme $M = \{0, 1, 2, ..., n\}$ (teda $x_1 = 0, x_2 = 1, ..., x_{n+1} = n$) a $p_{x+1} = p(x) = \binom{n}{x}\theta^x(1-\theta)^{n-x} > 0$ pre $x = 0, 1, 2, ..., n$, $\theta \in < 0, 1 >$. Zrejme

$$\sum_{j=1}^{n+1} p_j = \sum_{x=0}^{n} \binom{n}{x}\theta^x(1-\theta)^{n-x} = [\theta + (1-\theta)]^n = 1.$$

Náhodná veličina X, ktorá nadobúda hodnoty $\{0, 1, ..., n\}$ s pravdepodobnosťami $P(X = x) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$, $x = 0, 1, ..., n$ má *binomické rozdelenie pravdepodobnosti* s parametrami n, θ. Označujeme $X \sim \mathrm{Bi}(n, \theta)$.

Ak uvažujeme experiment, ktorý pozostáva z n nezávislých alternatívnych pokusov, v ktorých nás zaujíma len nastatie alebo nenastatie náhodného javu A (pravdepodobnosť nastatia javu A v jednotlivom alternatívnom pokuse je $\theta \in < 0, 1 >$), potom

$$X \ - \text{počet nastaní náhodného javu } A \text{ v experimente je } x$$

$x = 0, 1, 2..., n$, je diskrétna náhodná veličina a $X \sim Bi(n, \theta)$. Dokážte to ako cvičenie.

Poissonovo rozdelenie pravdepodobnosti

Majme $M = \{0, 1, 2, ..., \}$ a $p_{x+1} = p(x) = e^{-\lambda}\frac{\lambda^x}{x!} > 0$ pre $x = 0, 1, 2, ...$, $\theta > 0$. Zrejme

$$\sum_{j=1}^{\infty} p_j = \sum_{x=0}^{\infty} e^{-\lambda}\frac{\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} = e^{-\lambda}e^{\lambda} = 1.$$

Náhodná veličina X, ktorá nadobúda hodnoty $\{0, 1, ...\}$ s pravdepodobnosťami $P(X = x) = e^{-\lambda}\frac{\lambda^x}{x!}$, $x = 0, 1, ...,$ má *Poissonovo rozdelenie pravdepodobnosti* s parametrom λ. Označujeme $X \sim Po(\lambda)$. Takáto náhodná veličina popisuje napríklad výskyt "riedkych javov", počet organizmov v jednotke pôdy, počet listov na strome, počet havárií, počet prerušení výroby, počet hovorov v telefónnej sieti, atď.

Veta 5.1. (Poissonova) Ak $X_n \sim \mathrm{Bi}(n, p_n)$, kde $\lim_{n\to\infty} np_n = \lambda > 0$, $p_n \in (0, 1)$ a $X \sim \mathrm{Po}(\lambda)$, tak pre $k = 0, 1, 2, ...$ platí

$$\lim_{n\to\infty} P(\{X_n = k\}) = P(\{X = k\}).$$

Dôkaz: Pre $k = 0, 1, 2, ...$ platí

$$\lim_{n\to\infty} P(\{X_n = k\}) = \lim_{n\to\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k} =$$

$$= \lim_{n\to\infty}\frac{1}{k!}\frac{np_n(n-1)p_n...(n-k+1)p_n}{(1-p_n)^k}(1-p_n)^n = \frac{\lambda^k}{k!}e^{-\lambda},$$

lebo

$$(1-p_n)^n = \left(1 - \frac{1}{\frac{1}{p_n}}\right)^n = \left[\left(1 - \frac{1}{\frac{1}{p_n}}\right)^{\frac{1}{p_n}}\right]^{np_n} \to \left(e^{-1}\right)^{\lambda} = e^{-\lambda}. \ \clubsuit$$

Negatívne binomické rozdelenie pravdepodobnosti a

geometrické rozdelenie pravdepodobnosti

Majme $M = \{0, 1, 2, ...,\}$ a $p_{x+1} = p(x) = \binom{x+r-1}{r-1} p^r (1-p)^x =$
$= \binom{x+r-1}{x} p^r (1-p)^x > 0$ pre $x = 0, 1, 2, ...,$ $p \in (0, 1)$, $r \in \mathtt{N}$. Z Taylorovho rozvoja (MacLaurinov rad) funkcie $(1-z)^{-k} = \sum_{x=0}^{\infty} \frac{[(1-z)^{-k}]_{z=0}^{(x)}}{x!} z^x =$
$= \sum_{j=0}^{\infty} \binom{k+j-1}{j} z^j$, $k \in \mathtt{N}$, $|\mathtt{z}| < \mathtt{1}$ zrejme

$$\sum_{x=0}^{\infty} \binom{x+r-1}{r-1} p^r (1-p)^x = p^r \sum_{x=0}^{\infty} \binom{x+r-1}{r-1} (1-p)^x = p^r (1-(1-p))^{-r} = 1.$$

Náhodná veličina X, ktorá nadobúda hodnoty $\{0, 1, ...\}$ s pravdepodobnosťami $P(X = x) = \binom{x+r-1}{r-1} p^r (1-p)^x$, $x = 0, 1, ...,$ má *negatívne binomické rozdelenie pravdepodobnosti* s parametrami r, p. Označujeme $X \sim NeBi(r, p)$.

Ak uvažujeme experiment, ktorý pozostáva z nezávislých alternatívnych pokusov, v ktorých nás zaujíma len nastatie alebo nenastatie náhodného javu $A-$úspech (pravdepodobnosť nastatia úspechu v jednotlivom alternatívnom pokuse je $p \in (0, 1)$), potom

$$X \ - \text{počet neúspechov, ktoré predchádzajú } r-\text{tému úspechu}$$

je diskrétna náhodná veličina a $X \sim NeBi(r, p)$ s hodnotami $x = 0, 1, 2, ...,$. Dokážte to ako cvičenie.

Špeciálnym prípadom negatívneho binomického rozdelenia pre $r = 1$ je *geometrické rozdelenie pravdepodobnosti*. Náhodná veličina X, ktorá nadobúda hodnoty $\{0, 1, ...\}$ s pravdepodobnosťami $P(X = x) = p(1-p)^x$, $x = 0, 1, ...,$ $p \in (0, 1)$, má geometrické rozdelenie pravdepodobnosti s parametrom p. Označujeme $X \sim Ge(p)$. Ak uvažujeme experiment, ktorý pozostáva z nezávislých alternatívnych pokusov, v ktorých nás zaujíma len nastatie alebo nenastatie náhodného javu $A-$úspech (pravdepodobnosť nastatia úspechu v jednotlivom alternatívnom pokuse je $p \in (0, 1)$), potom

$$X \ - \text{počet neúspechov pred prvým úspechom}$$

je diskrétna náhodná veličina, $X \sim Ge(p)$ s hodnotami $x = 0, 1, 2, ...$.

Hypergeometrické rozdelenie pravdepodobnosti

Majme $N \in \mathtt{N}$, $(N \geq 2)$ súčiastok, z ktorých je $A \in \mathtt{N}$ chybných, pričom $N > A$. Zo všetkých N súčiastok náhodne vyberieme $n \in \mathtt{N}$ súčiastok (bez vrátenia), pričom $n \leq N$. Náhodná premenná

$$X \ - \ \text{počet chybných súčiastok medzi } n \text{ vytiahnutými}$$

má *hypergeometrické rozdelenie pravdepodobnosti* s parametrami N, A, n. Označujeme to $X \sim Hg(N, A, n)$. Samozrejme musíme sa presvedčiť, že takto popísaná funkcia X je skutočne náhodná premenná a definovať hypergeometrické rozdelenie pravdepodobnosti.

Najprv si uvedomme, že môžu nastať práve 4 prípady, a síce

(i) $n \leq A, \ \ n \leq N - A$ (počet dobrých súčiastok) vtedy X nadobúda hodnoty $x \in \{0, 1, ..., n\}$

(ii) $n \leq A, \ \ n > N - A$ vtedy X nadobúda hodnoty $x \in \{n - (N - A) = n - N + A$ (najmenej chybných)$, n - N + A + 1, ..., n$ (najviac chybných) $\}$

(iii) $n > A, \ \ n \leq N - A$ vtedy X nadobúda hodnoty $x \in \{0, 1, ..., A\}$

(iv) $n > A, \ \ n > N - A$ vtedy X nadobúda hodnoty $x \in \{n - N + A, n - N + A + 1, ..., A\}$

Teda $x-$ počet chybných súčiastok medzi n vytiahnutými je z intervalu $< k_1, k_2 >$, kde $k_1 = \max(0, n - N + A)$ a $k_2 = \min(A, n)$. Počet možných vytiahnutých $n-$tíc je $\binom{N}{n}$. Medzi n vybratými súčiastkami (teda vo vybratej $n-$tici) je x chybných $\binom{A}{x}$ spôsobmi a ku každému spôsobu je $\binom{N-A}{n-x}$ možností vybratia bezchybných, teda

$$P(\{X = x\}) = \frac{\binom{A}{x}\binom{N-A}{n-x}}{\binom{N}{n}}, \tag{5.2}$$

$x \in < \max(0, n - N + A), \min(A, n) >$.

Dôkaz toho, že (5.2) je rozdelenie pravdepodobnosti vyplýva z identity

$$\sum_{\alpha=\max\{0,n-N+A\}}^{\min\{n,A\}} \binom{N-A}{n-\alpha}\binom{A}{\alpha} = \binom{N}{n}, \tag{5.3}$$

ktorú dokažeme pomocou nasledujúcej lemy. Najprv si ale zadefinujeme klesajúci faktoriál reálneho čísla x. Ak $k \in \mathtt{N}_0, x \in \mathtt{R}$, tak klesajúci faktoriál $x_{(0)} = 0$ a pre $k \in \mathtt{N}$ je $x_{(k)} = x(x-1)...(x-k+1)$. Teraz kombinačné číslo $\binom{n}{k}$ môžeme písať ako $\frac{n_{(k)}}{k!}$ a "rozšírili" sme pojem kombinačného čísla $\binom{n}{k}$ tak, že namiesto $n \in \mathtt{N}_0$ môžeme uvažovať $n \in \mathtt{R}$ (samozrejme $k \in \mathtt{N}_0$ zostáva v platnosti).

Lema 5.1 Pre ľubovoľné reálne čísla x, y a $n \in \mathtt{N}_0$ platí

$$\sum_{k=0}^{n} \binom{x}{k}\binom{y}{n-k} = \binom{x+y}{n} \tag{5.4}$$

(Cauchyho kombinatorický vzorec).

Dôkaz:

1. Pre $n = 0$ je identita zrejmá.
2. Nech teda platí pre nejaké $n \in \mathtt{N}$ a dokážeme, že

$$\sum_{k=0}^{n+1} \binom{x}{k}\binom{y}{n+1-k} = \binom{x+y}{n+1}. \tag{5.5}$$

Vieme, že pre ľubovoľné $a \in \mathtt{R}, n \in \mathtt{N}_0$ je

$$\binom{a}{n+1} = \frac{a-n}{n+1}\binom{a}{n}, \quad \text{teda} \quad \binom{x+y}{n+1} = \frac{x+y-n}{n+1}\binom{x+y}{n},$$

preto počítajme

$$\frac{x+y-n}{n+1}\sum_{k=0}^{n}\binom{x}{k}\binom{y}{n-k} = \frac{1}{n+1}\sum_{k=0}^{n}[y-n+k+x-k]\binom{x}{k}\binom{y}{n-k} =$$

$$= \frac{1}{n+1}\left\{\sum_{k=0}^{n}\binom{x}{k}\binom{y}{n-k}\frac{y-n+k}{n-k+1}(n-k+1) + \right.$$

$$\left. + \sum_{k=0}^{n}\binom{x}{k}\frac{x-k}{k+1}\binom{y}{n-k}(k+1)\right\} =$$

$$= \frac{1}{n+1}\left\{\sum_{k=0}^{n}\binom{x}{k}\binom{y}{n-k+1}(n-k+1) + \sum_{k=0}^{n}\binom{x}{k+1}\binom{y}{n-k}(k+1)\right\} =$$

$$= \frac{1}{n+1}\left\{\binom{x}{0}\binom{y}{n+1}(n+1) + \sum_{k=1}^{n}\binom{x}{k}\binom{y}{n-k+1}(n-k+1) + \right.$$

$$\left. + \sum_{k=0}^{n-1}\binom{x}{k+1}\binom{y}{n-k}(k+1) + \binom{x}{n+1}\binom{y}{0}(n+1)\right\} =$$

$$= \frac{1}{n+1}\left\{\binom{x}{0}\binom{y}{n+1}(n+1) + \sum_{k=1}^{n}\binom{x}{k}\binom{y}{n-k+1}(n-k+1) + \right.$$

$$\left. + \sum_{j=1}^{n}\binom{x}{j}\binom{y}{n-j+1}j + \binom{x}{n+1}\binom{y}{0}(n+1)\right\} =$$

$$= \frac{1}{n+1}\left\{\binom{x}{0}\binom{y}{n+1}(n+1) + \sum_{k=1}^{n}\binom{x}{k}\binom{y}{n-k+1}(n+1) + \right.$$

$$\left. + \binom{x}{n+1}\binom{y}{0}(n+1)\right\} = \sum_{k=0}^{n+1}\binom{x}{k}\binom{y}{n+1-k} = \binom{x+y}{n+1}. \quad \clubsuit$$

6. Spojité náhodné veličiny

(náhodné veličiny (absolútne) spojitého typu)

Najprv si zopakujeme určité tvrdenia z matematickej analyzy týkajúce sa absolútne spojitej funkcie.

Definícia 6.1. Funkcia $F(\cdot)$ je absolútne spojitá (na $\mathtt{R}$), ak k ľubovoľnému $\epsilon > 0$ existuje také $\delta > 0$, že pre každú postupnosť $a_1 < b_1 < a_2 < b_2 < \ldots < a_n < b_n$ takú že $\sum_{i=1}^{n}(b_i - a_i) < \delta$ platí $\sum_{i=1}^{n}|F(b_i) - F(a_i)| < \epsilon$.

Vlastnosti absolútne spojitej funkcie:

(i) Ak je F absolútne spojitá, tak je spojitá.

(ii) Ak je F absolútne spojitá, tak má skoro všade (vzhľadom na Lebesgueovu mieru) vlastnú deriváciu. Táto derivácia je integrovateľná v Lebesgueovom zmysle a platí $F(x) = \int_a^x F'(t)dt + F(a)$ pre každé $a \in \mathtt{R}$.

(iii) Ak je F absolútne spojitá a platí $F'(x) = 0$ skoro všade (vzhľadom na Lebesgueovu mieru), potom je F konštantná skoro všade (vzhľadom na Lebesgueovu mieru).

(iv) Ak je F neurčitým integrálom funkcie f (v Lebesgueovom zmysle, teda $F(x) = \int f(x)dx$), potom je F absolútne spojitá a platí $F'(x) = f(x)$ skoro všade (vzhľadom na Lebesgueovu mieru).

(v) Ak je F absolútne spojitá, tak má na každom konečnom intervale $< a, b >$ konečnú variáciu, t.j. $\sup \sum_{j=1}^{N} |F(x_j) - F(x_{j-1})| < \infty$, pričom supremum sa berie cez všetky N a konečné postupnosti $a = x_0 < x_1 < \ldots < x_N = b$.

Teraz si zadefinujeme absolútne spojitú náhodnú veličinu.

Definícia 6.2. Povieme, že náhodná veličina X definovaná na $(\Omega, \mathcal{A}, P)$ je *absolútne spojitého typu (spojitá)*, ak existuje (nezáporná) integrovateľná funkcia $f(\cdot)$ taká, že pre každú borelovskú množinu $B \in \mathcal{B}$ je

$$P_X(B) = \int_B f(x)dx.$$

Funkciu f nazývame *hustotou rozdelenia pravdepodobnosti (hustotou)* náhodnej veličiny X.

Veta 6.1. (Vlastnosti hustoty.) Nech X je náhodná veličina absolútne spojitého typu, f je jej hustota a F jej distribučná funkcia. Potom

(i) $\int_{-\infty}^{\infty} f(x)dx = 1$;

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

(iii) $F(\cdot)$ je absolútne spojitá funkcia;

(iv) hustota $f(\cdot)$ je určená jednoznačne skoro všade vzhľadom k Lebesgueovej miere, t.j. ak f a g sú hustoty náhodnej veličiny X, tak $\mu(\{x : f(x) \neq g(x)\}) = 0$, kde μ je Lebesgueova miera;

(v) existuje $F'(x)$ skoro všade vzhľadom k Lebesgueovej miere μ a funkcia $g(x) = F'(x)$ je hustota náhodnej veličiny X;

(vi) pre $a < b$ platí $F(b) - F(a) = \int_a^b f(x)dx$ a tiež $P(\{a \leq X < b\}) = P(\{a < X < b\}) = P(\{a < X \leq b\}) = P(\{a \leq X \leq b\}) = \int_a^b f(x)dx$;

(vii) ak existuje v bode x derivácia $F'(x) = f(x)$, potom $P(x - \frac{h}{2} \leq X < x + \frac{h}{2}) = hf(x) + o(h)$, kde $o(h)$ je taká funkcia, pre ktorú platí $\lim_{h\to 0} \frac{o(h)}{h} = 0$;

(viii) $f(x) \geq 0$ pre každé $x \in \mathtt{R}$ skoro všade vzhľadom k Lebesgueovej miere.

Dôkaz:

(i) Ak má X hustotu f, tak z definície aboslútne spojitej náhodnej veličiny vyplýva, že $\forall\, B \in \mathcal{B}\ \ P_X(B) = \int_B f(x)dx$. Ak vezmeme $B = \mathtt{R}$, tak $1 = P_X(\{\mathtt{R}\}) = \int_{\mathtt{R}} f(x)dx = \int_{-\infty}^{\infty} f(x)dx$.

(ii) Vieme, že $F(x) = P(\{X < x\}) = P_X((-\infty, x)) = \int_{-\infty}^{x} f(t)dt$.

(iii) Tvrdenie z matematickej analyzy: Ak pre funkciu F platí $F(x) = \int_{-\infty}^{x} f(t)dt$ pre každé $x \in \mathtt{R}$, tak F je absolútne spojitá na $\mathtt{R}$.

(iv) Ak f a g sú hustoty náhodnej veličiny X, tak pre každú $B \in \mathcal{B}$ platí $P_X(B) = \int_B f(x)dx = \int_B g(x)dx$. Z toho dostávame, že pre každú $B \in \mathcal{B}$ platí $\int_B (f(x) - g(x))dx = 0$, čiže $\mu(\{x :\ f(x) \neq g(x)\}) = 0$.

(v) Tvrdenie je dôsledkom absolútnej spojitosti ((ii) vlastnosť absolútne spojitej funkcie).

(vi) Podľa Vety 4.1., Vety 6.1.(ii) a aditívnej vlastnosti integrálu platí $P(\{a \leq X < b\}) = F(b) - F(a) = \int_{-\infty}^{b} f(t)dt - \int_{-\infty}^{b} f(t)dt = \int_{a}^{b} f(t)dt$. Distribučná funkcia F je absolútne spojitá, preto je spojitá a pre každé $x \in \mathtt{R}$ podľa (4.3) platí $P(X = x) = 0$, z čoho ľahko dostaneme ostatné vzťahy.

(vii) Ak napíšeme pre ľubovoľné $h > 0$ a x také, že existuje $F'(x) = f(x)$
$\frac{o(h)}{h} = \frac{P(\{x - \frac{h}{2} \leq X < x + \frac{h}{2}\})}{h} - f(x)$, tak platí $\lim_{h \to 0} \frac{o(h)}{h} = \lim_{h \to 0} \frac{F(x + \frac{h}{2}) - F(x - \frac{h}{2})}{h} - \lim_{h \to 0} f(x) = f(x) - f(x) = 0$.

(viii) Funkcia f je nezáporná, lebo distribučná funkcia F je neklesajúca - dôkaz v matematickej analýze.

♣

Predstavu o hustote dá nasledujúci vzťah

$$\int_{x}^{x+\Delta x} f(t)dt \doteq f(x)\Delta x,$$

alebo aj

$$P(x \leq X < x + dx) \doteq f(x)dx.$$

Príklady (absolútne) spojitých náhodných veličín

Náhodná veličina s rovnomerným rozdelením

Náhodná veličina X má *rovnomerné rozdelenie pravdepodobnosti* na intervale (a, b) (pričom $-\infty < a < b < \infty$), ak jej hustota je

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & \text{ak } a < x < b \\ 0, & \text{ak } x \notin (a, b). \end{cases}$$

Značíme $X \sim Ro(a, b)$. Distribučná funkcia X je

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

Interpretácia takejto náhodnej veličiny je názorná: Istota (jednotková pravdepodobnosť) je na intervale (a, b) rovnomerne "rozprestrená".

Náhodná veličina s exponenciálnym rozdelením

Nech náhodný jav A sa vyskytuje v náhodných okamžikoch (napr. prerušenie výroby, vyhorenie žiarovky, prelet častice, atď.) Výskyty tohto náhodného javu A v neprekrývajúcich sa časových intervaloch sú nezávislé. Označme

$Q(t)$ – pravdepodobnosť, že sledovaný jav A nenastane v priebehu časového intervalu dĺžky t

Ak t_1, t_2 sú dĺžky dvoch na seba nadväzujúcich časových intervalov, tak

$$Q(t_1 + t_2) = Q(t_1)Q(t_2)$$

$P\{A$ nenastane v priebehu času t_1 a súčasne nenastane v priebehu času $t_2\}$. Nech Q je diferencovateľná funkcia času a pre $t = 0$ nadobúda maximum, teda $Q(0) = 1$.

Pre $t \geq 0,\ \Delta t > 0$ je

$$\ln Q(t + \Delta t) = \ln Q(t) + \ln Q(\Delta t),$$

čiže pre $t \geq 0$ je

$$(\ln Q(t))' = \lim_{\Delta t \to 0^+} \frac{\ln Q(t + \Delta t) - \ln Q(t)}{\Delta t} = \lim_{\Delta t \to 0^+} \frac{\ln Q(\Delta t)}{\Delta t} =$$

$$= \lim_{\Delta t \to 0^+} \frac{\ln Q(0 + \Delta t) - \ln Q(0)}{\Delta t} = [\ln Q(t)]'_{t=0} = -\lambda$$

(ide o deriváciu sprava, ktorú označíme $-\lambda$, pričom $\lambda > 0$). Máme teda diferenciálnu rovnicu s počiatočnou podmienkou

$$\frac{d \ln Q(t)}{dt} = -\lambda$$

$$Q(0) = 1.$$

Jej riešenie je $Q(t) = e^{-\lambda t}$. Označme

X – náhodnú veličinu – čas, keď nastane prvýkrát sledovaný jav

Zrejme

$$F_X(t) = P(\{X < t\}) = P(\text{jav } A \text{ nastane v čase } (0, t)) = 1 - Q(t)$$

(tuná $Q(t)$ je pravdepodobnosť, že sledovaný jav nenastane v intervale $(0, t)$), teda

$$F_X(t) = \begin{cases} 1 - e^{-\lambda t}, & \text{ak } t > 0 \\ 0, & \text{ak } t \leq 0. \end{cases}$$

Náhodná veličina X má *exponenciálne rozdelenie pravdepodobnosti* s parametrom λ a označujeme $X \sim ex(\lambda)$. Jej hustota je

$$f_X(t) = \begin{cases} \lambda e^{-\lambda t}, & \text{ak } t > 0 \\ 0, & \text{ak } t \leq 0 \end{cases}$$

(dostaneme derivovaním F).

Náhodná veličina s normálnym rozdelením (normálna náhodná veličina,
gaussovská náhodná veličina)

Ak má náhodná veličina X hustotu

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

$\mu \in (-\infty, \infty)$, $\sigma^2 > 0$, tak povieme, že X má *normálne (Gaussovo, gaussovské) rozdelenie pravdepodobnosti* s parametrami μ, σ^2 a píšeme $X \sim N(\mu, \sigma^2)$.

V prípade $\mu = 0$ a $\sigma = 1$ ide o *standardizovanú normálnu náhodnú veličinu*, čo označujeme $X \sim N(0,1)$. Jej hustota je

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

Normálne rozdelenie má náhodná veličina, ktorá vznikla súčtom veľkého počtu nezávislých náhodných veličín (o rozdelenia ktorých stači predpokladať určité veľmi všeobecné predpoklady). Normálne rozdelenie má veľmi dôležitú úlohu v teórii pravdepodobnosti a matematickej štatistike. Napríklad normálne rozdelená je náhodná chyba meracieho prístroja, chyba pri streľbe na cieľ, telesná výška jedincov homogénnej populácie, atď. Poznamenávame len, že skutočnosť, že $f(x)$ je hustota vyplýva z rovnosti $\int_0^\infty e^{-a^2x^2}dx = \dfrac{\sqrt{\pi}}{2a}$, $a > 0$.

Náhodná veličina s gama rozdelením

Ak má náhodná veličina X hustotu

$$f(x) = \begin{cases} \frac{a^p}{\Gamma(p)} e^{-ax} x^{p-1}, & \text{ak } x > 0 \\ 0, & \text{ak } x \le 0 \end{cases}$$

$a > 0$, $p > 0$, tak povieme, že X má *gama rozdelenie pravdepodobnosti* s parametrami a, p.

Gama funkcia $\Gamma(a)$ je definovaná predpisom $\Gamma(a) = \int_0^\infty x^{a-1}e^{-x}dx$, $a > 0$. Jej najčastejšie používané vlastnosti sú

$\Gamma(a+1) = a\Gamma(a)$, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, $\Gamma(n) = (n-1)!$ pre $n \in \mathtt{N}$.

5. Náhodná veličina s beta rozdelením

Ak má náhodná veličina X hustotu

$$f(x) = \begin{cases} \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}, & \text{ak } 0 < x < 1 \\ 0, & \text{inak} \end{cases}$$

$a > 0$, $b > 0$, tak povieme, že X má *beta rozdelenie pravdepodobnosti* s parametrami a, b.

Beta funkcia $B(a,b)$ je definovaná predpisom $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$, $a > 0, b > 0$. Vzťah medzi gama a beta funkciou je vyjadrený nasledovne: $B(a,b) = \dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

Poznámka o distribučných funkciách

Diskrétne a spojité náhodné veličiny (resp. distribučné funkcie diskrétnych a spojitých náhodných veličín) predstavujú dve prakticky veľmi dôležité triedy. Vo všeobecnosti ale o distribučných funkciách platí veta (nebudeme ju dokazovať, pozri napr. Rudin, W., Analýza v reálném a komplexním oboru, Academia, Praha, 1977)

Veta 6.2. Nech X je náhodná veličina s distribučnou funkciou F. Potom F sa dá napísať v tvare

$$F(x) = a_1 F_\alpha(x) + a_2 F_a(x) + a_3 F_s(x)$$

$a_1, a_2, a_3 \geq 0, a_1 + a_2 + a_3 = 1$, pričom $F_\alpha(\cdot)$ je distribučná funkcia diskrétnej náhodnej veličiny, $F_a(\cdot)$ je distribučná funkcia absolútne spojitej náhodnej veličiny a $F_s(\cdot)$ je distribučná funkcia singulárne spojitej náhodnej veličiny.

Povieme, že F je singulárne spojitá, ak je spojitá a pritom existuje borelovská množina B Lebesgueovej miery 0 a μ_F miery 1. Takáto funkcia má skoro všade (vzhľadom na Lebesgueovu mieru) deriváciu rovnú 0 a je spojitá v $\mathtt{R}$. Napríklad Cantorova funkcia je spojitá, diferencovateľná, rastúca, deriváciu má nulovú s výnimkou množiny Lebesgueovej miery 0. Takáto funkcia funkcia je spojitá a nie je absolútne spojitá.

7. Náhodné vektory

Máme často nielen jednu náhodnú veličinu, ale súčasne niekoľko náhodných veličín. Zaujíma nás, či niektoré z nich spolu "akosi" súvisia, či (zname) hodnoty jednej náhodnej veličiny (resp. určitej skupiny náhodných veličín) vedia niečo povedať o hodnote inej náhodnej veličiny (iných náhodných veličín). Snažíme sa vyšetrovať (aj) závislosť. Potrebujeme model, v ktorom pracujeme s niekoľkými náhodnými veličinami súčasne.

Zopakujme si: $(\Omega, \mathcal{A}, P)$ je pravdepodobnostný priestor

$X: \quad \Omega \to \mathtt{R}$, pre ktorú platí $x \in \mathtt{R} \Longrightarrow \; \{\omega \in \Omega : \quad X(\omega) < x\} \in \mathcal{A}$

je náhodná veličina.

Rozšírme na mnohorozmerný prípad: $(\Omega, \mathcal{A}, P)$ je pravdepodobnostný priestor

$$\mathbf{X} = \begin{pmatrix} X_1(\cdot) \\ X_2(\cdot) \\ \vdots \\ X_n(\cdot) \end{pmatrix} : \quad \Omega \to \mathtt{R}^n,$$

a označme

$$[\mathbf{X} < \mathbf{x}] = \{\omega \in \Omega : \quad X_1(\omega) < x_1, ..., X_n(\omega) < x_n\}.$$

$\mathcal{B}^n$ nech je najmenšia $\sigma-$algebra nad intervalmi tvaru

$$(-\infty, x_1) \times (-\infty, x_2) \times ... \times (-\infty, x_n)$$

pre ľubovoľné $\mathbf{x} \in \mathtt{R}^n$ (t.j. $\mathbf{x} = (x_1, ..., x_n)'$). Nazývame ju borelovská $\sigma-$algebra v $\mathtt{R}^n$.

Definícia 7.1. Majme pravdepodobnostný priestor $(\Omega, \mathcal{A}, P)$. Reálna vektorová funkcia $\mathbf{X}(\cdot)$ definovaná na Ω s hodnotami v $\mathtt{R}^n$, pre ktorú platí

$$\mathbf{x} \in \mathtt{R}^n \Longrightarrow [\mathbf{X} < \mathbf{x}] \in \mathcal{A}$$

sa nazýva *náhodný vektor* (vektor náhodných veličín, $n-$rozmerná náhodná veličina, vektorová náhodná veličina).

Definícia 7.2. Nech $\mathbf{X} = (X_1, ..., X_n)'$ je $n-$rozmerný náhodný vektor definovaný na pravdepodobnostnom priestore $(\Omega, \mathcal{A}, P)$. Reálnu funkciu

$$F(x_1, x_2, ..., x_n) = P(X_1 < x_1, ..., X_n < x_n) = P([\mathbf{X} < \mathbf{x}])$$

definovanú pre každé $\mathbf{x} \in \mathtt{R}^n$ nazývame *distribučnou funkciou náhodného vektora* $\mathbf{X}$.

Označenie:

$\Delta_h^{(i)} F(x_1, ..., x_n) = F(x_1, ..., x_{i-1}, x_i + h, x_{i+1}, ..., x_n) - F(x_1, ..., x_n)$ je *diferencia* funkcie F v premennej x_i s krokom $h \geq 0$. ďalej označme rekurentne

$$\Delta_{h_j}^{(j)} \Delta_{h_i}^{(i)} F(x_1, ..., x_n) = \Delta_{h_j}^{(j)} [\Delta_{h_i}^{(i)} F(x_1, ..., x_n)] =$$

$$= \Delta_{h_j}^{(j)} [F(x_1, ..., x_{i-1}, x_i + h_i, x_{i+1}, ..., x_n) - F(x_1, ..., x_n)] =$$

$$= F(x_1, ..., x_i + h_i, ..., x_j + h_j, ..., x_n) - F(x_1, ..., x_j + h_j, ..., x_n) -$$

$$-[F(x_1,...,x_i+h_i,...,x_n)-F(x_1,...,x_n)] = \Delta_{h_i}^{(i)}\Delta_{h_j}^{(j)}F(x_1,...,x_n).$$

Vlastnosti distribučnej funkcie popisuje nasledujúca veta

Veta 7.1. Distribučná funkcia $F_{\mathbf{X}}$ $n-$rozmerného náhodného vektora má tieto vlastnosti:

(i) $\lim_{x_i\to\infty,\ 1\le i\le n} F_{\mathbf{X}}(x_1,...,x_n) = 1$,

(ii) pre $i = 1,2,...,n$ je $\lim_{x_i\to-\infty} F_{\mathbf{X}}(x_1,...,x_n) = 0, \quad \forall x_1,...,x_{i-1},x_{i+1},...x_n$,

(iii) $F_{\mathbf{X}}$ je spojitá zľava v každej premennej,

(iv) pre ľubovoľné reálne $x_1,...,x_n$ a ľubovoľné $h_k \ge 0,\ (k = 1,2,...,n)$ platí $\Delta_{h_1}^{(1)}\Delta_{h_2}^{(2)}...\Delta_{h_n}^{(n)}F(x_1,...,x_n) \ge 0$.

Dôkaz nájdeme napr. v (Dupač, V., Hušková, M., Pravděpodobnost a matematická statistika, Karolinum, Praha, 2001 alebo Rényi, A., Teorie pravděpodobnosti, Academia, Praha, 1972).

Poznámka. Platí

$$\Delta_{h_1}^{(1)}\Delta_{h_2}^{(2)}...\Delta_{h_n}^{(n)}F(x_1,...,x_n) = P\left(\bigcap_{i=1}^{n}\{x_i \le X_i < x_i+h_i\}\right)$$

(dôkaz pozrite napr. v Dupač, V., Hušková, M., Pravděpodobnost a matematická statistika, Karolinum, Praha, 2001). Poznamenávame, že z (iv) a (ii) plynie, že $F_{\mathbf{X}}$ je neklesajúca funkcia v každej premennej. Naopak to neplatí, t.j. ak je nejaká funkcia neklesajúca v každej premennej, neplynie z toho ešte (iv), lebo napr. vezmeme $n = 2$ a $F(x_1,x_2) = 1$ pre $x_1 \ge 0, x_2 \ge 0, x_1+x_2 \ge 1$ a $F(x_1,x_2) = 0$ inak, potom $F(x_1,x_2)$ je neklesajúca v každej premennej a $\Delta_1^{(1)}\Delta_1^{(2)}F(0,0) = \Delta_1^{(1)}[F(0,1)-F(0,0)] = F(1,1)-F(0,1)-F(1,0)+F(0,0) = 1-1-1+0 = -1$, čo nemôže byť (podľa predchádzajúcej poznámky) $P(0 \le X_1 < 1, 0 \le X_2 < 1)$, čiže táto funkcia F nie je distribučnou funkciou.

Analogicky ako v jednorozmernom prípade definujeme Lebesgueovu-Stieltjesovu mieru μ_F indukovanú distribučnou funkciou F na borelovských množinách $\mathcal{B}^n$ (položíme pre $n-$rozmerný interval $< a_1,b_1) \times < a_2,b_2) \times ... \times < a_n,b_n)$, kde $a_i < b_i,\ i = 1,2,...,n$, mieru $\mu_F(\{< a_1,b_1) \times < a_2,b_2) \times ... \times < a_n,b_n)\}) = \Delta_{b_1-a_1}^{(1)}\Delta_{b_2-a_2}^{(2)}...\Delta_{b_n-a_n}^{(n)}F(a_1,...,a_n)$ a jednoznačne ju rozšírime na všetky borelovské množiny v $\mathtt{R}^n$ tak, aby miera $n-$rozmerných intervalov bola zachovaná).

Platí aj nasledujúca veta:

Veta 7.2. Nech funkcia $F(x_1,...,x_n)$ spĺňa podmienky (i)-(iv) Vety 7.1. Potom existuje pravdepodobnostný priestor $(\Omega,\mathcal{A},P)$ a $n-$rozmerný náhodný vektor $\mathbf{X}$ tak, že $F_{\mathbf{X}} = F$.

Dôkaz vety je analogický ako v jednorozmernom prípade.

