Pravděpodobnost a statistika

1 Axiomy teorie pravděpodobnosti

V roce 1933 zavedl Andrej N. Kolmogorov.

Definice (Pravděpodobnostní prostor)**.** Mějme neprázdnou množinu Ω a na ní σ-algebru $\mathcal{F}$. Na $\mathcal{F}$ uvažujme pravděpodobnostní míru P.

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

Definice 1 (σ-algebra)**.** Buď Ω neprázdná množina. Systém $\mathcal{F}$ podmnožin Ω se nazývá σ-algebra, pokud

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \Rightarrow \Omega \setminus A \in \mathcal{F}$ (uzavřenost na doplňku)
3. $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} \in \mathcal{F}$ (uzavřenost na spočetném sjednocení)

Příklad. Ukažme si různé příklady σ-algeber:

- triviální: $\{\emptyset, \Omega\}$
- největší: 2^{Ω}
- $B \subset \Omega : \{\emptyset, B, B^c, \Omega\}$
- $A, B : \{\emptyset, A, B, A \cup B, A \cap B, A^c, B^c, \ldots\}$

Definice 2 (Pravděpodobnostní míra)**.** Funkce $P : \mathcal{F} \to [0,1]$, kde prvky $\mathcal{F}$ jsou podmnožiny Ω nazveme *pravděpodobnostní míra*, pokud splňuje následující podmínky:

1. $P(\Omega) = 1$
2. $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ a $A_i \cap A_j = \emptyset \ \forall i \neq j \Rightarrow P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} A_i$ (spočetná aditivita)

Definice. $\omega \in \Omega$ se nazývá *elementární jev.* $A \in \mathcal{F}$ se nazývá *náhodný jev.*

Elementární jevy většinou nelze pozorovat přímo, je potřeba pozorovat náhodné jevy.

Příklad. Mějme hod kostkou. Potom:
$\Omega = \{1, 2, 3, 4, 5, 6\}$
$\omega = 4$
$A = \{2, 4, 6\}$

P každému náhodnému jevu A jen dá jeho pravděpodobnost.

1. Jestliže $P(A) = 1$, potom A je jev *jistý.*
2. Jestliže $P(A) = 0$, potom A je jev *nemožný.*

Věta 1. *Buď P pravděpodobnostní míra na $\mathcal{F}$. Pak platí následující:*

1. $P(A^c) = 1 - P(A) \ \forall A \in \mathcal{F}$
2. $A, B \in \mathcal{F} : A \subset B$, *pak* $P(A) \leq P(B)$ *a* $P(B \setminus A) = P(B) - P(A)$

Důkaz. Jednotlivé body:

1. $P(A \cup A^c) = P(A) + P(A^c)$. Jelikož $A \cup A^c = \Omega$ a podle vlastností pravděpodobnostní míry $P(\Omega) = 1$, $P(A) + P(A^c) = 1$.
2. $A \subset B \Rightarrow B = A \cup (B \cap A^c)$. To jsou dvě disjunktní množiny. Potom $P(B) = P(A) + P(B \setminus A) \geq P(A)$.

$\square$

Vezměme si systém $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$. Budeme předpokládat $A_i \subset A_{i+1}$. Potom budeme myslet $A_i \nearrow A$, kde $A = \bigcup_{i=1}^{\infty} A_i$, tedy že systém konverguje k A.

Analogicky systém, kde $A_i \supset A_{i+1}$ potom $A_i \searrow A$, kde $A = \bigcap_{i=1}^{\infty} A_i$

Věta 2 (Spojitost pravděpodobnostní míry)**.** *Nechť máme* $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ *takovou, že* $A_i \searrow \emptyset$. *Pak*

$$\lim_{i\to\infty} P(A_i) = 0.$$

Důkaz. Víme, že $P(\bigcap_{i=1}^{\infty} A_i) = P(\emptyset) = 0$. To je ekvivalentní s $1 - P((\bigcap_{i=1}^{\infty} A_i)^c) = 1 - P(\bigcup_{i=1}^{\infty} A_i^c) = 1$.

Dále víme, že $A_i \supseteq A_{i+1}$ a díky tomu $A_i^c \subseteq A_{i+1}^c$, tedy $A_i^c \nearrow \Omega$.

Nadefinujme si posloupnost $B_1 = A_1^c, B_{i+1} = A_{i+1}^c \setminus A_i^c$.

Potom B_i jsou disjunktní, $\bigcup_{i=1}^{\infty} B_i = \Omega$ a $P(\bigcup_{i=1}^{\infty} B_i) = P(\omega) = 1$ a z toho $\sum_{i=1}^{\infty} P(B_i) = 1$.

Ten si rozložíme na dva součty $\sum_{i=1}^{n} P(B_i)_{\to 1} + \sum_{i=n+1}^{\infty} P(B_i)_{\to 0} = 1$ pro $n \to \infty$.

Potom $P(\bigcup_{i=1}^{n} B_i) = P(\bigcup_{i=1}^{n} A_i^c) = 1 - P(\bigcap_{i=1}^{n} A_i)$. Z konvergence se toto rovná $1 - P(A_n)$. Víme, že tento výsledek konverguje k 1, tedy proto $P(A_n) \to 0$. $\square$

Definice 3 (Klasický pravděpodobnostní prostor)**.** Buď Ω neprázdná konečná množina, $\mathcal{F} = 2^{\Omega}$. Definujeme $P(A) = \frac{|A|}{|\Omega|}$ $\forall A \subset \Omega$. Potom $(\Omega, \mathcal{F}, P)$ nazýváme *klasický pravděpodobnostní prostor*.

Definice 4 (Diskrétní pravděpodobnostní prostor)**.** Buď Ω neprázdná konečná nebo spočetná množina, $\mathcal{F} = 2^{\Omega}$. Definujeme $p : \omega \to \mathbb{R}$ taková, že $p(\omega) \in [0,1]$ $\forall \omega \in \Omega$ a $\sum_{\omega\in\Omega} p(\omega) = 1$. Dále definujme $P(a) = \sum_{\omega\in A} p(\omega)$. Potom $(\Omega, \mathcal{F}, P)$ nazýváme *diskrétní pravděpodobnostní prostor*.

Jev A nastane s $P(A)$ nebo nenastane s $P(A^c)$, jev B nastal. Pomůže nám tato znalost zpřesnit znalost o jevu A?

- Jestliže $B \supset A$, potom jev A nastane vždy.
- Jestliže $B \cap A = \emptyset$, jev A nenastane.
- Jinak se to může chovat všelijak:

Definice 5 (Podmíněná pravděpodobnost)**.** Mějme jevy $A, B \in \mathcal{F}, P(B) > 0$. Definujeme podmíněnou pravděpodobnost jevu A při B jako $P(A|B) = \frac{P(A\cap B)}{P(B)}$.

Podmíněná pravděpodobnost splňuje vlastnosti pravděpodobnosti míry.

Pozor, $P(A|B\cup C) \neq P(A|B) + P(A|C)$!

Definice 6 (Nezávislost)**.** Jevy $A, B \in \mathcal{F}$ jsou *nezávislé*, pokud $P(A\cap B) = P(A)\cdot P(B)$.

Když jevy A, B jsou nezávislé, potom $P(A|B) = \frac{P(A)P(B)}{P(B)} = P(A)$. Tato vlastnost nezávislosti už nemusí platit pro tři jevy.

Definice 7 (Vzájemná nezávislost)**.** Buďte $A_i, i \in I$ náhodné jevy, I je libovolná indexová množina. A_i jsou *vzájemně nezávislé*, pokud $\forall n \in \mathbb{N}$ $\forall i_1, \ldots, i_n \subset I$ platí

$$P(\bigcap_{j=1}^{n} A_{i_j}) = \prod_{j=1}^{n} P(A_{i_j})$$

Věta 3 (O postupném podmiňování)**.** *Mějme* $A_1, \ldots, A_n$ *náhodné jevy (prvky* $\mathcal{F}$*) takové, že* $P(\bigcap_{i=1}^{n} A_i) > 0$. *Potom*

$$P(\bigcap_{i=1}^{n} A_i) = P(A_1|\bigcap_{i=2}^{n} A_i)\cdot P(A_2|\bigcap_{i=3}^{n} A_i)\cdots P(A_{n-1}|A_n)\cdot P(A_n).$$

Důkaz. Matematickou indukcí.

První krok pro $n = 2$: $P(A_1 \cap A_2) = P(A_1|A_2)\cdot P(A_2)$.

Nyní pro $n-1 \to n$: $P(\bigcap_{i=1}^{n} A_i) = P((\bigcap_{i=1}^{n-1} A_i) \cap A_n) = P(A_n|\bigcap_{i=1}^{n-1} A_i)\cdot P(\bigcap_{i=1}^{n-1} A_i)$. $\square$

Věta 4 (Inkluze a exkluze)**.** *Mějme* $A_1, \ldots, A_n$ *náhodné jevy. Pak*

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) - \sum_{1\le i\le j\le n} P(A_i\cap A_j) + \sum_{1\le i\le j\le k\le n} P(A_i\cap A_j\cap A_k) - \cdots + (-1)^{n-1} P(\bigcap_{i=1}^{n} A_i)$$

Důkaz. Matematickou indukcí.

První krok pro $n = 2$: $A = (A \setminus B) \cup (A \cap B)$, $B = (B \setminus A) \cup (A \cap B)$.
Proto $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Nyní pro $n-1 \to n$: $P(\bigcup_{i=1}^n A_i) = P((\bigcup_{i=1}^{n-1} A_i) \cup A_n) = P(\bigcup_{i=1}^{n-1} A_i) + P(A_n) - P(\bigcup_{i=1}^{n-1} (A_i \cap A_n)) =$
$= \sum_{i=1}^{n-1} P(A_i) - \sum_{1 \leq i \leq j \leq n-1} P(A_i \cap A_j) + \ldots + (-1)^{n-2} P(\bigcap_{i=1}^{n-1} A_i) + P(A_n) -$
$- \sum_{1 \leq i \leq j \leq n-1} P(A_i \cap A_n) + \ldots + (-1)^{n-2} P(\bigcap_{i=1}^{n} A_i)$.
Nyní dáme dohromady a dostaneme $\sum_{i=1}^n P(A_i) - \sum_{1 \leq i \leq j \leq n} P(A_i \cap A_j) + \ldots + (-1)^{n-1} P(\bigcap_{i=1}^n A_i)$. □

Z vlastností množin víme, že $(\bigcup_{i=1}^n A_i)^C = \bigcap_{i=1}^n A_i^C$ a z toho $P(\bigcup_{i=1}^n A_i) = 1 - P(\bigcap_{i=1}^n A_i^C)$.

Definice 8 (Disjunktní rozklad)**.** Spočetný systém náhodných jevů $\{B_i\}_{i=1}^\infty \subset \mathcal{F}$ nazveme *disjunktním rozkladem* Ω, pokud:

1. $B_i \cap B_j = \emptyset \ \forall i \neq j$
2. $\bigcup_i B_i = \Omega$ (stačilo by $P(\bigcup_i B_i) = 1$).
3. $P(B_i) > 0 \ \forall i$

Věta 5 (O úplné pravděpodobnosti)**.** *Mějme A náhodný jev a $\{B_i\}_i$ disjunktní rozklad. Pak*

$$P(A) = \sum_i P(A|B_i) \cdot P(B_i)$$

Důkaz. Víme, že $\{A \cap B_i\}$ jsou po dvou disjunktní množiny. Potom $\bigcup_i A \cap B_i = A \cap \bigcup_i B_i = A \cap \Omega = A$.
Potom $P(A) = P(\bigcup_i A \cap B_i) = \sum_i P(A \cap B_i) = \sum_i P(A|B_i) \cdot P(B_i)$. □

Věta 6 (Bayesova)**.** *Mějme A náhodný jev a $\{B_i\}_i$ disjunktní rozklad Ω a $P(A) > 0$. Potom*

$$P(B_i|A) = \frac{P(A|B_i) \cdot P(B_i)}{\sum_j P(A|B_j) \cdot P(B_j)}.$$

Důkaz. Víme, že $P(B_i|A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(A|B_i) \cdot P(B_i)}{\sum_j P(A|B_j) \cdot P(B_j)}$ využitím věty o úplné pravděpodobnosti. □

Názvosloví... $P(B_i)$ je apriorní pravděpodobnost, $P(B_i|A)$ je aposteriorní pravděpodobnost.

Příklad. Máme tři zásuvky a 6 mincí, ze kterých jsou 3 zlaté a 3 stříbrné. Do zásuvek náhodně vložíme dvojice ZS, ZZ a SS. Ze zásuvky 2 náhodně vybereme minci náhodně a zjistíme, že je zlatá. Jaká je pravděpodobnost, že i druhá z mincí je zlatá?

Označme si B_i jako ZZ v zásuvce i. Toto je disjunktní rozklad. Jev A, náhodný výběr mince ze zásuvky 2, nastal. Potom $P(A|B_1) = \frac{1}{4}$, $P(A|B_2) = 1$, $P(A|B_3) = \frac{1}{4}$.
Spočítejme tedy $P(B_2|A) = \frac{P(A|B_2) \cdot P(B_2)}{\sum_{i=1}^3 P(A|B_i) \cdot P(B_i)} = \frac{1}{\frac{1}{4} + 1 + \frac{1}{4}} = \frac{2}{3}$!

Věta 7 (Bonferroniho nerovnost)**.** *Mějme $A_1, \ldots, A_n$ náhodné jevy. Pak*

$$P(\bigcap_{i=1}^n A_i) \geq 1 - \sum_{i=1}^n (1 - P(A_i))$$

Důkaz. $P(\bigcap_{i=1}^n A_i) = 1 - P((\bigcap_{i=1}^n A_i)^C) = 1 - P(\bigcup_{i=1}^n A_i^C) \geq 1 - \sum_{i=1}^n P(A_i^C) = 1 - \sum_{i=1}^n (1 - P(A_i))$. □

2 Náhodné veličiny a jejich rozdělení

Definice 9. Mějme $(\Omega, \mathcal{F}, P)$ pravděpodobnostní prostor. Zobrazení $X : \Omega \to \mathbb{R}$ takové, že $X^{-1}(-\infty, a] = \{\omega, X(\omega) \leq a\} \in \mathcal{F} \ \forall a \in \mathbb{R}$ se nazývá *Náhodná veličina.*

Dohoda o značení. $[X \leq A]$ se bude myslet jako $\{\omega, X(\omega) \leq a\}$.

