Tahák

Pravděpodobnost

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

$P(A \cap B) = P(A) \cdot P(B)$ ⇕ A, B nezávislé

$P(\bar{A}) = 1 - P(A)$ - opač.

$P(A) = \sum_{\omega \in A} P(\omega)$ - diskr.

$P(A) = \frac{\mu(A)}{\mu(\Omega)}$ - geom.

$$\boxed{P(B|A) = \frac{P(A \cap B)}{P(A)} \quad P(A \cap B) = P(A) \cdot P(B|A)}$$

Bayesův vzorec

$P(A) = \sum_{i=1}^{n} P(H_i) \cdot P(A|H_i)$

$P(H_1|A) = \frac{P(H_1 \cap A)}{P(A)} = \frac{P(H_1) \cdot P(A|H_1)}{P(A)}$

Náhodná veličina (diskr.)

$F(x) = P(X \le x)$ - distr. f-ce

$p(x) = P(X = x)$ - pravd. f-ce

$\boxed{E(X) = \sum_{x \in M} x \cdot p(x)}$ - střed. hod-ta

$D(X) = E[X - E(X)]^2 = \sum_{x \in M} [x - E(X)]^2 \cdot p(x)$;

$\boxed{D(X) = E(X^2) - [E(X)]^2 \quad \text{kde } E(X^2) = \sum_{x \in M} x^2 \cdot p(x)}$ rozptyl

$\boxed{\sigma(X) = \sqrt{D(X)}}$ směrodat. odchylka

Rozdělení (diskr.)

Alternativní

(střelec střílí)

p-st π, $\pi \in (0,1)$

$X \sim A(\pi)$

$p(x) = \begin{cases} \pi^x (1-\pi)^{1-x} & x = 0,1 \\ 0 & \text{jinak} \end{cases}$

$E(X) = \pi$, $D(X) = \pi(1-\pi)$

pokus. 1. n

Binomické

(házíme kostkou)

$X \sim Bi(n, \pi)$

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

$E(X) = n\pi$

$D(X) = n\pi(1-\pi)$

Hypergeometrické

(vadné výrobky) N prvků, M vadných, n bereme

$X \sim Hg(N, M, n)$

$p(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$

$E(X) = n\frac{M}{N}$

$D(X) = n\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}$

Poissonovo

(narození dětí)

λ udál. za jednot. čas

$X \sim Po(\lambda)$

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

$E(x) = \lambda$, $D(X) = \lambda$

Geometrické

(Pepa hraje Frisbee, dokud netrefí Kolem j.d.)

počet úsp. před 1. neús.

$X \sim Ge(\pi)$; $p(x) = \pi^x (1-\pi)$, $x = 0, 1, \dots$

$E(X) = \frac{\pi}{1-\pi}$; $D(X) = \frac{\pi}{(1-\pi)^2}$

Rozdělení (spojité)

Rovnoměrné

$X \sim Ro(a, b)$; $f(x) = \begin{cases} \frac{1}{b-a} & x \in \langle a, b \rangle \\ 0 & \text{jinak} \end{cases}$

$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & x \in \langle a, b \rangle \\ 1 & x > b \end{cases}$ $\quad E(X) = \frac{a+b}{2}$ $\quad D(X) = \frac{(b-a)^2}{12}$

Exponenciální (předpok. Po)

X - doba mezi 2 udál.

$X \sim Exp(\lambda)$; $f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \le 0 \end{cases}$

$F(x) = \begin{cases} 1 - e^{-\lambda x} & x > 0 \\ 0 & x \le 0 \end{cases}$ $\quad E(X) = \frac{1}{\lambda}$ $\quad D(X) = \frac{1}{\lambda^2}$

Normální

$X \sim N(\mu, \sigma^2)$

Standardiz. norm.

$\mu = 0$, $\sigma^2 = 1$

$U \sim N(0,1)$

$\Rightarrow$ $X \sim N(\mu, \sigma^2)$, $U = \frac{X - \mu}{\sigma}$, $U \sim N(0,1)$

$$\boxed{P(X \le x) = P(\mu + \sigma U \le x) = P\left(U \le \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right) \quad \Phi(-u) = 1 - \Phi(u)}$$

Součet a průměr

$X_1, X_2, \dots X_n$ mají rozd. $N(\mu, \sigma^2)$

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

$Y \sim N(n\mu, n\sigma^2)$ $\quad \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

+ centrální limitní věta (všechny stejné rozd.)

