

Pravděpodobnost

Klasická pravděpodobnost. Pravděpodobnost $P(A)$ jevu A v konečném základním prostoru $\Omega = \{\omega_1, \dots, \omega_n\}$ je $P(A) = \frac{m}{n}$, kde m je počet elementárních jevů příznivých jevu A .

Některé vlastnosti axiomatické pravděpodobnosti:

- (i) $0 \leq P(A) \leq 1$,
- (ii) $P(\Omega) = 1, P(\emptyset) = 0$,
- (iii) pro neslučitelné jevy A_1, A_2, \dots platí

$$P(A_1 \cup \dots \cup A_n) = P(A_1) + \dots + P(A_n),$$
- (iv) $P(A) = 1 - P(\bar{A})$,
- (v) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$,
- (vi) $P(A) \leq P(B)$ pro $A \subseteq B$,
- (vii) $P(B - A) = P(B) - P(A)$ pro $A \subseteq B$,

Podmíněná pravděpodobnost. Pravděpodobnost jevu A za podmínky, že nastal jev B , je $P(A|B) = \frac{P(A \cap B)}{P(B)}$, pokud $P(B) \neq 0$.

Pravděpodobnost průniku. Platí $P(A_1 \cap A_2 \cap \dots \cap A_m) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \dots P(A_m|A_1 \cap \dots \cap A_{m-1})$.

Nezávislost náhodných jevů. Náhodné jevy A_1, \dots, A_m jsou vzájemně nezávislé, pokud platí

$$\begin{aligned} P(A_i \cap A_j) &= P(A_i)P(A_j), \text{ pro } i \neq j, & (\text{párově nezávislé}) \\ P(A_i \cap A_j \cap A_k) &= P(A_i)P(A_j)P(A_k), \text{ pro } i \neq j, j \neq k, \\ &\vdots \\ P(A_1 \cap A_2 \cap \dots \cap A_m) &= P(A_1)P(A_2) \dots P(A_m). \end{aligned}$$

Úplná pravděpodobnost. Pokud B_1, \dots, B_n je úplný systém neslučitelných jevů a $P(B_i) \neq 0$ pro všechna $i = 1, \dots, n$, pak pro libovolný jev A platí

$$P(A) = \sum_{i=1}^n P(B_i)P(A|B_i).$$

Bayesův vzorec. Pokud B_1, \dots, B_n je úplný systém neslučitelných jevů, $P(B_i) \neq 0$ pro všechna $i = 1, \dots, n$ a $P(A) > 0$, pak

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^n P(B_j)P(A|B_j)}$$

Náhodné veličiny

Funkční charakteristiky.

Distribuční funkce	$F(x) = P(X \leq x).$	$0 \leq F(x) \leq 1,$ F je neklesající a zprava spojitá, $\lim_{x \rightarrow \infty} F(x) = 1, \lim_{x \rightarrow -\infty} F(x) = 0,$ $P(a < X \leq b) = F(b) - F(a),$ $P(X = a) = F(a) - \lim_{x \rightarrow a-} F(x).$
Pravděpodobnostní funkce	$p(x) = P(X = x).$	$\sum_{x \in M} p(x) = 1,$ $F(x) = \sum_{t \leq x, t \in M} p(t).$
Hustota pravděpodobnosti	$f(x) : F(x) = \int_{-\infty}^x f(t) dt.$	$\int_{-\infty}^{\infty} f(x) dx = 1,$ $f(x) = F'(x),$ $P(a \leq X \leq b) = \int_a^b f(x) dx.$

Číselné charakteristiky.

Střední hodnota	$E(X) = \begin{cases} \sum_M x \cdot p(x), & \text{pro } X \text{ diskrétní,} \\ \int_{-\infty}^{\infty} x \cdot f(x) dx, & \text{pro } X \text{ spojitou.} \end{cases}$	$E(X \pm Y) = E(X) \pm E(Y),$ $E(a + bX) = a + b \cdot E(X),$ $E(X \cdot Y) = E(X) \cdot E(Y)$ pro X, Y nezávislé.
Rozptyl a směrodatná odchylka	$D(X) = E[(X - E(X))^2],$ $\sigma(X) = \sqrt{D(X)}.$	$D(X) \geq 0,$ $D(a + bX) = b^2 \cdot D(X),$ $D(X) = E(X^2) - E^2(X),$ $D(X + Y) = D(X) + D(Y)$ pro X, Y nezávislé.
α -kvantil x_α (spojité náh. vel.)	$F(x_\alpha) = \alpha$ pro $F(x)$ rostoucí, případně $\int_{-\infty}^{x_\alpha} f(x) dx = \alpha.$	