Definice 10 (Rozdělení náhodné veličiny)**.** Buď $X : \Omega \to \mathbb{R}$ náhodná veličina. Pravděpodobnostní míra P_X definovaná na $\mathbb{R}$ předpisem $P_X(-\infty, a] = P[X \leq a]$ se nazývá *rozdělení náhodné veličiny X.*

Definice. Označme $\mathcal{B}$ nejmenší σ-algebru na $\mathbb{R}$ obsahující všechny intervaly $(-\infty, a]$. Nazýváme ji *Borelova (nebo borelovská).* $\mathcal{B}$ obsahuje všechny otevřené i uzavřené množiny.
U této množiny poté platí $\forall B \in \mathcal{B} : X^{-1}(B) \in \mathcal{F}, P_x(B) = P[X \in B]$.

Náhodná veličina X nám tedy zobrazuje $(\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B}, P_X)$.

Nabývá-li X jen spočetně mnoha hodnot ($X(\omega) \in \mathbb{S}\ \forall \omega$, $\mathbb{S}$ je spočetná podmnožina $\mathbb{R}$), potom X splňuje podmínky z definice 9 a X nazveme *diskrétní* náhodnou veličinou a typicky lze volit $\mathbb{S} = \mathbb{N}$.

Rozdělení diskrétní náhodné veličiny X je plné charakterizováno souborem čísel $\{p_i\}_{i=1}^{\infty}$, $p_i \geq 0$, $\sum_i p_i = 1$ a $P_X(A) = \sum_{i \in A} p_i$.

Příklad. Rozdělení diskrétní náhodné veličiny můžou být:

1. Alternativní (Bernoulliho). X nabývá $\{0, 1\}$, $P[X = 1] = P_x(\{1\}) = p \in (0, 1)$ a $P[X = 0] = 1 - p$
2. Binomické. Rozdělení počtu úspěchů do n nezávislých pokusů. $P[X = k] = \binom{n}{k} p^k (1 - p)^{n-k}$
3. Geometrické. Počet neúspěchů před prvním úspěchem v nezávislých pokusech. $P[X = k] = p(1 - p)^k$

Pro jaké množiny $A \subset \mathbb{R}$ umíme najít P_X?

1. $(\infty, a]$
2. $(a, b], a < b = (-\infty, b] \setminus (-\infty, a]$
3. $(a, b) = \bigcup_{n=1}^{\infty} (a, b - \frac{1}{n}]$
4. Všechny otevřené množiny
5. $A \in \mathcal{B} \Rightarrow X^{-1} \in \mathcal{F}$ pro všechny $A \in \mathcal{B}$, X je *Borelovsky měřitelná*

Definice 11 (Náhodné jevy generované X)**.** Mějme X náhodnou veličinu, označme množinu $\mathcal{F}_X$ takovou, že $\mathcal{F}_X = \{B : B = X^{-1}(A)$ pro nějakou $A \in \mathcal{B}\}$.

$\mathcal{F}_x$ je také σ-algebra náhodných jevů generovaných v X.

$\mathcal{F}_X$ je σ-algebra a $\mathcal{F}_X \subset \mathcal{F}$. Potom $P_X(A) = P[X \in A] = P(X^{-1}(A))_{\in \mathcal{F}_X}$. Levá strana je míra na $(\mathbb{R}, \mathcal{B})$, pravá strana je míra na $(\Omega, \mathcal{F})$.

Příklad. Mějme $\Omega = \{1, 2, 3, 4, 5, 6\}^2$. Dále $\mathcal{F} = 2^{\Omega}$ a $P(\{\omega\}) = \frac{1}{36}$.

Nadefinujme si $X : (\omega_1, \omega_2) \to \omega_1$. Potom $x \in \{1, 2, 3, 4, 5, 6\}$.

Podívejme se na $X^{-1}(1) = \{(1, 1), (1, 2), \ldots, (1, 6)\}$. Podobně $X^{-1}(2) = \{(2, 1), (2, 2), \ldots, (2, 6)\}$. Z toho je vidět, že X^{-1} jsou navzájem disjunktní. A proto $\mathcal{F}_x \subsetneq \mathcal{F}$ a $P_X(1) = \frac{1}{6}$.

Nyní si nadefinujme $Y : (\omega_1, \omega_2) \to \omega_1 + \omega_2$. Potom $Y^{-1}(1) = \emptyset$, dále $Y^{-1}(2) = \{(1, 1)\}$ a tak dále.

Tyto náhodné veličiny jsou různé a $\mathcal{F}_X \neq \mathcal{F}_Y$. Na jednom pravděpodobnostním prostoru tedy můžeme definovat více různých modelů.

Definice 12 (Distribuční funkce a hustota)**.** Mějme náhodnou veličinu X a P_X její rozdělení. Funkce $F_X : \mathbb{R} \to (0, 1)$ definovaná jako $F_X(x) = P_X[-\infty, x] = P[X \leq x]$ se nazývá *distribuční funkce* náhodné veličiny X a plně popisuje P_X.

V angličtině cummulative distribution function C.D.F.

Je-li X diskrétní náhodná veličina (s hodnotami v $\mathbb{N}_0$), potom funkci $p_X(k)$ takovou, že $p_X(k) = P[X = k]$ nazveme *hustotou* náhodné veličiny X vůči aritmetické míře.

V angličtině probability density function P.D.F.

Příklad. Mějme alternativní rozdělení. Potom $P[X = 1] = p = 1 - P[X = 0]$. Potom hustota má tvar $p_X(0) = 1 - p$, $p_X(1) = p$. Distribuční funkce má tvar

$$F_X(x) = \begin{cases} 0 & x \in (-infty, 0) \\ 1 - p & x \in [0, 1) \\ 1 & x \in (1, \infty) \end{cases}$$

Obecně má $F_X(a)$ pro diskrétní náhodnou veličinu X tvar $\sum_{k=0}^{\lfloor a \rfloor} p_X(k)$ pro $a \geq 0$, $p_x(k)$ je $F_x(k) - F_x(k-1)$ pro $k \in \mathbb{N}_0$.

Věta 8 (Vlastnosti distribuční funkce)**.** *Mějme náhodnou veličinu X a její distribuční funkci F_X. Pak*

1. $\lim_{x \to -\infty} F_X(x) = 0$

2. $\lim_{x \to \infty} F_X(x) = 1$

3. F_X je neklesající a zprava spojitá

Důkaz. Nejprve ukážeme, že F_X je neklesající. Z definice $F_X(x) = P[X \leq x] \leq P[X \leq x] + P[x < X \leq y]$. Tyto dva jevy jsou disjunktní. Proto se předchozí výraz rovná $P([X \leq x] \cup [x < X \leq y]) = P[X \leq y] = F_X(y)$. Podobně to platí pro $F_X(x) = P_X(-\infty, x] \leq P_X(-\infty, x) + P_X(x, y] = P_X(-\infty, y) = F_X(y)$.

Nyní ukážeme 1. Z definice distribuční funkce $\lim_{x\to-\infty} F_X(x) = \lim_{x\to-\infty} P_X(-\infty, x]$. Z monotonie můžeme tento výraz přepsat jako $\lim_{n\nearrow\infty} P_X(-\infty, -n] = \lim_{n\nearrow\infty} P[X \leq -n]$; nazvěme $[X \leq -n] = A_n$.

Je vidět, že $A_n \supset A_{n+1}$. Podle věty 2 dostaneme $\bigcap_{n=1}^{\infty} A_n = \emptyset$. Platí tedy, že $\lim_{n\to\infty} P(A_n) = 0$.

Nyní 2. Podobně $\lim_{x\to+\infty} F_X(x) = \lim_{n\nearrow\infty} F_X(n) = \lim_{n\nearrow\infty} P_X(-\infty, n) = \lim_{n\nearrow\infty} P[X \leq n] = B_n$.

Vidíme, že $B_n \subset B_{n+1}$ a $\bigcup_{n=1}^{\infty} B_n = \Omega$. Podle věty 2 se limita rovná 1.

Nakonec ukážeme spojitost zprava. $\lim_{h\to0^+} F_X(f+h) = \lim_{n\to\infty} P_x(-\infty, x] + P_X(x, x+\frac{1}{n}] = F_X(x) + \lim_{n\to\infty} P_X(x, x+\frac{1}{n}) = F_X(x)$. □

Věta 9. *Nechť F splňuje vlastnosti z věty 8. Pak existuje pravděpodobnostní prostor $(\Omega, \mathcal{F}, P)$ a náhodná veličina X takové, že F je distribuční funkcí na veličině X.*

Důkaz. Kanonickou konstrukcí. Volme $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}$ a P jako míru na $\mathcal{B}$ definovanou předpisem $P(-\infty, x] = F(x)$.

Nyní musíme umět zadefinovat $P(-a, b) = F(b_-) - F(a)$. Tímto se dostaneme na míru na $\mathcal{B}$. Spojitost P je důsledkem spojitosti zprava F.

Nyní zvolíme $X : \Omega \to \mathbb{R}$ takové, že $x(\omega) = \omega$. Potom $F_X(a) = P[X \leq a] = P(-\infty, a] = F(a)$. □

Definice 13 (Nezávislost náhodných veličin)**.** Mějme $X_1, X_2, \ldots$ náhodné veličiny definované na $(\Omega, \mathcal{F}, P)$. Nazveme je *vzájemně nezávislými*, pokud pro každou konečnou $I \subset \mathbb{N}$ a každé $\{x_i\}_{i\in I}$ platí $P(\bigcap_{i\in I}[X_i \leq x_i]) = \prod_{i\in I} P[X_i \leq x_i]$.

3 Střední hodnota a další momenty

Snahou je najít číselné charakteristiky, které vypovídají o chování náhodné veličiny, jelikož samotná náhodná míra nám moc neřekne.

Definice 14 (Střední hodnota, obecná definice)**.** Buď X náhodná veličina (definovaná na $(\Omega, \mathcal{F}, P)$). Hodnotu $EX = \int_\Omega X(\omega)\, dP(\omega)$, existuje-li výraz napravo, nazveme *střední hodnotou* X.

Definice. Buď $\mathbb{S}$ nejvýše spočetná množina taková, že $P[X \in \mathbb{S}] = 1$. Pokud navíc $p_X(s) = P[x = s] > 0\ \forall s \in \mathbb{S}$, pak $\mathbb{S}$ nazveme *nosičem* P_X.

Věta 10 (Střední hodnota diskrétní náhodné veličiny)**.** *Buď X diskrétní náhodná veličina s hodnotami v $\mathbb{S}$. Pak*

$$EX = \sum_{s\in\mathbb{S}} sP[X = s] = \sum_{s\in\mathbb{S}} sP_X(\{s\}) = \sum_{s\in\mathbb{S}} sp_X(s) \quad \textit{MLPSS}$$

Důkaz. Využijeme definici: $EX = \int_\Omega X(\omega)\, dP(\omega)$. Množinu Ω si rozdělíme na $\bigcup_{s\in\mathbb{S}}\{\omega : X(\omega) = s\} = A_s$. Tyto množiny A_s jsou navzájem disjunktní a je jich spočetně mnoho.
$\int_{\bigcup A_s} X(\omega)\, dP(\omega) = \sum_{s\in\mathbb{S}} \int_{A_s} X(\omega)\, dP(\omega) = \sum_{s\in\mathbb{S}} s \int_{A_s} dP(\omega) = \sum_{s\in\mathbb{S}} s \int_\Omega \chi_{A_s}(\omega)\, dP(\omega) = \sum_{s\in\mathbb{S}} sP(A_s) = \sum_{s\in\mathbb{S}} sP[X = s]$. □

Příklad. Zkusme najít funkci $p_X(s), s \in \mathbb{Z}$ takovou, že:

1. $p_X(s) \geq 0$
2. $\sum_{s\in\mathbb{Z}} p_X(s) = 1$
3. $\sum_{s>0} sp_X(s) = \infty$, $\sum_{s<0} sp_X(s) = -\infty$

Pro tuto funkci neexistuje střední hodnota, jelikož máme nedefinovanou sumu $(\infty - \infty)$.

Pokud EX existuje a je konečná, pak mluvíme o náhodné veličině s konečnou střední hodnotou.
Pokud $P[X \geq b] = 1$ pro b konečnou konstantu, pak EX existuje a je $> b$ (analogicky pro $P[X \leq b]$).
Pokud $\exists a, b$ konečné a $P[a \leq X \leq b] = 1$, pak EX existuje konečná a $a \leq EX \leq b$.
Můžeme definovat $E|X| = \int_\Omega |X(\omega)|\, dP(\omega)$, která existuje vždy. Pokud $E|X| < \infty$, pak $X \in L_1$ a v tom případě existuje $|EX| < \infty$.
Střední hodnota charakterizuje polohu X (P_X).

Mějme a, b reálné a transformaci $X \to a + bX$. Potom $E(a + bX) = \sum_s (a + bs)p_X(s)$. Pokud $|EX| < \infty$, můžeme toto přepsat jako $\sum_s ap_X(s) + \sum_s bsp_X(s) = a + bEX$. Dá se tedy říct, že v tomto případě se EX chová lineárně.

Definice 15 (Obecné momenty náhodné veličiny)**.** Buď $X : \Omega \to \mathbb{R}$ náhodná veličina a $g : \mathbb{R} \to \mathbb{R}$ funkce taková, že $g(x)$ je náhodná veličina. Pak $Eg(x) = \int_\Omega g(X(\omega))\,dP(\omega)$ MLPSS.

Pro diskrétní náhodné veličiny potom $Eg(x) = \sum_{s\in\mathbb{S}} g(s)P[X = s]$ MLPSS.