Definícia 7.3. Distribučná funkcia F (n premenných) sa nazýva diskrétna, ak existuje konečná alebo spočítateľná postupnosť $M = \{\mathbf{x}_m\}_{m\in J}$, kde J je konečná alebo spočítateľná indexová množina (pričom $\mathbf{x}_m \in \mathtt{R}^n$ sú navzájom rôzne) a zodpovedajúca postupnosť kladných čísel $\{p_m\}_{m\in J}$ tak, že $\sum_{m\in J} p_m = 1$ a $F(\mathbf{x}) = \sum_{\mathbf{x}_m<\mathbf{x}} p_m$ pre všetky $\mathbf{x} \in \mathtt{R}^n$. (Nerovnosť $\mathbf{x}_m < \mathbf{x}$ znamená, že každá zložka vektora $\mathbf{x}_m$ je menšia ako príslušná zložka vektora $\mathbf{x}$.)

Platí (podobne ako v jednorozmernom prípade) pre každú $B \in \mathcal{B}^n$

$$\mu_F(B) = P(\mathbf{X} \in B) = P_{\mathbf{X}}(B) = \sum_{\mathbf{x}_m\in B} p_m.$$

Funkcia $(\mathbf{x}_m, p_m)_{m\in J}$ sa nazýva *pravdepodobnostná funkcia* náhodného vektora $\mathbf{X}$, ktorý nadobúda s nenulovými pravdepodobnosťami hodnoty $\{\mathbf{x}_m :\ m \in J\}$, pričom $P\{\mathbf{X} = \mathbf{x}_m\} = p_m,\ m \in J$.

Príklad 7.1.: Multinomické rozdelenie.

Uvažujme pokus, ktorý môže mať n rôznych disjunktných výsledkov $A_1, A_2, ..., A_n$. Nech $\theta_i = P(A_i),\ i = 1, 2, ..., n;\ 0 < \theta_i < 1,\ \sum_{i=1}^{n} \theta_i = 1$. Tento pokus budeme $k-$krát nezávisle opakovať. Označme

X_i-počet nastatí javu (výsledku) A_i v týchto k pokusoch.

Zrejme $X_i \in \{0, 1, 2, ..., k\},\ i = 1, 2, ..., n$, teda obor hodnôt X_i je $\{0, 1, 2, ..., k\}$. Pravdepodobnostná funkcia náhodného vektora $\mathbf{X}$ je

$$p(\mathbf{x}) = P(X_1 = x_1, ..., X_n = x_n) =$$
$$= \binom{k}{x_1}\binom{k - x_1}{x_2}...\binom{k - x_1 - x_2 - ... - x_{n-2}}{x_{n-1}}\theta_1^{x_1}...\theta_n^{x_n} =$$
$$= \frac{k!}{x_1!x_2!...x_n!}\theta_1^{x_1}...\theta_n^{x_n},$$

kde $x_i \in \{0, 1, ..., k\},\ \sum_{i=1}^{n} x_i = k$. Náhodný vektor $\mathbf{X}$ má multinomické rozdelenie pravdepodobnosti. Označujeme $\mathbf{X} \sim Mu_n(k, \theta_1, ..., \theta_n)$.

Definícia 7.4. Distribučná funkcia F (n premenných) sa nazýva absolútne spojitá, ak existuje funkcia $f:\ \mathtt{R}^n \to <0, \infty)$ taká, že

$$F(\mathbf{x}) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}...\int_{-\infty}^{x_n} f(t_1, ..., t_n)dt_n...dt_1$$

pre všetky $\mathbf{x} \in \mathtt{R}^n$. Funkcia f sa nazýva hustota.

Poznámka. Pre hustotu platí

(i) $f(\mathbf{x}) \geq 0$ pre skoro všetky $\mathbf{x} \in \mathtt{R}^n$ (vzhľadom na Lebesgueovu mieru), $\int_{\mathtt{R}^n} f(\mathbf{x})d\mathbf{x} = 1$,

(ii) pre každú $B \in \mathcal{B}^n$ platí $P_{\mathbf{X}}(B) = \int_B f(\mathbf{x})d\mathbf{x}$,

(iii) $f(x_1, ..., x_n) = \dfrac{\partial^n}{\partial x_1...\partial x_n}F(x_1, ..., x_n)$, pričom derivácia existuje skoro všade vzhľadom k Lebesgueovej miere.

Príklad 7.2. Náhodné vektory absolútne spojitého typu (s distribučnými funkciami absolútne spojitými).

$n-$rozmerný rovnomerný náhodný vektor $\mathbf{X}$ má hustotu

$$f(x_1, x_2, ..., x_n) = \begin{cases} \prod_{i=1}^{n} \frac{1}{\beta_i - \alpha_i}, & \text{pre } x_i \in (\alpha_i, \beta_i),\ i = 1, 2, ..., n \\ 0 & \text{inak,} \end{cases}$$

pričom $\alpha_i, \beta_i \in \mathtt{R},\quad \alpha_i < \beta_i,\ i = 1, 2, ..., n$. Označujeme $\mathbf{X} \sim Ro_n(\alpha_1, \beta_1, ..., \alpha_n, \beta_n)$.

$n-$rozmerný normálne rozdelený náhodný vektor $\mathbf{X}$ (náhodný vektor s regulárnym normálnym rozdelením) má hustotu

$$f(x_1, x_2, ..., x_n) = f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det \boldsymbol{\Sigma}}}e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})},$$

kde $\boldsymbol{\mu} = (\mu_1, ..., \mu_n)',\quad \mu_i \in \mathtt{R},\quad i = 1, 2, ..., n$ a $\boldsymbol{\Sigma}$ je pozitívne definitná matica. Značime $\mathbf{X} \sim N_n\ (\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Marginálne náhodné vektory

Náhodný jav $[\mathbf{X} < \mathbf{x}]$ môžeme písať aj ako $\bigcap_{i=1}^{n}[X_i < x_i]$. Ak si zvolíme pevné $j \in \{1, 2, ..., n\}$, tak pre (ľubovoľnú) postupnosť $x_j^{(k)} \to \infty$ je

$$A_1 = \bigcap_{\substack{i=1 \\ i \neq j}}^{n} [X_i < x_i] \cap [X_j < x_j^{(1)}] \subseteq A_2 \subseteq ...$$

a postupnosť náhodných javov $A_1, A_2, ...$ má $\lim_{n\to\infty} A_n = \bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$, pričom

$$\lim_{k\to\infty} A_k = \bigcup_{k=1}^{\infty} A_k = \bigcup_{k=1}^{\infty} \left(\bigcap_{\substack{i=1 \\ i \neq j}}^{n} [X_i < x_i] \cap [X_j < x_j^{(k)}] \right) = \bigcap_{\substack{i=1 \\ i \neq j}}^{n} [X_i < x_i] = A. \tag{7.1}$$

Preto $n-1$ rozmerný náhodný vektor $\mathbf{X}^{(-j)} = (X_1, ..., X_{j-1}, X_{j+1}, ..., X_n)'$ je opäť náhodným vektorom. Ak $F_{X_1,...,X_n}(x_1, ..., x_n) = P(X_1 < x_1, ..., X_n < x_n)$ je distribučná funkcia náhodného vektora $\mathbf{X}$, tak (zo spojitosti zdola pravdepodobnosti)

$$F_{X_1,...,X_{j-1},X_{j+1},...,X_n}(x_1, ..., x_{j-1}, x_{j+1}, ..., x_n) = P\left(\bigcap_{\substack{i=1 \\ i \neq j}}^{n} [X_i < x_i] \right) = P(\lim_{k\to\infty} A_k) = P(A) =$$

$$= \lim_{k\to\infty} P(A_k) = \lim_{x_j\to\infty} F_{X_1,...,X_n}(x_1, ..., x_j, ..., x_n).$$

Ak si teraz zvolíme $\{i_1, ..., i_k\} \subset \{1, 2, ..., n\}$ a vyberieme $X_{i_1}, ..., X_{i_k}$, tak $k-$rozmernému náhodnému vektoru $\mathbf{X}^* = (X_{i_1}, ..., X_{i_k})'$ hovoríme *marginálny náhodný vektor*. úplne analogicky ako pre jednu "vybratú" X_j z vektora $\mathbf{X}$, odvodíme distribučnú funkciu v prípade "vybratej" $k-$tice $X_{i_1}, ..., X_{i_k}$ z náhodného vektora $\mathbf{X}$. Distribučná funkcia marginálneho náhodného vektora $\mathbf{X}^* = (X_{i_1}, ..., X_{i_k})'$ teda je

$$F^*(\mathbf{x}^*) = F^*(x_{i_1}, ..., x_{i_k}) = \lim_{\substack{x_{y_1} \to \infty \\ \vdots \\ x_{y_{n-k}} \to \infty}} F(x_1, ..., x_n),$$

kde $\{y_1, ..., y_{n-k}\} = \{1, 2, ..., n\} - \{i_1, ..., i_k\}$. Dokázali sme vlastne tvrdenie

Veta 7.3. Všetky marginálne rozdelenia pravdepodobnosti náhodného vektora $\mathbf{X}$ sú jednoznačne určené rozdelením pravdepodobnosti náhodného vektora $\mathbf{X}$.

Veta 7.4. (a) Nech $(\mathbf{x}_m, p_m)_{m\in J}$, ($J$ je konečná alebo apočitateľná indexová množina) je pravdepodobnostná funkcia náhodného vektora $\mathbf{X} = (X_1, ..., X_n)'$, pričom $M = \{\mathbf{x}_m = (x_1^{(m)}, ..., x_n^{(m)})'\}_{m\in J}$. Potom marginálny náhodný vektor $\mathbf{X}^* = (X_{i_1}, ..., X_{i_k})'$ má marginálnu pravdepodobnostnú funkciu $(\mathbf{x}_s^*, p_s^*)_{s\in\mathcal{S}}$, kde

$$M^* = \{\mathbf{x}_s^*\}_{s\in\mathcal{S}} = \{(x_{i_1}^{(s)}, ..., x_{i_k}^{(s)})' : \ \exists\, m \in J, \text{ že } x_{i_1}^{(m)} = x_{i_1}^{(s)}, ..., x_{i_k}^{(m)} = x_{i_k}^{(s)}\},$$

pričom všetky $\mathbf{x}_s^*$ sú navzájom rôzne a

$$p_s^* = P(\{X_{i_1} = x_{i_1}^{(s)}, ..., X_{i_k} = x_{i_k}^{(s)}\}) = \sum_{\{\mathbf{x}_m\in M:\ x_{i_1}^{(m)}=x_{i_1}^{(s)},...,x_{i_k}^{(m)}=x_{i_k}^{(s)}\}} p_m.$$

(b) Nech $\mathbf{X}$ je spojitý náhodný vektor s hustotou $f(\mathbf{x})$. Potom marginálny náhodný vektor $\mathbf{X}^*$ má hustotu

$$f^*(x_{i_1},...,x_{i_k}) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(x_1,...,x_n)dx_{y_1}...dx_{y_{n-k}},$$

kde $\{y_1,...,y_{n-k}\} = \{1,2,...,n\} - \{i_1,...,i_k\}$.

Dôkaz: Dokážeme si len tvrdenie (b) (tvrdenie (a) si dokážte ako cvičenie). Platí

$$F^*(\mathbf{x}^*) = \lim_{\substack{x_{y_1}\to\infty \\ \vdots \\ x_{y_{n-k}}\to\infty}} F(x_1,...,x_n) = \lim_{\substack{x_{y_1}\to\infty \\ \vdots \\ x_{y_{n-k}}\to\infty}} \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_n} f(t_1,...,t_n)dt_1...dt_n =$$

$$= \int_{-\infty}^{x_{i_1}} ... \int_{-\infty}^{x_{i_k}} \left[\int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(t_1,...,t_n)dt_{y_1}...dt_{y_{n-k}}\right] dt_{i_1}...dt_{i_k},$$

teda

$$\int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(t_1,...,t_n)dt_{y_1}...dt_{y_{n-k}} = f^*(t_{i_1},...,t_{i_k}). \quad \clubsuit$$

8. Nezávislé náhodné veličiny

Definícia 8.1. Povieme, že náhodné veličiny $X_1, X_2, ..., X_n$ sú združene nezávislé, ak pre každú $n-$ticu reálnych čísel $x_1, x_2, ..., x_n$ platí

$$P(X_1 < x_1, X_2 < x_2, ..., X_n < x_n) = P(X_1 < x_1)P(X_2 < x_2)...P(X_n < x_n).$$

Veta 8.1. Nech náhodný vektor $\mathbf{X} = (X_1, ..., X_n)$ má distribučnú funkciu $F(\mathbf{x})$ a nech $F_{X_i}(x_i)$ je distribučná funkcia marginálnej náhodnej veličiny $X_i,\ i = 1,2,...,n$. Potom $X_1,...,X_n$ sú združene nezávislé práve vtedy, ak $F(\mathbf{x}) = \prod_{i=1}^n F_{X_i}(x_i)$ pre $\forall \mathbf{x} \in \mathtt{R}^n$.

Dôkaz: $X_1,...X_n$ sú združene nezávislé $\Longleftrightarrow$

$\forall \mathbf{x} = (x_1,...,x_n)' \in \mathtt{R}^n$ je $P\left(\bigcap_{i=1}^n \{X_i < x_i\}\right) = \prod_{i=1}^n P(\{X_i < x_i\}) \Longleftrightarrow$

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

$$F(\mathbf{x}) = F(x_1,...,x_n) = P(\{X_1 < x_1, X_2 < x_2, ..., X_n < x_n\}) =$$

$$= \prod_{i=1}^n P(\{X_i < x_i\}) = \prod_{i=1}^n F_{X_i}(x_i). \quad \clubsuit$$

Platia nasledujúce vety (ich dôkazy nájdete napr. v Rényi, A., Teorie pravděpodobnosti, Academia, Praha, 1972)

Veta 8.2. Nech náhodný vektor $\mathbf{X} = (X_1, ..., X_n)$ má diskrétnu distribučnú funkciu F a pravdepodobnostnú funkciu $(\mathbf{x}_m, p_m)_{m\in J}$. Potom $X_1,...,X_n$ sú združene nezávislé práve vtedy, ak pre $\forall \mathbf{x} = (x_1,...,x_n)'$

$$P(\{X_1 = x_1, ..., X_n = x_n\}) = \prod_{i=1}^n p_{X_i}(x_i),$$

kde $p_{X_i}(x_i) = P(\{X_i = x_i\})$.

Veta 8.3. Nech náhodný vektor $\mathbf{X} = (X_1, ..., X_n)$ má absolútne spojité rozdelenie pravdepodobnosti s hustotou $f(x_1, ..., x_n)$. Potom $X_1, ..., X_n$ sú združene nezávislé práve vtedy, ak pre $\forall \mathbf{x} = (x_1, ..., x_n)'$ až na množinu Lebesgueovej miery 0 platí

$$f(x_1, ..., x_n) = \prod_{i=1}^{n} f_{X_i}(x_i),$$

kde $f_{X_i}(x_i)$ je hustota marginálneho X_i.

Poznámka. Ak sú náhodné premenné $X_1, ..., X_n$ združene nezávislé, tak sú po dvoch nezávislé. Naopak neplatí, t.j. ak sú $X_1, ..., X_n$ po dvoch nezávislé ešte nemusia byť združene nezávislé.

Veta 8.4. Ak sú $X_1, ..., X_n$ združene nezávislé a $g_k(x),\ k = 1, 2, ..., n$ borelovsky merateľné funkcie reálnej premennej, tak sú náhodné veličiny $g_k(X_k),\ k = 1, 2, ..., n$ združene nezávislé.

Veta 8.5. Ak sú $X_1, ..., X_n$ združene nezávislé a $h(x_1, ..., x_k), k < n$ borelovsky merateľná funkcia, tak sú náhodné veličiny $h(X_1, ..., X_k), X_{k+1}, ..., X_n$ združene nezávislé.

Príklad 8.1. (Usporiadaný náhodný výber.) Nech $X_1, ..., X_n$ sú nezávislé náhodné veličiny na $(\Omega, \mathcal{A}, P)$, ktoré majú rovnaké rozdelenie pravdepodobnosti s distribučnou funkciou F (voláme ich *náhodný výber*). Najmenšiu z nich označme $X_{(1)}$, druhú najmenšiu $X_{(2)}$, až najväčšiu $X_{(n)}$. Teda $X_{(1)} \le X_{(2)} \le ... \le X_{(n)}$. Veličinám $X_{(1)}, X_{(2)}, ..., X_{(n)}$ hovoríme *usporiadaný výber*. Distribučná funkcia $G_r(x)$ náhodnej veličiny $X_{(r)},\ r \in \{1, 2, ..., n\}$ je

$$G_r(x) = \sum_{i=r}^{n} \binom{n}{i} [F(x)]^i [1 - F(x)]^{n-i}, \quad x \in \mathtt{R},$$

lebo pravdepodobnosť, že medzi náhodnými veličinami $X_1, ..., X_n$ bude práve i takých, že nadobudnú hodnotu menšiu než x je rovná

$$\binom{n}{i} [F(x)]^i [1 - F(x)]^{n-i}$$

a $X_{(r)}$ bude menšia než x práve vtedy, ak bude medzi $X_1, ..., X_n$ r, alebo $r + 1$, alebo, atď. n takých, ktoré majú hodnotu menšiu ako x. Spomenuté prípady sú disjunktné, teda pravdepodobnosť, že niektorý z nich nastane je súčtom pravdepodobností a dostávame $G_r(x)$.

Ak sú X_i absolútne spojité s hustotou $f(x)$, tak derivovaním $G_r(x)$ dostaneme hustotu $g_r(x)$ veličiny $X(r)$

$$g_r(x) = n\binom{n-1}{r-1} f(x)[F(x)]^{r-1}[1 - F(x)]^{n-r}$$

(pozri Anděl, J., Matematická statistika, SNTL/ALFA, Praha, 1985). Tam nájdeme aj distribučnú funkciu $G_{r,s}(x, y)$ náhodného vektora $(X_{(r)}, X_{(s)})',\ \ 1 \le r < s \le n$.

Príklad 8.2. Nech náhodný vektor $(X, Y)'$ má rovnomerné rozdelenie na $G \subset \mathtt{R}^2$, kde

(a) $G = \{(x, y)' \in \mathtt{R}^2 :\ \ 0 \le x < 1,\ 0 \le y \le 1\}$

(b) $G = \{(x, y)' \in \mathtt{R}^2 :\ \ 0 \le x < 1,\ 0 \le y \le 1,\ x + y \le 1\}$

V obidvoch prípadoch rozhodnite, či sú náhodné veličiny X a Y nezávislé.

Riešenie:

(a) hustota (X, Y) je

$$f_{X,Y}(x, y) = \begin{cases} c, & \text{ak } (x.y) \in G \\ 0, & \text{inak.} \end{cases}$$

Teda

$$1 = \int_G \int f_{X,Y}(x, y) dx dy = \int_0^1 \int_0^1 c dx dy = c \Rightarrow c = 1.$$

Marginálne hustoty sú

$$f_X(x) = \begin{cases} \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_0^1 dy = 1, & \text{ak } x \in< 0, 1) \\ 0, & \text{inak.} \end{cases}$$

Analogicky

$$f_Y(y) = \begin{cases} 1, & \text{ak } y \in< 0, 1 > \\ 0, & \text{inak.} \end{cases}$$

Pretože $\forall\ (x,y)' \in \mathtt{R}^2$ platí $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$, sú X a Y nezávislé.

(b) hustota (X,Y) je

$$f_{X,Y}(x,y) = \begin{cases} c, & \text{ak } (x.y) \in G \\ 0, & \text{inak.} \end{cases}$$

Teda

$$1 = \int_G \int f_{X,Y}(x,y)dxdy = c\int_0^1 \left[\int_0^{1-x} dy\right] dx = c\int_0^1 (1-x)dx =$$

$$= c\left[x - \frac{x^2}{2}\right]_0^1 = \frac{c}{2} \Rightarrow c = 2.$$

Marginálne hustoty sú

$$f_X(x) = \begin{cases} \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = 2\int_0^{1-x} dy = 2\,[y]_0^{1-x} = 2(1-x), & \text{ak } x \in< 0, 1) \\ 0, & \text{inak.} \end{cases}$$

Analogicky

$$f_Y(y) = \begin{cases} 2(1-y), & \text{ak } y \in< 0, 1 > \\ 0, & \text{inak.} \end{cases}$$

Platí, že $\forall\ (x,y)' \in< 0,1) \times < 0, 1 >$ okrem $\left\{(x,y) :\ x \in< 0, \frac{1}{2} >,\ y = \frac{1-2x}{2(1-x)}\right\}$ je $f_{X,Y}(x,y) \neq f_X(x) \cdot f_Y(y)$. Množina na ktorej $f_{X,Y}(x,y) \neq f_X(x) \cdot f_Y(y)$ je kladnej Lebesgueovej miery (nie je Lebesgueovej miery 0). Preto X a Y nie sú nezávislé.

9. Rozdelenie pravdepodobnosti transformovaných náhodných veličín

Veta 9.1. Nech X je náhodná veličina a h borelovsky merateľná funkcia. Potom $h(X)$ je náhodná veličina.

Dôkaz: Nech $B \in \mathcal{B}$ je ľubovoľná borelovská množina. Označme $h^{-1}(B) = \{t \in \mathtt{R} :\ h(t) \in B\}$. Pretože h je borelovsky merateľná, je $h^{-1}(B) \in \mathcal{B}$. Potom ale

$$\{\omega \in \Omega :\ h(X(\omega)) \in B\} = \{\omega \in \Omega :\ X(\omega) \in h^{-1}(B)\} \in \mathcal{A}. \quad \clubsuit$$

Veta 9.2. Nech zobrazenie $\mathbf{h} :\ \mathtt{R}^n \to \mathtt{R}^m$ je borelovsky merateľné, t.j. $\forall B \in \mathcal{B}^m$ je $\{\mathbf{x} = (x_1, ..., x_n)' \in \mathtt{R}^n :\ \mathbf{h}(x_1, ..., x_n)' \in B\} \in \mathcal{B}^n$. Nech $\mathbf{X} = (X_1, ..., X_n)'$ je $n-$rozmerný náhodný vektor na $(\Omega, \mathcal{A}, P)$. Potom $\mathbf{Y} = \mathbf{h}(\mathbf{X})$ je $m-$rozmerný náhodný vektor.

Dôkaz: Nech $B \in \mathcal{B}^m$. Potom z merateľnosti $\mathbf{h}$ vyplýva, že $\mathbf{h}^{-1}(B) = \{\mathbf{X} \in \mathtt{R}^n :\ \mathbf{h}(\mathbf{X}) \in B\} \in \mathcal{B}^n$. Preto

$$\{\omega \in \Omega :\ \mathbf{h}(\mathbf{X}(\omega)) \in B\} = \{\omega \in \Omega :\ \mathbf{X}(\omega) \in \mathbf{h}^{-1}(B)\} \in \mathcal{A}. \quad \clubsuit$$

V ďalšom sa budeme zaoberať rozdelením pravdepodobnosti transformovaných náhodných veličín, resp. transformovaných náhodných vektorov.

Poznámka. Pracovať budeme s Lebesgueovým integrálom z borelovsky merateľnej funkcie φ vzhľadom k Lebesgueovej-Sieljesovej miere μ_F na borelovskej množine A, t.j. budeme pracovať s integrálom

$$I = \int_A \varphi(t) d\mu_F(t) \stackrel{\text{ozn.}}{=} \int_A \varphi(t) dF(t).$$

Keď pracujeme s Lebesgueovym integrálom vzhľadom k Lebesgueovej miere, tak

$$I = \int_A \varphi(t) d\mu(t) \stackrel{\text{ozn.}}{=} \int_A \varphi(t) dt.$$

Poznámka. Pokiaľ je distribučná funkcia F funkciou "skokovitou", t.j. je to distribučná funkcia diskrétnej náhodnej veličiny s pravdepodobnostnou funkciou $(x_i, p_i)_{i\in J}$ (J je konečná alebo spočitateľná), je

$$I = \int_A \varphi(t) dF(t) = \sum_{x_i \in A} \varphi(x_i) p_i.$$

Ak je F distribučná funkcia spojitej náhodnej veličiny s hustotou $f(\cdot)$, tak

$$I = \int_A \varphi(t) dF(t) = \int_A \varphi(t) f(t) dt,$$

pričom posledný integrál je Lebesgueov integrál s Lebesgueovou mierou.

Veta 9.3. Nech náhodná veličina X má distribučnú funkciu F_X a h je borelovsky merateľná funkcia. Ak označíme F_Y distribučnú funkciu náhodnej veličiny $Y = h(X)$, potom $\forall y \in \mathtt{R}$ je $F_Y(y) = \int_{B_y} dF_X(x)$, kde $B_y = \{x \in \mathtt{R}: \ h(x) < y\}$.

Dôkaz: Pre ľubovoľné $y \in \mathtt{R}$ položme $B_y = \{x \in \mathtt{R}: \ h(x) < y\}$ a dostávame

$$F_Y(y) = P(Y < y) = P(h(X) < y) = P(X \in B_y) = P_X(B_y) =$$
$$= \int_{B_y} d\mu_{F_X}(x) = \int_{B_y} dF_X(x). \quad \clubsuit$$

Poznámka.

(a) Majme diskrétnu náhodnú veličinu X s pravdepodobnostnou funkciou $(x_i, p_i^{(X)})_{i\in J}$ a h nech je (borelovsky) merateľná. Označme ďalej $B_y^* = \{x \in \mathtt{R}: \ y = h(x)\}$. Pravdepodobnostná funkcia náhodnej veličiny $Y = h(X)$ je $(y_j, p_j^{(Y)})_{j\in K}$, kde $M_Y = \{y_j\}_{j\in K} = \{h(x_i): \ i \in J, \ h(x_i)$ sú navzájom rôzne $\}$ a

$$p_j^{(Y)} = P(Y = y_j) = P(h(X) = y_j) = P(X \in B_{y_j}^*) = P_X(B_{y_j}^*) = \sum_{\{x_i \in B_{y_j}^*\}} p_i^{(X)}.$$

(b) Ak X je (absolútne) spojitá náhodná veličina s hustotou f_X a distribučnou funkciou F_X, $Y = h(X)$, kde h je merateľná a $B_y = \{x \in \mathtt{R}: \ h(x) < y\}$, tak $\forall y \in \mathtt{R}$ je

$$F_Y(y) = P(Y < y) = P(h(X) < y) = P(X \in B_y) = \int_{B_y} f_X(x) dx.$$

Jednoducho sa dá určiť hustota f_Y v prípade, že transformácia $y = h(x)$ (teda funkcia h) je vzájomne jednoznačná (prostá a na) a teda existuje inverzná funkcia h^{-1} (teda $x = h^{-1}(y)$), pričom existuje aj derivácia $\frac{d}{dy} h^{-1}(y)$ a je spojitá. Potom z vety o substitúcii plynie

$$F_Y(y) = \int_{\{x:\ h(x) < y\}} f_X(x) dx = \int_{-\infty}^{y} f_X(h^{-1}(t)) \left| \frac{dh^{-1}(t)}{dt} \right| dt,$$

teda

$$f_Y(y) = f_X(h^{-1}(y)) \left| \frac{dh^{-1}(y)}{dy} \right|. \tag{9.1}$$

Príklad 9.1. Majme diskrétnu náhodnú X veličinu s pravdepodobnostnou funkciou $(x_i, p_i^{(X)})_{i\in J}$. Diskrétna náhodná veličina $Y = X^2$ má pravdepodobnostnú funkciu $(y_j, p_j^{(Y)})_{j\in K}$, kde $\{y_j\}_{j\in K} = \{x_i^2 : i \in J,\ x_i^2 \text{ sú rôzne}\}$ a

$$p_j^{(Y)} = P(Y = y_j) = P(X^2 = y_j) = P(X \in \{x : \ y_j = x^2\}) =$$

$$= P_X\{B^*_{y_j}\} = \sum_{\{x_i:\ y_j = x_i^2\}} p_i^{(X)}.$$

Ak $\mathtt{Z}$ je množina celých čísel, $x_z = z \quad (z \in \mathtt{Z})$, $p_0^{(X)} = e^{-\lambda}$, $p_z^{(X)} = e^{-\lambda} \frac{\lambda^{|z|}}{2|z|!}$ pre $z = \pm 1, \pm 2, ...$, tak $(x_z, p_z^{(X)})_{z\in\mathtt{Z}}$ je pravdepodobnostná funkcia diskrétnej náhodnej veličiny X a pravdepodobnostná funkcia diskrétnej náhodnej veličiny $Y = X^2$ je $(t^2, p_{t^2}^{(Y)})_{t\in \mathtt{N}_0}$, kde $p_{t^2}^{(Y)} = e^{-\lambda}\frac{\lambda^t}{t!}$ pre $t = 0, 1, 2, ...$

Príklad 9.2. Nech je náhodná veličina X absolútne spojitá s hustotou f_X. Položme $Y = a + bX$, kde $a, b \in \mathtt{R}$, $b \neq 0$. Nájdite hustotu f_Y náhodnej veličiny Y.

Riešenie: Transformácia $y = h(x) = a + bx$ je vzájomne jednoznačná, inverzná transformácia je $x = h^{-1}(y) = \frac{y-a}{b}$, ktorá má deriváciu $\frac{dx}{dy} = \frac{1}{b}$, teda podľa (9.1) je

$$f_Y(y) = f_X(h^{-1}(y)) \left| \frac{dh^{-1}(y)}{dy} \right| = \frac{1}{|b|} f_X\left(\frac{y-a}{b}\right).$$

Ak predpokladáme $X \sim N(\mu, \sigma^2)$, tak $f_X = \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $x \in (-\infty, \infty)$ a

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma|b|} e^{-\frac{\left(\frac{y-a}{b} - \mu\right)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}\sigma|b|} e^{-\frac{[y-(a+b\mu)]^2}{2(b\sigma)^2}}, \quad y \in (-\infty, \infty).$$

Predchádzajúce úvahy v prípade spojitej náhodnej veličiny rozšírime na mnohorozmerný prípad. Niekoľko základných pojmov:

Nech $\mathbf{h} = (h_1, ..., h_n)'$ je zobrazenie $\mathtt{R}^n \to \mathtt{R}^n$, teda $h_1, ..., h_n$ sú reálne funkcie n premenných $x_1, ..., x_n$. Jakobián (Jacobiho determinant) zobrazenia $\mathbf{h}$ je

$$D_{\mathbf{h}}(\mathbf{x}) = \det \frac{\partial \mathbf{h}}{\partial \mathbf{x}'} = \det \begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial h_n}{\partial x_1} & \cdots & \frac{\partial h_n}{\partial x_n} \end{pmatrix}.$$

Ak označíme $\mathbf{y} = \mathbf{h}(\mathbf{x})$, teda $y_i = h_i(\mathbf{x})$, $i = 1, 2, ..., n$, tak povieme, že zobrazenie $\mathbf{h}$ je *regulárne* na množine $M \subseteq \mathtt{R}^n$ ak

(i) M je otvorená,

(ii) funkcie $h_1, ..., h_n$ majú spojité prvé parciálne derivácie na M,

(iii) $\forall \mathbf{x} \in M$ platí, že $D_{\mathbf{h}}(\mathbf{x}) \neq 0$.

Zobrazenie $\mathbf{h}$ je *prosté* na M, ak platí

$\mathbf{x}_1 \in M, \mathbf{x}_2 \in M, \mathbf{x}_1 \neq \mathbf{x}_1 \Longrightarrow \mathbf{h}(\mathbf{x}_1) \neq \mathbf{h}(\mathbf{x}_2)$.

Veta 9.4. (Veta o substitúcii.) Nech $\mathbf{h}$ je zobrazenie otvorenej množiny $P \subseteq \mathtt{R}^n$ na $Q \subseteq \mathtt{R}^n$, nech $\mathbf{h}$ je regulárne a prosté zobrazenie na P s Jakobiánom $D_{\mathbf{h}}$. Nech $M \subset Q$ je borelovská a $H : \ \mathtt{R}^n \to \mathtt{R}$ reálna merateľná a integrovateľná funkcia. Potom platí

$$\int_M H(\mathbf{y})d\mathbf{y} = \int_{\mathbf{h}^{-1}(M)} H(\mathbf{h}(\mathbf{x}))|D_{\mathbf{h}}(\mathbf{x})|d\mathbf{x}.$$

Dôkaz: Jarník, V., Integrální počet I,II, NČSAV, Praha, 1955.

Bezprostredným dôsledkom tejto vety sú nasledujúce dve vety. Ich dôkazy nájdeme napr. v Anděl, J., Matematická statistika, SNTL/Alfa, Praha, 1985.

Veta 9.5. (Veta o hustote transformovaného náhodného vektora.) Nech náhodný vektor $\mathbf{X} = (X_1, ..., X_n)'$ má hustotu $f_{\mathbf{X}}(\mathbf{x})$, $\mathbf{x} \in \mathtt{R}^n$. Nech $\mathbf{h}$ je zobrazenie $\mathtt{R}^n$ do $\mathtt{R}^n$, ktoré je regulárne a prosté na otvorenej množine G, pre ktorú platí $\int_G f_{\mathbf{X}}(\mathbf{x})d\mathbf{x} = 1$. Ak $\mathbf{h}^{-1}$ je inverzné zobrazenie k $\mathbf{h}$, potom má náhodný vektor $\mathbf{Y} = \mathbf{h}(\mathbf{X})$ hustotu $f_{\mathbf{Y}}(\mathbf{y})$ tvaru

$$f_{\mathbf{Y}}(\mathbf{y}) = \begin{cases} f_{\mathbf{X}}(\mathbf{h}^{-1}(\mathbf{y}))|D_{\mathbf{h}^{-1}}(\mathbf{y})|, & \text{ak } \mathbf{y} \in \mathbf{h}(G) \\ 0, & \text{inak.} \end{cases}$$

Veta 9.5a. (Zovšeobecnená veta o hustote transformovaného náhodného vektora.) Nech náhodný vektor $\mathbf{X} = (X_1, ..., X_n)'$ má hustotu $f_{\mathbf{X}}(\mathbf{x})$, $\mathbf{x} \in \mathtt{R}^n$. Nech $\mathbf{h}$ je zobrazenie $\mathtt{R}^n$ do $\mathtt{R}^n$, ktoré je regulárne a prosté na disjunktných otvorených množinách $G_1, G_2, ...$ a zobrazuje ich na $\mathbf{h}(G_1), \mathbf{h}(G_2), ...,$ pričom platí platí $\int_G f_{\mathbf{X}}(\mathbf{x})d\mathbf{x} = 1$, kde $G = \bigcup_{j=1}^{\infty} G_j$. Ak označíme $\mathbf{h}_j^{-1}$ inverzné zobrazenie k $\mathbf{h}: \ G_j \to \mathbf{h}(G_j), \ j = 1, 2, ...,$, potom má náhodný vektor $\mathbf{Y} = \mathbf{h}(\mathbf{X})$ hustotu $f_{\mathbf{Y}}(\mathbf{y})$ tvaru $f_{\mathbf{Y}}(\mathbf{y}) = \sum_{j=1}^{\infty} f_j(\mathbf{y})$, kde

$$f_j(\mathbf{y}) = \begin{cases} f_{\mathbf{X}}(\mathbf{h}_j^{-1}(\mathbf{y}))|D_{\mathbf{h}_j^{-1}}(\mathbf{y})|, & \text{ak } \mathbf{y} \in \mathbf{h}(G_j) \\ 0, & \text{inak.} \end{cases}$$

Ukážeme si dva príklady.

Príklad 9.3. Nech $\mathbf{X} = (X_1, ..., X_n)'$ je náhodný vektor absolútne spojitého typu s hustotou $f_{\mathbf{X}}$. Nech $\mathbf{A}$ je regulárna matica typu $n \times n$. Nájdite hustotu náhodného vektora $\mathbf{Y} = \mathbf{AX}$.