Vybraná rozdělení pravděpodobnosti

Binomické rozdělení. $X \sim \text{Bi}(n, \pi)$, $n \in \mathbb{N}, 0 < \pi < 1$

$$p(x) = \binom{n}{x} \pi^x (1 - \pi)^{n-x}, \quad x = 0, 1, \dots, n.$$

$$\begin{aligned} E(X) &= n\pi, \\ D(X) &= n\pi(1 - \pi), \\ (n + 1)\pi - 1 &\leq \text{Mod}(X) \leq (n + 1)\pi. \end{aligned}$$

Aproximace:

- (a) Pro $n > 30$ a $\pi < \frac{1}{10}$ lze použít $\text{Bi}(n, \pi) \approx \text{Po}(n\pi)$.
- (b) Pro $n\pi(1 - \pi) > 9$ lze použít $\text{Bi}(n, \pi) \approx \text{N}(n\pi, n\pi(1 - \pi))$.

Hypergeometrické rozdělení. $X \sim \text{Hg}(N, M, n)$, $N, M, n \in \mathbb{N}$

$$p(x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}, \quad \max\{0, n-N+M\} \leq x \leq \min\{n, M\}, x \in \mathbb{N}_0.$$

$$\begin{aligned} E(X) &= n\pi, \quad \text{kde } \pi = \frac{M}{N}, \\ D(X) &= n\pi(1 - \pi) \frac{N-n}{N-1}, \\ a - 1 &\leq \text{Mod}(X) \leq a, \quad \text{kde } a = \frac{(M+1)(n+1)}{N+2}. \end{aligned}$$

Poissonovo rozdělení. $X \sim \text{Po}(\lambda)$, $\lambda > 0$

$$p(x) = \frac{\lambda^x}{x!} e^{-\lambda}, \quad x = 0, 1, 2, \dots$$

$$\begin{aligned} E(X) &= D(X) = \lambda, \\ \lambda - 1 &\leq \text{Mod}(X) \leq \lambda. \end{aligned}$$

Rovnoměrné spojité rozdělení. $X \sim \text{Ro}(a, b)$, $a < b, a, b \in \mathbb{R}$

$$\begin{aligned} f(x) &= \frac{1}{b-a}, \quad a \leq x \leq b, \\ F(x) &= \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x < b, \\ 1, & x \geq b. \end{cases} \end{aligned}$$

$$\begin{aligned} E(X) &= x_{0.5} = \frac{a+b}{2}, \\ D(X) &= \frac{(b-a)^2}{12}, \\ \text{Mod}(X) &= \langle a, b \rangle. \end{aligned}$$

Normální rozdělení. $X \sim \text{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}, \sigma^2 > 0$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2} \frac{(x - \mu)^2}{\sigma^2}\right), \quad x \in \mathbb{R}.$$

$$\begin{aligned} E(X) &= x_{0.5} = \text{Mod}(X) = \mu, \\ D(X) &= \sigma^2. \end{aligned}$$

Platí:

$$\begin{aligned} F(x) &= \Phi\left(\frac{x - \mu}{\sigma}\right), & f(x) &= \frac{1}{\sigma} \varphi\left(\frac{x - \mu}{\sigma}\right), \\ \Phi(-x) &= 1 - \Phi(x), & \varphi(-x) &= \varphi(x) \end{aligned}$$

Dvourozměrný diskretní náhodný vektor

Diskretní vektor $\mathbf{X} = (X, Y)$ s oborem hodnot $M = M_X \times M_Y \subseteq \mathbb{R}^2$ je popsán pravděpodobnostní funkcí $p(x, y)$, pro kterou platí $p(x, y) \geq 0$, $\sum p(x, y) > 0$ na M a dále

$$\sum_{M_X} \sum_{M_Y} p(x, y) = 1.$$

Marginální pravděpodobnostní funkce náhodných veličin X a Y se dostanou

$$\begin{aligned} X : p_X(x) &= \sum_{y \in M_Y} p(x, y), \\ Y : p_Y(y) &= \sum_{x \in M_X} p(x, y). \end{aligned}$$

Náhodné veličiny X, Y jsou stochasticky nezávislé, pokud platí $p_X(x) \cdot p_Y(y) = p(x, y)$ pro všechna $(x, y) \in R^2$.