Definice 16. Buď X náhodná veličina. Pak

1. pro $r \in \mathbb{N}$ EX^r je r-tý moment X.
2. $E|X|^r$ je r-tý absolutní moment X. Také pro $E|x|^r < \infty$ máme $X \in L_r$.
3. pro $r \in \mathbb{N}$ $E(X - EX)^r$ je r-tý centrální moment X.
4. pro $r = 2$ máme $\operatorname{var} x = E(X - EX)^2$ rozptyl (variance) X.
5. $\mu_3 = \frac{E(X-EX)^3}{(E(X-EX)^2)^{\frac{3}{2}}}$ je šikmost rozdělení, měří asymetrii.
6. Momentová vytvořující funkce $\Psi_X(t), t \in \mathbb{R}$ je $\Psi_X(t) = Ee^{tX}$ MLPSS. Vždy existuje minimálně $\Psi_X(0) = 1$.

Poznámky ohledně momentů:

1. Pokud $r > t \geq 1$ a $E|X|^r < \infty$, pak $E|X|^t < \infty$.
2. Jak spočítat rozptyl... $\operatorname{var} X = \sum_{s\in\mathbb{S}}(s - EX)^2 P[X = s]$.
   Nebo také $\operatorname{var} X = E(X^2 - 2X \cdot EX + (EX)^2) = EX^2 - E(2X \cdot EX) + E(EX)^2 = EX^2 - (EX)^2$.
   Nebo také $\operatorname{var} X = E(X(X - 1)) - EX(EX - 1)$.
3. $\operatorname{var} X \geq 0$. Pokud $\operatorname{var} X = 0$, potom $s = EX, p_X(EX) = 1$, tedy $P[X = EX] = 1$.

Věta 11 (momenty a ψ_X)**.** *Buď X náhodná veličina a ψ_X její momentová vytvořující funkce. Nechť $\exists\delta > 0$ takové, že $\psi_X(t)$ existuje konečná $\forall|t| < \delta$. Pak $\forall r \in \mathbb{N}$ $EX^r = \psi_X^{(r)}(0)$.*

Důkaz. V knize Pravděpodobnost a matematická statistika (Dupáč, Hušková). □

Věta 12 (M.v.f a charakterizace rozdělení)**.** *Nechť X a Y jsou náhodné veličiny. Nechť $\exists\delta > 0$ tak, že:*

1. *$\psi_X(t)$, $\psi_Y(t)$ existuje konečné $\forall|t| < \delta$*
2. *$\psi_X(t) = \psi_Y(t)\forall|t| < \delta$*

Pak $P_X = P_Y$

Věta 13 (Jensenova nerovnost)**.** *Buď X náhodná veličina s konečnou střední hodnotou EX. Buď φ konvexní funkce. Pak $E\varphi(x) \geq \varphi(EX)$.*

Důkaz. Pro $\varphi \in C^2$. Potom můžeme použít Taylorův rozvoj. Tedy $\varphi(X) = \varphi(EX) + \varphi'(EX)(X - EX) + \frac{\varphi''(c)}{2}(X - EX)^2$. Použijeme střední hodnotu. Dostáváme $E\varphi(x) = \varphi(EX) + \varphi'(EX) \cdot E(X - EX)_{=0} + \frac{\varphi''(c)}{2} \cdot \operatorname{var} X \geq \varphi(EX)$. □

Důkaz podle Matěje Konečného. Z definice konvexity máme

$$\varphi(X) \geq \varphi(EX) + \varphi'(EX) \cdot (X - EX),$$

a protože očekávaná hodnota zachovává nerovnosti, tak použitím E na tuhle nerovnost dostaneme přesně to, co chceme. □

Věta 14. *Nebyla přednesena.*

4 Náhodné vektory a jejich rozdělení

Definice 17 (Náhodný vektor)**.** Zobrazení $\boldsymbol{X} : \Omega \to \mathbb{R}^d$, kde $(\Omega, \mathcal{F}, P)$ je pravděpodobnostní prostor, $d \in \mathbb{N}, d \geq 2$ takové, že $\{\omega : \boldsymbol{X}(\omega) \leq a\} \in \mathcal{F}\ \forall a \in \mathbb{R}^d$ nazveme *náhodný vektor.*

Diskrétní náhodný vektor nabývá nejvýše spočetně mnoha hodnot. Neřekneme-li jinak, budou tyto hodnoty vždy podmnožinou $\mathbb{N}_0^d$.

Definice 18. Buď $\boldsymbol{X}$ d-rozměrný náhodný vektor. Pravděpodobnostní míra $P_{\boldsymbol{X}}$ definovaná na $\mathbb{R}^d$ předpisem $P_{\boldsymbol{X}}(\prod_{i=1}^d(-\infty, a_i]) = P[\boldsymbol{X} \leq \boldsymbol{a}]\ \forall a \in \mathbb{R}^d = P(\bigcap_{i=1}^d [X_i \leq a_i])$ nazveme rozdělení náhodného vektoru.

Věta 15 (Rozdělení diskrétní náhodné veličiny)**.** *Buď $\boldsymbol{X}$ diskrétní náhodný vektor a $P_{\boldsymbol{X}}$ jeho rozdělení. Pak existuje funkce $p_{\boldsymbol{X}} : \mathbb{N}^d \to [0,1]$ jednoznačně daná taková, že*

$$P_{\boldsymbol{X}}(\prod_{i=1}^{d}(-\infty, a_i]) = \sum_{\boldsymbol{z} \leq \boldsymbol{a}} p_{\boldsymbol{X}}(\boldsymbol{z})$$

Definice 19 (Distribuční funkce náhodného vektoru)**.** Buď $\boldsymbol{X}$ náhodný vektor a $P_{\boldsymbol{X}}$ jeho rozdělení. $F_{\boldsymbol{X}} : \mathbb{R}^d \to [0,1]$ definovaná jako $F_{\boldsymbol{X}}(\boldsymbol{a}) = P[\boldsymbol{X} \leq \boldsymbol{a}]$ se nazývá *distribuční funkce* náhodného vektoru $\boldsymbol{X}$.

Definice 20 (Marginální rozdělení)**.** Buď $\boldsymbol{X}$ náhodný vektor a $P_{\boldsymbol{X}}$ jeho rozdělení takové, že $P_{X_i}(-\infty, a] = \lim_{a_j \to \infty, j \neq i} P_{\boldsymbol{X}}(X_{j=1}^d(-\infty, a_j])$ se nazývá *marginální rozdělení* X_i a $F_{X_i} = \lim_{a_j \to \infty, j \neq i} F_{\boldsymbol{X}}(\boldsymbol{a})$ se nazývá jeho marginální distribuční funkce X_i.

Vezměme si, $\boldsymbol{X} : \Omega \to \mathbb{R}^d$. Potom $\boldsymbol{X} = (X_1, X_2, \ldots, X_d)$, kde

Terminologie. $P_{\boldsymbol{X}}, F_{\boldsymbol{X}}$ jsou sdružené rozdělení, distribuční funkce, zatímco P_{X_i}, F_{X_i} jsou marginální rozdělení. Nemáme však speciální termín pro $P_{(X_1, X_2, X_8)}$ subvektor.

Věta 16 (O Marginálním rozdělení)**.** *Máme-li náhodný vektor $\boldsymbol{X}$ a jeho rozdělení $P_{\boldsymbol{X}}$, pak marginální rozdělení složek $X_1, \ldots, X_d$ jsou jednoznačně určeny $P_{\boldsymbol{X}}$.*
Naopak ne!

Příklad. Hoďme dvěma kostkami. Jejich výsledky jsou A, B. Dále mějme $C = A - 1$ pro B sudé a $C = A + 1$ pro B liché. Zapišme si je do tabulky:

$A \backslash B$	1	2	3	4	5	6	
1	1/36	1/36	1/36	1/36	1/36	1/36	$\frac{1}{6}$
2	1/36	1/36	1/36	1/36	1/36	1/36	$\frac{1}{6}$
3	1/36	1/36	1/36	1/36	1/36	1/36	$\frac{1}{6}$
4	1/36	1/36	1/36	1/36	1/36	1/36	$\frac{1}{6}$
5	1/36	1/36	1/36	1/36	1/36	1/36	$\frac{1}{6}$
6	1/36	1/36	1/36	1/36	1/36	1/36	$\frac{1}{6}$
	1/6	1/6	1/6	1/6	1/6	1/6	

$A \backslash C$	1	2	3	4	5	6	
1	0	1/12	0	0	0	1/12	$\frac{1}{6}$
2	1/12	0	1/12	0	0	0	$\frac{1}{6}$
3	0	1/12	0	1/12	0	0	$\frac{1}{6}$
4	0	0	1/12	0	1/12	0	$\frac{1}{6}$
5	0	0	0	1/12	0	1/12	$\frac{1}{6}$
6	1/12	0	0	0	1/12	0	$\frac{1}{6}$
	1/6	1/6	1/6	1/6	1/6	1/6	

Vidíme, že pro náhodné vektory (A, B) a (A, C) jsou jejich marginální rozdělení stejná, ale uvnitř se chovají úplně jinak.

Zaveďme si značení: $\boldsymbol{a} < \boldsymbol{b}$ právě, když $a_i < b_i\ \forall i = 1, \ldots, d$. Dále nechť $\Delta_k(\boldsymbol{a}, \boldsymbol{b})$ je množina těch $\boldsymbol{c}$, pro které existuje právě k indexů $i_1, i_2, \ldots, i_k$ takových, že $c_{i_j} = a_{i_j}$ a pro zbytek indexů je $c_l = b_l$.
Tedy například pro $\boldsymbol{a} = (a_1, a_2, a_3), \boldsymbol{b} = (b_1, b_2, b_3)$ je $\Delta_1(\boldsymbol{a}, \boldsymbol{b}) = \{(a_1, b_2, b_3), (b_1, a_2, b_3), (b_1, b_2, a_3)\}$.

Věta 17 (Vlastnosti distribuční funkce)**.** *Buď $F_{\boldsymbol{X}}$ sdružená distribuční funkce náhodného vektoru $\boldsymbol{X}$. Pak:*

- $\lim_{a_i \to -\infty} F_{\boldsymbol{X}}(\boldsymbol{a}) = 0$ *pro libovolné i.*
- $\lim_{a_i \to \infty} F_{\boldsymbol{X}}(\boldsymbol{a}) = 1$ *pro každé i.*
- *V každé složce argumentu je $F_{\boldsymbol{X}}$ zprava spojitá a neklesající*
- $\forall \boldsymbol{a} < \boldsymbol{b}$ *platí* $\sum_{k=0}^d (-1)^k \sum_{\boldsymbol{c} \in \Delta k(\boldsymbol{a}, \boldsymbol{b})} F_{\boldsymbol{X}}(\boldsymbol{c}) \geq 0$

Podívejme se na čtvrtou podmínku pro $\boldsymbol{a} = (a_1, a_2), \boldsymbol{b} = (b_1, b_2)$. Pak tato podmínka říká, že pro $\boldsymbol{a} < \boldsymbol{b}$ platí, že $(b_1, b_2) - ((a_1, b_2) + (b_1, a_2)) + (a_1, a_2) \geq 0$. Jedná se tedy o výřez na boxu $(a_1, b_1) \times (a_2, b_2)$.

Věta 18 (Charakterizace distribuční funkce)**.** *Nechť $F : \mathbb{R}^d \to \mathbb{R}$ splňuje vlastnosti z věty 17. Pak existují $(\Omega, \mathcal{F}, P)$ a $\boldsymbol{X} : \Omega \to \mathbb{R}^d$ takové, že F je distribuční funkcí $\boldsymbol{X}$.*

Důkaz. Stejně, jako u věty 9. □

Náhodné veličiny $X_1, \dots, X_d$ jsou nezávislé, když $\forall A_1, \dots, A_d \in \mathcal{B}$ platí $P(\bigcap_{i=1}^d [X_i \in A_i]) = \prod_{i=1}^d P[X_i \in A_i]$.

Věta 19 (Charakterizace nezávislosti)**.** *Buď $\boldsymbol{X} = (X_1, \dots, X_d)$ náhodný vektor a $F_{\boldsymbol{X}}$ jeho sdružená distribuční funkce. Pak $X_1, \dots, X_d$ jsou nezávislé náhodné veličiny, pokud $\forall a_1, \dots, a_d \in \mathbb{R}$ platí*

$$F_{\boldsymbol{X}}(a_1, \dots, a_d) = \prod_{i=1}^d F_{X_i}(a_i).$$

Pro diskrétní náhodný vektor je nezávislost $X_1, \dots, X_d$ ekvivalentní tomu, že $p_{\boldsymbol{X}}(a_1, \dots, a_d) = \prod_{i=1}^d p_{X_i}(a_i)$.

5 Náhodné vektory a momenty

Definice 21 (Střední hodnota náhodného vektoru)**.** Buď $\boldsymbol{X}$ náhodný vektor. Potom definujeme:

1. $E\boldsymbol{X} = (EX_1, \dots, EX_d)$
2. Nechť $g : \mathbb{R}^d \to \mathbb{R}$ (taková, že $g(\boldsymbol{X})$ je náhodná veličina). Pak definujeme $Eg(\boldsymbol{X}) = \int_\Omega g(\boldsymbol{X}(\omega))\, dP(\omega)$, pokud všechny integrály existují.

Je-li $\boldsymbol{X}$ *diskrétní*, pak $Eg(\boldsymbol{X}) = \sum_{\boldsymbol{z} \in \mathbb{N}_0^d} g(\boldsymbol{z}) P[\boldsymbol{X} = \boldsymbol{z}]$, existuje-li řada napravo.

Když si vezmeme funkci $g_i : \mathbb{R}^d \to \mathbb{R}$, která splňuje $g_i(\boldsymbol{X}) = X_i$, potom $Eg_i(\boldsymbol{X}) = EX_i$.

Věta 20 (Linearita střední hodnoty)**.** *Buď $\boldsymbol{X} = (X_1, \dots, X_d)$ náhodný vektor a konstanty $a \in \mathbb{R}, b_1, \dots, b_d \in R$ a $E|X_i| < \infty$ $\forall i$. Pak*

$$E\left(a + \sum_{i=1}^d b_i X_i\right) = a + \sum_{i=1}^d b_i E X_i.$$

Důkaz. Jen pro diskrétní náhodný vektor, indukcí.