Riešenie. $\mathbf{A}$ je regulárna a preto zobrazenie $\mathbf{y} = \mathbf{h}(\mathbf{x}) = \mathbf{Ax}$ je regulárne na otvorenej $G = \mathtt{R}^n$, Inverzné zobrazenie je $\mathbf{x} = \mathbf{h}^{-1}(\mathbf{y}) = \mathbf{A}^{-1}\mathbf{y}$. Ľahko sa vidí, že $D_{\mathbf{h}^{-1}}(\mathbf{y}) = \det \mathbf{A}^{-1} = \dfrac{1}{\det \mathbf{A}}$ a preto $f_{\mathbf{Y}}(\mathbf{y}) = \dfrac{1}{|\det \mathbf{A}|} f_{\mathbf{x}}(\mathbf{A}^{-1}\mathbf{y}), \quad \mathbf{y} \in \mathtt{R}^n$.

Príklad 9.4. Nech X je spojitá náhodná veličina s hustotou $f_X(x)$. Nájdite hustotu náhodnej veličiny $Y = X^2$.

Riešenie. Použijeme Vetu 9.5. Funkcia $h: \ \mathtt{R} \to \mathtt{R}$ daná predpisom $y = h(x) = x^2$ je regulárna a prostá na disjunktných otvorených množinách $G_1 = (-\infty, 0), G_2 = (0, \infty)$, pričom tieto množiny zobrazuje na $h(G_1) = (0, \infty)$ a $h(G_2) = (0, \infty)$ a $\int_{G=G_1 \cup G_2} f_X(x)dx = 1$. h_1^{-1} dané predpisom $h_1^{-1}(y) = -\sqrt{y}$ je inverzné zobrazenie k zobrazeniu $h: \ G_1 = (-\infty, 0) \to h(G_1)$ a $h_2^{-1}(y) = \sqrt{y}$ je inverzné zobrazenie k zobrazeniu $h: \ G_2 = (0, \infty) \to h(G_2)$. Pre $y \in h(G_1)$ je $D_{h_1^{-1}}(y) = -\dfrac{1}{2\sqrt{y}}$ a pre $y \in h(G_2)$ je $D_{h_2^{-1}}(y) = \dfrac{1}{2\sqrt{y}}$. Náhodná premenná $Y = h(X)$ má preto hustotu

$$f_Y(y) = f_1(y) + f_2(y) = \begin{cases} \frac{1}{2\sqrt{y}} \left(f_X(-\sqrt{y}) + f_X(\sqrt{y}) \right), & \text{ak } y \in (0, \infty) \\ 0, & \text{inak.} \end{cases}$$

Poznámka. Často potrebujeme spočítať hustotu náhodnej veličiny $Y = h(\mathbf{X})$, kde $\mathbf{X} = (X_1, ..., X_n)'$ a h je reálna borelovsky merateľná funkcia n premenných. Ak $F_{\mathbf{X}}$ je distribučná funkcia náhodného vektora

$\mathbf{X}$, tak distribučná funkcia náhodnej veličiny Y je

$$F_Y(y) = P(Y < y) = P(h(\mathbf{X}) < y) = \int_{B_y} dF_{\mathbf{X}}(x_1, ..., x_n),$$

kde $B_y = \{(x_1, ..., x_n)' \in \mathtt{R}^n : \ h(x_1, ..., x_n) < y\}$.

- ak $\mathbf{X} = (X_1, ..., X_n)'$ je diskrétny náhodný vektor s pravdepodobnostnou funkciou $(\mathbf{x}_m, p_m^{(\mathbf{X})})_{m \in J}$, tak $Y = h(\mathbf{X})$ má pravdepodobnostnú funkciu
$(y_j, p_j^{(Y)})_{j \in K}$, kde $\{y_j : \ j \in K\} = \{h(\mathbf{x}_m) : \ m \in J, \ h(\mathbf{x}_m)$ rôzne $\}$ a

$$p_j^{(Y)} = P(Y = y_j) = P(h(\mathbf{X}) = y_j) = P(\{\mathbf{X} \in B^*_{y_j}\}) =$$

$$= P_{\mathbf{X}}(B^*_{y_j}) = \sum_{\{\mathbf{x}_i : \ h(\mathbf{x}_i) = y_j\}} p_i^{(\mathbf{X})},$$

kde $B^*_{y_j} = \{\mathbf{x}_i \ : \ h(\mathbf{x}_i) = y_j\}$.

- ak $\mathbf{X} = (X_1, ..., X_n)'$ je náhodný vektor absolútne spojitého typu, sú dve možnosti.

(i) $F_Y(y) = \int_{\{(x_1,...,x_n)' \in \mathtt{R}^n : \ h(x_1,...,x_n) < y\}} dF_{\mathbf{X}}(x_1, ..., x_n) =$
$= \int_{\{(x_1,...,x_n)' \in \mathtt{R}^n : \ h(x_1,...,x_n) < y\}} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$.

V tomto prípade je hustota $f_Y(y) = \dfrac{dF_Y(y)}{dy}$.

(ii) Rozšírime $h(\mathbf{x})$ na $\mathbf{h}(\mathbf{x}) \ \ \mathtt{R}^n \to \mathtt{R}^n$, aby spĺňala predpoklady Vety 9.4, nasledovným spôsobom:

$$\begin{aligned} y &= h(\mathbf{x}) \\ y_2 &= x_2 \\ &\vdots \\ y_n &= x_n \end{aligned}$$

Takto dostávame náhodný vektor $\mathbf{Y} = (Y, Y_2, Y_3, ..., Y_n)' = \mathbf{h}(\mathbf{X})$. Spočítame jeho hustotu $f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(\mathbf{h}^{-1}(\mathbf{y}))|D_{\mathbf{h}^{-1}}(\mathbf{y})|$ a nakoniec marginálnu hustotu

$$f_Y(y) = \int_{\mathtt{R}^{n-1}} f_{\mathbf{Y}}(y, y_2, ..., y_n) dy_2 ... dy_n$$

s.v. vzhľadom na Lebesgueovu mieru.

Príklad 9.5. Nech $X_1 \sim Po(\lambda_1), X_2 \sim Po(\lambda_2)$ a X_1, X_2 sú stochasticky nezávislé (niekedy sa značí $X_1 \perp X_2$). Aká je pravdepodobnostná funkcia náhodnej veličiny $Y = X_1 + X_2$?

Riešenie. $X_i, \ \ i = 1, 2$ má pravdepodobnostnú funkciu $(j, p_j^{(X_i)})_{j \in \mathtt{N}_0}$, kde $p_j^{(X_i)} = e^{-\lambda_i} \frac{\lambda_i^j}{j!}, \ \ j = 0, 1, 2, ...$
Náhodný vektor $\mathbf{X} = (X_1, X_2)'$ má pravdepodobnostnú funkciu $\left((i, j), p_{(i,j)}^{(\mathbf{X})}\right)_{(i,j) \in \mathtt{N}_0 \times \mathtt{N}_0}$, kde

$$p_{(i,j)}^{(\mathbf{X})} = P\{X_1 = i, X_2 = j\} = P\{X_1 = i\} P\{X_2 = j\} = p_i^{(X_1)} p_j^{(X_2)} =$$

$$= e^{-(\lambda_1 + \lambda_2)} \frac{\lambda_1^i \lambda_2^j}{i! j!}.$$

Náhodná veličina $Y = X_1 + X_2$ má pravdepodobnostnú funkciu
$(y_s, p_s^{(Y)})$ a $M_Y = \{y_k\}_{k \in K} = \{h(i, j) = i + j : \ (i, j) \in \mathtt{N}_0 \times \mathtt{N}_0\}$ (rôzne) teda $M_Y = \{k : \ k \in \mathtt{N}_0\}$ (rôzne) a

$$p_k^{(Y)} = \sum_{\{(i,j) \in \mathtt{N}_0 \times \mathtt{N}_0 : \ h(i,j) = i + j = k\}} p_{(i,j)}^{(\mathbf{X})} = \sum_{\{(i,j) \in \mathtt{N}_0 \times \mathtt{N}_0 : \ i + j = k\}} p_i^{(X_1)} p_j^{(X_2)} =$$

$$= \sum_{i=0}^{k} p_i^{(X_1)} p_{k-i}^{(X_2)} = \sum_{i=0}^{k} e^{-(\lambda_1+\lambda_2)} \frac{\lambda_1^i \lambda_2^{k-i}}{i!(k-i)!} =$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{k!} \sum_{i=0}^{k} \binom{k}{i} \lambda_1^i \lambda_2^{k-i} = e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1+\lambda_2)^k}{k!},$$

čiže $Y \sim Po(\lambda_1 + \lambda_2)$.

Pre súčet dvoch náhodných veličín dostávame pomocou predchádzajúcej Poznámky nasledujúcu vetu.

Veta 9.6. Nech náhodný vektor $\mathbf{X} = (X_1, X_2)'$ je spojitého typu s hustotou $f_{(X_1,X_2)}(x_1, x_2)$. Potom je náhodná veličina $Y = X_1 + X_2$ absolútne spojitého typu a jej hustota je

$$f_Y(y) = \int_{-\infty}^{\infty} f_{(X_1,X_2)}(y-x, x)dx = \int_{-\infty}^{\infty} f_{(X_1,X_2)}(x, y-x)dx \quad \text{s.v.}$$

Dôkaz: Ak zvolíme transformáciu $\mathbf{h}(\mathbf{x})$:

$$y = x_1 + x_2$$
$$y_2 = x_2,$$

tak $\mathbf{h}^{-1}(\mathbf{y})$ je

$$x_1 = y - y_2$$
$$x_2 = y_2,$$

a $D_{\mathbf{h}^{-1}}(y, y_2) = \det \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = 1$, čiže $f_{(Y,Y_2)}(y, y_2) = f_{(X_1,X_2)}(y - y_2, y_2)$ a

$$f_Y(y) = \int_{-\infty}^{\infty} f_{(X_1,X_2)}(y-x, x)dx.$$

Ak zvolíme transformáciu

$$y_1 = x_1$$
$$y = x_1 + x_2,$$

tak úplne analogicky dostaneme

$$f_Y(y) = \int_{-\infty}^{\infty} f_{(X_1,X_2)}(x, y-x)dx \quad \text{s.v.} \clubsuit$$

Dôsledok. Ak sú vo Vete 9.6 náhodné veličiny X_1 a X_2 nezávislé, tak $f_{(X_1,X_2)}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$ a náhodná veličina $Y = X_1 + X_2$ má hustotu

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1}(y-x) f_{X_2}(x)dx = \int_{-\infty}^{\infty} f_{X_1}(x) f_{X_2}(y-x)dx \quad \text{s.v.}$$

Hustotu $f_Y(y)$ nazývame v tomto prípade *konvolúciou* hustôt f_{X_1} a f_{X_2} a označujeme $f_Y = f_{X_1} * f_{X_2}$.

Nasledujúcu vetu dokážeme úplne analogicky ako Vetu 9.6. a jej dôsledok.

Veta 9.7. Nech X_1, X_2 sú nezávislé náhodné veličiny s hustotami f_1 a f_2. Potom

(i) $Y = X_1 X_2$ má hustotu

$$g(y) = \int_{-\infty}^{\infty} f_1\left(\frac{y}{x}\right) f_2(x) \frac{1}{|x|} dx, \quad \text{s.v.};$$

(ii) ak je $f_2(x) = 0$ pre $x \le 0$ a $c > 0$ daná konštanta, tak náhodná veličina $Z = \frac{cX_1}{X_2}$ má hustotu

$$h(z) = \frac{1}{c} \int_0^{\infty} f_1\left(\frac{zx}{c}\right) f_2(x) x dx, \quad \text{s.v.} \; .$$

Predchádzajúce vety využijeme na odvodenie najdôležitejších rozdelení (okrem už spomenutého normálneho rozdelenia), ktoré budeme používať v štatistike.

Veta 9.8. Nech $X_1, ..., X_n$ sú nezávislé $N(0,1)$ rozdelené náhodné veličiny. Náhodná veličina

$$Y = X_1^2 + X_2^2 + ... + X_n^2$$

má χ^2 rozdelenie s n stupňami voľnosti (označujeme $Y \sim \chi_n^2$) s hustotou

$$f(y) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} y^{\frac{n}{2}-1} e^{-\frac{y}{2}} \quad \text{pre } y > 0$$

a $f(y) = 0$ pre $y \leq 0$.

Dôkaz: Vetu dokážeme indukciou. Pre $n = 1$ je pre $x \geq 0$

$$F_{X_1^2}(x) = P\{X_1^2 < x\} = P\{-\sqrt{x} \leq X_1 < \sqrt{x}\} - P\{-\sqrt{x} = X_1\} =$$

$$= \int_{-\sqrt{x}}^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt,$$

preto

$$f_{X_1^2}(x) = \frac{d}{dx} F_{X_1^2}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(\sqrt{x})^2}{2}} (\sqrt{x})' - \frac{1}{\sqrt{2\pi}} e^{-\frac{(-\sqrt{x})^2}{2}} (-\sqrt{x})' =$$

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

(lebo $\Gamma(\frac{1}{2}) = \sqrt{\pi}$). Teda veta platí pre $n = 1$. Nech platí pre n, potom pre $n+1$ je

$$f_{X_1^2+...+X_{n+1}^2}(x) = \int_0^\infty f_{X_1^2+...+X_n^2}(x-u) f_{X_{n+1}^2}(u) du =$$

$$= \int_0^x \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} (x-u)^{\frac{n}{2}-1} e^{-\frac{x-u}{2}} \frac{1}{2^{\frac{1}{2}}\Gamma(\frac{1}{2})} u^{-\frac{1}{2}} e^{-\frac{u}{2}} du =$$

$$= \frac{e^{-\frac{x}{2}}}{2^{\frac{n+1}{2}}\Gamma(\frac{n}{2})\Gamma(\frac{1}{2})} \int_0^x (x-u)^{\frac{n}{2}-1} u^{-\frac{1}{2}} du =$$

(substitúcia $\frac{u}{x} = w,\ du = xdw$, pričom $B(\alpha,\beta) = \int_0^1 (1-x)^{\alpha-1} x^{\beta-1} dx$)

$$= \frac{e^{-\frac{x}{2}} x^{\frac{n}{2}-1} x^{-\frac{1}{2}} x}{2^{\frac{n+1}{2}}\Gamma(\frac{n}{2})\Gamma(\frac{1}{2})} \int_0^1 (1-w)^{\frac{n}{2}-1} w^{\frac{1}{2}-1} dw = \frac{e^{-\frac{x}{2}} x^{\frac{n+1}{2}-1}}{2^{\frac{n+1}{2}}\Gamma(\frac{n}{2})\Gamma(\frac{1}{2})} B(\tfrac{n}{2}, \tfrac{1}{2}) =$$

$$= \frac{1}{2^{\frac{n+1}{2}}\Gamma(\frac{n+1}{2})} x^{\frac{n+1}{2}-1} e^{-\frac{x}{2}}. \quad \clubsuit$$

Veta 9.9. Nech X a Y sú nezávislé náhodné veličiny, $X \sim \chi_k^2,\ Y \sim \chi_m^2$. Náhodná veličina $U = \frac{\frac{X}{k}}{\frac{Y}{m}}$ má Fisherovo-Snedecorovo F rozdelenie s k a m stupňami voľnosti (značíme $U \sim F_{k,m}$) a hustotu

$$f_U(u) = \frac{\Gamma(\frac{k+m}{2})}{\Gamma(\frac{k}{2})\Gamma(\frac{m}{2})} \left(\frac{k}{m}\right)^{\frac{k}{2}} u^{\frac{k}{2}-1} \left(1 + \frac{ku}{m}\right)^{-\frac{k+m}{2}} \qquad \text{pre } u > 0$$

a $f_U(u) = 0$ pre $u \leq 0$.

Dôkaz: Platí $U = \dfrac{\frac{m}{k}X}{Y}$. Využijeme Vetu 9.7(ii) a Vetu 9.8. Dostávame pre $u > 0$ (pre $u \leq 0$ je hustota χ_k^2 rovná 0)

$$f_U(u) = \frac{k}{m}\int_0^\infty y \frac{1}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})}\left(\frac{uyk}{m}\right)^{\frac{k}{2}-1} e^{-\frac{uyk}{2m}} \frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} y^{\frac{m}{2}-1}e^{-\frac{y}{2}}dy =$$

$$= \left(\frac{k}{m}\right)^{\frac{k}{2}} \frac{u^{\frac{k}{2}-1}}{2^{\frac{k+m}{2}}\Gamma(\frac{k}{2})\Gamma(\frac{m}{2})}\int_0^\infty y^{\frac{k+m}{2}-1}e^{-\frac{y}{2}(\frac{ku}{m}+1)}dy =$$

(substitucia $\frac{y}{2}(\frac{ku}{m}+1) = t$)

$$= \left(\frac{k}{m}\right)^{\frac{k}{2}} \frac{u^{\frac{k}{2}-1}\left(\frac{ku}{m}+1\right)^{-1}}{2^{\frac{k+m}{2}}\Gamma(\frac{k}{2})\Gamma(\frac{m}{2})}\int_0^\infty \frac{2^{\frac{k+m}{2}}t^{\frac{k+m}{2}-1}}{\left(\frac{ku}{m}+1\right)^{\frac{k+m}{2}-1}}e^{-t}dt =$$

$$= \frac{\Gamma(\frac{k+m}{2})}{\Gamma(\frac{n}{2})\Gamma(\frac{1}{2})}\left(\frac{k}{m}\right)^{\frac{k}{2}} u^{\frac{k}{2}-1}\left(1+\frac{ku}{m}\right)^{-\frac{k+m}{2}}. \quad \clubsuit$$

Veta 9.10. Nech $X \sim \chi_n^2$. Náhodná veličina $Y = \sqrt{X}$ má χ rozdelenie a n stupňami voľnosti (značíme $Y \sim \chi_n$) a hustotu

$$f_Y(y) = \frac{1}{2^{\frac{n}{2}-1}\Gamma(\frac{n}{2})}y^{n-1}e^{-\frac{y^2}{2}} \qquad \text{pre } y > 0$$

a $f_Y(y) = 0$ pre $y \leq 0$.

Dôkaz: Náhodná veličina Y nadobúda (rovnako ako X) len kladné hodnoty. Pre $y > 0$ je

$$F_Y(y) = P\{\sqrt{X} < y\} = P\{X < y^2\} = \int_0^{y^2} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}x^{\frac{n}{2}-1}e^{-\frac{x}{2}}dx =$$

(použijeme substitúciu $x = t^2,\ dx = 2tdt$)

$$= \int_0^y \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}t^{n-2}e^{-\frac{t^2}{2}}2tdt = \int_0^y \frac{1}{2^{\frac{n}{2}-1}\Gamma(\frac{n}{2})}t^{n-1}e^{-\frac{t^2}{2}}dt. \quad \clubsuit$$

Veta 9.11. Nech náhodné veličiny $Z \sim N(0,1)$ a $X \sim \chi_n^2$ sú nezávislé. Náhodná veličina $T = \dfrac{Z}{\sqrt{\frac{X}{n}}}$ má Studentovo t rozdelenie s n stupňami voľnosti (značíme t_n) a hustotu

$$f_T(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \qquad t \in (-\infty, \infty).$$

Dôkaz: $T = \frac{\sqrt{n}Z}{\sqrt{X}}$ a podľa Vety 9.7(ii) a Vety 9.10 dostávame (pre $t \in (-\infty,\infty)$)

$$f_T(t) = \frac{1}{\sqrt{n}}\int_0^\infty y\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2y^2}{2n}}\frac{1}{2^{\frac{n}{2}-1}\Gamma(\frac{n}{2})}y^{n-1}e^{-\frac{y^2}{2}}dy =$$

$$= \frac{1}{2^{\frac{n-1}{2}}\Gamma(\frac{n}{2})\sqrt{n\pi}}\int_0^\infty y^n e^{-\frac{y^2}{2}(\frac{t^2}{n}+1)}dy =$$

(substitúcia $\frac{y^2}{2}\left(\frac{t^2}{n}+1\right) = x,\ y = x^{\frac{1}{2}}\left(\frac{t^2}{n}+1\right)^{-\frac{1}{2}}2^{\frac{1}{2}},\ y\left(\frac{t^2}{n}+1\right)dy = dx$)

$$= \frac{1}{\Gamma(\frac{n}{2})\sqrt{n\pi}\left(\frac{t^2}{n}+1\right)^{\frac{n+1}{2}}}\int_0^\infty x^{\frac{n+1}{2}-1}e^{-x}dx = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}. \quad \clubsuit$$

10. Charakteristiky rozdelenia pravdepodobnosti

Stredná hodnota a rozptyl

Definícia 10.1. Nech X je náhodná veličina na $(\Omega, \mathcal{A}, P)$ a nech existuje

$$\int_{\Omega} X(\omega) dP(\omega) < \infty.$$

Potom číslo

$$\mathcal{E}(X) = \int_{\Omega} X(\omega) dP(\omega)$$

nazývame *strednou hodnotou* náhodnej veličiny X. Ak uvedený integrál nie je konečný alebo neexistuje, hovoríme, že stredná hodnota náhodnej veličiny X neexistuje.

Poznámka. Z definície strednej hodnoty náhodnej veličiny X vyplýva, že $\mathcal{E}(X)$ existuje práve vtedy ak je X borelovsky merateľná funkcia a integrovateľná na Ω vzhľadom k pravdepodobnostnej miere P. $\mathcal{L}_1(\Omega, \mathcal{A}, P) = \mathcal{L}_1$ označujeme množinu (priestor) všetkých náhodných veličín, ktoré majú konečnú strednú hodnotu na $(\Omega, \mathcal{A})$.

Základné vlastnosti strednej hodnoty vyplývajú zo základných vlastností integrovateľných funkcií (z teórie integrálu).

Veta 10.1. (Základné vlastnosti strednej hodnoty.) Nech X, X_1, X_2 sú náhodné veličiny definované na $(\Omega, \mathcal{A}, P)$, $\;a, a_1, a_2 \in \mathtt{R}$. Potom

(i) $\mathcal{E}(X)$ existuje (t.j. $X \in \mathcal{L}_1$) $\iff$ $\mathcal{E}|X|$ existuje;

(ii) ak $P(X = a) = 1 \implies \mathcal{E}(X) = a$;

(iii) ak existujú $\mathcal{E}(X_1),\ \mathcal{E}(X_2) \implies \mathcal{E}(a_1 X_1 + a_2 X_2) = a_1 \mathcal{E}(X_1) + a_2 \mathcal{E}(X_2)$;

(iv) ak existujú $\mathcal{E}(X_1),\ \mathcal{E}(X_2)$ a $X_1 \leq X_2 \implies \mathcal{E}(X_1) \leq \mathcal{E}(X_2)$;

(v) ak $|X_1| \leq X_2$ a existuje $\mathcal{E}(X_2)$, tak existuje $\mathcal{E}(X_1)$;

(vi) nech $P(X \geq 0) = 1$ a existuje $\mathcal{E}(X) \implies \mathcal{E}(X) \geq 0$.

Dôkaz vyplýva z vlastností Lebesgueovho integrálu.

Ďalšie vlastnosti strednej hodnoty, hlavne vzorce vhodné na jej výpočet vyplývajú z *vety o prenose integrácie z merateľného priestoru* $(\Omega, \mathcal{A})$ *na merateľný priestor* $(\Lambda, \mathcal{D})$ *pomocou merateľnej funkcie* g. Táto veta v prípade, že $(\Lambda, \mathcal{D}) = (\mathtt{R}^n, \mathcal{B}^n)$ a g je $n-$rozmerný náhodný vektor znie:

Veta 10.2. (O prenose integrácie.) Nech $\mathbf{X} = (X_1, ..., X_n)'$ je náhodný vektor definovaný na pravdepodobnostnom priestore $(\Omega, \mathcal{A}, P)$, g je borelovsky merateľná funkcia na $(\mathtt{R}^n, \mathcal{B}^n)$, $P_{\mathbf{X}}$ je rozdelenie pravdepodobnosti náhodného vektora $\mathbf{X}$. Potom

$$\int_{\Omega} g(\mathbf{X}(\omega)) dP(\omega) = \int_{\mathtt{R}^n} g(\mathbf{x}) dP_{\mathbf{X}}(\mathbf{x})$$

v zmysle, že ak jeden z integrálov existuje, tak existuje aj druhý a rovnajú sa.

Poznámka. Ak má náhodný vektor $\mathbf{X} = (X_1, ..., X_n)'$ distribučnú funkciu $F(\cdot)$, potom rozdelenie pravdepodobnosti $P_{\mathbf{X}} = \mu_F$, kde μ_F je Lebesgueova-Stieltjesova miera indukovaná distribučnou funkciou F a môžeme písať

$$\int_{\Omega} g(\mathbf{X}(\omega)) dP(\omega) = \int_{\mathtt{R}^n} g(\mathbf{x}) d\mu_F(\mathbf{x}) \overset{\text{píšeme}}{=} \int_{\mathtt{R}^n} g(\mathbf{x}) dF(\mathbf{x}).$$

Priamym dôsledkom vety o prenose integrácie je nasledujúci dôsledok, pomocou ktorého spočítame strednú hodnotu náhodnej veličiny $Y = g(X)$, keď g je borelovská funkcia a X náhodná veličina.

Dôsledok. Nech X je náhodná veličina a g borelovská funkcia. Potom stredná hodnota náhodnej veličiny $Y = g(X)$ existuje práve vtedy, ak existuje a je konečný integrál $\int_{-\infty}^{\infty} |g(x)|dF(x) < \infty$. V tomto prípade platí

$$\mathcal{E}(Y) = \int_{-\infty}^{\infty} g(x)dF(x)$$

(teda $Y = g(X) \in \mathcal{L}_1(\Omega, \mathcal{A}, P) \;\Leftrightarrow\; \int_{-\infty}^{\infty} |g(x)|dF(x)dx < \infty$). Špeciálne

(a) ak je X diskrétna s pravdepodobnostnou funkciou $(x_i, p_i)_{i\in J}$, potom $\mathcal{E}(Y)$ existuje práve vtedy ak $\sum_{i\in J} |g(x_i)|p_i < \infty$ a platí

$$\mathcal{E}(Y) = \sum_{i\in J} g(x_i)p_i$$

(teda $Y = g(X) \in \mathcal{L}_1(\Omega, \mathcal{A}, P) \;\Leftrightarrow\; \sum_{i\in J} |g(x_i)|p_i < \infty$).

(b) ak je X spojitá s hustotou f, potom $\mathcal{E}(Y)$ existuje práve vtedy ak existuje $\int_{-\infty}^{\infty} |g(x)|f(x)dx < \infty$ a platí

$$\mathcal{E}(Y) = \int_{-\infty}^{\infty} g(x)f(x)dx$$

(teda $Y = g(X) \in \mathcal{L}_1(\Omega, \mathcal{A}, P) \;\Leftrightarrow\; \int_{-\infty}^{\infty} |g(x)|f(x)dx < \infty$).

V prípade, že v predchádzajúcom Dôsledku uvažujeme funkciu $g(x) = x$, vieme spočítať strednú hodnotu náhodnej veličiny X nasledovne:

Dôsledok. Nech X je náhodná veličina na $(\Omega, \mathcal{A}, P)$. Potom stredná hodnota náhodnej veličiny X existuje práve vtedy, ak existuje a je konečný integrál $\int_{-\infty}^{\infty} |x|dF(x) < \infty$. V tomto prípade platí

$$\mathcal{E}(Y) = \int_{-\infty}^{\infty} xdF(x)$$

(teda $X \in \mathcal{L}_1(\Omega, \mathcal{A}, P) \;\Leftrightarrow\; \int_{-\infty}^{\infty} |x|dF(x)dx < \infty$).

Špeciálne

(a) ak je X diskrétna s pravdepodobnostnou funkciou $(x_i, p_i)_{i\in J}$, potom $\mathcal{E}(X)$ existuje práve vtedy ak $\sum_{i\in J} |x_i|p_i < \infty$ a platí

$$\mathcal{E}(X) = \sum_{i\in J} x_i p_i$$

(teda $X \in \mathcal{L}_1(\Omega, \mathcal{A}, P) \;\Leftrightarrow\; \sum_{i\in J} |x_i|p_i < \infty$).

(b) ak je X spojitá s hustotou f, potom $\mathcal{E}(X)$ existuje práve vtedy ak existuje $\int_{-\infty}^{\infty} |x|f(x)dx < \infty$ a platí

$$\mathcal{E}(Y) = \int_{-\infty}^{\infty} xf(x)dx$$

(teda $X \in \mathcal{L}_1(\Omega, \mathcal{A}, P) \;\Leftrightarrow\; \int_{-\infty}^{\infty} |x|f(x)dx < \infty$).

V prípade, že máme náhodný vektor, tak použijeme nasledujúci dôsledok.

Dôsledok. Nech $\mathbf{X} = (X_1, ..., X_n)'$ je náhodný vektor definovaný na $(\Omega, \mathcal{A}, P)$ a $g(x_1, ..., x_n)$ borelovská funkcia. Potom stredná hodnota náhodnej veličiny $Y = g(\mathbf{X})$ existuje práve vtedy, ak existuje a je konečný integrál $\int_{\mathrm{R^n}} |g(\mathbf{x})|dF(\mathbf{x}) < \infty$. V tomto prípade platí

$$\mathcal{E}(Y) = \int_{\mathrm{R^n}} g(\mathbf{x})dF(\mathbf{x}).$$

Špeciálne

(a) ak je $\mathbf{X}$ je diskrétneho typu s pravdepodobnostnou funkciou $(\mathbf{x}_i, p_i)_{i\in J}$, potom $\mathcal{E}(Y)$ existuje práve vtedy ak $\sum_{i\in J} |g(\mathbf{x}_i)|p_i < \infty$ a platí

$$\mathcal{E}(Y) = \sum_{i\in J} g(\mathbf{x}_i)p_i.$$

(b) ak je $\mathbf{X}$ spojitá s hustotou $f(x_1, ..., x_n)$, potom $\mathcal{E}(Y)$ existuje práve vtedy ak existuje $\int_{\mathbb{R}^n} |g(\mathbf{x})|f(\mathbf{x})d\mathbf{x} < \infty$ a platí

$$\mathcal{E}(Y) = \int_{\mathbb{R}^n} g(\mathbf{x})f(\mathbf{x})d\mathbf{x}.$$

Príklad 10.1. Stredná hodnota náhodnej veličiny s poissonovským rozdelením (stredná hodnota Poissonovho rozdelenia). Nech $X \sim Po(\lambda)$, teda X má pravdepodobnostnú funkciu $(x_i, p_i)_{i\geq 1}$, kde $x_i = 0, 1, 2, ...$ a $p_i = e^{-\lambda}\dfrac{\lambda^{x_i}}{x_i!}$.

Preto

$$\mathcal{E}(X) = \sum_{j=0}^{\infty} je^{-\lambda}\frac{\lambda^j}{j!} = e^{-\lambda}\sum_{j=1}^{\infty}\frac{\lambda^j}{(j-1)!} = \lambda e^{-\lambda}\sum_{j=1}^{\infty}\frac{\lambda^{j-1}}{(j-1)!} = \lambda e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = \lambda.$$

Príklad 10.2. Stredná hodnota náhodnej veličiny s normálnym rozdelením. Nech $X \sim N(\mu, \sigma^2)$, $\sigma > 0$, teda jej hustota je $f(x) = \dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in (-\infty, \infty)$. Potom

$$\mathcal{E}(X) = \int_{-\infty}^{\infty} xf(x)dx = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} xe^{-\frac{(x-\mu)^2}{2\sigma^2}}dx =$$

(substitúcia $y = \frac{x-\mu}{\sigma}$, $x = \sigma y + \mu$, $dy = \frac{dx}{\sigma}$)

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\sigma y + \mu)e^{-\frac{y^2}{2}}dy = \frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty} ye^{-\frac{y^2}{2}}dy + \mu\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy = \mu,$$

lebo $ye^{-\frac{y^2}{2}}$ je nepárna (lichá) funkcia.

Veta 10.3. (Stredná hodnota súčinu nezávislých náhodných veličín.) Nech $X_1, ..., X_n$ sú nezávislé náhodné veličiny na $(\Omega, \mathcal{A}, P)$ a nech existujú stredné hodnoty $\mathcal{E}(X_i), i = 1, 2, ..., n$, (t.j. $X_i \in \mathcal{L}_1(\Omega, \mathcal{A}, P)$). Potom platí

$$\mathcal{E}(\prod_{i=1}^{n} X_i) = \prod_{i=1}^{n}\mathcal{E}(X_i).$$

Dôkaz: Položme $Y = \prod_{i=1}^{n} X_i$, teda $g(x_1, ..., x_n) = x_1x_2...x_n$. Podľa posledného Dôsledku je

$$\mathcal{E}(Y) = \int_{\mathbb{R}^n} g(x_1, ..., x_n)dF_{\mathbf{X}}(x_1, ..., x_n) = \int_{\mathbb{R}^n} x_1...x_n d[F_{X_1}(x_1)...F_{X_n}(x_n)] =$$

$$= \int_{\mathbb{R}} x_1 dF_{X_1}(x_1)...\int_{\mathbb{R}} x_n dF_{X_n}(x_n) = \prod_{i=1}^{n}\mathcal{E}(X_i). \quad \clubsuit$$

Počiatočne, centrálne a absolútne momenty

Nech X je náhodná veličina na pravdepodobnostnom priestore $(\Omega, \mathcal{A}, P)$. Potom (číslo)

$\mu'_n = \mathcal{E}(X^n)$ nazývame $n-$tým počiatočným (obecným) momentom náhodnej veličiny X,

$\mu_n = \mathcal{E}\left((X - \mathcal{E}(X))^n\right)$ nazývame $n-$tým centrálnym momentom náhodnej veličiny X,

$\overline{\mu}_n = \mathcal{E}(|X|^n)$ nazývame $n-$tým absolútnym momentom náhodnej veličiny X, ak uvedené stredné hodnoty existujú.

Poznámka. Ak je $n-$tý moment konečný, t.j. $\mathcal{E}(X^n) < \infty$, tak píšeme $X \in \mathcal{L}_n(\Omega, \mathcal{A}, P)$, alebo skrátene $X \in \mathcal{L}_n$.

Definícia 10.2. Druhý centrálny moment $\mu_2 = \mathcal{E}\left(X - \mathcal{E}(X)\right)^2$ náhodnej veličiny X (ak existuje) voláme *rozptyl* alebo *disperzia* a označujeme

$$\mathcal{D}(X) = \mathcal{E}\left(X - \mathcal{E}(X)\right)^2 = \mu_2.$$

číslo $\sigma_X = \sqrt{\mathcal{D}(X)}$ nazývame *smerodajnou odchýlkou* náhodnej veličiny X.

Poznámka. Ak $X \in \mathcal{L}_2$, potom $X \in \mathcal{L}_1$, lebo zo Schwarzovej nerovnosti

$$|\mathcal{E}(X)| = \left|\int_{\mathbb{R}} x dF_X(x)\right| = \left|\int_{\mathbb{R}} 1x dF_X(x)\right| \leq$$

$$\leq \sqrt{\int_{\mathbb{R}} x^2 dF_X(x)} \sqrt{\int_{\mathbb{R}} 1^2 dF_X(x)} = \sqrt{\mathcal{E}(X^2)}.$$

Veta 10.4. (Vlastnosti rozptylu.) Nech X, X_1, X_2 sú náhodné veličiny definované na $(\Omega, \mathcal{A}, P)$ s konečnými druhými momentami, $a, a_1, a_2 \in \mathbb{R}$. Potom

(i) $\mathcal{D}(X) \geq 0$,

(ii) $\mathcal{D}(X) = \mathcal{E}(X^2) - \mathcal{E}^2(X)$,

(iii) ak $P(X = a) = 1$, tak $\mathcal{D}(X) = 0$,

(iv) $\mathcal{D}(a_1 + a_2X) = a_2^2\mathcal{D}(X)$,

(v) ak X_1 a X_2 sú nezávislé, tak $\mathcal{D}(X_1 \pm X_2) = \mathcal{D}(X_1) + \mathcal{D}(X_2)$.