$Y \overset{A}{\sim} N(n\mu, n\sigma^2)$ $\quad \bar{X} \overset{A}{\sim} N\left(\mu, \frac{\sigma^2}{n}\right)$

Náhodná veličina (spojitá)

$\int_{-\infty}^{\infty} f(x)\,dx = 1$; $F(x) = \int_{-\infty}^{x} f(t)\,dt$

$P(a \le X \le b) = \int_a^b f(x)\,dx$

$$\boxed{E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx \quad D(X) = E(X^2) - [E(X)]^2 \quad E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)\,dx}$$

$x_\alpha = F^{-1}(\alpha)$ - α-kvantil

0.5-kvantil - medián

Diskr. n. vektor

$\sum\sum p(x,y) = 1$; $p(x|y) = P(X = x \mid Y = y) = \frac{p(x,y)}{p_Y(y)}$, $p_Y(y) > 0$

$F(x,y) = \sum_{u \le x} \sum_{v \le y} p(u,v)$

nezávisl. $F(x,y) = F_X(x) \cdot F_Y(y)$

+ jen pro diskr. $p(x,y) = p_X(x) \cdot p_Y(y)$

$$\boxed{C(X,Y) = E(XY) - E(X)E(Y) \quad E(XY) = \sum_x \sum_y xy\,p(x,y)}$$

Kovariance

$$\boxed{\text{var}(X) = \begin{pmatrix} D(X) & C(X,Y) \\ C(Y,X) & D(Y) \end{pmatrix} \quad X \text{ - n.v. } (X,Y)'}$$

variační (kovarianční) matice

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

Korelační koeficient

$$\boxed{\text{cor}(X) = \begin{pmatrix} 1 & \rho(X,Y) \\ \rho(Y,X) & 1 \end{pmatrix}}$$

Korelační matice

Transformace n.v.

$Y = T(X) \rightsquigarrow F_Y(y) = P(Y \le y) = P(T(X) \le y) = P(X \le \text{něco}) = F_X(\text{něco})$

$p(x,y) = P(X = x, Y = y)$

$p_X(x) = P(X = x) = \sum_y p(x,y)$, $x \in \mathbb{R}$

$p_Y(y) = P(Y = y) = \sum_x p(x,y)$, $y \in \mathbb{R}$

Spojitý n. vektor

$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dy\,dx = 1$

$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,dv\,du$

$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$

$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$

E, D viz n.v. (s.)

Kov, kor viz D.n.v.

$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f(x,y)$ dxdy

$f(x|y) = \frac{f(x,y)}{f_Y(y)}$

$f(y|x) = \frac{f(x,y)}{f_X(x)}$

nezávisl.:

$F(x,y) = F_X(x) \cdot F_Y(y)$

$f(x,y) = f_X(x) \cdot f_Y(y)$

Bod. a inter. odhady

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ - výběr. průměr

$\hat{\sigma}^2 = s^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)$ - výběrový rozptyl

IS pro μ u N.r.

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

$\left(-\infty, \bar{x} + t_{1-\alpha}(n-1)\frac{s}{\sqrt{n}}\right)$ pravostr.

$\left(\bar{x} - t_{1-\alpha}(n-1)\frac{s}{\sqrt{n}}, \infty\right)$ levstr.

$\boxed{t_{1-\frac{\alpha}{2}}(\infty) = u_{1-\frac{\alpha}{2}}}$

IS pro σ^2 u N.r.

Ob. $\left(\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}, \frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right)$

Pr. $\left(0, \frac{(n-1)s^2}{\chi^2_{\alpha}(n-1)}\right)$

Lev. $\left(\frac{(n-1)s^2}{\chi^2_{1-\alpha}(n-1)}, \infty\right)$

IS pro μ (vel. roz. výb.)

$\left(\bar{x} - u_{1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}, \bar{x} + u_{1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\right)$ ob.

$\left(-\infty, \bar{x} + u_{1-\alpha}\frac{s}{\sqrt{n}}\right)$ pr.

$\left(\bar{x} - u_{1-\alpha}\frac{s}{\sqrt{n}}, \infty\right)$ lev.

IS pro μ u N.r. (známe σ^2)

stejně jako vel. roz. výb., ale místo s bude σ.

Testování hypotéz

T. μ N.r. (μ, σ^2) při zn. σ^2

$H_0: \mu = \mu_0$ $\quad U = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n}$