Vezměme si (X_1, X_2) a a, b_1, b_2. Potom $E(a + b_1 X_1 + b_2 X_2) = \sum_{\boldsymbol{z} \in \mathbb{N}_0^d} (a + b_1 z_1 + b_2 z_2) P[X_1 = z_1, X_2 = z_2] = \sum_{\boldsymbol{z} \in \mathbb{N}_0^d} aP[X_1 = z_1, X_2 = z_2] + \sum_{\boldsymbol{z} \in \mathbb{N}_0^d} b_1 z_1 P[X_1 = z_1, X_2 = z_2] + \sum_{\boldsymbol{z} \in \mathbb{N}_0^d} b_2 z_2 P[X_1 = z_1, X_2 = z_2]$.

Pro první sumu platí, že $\sum_{\boldsymbol{z} \in \mathbb{N}_0^d} aP[X_1 = z_1, X_2 = z_2] = a$.

Pro druhou sumu dostaneme $\sum_{\boldsymbol{z} \in \mathbb{N}_0^d} b_1 z_1 P[X_1 = z_1, X_2 = z_2] = \sum_{z_1=0}^\infty \sum_{z_2=0}^\infty b_1 z_1 P[X_1 = z_1, X_2 = z_2] = \sum_{z_1=0}^\infty b_1 z_1 P[X_1 = z_1, X_2 = z_2]$. Můžeme použít větu 5, kde $\bigcup_{z_2=0}^\infty [X_2 = z_2] = \Omega$. Nakonec se suma bude rovnat $b_1 \sum_{z_1=0}^\infty z_1 P[X_1 = z_1] = b_1 EX_1$. Analogicky pro třetí sumu.

Celá suma se tedy nakonec rovná $a + b_1 EX_1 + b_2 EX_2$.

Takto můžeme indukcí projít všechny dimenze. □

Poznámka. Předpoklad $\boldsymbol{X} = (X_1, \dots, X_d)$ je náhodný vektor a zaručuje, že $X_1, \dots, X_d$ jsou definované na stejném pravděpodobnostním prostoru a mají sdružené rozdělení $P_{\boldsymbol{X}} = P_{(X_1, \dots, X_d)}$.

Věta 21 (Střední hodnota součinu)**.** *Buď $\boldsymbol{X} = (X_1, \dots, X_d)$ náhodný vektor, $E|X_i| < \infty$ $\forall i$. Jsou-li $X_1, \dots, X_d$* nezávislé*, pak*

$$E \prod_{i=1}^d X_i = \prod_{i=1}^d EX_i.$$

Důkaz. Jen pro diskrétní náhodný vektor, indukcí. Postup podobný, jako u předchozí věty.

Máme (X_1, X_2). Potom $EX_1 X_2 = \sum_{\boldsymbol{z} \in \mathbb{N}_0^d} z_1 z_2 P[X_1 = z_1, X_2 = z_2] = \sum_{\boldsymbol{z} \in \mathbb{N}_0^d} z_1 P[X_1 = z_1, X_2 = z_2] \cdot z_2 P[X_1 = z_1, X_2 = z_2]$.

Máme tedy součet $\sum_{z_1=0}^\infty \sum_{z_2=0}^\infty z_1 P[X_1 = z_1, X_2 = z_2] \cdot z_2 P[X_1 = z_1, X_2 = z_2]$. Díky nezávislosti se tato suma rovná $\sum_{z_1=0}^\infty z_1 P[X_1 = z_1, X_2 = z_2] \cdot \sum_{z_2=0}^\infty z_2 P[X_1 = z_1, X_2 = z_2] = EX_1 \cdot EX_2$. □

Dříve jsme zjistili, že X_1, X_2 jsou nezávislé právě, když $F_{(X_1,X_2)(X_1,X_2)} = F_{X_1}(X_1) F_{X_2}(X_2)$. Nyní díky větě 21 víme, že X_1, X_2 jsou nezávislé $\Rightarrow EX_1 X_2 = EX_1 EX_2$. Opačná implikace však obecně neplatí.

Postačí si představit náhodné veličiny $X = \{-1, 0, 1\}, Y = \{1, 2\}$, pro které platí: $EX_1 X_2 = 0$, $EX_1 = 0$. Tedy $EX_1 X_2 = EX_1 EX_2$. Samotné pravděpodobnosti ale jsou $P[X = -1] = \frac{1}{3}, P[Y = 1] = \frac{1}{3}$. Avšak $P[X = -1, Y = 1] = 0 \neq \frac{1}{9}$. Nejsou tedy nezávislé.

Definice 22 (Kovariance náhodných veličin)**.** Buď (X_1, X_2) náhodný vektor takový, že $\operatorname{var} X_1 < \infty, \operatorname{var} X_2 < \infty$. Definujeme *kovarianci* $\operatorname{cov}(X_1, X_2) = E(X_1 - EX_1)(X_2 - EX_2)$.

Co tato kovariance znamená? $X_1 - EX_1, X_2 - EX_2$ je odchylka od střední hodnoty. $(X_1 - EX_1)(X_2 - EX_2)$ je kladné, když násobíme odchylky stejného znaménka, či záporné, když odchylky mají různé znaménko. Dále tento součin říká, jak jsou obě veličiny daleko.

Co se stane, když si vezmeme $E(X_1 - EX_1)(X_2 - EX_2)$? Pokud je kladná, můžeme očekávat, že obě hodnoty budou růst. Například u vysokého člověka můžeme očekávat, že bude mít i vysokou hmotnost. Když je záporná, je tomu naopak.

Jak se dá kovariance počítat? Můžeme využít přímou definici. Ta však není moc pohodlná.

Podobně jako pro rozptyl můžeme tento vzoreček zjednodušit: $E\big((X_1 - EX_1)(X_2 - EX_2)\big) = E(X_1X_2 - X_1EX_2 - X_2EX_1 + EX_1EX_2)$.

Díky větě 20 můžeme tento výraz rozepsat jako $EX_1X_2 - E(X_1EX_2) - E(X_2EX_1) + E(EX_1EX_2) = EX_1X_2 - EX_1EX_2$.

Důsledkem věty 21 je, že nezávislé X_1, X_2 mají nulovou kovarianci.

Platí, že $\operatorname{cov}(X_1, X_1) = \operatorname{var} X_1$ a kovariance je symetrická.

Věta 22 (Rozptyl součtu náhodných veličin)**.** *Buď $\boldsymbol{X} = (X_1, \dots, X_d)$ náhodný vektor, $\operatorname{var} x_i < \infty\ \forall i$. Pak*

$$\operatorname{var}\left(\sum_{i=1}^d X_i\right) = \sum_{i=1}^d \operatorname{var} X_i + \sum_{1\le i\ne j\le d} \operatorname{cov}(X_i, X_j) = \sum_{i=1}^d \operatorname{var} X_i + 2\sum_{1\le i<j\le d} \operatorname{cov}(X_i, X_j).$$

Speciálně, jsou-li $X_1, \dots, X_d$ nezávislé, pak

$$\operatorname{var}\left(\sum_{i=1}^d X_i\right) = \sum_{i=1}^d \operatorname{var} X_i.$$

Důkaz. Podle definice rozepíšeme a postupnou úpravou: $\operatorname{var}\left(\sum_{i=1}^d X_i\right) = E\left(\sum_{i=1}^d X_i - E\left(\sum_{i=1}^d X_i\right)\right)^2 = E\left(\sum_{i=1}^d (X_i - EX_i)^2 + 2\sum_{1\le i<j\le d}(X_i - EX_i)(X_j - EX_j)\right) = \sum_{i=1}^d \operatorname{var} X_i + 2\sum_{1\le i<j\le d} \operatorname{cov}(X_i, X_j)$. $\square$

Příklad. Mějme X_1, X_2 nezávislé, kde $\sigma^2 = \operatorname{var} X_1 = \operatorname{var} X_2$. Podle věty 22:

1. $\operatorname{var}(X_1 + X_2) = \operatorname{var} X_1 + \operatorname{var} X_2 = 2\sigma^2$
2. $\operatorname{var}(X_1 - X_2) = \operatorname{var} X_1 + \operatorname{var} X_2 = 2\sigma^2$
3. $\operatorname{var}(X_1 + X_1) = \operatorname{var} X_1 + \operatorname{var} X_1 + 2\operatorname{cov}(X_1, X_1) = 4\operatorname{var} X_1 = 4\sigma^2$
4. $\operatorname{var}(X_1 - X_1) = \operatorname{var} X_1 + \operatorname{var} X_1 - 2\operatorname{cov}(X_1, X_1) = 0$

Již umíme spočítat $\operatorname{cov}(X_i, X_j)$. Představme si, že $Z_i = 1000X_i$ (kg $\to$ g). Potom ale získáváme rovnost $\operatorname{cov}(Z_i, X_j) = 1000\operatorname{cov}(X_i, X_j)$, což je mnohem vyšší číslo.

Chtěli bychom tedy, aby kovariance byla bezrozměrná, tedy chceme $\varphi(X) : \varphi(aX) = |a|\varphi(X)$.

Potom $\frac{\operatorname{cov}(aX, bY)}{\varphi(aX)\varphi(bY)} = \frac{\operatorname{cov}(X,Y)}{\varphi(X)\varphi(Y)}$.

Definice. *Korelací* nazveme $\operatorname{corr}(X, Y) = \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{var} X \operatorname{var} Y}}$.

Definice 23 (Varianční a korelační matice)**.** Buď $\boldsymbol{X} = (X_1, \dots, X_d)$ náhodný vektor, $\operatorname{var} X_i < \infty\ \forall i = 1, \dots, d$. Potom varianční matice je $\operatorname{Var} \boldsymbol{X} = \{\operatorname{cov}(X_i, X_j)\}_{i,j}$ a korelační matice je $\operatorname{Corr} \boldsymbol{X} = \{\operatorname{corr}(X_i, X_j)\}_{i,j}$.

Varianční matice $\operatorname{Var} \boldsymbol{X}$ má díky definici na diagonále právě $\operatorname{var} X_i$. Podobně korelační matice má na diagonále samé jedničky.

Značení.

- $\operatorname{corr}(X, Y) = \rho_{X,Y}$
- $\operatorname{var} X = \sigma_X^2$
- $\operatorname{cov}(X, Y) = \sigma_X \sigma_y \rho_{X,Y}$
- $EX = \mu_X$

Věta 23 (Vlastnosti kovariance a korelace)**.** *Buďte* $\boldsymbol{X}$ *náhodný vektor s variační matici* $\operatorname{Var}\boldsymbol{X}$ *(existuje, konečné prvky). Dále mějme* X, Y *náhodné veličiny s konečným rozptylem. Potom:*

1. $-1 \leq corr(X,Y) \leq 1$, $\operatorname{corr}(X,X) = 1$
2. $|\operatorname{corr}(X,Y)| = 1 \Leftrightarrow \exists a \neq 0, b \in \mathbb{R} : X = aY + b$
3. $\operatorname{cov}(aX + c, bY + d) = ab\operatorname{cov}(X,Y)$
4. $\operatorname{corr}(aX + c, bY + d) = \operatorname{sgn}(ab)\operatorname{corr}(X,Y)$
5. *jsou-li* X, Y *nezávislé,* $\operatorname{cov}(X,Y) = 0$
6. $\operatorname{Var}\boldsymbol{X}, \operatorname{Corr}\boldsymbol{X}$ *jsou symetrické pozitivně semidefinitní*
7. $\forall A \in \mathbb{R}^{l\times d}, B \in \mathbb{R}^l, \boldsymbol{X} \in \mathbb{R}^d : \operatorname{Var}(A\boldsymbol{X} + B) = A(\operatorname{Var}\boldsymbol{X})A^T$

Důkaz. Jen některé body.

1. Máme vlevo $E((X - EX)(Y - EY))$, vpravo $\sqrt{E(X - EX)^2 E(Y - EY)^2}$.

   Pro diskrétní náhodné veličiny získáváme vlevo $\sum_u \sum_v (u - EX)(v - EY)P[X = u, Y = v]$, vpravo $\left(\sum_u \sum_v (u - EX)^2 P[X = u, Y = v] \cdot \sum_u \sum_v (v - EY)^2 P[X = u, Y = v]\right)^{\frac{1}{2}}$.

   Nyní využijeme Cauchyho-Schwarzovy nerovnosti. Díky ní poté platí, že levá strana $\leq$ pravá strana.

   Pro integrály existuje zobecnění zvané Hölderova nerovnost $\to$ z ní plyne platnost pro všechny náhodné veličiny X, Y s konečnými rozptyly.

2. Je-li $X = aY + B$, pod ze bodu 3 a 4 získáváme, že $\operatorname{corr}(X,Y) = \operatorname{corr}(aY + B, Y) = \operatorname{sgn}(a)\operatorname{corr}(Y,Y) = \pm 1$.

   Naopak, je-li $|\operatorname{corr}(X,Y)| = 1$, pak C-S nerovnost vychází jako rovnost a to je možné jen, pokud $X = aY + b$.

3. Přímo z definice $E((aX + c - E(aX + c))(bY + d - E(bY + d))) = E((aX + c - aEX))(bY + bEY)) = ab\,E((X - EX)(Y - EY))$

4. Víme, že $\operatorname{var}(aX) = a^2 \operatorname{var} X$ a $\operatorname{var}(bY) = b^2 \operatorname{var} Y$. Potom $\frac{ab\operatorname{cov}(X,Y)}{\sqrt{a^2 \operatorname{var} X b^2 \operatorname{var} Y}} = \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{var} X \operatorname{var} Y}}$

6. $\forall a \in \mathbb{R}^d : a^T \operatorname{Var}\boldsymbol{X} a \geq 0$. Po rozepsání $a^T \operatorname{Var}\boldsymbol{X} a = \sum_{i=1}^d \sum_{j=1}^d a_i a_j \operatorname{cov}(X_i, X_j) = \sum_{i=1}^d \operatorname{var}(a_i X_i) + 2\sum\sum_{1\leq i<j\leq d} \operatorname{cov}(a_i X_i, a_j X_j) = \operatorname{var}(a^T \boldsymbol{X}) \geq 0$.