Dôkaz:

(i) Pre náhodnú veličinu $Y = (X - \mathcal{E}(X))^2$ platí, že $P(Y \geq 0) = 1$, preto z vlastnosti strednej hodnoty $\mathcal{E}(Y) = \mathcal{D}(X) \geq 0$,

(ii) $\mathcal{D}(X) = \mathcal{E}\left(X - \mathcal{E}(X)\right)^2 = \mathcal{E}[X^2 - 2X\mathcal{E}(X) + (\mathcal{E}(X))^2] =$
$= \mathcal{E}(X^2) - 2\mathcal{E}(X)\mathcal{E}(X) + (\mathcal{E}(X))^2 = \mathcal{E}(X^2) - \mathcal{E}^2(X)$,

(iii) ak je $P(X = a) = 1$, tak X je diskrétna náhodná veličina s pravdepodobnostnou funkciou $(a, 1)$, teda $\mathcal{E}(X) = a1 = a$ a $\mathcal{D}(X) = \mathcal{E}(X - \mathcal{E}(X))^2 = (a - a)^2 \cdot 1 = 0$,

(iv) $\mathcal{D}(a_1 + a_2X) = \mathcal{E}[a_1 + a_2X - \mathcal{E}(a_1 + a_2X)]^2 = \mathcal{E}(a_1 + a_2X - a_1 - a_2\mathcal{E}(X))^2 = \mathcal{E}[a_2^2(X - \mathcal{E}(X))^2] = a_2^2\mathcal{E}(X - \mathcal{E}(X))^2 = a_2^2\mathcal{D}(X)$,

(v) $\mathcal{D}(X_1 \pm X_2) = \mathcal{E}[X_1 \pm X_2 - \mathcal{E}(X_1 \pm X_2)]^2 = \mathcal{E}[X_1 \pm X_2 - \mathcal{E}(X_1) \mp \mathcal{E}(X_2)]^2 = \mathcal{E}[(X_1 - \mathcal{E}(X_1))^2 \pm 2(X_1 - \mathcal{E}(X_1))(X_2 - \mathcal{E}(X_2)) + (X_2 - \mathcal{E}(X_2))^2] = \mathcal{E}(X_1 - \mathcal{E}(X_1))^2 + \mathcal{E}(X_2 - \mathcal{E}(X_2))^2 \pm 2\mathcal{E}[(X_1 - \mathcal{E}(X_1))(X_2 - \mathcal{E}(X_2))]$. Pretože sú X_1 a X_2 nezávislé, platí $\mathcal{E}(X_1X_2) = \mathcal{E}(X_1)\mathcal{E}(X_2)$. Ale tiež $(X_1 - \mathcal{E}(X_1))$ a $(X_2 - \mathcal{E}(X_2))$ sú nezávislé a tiež $\mathcal{E}[(X_1 - \mathcal{E}(X_1))(X_2 - \mathcal{E}(X_2))] = \mathcal{E}(X_1 - \mathcal{E}(X_1))\mathcal{E}(X_2 - \mathcal{E}(X_2)) = 0$. Dostávame, že $\mathcal{D}(X_1 \pm X_2) = \mathcal{D}(X_1) + \mathcal{D}(X_2)$. ♣

Príklad 10.3. Rozptyl náhodnej veličiny s poissonovským rozdelením (rozptyl Poissonovho rozdelenia). Nech $X \sim Po(\lambda)$, teda X má pravdepodobnostnú funkciu $(x_i, p_i)_{i \geq 1}$, kde $x_i = 0, 1, 2, ...$ a $p_i = e^{-\lambda}\dfrac{x_i^\lambda}{x_i!}$. V Príklade 10.1. sme spočítali, že $\mathcal{E}(X) = \lambda$. Platí $\mathcal{D}(X) = \mathcal{E}(X^2) - \mathcal{E}^2(X)$. Spočítame

$$\mathcal{E}(X^2) = \sum_{j=0}^{\infty} j^2 e^{-\lambda}\frac{\lambda^j}{j!} = e^{-\lambda}\left\{\frac{\lambda^1}{1!} + \sum_{j=2}^{\infty}[j(j-1) + j]\frac{\lambda^j}{j!}\right\} =$$

$$= e^{-\lambda} \sum_{j=2}^{\infty} \frac{\lambda^j}{(j-2)!} + e^{-\lambda} \sum_{j=0}^{\infty} j \frac{\lambda^j}{j!} = e^{-\lambda} \lambda^2 \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} + \lambda = \lambda^2 + \lambda.$$

Preto

$$\mathcal{D}(X) = \mathcal{E}(X^2) - \mathcal{E}^2(X) = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Príklad 10.4. Rozptyl náhodnej veličiny s normálnym rozdelením. Nech $X \sim N(\mu, \sigma^2),\ \sigma > 0$, teda jej hustota je $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in (-\infty, \infty)$. V Príklade 10.2. sme spočítali, že $\mathcal{E}(X) = \mu$. Preto

$$\mathcal{D}(X) = \mathcal{E}(X-\mu)^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx =$$

(substitúcia $u = \frac{x-\mu}{\sigma},\quad x = \sigma u + \mu,\ du = \frac{dx}{\sigma}$)

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u^2 e^{-\frac{u^2}{2}} du = 2\frac{\sigma^2}{\sqrt{2\pi}} \int_0^{\infty} u^2 e^{-\frac{u^2}{2}} du =$$

(substitúcia $\frac{u^2}{2} = t,\quad u = \sqrt{2t},\quad udu = dt$)

$$= \frac{2\sigma^2}{\sqrt{2\pi}} \int_0^{\infty} \sqrt{2t} e^{-t} dt = \frac{2\sigma^2}{\sqrt{\pi}} \int_0^{\infty} t^{\frac{3}{2}-1} e^{-t} dt = \frac{2\sigma^2}{\sqrt{\pi}} \Gamma(\tfrac{3}{2}) = \frac{2\sigma^2}{\sqrt{\pi}} \frac{1}{2} \Gamma(\tfrac{1}{2}) = \sigma^2.$$

Medián, módus a kvantily

K charakterizácii rozdelenia pravdepodobnosti náhodnej veličiny X s distribučnou funkciou F sa používajú aj iné charakteristiky. Jednou z nich je *medián* $\tilde{x}$. Je to (ľubovoľné) číslo, pre ktoré platí

$$F(\tilde{x}) \leq \frac{1}{2}, \quad F(\tilde{x}+0) \geq \frac{1}{2}.$$

Vo všeobecnosti tieto podmienky neurčujú medián jednoznačne.

Ďalšia charakteristika je *módus* $\hat{x}$. Ak je náhodná veličina diskrétneho typu s pravdepodobnostnou funkciou $(x_i, p_i)_{i \geq 1}$, tak $\hat{x}$ je to číslo x_j, pre ktoré platí $P(X = \hat{x}) \geq P(X = x_i),\ i = 1, 2, ...$ Ak má X spojité rozdelenie s hustotou f, za módus považujeme tú hodnotu $\hat{x} \in \mathtt{R}$, pre ktorú platí $f(\hat{x}) \geq f(x),\ -\infty < x < \infty$. Ani módus nie je vo všeobecnosti určený jednoznačne.

Zaveďme si funkciu F^{-1} predpisom

$$F^{-1}(u) = \inf\{x \in \mathtt{R}:\ F(x) \geq u\}, \quad 0 < u < 1.$$

Funkcia F^{-1} sa nazýva *kvantilová funkcia* zodpovedajúca distribučnej funkcii F. Hodnoty $F^{-1}(u)$ sú *kvantily*. Teda $\alpha-$kvantilom je $F^{-1}(\alpha)$. Ak je F rastúca a spojitá, potom F^{-1} je inverzná funkcia k distribučnej funkcii F.

Veta 10.5. (Čebyševova nerovnosť.) Nech X je náhodná veličina s konečným druhým momentom. Potom pre ľubovoľné $\varepsilon > 0$ platí

$$P(|X - \mathcal{E}(X)| \geq \varepsilon) \leq \frac{\mathcal{D}(X)}{\varepsilon^2}.$$

Dôkaz: Pre ľubovoľné $\varepsilon > 0$ položme $M_\varepsilon = \{\omega\ :\ |X(\omega) - \mathcal{E}(X)| \geq \varepsilon\}$.

$$P\{\omega\ :\ |X(\omega) - \mathcal{E}(X)| \geq \varepsilon\} = \int_{M_\varepsilon} dP(\omega) = \int_{M_\varepsilon} 1 dP(\omega) \leq \int_{M_\varepsilon} \frac{(X(\omega) - \mathcal{E}(X))^2}{\varepsilon^2} dP(\omega) \leq$$

$$\leq \frac{1}{\varepsilon^2} \int_\Omega (X(\omega) - \mathcal{E}(X))^2 dP(\omega) = \frac{\mathcal{D}(X)}{\varepsilon^2},$$

teda $P(|X - \mathcal{E}(X)| \geq \varepsilon) \leq \frac{\mathcal{D}(X)}{\varepsilon^2}$. ♣

Poznámka. Z Čebyševovej nerovnosti dostávame

$$P(|X - \mathcal{E}(X)| < \varepsilon) = 1 - P(|X - \mathcal{E}(X)| \geq \varepsilon) \geq 1 - \frac{\mathcal{D}(X)}{\varepsilon^2}.$$

V prípade, že zvolíme $\varepsilon = k\sqrt{\mathcal{D}(X)}$, je

$$P(|X - \mathcal{E}(X)| < k\sqrt{\mathcal{D}(X)}) \geq 1 - \frac{1}{k^2},$$

špeciálne pre $k = 3$

$$P(|X - \mathcal{E}(X)| < 3\sqrt{\mathcal{D}(X)}) \geq 1 - \frac{1}{9} \doteq 0.89.$$

Kovariancia a korelačný koeficient

V nasledujúcom budeme predpokladať, že náhodné veličiny majú konečné druhé momenty.

Definícia 10.3. *Kovariancia* náhodných veličín X a Y je (číslo)

$$C(X,Y) = \mathcal{E}\left[(X - \mathcal{E}(X))(Y - \mathcal{E}(Y))\right]$$

a *korelačný koeficient*

$$R(X,Y) = \frac{C(X,Y)}{\sqrt{\mathcal{D}(X)\mathcal{D}(Y)}},$$

ak $\mathcal{D}(X) > 0, \mathcal{D}(Y) > 0$. Niekedy značíme $R(X,Y)$ ako $\varrho_{X,Y}$.

Pomocou vety o strednej hodnote transformovaného náhodného vektora dostávame

Veta 10.6. Ak náhodné veličiny X a Y majú združenú distribučnú funkciu $F(x,y)$, potom

$$C(X,Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mathcal{E}(X))(y - \mathcal{E}(Y)) dF(x,y),$$

teda

(a) v prípade, že $(X,Y)'$ je náhodný vektor s pravdepodobnostnou funkciou $((x_m, y_m), p_m)_{m \in J}$, tak

$$C(X,Y) = \sum_{m \in J} (x_m - \mathcal{E}(X))(y_m - \mathcal{E}(Y)) p_m;$$

(b) v prípade, že $(X,Y)'$ je spojitý náhodný vektor so združenou hustotou $f(x,y)$, tak

$$C(X,Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mathcal{E}(X))(y - \mathcal{E}(Y)) f(x,y) dx dy.$$

Veta 10.7. (Vlastnosti kovariancie a korelačného koeficienta.) Nech X a Y sú náhodné veličiny, s konečnými nenulovými rozptylmi, $a_1, a_2, b_1, b_2 \in \mathtt{R}$. Potom

(i) $C(X,X) = \mathcal{D}(X)$ a $R(X,X) = 1$;

(ii) $C(X,Y) = C(Y,X)$ a $R(X,Y) = R(Y,X)$;

(iii) $C(X,Y) = \mathcal{E}(XY) - \mathcal{E}(X)\mathcal{E}(Y)$;

(iv) ak sú X a Y nezávislé náhodné veličiny, tak $C(X,Y) = R(X,Y) = 0$;

(v) $|C(X,Y)| \le \sqrt{\mathcal{D}(X)\mathcal{D}(Y)} = \sigma_X \sigma_Y$ a $|R(X,Y)| \le 1$;

(vi) $C(a_1 + a_2X, b_1 + b_2Y) = a_2b_2C(X,Y)$ a ak $a_2 \neq 0, b_2 \neq 0$, tak

$R(a_1 + a_2X, b_1 + b_2Y) = R(X,Y)\mathrm{sign}(a_2b_2)$;

(vii) $\mathcal{D}(X \pm Y) = \mathcal{D}(X) + \mathcal{D}(Y) \pm 2C(X,Y)$;

(viii) $R(X,Y) = 1 \Longleftrightarrow$ existujú konštanty a a $b > 0$ také, že $P(Y = a + bX) = 1$

a

$R(X,Y) = -1 \Longleftrightarrow$ existujú konštanty a a $b < 0$ také, že $P(Y = a + bX) = 1$.

Dôkaz:

(i) $C(X,X) = \mathcal{E}(X - \mathcal{E}(X))^2 = \mathcal{D}(X)$ a $R(X,X) = \dfrac{C(X,X)}{\sqrt{\mathcal{D}(X)}\sqrt{\mathcal{D}(X)}} = 1$;

(ii) $C(X,Y) = \mathcal{E}[(X - \mathcal{E}(X))(Y - \mathcal{E}(Y))] = \mathcal{E}[(Y - \mathcal{E}(Y))(X - \mathcal{E}(X))] = C(Y,X)$,

teda aj $R(X,Y) = \dfrac{C(X,Y)}{\sqrt{\mathcal{D}(X)}\sqrt{\mathcal{D}(Y)}} = \dfrac{C(Y,X)}{\sqrt{\mathcal{D}(Y)}\sqrt{\mathcal{D}(X)}} = R(Y,X)$;

(iii) $C(X,Y) = \mathcal{E}[(X - \mathcal{E}(X))(Y - \mathcal{E}(Y))] = \mathcal{E}[XY - X\mathcal{E}(Y) - Y\mathcal{E}(X) + \mathcal{E}(X)\mathcal{E}(Y)]$

$= \mathcal{E}(XY) - \mathcal{E}(X)\mathcal{E}(Y)$;

(iv) ak su X a Y nezávislé, tak $\mathcal{E}(XY) = \mathcal{E}(X)\mathcal{E}(Y)$ a teda $C(X,Y) = \mathcal{E}(XY)$

$-\mathcal{E}(X)\mathcal{E}(Y) = \mathcal{E}(X)\mathcal{E}(Y) - \mathcal{E}(X)\mathcal{E}(Y) = 0$ a preto aj $R(X,Y) = 0$;

(v) podľa Schwarzovej nerovnosti

$$\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - \mathcal{E}(X))(y - \mathcal{E}(Y))dF(x,y)\right]^2 \le$$

$$\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - \mathcal{E}(X))^2 dF(x,y)\right]\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(y - \mathcal{E}(Y))^2 dF(x,y)\right], \tag{10.1}$$

teda $|C(X,Y)| \le \sqrt{\mathcal{D}(X)\mathcal{D}(Y)}$ a podelením tejto rovnosti výrazom $\sqrt{\mathcal{D}(X)\mathcal{D}(Y)}$ dostávame $|R(X,Y)| = \dfrac{|C(X,Y)|}{\sqrt{\mathcal{D}(X)\mathcal{D}(Y)}} \le 1$;

(vi) $C(a_1 + a_2X, b_1 + b_2Y) = \mathcal{E}[(a_1 + a_2X - \mathcal{E}(a_1 + a_2X))(b_1 + b_2Y - \mathcal{E}(b_1 + b_2Y))] = \mathcal{E}\{[a_2(X - \mathcal{E}(X))][b_2(Y - \mathcal{E}(Y))]\} = a_2b_2\mathcal{E}[(X - \mathcal{E}(X))(Y - \mathcal{E}(Y))] = a_2b_2C(X,Y)$ a ak $a_2 \neq 0, b_2 \neq 0$, tak

$$R(a_1 + a_2X, b_1 + b_2Y) = \frac{a_2b_2C(X,Y)}{\sqrt{a_2^2\mathcal{D}(X)}\sqrt{b_2^2\mathcal{D}(Y)}} =$$

$$\frac{a_2b_2}{|a_2||b_2|}R(X,Y) = \mathrm{sign}(a_2b_2)R(X,Y);$$

(vii) $\mathcal{D}(X \pm Y) = \mathcal{E}[X \pm Y - \mathcal{E}(X \pm Y)]^2 = \mathcal{E}[(X - \mathcal{E}(X)) \pm (Y - \mathcal{E}(Y))]^2 = \mathcal{E}[(X - \mathcal{E}(X))^2 \pm 2(X - \mathcal{E}(X))(Y - \mathcal{E}(Y)) + (Y - \mathcal{E}(Y))^2] = \mathcal{D}(X) \pm 2C(X,Y) + \mathcal{E}(Y)$;

(viii) $R(X,Y) = 1 \Rightarrow |C(X,Y)| = \sqrt{\mathcal{D}(X)\mathcal{D}(Y)} > 0$, t.j. nastala rovnosť v Schwarzovej nerovnosti (10.1), ktorá môže nastať práve vtedy keď

1. $\exists b \neq 0$, že $\mu_F\{(x,y) \in \mathrm{R}^2 :\ y - \mathcal{E}(Y) = b(x - \mathcal{E}(X))\} = 1$,

alebo keď

2. $\mu_F\{(x,y) \in \mathrm{R}^2 :\ x - \mathcal{E}(X) = 0\} = 1$ alebo $\mu_F\{(x,y) \in \mathrm{R}^2 :\ y - \mathcal{E}(Y) = 0\} = 1$.

Pretože v druhom prípade by bola $\mathcal{D}(X) = 0$ alebo $\mathcal{D}(Y) = 0$ (čo nemôže byť), nastáva iba 1. prípad a teda

$$\exists b \neq 0 \ \ P\{\omega :\ Y(\omega) - \mathcal{E}(Y) = b(X(\omega) - \mathcal{E}(X))\} = 1,$$

čiže

$$\exists b \neq 0 \ \ P\{Y = \mathcal{E}(Y) - b\mathcal{E}(X) + b(X) = a + bX\} = 1.$$

Preto $C(X,Y) = C(X, a+bX) = bC(X,X) > 0$ (podľa (vi)) a $b > 0$.

Prípad $R(X,Y) = -1$ dokážeme úplne analogicky. ♣

Poznámka. Ak je $C(X,Y) = 0$, teda ak je $R(X,Y) = 0$, potom povieme, že náhodné veličiny X a Y su *nekorelované*.

Príklad 10.5. Nech (X,Y) je diskrétny náhodný vektor s pravdepodobnostnou funkciou $((x,y)_i, p_i)_{i\in J}$, pričom $M = \{(x,y)_i\}_{i\in J} = \{-1,0,1\} \times \{-1,0,1\}$ a $p(-1,1) = p(-1,-1) = p(1,1) = p(1,-1) = \frac{1}{6},\ p(0,0) = \frac{1}{3},\ p(-1,0) = p(0,1) = p(0,-1) = p(1,0) = 0$. Vypočítajte $R(X,Y)$ a rozhodnite, či X a Y sú nezávislé.

$x \setminus y$	-1	0	1	$p_X(x)$
-1	$1/6$	0	$1/6$	$2/6$
0	0	$1/3$	0	$1/3$
1	$1/6$	0	$1/6$	$2/6$
$p_Y(y)$	$2/6$	$1/3$	$2/6$	1

Riešenie:
$\mathcal{E}(X) = \sum_{x\in\{-1,0,1\}} x p_X(x) = (-1)\frac{1}{3} + 0\frac{1}{3} + 1\frac{1}{3} = 0$ a rovnako $\mathcal{E}(Y) = 0$. Ďalej $\mathcal{E}(XY) =$
$= \sum_{(x,y)\in\{-1,0,1\}\times\{-1,0,1\}} xy\ p_{X,Y}(x,y) = (-1)(-1)\frac{1}{6} + 1(-1)\frac{1}{6} + (-1)1\frac{1}{6} + 1\cdot 1\frac{1}{6} = 0$, teda $C(X,Y) = \mathcal{E}(XY) - \mathcal{E}(X)\mathcal{E}(Y) = 0$. Náhodné veličiny X a Y sú nekorelované. Ale $p_{X,Y}(-1,-1) = \frac{1}{6} \neq p_X(-1)p_Y(-1) = \frac{2}{6}\frac{2}{6} = \frac{1}{9}$.

11. Charakteristická funkcia

Pravdepodobnostné správanie sa náhodných veličín a vektorov úplne charakterizuje ich rozdelenie pravdepodobnosti resp. distribučná funkcia. Mnoho vlastností náhodných veličín alebo vektorov je ale ťažkopánde a zdĺhavé dokazovať pomocou distribučnej funkcie. Pracujeme preto s iným analytickým vyjadrením rozdelenia pravdepodobnosti, a síce s Fourierovou-Stieltjesovou transformáciou, ktorá sa v teórii pravdepodobnosti volá *charakteristická funkcia*.

Definícia 11.1 Charakteristická funkcia náhodnej veličiny X je komplexná funkcia reálnej premennej $\psi(\cdot): \mathtt{R} \to \mathtt{C}$ definovaná ako

$$\psi(t) = \mathcal{E}\left(e^{itX}\right), \quad t \in \mathtt{R}.$$

V teórii pravdepodobnosti sa dokazuje množstvo vlastností charakteristických funkcií. Niektoré sú obsahom nasledujúcej vety. Dôkazy tejto aj nasledujúcich viet nájdeme v napr. knihe Rényi, A., Teorie pravděpodobnosti, ACADEMIA, Praha, 1972.

Veta 11.1. Nech X je náhodná veličina a $\psi(t)$ jej charakteristická funkcia. Potom

(i) $|\psi(t)| \leq 1 \quad \forall t \in \mathtt{R}$;

(ii) $\psi(0) = 1$;

(iii) $\forall t \in \mathtt{R} \ \ \overline{\psi(t)} = \psi(-t)$;

(iv) $\psi(t)$ je rovnomerne spojitá na $\mathtt{R}$. $(\forall_{\ \epsilon>0}\ \exists_{\ \delta>0}\ \forall_{\ t_1,t_2:\ |t_2-t_1|<\delta}\ |\psi(t_2) - \psi(t_1)| < \epsilon)$

Veta 11.2. Ak existuje prvých n momentov $\mu_1', ..., \mu_n'$ náhodnej veličiny X a tieto momenty sú konečné, potom charakteristická funkcia $\psi(t)$ náhodnej veličiny X má prvých n derivácií a platí

$$\psi^{(k)}(0) = i^k \mu_k', \quad k = 1, 2, ..., n.$$

Ďalej platí

$$\psi(t) = \sum_{k=0}^{n} \mu'_k \frac{(it)^k}{k!} + o(t^n),$$

kde $o(t^n)$ je taká fumkcia, že $\lim_{t\to 0} \frac{o(t^n)}{t^n} = 0$.

Veta 11.3. Ak je $\psi(t)$ charakteristická funkcia zodpovedajúca distribučnej fumkcii $F(x)$ a $a, b,\ a < b$ body spojitosti funkcie $F(x)$, tak platí

$$F(b) - F(a) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\psi(t) \frac{e^{-ita} - e^{-itb}}{2it} - \psi(-t) \frac{e^{ita} - e^{itb}}{2it} \right] dt.$$

Veta 11.4. Ak pre charakteristickú funkciu $\psi(t)$ náhodnej premennej X platí $\int_{-\infty}^{\infty} |\psi(t)| dt < \infty$, tak má X spojitú hustotu $f(x)$ a môžeme ju vyjadriť v tvare

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi(t) e^{-itx} dt.$$

Príklad 11.1. Nech $X \sim A(\theta)$, teda X má alternatívne rozdelenie s pravdepodobnostnou funkciou $(x_i, p_i)_{i=1,2}$, pričom $x_1 = 0, x_2 = 1$ a $p_1 = 1 - \theta,\ p_2 = \theta$. Charakteristická funkcia tejto náhodnej premennej je

$$\psi_X(t) = \mathcal{E}(e^{itX}) = e^{itx_1}(1-\theta) + e^{itx_2}\theta = e^{it0}(1-\theta) + e^{it1}\theta = 1 - \theta + e^{it}\theta.$$

Charakteristická funkcia $Y \sim Bi(n, \theta)$ je $\psi(t) = (1 - \theta + e^{it}\theta)^n$.

Príklad 11.2. Nech $X \sim Ro(-a, a)$ (rovnomerne rozdelená na $(-a, a)$). Potom jej hustota je

$$f(x) = \begin{cases} \dfrac{1}{2a}, & \text{ak } -a < x < a \\ 0, & \text{ak } x \notin (-a, a). \end{cases}$$

Charakteristická funkcia tejto náhodnej veličiny je pre $t \neq 0$

$$\psi(t) = \mathcal{E}\left(e^{itX}\right) = \int_{-\infty}^{\infty} e^{itx} f(x) dx = \int_{-a}^{a} e^{itx} f(x) dx = \frac{1}{2a} \left[\frac{e^{itx}}{it} \right]_{-a}^{a} =$$

$$= \frac{1}{2a} \frac{e^{ita} - e^{-ita}}{it} = \frac{\cos(ta) + i\sin(ta) - \cos(-ta) + i\sin(ta)}{2at} = \frac{\sin at}{at}$$

a

$$\psi(0) = \mathcal{E}\left(e^{i0X}\right) = \int_{-\infty}^{\infty} e^{i0x} f(x) dx = 1.$$

Príklad 11.3. Nech $U \sim N(0, 1)$. Jej charakteristická funkcia je

$$\psi_U(t) = \mathcal{E}\left(e^{itU}\right) = \int_{-\infty}^{\infty} e^{itu} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(u^2 - 2itu)} du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}[(u-it)^2 + t^2]} du =$$

(substitúcia $u - it = s,\ du = ds$, môže sa použiť aj Dodatok na str. 90)

$$= e^{-\frac{1}{2}t^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}s^2} ds = e^{-\frac{t^2}{2}}.$$

Dokážte, že charakteristická funkcia $X \sim N(\mu, \sigma^2),\ \sigma^2 > 0$ je $\psi_X(t) = e^{i\mu t} e^{-\frac{\sigma^2 t^2}{2}}$. (Použite substitúciu $\frac{x-\mu}{\sigma} = u$.)

Dokážme si ešte niektoré vlastnosti charakteristickej funkcie.

Veta 11.4. Nech X je náhodná veličina, $\psi_X(t)$ jej charakteristická funkcia, a, b reálne čísla. Potom náhodná veličina $Y = a + bX$ má charakteristickú funkciu $\psi_Y(t) = e^{ita}\psi(tb)$.

Dôkaz:

$$\psi_Y(t) = \mathcal{E}\left(e^{itY}\right) = \mathcal{E}\left(e^{it(a+bX)}\right) = \mathcal{E}\left(e^{ita}e^{itbX}\right) = e^{ita}\mathcal{E}\left(e^{itbX}\right) = e^{ita}\psi_X(tb). \clubsuit$$

Najdôležitejšie aplikácie pre charakteristickú funkciu plynú z nasledujúcej vety.

Veta 11.5. Nech X_1 a X_2 sú nezávislé náhodné veličiny s charakteristickými funkciami $\psi_1(t)$ a $\psi_2(t)$. Potom náhodná veličina $X = X_1 + X_2$ má charakteristickú funkciu $\psi_X(t) = \psi_1(t)\psi_2(t)$.

Dôkaz:

$$\psi_X(t) = \mathcal{E}\left(e^{it(X_1+X_2)}\right) = \mathcal{E}\left(e^{itX_1}e^{itX_2}\right) = \mathcal{E}\left(e^{itX_1}\right)\mathcal{E}\left(e^{itX_2}\right) = \psi_1(t)\psi_2(t). \quad \clubsuit$$

Upozorňujeme len, že tvrdenie Vety 11.5 podľa nasledujúceho protipríkladu nemožno obrátiť.

Príklad 11.4. Nech X_1 má Cauchyho rozdelenie s hustotou $f(x) = \frac{1}{\pi}\frac{1}{1+x^2}$, $x \in \mathtt{R}$. Položme $X_2 = X_1$ a spočítajme charakteristickú funkciu náhodnej veličiny $X = X_1 + X_2 = 2X_1$. Charakteristická funkcia náhodnej veličiny X_1 je

$$\psi_{X_1}(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} e^{itx}\frac{1}{1+x^2}dx = e^{-|t|}$$

(podľa rezíduovej vety, môže sa použiť aj Dodatok na str. 98). Pretože $X = 2X_1$, je

$$\psi_X(t) = \psi_{0+2X_1}(t) = e^{it0}\psi_{X_1}(2t) = e^{-|2t|}.$$

Dostali sme $\psi_{X_1+X_1}(t) = \psi_{X_1}(t)\psi_{X_2}(t)$, ale X_1 a X_2 nie sú nezávislé.

Charakteristická funkcia náhodného vektora

Nech $\mathbf{X} = (X_1, ..., X_n)'$ je $n-$rozmerný náhodný vektor.

Definícia 11.2. Funkciu $\psi : \mathtt{R}^n \to \mathtt{C}$ definovanú predpisom

$$\psi(\mathbf{t}) = \psi(t_1, ..., t_n) = \mathcal{E}\left(e^{i\mathbf{t}'\mathbf{X}}\right) = \mathcal{E}\left(e^{i\sum_{j=1}^n t_jX_j}\right)$$

budeme nazývať *charakteristickou funkciou náhodného vektora* $\mathbf{X}$.

Analogicky ako v jednorozmernom prípade sa dajú odvodiť vlastnosti charakteristickej funkcie náhodného vektora.

Veta 11.6. Platí

(i) $|\psi(\mathbf{t})| \leq 1$ pre všetky $\mathbf{t} \in \mathtt{R}^n$;

(ii) $\psi(0, 0, ..., 0) = \psi(\mathbf{0}) = 1$;

(iii) $\psi(-t_1, -t_2, ..., -t_n) = \overline{\psi(t_1, ..., t_n)}$;

(iv) ψ je rovnomerne spojitá na $\mathtt{R}^n$;

(v) $\mathbf{b} \in \mathtt{R}^m$, $\mathbf{A}_{m,n}$ je matica reálnych čísel, $\mathbf{Y} = \mathbf{b} + \mathbf{AX}$, potom $\psi_{\mathbf{Y}}(\mathbf{u}) = e^{i\mathbf{u}'\mathbf{b}}\psi_{\mathbf{X}}(\mathbf{A}'\mathbf{u})$, $u \in \mathtt{R}^m$;

(vi) keď existujú stredné hodnoty $\mathcal{E}(X_j)$ pre $j = 1, 2, ..., n$, potom

$$\left[\frac{\partial\psi(\mathbf{t})}{\partial t_j}\right]_{\mathbf{t}=(0,0,...,0)} = i\mathcal{E}(X_j);$$

(vii) keď existujú stredné hodnoty $\mathcal{E}(X_jX_k)$ pre $j,k=1,2,...,n$, potom

$$\left[\frac{\partial^2\psi(\mathbf{t})}{\partial t_j\partial t_k}\right]_{\mathbf{t}=(0,0,...,0)} = -\mathcal{E}(X_jX_k);$$

(viii) ak $\psi_j(t)$ je charakteristická funkcia náhodnej veličiny X_j, potom $\psi_j(t_j)=\psi_{\mathbf{X}}(0,0,...,t_j,0,...,0)$;

(ix) nech $\mathbf{X}$ má charakteristickú funkciu $\psi_{\mathbf{X}}(t_1,...,t_n)$ a $\mathbf{Y}$ má charakteristickú funkciu $\psi_{\mathbf{Y}}(t_1,...,t_n)$, pričom $\mathbf{X},\mathbf{Y}$ sú nezávislé, potom $\mathbf{Z}=\mathbf{X}+\mathbf{Y}$ má charakteristickú funkciu $\psi_{\mathbf{Z}}(\mathbf{t})$ a platí $\psi_{\mathbf{Z}}(\mathbf{t})=\psi_{\mathbf{X}}(\mathbf{t})\psi_{\mathbf{Y}}(\mathbf{t})$;

(x) zložky náhodného vektora $\mathbf{X}=(X_1,X_2,...,X_n)'$ sú nezávislé práve vtedy ak $\psi_{\mathbf{X}}(\mathbf{t})=\prod_{i=1}^n\psi_{X_i}(t_i)$ (dôkaz pozri v Rényi, A. Teoria pravděpodobnosti, ACADEMIA, Praha, 1972).

12. Konvergencia náhodných veličín

Majme postupnosť náhodných veličín $X_1,X_2,...$ a náhodnú veličinu X. Nech sú všetky tieto veličiny definované na (tom istom) pravdepodobnostnom priestore $(\Omega,\mathcal{A},P)$.

Definícia 12.1. Povieme, že X_n konverguje k X *skoro iste*, ak

$$P(\{\omega:\ X_n(\omega)\to X(\omega)\})=1.$$

Ak pre každé $\epsilon>0$ platí

$$P(\{\omega:\ |X_n(\omega)-X(\omega)|>\epsilon\})\to 0,$$

potom povieme, že X_n konverguje k X *podľa pravdepodobnosti*. Nech $\mathcal{E}(X_n^2)<\infty$ pre $n=1,2,...$. Ak platí

$$\mathcal{E}\left((X_n-X)^2\right)\to 0,$$

potom povieme, že X_n konverguje k X *podľa (kvadratického) stredu*.

Nech X_n má distribučnú funkciu $F_n(\cdot)$ a nech Nech X má distribučnú funkciu $F(\cdot)$. Povieme, že X_n konverguje k X *v distribúcii* ak $F_n(x)$ konverguje k $F(x)$ v každom bode x, v ktorom je $F(\cdot)$ spojitá. Táto konvergencia sa často označuje aj ako $\mathcal{L}(X_n)\to\mathcal{L}(X)$ a hovorí sa, že X_n má asymptotické rozdelenie $\mathcal{L}(X)$. Niekedy sa táto konvergencia volá aj *slabá konvergencia*.

Lema 12.1. (i) Postupnosť X_n konverguje k X podľa pravdepodobnosti práve vtedy, ak $\forall\epsilon>0$ a $\forall\delta>0$ existuje n_0, že pre všetky $n\geq n_0$ platí

$$P(\{\omega:\ |X_n(\omega)-X(\omega)|>\epsilon\})<\delta.$$

(ii) Postupnosť X_n konverguje k X podľa pravdepodobnosti práve vtedy, ak $\forall k\in\mathbb{N}$ a $\forall\delta>0$ existuje n_0, že pre všetky $n\geq n_0$ platí

$$P(\{\omega:\ |X_n(\omega)-X(\omega)|\geq\frac{1}{k}\})<\delta.$$

(iii) Postupnosť X_n konverguje k X podľa pravdepodobnosti práve vtedy, ak $\forall\epsilon>0$

$$P(\{\omega:\ |X_n(\omega)-X(\omega)|\geq\epsilon\})\to 0.$$

Dôkaz je jednoduchý a spravte si ho ako cvičenie.

Veta 12.1. (Limitná veta pre charakteristické funkcie.) Nech je daná postupnosť distribučných funkcií $F_1(\cdot),F_2(\cdot),...$ a im zodpovedajúca postupnosť charakteristických funkcií $\psi_1(\cdot),\psi_2(\cdot),...$. K tomu, aby postupnosť $\{F_n(\cdot)\}$ konvergovala k nejakej distibučnej funkcii $F(\cdot)$ vo všetkých bodoch spojitosti tejto funkcie, je nutné a stačí, aby postupnosť $\{\psi_n(\cdot)\}$ konvergovala v každom bode k nejakej funkcii $\psi(\cdot)$, ktorá je spojitá v bode $t = 0$. Ak je táto podmienka splnená, tak $\psi(\cdot)$ je charakteristická funkcia odpovedajúca distribučnej funkcii $F(\cdot)$ a postupnosť $\{\psi_n(\cdot)\}$ konverguje k $\psi(\cdot)$ rovnomerne na každom konečnom intervale. $(\forall_{<a,b>} \forall_{\epsilon>0} \exists_N \forall_{n>N} \forall_{t\in<a,b>} |\psi_n(t) - \psi(t)| < \epsilon)$

Dôkaz vety nájdete napríklad v knihe Rényi, A. Teorie pravděpodobnosti, ACADEMIA, Praha, 1972.

