

- ① Volba správného modelu
- ② Odhady parametrů
 - ① bodové
 - ② intervalové
- ③ Testování hypotéz

- Necht' X_1, X_2, \dots, X_n jsou nezávislé náhodné veličiny s rozdělením $N(0, 1)$. Pak náhodná veličina

$$Y = \sum_{i=1}^n X_i^2$$

má rozdělení χ_n^2 (čti "chí-kvadrát rozdělení s n stupni volnosti").

- Necht' X je náhodná veličina s rozdělením $N(0, 1)$ a Y na ní nezávislá náhodná veličina s rozdělením χ_n^2 . Pak náhodná veličina

$$Z = \frac{X}{\sqrt{Y}} \sqrt{n}$$

má rozdělení t_n (čti "Studentovo t-rozdělení s n stupni volnosti").

- Necht' U a V jsou nezávislé náhodné veličiny s rozděleními χ_n^2 , resp. χ_m^2 . Pak náhodná veličina

$$W = \frac{U/n}{V/m}$$

má rozdělení $F_{n,m}$ (čti "Fisherovo-Snedecorovo rozdělení s parametry n a m ").

Definice

Náhodný vektor $\mathbb{X} = (X_1, X_2, \dots, X_n)^T$ nezávislých, stejně rozdělených náhodných veličin s distribuční funkcí F_θ , která závisí na parametru θ , se nazývá náhodný výběr.

Definice

Funkce

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

náhodného výběru $\mathbb{X} = (X_1, X_2, \dots, X_n)^T$ se nazývá výběrový průměr a

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

se nazývá výběrový rozptyl. $S_n = \sqrt{S_n^2}$ je pak výběrová směrodatná odchylka.

Věta

Nechť $\mathbb{X} = (X_1, X_2, \dots, X_n)^T$ je náhodný výběr z rozdělení $N(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$. Pak

- 1 výběrový průměr \bar{X}_n a výběrový rozptyl S_n^2 jsou nezávislé náhodné veličiny,
- 2 výběrový průměr \bar{X}_n má rozdělení $N(\mu, \sigma^2/n)$,
- 3 náhodná veličina $(n-1)S_n^2/\sigma^2$ má rozdělení $\chi_{(n-1)}^2$,
- 4 náhodná veličina $T = \frac{\bar{X}_n - \mu}{S_n} \sqrt{n}$ má rozdělení $t_{(n-1)}$.

Definice

Nechť F je spojitá a monotónní distribuční funkce a $0 < \beta < 1$. Pak hodnotu z_β takovou, že $F(z_\beta) = \beta$, nazýváme β -kvantil rozdělení s distribuční funkcí F .

Poznámka

- 1 Výraz β -kvantil používaný ve statistice je vlastně hodnota $q(\beta)$ kvantilové funkce q . Jestliže tedy distribuční funkce F není spojitá nebo není monotónní, pak β -kvantil můžeme dodefinovat analogicky k definici kvantilové funkce q .

- 2 Pro náhodnou veličinu X s distribuční funkcí F a kvantily z_β je

$$P(z_{\alpha/2} < X < z_{1-\alpha/2}) = F(z_{1-\alpha/2}) - F(z_{\alpha/2}) = 1 - \alpha.$$

- 3 β -kvantily rozdělení používaných ve statistice budeme značit u_β pro normované normální rozdělení, $t_{\beta,n}$ pro rozdělení t_n a $\chi_{\beta,n}^2$ pro rozdělení χ_n^2 .

Definice

Nechť $(x_1, x_2, \dots, x_n)^T$ je realizace náhodného výběru $(X_1, X_2, \dots, X_n)^T$. Pak

$$F_{\text{emp}}(x) = \frac{\#\{x_i : x_i \leq x\}}{n},$$

kde $\#$ značí počet prvků, se nazývá empirická distribuční funkce.

Definice

Nechť $(x_1, x_2, \dots, x_n)^T$ je realizace náhodného výběru $(X_1, X_2, \dots, X_n)^T$, $F_{\text{emp}}(x)$ příslušná empirická distribuční funkce a z_β značí β -kvantil náhodné veličiny s distribuční funkcí F_{emp} . Pak hodnoty $z_{1/4}$, $z_{1/2}$ a $z_{3/4}$ se nazývají 1.kvartil, 2.kvartil (též "medián"), resp. 3.kvartil. Nejčastěji zatoupený prvek v realizaci náhodného výběru se nazývá modus.