   Podobně $\operatorname{Corr}\boldsymbol{X} = B\operatorname{Var}\boldsymbol{X} B^T$, kde $B_{i,i} = (\operatorname{var} X_i)^{-1}$. Potom $a^t \operatorname{Corr}\boldsymbol{X} a = a^T B \operatorname{var}\boldsymbol{X} B^T a \geq 0$.

7. Máme $A\boldsymbol{X}$. Víme $\operatorname{var} a^t X = a^T \operatorname{var}\boldsymbol{X} a$. Podobně, jako u předchozího bodu $\operatorname{cov}(a^T \boldsymbol{X}, b^T \boldsymbol{Y}) = a^T \operatorname{Var}\boldsymbol{X} b$.

□

Korelace tedy měří míru linearity závislosti X a Y.
Důsledkem sedmého bodu je, že $\operatorname{cov}(\sum_{i=1}^k X_i, \sum_{j=1}^l Y_j) = \sum_{i=1}^k \sum_{j=1}^l \operatorname{cov}(X_i, Y_j)$ a $\operatorname{cov}(X + Y, Z) = \operatorname{cov}(X,Z) + \operatorname{cov}(Y,Z)$. Tedy kovariance je bilineární forma.

Řekněme, že $\operatorname{corr}(x,y) = 1$. Potom

$$\operatorname{Corr}(X,Y) = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \qquad \operatorname{Var}(X,Y) = \begin{pmatrix} \operatorname{var} X & -\sqrt{\operatorname{var} X \operatorname{var} Y} \\ -\sqrt{\operatorname{var} X \operatorname{var} Y} & \operatorname{var} Y \end{pmatrix}$$

Po vynásobení $\operatorname{Var}(X,Y)$ vektorem $(\sqrt{\operatorname{Var} Y}, \sqrt{\operatorname{var} X})$ dostaneme nulový vektor, tedy matice je singulární. Proto $\exists u, v \neq 0$ taková, že $(u,v)^T \operatorname{Var}(X,Y)(u,v) = 0$, tedy $\operatorname{var}(uX + vY) = 0$, neboli $uX + vY$ je konstanta. Tím jsme jinak dokázali implikaci $\Rightarrow$ v bodu 2 věty 23.

Definice 24. Náhodné veličiny X, Y takové, že $\operatorname{cov}(X,Y) = 0$ nazveme *nekorelované.*

Víme, že když jsou X, Y nezávislé, pak jsou X, Y nekorelované. Naopak to však obecně **neplatí**! Dobrým příkladem, proč to neplatí, jsou náhodné veličiny X, Y, kde $Y = X^2$.

Transformace

Asi nejčastější transformace, kterou používáme, je $\boldsymbol{X} \to \sum_{i=1}^{d} X_i$.

Transformace je tedy zobrazení $\varphi : \mathbb{R}^d \to \mathbb{R}^l$. My potřebujeme, aby $\varphi(\boldsymbol{X})$ byl opět náhodný vektor, který splňuje definici 17.

Mějme tedy vektor $\boldsymbol{Y} = \varphi(\boldsymbol{X})$. Jaké má tento vektor rozdělení?

Věta 24. *Buď $\boldsymbol{X}$ d-rozměrný náhodný vektor a $\varphi : \mathbb{R}^d \to \mathbb{R}^l$ (slušně vychovaná) funkce. Potom $\boldsymbol{Y} = \varphi(\boldsymbol{X})$ s distribuční funkcí*

$$F_{\boldsymbol{Y}}(u) = P_{\boldsymbol{X}}\left(\varphi^{-1}\left(\prod_{i=1}^{l}(-\infty, u_i]\right)\right).$$

Je-li $\boldsymbol{X}$ diskrétní náhodný vektor, pak $\boldsymbol{Y}$ je též diskrétní náhodný vektor a

$$P[\boldsymbol{Y} = \boldsymbol{u}] = \sum_{\boldsymbol{v}:\varphi(\boldsymbol{v})=\boldsymbol{u}} P[\boldsymbol{X} = \boldsymbol{v}]$$

Máme tedy dva náhodné vektory $\boldsymbol{X} : \Omega \to \mathbb{R}^d$ a $\boldsymbol{Y} : \Omega \to \mathbb{R}^l$ takové, že $Y = \varphi(\boldsymbol{X})$. Nás zajímá $F_{\boldsymbol{Y}}(\boldsymbol{u}) = P(\{\omega : \boldsymbol{Y}(\omega) \leq \boldsymbol{u}\}) = P(\{\omega : \boldsymbol{Y}(\omega) \in A\}) = P(\{\omega : \boldsymbol{X}(\omega) \in \varphi^{-1}(A)\}) = P_{\boldsymbol{X}}\left(\varphi^{-1}(A)\right)$, kde $A = \prod_{i=1}^{d}(-\infty, u_i]$.

Věta 25 (Rozdělení součtu a součinu diskrétního náhodného vektoru)**.** *Buďte X, Y diskrétní náhodné veličiny (na stejném $(\Omega, \mathcal{F}, P)$), pak:*

1. *$Z = X + Y$ je také diskrétní náhodná veličina a platí $P[Z = v] = \sum_u P[X = u, Y = v - u]$.*
2. *Pokud jsou X, Y kladné (tedy $P[X > 0] = P[Y > 0] = 1$), pak $V = XY$ je také kladná diskrétní náhodná veličina a $P[V = z] = \sum_u P[X = u, Y = \frac{z}{u}]$.*

   pro obecné d.n.v je třeba dát pozor na $ab = -(a)(-b)$ a $0b = a0 = 0$.

Důkaz. Podle věty o úplné pravděpodobnosti $P[Z = v] = \sum_u P[Z = v, X = u] = \sum_u P[X = u, X + Y = v] = \sum_u P[X = u, Y = v - u]$. Druhý bod analogicky.
Pro více, než dvě veličiny můžeme postupovat indukcí. □

Věta 26 (Momentová vytvořující funkce součtu)**.** *Buďte $X_1, \dots, X_n$ nezávislé náhodné veličiny a Y jejich součet. Pak*

$$\psi_Y(t) = \prod_{i=1}^{n} \psi_{X_i}(t).$$

Důkaz. Z nezávislosti plyne, že $\psi_Y(t) = Ee^{tY} = Ee^{t\sum X_i} = E\prod e^{tX_i}$. Z nezávislosti poté máme $\prod \psi_{X_i}(t)$. □

Mějme náhodný vektor $\boldsymbol{X}$ a $\boldsymbol{Y} = \varphi(\boldsymbol{X})$. Chceme znát $E\boldsymbol{Y}$. Můžeme třeba určit rozdělení $\boldsymbol{Y}$ a pak počítat střední hodnotu podle definice. Toto však nemusí být jednoduché.

Druhá možnost $Eg(\boldsymbol{Y}) = Eg(\varphi(\boldsymbol{X})) = Eh(\boldsymbol{X})$. Poté není třeba určovat rozdělení, je však potřeba najít h.

Příklad. Mějme $Y = \sum_{i=1}^{d} X_i$. Najít rozdělení součtu může být těžké a lze v něm nadělat spoustu chyb. Avšak $E\sum_{i=1}^{d} X_i$ je vždy $\sum_{i=1}^{d} EX_i$, a to díky linearitě střední hodnoty (věta 20).

6 Absolutně spojité náhodné vektory

Definice 17,TODO[18,19,20] zde stále platí.

Definice 25 (Absolutně spojitý (AC) náhodný vektor)**.** Náhodný vektor $\boldsymbol{X} = (X_1, \dots, X_d)$ je *absolutně spojitý* (nebo jeho rozdělení je absolutně spojité), existuje-li nezáporná funkce $f : \mathbb{R}^d \to \mathbb{R}$ taková, že

$$F_{\boldsymbol{X}} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(t_1, \dots, t_d)\, \mathrm{d}t_d \dots \mathrm{d}t_1$$

Funkce f se nazývá sdruženou hustotou rozdělení $\boldsymbol{X}$.

Pro diskrétní náhodný vektor jsme hustotou označovali $p_{\boldsymbol{X}}(\boldsymbol{u}) = P[\boldsymbol{X} = \boldsymbol{u}]$. Pro AC náhodný vektor však platí $P[\boldsymbol{X} = \boldsymbol{u}] = 0\ \forall \boldsymbol{u} \in \mathbb{R}^d$.

Pravděpodobnost $P[\boldsymbol{a} \leq \boldsymbol{X} \leq \boldsymbol{b}] = \int_{a_1}^{a_n} \cdots \int_{a_d}^{b_d} f(t_1, \ldots, t_d)\, \mathrm{d}t_d \ldots \mathrm{d}t_1$.

Zhruba řečeno $f_{\boldsymbol{X}}(\boldsymbol{u}) \approx P[\boldsymbol{X} \in U(\boldsymbol{u}, \delta)]/|U(\boldsymbol{u}, \delta)|$, kde $U(\boldsymbol{u}, \delta)$ je dostatečně „malé“ okolí $\boldsymbol{u}$.

Jestliže máme f spojitou funkci, která se však nedá zapsat jako integrál, potom náhodné rozdělení definované touto funkcí je *singulárně spojité*. Takové rozdělení se však chová dost ošklivě.

V reálném světě nebude platit, že všechny vektory jsou buďto celé diskrétní, nebo absolutně spojité. Můžou existovat vektory se složkami různých „typů“, například vektor (X, Y), kde X je diskrétní veličina, zatímco Y je spojitá.

Příklad. Představme si, že máme pozorování směru větru z meteorologie. Nejčastěji vítr fouká směrem na západ. Potom tento náhodný vektor není ani absolutně spojitý (nelze napsat četnost na kružnici jako integrál), ani diskrétní (plocha kružnice je nulová).

Věta 27 (Marginální rozdělení AC náhodného vektoru)**.** *Buď $\boldsymbol{X} = (X_1, \ldots, X_d)$ AC náhodný vektor se sdruženou hustotou $f_{\boldsymbol{X}}$. Pak marginální distribuční funkce F_{X_i} je daná*

$$F_{X_i}(t) = \int_{-\infty}^{\infty} \cdots \underbrace{\int_{-\infty}^{t}}_{i\text{-tý}} \cdots \int_{-\infty}^{\infty} f(u_1, \ldots, u_d)\, \mathrm{d}u_d \ldots \mathrm{d}u_1 = \int_{-\infty}^{t} f_{X_i}(u)\, \mathrm{d}u$$

kde

$$f_{X_i}(u_i) = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{d-1 \text{ integrálů}} f(u_1, \ldots, u_{i-1}, u_i, u_{i+1}, \ldots, u_d)\, \mathrm{d}u_d \ldots \mathrm{d}u_{i+1} \mathrm{d}u_{i-1} \ldots \mathrm{d}u_1$$

je marginální hustota náhodné veličiny X_i.

Důkaz. Podle definice marginální funkce $F_{X_i}(t) = \lim_{t_j \to \infty, j \neq i} F_{\boldsymbol{X}}(t_1, \ldots, t_d)$. Podle definice AC vektoru je to dále rovno $\lim_{t_j \to \infty, j \neq i} \int_{-\infty}^{t_1} \cdots \int_{-\infty}^{t_{i-1}} \int_{-\infty}^{t_i} \int_{-\infty}^{t_{i+1}} \cdots \int_{-\infty}^{t_d} f_{\boldsymbol{X}}$ neklesající v t_j omezena shora 1. Tedy $\lim F_{\boldsymbol{X}}(\boldsymbol{t}) = 1$. Potom

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{t} \cdots \int_{-\infty}^{\infty} f_{\boldsymbol{X}} = \int_{-\infty}^{t} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\boldsymbol{X}}(u_1, \ldots, u_d)\, \mathrm{d}u_d \ldots \mathrm{d}u_{i+1} \mathrm{d}u_{i-1} \ldots \mathrm{d}u_1}_{\geq 0} \mathrm{d}u_i$$

□

Věta 28 (Nezávislost AC)**.** *Buď $\boldsymbol{X} = (X_1, \ldots, X_d)$ AC náhodný vektor se sdruženou hustotou $f_{\boldsymbol{X}}$. Pak náhodné veličiny $X_1, \ldots, X_d$ jsou nezávislé právě tehdy, když*

$$f_{\boldsymbol{X}}(\boldsymbol{a}) = \prod_{i=1}^{d} f_{X_i}(a_i)\ \forall \boldsymbol{a} = (a_1, \ldots, a_d) \in \mathbb{R}^d$$

Důkaz této věty vychází ze vztahu hustoty a rozdělení.

Náhodné rozdělení, distribuční funkci i hustotu umíme mezi sebou vyvodit. Také umíme vyvodit marginální distribuční funkci, naopak však ne.

Věta 29 (Momenty AC náhodného vektoru)**.** *Buď $\boldsymbol{X} = (X_1, \ldots, X_d)$ AC náhodný vektor se sdruženou hustotou $f_{\boldsymbol{X}}$ a $g : \mathbb{R}^d \to \mathbb{R}$. Pak*

$$Eg(\boldsymbol{X}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(u_1, \ldots, u_d) f_{\boldsymbol{X}}(u_1, \ldots, u_d)\, \mathrm{d}u_d \ldots \mathrm{d}u_1$$

pokud tento integrál existuje. Speciálně pro $g(\boldsymbol{X}) = X_1$ je

$$Eg(x) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} u_1 f_{\boldsymbol{X}}(u_1, \ldots, u_d)\, \mathrm{d}u_d \ldots \mathrm{d}u_1 = \int_{-\infty}^{\infty} u_1 f_{X_1}(u_1)\, \mathrm{d}u_1.$$

Transformace

Věta 30 (Konvoluce)**.** *Buď (X, Y) AC náhodný vektor s hustotou $f_{(X,Y)}$. Pak $Z = X + Y$ je AC náhodná veličina a platí*

$$f_Z(t) = \int_{-\infty}^{\infty} f_{(X,Y)}(x, t - x)\,\mathrm{d}x$$

Tato věta je v podstatě spojitá verze diskrétního součtu.