Veta 12.2. a) Z konvergencie skoro iste plynie konvergencia podľa pravdepodobnosti.

b) Z konvergencie podľa stredu plynie konvergencia podľa pravdepodobnosti.

c) Z konvergencie podľa pravdepodobnosti plynie konvergencia v distribúcii.

Dôkaz: pozri Anděl, J., Matematická statistika, SNTL, Praha, 1985.

Poznámka. Bez ďalších podmienok sa tvrdenie Vety 12.2 nedá zosilniť. Z konvergencie skoro iste neplynie konvergencia podľa stredu a z konvergencie podľa stredu neplynie konvergencia skoro iste. Z konvergencie podľa pravdepodobnosti neplynie konvergencia skoro iste ani konvergencia podľa stredu. Z konvergencie v distribúcii neplynie konvergencia podľa pravdepodobnosti ani konvergencia skoro iste ani konvergencia podľa stredu. Protipríklady nájdeme v knižkách o teórii pravdepodobnosti.

13. Zákon veľkých čísel

Ak máme postupnosť náhodných veličín $X_1, X_2, ...$, ktoré sú nezávislé a rovnako rozdelené, potom "výberový priemer", teda náhodná veličina $\frac{1}{n}\sum_{i=1}^{n} X_i$ "sa blíži" (teda jej realizácia vždy "lepšie a lepšie" vyjadruje) strednú hodnotu $\mathcal{E}(X_1)$ (len upozorňujeme, že stredné hodnoty náhodných veličín $X_1, X_2, ...$ sú rovnaké). Tento fakt matematicky vyjadruje *zákon veľkých čísel.* Jeho snáď najjednoduchšia podoba je:

Veta 13.1. (Zákon veľkých čísel.) Nech $X_1, X_2, ...$ sú (po dvoch) nezávislé náhodné veličiny s rovnakými strednými hodnotami μ (konečnými) a rovnakými (konečnými) rozptylmi σ^2 definované na (rovnakom) pravdepodobnostnom priestore $(\Omega, \mathcal{A}, P)$. Potom pre $n \to \infty$ platí

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \to \mu$$

podľa pravdepodobnosti.

Dôkaz: Ľahko sa vidí, že platí

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

Z Čebyševovej nerovnosti (Veta 10.5) dostávame pre $\forall\ \epsilon > 0$

$$P(|\overline{X} - \mu| \geq \epsilon) \leq \frac{\sigma^2}{n\epsilon^2},$$

pričom samozrejme pre $n \to \infty$ platí $\dfrac{\sigma^2}{n\epsilon^2} \to 0$, takže $P(|\overline{X} - \mu| \geq \epsilon) \to 0$. ♣

Iná modifikácia tohto zákona, ktorá sa často používa v štatistike je

Veta 13.2. (Chinčinova) Nech $X_1, X_2, ...$ sú nezávislé náhodné veličiny rovnako rozdelené s konečnou strednou hodnotou μ a definované na (rovnakom) pravdepodobnostnom priestore $(\Omega, \mathcal{A}, P)$. Potom pre $n \to \infty$ platí

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \to \mu$$

podľa pravdepodobnosti.

Dôkaz nájdeme napríklad v knižke Anděl, J., Matematická statistika, SNTL, Praha, 1985.

Niektoré dôsledky uvedených zákonov veľkých čísel sú napr.

Dôsledok. Nech $X_1, X_2, ...$ sú (po dvoch) nezávislé náhodné veličiny s rovnakými strednými hodnotami μ (konečnými) a s rozptylmi $\mathcal{D}(X_i) \leq c, i = 1, 2, ...$ Potom $\{X_n\}_{n=1}^{\infty}$ spĺňa zákon veľkých čísel.

Dôsledok (Markovova veta). Nech $X_1, X_2, ...$ sú (po dvoch) nezávislé náhodné veličiny s rovnakými strednými hodnotami μ (konečnými) a s rozptylmi $\mathcal{D}(X_i)$, pričom $\lim_{n\to\infty} \frac{1}{n^2}\mathcal{D}(\sum_{i=1}^{n} X_i) = 0$. Potom $\{X_n\}_{n=1}^{\infty}$ spĺňa zákon veľkých čísel.

Dôsledok (Bernoulliho veta). Majme postupnosť nezávislých pokusov, pričom každý može končiť úspechom s pravdepodobnosťou θ alebo neúspechom s pravdepodobnosťou $1 - \theta$, ($\theta \in (0, 1)$). Označme náhodnú veličinu

$$Y_n \; - \; \text{počet úspechov v } n \text{ nezávislých pokusoch.}$$

$Z_n = \dfrac{1}{n} Y_n$ je relatívna početnosť úspechov v n pokusoch. Platí, že

$$Z_n = \frac{1}{n} Y_n \to \theta$$

podľa pravdepodobnosti.

Dôkaz: Ak označíme náhodnú veličinu

$$X_i = \begin{cases} 0, & \text{ak v } i\text{−tom pokuse bol neúspech} \\ 1, & \text{ak v } i\text{−tom pokuse bol úspech.} \end{cases}$$

$X_1, X_2, ...$ sú nezávislé, $P(X_i = 0) = 1 - \theta$, $P(X_i = 1) = \theta$, $\mathcal{E}(X_i) = \theta$, $\mathcal{D}(X_i) = \theta(1 - \theta) \leq \dfrac{1}{4}$. Platí $Y_n = \sum_{i=1}^{n} X_i$ a $Z_n = \dfrac{1}{n} Y_n = \dfrac{1}{n} \sum_{i=1}^{n} X_i$. Podľa Dôsledku pred Markovou vetou $Z_n \to \theta$ podľa pravdepodobnosti. ♣

Vyššieuvedené tvary zákona veľkých čísel zaručovali konvergenciu (výberového priemeru) $\overline{X}_n$ k strednej hodnote μ podľa pravdepodobnosti. Preto sa volajú *slabé zákony veľkých čísel.* Dajú sa odvodiť vety, ktoré zaručujú takúto konvergenciu skoro iste. Volajú sa *silné zákony veľkých čísel.*

Poznámka. K tomu, aby postupnosť náhodných veličín $X_1, X_2, ...$ spĺňala silný zákon veľkých čísel, stačí, aby táto postupnosť spĺňala podmienky Chinčinovej vety. Toto tvrdenie sa volá II. Kolmogorova veta a jej dôkaz je napr. v knižke Dupač, V., Hušková, M., Pravděpodobnost a matematická statistika, KAROLINUM, Praha, 2001. Tam nájdeme aj iné formulácie silného zákona veľkých čísel.

14. Centrálne limitné vety

Majme postupnosť nezávislých náhodných veličín $X_1, X_2, ...$, ktoré sú definované na tom istom pravdepodobnostnom priestore $(\Omega, \mathcal{A}, P)$. Ak $\mathcal{E}(X_i) = \mu_i$ a $\mathcal{D}(X_i) = \sigma_i^2$, tak náhodné veličiny

$C_i = X_i - \mu_i$ nazývame centrované (majú nulovú strednú hodnotu);

$U_i = \dfrac{X_i - \mu_i}{\sigma_i}$ nazývame štandardizované (majú nulovú strednú hodnotu a jednotkový rozptyl).

Čo je štandardizovaný priemer nezávislých náhodných veličín $X_1, X_2, ...$?

$$\mathcal{E}(\overline{X}_n) = \mathcal{E}(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu_i, \quad \mathcal{D}(\overline{X}_n) = \mathcal{D}(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma_i^2,$$

preto štandardizovaný priemer nezávislých náhodných veličín $X_1, X_2, ...$ je

$$U_{\overline{X}_n} = \frac{\overline{X}_n - \mathcal{E}(\overline{X}_n)}{\sqrt{\mathcal{D}(\overline{X}_n)}} = \frac{\frac{1}{n}\sum_{i=1}^n X_i - \frac{1}{n}\sum_{i=1}^n \mu_i}{\sqrt{\frac{1}{n^2}\sum_{i=1}^n \sigma_i^2}} = \frac{\sum_{i=1}^n (X_i - \mu_i)}{\sqrt{\sum_{i=1}^n \sigma_i^2}}.$$

Ak $\mathcal{E}(X_i) = \mu$ a $\mathcal{D}(X_i) = \sigma^2$, tak

$$U_{\overline{X}_n} = \frac{\sum_{i=1}^n (X_i - \mu)}{\sqrt{n\sigma^2}} = \frac{\sum_{i=1}^n (X_i - \mu)}{\sigma\sqrt{n}} = \frac{1}{\sigma\sqrt{n}}(X_1 + ... + X_n - n\mu).$$

Centrálne limitné vety tvrdia, že za dosť všeobecných podmienok má štandardizovaný priemer nezávislých náhodných veličín asymptoticky normované normálne rozdelenie. Teda konverguje v distribúcii k náhodnej veličine s $N(0,1)$ rozdelením.

Veta 14.1. (Lindebergova CLV.) Nech $X_1, X_2, ...$ sú nezávislé náhodné veličiny s rovnakým rozdelením pravdepodobnosti so strednou hodnotou μ a konečným nenulovým rozptylom σ^2. Potom

$$U_{\overline{X}_n} = \frac{1}{\sigma\sqrt{n}}(X_1 + ... + X_n - n\mu)$$

konverguje k distribúcii k náhodnej veličine $X \sim N(0,1)$.

Dôkaz: Položme $Y_i = \frac{X_i - \mu}{\sigma}$, $i = 1, 2, ...$ Náhodné veličiny $Y_1, Y_2, ...$ sú nezávislé a štandardizované, teda $\mathcal{E}(Y_i) = \mu_1' = 0$. Ich rozptyl je 1. Je to aj ich počiatočný moment druhého rádu, teda μ_2'. Nech charakteristická funkcia ich rozdelenia je $\psi(\cdot)$. Podľa Vety 11.2 je

$$\psi(t) = \mu_0' \frac{(it)^0}{0!} + \mu_1' \frac{(it)^1}{1!} + \mu_2' \frac{(it)^2}{2!} + o(t^2) = 1 - \frac{t^2}{2} + o(t^2),$$

kde $o(t^2)$ je (nejaká) funkcia $R(t)$, pričom $\lim_{t \to 0} \frac{R(t)}{t^2} = 0$.
Charakteristická funkcia $\psi_j(t)$ náhodnej veličiny $\dfrac{Y_j}{\sqrt{n}}$ je

$$\mathcal{E}\left(e^{\frac{itY_j}{\sqrt{n}}}\right) = \mathcal{E}\left(e^{i\frac{t}{\sqrt{n}}Y_j}\right) = \psi\left(\frac{t}{\sqrt{n}}\right) = 1 - \frac{t^2}{2n} + R\left(\frac{t}{\sqrt{n}}\right).$$

Pretože $U_{\overline{X}_n} = \dfrac{Y_1}{\sqrt{n}} + ... + \dfrac{Y_n}{\sqrt{n}}$, je charakteristická funkcia $\psi_n(t)$ náhodnej veličiny $U_{\overline{X}_n}$ rovná

$$\psi_n(t) = \left[1 - \frac{t^2}{2n} + R\left(\frac{t}{\sqrt{n}}\right)\right]^n,$$

pričom pre každé pevné t je

$$\lim_{n\to\infty} nR\left(\frac{t}{\sqrt{n}}\right) = \lim_{n\to\infty} t^2 \frac{R\left(\frac{t}{\sqrt{n}}\right)}{\frac{t^2}{n}} = t^2 \lim_{\frac{t}{\sqrt{n}} \to 0} \frac{R\left(\frac{t}{\sqrt{n}}\right)}{\left(\frac{t}{\sqrt{n}}\right)^2} = 0.$$

Pre každé pevné t dostávame

$$\lim_{n\to\infty} \psi_n(t) = \lim_{n\to\infty} \left[1 - \frac{\frac{t^2}{2} - nR\left(\frac{t}{\sqrt{n}}\right)}{n}\right]^n = e^{-\frac{t^2}{2}},$$

čo je charakteristická funkcia náhodnej veličiny s $N(0,1)$ rozdelením. Podľa Vety 12.1 máme vetu dokázanú.
♣

Veta 14.2 (Ljapunovova CLV.) Nech $X_1, X_2, ...$ sú nezávislé náhodné veličiny pre ktoré existujú konečné momenty $\mathcal{E}(X_k) = \mu_k,\ \mathcal{D}(X_k) = \sigma_k^2 > 0,\ \mathcal{E}|X_k - \mu_k|^3 = H_k^3,\ \ k = 1, 2, ...$ Položme $S_n = \sqrt{\sum_{k=1}^n \sigma_k^2},\ K_n = \sqrt{\sum_{k=1}^n H_k^3}$. Potom Ljapunovova podmienka $\lim_{n\to\infty} \dfrac{K_n}{S_n} = 0$ je postačujúca k tomu, aby pre každé $x \in \mathtt{R}$

$$\lim_{n\to\infty} P(U_{\overline{X}_n} < x) = \lim_{n\to\infty} P\left(\frac{\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}} < x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

Dôkaz nájdeme napr. v knihe Rényi, A., Teorie pravděpodobnosti, ACADEMIA, Praha, 1972.

Poznámka. Existuje veľa modifikácií CLV. Mnohé nájdeme v knihe Rényi, A., Teorie pravděpodobnosti, ACADEMIA, Praha, 1972.

Veta 14.3 (Moivreova-Laplaceova integrálna veta.) Nech $p \in (0, 1)$ a $Z_1, Z_2, ...$ sú náhodné veličiny s binomickým rozdelením, teda $Z_n \sim Bi(n, p)$. Potom platí pre každé $x \in \mathtt{R}$

$$\lim_{n\to\infty} P\left(\frac{Z_n - np}{\sqrt{np(1-p)}} < x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

Dôkaz: Veta je špeciálnym prípadom Vety 14.1 (Lindebergova CLV) ak $X_i,\ i = 1, 2, ...,$ sú nezávislé, $X_i \sim A(p)$ ($A(p)$ je alternatívne rozdelenie s parametrom p). Potom $\mathcal{E}(X_i) = \mu$ a $\mathcal{D}(X_i) = p(1-p) = \sigma^2$. Platí $\sum_{j=1}^n X_i = Z_n \sim Bi(n, p)$ a $U_{\overline{X}_n} = \dfrac{1}{\sigma\sqrt{n}}(X_1 + ... + X_n - n\mu) = \dfrac{Z_n - np}{\sqrt{np(1-p)}}$ konverguje v distribúcii k náhodnej veličine s $N(0, 1)$ rozdelením. ♣

Poznámka. Veta sa dá sformulovať aj nasledovne:

Pre $p \in (0, 1),\ -\infty \le a < b \le \infty$ nech $Z_1, Z_2, ...$ sú náhodné veličiny s binomickým rozdelením, teda $Z_n \sim Bi(n, p)$. Potom platí

$$\lim_{n\to\infty} P\left(a \le \frac{Z_n - np}{\sqrt{np(1-p)}} < b\right) = \Phi(b) - \Phi(a),$$

kde $\Phi(\cdot)$ je distribučná funkcia normovaného normálneho rozdelenia.

Príklad 14.1. Nájdite približnú hodnotu pravdepodobnosti toho, že počet šestiek, ktoré padnú v 12000 hodoch homogénnou kockou bude medzi 1900 a 2150.

Riešenie: ťažko by sme spočítali $\sum_{i=1900}^{2150} \binom{12000}{i} (\frac{1}{6})^i (\frac{5}{6})^{12000-i}$. Pretože $n = 12000,\ Z_n \sim Bi(12000, \frac{1}{6}), \mathcal{E}(Z_n) = np = \frac{12000}{6} = 2000,\ \mathcal{D}(Z_n) = np(1-p) = 2000\frac{5}{6} = \frac{10000}{6}$, dostávame

$$P(1900 \le Z_n < 2150) = P(1900 - \mathcal{E}(Z_n) \le Z_n - \mathcal{E}(Z_n) < 2150 - \mathcal{E}(Z_n)) =$$

$$= P\left(\frac{1900 - \mathcal{E}(Z_n)}{\sqrt{\mathcal{D}(Z_n)}} \le \frac{Z_n - \mathcal{E}(Z_n)}{\sqrt{\mathcal{D}(Z_n)}} < \frac{2150 - \mathcal{E}(Z_n)}{\sqrt{\mathcal{D}(Z_n)}}\right) =$$

$$= P\left(\frac{1900 - 2000}{\sqrt{\frac{10000}{6}}} \le \frac{Z_n - np}{\sqrt{np(1-p)}} < \frac{2150 - 2000}{\sqrt{\frac{10000}{6}}}\right) =$$

$$= P(-2.45 \le \frac{Z_n - np}{\sqrt{np(1-p)}} < 3.67) \doteq \Phi(3.67) - \Phi(-2.45) =$$

$$= 0.9998 - 0.0071 = 0.9927$$

(lebo $\Phi(-u) = 1 - \Phi(u)$, kde $\Phi(\cdot)$ je distribučná funkcia $N(0, 1)$ rozdelenia).

15. Popisná štatistika
(podľa Zvára, K., Štěpán, J. Pravděpodobnost a matematická štatistika, Matfyzpress, Praha, 2001)

Štatistika skúma javy na rozsiahlom súbore prípadov a zaujímajú ju tie vlastnosti javov, ktoré sa prejavujú vo veľkom súbore prípadov, nie v jednotlivých prípadoch. Základný pojem je štatistický súbor (základný súbor). Je to dobre definovaná (určená) množina štatistických jednotiek. Štatistický súbor môže byť určený zoznamom svojich prvkov (jednotiek), alebo pomocou nejakého pravidla, predpisu. V prípade pochybností sa dá overiť, či skúmaná jednotka patrí do štatistického súboru alebo nie. Na štatistických jednotkách sa meria (určuje, pozoruje) jeden alebo viac štatistických znakov. Znaky podľa typov delíme na

Nominálne znaky, ktorých hodnoty sú disjungtné kategórie. Medzi hodnotami nie je žiaden vzťah, usporiadanie. Napríklad farba očí, politická príslušnosť, atď.

Ordinálne znaky sú vlastne nominálne znaky, ale ich hodnoty sa dajú usporiadať. Napríklad najvyššie dosiahnuté vzdelanie, hodnosť u vojska, počet hviezdičiek v hotelovej kategórii, atď. Poznáme len poradie hodnoty znaku, neexistuje "vzdialenosť" medzi hodnotami.

Intervalové znaky nadobúdajú číselné hodnoty. Sú teda usporiadané, ale poznáme u nich aj (prirodzenú) vzdialenosť medzi hodnotami. Sú charakteristické tým, že nula je u nich len dohodnutá (napr. teplotné stupnice).

Pomerové znaky, ktorých hodnoty sa vzťahujú na nejakú dohodnutú jednotku. Hodnoty znaku udávajú násobok dohodnutej jednotky. Nula znamená neexistenciu meranej vlastnosti. Sem patrí napr. väčšina fyzikálnych veličín.

Štatistické znaky nominálne, či ordinálne sa nazývajú kvalitatívne, intervalové či pomerové znaky sa nazývajú kvantitatívne (niekedy kardinálne).

Kvantitatívne znaky delíme na diskrétne a spojité.

Predpokladajme, že sme na n štatistických jednotkách namerali súbor hodnôt $x_1, x_2, ..., x_n$ daného znaku. Celkovému počtu prvkov súboru hovoríme rozsah súboru.

Ako spracovávame, zhrnieme, oznamujeme hodnoty súboru ?

Ak jednotlivé hodnoty (ordinálneho resp. kvantitatívnho) znaku usporiadame do neklesajúcej postupnosti $x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n)}$, dostaneme usporiadaný súbor hodnôt. Indexy v dolných zátvorkách udávajú poradie jednotlivých zistených hodnôt znaku. Najmenšia je $x_{(1)}$, najväčšia je $x_{(n)}$. Keď je súbor veľký a hodnoty sa často opakujú, prehľadnejšie ich zapíšeme do tabuľky početností, v ktorej $a_1 < a_2 < ... < a_m$ sú navzájom rôzne usporiadané hodnoty znaku v súbore (v prípade nominálneho znaku len rôzne hodnoty) a $n_1, n_2, ..., n_m$ sú zistené (absolútne) početnosti týchto hodnôt (t.j. n_i-krát bola nameraná v súbore hodnota znaku a_i). Zrejme $\sum_{i=1}^m n_i = n$. Takýmto spôsobom sa typicky spracovávajú kvalitatívne znaky a diskrétne znaky. V prípade kvantitatívneho spojitého znaku postupujeme nasledovne. Keď meraný znak nadobúda príliš veľa rôznych číselných hodnôt, umelo zmenšíme počet rozlišovaných hodnôt tak, že obor všetkých hodnôt rozdelíme na disjunktné intervaly. Zvolíme napr. hraničné body $(-\infty \leq) t_0 < t_1 < ... < t_m (\leq \infty)$ a všetky hodnoty znaku z $j-$teho intervalu $(t_{j-1}, t_j >$ (niekedy je vhodnejšie $< t_{j-1}, t_j)$) stotožníme so stredom tohto intervalu $a_j = \dfrac{t_{j-1} + t_j}{2}$. Ak je $t_0 = -\infty$, spravidla zvolíme $a_1 = t_1 - \dfrac{t_2 - t_1}{2}$, takže t_1 je v strede intervalu (a_1, a_2). Podobne pre $t_m = \infty$ spravidla volíme $a_m = t_{m-1} + \dfrac{t_{m-1} - t_{m-2}}{2}$. Najčastejšie sa volia $t_1, t_2, ..., t_{m-1}$ tak, aby intervaly boli rovnako dlhé (až na krajné). Teda $t_j - t_{j-1} = h, \quad j = 2, 3, ..., m-1$.

Teraz určíme počty n_j hodnôt x_i, ktoré patria do jednotlivých intervalov (tried), tzv. triedne početnosti. Potom napíšeme tabuľku početností. Tabuľku početností znázorníme graficky pomocou polygónu početností, keď lomenou čiarou spojíme body o súradniciach (a_j, n_j), $j = 1, 2, ..., m$. častejšie znázorníme tabuľku početností histogramom, keď nad intervalmi $(a_j - \frac{h}{2}, a_j + \frac{h}{2} >$, $j = 1, 2, ..., m$ kreslíme obdĺžnik, ktorého výška je rovná n_j. Ak triedne intervaly nemajú rovnakú šírku, je n_j výška obdĺžnika nad zodpovedajúcim intervalom. Do uvedených grafov sa dá namiesto absolútnej početností n_j znázorniť aj relatívna početnosť $f_j = \frac{n_j}{n}$, prípadne sa dajú absolútne resp. relatívne početnosti sčítať (kumulovať) a použiť buď $\sum_{i=1}^{j} n_i$ alebo $\sum_{i=1}^{j} f_i$ (kumulatívne diagramy).

Príklad 15.1. V tabuľke 15.1 sú uvedené triedne početnosti priemerných známok na koncoročnom vysvedčení u 372 detí.

Zodpovedajúce histogramy (triednych početností a kumulatívnych triednych početností) sú (obr. 10.1 v Zvára, K., Štěpán, J. Pravděpodobnost a matematická štatistika)

Histogram triednych početností

Histogram kumulatívnych triednych početností

interval $< t_{j-1}, t_j)$	stred a_j	početnosť n_j	kumul. početnosť N_j
$< 1,0; 1,2)$	1,1	31	31
$< 1,2; 1,4)$	1,3	48	79
$< 1,4; 1,6)$	1,5	29	108
$< 1,6; 1,8)$	1,7	37	145
$< 1,8; 2,0)$	1,9	27	172
$< 2,0; 2,2)$	2,1	41	213
$< 2,2; 2,4)$	2,3	32	245
$< 2,4; 2,6)$	2,5	19	264
$< 2,6; 2,8)$	2,7	28	292
$< 2,8; 3,0)$	2,9	23	315
$< 3,0; 3,2)$	3,1	24	339
$< 3,2; 3,4)$	3,3	25	364
$< 3,4; 3,6)$	3,5	4	368
$< 3,6; 3,8)$	3,7	4	372

Tabuľka 15.1

Poznámka. Obyčajne používame triedne intervaly konštantnej šírky. Pri voľbe počtu intervalov môžeme vyjsť zo *Sturgesovho* pravidla, podľa ktorého

$$m \doteq 1 + 3,3 \log_{10} n \doteq 1 + 1,43 \ln n.$$

Tejto hodnoty sa pridržiavame "približne".

Miery polohy

Miery polohy udávajú hodnotu, okolo ktorej sa nachádzajú jednotlivé pozorovania (hodnoty znaku).

Priemer (tiež výberový, či empirický aritmetický priemer)

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{1}{n} \sum_{j=1}^{m} n_j a_j.$$

Priemer sa určuje u kvantitatívnych znakoch a rovnakým spôsobom závisí od každej hodnoty znaku. Zrejme pre ľubovoľné a, b je

$$\overline{a + bx} \ (= \frac{1}{n} \sum_{i=1}^{n} (a + bx_i)) = a + b\overline{x},$$

takže sa prirodzene mení so zmenou merítka.

Geometrický priemer

$$\overline{x}_G = \sqrt[n]{x_1 x_2 ... x_n}.$$

Geometrický priemer má zmysel len keď všetky hodnoty znaku sú kladné. Nie je invariantný voči lineárnej transformácii údajov. Používa sa v prípade, že ide o násobenie. Častejšie sa používa v ekonómii. Napr. ak je inflácia $20\%, 50\%, 30\%, 20\%$ a 5% (v jednotlivých rokoch), tak je to to isté, ako keby bola inflácia každý rok 24%, lebo výsledná inflácia je $1,2.1,5.1,3.1,2.1,05 = 2,9484$ a to je to isté ako keby v každom roku bola $\sqrt[5]{1,2.1,5.1,3.1,2.1,05} = 1,24$.

Harmonický priemer

$$\overline{x}_H = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}} = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}}.$$

Tiež nie je invariantný voči lineárnej transformácii. Dá sa ukazať, že ak sú všetky hodnoty znaku kladné, tak platí

$$\overline{x}_H \le \overline{x}_G \le \overline{x}$$

(pozri napr. Anděl, J., Statistické metódy, Matfyzpress, Praha, 1993).

Medián (rozumie sa výberový medián) je definovaný pomocou usporiadaného súboru hodnôt $x_{(1)} \le x_{(2)} \le ... \le x_{(n)}$ ako

$$\tilde{x} = \begin{cases} x_{\left(\frac{n+1}{2}\right)}, & \text{ak } n \text{ je nepárne (liché)} \\ \frac{1}{2}\left(x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}\right) & \text{ak } n \text{ je párne (sudé).} \end{cases}$$

Je to taká hodnota, ktorá delí usporiadané hodnoty $x_{(1)}, x_{(2)}, ..., x_{(n)}$ na dva rovnako početné diely. Preto nezáleží, aké veľké (malé) sú prvé resp. posledné členy usporiadaného súboru $x_{(1)}, x_{(2)}, ..., x_{(n)}$. Platí

$$\widetilde{(a+bx)} = a + b\tilde{x}.$$

Ak $g(\cdot)$ je monotónna funkcia, potom analogická vlastnosť platí pre transformované hodnoty. Ak je počet hodnôt nepárny (lichý), platí táto vlastnosť presne, ak je počet hodnôt párny (sudý), platí "skoro presne" ($g(\tilde{x})$ nie je vo všeobecnosti v tomto prípade priemerom hodnôt $g(x_{\left(\frac{n}{2}\right)}), g(x_{\left(\frac{n+2}{2}\right)})$. Pre nepárny (lichý) počet meraní má medián zmysel už pri ordinálnom znaku, pri párnom (sudom) počte meraní potrebujeme kvantitatívny znak. Keby sme pre párny (sudý) počet meraní definovali medián ako ľubovoľné číslo, pre ktoré platí $x_{\left(\frac{n}{2}\right)} \le \tilde{x} \le x_{\left(\frac{n}{2}+1\right)}$, nebol by síce definovaný jednoznačne, ale existoval by aj pre ordinálny znak (s číselnými hodnotami).

Medián môžene zovšeobecniť. Namiesto toho, aby oddeľoval polovicu najmenších údajov od ostatných, môže oddeľovať $p-$ty diel údajov. Zvolíme p, $0 < p < 1$. Definujeme $p-$ty výberový kvantil (percentil) vzťahom

$$x_p = \begin{cases} x_{([np]+1)}, & \text{ak } np \ne [np] \\ \frac{1}{2}\left(x_{(np)} + x_{(np+1)}\right) & \text{ak } np = [np], \end{cases}$$

kde $[np]$ je celá časť np, t.j. najväčšie celé číslo nie väčšie ako np. Napr. pre $p = 0,12$, $n = 24$ je $[np] = [2,88] = 2$, teda $x_{0,12} = x_{(3)}$ a pre $p = 0,4$, $n = 50$ je $[np] = [20] = 20$, teda $x_{0,4} = \frac{1}{2}\left(x_{(20)} + x_{(21)}\right)$. U ordinálneho znaku (číselného) ak $np = [np]$, môžeme ako x_p použiť ľubovoľnú hodnotu, ktorá leží medzi $x_{([np])}$ a $x_{([np]+1)}$. Medián je špeciálny prípad výberového kvantilu, a síce $\tilde{x} = x_{0,5}$.

V grafických zobrazeniach sa používajú dolný kvartil a horný kvartil

$$Q_1 = x_{0,25}, \quad Q_3 = x_{0,75}.$$

Módus je najčastejšou hodnotou. Má zmysel najmä vtedy, ak je počet m skutočne sa vyskytujúcich rôznych hodnôt podstatne menší ako rozsah n súboru. Módus je použiteľný pre každý typ znaku (aj keď v prípade nominálneho znaku je ťažko hovoriť o miere polohy). Nemusí byť určený jednoznačne (bimodálne súbory).

Miery variability

Miery variability charakterizujú veľkosť variability hodnôt znaku okolo nejakej "miery jej polohy", alebo "roztrúsenosť" hodnôt znaku. Miera variability by mala byť invariantná voči "posunutiu" všetkých hodnôt znaku, resp. voči lineárnej transformácii hodnôt znaku.

Rozptyl (empirický rozptyl)

$$s_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2,$$

resp. ak $a_1 < a_2 < ... < a_m$ sú rôzne hodnoty znaku, tak

$$s_x^2 = \frac{1}{n}\sum_{j=1}^{m} n_j(a_j - \overline{x})^2.$$

Platí

$$s_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i^2 - 2x_i\overline{x} + \overline{x}^2) = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - 2\frac{1}{n}\sum_{i=1}^{n} x_i\overline{x} + \frac{1}{n}\sum_{i=1}^{n}\overline{x}^2 =$$

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

Niekedy sa používa

$$s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

(neskôr budeme analyzovať prečo).

Keď sa používajú triedne početnosti, doporučuje sa *Sheppardova korekcia*, čo znamená zmenšiť výraz $s_x^2 = \frac{1}{n}\sum_{j=1}^{m} n_j(a_j - \overline{x})^2$ o hodnotu $\frac{h^2}{12}$, kde h je šírka rovnako širokých triednych intervalov.

Smerodajná odchýlka (empirická smerodajná odchýlka)

$$s_x = \sqrt{s_x^2}.$$

Jej dôležitá vlastnosť je, že je vyjadrená v rovnakých jednotkách ako namerané údaje. Rozptyl aj smerodajná odchýlka záležia na všetkých údajoch (sú citlivé na hodnoty najmä "krajných" údajov).

Rozpätie je rozdiel maximálnej a minimálnej hodnoty

$$R = x_{(n)} - x_{(1)}.$$

Rozpätie záleží len na veľkosti maximálnej a minimálnej hodnoty.

Kvartilové rozpätie

$$R_Q = Q_3 - Q_1 = x_{0{,}75} - x_{0{,}25}$$

Kvartilová odchýlka je polovica kvartilového rozpätia

$$\frac{Q_3 - Q_1}{2} = \frac{x_{0{,}75} - x_{0{,}25}}{2}.$$

Priemerná odchýlka je

$$d = \frac{1}{n}\sum_{i=1}^{n}|x_i - \tilde{x}|$$

(niekedy sa namiesto mediánu $\tilde{x}$ použije priemer $\overline{x}$).
Všetky uvedené miery variability predpokladajú kvantitatívny znak. Pre znaky nominálne (aj ordinálne) sa variabilita dá charakterizovať pomocou entropie

$$H = -\sum_{j=1}^{m}\frac{n_j}{n}\log\frac{n_j}{n},$$

pričom predpokladáme m rôznych hodnôt znaku s nenulovými početnosťami
$n_1, n_2, ..., n_m$.

Miery šikmosti a špicatosti

Výberový koeficient šikmosti (skewness)

$$g_1 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^3}{ns^3},$$

ktorý môže byť kladný aj záporný (Kladná (pravá) šikmosť je ak hustota je koncentrovaná v "ľavej" časti grafu a "pomaly dlho graf klesá). Kvantilový koeficient šikmosti

$$\frac{(x_{1-p} - \tilde{x}) - (\tilde{x} - x_p)}{x_{1-p} - x_p} = \frac{x_{1-p} - 2\tilde{x} + x_p}{x_{1-p} - x_p}$$

pre $0 < p < 0,5$. špeciálne pre $p = 0,25$ to je kvartilový koeficient šikmosti

$$\frac{(Q_3 - \tilde{x}) - (\tilde{x} - Q_1)}{Q_3 - Q_1}.$$

Výberový koeficient špicatosti (excess)

$$g_2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^4}{ns^4} - 3$$

a kvantilový koeficient špicatosti

$$\frac{x_{(n)} - x_{(1)}}{x_{1-p} - x_p}$$

pre $0 < p < 0,5$.

Diagramy

Veľmi názorné a obľúbené sú vedľa histogramov aj iné grafické znázornenia nameraných údajov a ich vlastností. Zaraďujeme ich medzi exploračné (výskumné) štatistické metódy (EDA - Exploratory Data Analysis).

Krabicový (fúzatý) diagram (box plot, box and whisker plot). Má mnohé modifikácie, napr. (obr. 10.2, Zvára, K., štěpán, J. Pravděpodobnost a matematická štatistika) Na krabicovom diagrame sú $Q_1, Q_3, \tilde{x}, R_Q$, tykadlá (fúzy) siahajú k takému najvzdialenejšiemu (od odpovedajúceho kvartila) pozorovaniu, ktoré nie je od neho vzdialené viac ako 1,5 násobok kvartilového rozpätia. Jednotlivo sú znázorňované pozorovania, ktoré sú viac vzdialené. U niektorých programov siahajú "fúzy" k najmenšiemu resp. k najväčšiemu pozorovaniu. Inokedy k výberovému 10% resp. 90% kvantilu.

Príklad 15.2. (Jednoduchý prípad.) Namerali sa údaje 21,24,24,25,25,25,25,25, 26,26,27,27, teda $n = 12$. ľahko vidíme, že $\tilde{x} = 25$, $Q_1 = x_{0,25} = \frac{1}{2}[x_{(3)} + x_{(4)}] = \frac{1}{2}(24 + 25) = 24,5$, $Q_3 = x_{0,75} = \frac{1}{2}[x_{(9)} + x_{(10)}] = 26$. $R_Q = Q_3 - Q_1 = 26 - 24,5 = 1,5$. Najmenšie pozorovanie je odľahlé, lebo $21 < 24,5 - 1,5.1,5 = 22,25$. Krabicový diagram je uvedený vyššie.