Kovariance

$$C(X, Y) = E[(X - E(X))(Y - E(Y))],$$

$$\begin{aligned} C(X, X) &= D(X), \\ C(X, Y) &= E(XY) - E(X) \cdot E(Y), \\ D(aX + bY) &= a^2 D(X) + b^2 D(Y) + 2ab C(X, Y), \\ C(X, Y) &= 0 \text{ pro } X, Y \text{ stochasticky nezávislé.} \end{aligned}$$

Korelační koeficient

$$\rho(X, Y) = \frac{C(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

$$\begin{aligned} -1 &\leq \rho(X, Y) \leq 1, \\ \rho(X, X) &= 1, \\ \rho(X, Y) = \pm 1 &\Rightarrow X, Y \text{ jsou lineárně závislé,} \\ \rho(X, Y) = 0 &\text{ pro } X, Y \text{ stochasticky nezávislé.} \end{aligned}$$

Bodové a intervalové odhady parametrů normálního a alternativního rozdělení

výběrový průměr	$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i,$	\bar{X} je nestranným konzistentním odhadem střední hodnoty, S^2 je nestranným konzistentním odhadem rozptylu, S_n^2 je konzistentním odhadem rozptylu, $S_n^2 = \overline{X^2} - \bar{X}^2$.
výběrový rozptyl	$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$	
	$S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$	
výběrová směrodatná odchylka	$S = \sqrt{S^2}, \quad S_n = \sqrt{S_n^2}.$	Pro náhodný výběr z $N(\mu, \sigma^2)$ jsou odhady \bar{X}, S^2 navíc nejlepší, pro náhodný výběr z $A(\pi)$ je \bar{X} nejlepší také.

Intervalové odhady parametrů normálního rozdělení. Nechtě X_1, \dots, X_n je náhodný výběr z $N(\mu, \sigma^2)$ a $\alpha \in (0, 1)$ je dané číslo. Symboly $t_\alpha(n), \chi_\alpha^2(n)$ značí α -kvantily Studentova $t(n)$ a Pearsonova $\chi^2(n)$ rozdělení s n stupni volnosti.

parametr	100(1 - α)% interval spolehlivosti		
	pravostranný	oboustranný	levostranný
$(\sigma^2 \text{ známý})$	$\left(-\infty, \bar{X} + u_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}\right)$	$\left\langle \bar{X} - u_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + u_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right\rangle$	$\left\langle \bar{X} - u_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}, \infty \right\rangle$
$(\sigma^2 \text{ neznámý})$	$\left(-\infty, \bar{X} + t_{1-\alpha}(n-1) \cdot \frac{S}{\sqrt{n}}\right)$	$\left\langle \bar{X} - t_{1-\frac{\alpha}{2}}(n-1) \cdot \frac{S}{\sqrt{n}}, \bar{X} + t_{1-\frac{\alpha}{2}}(n-1) \cdot \frac{S}{\sqrt{n}} \right\rangle$	$\left\langle \bar{X} - t_{1-\alpha}(n-1) \cdot \frac{S}{\sqrt{n}}, \infty \right\rangle$
$(\mu \text{ neznámý})$	$\left\langle 0, \frac{(n-1) \cdot S^2}{\chi_\alpha^2(n-1)} \right\rangle$	$\left\langle \frac{(n-1) \cdot S^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)}, \frac{(n-1) \cdot S^2}{\chi_{\frac{\alpha}{2}}^2(n-1)} \right\rangle$	$\left\langle \frac{(n-1) \cdot S^2}{\chi_{1-\alpha}^2(n-1)}, \infty \right\rangle$

Intervalový odhad parametru alternativního rozdělení. Nechtě X_1, \dots, X_n je náhodný výběr z rozdělení $A(\pi)$ a $\alpha \in (0, 1)$ dané číslo. Označme $p = \bar{X}$ bodový odhad parametru π .

parametr	100(1 - α)% interval spolehlivosti pro $n > 30$		
	pravostranný	oboustranný	levostranný
π	$\left\langle 0, p + u_{1-\alpha} \cdot \sqrt{\frac{p \cdot (1-p)}{n}} \right\rangle$	$\left\langle p - u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{p \cdot (1-p)}{n}}, p + u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{p \cdot (1-p)}{n}} \right\rangle$	$\left\langle p - u_{1-\alpha} \cdot \sqrt{\frac{p \cdot (1-p)}{n}}, 1 \right\rangle$

Testy statistických hypotéz

Jednovýběrové testy hypotéz o parametrech.