Důkaz. Chceme najít $P[Z \le z] = P[X + Y \le z] = P[(X, Y) \in H]$ kde H je polorovina rozdělená přímkou $Z = X + Y$. To se rovná $\iint_H f_{(X,Y)}(u, v)\,\mathrm{d}v\mathrm{d}u = \int_{-\infty}^{\infty}\int_{-\infty}^{z-u} f_{(X,Y)}(u, v)\,\mathrm{d}v\mathrm{d}u = F_z(z)$.

Využijme vztahu distribuční funkce a hustoty: $F_Z(z) = \int_{-\infty}^{z} f(z)(t)\,\mathrm{d}t$. Využijme vztah $f_Z(z) = F_Z'(z)$, který ale ne nutně pro všechny z platí, jelikož F_z nemusí být derivovatelná. Těchto problematických bodů však není „mnoho“ a v nich můžeme f dodefinovat podle potřeby.

Potřebujeme tedy najít $\frac{\partial}{\partial z}\int_{-\infty}^{\infty}\int_{-\infty}^{z-u} f_{(X,Y)}(u, v)\,\mathrm{d}v\mathrm{d}u$. Tato derivace se prvního integrálu nedotýká, dostáváme se tedy na $\int_{-\infty}^{\infty}\left(\frac{\partial}{\partial z}\int_{-\infty}^{z-u} f_{(X,Y)}(u, v)\,\mathrm{d}v\right)\mathrm{d}u$. Vnitřek můžeme zderivovat, kde získáme $f_{(X,Y)}(u, z-u)\cdot 1$. Dostáváme tedy výraz ve větě. □

Věta 31 (Hustota podílu a součinu)**.** *Buď (X, Y) AC náhodný vektor s hustotou $f_{(X,Y)}$ a nechť $P[Y > 0] = 1$. Pak $V = \frac{X}{Y}$ je AC náhodná veličina s hustotou*

$$f_V(v) = \int_{-\infty}^{\infty} f_{(X,Y)}(vy, y)\cdot y\,\mathrm{d}y.$$

Dále $W = XY$ je AC náhodná veličina s hustotou

$$f_W(w) = \int_{-\infty}^{\infty} f_{(X,Y)}(\frac{w}{y}, y)\cdot\frac{1}{y}\,\mathrm{d}y.$$

Důkaz. Chceme najít $P[V \le v] = P[\frac{X}{Y} \le v] = P[X \le vY] = P[(X, Y) \in H]$ kde H je polorovina rozdělená přímkou $x = vY$. Postupujme tedy podobně, jako u předchozího důkazu.

Derivováním vnitřního integrálu tentokrát dostaneme $F_{(X,Y)}(vy, y)\cdot y = \frac{\partial(vy)}{\partial y}$.

Hustotu součinu dostaneme přímým dosazením obrácené hodnoty Y do podílu. □

Normální rozdělení (taky Gaussovo, či po Germánsku Gaußovo)

Zapisuje se $X \sim N(0, 1)$ (X má normální rozdělení s parametry 0 a 1).

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

Dále vezměme si $Y = \mu + \sigma X$, potom $f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(y-\mu)^2}{2\sigma^2}}$. Dále $EY = \mu$, $\operatorname{var} Y = \sigma^2$.

Pokud máme $Z \sim N(\mu, \sigma^2)$, potom normované či standardní rozdělení je $\frac{Z-\mu}{\sigma} \sim N(0, 1)$.

Mnohorozměrné normální rozdělení

1. Mějme $X, Y \sim N(0, 1)$, které jsou navzájem nezávislé. Potom $f_{X,Y} = \frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}$.

2. Mějme $X \sim N(\nu_1, \sigma_1^2)$ a $Y \sim N(\mu_2, \sigma_2^2)$, které jsou nezávislé. $f_{X,Y} = \frac{1}{2\pi\sqrt{\sigma_1^2\sigma_2^2}}\exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2}\right)$.

Dále $\operatorname{Var}(X, Y) = \begin{pmatrix}\sigma_1^2 & 0\\ 0 & \sigma_2^2\end{pmatrix} = \Sigma$.

Potom $\det\Sigma = \sigma_1^2\sigma_2^2$ a Σ je nějaká kvadratická forma, tedy

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sqrt{\det\Sigma}}\exp\left(-\frac{1}{2}(x - \mu_1, y - \mu_2)^T\Sigma^{-1}(x - \mu_1, y - \mu_2)\right).$$

Tato hustota lze definovat pro libovolnou $\Sigma = \begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$ symetrickou pozitivně definitní. Potom $EX = \mu_1$, $EY = \mu_2$, $\operatorname{var} X = \sigma_1^2$, $\operatorname{var} Y = \sigma_2^2$, $\operatorname{corr}(X, Y) = \rho$.

Tato definice lze zobecnit na d rozměrů. Mějme $\Sigma^{d\times d}$ symetrickou pozitivně definitní, $\boldsymbol{\mu} \in \mathbb{R}^d$. Potom $f_{\boldsymbol{X}}(\boldsymbol{u}) = \frac{1}{(2\pi)^{\frac{d}{2}}\sqrt{\det\Sigma}}\exp\{-\frac{1}{2}(\boldsymbol{u} - \boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{u} - \boldsymbol{\mu})\}$.

V tomto náhodném vektoru je potom $\operatorname{Var}\boldsymbol{X} = \Sigma$, $E\boldsymbol{X} = \boldsymbol{\mu}$. Každé marginální rozdělení bude také odpovídat normálnímu rozdělení.

Normální rozdělení je jediné absolutně spojité rozdělení, kde platí, že když $(X, Y) \sim N_2(\dots)$ a $\operatorname{cov}(X, Y) = 0$, X, Y jsou nezávislé.

7 Podmíněné rozdělení a podmíněná střední hodnota

Definice 26 (Podmíněné rozdělení diskrétního náhodného vektoru)**.** Buď $X = (Y, Z)$ diskrétní náhodný vektor (s hodnotami v $\mathbb{N}_0^2$). Pak pro každé $n \in N_0$ takové, že $P[Z = n] > 0$ definujeme podmíněné rozdělení Y za podmínky $Z = n$:

$$P[Y = k|Z = n] = \frac{P[Y = k, Z = n]}{P[Z = n]}.$$

Podmínku $Z = n$ chápeme jako další informaci o chování náhodné veličiny Y, dává nám na něj přesnější pohled, platí-li podmínka.

Příklad. Mějme 6 mincí a jednu kostku. Hodíme kostkou a podle výsledku hodíme odpovídající počet mincí. Jaká je pravděpodobnost, že nepadne žádný líc?

Jaké je rozdělení počtu líců? Víme, že počet líců je $0, \ldots, 6$. To však nedokážeme spočítat přímo, musíme vždy spočítat rozdělení za podmínky hození kostkou.

Definice 27 (Podmíněná střední hodnota)**.** Buď $X = (Y, Z)$ diskrétní náhodný vektor takový, že

1. $P[Z = n] > 0$
2. $Eg(Y, n) < \infty$ pro zvolenou funkci g.

Pak definujeme $Eg(Y, Z|Z = n) = \sum_{k=0}^{\infty} g(k, n)P[Y = k|Z = n]$.

Speciálně, pokud $E|Y| < \infty$, potom $E(Y|Z = n) = \sum_{k=0}^{\infty} kP[Y = k|Z = n]$. a zveme toto *podmíněnou střední hodnotu* $g(Y, Z)$ za podmínky $Z = n$.

Věta 33 (O úplné střední hodnotě)**.** *Buď $X = (Y, Z)$ diskrétní náhodný vektor a EY existuje konečná. Pak*

$$EY = \sum_n E[Y|Z = n]P[Z = n]$$

pro ta n, pro které $P[Z = n] > 0$.

Důkaz. Začněme s pravou stranou. $\sum_n E[Y|Z = n]P[Z = n] = \sum_n \sum_{k=0}^{\infty} kP[Y = k|Z = n]P[Z = n] = \sum_n \sum_{k=0}^{\infty} kP[Y = k, Z = n]$. Jelikož počítáme s existující EY, můžeme dále pokračovat na $\sum_{k=0}^{\infty} k \sum_n P[Y = k, Z = n] = EY$. □

Všimněme si, že podmíněná středí hodnota je konstanta závislá na n a při sečtení všech n dostaneme přímou střední hodnotu EY, která je již konstantní v daném rozdělení.

Můžeme si také všimnout, že podmíněná pravděpodobnost nám faktoruje množinu Ω podle hodnoty z. Na této faktorové množině poté můžeme zavést novou náhodnou veličinu se zajímavými vlastnostmi.

Definice 28 (Podmíněná střední hodnota jako náhodná veličina)**.** Buď $X = (Y, Z)$ diskrétní náhodný vektor, $E|Y| < \infty$. Definujeme náhodnou veličinu $E(Y|Z)$ předpisem $E(Y|Z)(\omega) = E(Y|Z = n)\ \forall \omega \in \{\omega : Z(\omega) = n\}$.

Rozdělení této náhodné veličiny je zřejmě $P[E(Y|Z) = E(Y|Z = n)] = P[Z = n]$.

EY má tuto vlastnost: $E(Y - EY)^2 = \min_{a \in \mathbb{R}} E(Y - a)^2$. Střední hodnota je tedy nejlepší konstanta, která reprezentuje nejmenší čtvercovou chybu.

Na množině $\{\omega : Z(\omega) = n\}$ platí $E(Y - E(Y|Z = n))^2 = \min_{b \in \mathbb{R}} E(Y - b)^2$. Můžeme tedy najít ještě lepší konstantu, než EY, reprezentující tuto informaci.

Věta 34. *Buď $X = (Y, Z)$ diskrétní náhodný vektor, $E|Y| < \infty$. Pak $E(E(Y|Z)) = EY$.*

Důkaz. Přímý důsledek věty 33. □

Poznámka (Podmíněný rozptyl)**.** $\operatorname{var}(Y|Z = n) = E\left((Y - E(Y|Z = n))^2 \middle| Z = n\right)$

Definice 29 (Podmíněná hustota)**.** Buď $X = (Y, Z)$ absolutně spojitý náhodný vektor se sdruženou hustotou $f_{Y,Z}$. Definujme $f_{Y|Z}(v|w)$ předpisem:

$$f_{Y|Z}(v|w) = \begin{cases} \frac{f_{Y,Z}(v,w)}{f_Z(w)} & f_Z(w) > 0 \\ 0 & \text{jinak} \end{cases}$$

kterouž funkci nazveme hustotou podmíněného rozdělení Y za podmínky $Z = w$.

Nyní pojďme ověřit, že takto definovaná hustota je opravdu hustotou.
Marginální hustota $f_Z = \int_{-\infty}^{\infty} f_{Y,Z}(v,w)\,\mathrm{d}v$. Pokud je rovna nule, potom $f_{Y,Z}(v,w) = 0$.
Nyní $\int_{-\infty}^{\infty} f_{Y|Z}(v|w)\,dv = \int \frac{f_{Y,Z}(v,w)}{f_Z w} = \frac{f_Z(w)}{f_Z(w)}$, kdykoliv $f_Z(w) > 0$.

$P[Y \in A, Z \in B] = \int_A \int_B f_{Y,Z}(v,w)\,\mathrm{d}w\mathrm{d}v = \int_A \int_B f_{Y|Z}(v|w) f_Z(w)\,\mathrm{d}w\mathrm{d}v = \int_B \left(\int_A f_{Y|Z}(v|w)\,\mathrm{d}v \right) f_Z(w)\,\mathrm{d}w$. Zde můžeme vidět náznaky věty o úplně pravděpodobnosti pro absolutně spojitý náhodný vektor.

Definice 30 (Podmíněná střední hodnota)**.** Buď $X = (Y,Z)$ absolutně spojitý náhodný vektor, $g : \mathbb{R}^2 \to \mathbb{R}$ taková, že $E|G(Y,w)| < \infty$. Pak definujeme $E(g(Y,Z)|Z = w) = \int_{-\infty}^{\infty} g(v,w) f_{Y|Z}(v|w)\,\mathrm{d}v$ a $E(Y|Z = w) = \int_{-\infty}^{\infty} v f_{Y|Z}(v|w)\,\mathrm{d}v$.

8 Nerovnosti a meze

Nerovnosti se hodí k odhadování určování pravděpodobností určitých jevů či veličin.

Věta 35 (Markovova nerovnost)**.** *Buď X nezáporná náhodná veličina. Potom*

$$P[X \geq \varepsilon] \leq \frac{EX}{\varepsilon} \qquad \forall \varepsilon > 0$$

Důkaz. $[X \geq \varepsilon] = P\{\omega : X(\omega) \geq \varepsilon\} = \int_{\omega : X(\omega) \geq \varepsilon} 1\,\mathrm{d}P \leq \int_{\omega : X(\omega) \geq \varepsilon} \frac{X(\omega)}{\varepsilon}_{\text{Vždy kladné}}\,\mathrm{d}P \leq \int_\Omega \frac{X(\omega)}{\varepsilon}\,\mathrm{d}P = \frac{EX}{\varepsilon}$. □

Věta 36 (Markovova nerovnost podruhé)**.** *Buď X nezáporná náhodná veličina. Potom*

$$P[X \geq \varepsilon] \leq \frac{E(X^k)}{\varepsilon^k} \qquad \forall \varepsilon > 0, k > 0$$

Důkaz. Jediná změna oproti předchozímu: $X(\omega) \geq \varepsilon \Rightarrow \left(\frac{X(\omega)}{\varepsilon}\right)^k \geq 1$. Jinak stejně. □

Věta 37 (Čebyševova nerovnost)**.** *Buď X náhodná veličina a $E|X| < \infty$. Pak*

$$P[|X - EX| \geq \varepsilon] \leq \frac{\operatorname{var} X}{\varepsilon^2}$$

Důkaz. Věta 36 použitá na $|X - EX|$ a $k = 2$. □

Pojďme si na chvíli povídat o náhodných procházkách. Představme si, že jsme zrovna vyšli z hospůdky ve tři ráno a náhodně v ulicích jdeme na sever nebo na jih pár minut. Kam až se můžeme dostat? Za jak dlouho se vrátíme zpět k hospodě, abychom tam strávili čas tentokrát až do rána?