Príklad 15.3. Zisťovali sa hmotnosti detí v 12. mesiaci ich veku. Histogramy početností dievčat a chlapcov sú (obr. 10.3, Zvára, K., Štěpán, J. Pravděpodobnost a matematická štatistika)

Oba histogramy ukazujú na kladnú šikmosť. Vidieť, že hmotnosti chalpcov sú v priemere väčšie ako dievčat. Odpovedajúce krabicové diagramy sú (obr. 10.4, Zvára, K., Štěpán, J. Pravděpodobnost a matematická štatistika)

Keď chceme vyjadriť závislosť (súvislosť) dvoch kvantitatívnych znakov s nameranými hodnotami $(x_1, y_1), ..., (x_n, y_n)$, použijeme rozptylový diagram, na ktorom sú znázornené body $[x_i, y_i]$. Tri rôzne typy závislostí sú na nasledujúcich obrázkoch (obr. 10.5, Zvára, K., Štěpán, J. Pravděpodobnost a matematická štatistika)

Ak sledujeme súčasné správanie sa niekoľkých znakov, užitočný sa ukazuje maticový diagram, v ktorom sú znázornené súčasne histogramy pre jednotlivé znaky a (vzájomné) rozptylové diagramy (obr. 10.6, Zvára, K., Štěpán, J. Pravděpodobnost a matematická štatistika)

16. Náhodný výber

Štatistike sa niekedy hovorí, že je *metodologická náuka*, ktorá *objektivizuje proces poznania*. Skúsme si popísať, ako sa to dosahuje.

Základný súbor (štatistický súbor) voláme tiež populácia. Predpokladáme, ze má N jednotiek ($N \gg 0$). Principiálne môžeme zmerať hodnotu kvantitatívneho znaku Z na každej jednotke a dostať hodnoty $z_1, z_2, ..., z_N$. Priemer hodnôt z celej populácie označme μ a budeme ho nazývať populačný priemer. Označme σ^2 populačný rozptyl, teda

$$\mu = \frac{1}{N}\sum_{i=1}^{N} z_i, \quad \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(z_i-\mu)^2.$$

Pretože N je veľmi veľké, nie je možné (resp. je veľmi nehospodárne) zmerať hodnotu znaku na každej jednotke. Preto vyberieme skupinu n jednotiek a zistíme hodnotu znaku len na týchto jednotkách. Tento výber (výberový súbor) musí byť taký, aby dobre reprezentoval celú populáciu (celý základný súbor). Budeme vždy predpokladať, že $n < N$.

Jeden zo spôsobov dosiahnuť "dobre reprezentujúci" výberový súbor je urobiť náhodný výber bez vrátenia (prostý náhodný výber). To znamená, že vyberieme jeden z N prvkov základného súboru, potom náhodne vyberieme jeden z $N-1$ zostávajúcich, atď., až jeden z $N-n+1$ zostávajúcich. Výberový súbor môžeme vybrať $\binom{N}{n}$ spôsobmi. Keď budeme prvky výberového súboru vyberať náhodne, tak dosiahneme požadovanú reprezentatívnosť a každá $n-$tica bude mať rovnakú pravdepodobnosť, že bude vybraná..

My sa sústredíme na tzv. výber s vrátením z konečnej populácie. Náhodne vyberieme z populácie nejaký prvok (nejakú štatistickú jednotku), zistíme hodnotu meraného znaku a vrátime ho späť.

Označme $X -$ hodnotu znaku na náhodne vybranej štatistickej jednotke. Zrejme X je náhodná veličina, ktorá nadobúda hodnoty $b_1 < b_2 < ... < b_m$ s pravdepodobnosťami $P\{X = b_j\} = \frac{n_j}{N}$, $j = 1, 2, ..., m$, kde n_j je počet tých štatistických jednotiek v základnom súbore, na ktorých je hodnota znaku rovná b_j. Platí

$$\mathcal{E}(X) = \sum_{j=1}^{m} b_j P\{X = b_j\} = \frac{1}{N}\sum_{j=1}^{m} n_j b_j = \mu$$

a

$$\mathcal{D}(X) = \mathcal{E}(X-\mu)^2 = \sum_{j=1}^{m}(b_j-\mu)^2 P\{X = b_j\} = \frac{1}{N}\sum_{j=1}^{m} n_j(b_j-\mu)^2 = \sigma^2.$$

Keď nezávisle vyberáme $n-$ticu štatistických jednotiek (po náhodnom vybratí štatistickú jednotku vždy vrátime späť do súboru), tak tento výber modelujeme $n-$ ticou $(X_1, ..., X_n)$ náhodných veličín, pričom sú (združene) nezávislé a rovnako rozdelené (ako náhodná veličina X).

Teraz sa pozrime trochu ináč na výber s vrátením. Nemusíme sa starať o hodnoty $z_1, z_2, ..., z_N$ (resp. $b_1, ..., b_m$) znaku, ale stačí nám vedieť, aký je pre dané (ale ľubovoľné) $z \in \mathtt{R}$ pomer $p(z)$ množstva tých štatistických jednotiek, na ktorých hodnota znaku je menšia ako z k celkovému počtu jednotiek, teda k N. čiže

$$p(z) = \frac{\{\text{počet tých } z_i, \text{ ktoré sú menšie ako } z\}}{N}.$$

Vyberme náhodne jednu štatistickú jednotku (zrealizujme náhodnú veličinu X) a označme nameranú hodnotu x. Teda ak hodnota znaku na tejto jednotke je x, je to (konkrétna) realizácia náhodnej veličiny X. Platí

$$F_X(x) = P(X < x) = p(x).$$

Preto X má distribučnú funkciu $F_X(\cdot) = p(\cdot)$. Už nás nezaujíma, či populácia je konečná, alebo nekonečná. V reálnom živote je vždy konečná, ale keď je N veľmi veľké, považujeme ju za nekonečnú. V takomto "veľkom" základnom súbore aj keď realizujeme výber bez vrátenia, môžeme považovať vybrané hodnoty za realizácie nezávislých náhodných veličín. (Intuitívne to znamená, že pri "nekonečne" veľkom základnom subore odobratie niekoľkých jednotiek prakticky nezmení funkciu $p(x)$.)

V prípade náhodného výberu s vrátením (z konečnej alebo "nekonečnej" populácie) alebo náhodného výberu bez vrátenia z "nekonečnej" populácie je výsledkom pokusu $n-$tica nezávislých náhodných veličín $X_1, X_2, ..., X_n$ rovnako rozdelených, ktoré majú (rovnakú) distribučnú funkciu $F_X(\cdot)$. Takáto $n-$tica náhodných veličín sa nazýva náhodný výber rozsahu n (n nezávislých kópií náhodnej veličiny X).

Predpokladajme, že nahodná veličina X má konečnú strednú hodnotu μ a disperziu σ^2. V prípade konečnej populácie sa táto stredná hodnota rovná populačnému priemeru a disperzia rovná populačnému rozptylu. Náhodnú veličinu

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

nazývame výberový priemer a náhodnú veličinu

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

výberový rozptyl.

Aké vlastnosti majú výberový priemer a výberový rozptyl ?

Veta 16.1. Pre výberový priemer platí

$$\mathcal{E}(\overline{X}) = \mu,$$

$$\mathcal{D}(\overline{X}) = \frac{\sigma^2}{n}.$$

Dôkaz:

$$\mathcal{E}(\overline{X}) = \mathcal{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mathcal{E}(X_i) = \frac{1}{n}\sum_{i=1}^{n}\mu = \mu,$$

$$\mathcal{D}(\overline{X}) = \mathcal{D}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathcal{D}(X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{\sigma^2}{n}. \quad \clubsuit$$

Veta 16.2. Pre náhodný výber rozsahu n z rozdelenia s konečným rozptylom σ^2 platí

$$\mathcal{E}(S^2) = \sigma^2.$$

Dôkaz: Platí

$$\mathcal{E}(S^2) = \mathcal{E}\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \mathcal{E}\left(\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \mu + \mu - \overline{X})^2\right) =$$

$$= \frac{1}{n-1}\left(\mathcal{E}\left[\sum_{i=1}^{n}(X_i-\mu)^2 + 2\sum_{i=1}^{n}(X_i-\mu)(\mu-\overline{X}) + n\sum_{i=1}^{n}(\mu-\overline{X})^2\right]\right) =$$

$$= \frac{1}{n-1}\left(\sum_{i=1}^{n}\sigma^2 + 2\mathcal{E}\left[(\mu-\overline{X})\sum_{i=1}^{n}(X_i-\mu)\right] + \mathcal{E}\sum_{i=1}^{n}(\overline{X}-\mu)^2\right) =$$

$$= \frac{1}{n-1}n\sigma^2 + \frac{1}{n-1}2\mathcal{E}\left[n(\mu-\overline{X})\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)\right] + \frac{1}{n-1}\mathcal{E}\sum_{i=1}^{n}(\overline{X}-\mu)^2 =$$
$$= \frac{n}{n-1}\sigma^2 - 2\frac{n}{n-1}\mathcal{E}(\overline{X}-\mu)^2 + \frac{n}{n-1}\mathcal{E}(\overline{X}-\mu)^2 = \frac{n}{n-1}\sigma^2 - \frac{n}{n-1}\mathcal{D}(\overline{X}) =$$
$$= \frac{n}{n-1}\sigma^2 - \frac{n}{n-1}\frac{\sigma^2}{n} = \sigma^2. \quad \clubsuit$$

Ak je základný súbor rozsiahly, niekedy ho rozdelíme na L "neprekrývajúcich" sa častí, ktoré nazývame <u>oblasti</u>. Z každej oblasti vykonáme prostý náhodný výber (bez vrátenia). Každú oblasť považujeme za "menší" základný súbor. <u>Oblastné usporiadanie výberu (oblastný výber)</u> je motivované napr. tým, že celý základný súbor pozostáva z "prirodzených" podsúborov, že zber dát v určitých podoblastiach je špecifický (finančne, časovo), atď. Oblasti môžu byť aj "umelo vytvorené". Ak sú rozsahy oblastí $N_1, N_2, ..., N_L$ a oblastné výberové súbory majú rozsahy $n_1, n_2, ..., n_L$, potom celý základný súbor má rozsah $N = N_1 + ... + N_L$ a celý výberový súbor má rozsah $n = n_1 + ... + n_L$. Ak

$$\frac{n_1}{N_1} = \frac{n_2}{N_2} = ... = \frac{n_L}{N_L}(= k),$$

tak hovoríme, že oblastný výber je <u>rovnomerný</u>. V takomto prípade má každá jednotka rovnakú pravdepodobnosť $\frac{n}{N}$ zahrnutia do výberu (nezávisle od toho, do ktorej oblati patrí).

Ak sa základný súbor skladá z veľmi veľkého množstva jednotiek (roztrúsených), ťažko uskutočníme aj oblastný výber. Vzniká potreba vyberať jednotky vždy po celých <u>skupinách</u>. Skupiny môžeme považovať za nové jednotky vzniknuté zlučovaním pôvodných jednotiek. Môžu to byť malé skupiny (napr. rodiny), alebo aj veľmi veľké (okresy, školy, závody v podniku). Tento spôsob výberu nazývame <u>dvojstupňový výber</u>. Najprv vyberieme skupinky, z ktorých potom vyberáme "prvotné" jednotky. Výber skupiniek nazývame <u>prvým výberovým stupňom</u>. Výber prvotných jednotiek nazývame <u>druhým výberovým stupňom</u>. Nech oblasti obsahujú po rade $M_1, M_2, ..., M_L$ skupiniek, z ktorých v prvom výberovom stupni vyberieme postupne $m_1, m_2, ..., m_L$ skupiniek. V druhom výberovom stupni z každej výberovej skupinky v $h-$tej oblasti (ak táto skupinka bola v prvom stupni vybraná) vyberieme $100\pi_h$ percent prvotných jednotiek. Ako zvoliť čísla $m_1, ..., m_L, \pi_1, ..., \pi_L$, aby každá prvotná jednotka mala rovnakú pravdepodobnosť dostať sa do výberového súboru, bez ohľadu na to, do ktorej oblasti resp. skupinky patrí.

Označme náhodný javy

$A-$ štatistická jednotka J z $h-$tej oblasti bola vybratá do výberového súboru v druhom výberovom stupni

$B-$ skupina, do ktorej patrí štatistická jednotka J, bola vybratá v $h-$tej oblasti v prvom výberovom stupni

Je zrejmé, že štatistická jednotka J bola vybratá do výberového súboru práve vtedy, ak nastal náhodný jav $A \cap B$. Platí $P(A \cap B) = P(A|B)P(B)$. Zrejme $P(B) = \frac{m_h}{M_h}$ a $P(A|B) = \pi_h$. Preto pravdepodobnosti, že prvotná jednotka z $h-$tej oblasti ($h = 1, 2, ..., L$) sa dostane do výberového súboru sú postupne

$$\frac{m_1}{M_1}\pi_1, \frac{m_2}{M_2}\pi_2, ..., \frac{m_L}{M_L}\pi_L.$$

Výber bude rovnomerný, ak

$$\frac{m_1}{M_1}\pi_1 = \frac{m_2}{M_2}\pi_2 = ... = \frac{m_L}{M_L}\pi_L.$$

V tomto prípade každá jednotka v základnom súbore bude mať rovnakú pravdepodobnosť dostať sa do výberového súboru.

Príklad 16.1. Pri štatistickom šetrení týkajúcom sa zisťovania sociálnych pomerov v rodinách školákov do 15 rokov prvý stupeň záležal od výberu škôl a druhý od výberu žiakov vybranej školy. školy boli rozdelené na 3 druhy (oblasti), a síce, (i) päťročné školy, (ii) deväťročné školy, (iii) osemročné gymnázia. Z päťročných škôl bola vybratá každá stá škola, teda $\dfrac{m_1}{M_1} = \dfrac{1}{100}$ a z tejto školy boli zahrnuté do výberu všetci žiaci, teda $\pi_1 = 1$. Z deväťročných škôl bola vybratá každá päťdesiata a do výberu z nej vybratý každý druhý žiak. Z osemročných gymnázií bolo vybraté každé dvadsiate piate a z neho vybratý každý štvrtý žiak. Teda

$$\frac{m_1}{M_1} = \frac{1}{100},\ \pi_1 = 1,\quad \frac{m_1\pi_1}{M_1} = \frac{1}{100},$$

$$\frac{m_2}{M_2} = \frac{1}{50},\ \pi_2 = \frac{1}{2},\quad \frac{m_2\pi_2}{M_2} = \frac{1}{100},$$

$$\frac{m_3}{M_3} = \frac{1}{25},\ \pi_3 = \frac{1}{4},\quad \frac{m_3\pi_3}{M_3} = \frac{1}{100}.$$

Každý žiak bez ohľadu na druh školy mal pravdepodobnosť $\dfrac{1}{100}$ dostať sa do výberu.

V každom druhu výberu majú výberový priemer, výberový rozptyl a iné výberové charakteristiky špecifické (pravdepodobnostné, štatistické) vlastnosti. O tom pojednáva teória výberových šetrení.

17. Odhady parametrov

(hlavne podľa Anděl, J., Matematická statistika, SNTL/ALFA, Praha, 1985)

Predpokladajme, že náhodný vektor $\mathbf{X} = (X_1, X_2, ..., X_n)'$ má hustotu (v prípade diskrétneho náhodného vektora pravdepodobnostnú funkciu) $f(\mathbf{x}, \boldsymbol{\theta})$, kde $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_m)'$ je neznámy parameter. Na základe $\mathbf{X}$ je treba získať "čo možno najlepší" odhad tohto parametra. Vieme len toľko, že $\boldsymbol{\theta}$ sa nachádza v parametrickom priestore Ω (pozor, nie je to tentokrát priestor elementárnych javov).

Definícia 17.1. Bodový odhad parametra $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_m)'$ je merateľné zobrazenie $\mathbf{g} :\ (\mathtt{R}^n, \mathcal{B}_n) \rightarrow (\mathtt{R}^m, \mathcal{B}_m)$ (nezávisiace od $\boldsymbol{\theta}$) také, že $m-$rozmerný náhodný vektor $\mathbf{t} = \mathbf{g}(\mathbf{X})$ v nejakom "rozumnom zmysle" aproximuje neznámy vektor parametrov $\boldsymbol{\theta}$).

Poznámka. Obyčajne predpokladáme, že náhodný vektor $\mathbf{X} = (X_1, X_2, ..., X_n)'$ je náhodným výberom z rozdelenia s distribučnou funkciou $F(\cdot; \boldsymbol{\theta})$. Preto sa niekedy pre upresnenie povie, že odhad $\mathbf{T}$ je založený na náhodnom vektore $\mathbf{X}$.

Definícia 17.2. Intervalový odhad parametra $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_m)'$ je taká (náhodná) množina z $\mathcal{B}_m$, ktorá s "dostatočne veľkou" pravdepodobnosťou pokrýva $\boldsymbol{\theta}$.

Poznámka. Namiesto parametra $\boldsymbol{\theta}$ môžeme uvažovať aj odhad určitej (konkrétnej) parametrickej funkcie $\mathbf{h}(\boldsymbol{\theta})$.

Niekedy sa najprv vezmú nejaké merateľné funkcie $S_1(\mathbf{x}), ..., S_k(\mathbf{x})$, vytvorí sa náhodný vektor $\mathbf{S}(\mathbf{X}) = (S_1(\mathbf{X}), S_2(\mathbf{X}), ..., S_k(\mathbf{X}))'$ (pre $m \leq k \leq n$). Každý takýto náhodný vektor sa volá štatistika. Ak $k = m$, tak takáto štatistika je (bodovým) odhadom.

Definícia 17.3. Povieme, že odhad $\mathbf{T}$ parametra $\boldsymbol{\theta}$ je nestranný, ak platí

$$\forall \boldsymbol{\theta} \in \Omega \quad \mathcal{E}_\theta(\mathbf{T}) = \boldsymbol{\theta}.$$

Poznámka. Odhad $\mathbf{T}$ (ako predpis) nezávisí od $\boldsymbol{\theta}$, ale jeho rozdelenie pravdepodobnosti od $\boldsymbol{\theta}$ závisí. Preto sa v Definícii 17.3 píše $\mathcal{E}_\theta(\mathbf{T})$. Zdôrazňuje sa tým, že stredná hodnota odhadu $\mathbf{T}$ sa ráta za predpokladu, že hodnota parametra rozdelenia je rovná $\boldsymbol{\theta}$.

Niekedy nestranný odhad vôbec neexistuje, alebo existuje iný odhad ako nestranný, ktorý je z určitého hľadiska výhodnejší.

Príklad 17.1. Majme náhodný výber $X_1, X_2, ..., X_n$ z rozdelenia s distribučnou funkciou $F(\cdot)$ a konečnou strednou hodnotou μ a disperziou σ^2. Náhodná veličina

$$T(X_1, ..., X_n) = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

je podľa Vety 16.1 nestranným odhadom parametra μ. Podľa Vety 16.2 je

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

nestranným odhadom σ^2.

Iným kritériom pre odhad $T(X_1, ..., X_n)$ jednorozmerného parametra θ je veľkosť jeho <u>strednekvadratickej odchýlky</u>, teda

$$\mathcal{E}(T-\theta)^2.$$

Platí

$$\mathcal{E}(T-\theta)^2 = \mathcal{E}\left((T-\mathcal{E}(T)) + (\mathcal{E}(T)-\theta)\right)^2 =$$
$$= \mathcal{E}[(T-\mathcal{E}(T))^2] + 2\mathcal{E}[(T-\mathcal{E}(T))(\mathcal{E}(T)-\theta)] + \mathcal{E}[(\mathcal{E}(T)-\theta)^2] = \mathcal{D}(T) + (\mathcal{E}(T)-\theta)^2,$$

čo je rozumná charakteristika odhadu. Ak platí

$$\mathcal{E}(\mathbf{T}) = \boldsymbol{\theta} + \mathbf{b}(\boldsymbol{\theta}),$$

pričom (vektorová) funkcia $\mathbf{b}$ nie je identicky rovná $\mathbf{0}$ na množine Ω, tak odhad $\mathbf{T}$ je <u>vychýlený</u>. Vektoru $\mathbf{b}(\boldsymbol{\theta})$ sa hovorí vychýlenie odhadu T v bode $\boldsymbol{\theta}$.

Príklad 17.2. Nech X je diskrétna náhodná veličina s binomickým rozdelením pravdepodobnosti, teda $X \sim Bi(n,p)$, pričom n považujeme za známe. Pre funkciu $\phi(p) = \dfrac{1}{p}$ parametra p neexistuje nestranný odhad založený na náhodnej veličine X. Ukážte.

Riešenie: Sporom. Nech existuje odhad T, teda merateľná funkcia náhodnej veličiny X (kde $X \sim Bi(n,p)$), pre ktorú platí

$$\mathcal{E}_p(T(X)) = \frac{1}{p} \quad \forall p \in (0,1).$$

Teda platí

$$\forall p \in (0,1) \quad \mathcal{E}_p(T(X)) = \sum_{j=0}^{n} T(j)\binom{n}{j}p^j(1-p)^{n-j} =$$
$$= T(0)(1-p)^n + T(1)np(1-p)^{n-1} + ... + T(n)p^n(1-p)^0 = \frac{1}{p}.$$

Na ľavej strane predchádzajúcej rovnosti máme polynóm premennej p stupňa najviac n, tento nemôže byť rovný racionálnej lomenej funkcii $\dfrac{1}{p}$ pre všetky $p \in (0,1)$. Teda nestranný odhad parametrickej funkcie $\dfrac{1}{p}$ založený na náhodnej veličine $X \sim Bi(n,p)$ neexistuje.

Príklad 17.3. Majme náhodný výber $X_1, X_2, ..., X_n$ z rozdelenia $\mathrm{N}(\mu, \sigma^2)$, $n \geq 2, \sigma^2 > 0$. Pre výberový rozptyl $S^2 = \dfrac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$ platí $\mathcal{E}(S^2) = \sigma^2$ (pozri Vetu 16.2) a $\mathcal{D}(S^2) = \dfrac{2\sigma^4}{n-1}$ (dokážeme si v kapitole 19). Majme odhad parametra σ^2 (ktorý odhad je typu) $T(\boldsymbol{X}) = c\sum_{i=1}^n (X_i - \overline{X})^2$. Pre aké c má tento odhad minimálnu strednekvadratickú odchýlku ? Čomu sa táto odchýlka rovná ?

Riešenie: $T = (n-1)cS^2$, preto $\mathcal{E}(T) = (n-1)c\mathcal{E}(S^2) = (n-1)c\sigma^2$ a $\mathcal{D}(T) = (n-1)^2c^2\mathcal{D}(S^2) = 2\sigma^4c^2(n-1)$. Strednekvadratická odchýlka odhadu T je

$$\mathcal{E}(T-\sigma^2)^2 = \mathcal{D}(T) + (\mathcal{E}(T) - \sigma^2)^2 = 2\sigma^4c^2(n-1) + [(n-1)c\sigma^2 - \sigma^2]^2 =$$

$$= \sigma^4\{2c^2(n-1) + (n-1)^2c^2 - 2c(n-1) + 1\} = \sigma^4\{c^2(n^2-1) - 2c(n-1) + 1\}.$$

Vzhľadom na c to je kvadratická funkcia, ktorá má minimum (po derivácii) v bode $c = \dfrac{1}{n+1}$. Preto odhad $T(\boldsymbol{X}) = \dfrac{1}{n+1}\sum_{i=1}^n (X_i - \overline{X})^2$ má najmenšiu strednekvadratickú odchýlku zo všetkých odhadov typu $T(\boldsymbol{X}) = c\sum_{i=1}^n (X_i - \overline{X})^2$. Táto minimálna strednekvadratická odchýlka je

$$\sigma^4\left\{\frac{n^2-1}{(n+1)^2} - \frac{2(n-1)}{n+1} + 1\right\} = \frac{2\sigma^4}{n+1}.$$

Uvažujme teraz jednorozmerný parameter θ. Nech $X_1, X_2, ...$ sú nezávislé rovnako rozdelené náhodné veličiny definované na tom istom pravdepodobnostnom priestore $(\Omega^*, \mathcal{A}, P)$ s rozdelením pravdepodobnosti, ktoré má distribučnú funkciu $F(\cdot, \theta)$. Pre každé prirodzené n majme $T_n(X_1, X_2, ..., X_n)$ - odhad parametra θ.

Definícia 17.4. T_n je <u>konzistentným</u> odhadom θ, ak T_n konverguje podľa pravdepodobnosti k θ, t.j. $\forall \epsilon > 0 \;\; P\{\omega \in \Omega^* : \; |T_n(\omega) - \theta| > \epsilon\} \to 0$.

Veta 17.1 Nech pre každé prirodzené n je $\mathcal{E}(T_n^2) < \infty$. Ak

(i) $\mathcal{E}(T_n) \to \theta$ a

(ii) $\mathcal{D}(T_n) \to 0$,

tak T_n je konzistentným odhadom parametra θ.

Dôkaz: Využijeme dve nerovnosti, a síce

$\forall \gamma > 0 \; P\{|\xi - \mathcal{E}(\xi)| < \gamma\} \geq 1 - \frac{\mathcal{D}(\xi)}{\gamma^2}$ (Čebyševova nerovnosť) a

$|a+b| \leq |a| + |b|$ (nerovnosť platná pre všetky reálne čísla).

Preto

$$\{\omega : \; |T_n(\omega) - \theta| < \epsilon\} = \{\omega : \; |T_n(\omega) - \mathcal{E}(T_n) + \mathcal{E}(T_n) - \theta| < \epsilon\} \supset$$
$$\{\omega : \; |T_n(\omega) - \mathcal{E}(T_n)| < \frac{\epsilon}{2} \cap |\mathcal{E}(T_n) - \theta| < \frac{\epsilon}{2}\},$$

teda

$$P\{\omega : \; |T_n(\omega) - \theta| < \epsilon\} \geq P\{\omega : \; |T_n(\omega) - \mathcal{E}(T_n)| < \frac{\epsilon}{2} \cap |\mathcal{E}(T_n) - \theta| < \frac{\epsilon}{2}\}. \tag{17.1}$$

Platí, že ak A, B sú dva náhodné javy a $P(B) = 1 \;\Rightarrow P(A \cap B) = P(A)$ (lebo $A \cup B \supset B \Rightarrow 1 = P(B) \leq P(A \cup B) = 1$ a $0 = P(A \cup B) - P(B) = P(A) - P(A \cap B)$). Pretože $\mathcal{E}(T_n) \to \theta$ (s pravdepodobnosťou 1), teda

$$\forall \epsilon > 0 \;\; \exists n_{10} \;\; \forall n \geq n_{10} \quad P\{\omega : \; |\mathcal{E}(T_n) - \theta| < \frac{\epsilon}{2}\} = 1,$$

dostávame z (17.1)

$$\forall \epsilon > 0 \;\; \exists n_{10} \;\; \forall n \geq n_{10} \;\; P\{\omega : \; |T_n(\omega) - \theta| < \epsilon\} \geq P\{\omega : \; |T_n(\omega) - \mathcal{E}(T_n)| < \frac{\epsilon}{2}\}. \tag{17.2}$$

Pretože $\mathcal{D}(T_n) \to 0$ (s pravdepodobnosťou 1), dostávame, že

$$\forall \varepsilon > 0 \; \forall \vartheta > 0 \;\; \exists n_{20} \;\; \forall n > n_{20} \;\; \mathcal{D}(T_n) < \vartheta \frac{\varepsilon^2}{4}. \tag{17.3}$$

Z Čebyševovej nerovnosti zase platí pre každé n

$$\forall \epsilon > 0 \;\; P\{\omega : \; |T_n(\omega) - \mathcal{E}(T_n)| < \frac{\epsilon}{2}\} \geq 1 - \frac{\mathcal{D}(T_n)}{\frac{\epsilon^2}{4}}. \tag{17.4}$$

Zo vzťahov (17.3) a (17.4) dostávame, že

$$\forall \epsilon > 0 \;\; \forall \vartheta > 0 \;\; \exists n_{20} \;\; \text{že } \forall n > n_{20} \;\; P\{\omega : \; |T_n(\omega) - \mathcal{E}(T_n)| < \frac{\epsilon}{2}\} \geq 1 - \vartheta. \tag{17.5}$$

Zo vzťahov (17.2) a (17.5) dostávame, že
$\forall \epsilon \;\; \forall \vartheta > 0 \;\; \forall n > \max\{n_{10}, n_{20}\}$

$$P\{\omega : \; |T_n(\omega) - \theta| < \epsilon\} \geq P\{\omega : \; |T_n(\omega) - \mathcal{E}(T_n)| < \frac{\epsilon}{2}\} \geq 1 - \vartheta. \quad \clubsuit$$

Príklad 17.4. Nech $X_1, X_2, ...$ sú nezávislé náhodné veličiny, každá s rovnomerným rozdelením pravdepodobnosti na intervale $(0, \theta)$, $\theta > 0$ (neznáme). Náhodná veličina $X_{(n)} = \max\{X_1, X_2, ..., X_n\}$. Ukážme, že $X_{(n)}$ je konzistentný odhad parametra θ. Pritom $X_{(n)}$ nie je nestranný odhad parametra θ.

Riešenie:

Hustota rozdelenia pravdepodobnosti náhodnej veličiny X_i, $i \in \{1, 2, ..., n\}$ je

$$f_i(t) = \begin{cases} 0, & \text{ak } t \leq 0 \\ \frac{1}{\theta}, & \text{ak } t \in (0, \theta) \\ 0, & \text{ak } t \geq \theta. \end{cases}$$

Preto distribučná funkcia

$$F_{X_{(n)}}(x) = P\{X_{(n)} < x\} = P\{X_1 < x, ..., X_n < x\} =$$

$$= \prod_{i=1}^{n} P\{X_i < x\} = \begin{cases} 0, & \text{ak } x \leq 0 \\ \left(\int_0^x \frac{1}{\theta} dt\right)^n = \frac{x^n}{\theta^n}, & \text{ak } x \in (0, \theta) \\ 1, & \text{ak } x \geq \theta. \end{cases}$$

Hustota

$$f_{X_{(n)}}(x) = F'_{X_{(n)}}(x) = \begin{cases} 0, & \text{ak } x \leq 0 \\ \frac{nx^{n-1}}{\theta^n}, & \text{ak } x \in (0, \theta) \\ 0, & \text{ak } x \geq \theta \end{cases}$$

a

$$\mathcal{E}(X_{(n)}) = \int_0^\theta x \frac{nx^{n-1}}{\theta^n} dx = \frac{n}{\theta^n} \left[\frac{x^{n+1}}{n+1}\right]_0^\theta = \frac{n\theta}{n+1},$$

$$\mathcal{E}(X_{(n)}^2) = \int_0^\theta x^2 \frac{nx^{n-1}}{\theta^n} dx = \frac{n}{\theta^n} \left[\frac{x^{n+2}}{n+2}\right]_0^\theta = \frac{n\theta^2}{n+2}.$$

Dostávame, že

$$\mathcal{D}(X_{(n)}) = \mathcal{E}(X_{(n)}^2) - \mathcal{E}^2(X_{(n)}) = \frac{n\theta^2}{n+2} - \frac{n^2\theta^2}{(n+1)^2} = \frac{n\theta^2}{(n+1)^2(n+2)}.$$

Podľa Vety 17.1 je $X_{(n)}$ konzistentným odhadom θ ($\mathcal{E}(T_n) \to \theta$ a $\mathcal{D}(T_n) \to 0$). Ľahko v tomto prípade získame nestranný konzistentný odhad. Stačí zvoliť $T_n = \dfrac{n+1}{n} X_{(n)}$.

Teraz si zavedieme eficientný (výdatny) odhad. Majme jednorozmerný parameter θ a náhodný vektor $\mathbf{X} = (X_1, ..., X_n)'$ nech má hustotu $f(\mathbf{x}, \theta)$ (distribučnú funkciu $F(\mathbf{x}, \theta)$. Majme odhad $T = T(\mathbf{X})$ parametra θ. Aká je dolná hranica strednej kvadratickej chyby $\mathcal{E}(T - \theta)^2$? Kedy sa táto hranica dosiahne ?

Definícia 17.5. Systém hustôt $\{f(\mathbf{x}, \theta),\ \theta \in \Omega\}$ je regulárny, ak platí

a) Ω je neprázdna otvorená množina,

b) množina $M = \{\mathbf{x} :\ f(\mathbf{x}, \theta) > 0\}$ nezávisí od θ,

c) pre $\forall \theta \in \Omega$ a pre skoro všetky $\mathbf{x} \in M$ existuje konečná parciálna derivácia $f'(\mathbf{x}, \theta) = \dfrac{\partial f(\mathbf{x}, \theta)}{\partial \theta}$,

d) pre $\forall \theta \in \Omega$ platí $\int_M \dfrac{f'(\mathbf{x}, \theta)}{f(\mathbf{x}, \theta)} dF(\mathbf{x}, \theta) = 0$,

e) pre $\forall \theta \in \Omega$ je integrál (výraz) $J(\theta) = \int_M \left[\dfrac{f'(\mathbf{x}, \theta)}{f(\mathbf{x}, \theta)}\right]^2 dF(\mathbf{x}, \theta)$ konečný a kladný ($0 < J(\theta) < \infty$).

Veličinu $J(\theta)$ voláme Fisherova informácia o parametri θ (Fisherova miera informácie o parametri θ, ktorá (informácia) je obsiahnutá v danej regulárnej triede hustôt) .

Fisherovu informáciu môžeme chápať aj ako

$$J(\theta) = \mathcal{E}_\theta \left[\left(\frac{f'(\mathbf{X})}{f(\mathbf{X})} \right)^2 \right] = \mathcal{E}_\theta \left[\left(\frac{\partial \ln f(\mathbf{X})}{\partial \theta} \right)^2 \right],$$

lebo mimo množiny M môžeme definovať $\dfrac{f'}{f}$ ľubovoľne, teda aj ako 0 a za integračný obor vziať $\mathtt{R}^n$.

Príklad 17.5. Systém hustôt $\{f(x, \theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2}}, -\infty < x < \infty, \theta \in \Omega = \mathtt{R}\}$ (vzhľadom k Lebesguovej miere) je regulárny (ide o hustoty $N(\theta, 1)$).

Dokážte ako cvičenie.

Veta 17.2. (Raova-Cramerova) Nech T je taký odhad θ, že $\forall\ \theta \in \Omega$ je $\mathcal{E}_\theta(T^2) < \infty$. Nech $b(\theta) = \mathcal{E}_\theta(T) - \theta$ je vychýlenie (bias) odhadu T. Nech platí

(i) systém hustôt $\{f(\mathbf{x}, \theta),\ \theta \in \Omega\}$ je regulárny,

(ii) pre $\forall \theta \in \Omega$ existuje derivácia $b'(\theta)$,

(iii) $\forall \theta \in \Omega$ $\dfrac{\partial}{\partial \theta} \int_M T(\mathbf{x}) dF(\mathbf{x}, \theta) = \int_M T(\mathbf{x}) \dfrac{f'(\mathbf{x}, \theta)}{f(\mathbf{x}, \theta)} dF(\mathbf{x}, \theta)$.

Potom pre $\forall \theta \in \Omega$ platí

$$\mathcal{E}_\theta(T - \theta)^2 \geq \frac{[1 + b'(\theta)]^2}{J(\theta)}.$$

Dôkaz:

$$b(\theta) = \mathcal{E}_\theta(T) - \theta = \int_M T(\mathbf{x}) dF(\mathbf{x}, \theta) - \theta,$$

teda

$$b(\theta) + \theta = \int_M T(\mathbf{x}) dF(\mathbf{x}, \theta).$$

Podľa (iii)

$$\frac{\partial}{\partial \theta} \int_M T(\mathbf{x}) dF(\mathbf{x}, \theta) = \int_M T(\mathbf{x}) \frac{f'(\mathbf{x}, \theta)}{f(\mathbf{x}, \theta)} dF(\mathbf{x}, \theta) = b'(\theta) + 1.$$

Z podmienky d) regularity triedy hustôt $\{f(\mathbf{x},\theta),\ \theta\in\Omega\}$ platí

$$\int_M \theta \frac{f'(\mathbf{x},\theta)}{f(\mathbf{x},\theta)} dF(\mathbf{x},\theta)=0,$$

teda

$$\int_M (T(\mathbf{x})-\theta)\frac{f'(\mathbf{x},\theta)}{f(\mathbf{x},\theta)} dF(\mathbf{x},\theta)=b'(\theta)+1.$$

Podľa Schwarzovej nerovnosti

$$[b'(\theta)+1]^2\le \int_M (T(\mathbf{x})-\theta)^2 dF(\mathbf{x},\theta)\int_M \left[\frac{f'(\mathbf{x},\theta)}{f(\mathbf{x},\theta)}\right]^2 dF(\mathbf{x},\theta),$$

čiže

$$\frac{[1+b'(\theta)]^2}{J(\theta)}\le \mathcal{E}_\theta(T-\theta)^2. \quad \clubsuit$$

Kedy nastáva rovnosť $\dfrac{[1+b'(\theta)]^2}{J(\theta)}=\mathcal{E}_\theta(T-\theta)^2$? Vo Schwarzovej nerovnosti nastáva rovnosť práve vtedy ak

(α) $T(\mathbf{x})-\theta=0$ s.v. vzhľadom k Lebesgueovej-Stieltjesovej miere μ_F,

alebo ak

(β) $\exists K(\theta)$ nezávislá na $\mathbf{x}$, že $\dfrac{f'(\mathbf{x},\theta)}{f(\mathbf{x},\theta)}=K(\theta)[T(\mathbf{x})-\theta]$ s.v. vzhľadom k Lebesgueovej-Stieltjesovej miere μ_F.