$H_1: \mu > \mu_0$ $\quad W: \{U : U \ge u_{1-\alpha}\}$

$\mu < \mu_0$ $\quad \{U : U \le -u_{1-\alpha}\}$

$\mu \ne \mu_0$ $\quad \{U : |U| \ge u_{1-\frac{\alpha}{2}}\}$

T. μ N.r. (μ, σ^2) (t-test)

$H_0: \mu = \mu_0$

$T = \frac{\bar{X} - \mu_0}{S}\sqrt{n}$

$H_1: \mu > \mu_0$ $\quad W: \{T : T \ge t_{1-\alpha}(n-1)\}$

$\mu < \mu_0$ $\quad \{T : T \le -t_{1-\alpha}(n-1)\}$

$\mu \ne \mu_0$ $\quad \{T : |T| \ge t_{1-\frac{\alpha}{2}}(n-1)\}$

T. μ pro velk. výběr

$H_0: \mu = \mu_0$

$U = \frac{\bar{X} - \mu_0}{S}\sqrt{n}$

$H_1: \mu > \mu_0$ $\quad W: \{U : U \ge u_{1-\alpha}\}$

$\mu < \mu_0$ $\quad \{U : U \le -u_{1-\alpha}\}$

$\mu \ne \mu_0$ $\quad \{U : |U| \ge u_{1-\frac{\alpha}{2}}\}$

T. σ^2 N.r. (μ, σ^2)

$H_0: \sigma^2 = \sigma_0^2$

$K = \frac{(n-1)S^2}{\sigma_0^2}$

H_1: $\sigma^2 > \sigma_0^2$ $\quad W: \{k : K \ge \chi^2_{1-\alpha}(n-1)\}$

$\sigma^2 < \sigma_0^2$ $\quad \{k : K \le \chi^2_{\alpha}(n-1)\}$

$\sigma^2 \ne \sigma_0^2$ $\quad \{k : K \le \chi^2_{\frac{\alpha}{2}}(n-1) \vee K \ge \chi^2_{1-\frac{\alpha}{2}}(n-1)\}$

T. param. A.r. pro vel. výb.

$H_0: \pi = \pi_0$

$U = \frac{P - \pi_0}{\sqrt{\pi_0(1-\pi_0)}}\sqrt{n}$

$P = \frac{\text{chceme}}{\text{počet}}$

$H_1: \pi > \pi_0$ $\quad W: \{U : U \ge u_{1-\alpha}\}$

$\pi < \pi_0$ $\quad \{U : U \le -u_{1-\alpha}\}$

$\pi \ne \pi_0$ $\quad \{U : |U| \ge u_{1-\frac{\alpha}{2}}\}$

T. shody μ N.r. (t-test) (nezáv.)

$H_0: \mu_X = \mu_Y$

$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}}$

stupňů volnosti: (zaokr. ↓)

$\nu \approx \frac{\left(\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}\right)^2}{\frac{1}{n_X - 1}\left(\frac{S_X^2}{n_X}\right)^2 + \frac{1}{n_Y - 1}\left(\frac{S_Y^2}{n_Y}\right)^2}$

$H_1: \mu_X > \mu_Y$ $\quad W: \{T : T \ge t_{1-\alpha}(\nu)\}$

$\mu_X < \mu_Y$ $\quad \{T : T \le -t_{1-\alpha}(\nu)\}$

$\mu_X \ne \mu_Y$ $\quad \{T : |T| \ge t_{1-\frac{\alpha}{2}}(\nu)\}$

T. shody μ pro vel. výb. (nezáv.)

$H_0: \mu_X = \mu_Y$

$U = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}}$

$H_1: \mu > \mu_0$ $\quad W: \{U : U \ge u_{1-\alpha}\}$

$\mu < \mu_0$ $\quad \{U : U \le -u_{1-\alpha}\}$

$\mu \ne \mu_0$ $\quad \{U : |U| \ge u_{1-\frac{\alpha}{2}}\}$

T. shody μ N.r. (párový t-test) (závislý)

$D_i = X_i - Y_i$

$H_0: \mu_X = \mu_Y$

$T = \frac{\bar{D}}{S_d}\sqrt{n}$

$H_1: \mu_X > \mu_Y$ $\quad W: \{T : T \ge t_{1-\alpha}(n-1)\}$

$\mu_X < \mu_Y$ $\quad \{T : T \le -t_{1-\alpha}(n-1)\}$

$\mu_X \ne \mu_Y$ $\quad \{T : |T| \ge t_{1-\frac{\alpha}{2}}(n-1)\}$

T. shody σ^2 N.r. (F_test) (nezáv.)

$H_0: \sigma_x^2 = \sigma_y^2$ Fisherovo - Snedecorovo r.