Mějme $X_1, X_2, \dots$ nezávislé, jejichž pravděpodobnosti jsou $P[X_i = 1] = P[X_i = -1] = \frac{1}{2}$. Dále mějme $S_k = \sum X_i$. Čebyševova nerovnost říká $P[|S_n - ES_n \geq \varepsilon] \leq \frac{\operatorname{var} S_n}{\varepsilon^2}$.
Spočtěme nyní $ES_k = \sum EX_i = 0$ a $\operatorname{var} S_k = \operatorname{var} \sum S_i = \sum \operatorname{var} S_i = k$.
Podle nerovnosti tedy dostáváme, že $P[|S_n| \geq \varepsilon] \leq \frac{n}{\varepsilon^2}$. Dále $P[|S_n| \geq 2\sqrt{n}] \leq \frac{n}{4n} \leq \frac{1}{4}$.

Věta 38 (Kolmogorovova nerovnost)**.** *Buďte $X_1, X_2, \dots$ nezávislé náhodné veličiny a $E|X_i| < \infty$ $\forall i$. Pak*

$$P\left[\max_{1 \leq k \leq n} \left| \sum_{i=1}^{k} (X_i - EX_i) \right| \geq \varepsilon \right] \leq \frac{\sum_{i=1}^n \operatorname{var} X_i}{\varepsilon^2}$$

Důkaz. Označme si $S_k = \sum_{i=1}^k (X_i - EX_i)$. Potom $ES_k = 0$ a $\operatorname{var} S_k = \sum_{i=1}^k \operatorname{var} X_k$ díky nezávislosti. Hledáme tedy pravděpodobnost $P[\max_{1 \leq k \leq n} |S_k| \geq \varepsilon]$.

Zadefinujme si množinu $A_k = \{\omega : |S_j(\omega)| < \varepsilon, j < k, |S_k(\omega)| \geq \varepsilon\}$ a $S_0 = 0$. Můžeme si všimnout, že A_k jsou po dvou disjunktní. Rozložili jsme pravděpodobnostní prostor podle prvního překročení ε.

Potom $P[\max_{1 \leq k \leq n} |S_k| \geq \varepsilon] = P[\bigcup_{k=1}^n A_k] = \int_{\bigcup_{k=1}^n A_k} 1\,\mathrm{d}P = \sum_{k=1}^n \int_{A_k} 1\,\mathrm{d}P \leq \sum_{k=1}^n \int_{A_k} \frac{S_k^2(\omega)}{\varepsilon^2}\,\mathrm{d}P$.
Dále $\int_{A_k} S_k^2(\omega)\,\mathrm{d}P = \int_{A_k} (S_n - S_k + S_k)^2\,\mathrm{d}P = \int_{A_k} (S_n - S_k)^2\,\mathrm{d}P + \int_{A_k} 2(S_n - S_k) \cdot S_k\,\mathrm{d}P + \int_{A_k} (S_k)^2\,\mathrm{d}P$.

Z toho se podívejme na $\int_{A_k} 2(S_n - S_k) \cdot S_k\,\mathrm{d}P = \int_\Omega (S_n - S_k)(S_k \chi_{A_k})\,\mathrm{d}P = E\left((S_n - S_k) \cdot S_k \cdot \chi_{A_k}\right)$. Využijme nezávislosti a toho, že $S_n - S_k = X_{k+1} + X_{k+2} + \dots + X_n$ a $S_k = X_1 + X_2 + \dots + X_k$ a X_{A_k} závisí jen na $(X_1, \dots, X_k)$. Tedy celý integrál se rovná $E(S_n - S_k) E S_k \chi_{A_k} = 0$.

Tedy $\sum_{k=1}^n \int_{A_k} \frac{S_k^2(\omega)}{\varepsilon^2}\,\mathrm{d}P \leq \int_{A_k} \frac{S_n^2(\omega)}{\varepsilon^2}\,\mathrm{d}P$, kde potom můžeme postupovat jako u předchozích nerovností. □

Věta 39 (Chernoffovy meze). *Buď X náhodná veličina. Označme $\psi_X(t) = Ee^{tX}$ její momentovou vytvořující funkci. Pak*

$$P[X \geq a] \leq \min_{t>0} \frac{\psi_X(t)}{e^{ta}}, P[X \leq A] \leq \min_{t<0} \frac{\psi_X(t)}{e^{ta}}$$

Důkaz. Pro první mez $P[X \geq a] = P[e^{tX} \geq e^{ta}]$ pro $t > 0$. Dostali jsme nezápornou veličinu, využijeme Markovovu nerovnost, tedy $P[e^{tX} > e^{ta}] \leq \frac{Ee^{tx}}{e^{ta}} \forall t > 0$. Z toho potom jen vybereme minimu.

Druhá mez analogicky. $P[X \leq a] = P[e^{tX} \geq e^{ta}] \leq \frac{Ee^{tx}}{e^{ta}}$ pro $t < 0$. □

Definice 31 (Poissonovské pokusy). Buďte $X_1, X_2, \ldots$ nezávislé náhodné veličiny takové, že $P[X_i = 1] = p_i = 1 - P[X_i = 0], p_i \in (0,1)$. Takové pokusy se nazývají nezávislé poissonovské pokusy.

Jestliže u poissonovských pokusů bude každé $p_i = p$ pro nějaké konstantní p, dostaneme bernoulliovské pokusy.

Věta 40 (Horní Chernoffovy meze pro poissonovské pokusy). *Buďte $X_1, X_2, \ldots$ nezávislé poissonovské pokusy s parametry $p_1, p_2, \ldots$. Označme $S_n = \sum_{i=1}^n X_i, \mu_S = \sum_{i=1}^n p_i$.*

1. *Pro $\delta > 0$:*
$$P[S_n \geq (1+\delta)\mu_s] \leq \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^{\mu_s}$$

2. *Pro $0 < \delta \leq 1$:*
$$P[S_n \geq (1+\delta)\mu_s] \leq e^{-\mu_s \delta^2/3}$$

3. *Pro $\delta > 6\mu_s$:*
$$P[S_n > \delta] \leq 2^{-\delta}$$

Důkaz. Jen první. $P[S_n \geq (1+\delta)\mu_s] \leq \frac{\psi_{S_n}(t)}{e^{t(1+\delta)\mu_s}}$.

Z nezávislosti a věty 26 dostáváme $\psi_{s_n} = \prod_{i=1}^n \psi_{x_i}(t)$ a $\psi_{x_i}(t) = Ee^{tX_i} = p_i e^t + (1-p_i) = 1 + p_i(e^t - 1)$.
Dále $\prod_{i=1}^n (1 + p_i(e^t - 1)) = \exp\{\sum_{i=1}^n \log(1 + p_i(e^t - 1))\} \leq \exp\{(e^t - 1)\sum_{i=1}^n p_i\} = \exp\{(e^t - 1)\mu_s\}$.

Dosadíme do vzorce, dostáváme $\frac{\prod_{i=1}^n (1+p_i(e^t-1))}{e^{t(1+\delta)\mu_s}} \leq \frac{\exp\{(e^t-1)\mu_s\}}{\exp t(1+\delta)\mu_s} = \exp \mu_s(e^t - 1 - t(1+\delta)$.

Nakonec $P[S_n \geq (1+\delta)\mu_s] \leq \frac{\prod_{i=1}^n (1+p_i(e^t-1))}{e^{t(1+\delta)\mu_s}} = \exp \mu_s[\delta - \log(1+\delta)^{1+\delta}] =$ hledaný výraz. □

9 Náhodné výběry a limitní věty

Na této přednášce jsem z nemoci chyběl, proto může pár věcí chybět.

Definice 32 (Množinový lim sup a lim inf). Buď $A_n, n = 1, 2, \ldots$ množiny (náhodné jevy). Definujme

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^\infty \bigcup_{i=n}^\infty A_i \quad \liminf_{n\to\infty} A_n = \bigcup_{n=1}^\infty \bigcap_{i=n}^\infty A_i$$

Jestliže $a \in \limsup_{n\to\infty} A_n$, potom existuje nekonečně mnoho množin A_i, kde $a \in A_i$. Jestliže $b \in \limsup_{n\to\infty} A_n$, potom existuje nejvýše konečně mnoho množin A_i, kde $b \notin A_i$.

Věta 42 (Borel-Cantelliho 0-1 pravidlo).

1. *Buď $A_n, n = 1, 2, \ldots$ náhodné události. Pak*
$$\sum_{i=1}^\infty P(A_i) < \infty \Rightarrow P\left(\limsup_{n\to\infty} A_n\right) = 0.$$

2. *Buď $A_n, n = 1, 2, \ldots$* nezávislé *náhodné události. Potom*
$$\sum_{i=1}^\infty P(A_i) = \infty \Rightarrow P\left(\limsup_{n\to\infty} A_n\right) = 1.$$

Všimněme si, že $\liminf A_n = (\limsup A_n^C)^C$. Z tohoto důvodu můžeme ve větě 42 používat jak lim sup, tak i lim inf.

Definice 33 (Náhodný výběr). Posloupnost $X_1, X_2, \ldots$ náhodných veličin či vektorů takových, že jsou nezávislé, identicky rozložené, nazvěme náhodný výběr velikosti n náhodného rozložení P_X.

Definice 34 (Výběrové momenty)**.** Buď $X_1, \dots, X_n$ náhodný výběr. Potom

1. $\overline{X_n} = \frac{1}{n}\sum_{i=1}^n X_i$ se nazývá *výběrový průměr.*
2. $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X_n})^2$ se nazývá *výběrový rozptyl.*
3. $\widehat{F_n}(x) = \frac{1}{n}\sum_{i=1}^n \chi(X_i \le x)$, kde χ je indikátorová funkce, se nazývá *empirická distribuční funkce.*

Věta 43 ((Slabý) Zákon velkých čísel)**.** *Buď $X_1, X_2, \dots$ nezávislé a identicky rozložené náhodné veličiny s konečným průměrem $EX_1 = \mu$ a konečným rozptylem* $\operatorname{var} X_1 = \sigma^2$*. Potom*

$$P\left[\left|\frac{1}{n}\sum_{i=1}^n X_i - \mu\right| > \varepsilon\right] \to 0 \text{ pro } n \to \infty. \qquad \forall \varepsilon > 0$$

Důkaz. Vidíme, že střední hodnota z průměru $E\overline{X_n} = E\frac{1}{n}\sum X_i = \frac{1}{n}\sum EX_i = \mu$.. Využijme nyní Čebyšeovu nerovnost, tedy $P[|\overline{X_n} - \mu| > \varepsilon] \le \frac{\operatorname{var}\overline{X_n}}{\varepsilon^2}$.

Podívejme se nyní na $\operatorname{var}\overline{X_n} = \operatorname{var}\frac{1}{n}\sum X_i = \frac{1}{n^2}\operatorname{var}\sum X_i$. Díky nezávislosti získáme $\frac{1}{n^2}\operatorname{var}\sum X_i = \frac{1}{n^2}\sum \operatorname{var} X_i + \frac{1}{n^2}\sum_{i\neq j}\operatorname{cov}(X_i, X_j)$. Jelikož jsou každé X_i nekorelované, druhá suma je nulová, zatímco ta první limitně konverguje k 0. $\square$

Tato vlastnost se nazývá (slabá) konzistence výběrového průměru. Předpoklad $\sigma^2 < \infty$ se zdá příliš silný. Dal by se oslabit, ale potom by byl důkaz mnohem složitější. Tuto vlastnost nadále budeme označovat jako $\overline{X_n} - \mu \xrightarrow{P} 0$ nebo $\overline{X_n} \xrightarrow{P} \mu$.

Příklad. Mějme alternativní rozdělení v n pokusech, každý pokus nastává s pravděpodobností p. Potom výběrový průměr je počet úspěchů v n pokusech a střední hodnota je p.

Relativní počet úspěchů tedy díky zákonu velkých čísel konverguje k p.

Věta 44 (Konzistence empirické distribuční funkce)**.** *Buď $X_1, X_2, \dots$ nezávislé a identicky rozložené náhodné veličiny s distribuční funkcí F_X. Pak pro každé x platí*

$$P\left[\left|\widehat{F_n}(x) - F_X(x)\right| > \varepsilon\right] \to 0 \text{ pro } n \to \infty.$$

Příklad. Máme X s neznámým rozdělením. Chceme určit $P[0 < X < 1]$. Potřebujeme náhodný výběr $X_1, X_2, \dots$ s tímto neznámým náhodným rozdělením a víme, že $\widehat{F_n}(1) \to F_X(1)$ a $\widehat{F_n}(0) \to F_X(0)$. Tedy $P[0 < x \le 1] = F_X(1) - F_X(0)$. Dále $\chi(x \le a)$ dává jedničku, pokud je $x \le a$, jinak nulu.

Proto $\widehat{F_n}(1) - \widehat{F_n}(0) = \frac{1}{n}\sum \chi(0 < X \le 1) \xrightarrow{p} P[0 < X \le 1]$.

Důkaz. Podívejme se na střední hodnotu a rozptyl odhadu $\widehat{F_n}(x)$. Tedy $E\widehat{F_n}(x) = E\frac{1}{n}\sum \chi(X_i \le X) = \frac{1}{n}\sum E\chi(X_i \le X) = E\chi(X_1 \le X)$. Dále $\chi(X \le x) = 0$ s pravděpodobností $F(x)$ a 1 s pravděpodobností $1 - F(x)$. Máme tedy Bernoulliho náhodný pokus se střední hodnotou $F(x)$. Tedy $E\chi(X_1 \le x) = F(x)$ a $\operatorname{var}\chi(X_1 \le x) = F(x)(1 - F(x))$.

Využijme nyní Čebyšeovu nerovnost. $P[|\widehat{F_n}(x) - F(x)| > \varepsilon] \le \frac{\operatorname{var}\widehat{F_n}(x)}{\varepsilon^2} = \frac{1}{n\varepsilon^2}F(x)(1 - F(x))$, což pro $n \to \infty$ konverguje k nule. $\square$

Předchozí věta platí dokonce i rovnoměrně, tedy $P\left[\sup_x \left|\widehat{F_n}(x) - F_X(x)\right| > \varepsilon\right] \to 0$.