V prípade (α) $T(\mathbf{x})-\theta=0$, teda $P\{T(\mathbf{X})=\theta\}=1$ je $T(\mathbf{X})$ diskrétna náhodná veličina ktorá nadobúda hodnotu θ s pravdepodobnosťou 1, čiže $\mathcal{E}(T)=\theta$ a $b(\theta)=0$. Samozrejme vtedy $\mathcal{E}(T-\theta)^2=\mathcal{D}(T)=0$. Toto nemôže byť, lebo v uvažovanom prípade (α) podľa dôkazu vety je $\mathcal{E}(T-\theta)^2=\dfrac{[1+b'(\theta)]^2}{J(\theta)}=\dfrac{1}{J(\theta)}>0$ ($J(\theta)>0$ v regulárnej triede hustôt).

Dostávame, že rovnosť $\dfrac{[1+b'(\theta)]^2}{J(\theta)}=\mathcal{E}_\theta(T-\theta)^2$ nastáva práve vtedy ak platí (β), čiže $\exists K(\theta)$ nezávislá na $\mathbf{x}$, že s.v. vzhľadom k Lebesgueovej-Stieltjesovej miere μ_F

$$\frac{f'(\mathbf{x},\theta)}{f(\mathbf{x},\theta)}=K(\theta)[T(\mathbf{x})-\theta],$$

čo je to isté ako

$$\frac{\partial \ln f(\mathbf{x},\theta)}{\partial\theta}=K(\theta)T(\mathbf{x})-K(\theta)\theta,$$

alebo

$$\int \frac{\partial \ln f(\mathbf{x},\theta)}{\partial\theta}d\theta=T(\mathbf{x})\int K(\theta)d\theta-\int K(\theta)\theta d\theta. \tag{17.6}$$

Ak označíme $\int K(\theta)d\theta=Q(\theta)$ a $\int K(\theta)\theta d\theta=R(\theta)$ tak (17.6) môžeme napísať ako

$$\ln f(\mathbf{x},\theta)=Q(\theta)T(\mathbf{x})-R(\theta)+H(\mathbf{x}).$$

Keď ešte označíme $C(\theta)=e^{-R(\theta)},\ u(\mathbf{x})=e^{H(\mathbf{x})}$, tak dostávame pre hustotu $f(\mathbf{x},\theta)$

$$f(\mathbf{x},\theta)=C(\theta)e^{Q(\theta)T(\mathbf{x})}u(\mathbf{x}).$$

Definícia 17.6. Nech parametrický priestor Ω je totožný s nejakou borelovskou množinou v $\mathtt{R}^m$. Ak hustota $f(\mathbf{x},\boldsymbol{\theta})$ náhodného vektora $\mathbf{X}$ má tvar

$$f(\mathbf{x},\boldsymbol{\theta})=C(\boldsymbol{\theta})e^{\sum_{j=1}^s Q_j(\boldsymbol{\theta})T_j(\mathbf{x})}u(\mathbf{x}),$$

kde $C(\boldsymbol{\theta})$, $Q_j(\boldsymbol{\theta})$ sú merateľné funkcie parametra $\boldsymbol{\theta}$ a $T_j(\mathbf{x})$, $u(\mathbf{x})$ sú merateľné funkcie premennej $\mathbf{x}$, tak povieme, že $f(\mathbf{x}, \boldsymbol{\theta})$ je hustota exponenciálneho typu.

Poznámka. V Raovej-Cramerovej vete nastáva rovnosť $\dfrac{[1+b'(\theta)]^2}{J(\theta)} = \mathcal{E}_\theta(T-\theta)^2$ práve vtedy ak hustota náhodného vektora $\mathbf{X}$ je exponenciálneho typu.

Definícia 17.7. Ak pre odhad T parametra θ platí, že spĺňa všetky predpoklady Raovej-Cramerovej vety, čiže $\forall\ \theta \in \Omega$ je $\mathcal{E}_\theta(T^2) < \infty$ a

(i) systém hustôt $\{f(\mathbf{x}, \theta),\ \theta \in \Omega\}$ je regulárny,

(ii) pre $\forall \theta \in \Omega$ existuje derivácia $b'(\theta)$,

(iii) $\forall \theta \in \Omega\ \dfrac{\partial}{\partial \theta} \int_M T(\mathbf{x}) dF(\mathbf{x}, \theta) = \int_M T(\mathbf{x}) \dfrac{f'(\mathbf{x}, \theta)}{f(\mathbf{x}, \theta)} dF(\mathbf{x}, \theta)$

($b(\theta) = \mathcal{E}_\theta(T) - \theta$ je vychýlenie (bias) odhadu T), potom tento odhad nazývame regulárny.

Dôsledok. Pre každý regulárny nestranný odhad T parametra θ platí

$$\mathcal{D}(T) \geq \frac{1}{J(\theta)}.$$

číslu $\dfrac{1}{J(\theta)}$ sa hovorí dolná Raova-Cramerova hranica pre disperziu regulárneho nestranného odhadu.

Definícia 17.8. Eficienciu (výdatnosť) e regulárneho nestranného odhadu T definujeme ako

$$e = \frac{1}{\mathcal{D}(T)J(\theta)}.$$

Eficiencia sa dá písať aj ako

$$e = \frac{1}{\mathcal{D}(T)J(\theta)} = \frac{\frac{1}{J(\theta)}}{\mathcal{D}(T)}.$$

Pretože platí, že

$$\mathcal{D}(T) \geq \frac{1}{J(\theta)} > 0,$$

je

$$1 \geq \frac{1}{\mathcal{D}(T)J(\theta)} = e > 0,$$

čiže

$$0 < e \leq 1.$$

Definícia 17.9. Ak pre odhad T je $e = 1$, tak tento odhad sa nazýva eficientný (výdatný).

Poznámka. Eficiencia je definovaná len pre regulárne nestranné odhady. Ak T nie je regulárny, môže sa stať, že formálne sa dá spočítať jeho eficiencia a vyjde $e > 1$ (pozri Cramer, H., Mathematical methods of statistics, Princeton, 1946, §32.3).

Príklad 17.6. Nech $X_1, ..., X_n$ je náhodný výber z $N(\theta, 1)$. Odhad $T = \overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$ je nestranný, regulárny a eficientný odhad parametra θ. Dokážte.

Riešenie: $T = \overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$, pričom $X_1, ..., X_n$ sú nezávislé, teda $\mathcal{E}(T) = \theta$, $\mathcal{D}(T) = \frac{1}{n}$, $\mathcal{E}(T^2) = \theta^2 + \frac{1}{n}$. Združené rozdelenie $\mathbf{X} = (X_1, ..., X_n)'$ má hustotu $f(\mathbf{x}, \theta) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{i=1}^n (x_i-\theta)^2}$. Tento systém hustôt je regulárny (dokážte). ďalej

$$f'(\mathbf{x}, \theta) = \frac{\partial f(\mathbf{x}, \theta)}{\partial \theta} = \sum_{i=1}^n (x_i - \theta) f(\mathbf{x}, \theta),$$

$$J(\theta) = \int_{\mathbb{R}^n} \left[\frac{\sum_{i=1}^n (x_i - \theta) f(\mathbf{x}, \theta)}{f(\mathbf{x}, \theta)} \right]^2 f(\mathbf{x}, \theta) d\mathbf{x} =$$

$$= \int_{\mathbb{R}^n} \left[\sum_{i=1}^n (x_i - \theta) \right]^2 f(\mathbf{x}, \theta) d\mathbf{x} = \mathcal{E}[\sum_{i=1}^n (X_i - \theta)]^2 = \sum_{i=1}^n 1 = n.$$

Ešte preverme podmienku (iii) Raovej-Cramerovej vety. Platí

$$\frac{\partial}{\partial \theta} \int_{\mathbb{R}^n} T(\mathbf{x}) dF(\mathbf{x}, \theta) = \frac{\partial}{\partial \theta} \mathcal{E}(T(\mathbf{X})) = \frac{\partial}{\partial \theta} \theta = 1$$

a

$$\int_{\mathbb{R}^n} T(\mathbf{x}) \frac{f'(\mathbf{x}, \theta)}{f(\mathbf{x}, \theta)} dF(\mathbf{x}, \theta) = \int_{\mathbb{R}^n} T(\mathbf{x}) f'(\mathbf{x}, \theta) d\mathbf{x} =$$

$$= \int_{\mathbb{R}^n} \left(\frac{1}{n} \sum_{i=1}^n x_i \right) \sum_{i=1}^n (x_i - \theta) f(\mathbf{x}, \theta) d\mathbf{x} = \mathcal{E}\{\overline{X}(n(\overline{X} - \theta))\} =$$

$$= n\mathcal{E}[(\overline{X} - \theta + \theta)(\overline{X} - \theta)] = n\{\mathcal{D}(\overline{X}) + \theta\mathcal{E}(\overline{X} - \theta)\} = n \left\{ \frac{1}{n} \right\} = 1.$$

Vidíme, že $T = \overline{X}$ je nestranný odhad θ, regulárny a jeho disperzia je $\frac{1}{n}$, čo je $\frac{1}{J(\theta)}$. Je preto aj eficientný.

18. Metóda maximálnej vierohodnosti a momentová metóda

Popíšeme dve konkrétne cesty na odvodenie odhadu. Majme náhodný vektor $\mathbf{X} = (X_1, X_2, ..., X_n)'$ a poznáme jeho hustotu $p(\mathbf{x}, \theta)$ (resp. v diskrétnom prípade pravdepodobnostnú funkciu $(\mathbf{x}_m, p_m(\theta))_{m \in J}$). Predpokladajme, že $\theta \in \Omega$, kde Ω je otvorený interval v $\mathbb{R}$. Pri pevnej hodnote θ je hustota $p(\mathbf{x}, \theta)$ funkciou $\mathbf{x}$. Pravda pre ľubovoľné (pevné) $\mathbf{x}$ môžeme $p(\mathbf{x}, \theta)$ (resp. $P\{X_1 = x_1, ..., X_n = x_n; \theta\}$) chápať ako funkciu parametra θ. Pre túto funkciu budeme používať označenie $L(\mathbf{x}, \theta)$ a volať ju vierohodnostná funkcia (z anglického likelihood function). Samozrejme môžeme uvažovať aj funkciu náhodného vektora $L(\mathbf{X}, \theta)$ (ak napr. $L(\cdot, \theta)$ je pre dané θ merateľná funkcia, tak $L(\mathbf{X}, \theta)$ je náhodná veličina).

Definícia 18.1. Ak existuje $\theta^* \in \Omega$, že pre všetky $\theta \in \Omega$

$$L(\mathbf{X}, \theta) \leq L(\mathbf{X}, \theta^*), \tag{18.1}$$

potom hovoríme, že θ^* je *odhad parametra θ získaný metódou maximálnej vierohodnosti (ML odhad).*

Analyzujme predchádzajúcu definíciu. Pre dané $\mathbf{x}$ vieme (často aj explicitne) nájsť $\theta^*(\mathbf{x})$, ktoré maximalizuje $L(\mathbf{x}, \theta)$. Teda takto máme určenú (niekedy aj explicitne) funkciu $\theta^*(\mathbf{x})$. Ak ju chápeme ako funkciu náhodného vektora $\theta^*(\mathbf{X})$, tak toto je ML odhad parametra θ. Jeho realizácia je $\theta^*(\mathbf{x})$ (ak realizácia náhodného vektora $\mathbf{X}$ je $\mathbf{x}$). Zrejme ak $L(\mathbf{x}, \theta)$ (pri každom $\mathbf{x}$) je dostatočne hladká funkcia θ (napr. pre každé $\mathbf{x}$ existuje $\frac{\partial L(\mathbf{x}, \theta)}{\partial \theta}$), potom θ^* nutne musí byť riešením rovnice

$$\left. \frac{\partial L(\mathbf{X}, \theta)}{\partial \theta} \right|_{\theta = \theta^*} = 0. \tag{18.2}$$

Ak položíme $\ln 0 = -\infty$, potom $L(\mathbf{X}, \theta) \leq L(\mathbf{X}, \theta^*)$ bude platiť pre všetky $\theta \in \Omega$ práve vtedy ak

$$\forall \theta \in \Omega \quad \ln L(\mathbf{X}, \theta) \leq \ln L(\mathbf{X}, \theta^*).$$

Teda v prípade dostatočne hladkej funkcie L môžeme písať rovnicu (18.2) ako

$$\left.\frac{\partial \ln L(\mathbf{X},\theta)}{\partial\theta}\right|_{\theta=\theta^*} = \left.\frac{\partial \ln p(\mathbf{X},\theta)}{\partial\theta}\right|_{\theta=\theta^*} = \left.\frac{\partial l(\mathbf{X},\theta)}{\partial\theta}\right|_{\theta=\theta^*} = 0 \tag{18.3}$$

Rovnicu (18.3) voláme vierohodnostná rovnica.

Poznámka. Dôležitý pri úvahách okolo rovnice (18.3) je fakt, že Ω je otvorený interval. Keby k Ω patril jej krajný bod, mohlo by sa stať, že θ^* splňujúci (18.3) je práve v tomto bode. Vtedy ale θ^* nemusí byť koreňom (18.3).

V ďalšom sa ohraničíme na prípad, že $X_1, ..., X_n$ je nahodný výber zo spojitého rozdelenia s hustotou $f(\cdot,\theta)$ (v diskrétnom prípade s pravdepodobnostnou funkciou $(x_m, p_m)_{m\in J}$). Potom $\mathbf{X} = (X_1, ..., X_n)'$ má hustotu $p(\mathbf{x},\theta) = f(x_1,\theta)...f(x_n,\theta)$. Vierohodnostná rovnica (18.3) má preto tvar

$$\sum_{i=1}^{n} \left.\frac{\partial \ln f(X_i,\theta)}{\partial\theta}\right|_{\theta=\theta^*} = 0 \tag{18.4}$$

(v diskrétnom prípade $\sum_{i=1}^n \left.\frac{\partial \ln p_{X_i}(\theta)}{\partial\theta}\right|_{\theta=\theta^*} = 0$).

Poznámka. Dá sa dokázať, že za "rozumných predpokladov" existuje riešenie ${}^n\theta^* = \theta^*(X_1, ..., X_n)$ vierohodnostnej rovnice (18.4), (vlastne postupnosť riešení $\{{}^n\theta^*\}_{n\geq 1}$), ktoré je maximálne vierohodným odhadom. Tento odhad má veľmi význačnú pravepodobnostnú vlastnosť, a síce je konzistentným odhadom (teda podľa pravdepodobnosti ${}^n\theta^* \to \theta_0$, kde θ_0 je skutočná hodnota parametra θ) (pozri napr. Anděl, J., Základy matematické statistiky, MATFYZPRESS, Praha, 2005, §7.6). Navyše to je asymptoticky normálny vierohodný odhad, presnejšie

$$\sqrt{n}({}^n\theta^* - \theta_0) \to N\left(0, \frac{1}{J(\theta_0)}\right),$$

(konvergencia v distribúcii).

Pre praktické účely to znamená, že pri "dostatočne veľkom" n považujeme rozdelenie pravdepodobnosti náhodnej veličiny ${}^n\theta^*$ za $N\left(\theta_0, \frac{1}{nJ(\theta_0)}\right)$. Pretože $J(\theta_0)$ nepoznáme, "nahrádzame" ho (blízkou) hodnotou $J({}^n\theta^*)$.

Príklad 18.1. Majme náhodný výber $\mathbf{X} = (X_1, ..., X_n)$ z binomického rozdelenia s paramtrami m (známym) a $\pi \in (0,1)$. Parameter π odhadujeme metódou maximálnej vierohodnosti. Náhodná veličina $X_i, i \in \{1,2,...,n\}$ má pravdepodobnostnú funkciu $(j, p_j)_{j=0,1,...,m}$, kde $p_j = \binom{m}{j}\pi^j(1-\pi)^{m-j}$. Vierohodnostná rovnica (18.4) má tvar

$$\sum_{i=1}^{n} \left.\frac{\partial \ln p_{X_i}(\pi)}{\partial\pi}\right|_{\pi=\pi^*} = \sum_{i=1}^{n} \left.\frac{\partial \ln\left[\binom{m}{X_i}\pi^{X_i}(1-\pi)^{m-X_i}\right]}{\partial\pi}\right|_{\pi=\pi^*} =$$

$$= \sum_{i=1}^{n} \frac{\partial}{\partial\pi}\left[\ln\binom{m}{X_i} + X_i\ln\pi + (m-X_i)\ln(1-\pi)\right]_{\pi=\pi^*} =$$

$$= \frac{\partial}{\partial\pi}\left[n\overline{X}\ln\pi + n(m-\overline{X})\ln(1-\pi)\right]_{\pi=\pi^*} =$$

$$= \frac{1}{\pi^*}n\overline{X} - \frac{1}{1-\pi^*}n(m-\overline{X}) = 0 \quad \Longrightarrow \pi^* = \frac{\overline{X}}{m}$$

ak $\overline{X} \neq 0, \overline{X} \neq m$. Ľahko sa presvedčíme, že ide o maximum, lebo

$$\left.\frac{\partial^2 l(\mathbf{X},\theta)}{\partial \pi^2}\right|_{\pi=\pi^*} = -\frac{1}{\pi^{*2}} n\overline{X} - \frac{1}{(1-\pi^*)^2} n(m-\overline{X}) =$$

$$= -n\left[\frac{\overline{X}}{\pi^{*2}} + \frac{m-\overline{X}}{(1-\pi^*)^2}\right] < 0.$$

V prípade, že máme vektor parametrov $\boldsymbol{\theta} \in \Omega \subset \mathtt{R}^m$, tak namiesto jednej vierohodnostnej rovnice (18.4) riešime sústavu vierohodnostných rovníc

$$\sum_{i=1}^{n} \left.\frac{\partial \ln f(X_i, \boldsymbol{\theta})}{\partial \theta_j}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} = 0, \quad j = 1, 2, ..., m \tag{18.5}$$

(v diskrétnom prípade $\sum_{i=1}^{n} \left.\frac{\partial \ln p_{X_i}(\boldsymbol{\theta})}{\partial \theta_j}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} = 0, \quad j = 1, 2, ..., m$). Aj v mnohorozmernom prípade parametra majú ML odhady analogické vlastnosti ako v jednorozmernom prípade, bližšie pozri napr. v knihe Anděl, J., Základy matematické statistiky, MATFYZPRESS, Praha, 2005.

Príklad 18.2. Majme náhodný výber $\mathbf{X} = (X_1, ..., X_n)'$ z normálneho rozdelenia s parametrami μ a σ^2, teda $\boldsymbol{\theta} = (\mu, \sigma^2)'$ a priestor parametrov $\Omega = \mathtt{R} \times (0, \infty)$. Parametre odhadnime metódou maximálnej vierohodnosti.

Hustota náhodnej veličiny $X_i, \quad i \in \{1, 2, ..., n\}$ je

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

a sústava vierohodnostných rovníc (18.5) je

$$\frac{\partial}{\partial \sigma^2}\left[\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\right]_{\mu=\mu^*, \sigma^2=\sigma^{*2}} = 0$$

$$\frac{\partial}{\partial \mu}\left[\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\right]_{\mu=\mu^*, \sigma^2=\sigma^{*2}} = 0.$$

Teda

$$-\frac{n}{2}\frac{1}{2\pi\sigma^{*2}} 2\pi + \frac{1}{2\sigma^{*4}}\sum_{i=1}^{n}(X_i-\mu^*)^2 = 0 \tag{18.6}$$

$$\frac{1}{2\sigma^{*2}}\sum_{i=1}^{n} 2(X_i-\mu^*) = 0. \tag{18.7}$$

Z (18.7) dostávame, že $\mu^* = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$ a po dosadení do (18.6) máme $\sigma^{*2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$.

Dokážme ešte, že pre všetky $(\mu, \sigma^2)' \in \Omega = \mathtt{R} \times (0, \infty)$ je

$$l(\mathbf{X}, \mu, \sigma^2) \leq l(\mathbf{X}, \mu^*, \sigma^{*2}),$$

čiže pre každú realizáciu $\mathbf{x}$ náhodného vektora $\mathbf{X}$ je

$$l(\mathbf{x}, \mu, \sigma^2) \leq l(\mathbf{x}, \overline{x}, s^{*2}),$$

kde $\overline{x}$ je realizácia $\mu^* = \overline{X}$ a s^{*2} je realizácia $\sigma^{*2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$.

Upravujme

$$l(\mathbf{x}, \mu, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}[(x_i - \overline{x}) + (\overline{x} - \mu)]^2 =$$

$$= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) -$$

$$-\frac{1}{2\sigma^2}\left\{\sum_{i=1}^{n}(x_i-\overline{x})^2 + 2\sum_{i=1}^{n}(x_i-\overline{x})(\overline{x}-\mu) + \sum_{i=1}^{n}(\overline{x}-\mu)^2\right\} =$$

$$= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\left\{\sum_{i=1}^{n}(x_i-\overline{x})^2 + n(\overline{x}-\mu)^2\right\} =$$

$$(18.8) \qquad = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{ns^{*2} + n(\overline{x}-\mu)^2}{2\sigma^2}.$$

Samozrejme

$$l(\mathbf{x}, \overline{x}, s^{*2}) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(s^{*2}) - \frac{n}{2}.$$

Preto pre každú realizáciu $\mathbf{x}$ je

$$l(\mathbf{x}, \overline{x}, s^{*2}) - l(\mathbf{x}, \mu, \sigma^2) = \frac{n}{2}\left\{\left[\left(\frac{s^{*2}}{\sigma^2} - 1\right) - \ln\frac{s^{*2}}{\sigma^2}\right] + \frac{(\overline{x}-\mu)^2}{\sigma^2}\right\} \geq 0,$$

lebo pre všetky kladné čísla $t = \frac{s^{*2}}{\sigma^2}$ je $\varphi(t) = t - 1 - \ln t \geq 0$. (Funkcia $\varphi(t)$ nadobúda pre kladné t minimum v bode $t = 1$, pričom $\varphi(1) = 0$.)

Teraz si popíšeme relatívne najjednoduchšiu metódu získania odhadu - momentovu metódu. Nech $\mathbf{X} = (X_1, ..., X_n)'$ je náhodný výber z rozdelenia, ktoré závisí od $\boldsymbol{\theta} = (\theta_1, ..., \theta_m)'$. Nech pre všetky $\boldsymbol{\theta} \in \Omega$ existujú momenty

$$\mu_k' = \mathcal{E}(X_i^k), \quad k = 1, 2, ..., m.$$

Samozrejme tieto momenty tiež závisia od $\boldsymbol{\theta}$, čiže $\mu_k' = \mu_k'(\boldsymbol{\theta})$. Výberové momenty sú

$$M_k' = \frac{1}{n}\sum_{i=1}^{n} X_i^k, \quad k = 1, 2, ... \,.$$

Momentová metóda odhadu $\boldsymbol{\theta}$ spočíva v tom, že (momentový) odhad $\widetilde{\boldsymbol{\theta}}$ je riešením rovníc

$$(18.9) \qquad \mu_k'(\widetilde{\boldsymbol{\theta}}) = M_k', \quad k = 1, 2, ..., m.$$

Niekedy sa môže stať, že m rovníc (18.9) nestačí k (jednoznačnému) určeniu $\widetilde{\boldsymbol{\theta}}$. Potom sa obyčajne vezmú ďalšie rovnice $\mu_k'(\widetilde{\boldsymbol{\theta}}) = M_k'$, $\quad k = m+1, ...$ (samozrejme príslušné teoretické momenty μ_k' musia existovať). Podľa Chinčinovej vety (Veta 13.2.) M_k' konvergujú podľa pravdepodobnosti k μ_k'. Tento fakt spolu s inými limitnými vetami obyčajne umožňuje v konkrétnom prípade dokázať konzistenciu odhadov získaných momentovou metódou.

Príklad 18.3. Majme náhodný výber $X_1, X_2, ..., X_n$ z exponenciálneho rozdelenia s $f(x,\theta) = \frac{1}{\theta}e^{-\frac{x}{\theta}}$ pre $x > 0$ a $f(x,\theta) = 0$ pre $x \leq 0$. Platí

$$\mu_1'(\theta) = \theta, \quad M_1' = \overline{X},$$

teda dostávame odhad parametra θ momentovou metódou

$$\widetilde{\theta} = \overline{X}.$$

19. Bodové a intervalové odhady parametrov normálneho rozdelenia

Najprv si dokážme dve tvrdenia:

Veta 19.1. Nech náhodný vektor $\mathbf{X} = (X_1, ..., X_n)' \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, pričom $\boldsymbol{\Sigma}$ je pozitívne definitná matica (regulárna). (Teda $\mathbf{X}$ má regulárne normálne rozdelenie.) Ak $\mathbf{B}_{n,n}$ je regulárna matica a $\mathbf{a} \in \mathtt{R}^n$, tak náhodný vektor $\mathbf{Y} = \mathbf{a} + \mathbf{B}\mathbf{X} \sim N(\mathbf{a} + \mathbf{B}\mathbf{X}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}')$.

Dôkaz: Použijeme Vetu 9.4 (o hustote transformovaného náhodného vektora). Hustota náhodného vektora $\mathbf{X}$ je

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-\frac{n}{2}} (\det \boldsymbol{\Sigma})^{-\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})}.$$

Inverzné zobrazenie k $\mathbf{h}: \ \mathbf{h}(\mathbf{x}) = \mathbf{a} + \mathbf{B}\mathbf{x}$ je $\mathbf{h}^{-1}(\mathbf{y}) = \mathbf{B}^{-1}(\mathbf{y} - \mathbf{a})$ a Jakobián $D_{\mathbf{h}^{-1}}(\mathbf{y}) = \det \frac{\partial \mathbf{h}^{-1}}{\partial \mathbf{y}'} = \det \mathbf{B}^{-1}$. Preto hustota náhodného vektora $\mathbf{Y} = \mathbf{a} + \mathbf{B}\mathbf{X}$ je

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(\mathbf{B}^{-1}(\mathbf{y} - \mathbf{a})) |\det \mathbf{B}^{-1}| =$$

$$= (2\pi)^{-\frac{n}{2}} (\det \boldsymbol{\Sigma})^{-\frac{1}{2}} |\det \mathbf{B}|^{-1} e^{-\frac{1}{2}(\mathbf{B}^{-1}(\mathbf{y}-\mathbf{a})-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{B}^{-1}(\mathbf{y}-\mathbf{a})-\boldsymbol{\mu})} =$$

$$= (2\pi)^{-\frac{n}{2}} (\det(\mathbf{B}\boldsymbol{\Sigma}\mathbf{B}'))^{-\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{y}-\mathbf{a}-\mathbf{B}\boldsymbol{\mu})' (\mathbf{B}\boldsymbol{\Sigma}\mathbf{B}')^{-1} (\mathbf{y}-\mathbf{a}-\mathbf{B}\boldsymbol{\mu})},$$

čo je hustota $N(\mathbf{a} + \mathbf{B}\boldsymbol{\mu}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}')$. ♣

Veta 19.2. Nech $X_1, ..., X_n$ sú nezávislé, $X_i \sim N(\mu_i, \sigma^2)$, $i = 1, 2, ..., n$. $\mathbf{B}$ je ortonormálna $n \times n$ matica. Položme $\mathbf{X} = (X_1, ..., X_n)'$ a $\mathbf{Y} = (Y_1, ..., Y_n)' = \mathbf{B}(\mathbf{X} - \boldsymbol{\mu})$, kde $\boldsymbol{\mu} = (\mu_1, ..., \mu_n)'$. Potom $Y_1, ..., Y_n$ sú nezávislé a $Y_j \sim N(0, \sigma^2)$, $j = 1, 2, ..., n$.

Dôkaz: Pretože $X_1, ..., X_n$ sú nezávislé, $X_i \sim N(\mu_i, \sigma^2), i = 1, ..., n$, má $\mathbf{X}$ hustotu

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x_i - \mu_i}{\sigma}\right)^2} = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i - \mu_i}{\sigma}\right)^2},$$

čo je hustota $N(\boldsymbol{\mu}, \sigma^2 \mathbf{I})$. Ak je $\mathbf{B}$ ortonormálna (teda $\mathbf{B}\mathbf{B}' = \mathbf{B}'\mathbf{B} = \mathbf{I}$), tak z Vety 19.1 plynie, že $\mathbf{Y} = \mathbf{B}(\mathbf{X} - \boldsymbol{\mu}) \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$ a preto má $\mathbf{Y}$ hustotu

$$f_{\mathbf{Y}}(\mathbf{y}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{y_i}{\sigma}\right)^2} = \prod_{i=1}^{n} f_{Y_i}(y_i). \quad ♣$$

Teraz si dokážme nasledujúcu vetu:

Veta 19.3 Majme $X_1, ..., X_n$ náhodný výber z rozdelenia $N(\mu, \sigma^2)$. Pre výberový priemer $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ a výberový rozptyl $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ platí

(i) $\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$;

(ii) $\frac{n-1}{\sigma^2} S^2 \sim \chi^2_{n-1}$;

(iii) ak je $n > 1$, tak sú náhodné veličiny $\overline{X}$ a S^2 nezávislé.

Dôkaz: Uvažujme ortonormálnu maticu $\mathbf{B}$ (Helmertova matica)

$$\mathbf{B} = \begin{pmatrix} \mathbf{b}_1' \\ \mathbf{b}_2' \\ \mathbf{b}_3' \\ \vdots \\ \mathbf{b}_{n-1}' \\ \mathbf{b}_n' \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \cdots & \cdots & \frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{1\cdot 2}} & -\frac{1}{\sqrt{1\cdot 2}} & 0 & \cdots & \cdots & 0 \\ \frac{1}{\sqrt{2\cdot 3}} & \frac{1}{\sqrt{2\cdot 3}} & -\frac{2}{\sqrt{2\cdot 3}} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ \frac{1}{\sqrt{(n-2)(n-1)}} & \frac{1}{\sqrt{(n-2)(n-1)}} & \cdots & \frac{1}{\sqrt{(n-2)(n-1)}} & -\frac{n-2}{\sqrt{(n-2)(n-1)}} & 0 \\ \frac{1}{\sqrt{(n-1)n}} & \frac{1}{\sqrt{(n-1)n}} & \cdots & \cdots & \frac{1}{\sqrt{(n-1)n}} & -\frac{n-1}{\sqrt{(n-1)n}} \end{pmatrix}$$

(presvedčte sa, že je ortonormálna). Podľa Vety 19.2 je $\mathbf{Y} = (Y_1, ..., Y_n)' = \mathbf{B}(\mathbf{X} - \boldsymbol{\mu}) \sim N(\mathbf{0}, \sigma^2\mathbf{I})$, (tentokrát $\boldsymbol{\mu} = (\mu, ..., \mu)'$) a teda $Y_i \sim N(0, \sigma^2)$ $i = 1, ..., n$ sú združene nezávislé.

Počítajme

$$\mathbf{Y}'\mathbf{Y} = (\mathbf{X} - \boldsymbol{\mu})'\mathbf{B}'\mathbf{B}(\mathbf{X} - \boldsymbol{\mu}) = (\mathbf{X} - \boldsymbol{\mu})'(\mathbf{X} - \boldsymbol{\mu}) = \sum_{i=1}^{n}(X_i - \mu)^2, \tag{19.1}$$

$$Y_1 = \mathbf{b}_1'(\boldsymbol{X} - \boldsymbol{\mu}) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i - \mu) = \frac{1}{\sqrt{n}}(n\overline{X} - n\mu) = \sqrt{n}(\overline{X} - \mu), \tag{19.2}$$

$$\sum_{i=1}^{n}(X_i - \overline{X})^2 = \sum_{i=1}^{n}[(X_i - \mu) - (\overline{X} - \mu)]^2 = \tag{19.3}$$

$$= \sum_{i=1}^{n}(X_i - \mu)^2 - 2(\overline{X} - \mu)\sum_{i=1}^{n}(X_i - \mu) + n(\overline{X} - \mu)^2 = \mathbf{Y}'\mathbf{Y} - n(\overline{X} - \mu)^2 =$$

$$= \sum_{i=1}^{n}Y_i^2 - Y_1^2 = \sum_{i=2}^{n}Y_i^2.$$

Z (19.2) dostávame $\overline{X} = \frac{Y_1}{\sqrt{n}} + \mu$ a podľa Príkladu 9.2 (alebo Vety 19.1 pre $n = 1$) je $\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$. Podľa (19.3) je $\frac{n-1}{\sigma^2}S^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \sum_{j=2}^{n}\left(\frac{Y_j}{\sigma}\right)^2$ a je teda súčtom mocnín nezávislých náhodných veličín $\frac{Y_j}{\sigma}$, pričom každá z nich má $N(0,1)$ rozdelenie. Preto $\frac{n-1}{\sigma^2}S^2 \sim \chi^2_{n-1}$. Pretože $Y_1, ..., Y_n$ sú nezávislé sú aj $\overline{X} = \frac{Y_1}{\sqrt{n}} + \mu$ a $S^2 = \frac{1}{n-1}\sum_{i=2}^{n}Y_i^2$ nezávislé. ♣

K zostrojeniu bodových a intervalových odhadov parametrov normálneho rozdelenia budeme okrem náhodných veličín (štatistík) $\overline{X}$ a S^2 potrebovať ešte štatistiky

$$U = \frac{\overline{X} - \mu}{\sigma}\sqrt{n} \qquad \text{a} \qquad T = \frac{\overline{X} - \mu}{S}\sqrt{n}.$$

Veta 19.4 Majme $X_1, ..., X_n$ náhodný výber z rozdelenia $N(\mu, \sigma^2)$. Nech $\overline{X} = \frac{1}{n}\sum_{i=1}^{n}X_i$ je výberový priemer a $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ výberový rozptyl. Platí

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

(ii) $T = \frac{\overline{X} - \mu}{S}\sqrt{n} \sim t_{n-1}$ (Studentovo t rozdelenie s $n - 1$ stupňami voľnosti).

Dôkaz: Pretože podľa Vety 19.3 $\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$, podľa Príkladu 9.2 (alebo Vety 19.1) má $U = \overline{X}\frac{\sqrt{n}}{\sigma} - \frac{\mu\sqrt{n}}{\sigma} \sim N(0, 1)$ rozdelenie.

Podľa predchádzajúcej vety sú $\overline{X}$ a S^2 nezávislé, preto aj U a $\frac{n-1}{\sigma^2}S^2$ sú nezávislé, pričom $\frac{n-1}{\sigma^2}S^2 \sim \chi^2_{n-1}$. Náhodná veličina

$$T = \frac{U}{\sqrt{\frac{\frac{n-1}{\sigma^2}S^2}{n-1}}} = \frac{\frac{\overline{X} - \mu}{\sigma}\sqrt{n}}{\sqrt{\frac{\frac{n-1}{\sigma^2}S^2}{n-1}}} = \frac{\overline{X} - \mu}{S}\sqrt{n}$$

má Studentovo t_{n-1} rozdelenie. ♣

Veta 19.5 Majme $X_1, ..., X_n$ náhodný výber z rozdelenia $N(\mu, \sigma^2)$, kde μ je neznámy parameter (stredná hodnota) a σ^2 známe kladné číslo. Potom

$$\left\langle \overline{X} - u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{X} + u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \right\rangle \tag{19.4}$$

je $100(1 - \alpha)\%$–ný interval spoľahlivosti pre strednú hodnotu μ pri známom σ^2 ($u_{1-\frac{\alpha}{2}}$ je $(1 - \frac{\alpha}{2})$ kvantil $N(0, 1)$ rozdelenia (tabuľkovaná hodnota)).