$F_q(\nu_1, \nu_2) = \frac{1}{F_{1-q}(\nu_2, \nu_1)}$

$F = \frac{S_x^2}{S_y^2}$ H_1: $\sigma_x^2 > \sigma_y^2$ W: $\{F : F \geq F_{1-\alpha}(\nu_1, \nu_2)\}$

$\sigma_x^2 < \sigma_y^2$ $\{F : F \leq F_\alpha(\nu_1, \nu_2)\}$

$\sigma_x^2 \neq \sigma_y^2$ $\{F : F \leq F_{\frac{\alpha}{2}}(\nu_1, \nu_2) \vee F \geq F_{1-\frac{\alpha}{2}}(\nu_1, \nu_2)\}$

$\nu_1 = n_x - 1$

$\nu_2 = n_y - 1$

Regresní analýza

Přímka

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$

$n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$

$\beta_0 \sum_{i=1}^{n} x_i + \beta_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i$

$S_e = \sum_{i=1}^{n} y_i^2 - \hat{\beta}_0 \sum_{i=1}^{n} y_i - \hat{\beta}_1 \sum_{i=1}^{n} x_i y_i$

$\hat{\sigma}^2 = s^2 = \frac{S_e}{n-2}$

$S_T = \sum_{i=1}^{n} (Y_i - \bar{Y})^2 = (n-1) S_y^2$

$R^2 = 1 - \frac{S_e}{S_T}$

Rovina

$n\beta_0 + \beta_1 \sum x_i + \beta_2 \sum z_i = \sum Y_i$

$\beta_0 \sum x_i + \beta_1 \sum x_i^2 + \beta_2 \sum x_i z_i = \sum x_i Y_i$

$\beta_0 \sum z_i + \beta_1 \sum x_i z_i + \beta_2 \sum z_i^2 = \sum z_i Y_i$

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 z$

$S_e = \sum Y_i^2 - \hat{\beta}_0 \sum Y_i - \hat{\beta}_1 \sum x_i Y_i - \beta_2 \sum z_i Y_i$

$\hat{\sigma}^2 = s^2 = \frac{S_e}{n-3}$ $R^2 = 1 - \frac{S_e}{S_T}$

$S_T = (n-1) S_y^2$

Parabola

$n\beta_0 + \beta_1 \sum x_i + \beta_2 \sum x_i^2 = \sum Y_i$

$\beta_0 \sum x_i + \beta_1 \sum x_i^2 + \beta_2 \sum x_i^3 = \sum x_i Y_i$

$\beta_0 \sum x_i^2 + \beta_1 \sum x_i^3 + \beta_2 \sum x_i^4 = \sum x_i^2 Y_i$

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2$

$S_e = \sum Y_i^2 - \hat{\beta}_0 \sum Y_i - \hat{\beta}_1 \sum x_i Y_i - \hat{\beta}_2 \sum x_i^2 Y_i$

$\hat{\sigma}^2 = s^2 = \frac{S_e}{n-3}$ $R^2 = 1 - \frac{S_e}{S_T}$

$S_T = (n-1) S_y^2$

Korelační analýza

$\hat{\rho} = \frac{S_{XY}}{\sqrt{S_X^2 S_Y^2}}$ - Pearsonův výběrový korelační koeficient

$S_{XY} = \frac{1}{n-1} \sum (X_i - \bar{X})(Y_i - \bar{Y}) = \frac{1}{n-1} \left(\sum X_i Y_i - n\bar{X}\bar{Y}\right)$

Test nezávislosti (dvourozm. N.r.)

H_0: X_i a Y_i jsou nezávislé

H_1: X_i a Y_i nejsou nezávislé

$T = \frac{\hat{\rho}}{\sqrt{1 - \hat{\rho}^2}} \sqrt{n-2}$

$W = \{T : |T| \geq t_{1-\frac{\alpha}{2}}(n-2)\}$

$R_1, \dots, R_n$ - pořadí vel. $X_1, \dots, X_n$ a $Q_1, \dots, Q_n$ - pořadí vel. $Y_1, \dots, Y_n$.

$\hat{\rho}_s = 1 - \frac{6}{n(n^2-1)} \sum_{i=1}^{n} (R_i - Q_i)^2$ - Spearmanův výběrový korelační koeficient

Test nezávislosti (dvourozm. spojité r.)

H_0: X_i a Y_i jsou nezávislé

H_1: X_i a Y_i nejsou nezávislé

$T = \frac{\hat{\rho}_s}{\sqrt{1 - (\hat{\rho}_s)^2}} \sqrt{n-2}$

$W = \{T : |T| \geq t_{1-\frac{\alpha}{2}}(n-2)\}$

0	$F_X(x)$	1
0	$F(x,y)$	$F_Y(y)$
0	0	0

Kombinatorická pravidla

Uspořádané výběry	Bez opakování	Variace bez opakování	$V_k(n) = \frac{n!}{(n-k)!}$
		Permutace bez opakování	$P(n) = n!$
	S opakováním	Variace s opakováním	$V'_k(n) = n^k$
		Permutace s opakováním	$P'_{n_1,\dots,n_k}(n) = \frac{n!}{n_1! \dots n_k!}$
Neuspořádané výběry	Bez opakování	Kombinace bez opakování	$C_k(n) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$
	S opakováním	Kombinace s opakováním	$C'_k(n) = \binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-1)!}$