Věta 45 (Centrální limitiní věta)**.** *Buď $X_1, X_2, \dots$ nezávislé a identicky rozložené náhodné veličiny s konečným průměrem $EX_1 = \mu$ a konečným kladným rozptylem* $\operatorname{var} X_1 = \sigma^2 > 0$*. Pak*

$$P\left[\sqrt{n}\frac{\overline{X_n} - \mu}{\sigma} \le x\right] \to \Phi(x)$$

kde Φ je distribuční funkce standardního normálního rozdělení $N(0, 1)$.

Stejně tak můžeme říct

$$\sqrt{n}\frac{\overline{X_n} - \mu}{\sigma} \xrightarrow{d} N(0,1) \text{ nebo } \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{d} N(0,1) \text{ nebo } \sqrt{n}\left(\overline{X_n} - \mu\right) \xrightarrow{d} N(0, \sigma^2).$$

Příklad. Podívejme se na využití Centrální limitní věty.

1. Chceme vědět, s jakou pravděpodobností bude počet úspěchů v Bernoulliovských pokusech vyšší, než a. Máme $X_1, \dots, X_n$ nezávislé stejnoměrně rozdělené s alternativním rozdělením parametru p. Kolik bude $P[\sum X_i > a]$?

   Použijeme Centrální limitní větu a použijeme stejné úpravy na obě strany:

$$P[\Sigma X_i > a] = P\left[\frac{\sum X_i - np}{\sqrt{np(1-p)}} > \frac{a-np}{\sqrt{np(1-p)}}\right] = P[\dots < \dots] \approx \Phi\left(\frac{a-np}{\sqrt{np(1-p)}}\right)$$

2. Určení počtu pokusů. Chceme vědět, kolik pokusů máme udělat, aby se relativní počet úspěchů v Bernoulliovských pokusech (tedy podíl úspěchů k počtu pokusů) lišil od teoretické pravděpodobnosti nejvýše o ε s pravděpodobností nejvýše α. Tedy hledáme n tak, aby $P[|\overline{X_n} - p| > \varepsilon] \leq \alpha$.

   Znova využijeme CLV a upravujeme. Úpravou dostaneme

$$P\left[\sqrt{n}\frac{|\overline{X_n} - p|}{\sqrt{p(1-p)}} > a\right] = 1 - P\left[\dots \leq a\right] = 1 - (P[\dots \leq a] - P[\dots \leq -a]) = 1 - (\Phi(a) - \Phi(-a)).$$

   To znamená, že hledáme $a = \frac{\sqrt{n}\varepsilon}{\sqrt{p(1-p)}}$ tak, aby $1 - (\Phi(a) - \Phi(-a)) \leq \alpha$, tedy $\Phi(a) - \Phi(-a) \geq 1 - \alpha$. Víme, že $\Phi(-a) = 1 - \Phi(a)$, z toho dostáváme, že $2\Phi(a) \geq 2 - \alpha$, tedy $a \geq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$. Z a potom získáme n.

Věta 46 (Delta)**.** *Buď $Y_1, Y_2, \dots$ posloupnost náhodných veličin takových, že $\sqrt{n}(Y_n - \mu) \xrightarrow{d} N(0, \sigma^2)$ a diferencovatelnou funkci g. Pak*

$$\sqrt{n}\,(g(Y_n) - g(\mu)) \xrightarrow{d} N\big(0, (g'(\mu))^2\sigma^2\big)$$

Delta věta se používá, když chceme asymptotickou normalitu nějaké diferencovatelné transformace normálního rozdělení.

Příklad. Předpokládejme, že máme Poissonovo rozdělení s parametrem λ. Pak víme, že $EX = \operatorname{var} X = \lambda$. Z CLV víme, že $\sqrt{n}(\bar(X_n) - \lambda) \xrightarrow{d} N(0, \lambda)$. Dále víme, že $P[X = 0] = e^{-\lambda} = g(\lambda)$.

Delta věta říká $\sqrt{n}(e^{-\overline{X_n}} - e^{-\lambda}) \xrightarrow{d} N(0, \lambda(g'(\lambda))^2) = N(0, \lambda e^{-2\lambda})$.

Věta 47 (Cramér–Slutskij)**.** *Buď $X_n \xrightarrow{d} N(\mu, \sigma^2)$, $U_n \xrightarrow{P} a$ a $Z_n \xrightarrow{P} s > 0$. Pak*

$$Z_n X_n + U_n \xrightarrow{d} N(s\mu + a, s^2\sigma^2)$$

Cramér–Slutskij a Delta věty jsou základní nástroje pro asymptotické odhadovací techniky založené na CLV. Z věty C–S vyplývá, že konvergence $\xrightarrow{d}$ je *slabší*, než konvergence $\xrightarrow{P}$.

Tato věta se nejčastěji používá, když neznáme rozptyl nebo je práce s rozptylem příliš obtížná. V tom případě potom víme, že díky CLV $\sqrt{n}\frac{\overline{X_n} - \mu}{\sigma^2} \xrightarrow{d} N(0, 1)$ a také ze zákona velkých čísel $\frac{Sn}{\sigma} \xrightarrow{P} 1$ pokud $0 < \sigma^2 < \infty$.

Díky C–S větě pak $\frac{(\sqrt{n} \cdot \overline{X_n} - \mu)/\sigma}{S_n/\sigma} = \sqrt{n}\frac{\overline{X_n} - \mu}{s_n} \xrightarrow{d} N(0, 1)$, tedy jsme nahradili rozptyl jeho výběrovým odhadem.

10 Bodové a intervalové odhady, hypotézy

Jedním ze základních úkolů statistiky je udělat nějaký (dobrý) odhad.

Představme si, že máme pokus, jehož výsledkem může být $0, \dots, n$. Udělali jsme tedy n pokusů, zajímá nás $P[X = k]$. Jednou z možností je sestavit empirickou distribuční funkci, která bude po částech skoková. Tyto odhady nejsou ideální ani přesné, musíme odhadovat $n + 1$ hodnot.

Představme si ale, že předem víme, že $x \sim \mathrm{Bi}(n, p)$. Již známe n, stačí nám odhadnout pouze jeden parametr p, potom náhodné rozdělení již známe.

Definice 35 (Parametrické třídy)**.** Parametrickou třídou nazveme $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$, kde $\Theta \subset \mathbb{R}^d$ je prostor parametrů a P_θ je rozdělení náhodné veličiny (vektoru) známé až na parametr θ.

Příklad.

- $\mathrm{Bi}(n, p)$, rozdělení $P[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$. Potom $\Theta = (0, 1)$ a p je neznámá.
- $N(\mu, \sigma^2)$ s hustotou $\frac{1}{\sqrt{2\pi\sigma}} \exp(-\frac{(x-\mu)^2}{2\sigma^2})$, potom $\Theta = \mathbb{R} \times (0, \infty)$ a parametry jsou (μ, σ^2).
- Exponenciální rozdělení s hustotou $f(x) = \lambda e^{-\lambda(x-a)}$, potom $\Theta = (0, \infty) \times \mathbb{R}$, parametry (λ, a).

Definice 36 (Bodový odhad)**.** Buď $X_1, X_2, \dots, X_n$ náhodný výběr s rozdělením P_θ z parametrické třídy $\mathcal{P}$ a s θ neznámý parametrem. Bodovým odhadem θ je funkce $T : \mathbb{R}^n \to \Theta$, $T(X_1, \dots, X_n) \to \dots$, její předpis nezávisí na parametru θ.

Úkolem je najít T, které „dobře“ aproximuje θ, dále:

1. Bodový odhad T nazveme nestranným, pokud $\forall \theta \in \Theta$ platí $ET = \theta$ je-li θ skutečnou hodnotou parametru.
2. Bodový odhad T je konzistentní, pokud $\forall \theta \in \Theta$ platí $P[T - \theta] > \varepsilon \to 0$ je-li θ skutečnou hodnotou parametru. (Tedy $T \xrightarrow{P} \theta$).

Máme-li více konzistentních odhadů, chceme vybrat ten s nejmenším rozptylem. Není však zaručeno, aby existoval stejnoměrně nejlepší odhad.

Čím máme větší náhodný výběr, tím je odhad T přesnější. To nám zaručuje právě konzistence T.

Malá odbočka do angličtiny. „Estimator“ znamená funkci T, zatímco „Estimate“ je už konkrétní hodnota $T(X_1, \dots, X_n)$.

Jak získat bodový odhad? Existuje na to mnoho způsobů:

Definice 37 (Metoda momentů)**.** Buď $P_\theta \in \mathcal{P}$ a úlohu najít bodový odhad θ. Předpokládejme, že momenty $EX^j = h_j(\theta), j = 1, 2, \dots$ Potom řešení rovnic

$$\frac{1}{n} \underbrace{\sum_{i=1}^{n} X_i^j}_{\text{Výběrový moment}} = \underbrace{h_j(\theta)}_{\text{Skutečný moment}} \quad , j = 1, \dots, \dim(\Theta)$$

nazveme odhad metodou momentů.

Příklad. Mějme normální rozdělení. Potom $EX = \mu = h_1(\mu, \sigma^2)$ a $EX^2 = \sigma^2 + \mu^2 = h_2(\mu, \sigma^2)$. Odhad je poté $\widehat{\mu} = \overline{X_n}$ a $\widehat{\sigma^2} = \frac{1}{n} \sum_{i=1}^{n} X_i^2 - (\overline{X_i})^2$.

Ze zákonu velkých čísel potom $\overline{X_n} \xrightarrow{P} \mu$, $\frac{1}{n} \sum_{i=1}^{n} X_i^2 \to EX^2$, tedy odhad $\widehat{\sigma^2} \to \operatorname{var} X$.

Bodový odhad jde však zlepšit odhadem intervalovým, který je ale složitější.

Definice 38 (Intervalový odhad)**.** Buď $X_1, \dots, X_n$ náhodný výběr z P_θ, θ neznámý parametr a $\Theta \subset \mathbb{R}$. Intervalovým odhadem nazveme dvojici funkcí $L(X_1, \dots, X_b) \to \mathbb{R}$ a $U(X_1, \dots, X_b) \to \mathbb{R}$, jejichž předpis nezávisí na θ a které splňují $P[L(X_1, \dots, X_b) \leq \theta \leq U(X_1, \dots, X_b)] \geq 1 - \alpha$ (Interval spolehlivosti s hladinou spolehlivosti $1 - \alpha$) pro každé θ, je-li θ skutečnou hodnotou parametru.

Náhodné jsou L a U! Když provádíme náhodné experimenty, Počet intervalů (L, U) pokrývajících θ je v průměru $100(1 - \alpha)\%$. Naše snaha je získat $|UL|$ *co nejmenší.*

Jak najít L a U? Možné postupy:

1. Najdeme funkci $H : (X_1, \dots, X_n, \theta)$, jejíž rozdělení nezávisí na θ. (nebo alespoň nezávisí asymptoticky). Obvykle pomůže Centrální limitní věta.

   Příkladem je $X_1, \dots, X_n$ s alternativním rozdělením závislém na p, odhadem může být $\frac{\sum x_i - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0, 1)$ nezávislé na p.
2. Najdeme kvantily $l(\alpha), u(\alpha)$ rozdělení H takové, že $P[l(\alpha) \leq H \leq u(\alpha)] \geq 1 - \alpha$. Je-li (asymptotické) rozdělení H normální $N(0, 1)$, pak použijeme kvantily $l(\alpha) = q(\alpha/2) = \Phi^{-1}(\frac{\alpha}{2})$ a $u(\alpha) = q_{1-\alpha/2} = \Phi^-1(1 - \frac{\alpha}{2})$.
3. Osvobození parametru různými algebraickými úpravami dojdeme k nerovnostem $l(\alpha) \leq H \leq u(\alpha)$ právě, když $L(X_1, \dots, X_n, l, u) \leq \theta \leq U(X_1, \dots, X_n, l, u)$ *je-li to možné.*

   Avšak někdy ještě musíme zapojit Cramér–Slutskij větu.

Příklad. Mějme $\operatorname{Alt}(p)$, potom $\frac{\sum X_i - np}{\sqrt{np(1-p)}} \in (q_{\alpha/2}, q_{1-\alpha/2})$ s pravděpodobností přibližně pa. Dále z kvantilů získáme $q_{\alpha/2}\sqrt{np(1-p)} \leq \sum X_i - np \leq q_{1-\alpha/2}\sqrt{np(1-p)}$, máme kvadratickou nerovnost.

Nebo můžeme použít Cramér–Slutskij větu a máme $\overline{X_n} \xrightarrow{P} p \Rightarrow \overline{X_n}(1 - \overline{X_n}) \xrightarrow{P} p(1-p)$. Z toho dostáváme, že $\frac{\sum X_i - np}{\sqrt{n\overline{X_n}(1-\overline{X_n})}} \xrightarrow{d} N(0, 1)$. Potom $q_{\alpha/2}\sqrt{n\overline{X_n}(1 - \overline{X_n})} \leq p \leq q_{1-\alpha/2}\sqrt{n\overline{X_n}(1 - \overline{X_n})}$. Z toho již dostaneme $L = \overline{X_n} + \frac{q_{\alpha/2}}{\sqrt{n}} \sqrt{n\overline{X_n}(1 - \overline{X_n})}$ a $U = \overline{X_n} + \frac{q_{1-\alpha/2}}{\sqrt{n}} \sqrt{n\overline{X_n}(1 - \overline{X_n})}$, potom $P[L \leq p \leq U] \approx 1 - \alpha$.

Použití intervalového odhadu. Rozhodnutí o hypotéze: Mějme H_0 hypotézu, takovou, že $H_0 : \theta = \theta_0$. Dále mějme alternativu $H_1 : \theta \neq \theta_0$.

Intervalový odhad pro neznáme θ se spolehlivostí $1 - \alpha$. Máme (L, U) takové, že $P(L \leq \theta \leq U) = 1 - \alpha$. Je naše $\theta_0 \in (L, U)$? Pokud ano, H_0 nezamítneme na hladině statistické významnosti α. Pokud však ne, zamítneme ji.

Všimněme si, že se zvyšujícím se n se interval (L, U) zmenšuje. Je potřeba také vážit faktický význam!