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

$$1-\alpha = P\{u_{\frac{\alpha}{2}} \leq \frac{\overline{X}-\mu}{\sigma}\sqrt{n} \leq u_{1-\frac{\alpha}{2}}\} =$$

$$= P\left\{\overline{X} - u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right\},$$

(lebo $u_{\frac{\alpha}{2}} = -u_{1-\frac{\alpha}{2}}$). ♣

Interval (19.4) je náhodný interval s pevnou dĺžkou (jeho krajné hodnoty sú náhodné premenné). Chápať ho treba (frekventisticky) tak, že ak by sme realizovali napr. $M-$krát nezávisle náhodný výber rozsahu n z $N(\mu,\sigma^2)$ rozdelenia (pritom σ^2 poznáme a μ je vždy rovnaké), tak "približne" $\frac{M}{100}(1-\alpha)$ realizácií pokryje skutočnú neznámu hodnotu μ (teda $100(1-\alpha)\%$ z týchto realizácií pokryje μ).

Veta 19.6 Majme $X_1,...,X_n$ náhodný výber z rozdelenia $N(\mu,\sigma^2)$, kde μ ani σ^2 nepoznáme. Potom

$$\left\langle \overline{X} - t_{n-1}(1-\frac{\alpha}{2})\frac{S}{\sqrt{n}}, \overline{X} + t_{n-1}(1-\frac{\alpha}{2})\frac{S}{\sqrt{n}} \right\rangle \tag{19.5}$$

je $100(1-\alpha)\%-$ný interval spoľahlivosti pre strednú hodnotu μ pri neznámom σ^2 ($t_{n-1}(1-\frac{\alpha}{2})$ je $(1-\frac{\alpha}{2})$ kvantil Studentovho t_{n-1} rozdelenia) a

$$\left\langle \frac{(n-1)S^2}{\chi^2_{n-1}(1-\frac{\alpha}{2})}, \frac{(n-1)S^2}{\chi^2_{n-1}(\frac{\alpha}{2})} \right\rangle \tag{19.6}$$

je $100(1-\alpha)\%-$ný interval spoľahlivosti pre rozptyl σ^2 ($\chi^2_{n-1}(1-\frac{\alpha}{2})$ je $(1-\frac{\alpha}{2})$ kvantil χ^2_{n-1} rozdelenia).

Dôkaz: Pretože $T = \frac{\overline{X}-\mu}{S}\sqrt{n} \sim t_{n-1}$, platí

$$1-\alpha = P\left\{t_{n-1}(\frac{\alpha}{2}) \leq \frac{\overline{X}-\mu}{S}\sqrt{n} \leq t_{n-1}(1-\frac{\alpha}{2})\right\} =$$

$$= P\left\{\overline{X} - t_{n-1}(1-\frac{\alpha}{2})\frac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{n-1}(1-\frac{\alpha}{2})\frac{S}{\sqrt{n}}\right\}.$$

Vzhľadom na to, že $\frac{n-1}{\sigma^2}S^2 \sim \chi^2_{n-1}$, zase platí

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

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

V ďalšom sa budeme zaoberať prípadom, že máme dva nezávislé náhodné výbery.

Veta 19.7 Majme $X_1,...,X_{n_X}$ náhodný výber z rozdelenia $N(\mu_X,\sigma^2_X)$, $\overline{X}$ je jeho výberový priemer a S^2_X jeho výberový rozptyl. ďalej majme $Y_1,...,Y_{n_Y}$ náhodný výber z rozdelenia $N(\mu_Y,\sigma^2_Y)$, $\overline{Y}$ je jeho výberový priemer a S^2_Y jeho výberový rozptyl. Predpokladajme, že oba výbery sú nezávislé. Potom

(i) štatistika

$$U_{\overline{X}-\overline{Y}} = \frac{\overline{X}-\overline{Y}-(\mu_X-\mu_Y)}{\sqrt{\frac{\sigma^2_X}{n_X}+\frac{\sigma^2_Y}{n_Y}}} \sim N(0,1),$$

(ii) ak $\sigma^2_X = \sigma^2_Y = \sigma^2$, tak štatistika

$$T_{\overline{X}-\overline{Y}} = \frac{\overline{X}-\overline{Y}-(\mu_X-\mu_Y)}{\sqrt{\frac{(n_X-1)S^2_X+(n_Y-1)S^2_Y}{n_X+n_Y-2}}}\sqrt{\frac{n_Xn_Y}{n_X+n_Y}} \sim t_{n_X+n_Y-2},$$

(iii) štatistika

$$F = \frac{S_X^2}{S_Y^2}\frac{\sigma_Y^2}{\sigma_X^2} \sim F_{n_X-1,n_Y-1}.$$

Dôkaz: Z nezávislosti náhodných výberov vyplýva, že štatistiky $\overline{X}, \overline{Y}, S_X^2, S_Y^2$ sú nezávislé.

(i) Pretože $\overline{X} \sim N(\mu_X, \frac{\sigma_X^2}{n_X})$, $\overline{Y} \sim N(\mu_Y, \frac{\sigma_Y^2}{n_Y})$ a sú nezávislé, je $\overline{X} - \overline{Y} \sim N(\mu_X - \mu_Y, \frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y})$ (vyplýva to napr. z Vety 11.5 o charakteristickej funkcii súčtu nezávislých náhodných veličín). Potom ale štandardizovaná náhodná veličina

$$U_{\overline{X}-\overline{Y}} = \frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}} \sim N(0,1).$$

(ii) Ak je $\sigma_X^2 = \sigma_Y^2 = \sigma^2$, tak štatistika

$$U_{\overline{X}-\overline{Y}} = \frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}} = \frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sigma\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}} =$$

$$= \frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sigma}\sqrt{\frac{n_X n_Y}{n_X + n_Y}} \sim N(0,1).$$

Pretože $\frac{n_X-1}{\sigma^2}S_X^2 \sim \chi^2_{n_X-1}$, $\frac{n_Y-1}{\sigma^2}S_Y^2 \sim \chi^2_{n_Y-1}$ má náhodná veličina

$$\frac{n_X-1}{\sigma^2}S_X^2 + \frac{n_Y-1}{\sigma^2}S_Y^2 = \frac{1}{\sigma^2}[(n_X-1)S_X^2 + (n_Y-1)S_Y^2] \sim \chi^2_{n_X+n_Y-2}$$

(vyplýva to napr. z definície χ^2 rozdelenia) a je nezávislá s $U_{\overline{X}-\overline{Y}}$. Potom ale

$$\frac{U_{\overline{X}-\overline{Y}}}{\sqrt{\frac{\frac{1}{\sigma^2}[(n_X-1)S_X^2+(n_Y-1)S_Y^2]}{n_X+n_Y-2}}} = \frac{\frac{\overline{X}-\overline{Y}-(\mu_X-\mu_Y)}{\sigma}\sqrt{\frac{n_X n_Y}{n_X+n_Y}}}{\sqrt{\frac{\frac{1}{\sigma^2}[(n_X-1)S_X^2+(n_Y-1)S_Y^2]}{n_X+n_Y-2}}} =$$

$$= \frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{(n_X-1)S_X^2+(n_Y-1)S_Y^2}{n_X+n_Y-2}}}\sqrt{\frac{n_X n_Y}{n_X + n_Y}} = T_{\overline{X}-\overline{Y}} \sim t_{n_X+n_Y-2}.$$

(iii) ľahko vidíme, že

$$\frac{\frac{\frac{n_X-1}{\sigma_X^2}S_X^2}{n_X-1}}{\frac{\frac{n_Y-1}{\sigma_Y^2}S_Y^2}{n_Y-1}} = \frac{S_X^2}{S_Y^2}\frac{\sigma_Y^2}{\sigma_X^2} = F \sim F_{n_X-1,n_Y-1}. \quad \clubsuit$$

Teraz už ľahko dokážeme nasledujúcu vetu

Veta 19.8 Majme $X_1, ..., X_{n_X}$ náhodný výber z rozdelenia $N(\mu_X, \sigma_X^2)$, $\overline{X}$ je jeho výberový priemer a S_X^2 jeho výberový rozptyl. ďalej majme $Y_1, ..., Y_{n_Y}$ náhodný výber z rozdelenia $N(\mu_Y, \sigma_Y^2)$, $\overline{Y}$ je jeho výberový priemer a S_Y^2 jeho výberový rozptyl. Predpokladajme, že oba výbery sú nezávislé. Potom

(i) ak sú σ_X^2 a σ_Y^2 známe, tak $100(1-\alpha)\%$ interval spoľahlivosti pre $\mu_X - \mu_Y$ je

$$\left\langle \overline{X} - \overline{Y} - u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}, \overline{X} - \overline{Y} + u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}} \right\rangle,$$

(ii) ak sú σ_X^2 a σ_Y^2 neznáme, ale platí $\sigma_X^2 = \sigma_Y^2 = \sigma^2$, tak $100(1-\alpha)\%$ interval spoľahlivosti pre $\mu_X - \mu_Y$ je

$$\Bigg\langle \overline{X} - \overline{Y} - t_{n_X+n_Y-2}(1-\frac{\alpha}{2})\sqrt{\frac{(n_X-1)S_X^2 + (n_Y-1)S_Y^2}{n_X+n_Y-2}}\sqrt{\frac{n_X+n_Y}{n_X n_Y}},$$

$$\overline{X}-\overline{Y}+t_{n_X+n_Y-2}(1-\frac{\alpha}{2})\sqrt{\frac{(n_X-1)S_X^2+(n_Y-1)S_Y^2}{n_X+n_Y-2}}\sqrt{\frac{n_X+n_Y}{n_Xn_Y}}\Big\rangle,$$

(iii) ak sú $\mu_X,\mu_Y,\sigma_X^2,\sigma_Y^2$ neznáme, tak $100(1-\alpha)\%$ interval spoľahlivosti pre $\frac{\sigma_X^2}{\sigma_Y^2}$ je

$$\Big\langle\frac{S_X^2}{S_Y^2}\frac{1}{F_{n_X-1,n_Y-1}(1-\frac{\alpha}{2})},\frac{S_X^2}{S_Y^2}\frac{1}{F_{n_X-1,n_Y-1}(\frac{\alpha}{2})}\Big\rangle$$

($F_{n_X-1,n_Y-1}(1-\frac{\alpha}{2})$ je $(1-\frac{\alpha}{2})$ kvantil F_{n_X-1,n_Y-1} rozdelenia).

Dôkaz: Spravte ako cvičenie, využite štatistiky z Vety 19.7.

Ešte pre úplnosť si uvedieme bez dôkazu interval spoľahlivosti pre rozdiel stredných hodnôt u tzv. párových výberov.

Veta 19.9. Nech $\mathbf{X}_1=(X_1,Y_1)',...,\mathbf{X}_n=(X_n,Y_n)'$ je náhodný výber z dvojrozmerného normálneho rozdelenia $N(\boldsymbol{\mu},\boldsymbol{\Sigma})$ s paramertami $\boldsymbol{\mu}=(\mu_X,\mu_Y)'$ a $\boldsymbol{\Sigma}=\begin{pmatrix}\sigma_X^2 & \varrho\sigma_X\sigma_Y\\ \varrho\sigma_X\sigma_Y & \sigma_Y^2\end{pmatrix}$, pričom $\mu_X,\mu_Y\in\mathtt{R}$, $\sigma_X^2>0$, $\sigma_Y^2>0$, $\varrho\in(0,1)$. Pre $i=1,2,...,n$ označme $Z_i=X_i-Y_i$, $\overline{Z}=\frac{1}{n}\sum_{i=1}^n Z_i$, $S_Z^2=\frac{1}{n-1}\sum_{i=1}^n(Z_i-\overline{Z})^2$. Potom

$$\Big\langle\overline{Z}-t_{n-1}(1-\frac{\alpha}{2})\frac{S_Z}{\sqrt{n}},\overline{Z}+t_{n-1}(1-\frac{\alpha}{2})\frac{S_Z}{\sqrt{n}}\Big\rangle$$

je $100(1-\alpha)\%-$ný interval spoľahlivosti pre $\mu_X-\mu_Y$.

Niekedy sa takémuto intervalu spoľahlivosti hovorí aj intervalový odhad $\mu_X-\mu_Y$ o spoľahlivosti $(1-\alpha)$.

20. Testovanie hypotéz

Ukážeme si, v čom spočíva (v matematickej štatistike) podstata testovania hypotéz. Myslíme tým štatistické testovanie hypotéz, niekedy tiež hovoríme o testovaní štatistických hypotéz.

Majme náhodný výber $\mathbf{X}=(X_1,...,X_n)'$, pričom nevieme, ci pochádza z rozdelenia $N(\mu_0,\sigma^2)$ alebo z $N(\mu_1,\sigma^2)$, poznáme μ_0,μ_1, $(\mu_0\neq\mu_1)$ aj σ^2.

Máme hypotézu (tzv. nulovú hypotézu)

H_0: výber pochádza z rozdelenia $N(\mu_0,\sigma^2)$

Tzv. alternatívna hypotéza (konkurujúca) je

H_1: výber pochádza z rozdelenia $N(\mu_1,\sigma^2)$

Rozhodnutie bude také, že platnosť H_0 *nezamietneme* alebo *zamietneme*.

Pri rozhodovaní o platnosti H_0 alebo H_1 sa môzeme dopustiť jednej z dvoch chýb.

(i) Ak zamietneme H_0, hoci ona platí (je správna), urobíme tzv. *chybu prvého druhu*.

(ii) Ak nezamietneme H_0, hoci nie je správna (t.j. platí H_1), urobíme *chybu druhého druhu*.

Svoje rozhodovanie založíme na realizácii $\mathbf{x}=(x_1,x_2,...,x_n)'$ náhodného výberu $\mathbf{X}$. Preto bude "ovplyvnené" náhodou. Prirodzene požadujeme, aby rozhodovacie pravidlo, podľa ktorého zamietneme alebo nezamietneme H_0 bolo také, aby pravdepodobnosti oboch chýb boli čo najmenšie. Keď rozsah náhodného výberu n je pevne určený, nedajú sa pravdepodobnosti oboch horeuvedených chýb súčasne urobiť takými malými, ako by sme si priali. Zaužívalo sa trvať na požiadavke, aby pravdepodobnosť chyby prvého druhu bola rovná α, kde α je vopred zvolené číslo z intervalu $(0,1)$. V praxi sa ukázalo vhodné voliť $\alpha\in\{0,1;0,05;0,01\}$. číslu α sa hovorí *hladina významnosti testu*. Pravdepodobnosť chyby druhého druhu označme β.

Štatistické rozhodovanie prebieha tak, že sa dopredu určí tzv. *kritický obor (kritická oblasť)* W $(\in \mathtt{R}^n)$, t.j. množina realizácií $\mathbf{x}$, pri ktorých budeme H_0 zamietať. Teda ak sa realizuje $\mathbf{x} \in W$, tak H_0 zamietneme. Tvar kritického oboru stanovujeme tak, aby za platnosti H_0 padla realizácia $\mathbf{x}$ do kritického oboru "zriedka", ale za platnosti H_1 tam padla "čo najčastejšie". Veľkosť kritického oboru volíme tak, aby sme platnú H_0 zamietali s pravdepodobnosťou α.

Na testovanie (rozhodovanie) použijeme "vhodnú" štatistiku $T = T(\mathbf{X})$, ktorú nazyvame *testovacia štatistika*. V takom prípade popíšeme kritickú oblasť ako množinu $T(W)$. Teda H_0 zamietneme, ak $T(\mathbf{x}) \in T(W)$.

Vráťme sa k testovaniu H_0: výber pochádza z rozdelenia $N(\mu_0, \sigma^2)$
oproti alternatívnej hypotéze H_1: výber pochádza z rozdelenia $N(\mu_1, \sigma^2)$.

Použijeme testovaciu štatistiku $T(\mathbf{X}) = \overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$. Vieme, že (v našom prípade) za platnosti H_0 bude $\overline{X} \sim N(\mu_0, \frac{\sigma^2}{n})$, teda realizácie $\overline{x}$ budú (pri dosť veľkom n) blízo μ_0. Navrhneme také rozhodovacie (testovacie) pravidlo, že ak $|\overline{x} - \mu_0| \geq k$, tak zamietneme H_0. Teda "tvar" kritickej oblasti je $\{\mathbf{x} : \ \overline{x} \in (-\infty, \mu_0 - k) \cup (\mu_0 + k, \infty)\}$. "Veľkosť" kritickej oblasti (teda číslo k) volíme tak, aby pravdepodobnosť chyby prvého druhu bola α, teda aby realizácia $\overline{x}$ padla do kritickej oblasti za platnosti H_0 s pravdepodobnosťou α. Inými slovami chceme aby

$$P\{|\overline{X} - \mu_0| \geq k\} = \alpha,$$

pričom $\overline{X} \sim N(\mu_0, \frac{\sigma^2}{n})$. Zrejme

$$\alpha = P\{|\overline{X} - \mu_0| \geq k\} = P\{|\frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}| \geq \frac{k}{\sigma}\sqrt{n}\} \Rightarrow P\{-\frac{k}{\sigma}\sqrt{n} \leq \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \leq \frac{k}{\sigma}\sqrt{n}\} = 1 - \alpha.$$

Pretože $\frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$, je zrejmé, že $\frac{k}{\sigma}\sqrt{n} = u_{1-\frac{\alpha}{2}}$. Kritická oblasť testu H_0 oproti H_1 s hladinou významnosti α (pri použití testovacej štatistiky $\overline{X}$) je $W_\alpha = \{\mathbf{x} : \ |\overline{x} - \mu_0| \geq u_{1-\frac{\alpha}{2}} \dfrac{\sigma}{\sqrt{n}}\}$.

Treba si všimnúť, že nezamietnutie H_0 neznamená, že H_0 je správna. K tomu, aby sme považovali H_0 za správnu, potrebovali by sme mať ešte záruku, že β je dosť malé. Potom by sme mohli hovoriť, že H_0 prijímame. Testovať H_0 na hladine významnosti α len zaručuje, že zamietnutie nulovej hypotézy, hoci je správna nastane s pravdepodobnosťou α.

V sledovanom príklade sme mali tzv. jednoduchú hypotézu H_0 – testovaný parameter (stredná hodnota) v prípade platnosti H_0 mohol nadobudnúť len jednu hodnotu, a síce μ_0. Aj alternatíva H_1 bola jednoduchá. Pri testovaní hypotéz obyčajne predpokladáme, že parameter rozdelenia pravdepodobnosti náhodného výberu $\mathbf{X}$ je $\boldsymbol{\theta} = (\theta_1, ..., \theta_m)' \in \boldsymbol{\Theta} \subset \mathtt{R}^m$, kde $\boldsymbol{\Theta}$ je parametrický priestor – otvorená a neprázdna množina. $\boldsymbol{\Theta}_0 \subset \boldsymbol{\Theta}$ a $\boldsymbol{\Theta}_1 = \boldsymbol{\Theta} - \boldsymbol{\Theta}_0$ sú dve "konkurujúce si" možiny. $H_0 : \boldsymbol{\theta} \in \boldsymbol{\Theta}_0$ a $H_1 : \boldsymbol{\theta} \in \boldsymbol{\Theta} - \boldsymbol{\Theta}_0$. Pretože H_0 aj H_1 nie sú vo všeobecnosti jednoduché, hladina významnosti testu s kritickou oblasťou W je

$$\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_0} P_{\boldsymbol{\theta}}(\mathbf{X} \in W).$$

Tiež sa uvažuje sa funkcia $\beta(\boldsymbol{\theta}) = P_{\boldsymbol{\theta}}(\mathbf{X} \in W), \quad \boldsymbol{\theta} \in \boldsymbol{\Theta}$. Volá sa *silofunkcia testu* s kritickou oblasťou W. Niekedy sa pracuje s funkciou $1 - \beta(\boldsymbol{\theta})$, ktorá sa volá *operačná charakteristika testu.* Ak je H_1 jednoduchá, (teda $H_1 : \ \boldsymbol{\theta} = \boldsymbol{\theta}_1$, tak $1 - \beta(\theta_1)$ sa volá *sila testu.*

Prehľad niektorých vybraných testov pre jeden náhodný výber $\mathbf{X}$ z $N(\mu, \sigma^2)$ rozdelenia ($\overline{x}$ je realizácia $\overline{X}$ a s^2 je realizácia S^2):

$H_0 : \mu = \mu_0 \;\; H_1 : \mu \neq \mu_0 \quad W_\alpha = \{\mathbf{x} : |\overline{x} - \mu_0|\sqrt{n} \geq \sigma u_{1-\frac{\alpha}{2}}\} \quad \sigma^2$ je známe

$H_0 : \mu = \mu_0 \;\; H_1 : \mu > \mu_0 \quad W_\alpha = \{\mathbf{x} : (\overline{x} - \mu_0)\sqrt{n} \geq \sigma u_{1-\alpha}\} \quad \sigma^2$ je známe

$H_0 : \mu = \mu_0 \;\; H_1 : \mu < \mu_0 \quad W_\alpha = \{\mathbf{x} : (\overline{x} - \mu_0)\sqrt{n} \leq -\sigma u_{1-\alpha}\} \quad \sigma^2$ je známe

$H_0 : \mu = \mu_0 \;\; H_1 : \mu \neq \mu_0 \quad W_\alpha = \{\mathbf{x} : |\overline{x} - \mu_0|\sqrt{n} \geq st_{n-1}(1 - \frac{\alpha}{2})\} \quad \sigma^2$ neznáme

$H_0 : \mu = \mu_0 \;\; H_1 : \mu > \mu_0 \quad W_\alpha = \{\mathbf{x} : (\overline{x} - \mu_0)\sqrt{n} \geq st_{n-1}(1 - \alpha)\} \quad \sigma^2$ neznáme

$H_0 : \mu = \mu_0 \;\; H_1 : \mu < \mu_0 \quad W_\alpha = \{\mathbf{x} : (\overline{x} - \mu_0)\sqrt{n} \leq -st_{n-1}(1 - \alpha)\} \quad \sigma^2$ neznáme

$H_0 : \sigma^2 = \sigma_0^2 \; H_1 : \sigma^2 \neq \sigma_0^2 \; W_\alpha = \{\mathbf{x} : \frac{(n-1)}{\sigma_0^2}s^2 \notin (\chi^2_{n-1}(\frac{\alpha}{2}), \chi^2_{n-1}(1 - \frac{\alpha}{2}))\} \; \mu$ neznáme

$H_0 : \sigma^2 = \sigma_0^2 \; H_1 : \sigma^2 > \sigma_0^2 \quad W_\alpha = \{\mathbf{x} : \frac{(n-1)}{\sigma_0^2}s^2 \geq \chi^2_{n-1}(1 - \alpha)\} \quad \mu$ neznáme

$H_0 : \sigma^2 = \sigma_0^2 \; H_1 : \sigma^2 < \sigma_0^2 \quad W_\alpha = \{\mathbf{x} : \frac{(n-1)}{\sigma_0^2}s^2 \leq \chi^2_{n-1}(\alpha)\} \quad \mu$ neznáme

Prehľad niektorých vybraných testov v prípade dvoch nezávislých náhodných výberov, a síce $\mathbf{X} = (X_1, ..., X_{n_X})'$ z $N(\mu_X, \sigma_X^2)$ s výberovým priemerom $\overline{X}$ a výberovým rozptylom S_X^2 a $\mathbf{Y} = (Y_1, ..., Y_{n_Y})'$ z $N(\mu_Y, \sigma_Y^2)$ s výberovým priemerom $\overline{Y}$ a výberovým rozptylom S_Y^2, $S_{XY}^2 = \dfrac{(n_X - 1)S_X^2 + (n_Y - 1)S_Y^2}{n_X + n_Y - 2}$, $\overline{x}$ $(\overline{y})$ je realizácia $\overline{X}$ $(\overline{Y})$, s_X^2 (s_Y^2) je realizácia S_X^2 (S_Y^2) a s_{XY}^2 je realizácia S_{XY}^2:

σ_X^2, σ_Y^2 sú známe

$H_0 : \mu_X = \mu_Y \;\; H_1 : \mu_X \neq \mu_Y \quad W_\alpha = \{(\mathbf{x}', \mathbf{y}')' : |\overline{x} - \overline{y}| \geq u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}\}$

$\sigma_X^2 = \sigma_Y^2$ neznáme

$H_0 : \mu_X = \mu_Y \; H_1 : \mu_X \neq \mu_Y$

$$W_\alpha = \{(\mathbf{x}', \mathbf{y}')' : |\overline{x} - \overline{y}| \geq t_{n_X+n_Y-2}(1 - \tfrac{\alpha}{2})s_{XY}\sqrt{\tfrac{n_X+n_Y}{n_X n_Y}}\}$$

μ_X, μ_Y neznáme

$H_0 : \sigma_X^2 = \sigma_Y^2 \; H_1 : \sigma_X^2 \neq \sigma_Y^2$

$$W_\alpha = \{(\mathbf{x}', \mathbf{y}')' : \tfrac{s_X^2}{s_Y^2} \notin (F_{n_X-1, n_Y-1}(\tfrac{\alpha}{2}), F_{n_X-1, n_Y-1}(1 - \tfrac{\alpha}{2}))\}$$

Poznámka. Namiesto hladiny významnosti α bežný štatistický softvér (STATISTICA, S+, SAS) udáva dosiahnutú hladinu (anglicky $P-$*value*, *significance value*, *significance level*). Je to najmenšia hladina významnosti testu, pri ktorej by sme (pri danej realizácii testovacej štatistiky) hypotézu H_0 ešte zamietli. Vyjadruje pravdepodobnosť spočítanú za platnosti nulovej hypotézy, že dostaneme práve našu realizáciu alebo realizáciu ešte viac odporujúcu testovanej hypotéze.

Pri "vyberaní" vhodného testu postupujeme tak, že medzi testami na (požadovanej) hladine významnosti α sa snažíme zvoliť test s čo najmenšou pravdepodobnosťou chyby druhého druhu. To ale ukazuje práve funkcia $\beta(\boldsymbol{\theta})$. Obom požiadavkám sa (niekedy) dá vyhovieť v jednoduchom prípade, a síce ak máme jednoduchú hypotézu $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ oproti jednoduchej alternatíve $H_1 : \boldsymbol{\theta} = \boldsymbol{\theta}_1$. Hovorí o tom nasledujúca veta.

Veta 20.1. (Neymanova-Pearsonova lema) Majme náhodný vektor $\mathbf{X} = (X_1, ..., X_n)'$ s hustotou $f(\mathbf{x}, \boldsymbol{\theta})$. Nech k danému $\alpha \in (0, 1)$ existuje také $c > 0$, že pre množinu

$$W_c = \{\mathbf{x} : \; f(\mathbf{x}, \boldsymbol{\theta}_1) \geq cf(\mathbf{x}, \boldsymbol{\theta}_0)\} \tag{20.1}$$

platí

$$\int_{W_c} f(\mathbf{x}, \boldsymbol{\theta}_0)d\mathbf{x} = \alpha. \tag{20.2}$$

Potom pre každú merateľnú množinu W takú, že

$$\int_W f(\mathbf{x}, \boldsymbol{\theta}_0)d\mathbf{x} = \alpha$$

platí

$$\int_{W_c} f(\mathbf{x}, \boldsymbol{\theta}_1) d\mathbf{x} \geq \int_W f(\mathbf{x}, \boldsymbol{\theta}_1) d\mathbf{x}. \tag{20.3}$$

Dôkaz: Množiny W a W_c sa môžu napísať ako

$$W = (W - W_c) \cup (W \cap W_c), \quad W_c = (W_c - W) \cup (W \cap W_c).$$

Počítajme teraz

$$\int_{W_c} f(\mathbf{x}, \boldsymbol{\theta}_1) d\mathbf{x} - \int_W f(\mathbf{x}, \boldsymbol{\theta}_1) d\mathbf{x} =$$

$$= \int_{W_c - W} f(\mathbf{x}, \boldsymbol{\theta}_1) d\mathbf{x} + \int_{W \cap W_c} f(\mathbf{x}, \boldsymbol{\theta}_1) d\mathbf{x} -$$

$$- \int_{W - W_c} f(\mathbf{x}, \boldsymbol{\theta}_1) d\mathbf{x} - \int_{W \cap W_c} f(\mathbf{x}, \boldsymbol{\theta}_1) d\mathbf{x} \geq$$

$$\geq c \int_{W_c - W} f(\mathbf{x}, \boldsymbol{\theta}_0) d\mathbf{x} - c \int_{W - W_c} f(\mathbf{x}, \boldsymbol{\theta}_0) d\mathbf{x} =$$

$$= c \int_{W_c - W} f(\mathbf{x}, \boldsymbol{\theta}_0) d\mathbf{x} + c \int_{W \cap W_c} f(\mathbf{x}, \boldsymbol{\theta}_0) d\mathbf{x} -$$

$$-c \int_{W - W_c} f(\mathbf{x}, \boldsymbol{\theta}_0) d\mathbf{x} - c \int_{W \cap W_c} f(\mathbf{x}, \boldsymbol{\theta}_0) d\mathbf{x} =$$

$$= c \int_{W_c} f(\mathbf{x}, \boldsymbol{\theta}_0) d\mathbf{x} - c \int_W f(\mathbf{x}, \boldsymbol{\theta}_0) d\mathbf{x} = c\alpha - c\alpha = 0,$$

lebo na množine $W_c - W$ je

$$\int_{W_c - W} f(\mathbf{x}, \boldsymbol{\theta}_1) d\mathbf{x} \geq c \int_{W_c - W} f(\mathbf{x}, \boldsymbol{\theta}_0) d\mathbf{x}$$

a mimo množiny W_c je

$$\int_{W - W_c} f(\mathbf{x}, \boldsymbol{\theta}_1) d\mathbf{x} \leq c \int_{W - W_c} f(\mathbf{x}, \boldsymbol{\theta}_0) d\mathbf{x},$$

čiže

$$- \int_{W - W_c} f(\mathbf{x}, \boldsymbol{\theta}_1) d\mathbf{x} \geq -c \int_{W - W_c} f(\mathbf{x}, \boldsymbol{\theta}_0) d\mathbf{x}. \quad \clubsuit$$

Veta 20.1. teda tvrdí, že ak máme testovať jednoduchú hypotézu oproti jednoduchej alternatíve a sú splnené podmienky (20.1) a (20.2), tak test s kritickou oblasťou W_c má hladinu významnosti α a pre akýkoľvek test s hladinou významnosti α je podľa (20.3) sila testu s kritickou oblasťou W_c väčšia. Test s kritickou oblasťou W_c je najsilnejší možný medzi všetkými testami s hladinou významnosti α.

Príklad 20.1. Majme náhodný výber $\mathbf{X} = (X_1, ..., X_n)'$ z rozdelenia $N(\mu, \sigma^2)$, pričom σ^2 poznáme. Nájdite najsilnejší test nulovej hupotézy $H_0 : \ \mu = \mu_0$ oproti alternatívnej hypotéze $H_1 : \ \mu = \mu_1$, kde $\mu_0 < \mu_1$.

Riešenie: Pretože

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

je kritický obor z Neymanovej-Pearsonovej lemy

$$W_c = \left\{ \mathbf{x} : \ \frac{f(\mathbf{x}, \mu_1, \sigma^2)}{f(\mathbf{x}, \mu_0, \sigma^2)} = e^{-\frac{1}{2\sigma^2} \left(\sum_{i=1}^n (x_i - \mu_1)^2 - \sum_{i=1}^n (x_i - \mu_0)^2 \right.} \geq c \right\},$$

teda

$$W_c = \left\{ \mathbf{x} : \frac{1}{2\sigma^2} \left(\sum_{i=1}^{n} (x_i - \mu_0)^2 - \sum_{i=1}^{n} (x_i - \mu_1)^2 \right) \geq \ln c \right\}.$$

Úpravou dostávame

$$W_c = \left\{ \mathbf{x} :\ 2\overline{x}(\mu_1 - \mu_0) - (\mu_1^2 - \mu_0^2) \geq \frac{2\sigma^2}{n} \ln c \right\},$$

$$W_c = \left\{ \mathbf{x} :\ \overline{x} \geq \frac{\mu_1 + \mu_0}{2} + \frac{\sigma^2}{n(\mu_1 - \mu_0)} \ln c \right\},$$

$$W_c = \left\{ \mathbf{x} :\ \frac{\overline{x} - \mu_0}{\sigma} \sqrt{n} \geq \frac{\frac{\mu_1 + \mu_0}{2} + \frac{\sigma^2}{n(\mu_1 - \mu_0)} \ln c - \mu_0}{\sigma} \sqrt{n} \right\},$$

$$W_c = \left\{ \mathbf{x} :\ \frac{\overline{x} - \mu_0}{\sigma} \sqrt{n} \geq \frac{\mu_1 - \mu_0}{2\sigma} \sqrt{n} + \frac{\sigma \ln c}{\sqrt{n}(\mu_1 - \mu_0)} \right\}. \tag{20.4}$$

Treba nám ešte určiť c. Jeho hodnotu spočítame z podmienky

$$\alpha = \int_{W_c} f(\mathbf{x}, \mu_0, \sigma^2) d\mathbf{x} = P_{\mu_o}(\mathbf{X} \in W_c) = P_{\mu_0}\{\omega :\ \mathbf{X}(\omega) \in W_c\} =$$

$$= P \left\{ \omega :\ \frac{\overline{X}(\omega) - \mu_0}{\sigma} \sqrt{n} \geq \frac{\mu_1 - \mu_0}{2\sigma} \sqrt{n} + \frac{\sigma \ln c}{\sqrt{n}(\mu_1 - \mu_0)} \right\}.$$

Pretože $\frac{\overline{X} - \mu_0}{\sigma} \sqrt{n} \sim N(0, 1)$, dostávame, že

$$\frac{\mu_1 - \mu_0}{2\sigma} \sqrt{n} + \frac{\sigma \ln c}{\sqrt{n}(\mu_1 - \mu_0)} = u_{1-\alpha},$$

čize

$$c = e^{\dfrac{2\sqrt{n}(\mu_1 - \mu_0)\sigma u_{1-\alpha} - n(\mu_1 - \mu_0)^2}{2\sigma^2}}.$$

Jednoduchšie určenie W_c je z (20.4)

$$W_c = \left\{ \mathbf{x} :\ \frac{\overline{x} - \mu_0}{\sigma} \sqrt{n} \geq u_{1-\alpha} \right\} = \left\{ \mathbf{x} :\ \overline{x} \geq \mu_0 + u_{1-\alpha} \frac{\sigma}{\sqrt{n}} \right\}.$$

Poznámka. Podobne by sme v prípade hľadania najsilnejšieho testu nulovej hypotézy $H_0 :\ \mu = \mu_0$ oproti alternatívnej hypotéze $H_1 :\ \mu = \mu_1$, keď $\mu_0 > \mu_1$ odvodili, že v tomto prípade H_0 zamietame na hladine významnosti α ak

$$\frac{\overline{x} - \mu_0}{\sigma} \sqrt{n} \leq -u_{1-\alpha}.$$

Urobte ako cvičenie.

DODATOK Platí

$$\int_0^\infty (\cos bx) e^{-a^2x^2} dx = \frac{\sqrt{\pi}}{2a} e^{-\frac{b^2}{4a^2}}, \quad a > 0,$$

$$\int_0^\infty \frac{\cos bx}{1 + x^2} dx = \frac{\pi}{2} e^{-|b|},$$

$$\int_0^\infty x^n e^{-ax^2} dx = \frac{\Gamma\left(\frac{n+1}{2}\right)}{2a^{\frac{n+1}{2}}}, \quad a > 0, n > -1.$$

Dôkazy nájdete v učebnici matematickej analýzy.