Tabulka 1: Testy hypotéz o parametrech normálního rozdělení na hladině významnosti α .

Test H_0 proti H	Testové kritérium	Kritický obor
nulová hypotéza $H_0 : \mu = \mu_0$	$T = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$ pro σ^2 známé	$W_1 = \left(-\infty, -u_{1-\frac{\alpha}{2}}\right) \cup \left(u_{1-\frac{\alpha}{2}}, \infty\right)$ $W_2 = (u_{1-\alpha}, \infty)$ $W_3 = (-\infty, -u_{1-\alpha})$
alternativa $H_1 : \mu \neq \mu_0$ $H_2 : \mu > \mu_0$ $H_3 : \mu < \mu_0$	$T = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$ pro σ^2 neznámé	$W_1 = \left(-\infty, -t_{1-\frac{\alpha}{2}}(n-1)\right) \cup \left(t_{1-\frac{\alpha}{2}}(n-1), \infty\right)$ $W_2 = (t_{1-\alpha}(n-1), \infty)$ $W_3 = (-\infty, -t_{1-\alpha}(n-1))$
nulová hypotéza $H_0 : \sigma^2 = \sigma_0^2$	$T = \frac{(n-1) \cdot S^2}{\sigma_0^2}$ pro μ neznámé	$W_1 = \left\langle 0, \chi_{\frac{\alpha}{2}}^2(n-1) \right\rangle \cup \left\langle \chi_{1-\frac{\alpha}{2}}^2(n-1), \infty \right\rangle$ $W_2 = \left\langle \chi_{1-\alpha}^2(n-1), \infty \right\rangle$ $W_3 = \left\langle 0, \chi_\alpha^2(n-1) \right\rangle$
alternativa $H_1 : \sigma^2 \neq \sigma_0^2$ $H_2 : \sigma^2 > \sigma_0^2$ $H_3 : \sigma^2 < \sigma_0^2$		

Tabulka 2: Test hypotézy o parametru π alternativního rozdělení na hladině významnosti α . Statistika $p = \bar{X}$ značí nestranný odhad parametru.

Test H_0 proti H	Testové kritérium	Kritický obor
nulová hypotéza $H_0 : \pi = \pi_0$	$T = \frac{p - \pi_0}{\sqrt{\pi_0 \cdot (1 - \pi_0)}} \cdot \sqrt{n}$	$W_1 = \left(-\infty, -u_{1-\frac{\alpha}{2}}\right) \cup \left(u_{1-\frac{\alpha}{2}}, \infty\right)$ $W_2 = (u_{1-\alpha}, \infty)$ $W_3 = (-\infty, -u_{1-\alpha})$
alternativa $H_1 : \pi \neq \pi_0$		
$H_2 : \pi > \pi_0$		
$H_3 : \pi < \pi_0$		

Dvouvýběrové testy hypotéz o parametrech.

Tabulka 3: Test shody rozptylů dvou normálních rozdělení na hladině významnosti α . Statistiky S_x^2, S_y^2 značí nestranné odhady rozptylů příslušných náhodných výběrů.

Test H_0 proti H	Testové kritérium	Kritický obor
nulová hypotéza $H_0 : \sigma_x^2 = \sigma_y^2$	$F = \frac{S_x^2}{S_y^2}$	$W_1 = \langle 0, F_{\frac{\alpha}{2}}(n_x - 1, n_y - 1) \rangle \cup \left(F_{1-\frac{\alpha}{2}}(n_x - 1, n_y - 1), \infty \right)$ $W_2 = (F_{1-\alpha}(n_x - 1, n_y - 1), \infty)$ $W_3 = \langle 0, F_{\alpha}(n_x - 1, n_y - 1) \rangle$
alternativa		
$H_1 : \sigma_x^2 \neq \sigma_y^2$		
$H_2 : \sigma_x^2 > \sigma_y^2$ $H_3 : \sigma_x^2 < \sigma_y^2$		

Tabulka 4: Test shody středních hodnot dvou normálních rozdělení na hladině významnosti α . Statistiky S_x^2, S_y^2, S_d^2 značí nestranné odhady rozptylů příslušných náhodných výběrů.

Test H_0 proti H	Testové kritérium	Kritický obor
nulová hypotéza $H_0 : \sigma_x^2 = \sigma_y^2$	pro $\sigma_x^2 = \sigma_y^2$:	$W_1 = \left(-\infty, -t_{1-\frac{\alpha}{2}}(\nu) \right) \cup \left(t_{1-\frac{\alpha}{2}}(\nu), \infty \right)$
	$T = \frac{\bar{X} - \bar{Y}}{S} \sqrt{\frac{n_x \cdot n_y}{n_x + n_y}}$, kde $S = \sqrt{\frac{(n_x-1)S_x^2 + (n_y-1)S_y^2}{n_x + n_y - 2}}$	$W_2 = (t_{1-\alpha}(\nu), \infty)$ $W_3 = (-\infty, -t_{1-\alpha}(\nu))$ $\nu = n_x + n_y - 2$
alternativa $H_1 : \mu_x \neq \mu_y$ $H_2 : \mu_x > \mu_y$ $H_3 : \mu_x < \mu_y$	pro $\sigma_x^2 \neq \sigma_y^2$:	$W_1 = \left(-\infty, -t_{1-\frac{\alpha}{2}}(\nu) \right) \cup \left(t_{1-\frac{\alpha}{2}}(\nu), \infty \right)$
	$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}}$	$W_2 = (t_{1-\alpha}(\nu), \infty)$ $W_3 = (-\infty, -t_{1-\alpha}(\nu))$ $\nu \approx \frac{\left(\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y} \right)^2}{\frac{1}{n_x-1} \left(\frac{S_x^2}{n_x} \right)^2 + \frac{1}{n_y-1} \left(\frac{S_y^2}{n_y} \right)^2}$
	pro závislé výběry:	$W_1 = \left(-\infty, -t_{1-\frac{\alpha}{2}}(n-1) \right) \cup \left(t_{1-\frac{\alpha}{2}}(n-1), \infty \right)$
	$T = \frac{\bar{D}}{\sqrt{S_d^2}} \cdot \sqrt{n}$, kde $D_i = X_i - Y_i$	$W_2 = (t_{1-\alpha}(n-1), \infty)$ $W_3 = (-\infty, -t_{1-\alpha}(n-1))$

Tabulka 5: Test shody parametrů π_x, π_y dvou alternativních rozdělení na hladině významnosti α . Statistiky $p_x = \bar{X}, p_y = \bar{Y}$ značí nestranné odhady parametrů příslušných náhodných výběrů.

Test H_0 proti H	Testové kritérium	Kritický obor
nulová hypotéza $H_0 : \pi_x = \pi_y$	$U = \frac{p_x - p_y}{\sqrt{\frac{p_x \cdot (1-p_x)}{n_x} + \frac{p_y \cdot (1-p_y)}{n_y}}}$	$W_1 = \left(-\infty, -u_{1-\frac{\alpha}{2}} \right) \cup \left(u_{1-\frac{\alpha}{2}}, \infty \right)$ $W_2 = (u_{1-\alpha}, \infty)$ $W_3 = (-\infty, -u_{1-\alpha})$
alternativa		
$H_1 : \pi_x \neq \pi_y$		
$H_2 : \pi_x > \pi_y$ $H_3 : \pi_x < \pi_y$		

Pearsonův χ^2 test dobré shody. $H_0 : X \sim F(x; \hat{\theta})$ proti $H : X \not\sim F(x; \theta)$ pro libovolné $\theta = (\theta_1, \dots, \theta_m) \in \Theta$.
Testová statistika

$$T = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i} \sim \chi^2(k - m - 1),$$

Podmínky použitelnosti:

- (i) $n \cdot p_i \geq 1$ pro všechny třídy,
- (ii) $n \cdot p_i \geq 5$ pro 80 % tříd.

kritický obor $W = (\chi_{1-\alpha}^2(k - m - 1), \infty)$.