Poznámka

Občas se 1.kvartil definuje jako $z^ = \max(x_i : F_{\text{emp}}(x_i) \leq 1/4)$ nebo $z^{**} = \min(x_i : F_{\text{emp}}(x_i) \geq 1/4)$, popř. jako $z^{***} = z^* + \frac{1}{4}(z^{**} - z^*)$. Analogicky se pak definuje i 2. a 3.kvartil.*

Sledovali jsme doby mezi příchody zákazníků (v minutách) a naměřili jsme těchto 21 hodnot:

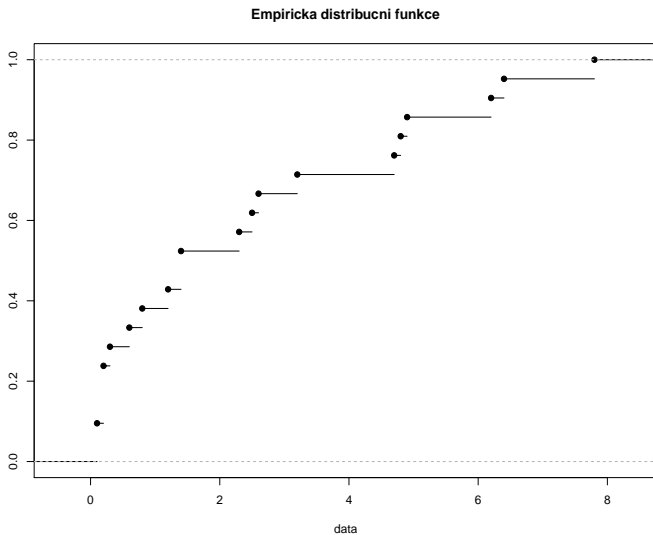
4.9, 6.2, 2.6, 0.6, 0.3, 2.3, 3.2, 1.4, 6.4, 4.8, 1.2
2.5, 0.2, 0.2, 0.8, 0.1, 0.1, 1.4, 7.8, 0.2, 4.7.

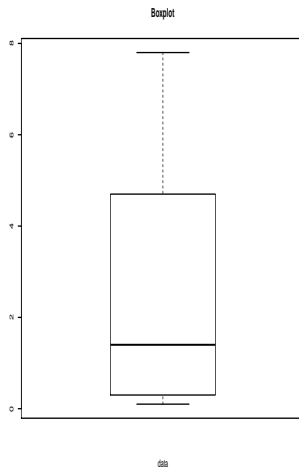
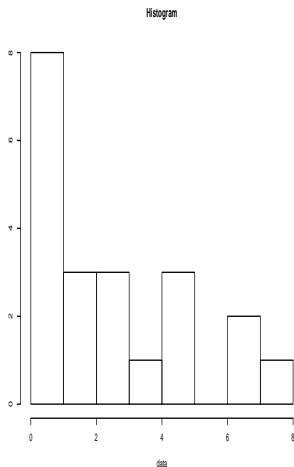
Pro přehlednost si hodnoty seřadíme od nejmenší po největší:

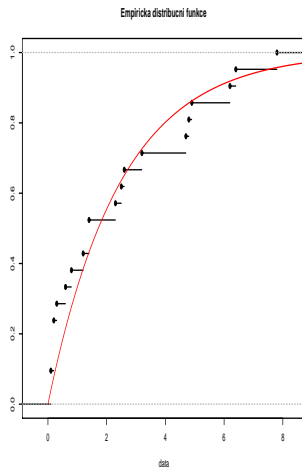
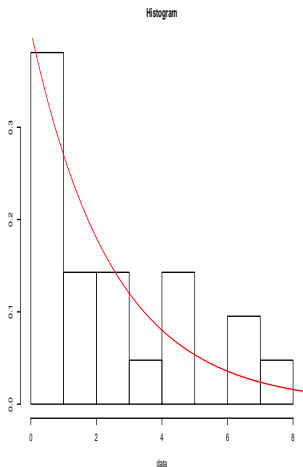
0.1, 0.1, 0.2, 0.2, 0.2, **0.3**, 0.6, 0.8, 1.2, 1.4, **1.4**,
2.3, 2.5, 2.6, 3.2, **4.7**, 4.8, 4.9, 6.2, 6.4, **7.8**.

Máme zde:

- výběrový průměr (pro danou realizaci) $\bar{X}_{21} = 2.471$,
- výběrový rozptyl (pro danou realizaci) $S_{21}^2 = 5.81$,
- výběrovou směrodatnou odchylku (pro danou realizaci) $S_{21} = 2.21$,
- 1.kvartil = 0.3, medián (tj. 2.kvartil) = 1.4 a 3.kvartil = 4.7,
- min = 0.1, max = 7.8, modus = 0.2.







Definice

Nechť $(x_1, x_2, \dots, x_n)^T$ je realizace náhodného výběru $(X_1, X_2, \dots, X_n)^T$ a rozdělení X_1, \dots, X_n závisí na parametru θ . Bodový odhad parametru θ je libovolná funkce $\hat{\theta}(X_1, X_2, \dots, X_n)$ náhodného výběru $(X_1, X_2, \dots, X_n)^T$, jejíž předpis nezávisí na θ .

Poznámka

Pro jednoduchost si můžeme bodový odhad představit jako hodnotu $\hat{\theta}$ získanou z realizace $(x_1, x_2, \dots, x_n)^T$ náhodného výběru $(X_1, X_2, \dots, X_n)^T$, kde tato hodnota co možná nejlépe odhaduje parametr θ . Jelikož však při opakování náhodného výběru získáváme různé realizace, můžeme pro každou z nich dostat jinou hodnotu odhadovaného parametru, tudíž bodový odhad je náhodná veličina.

Definice

Jestliže pro bodový odhad $\hat{\theta}(X_1, X_2, \dots, X_n)$ parametru θ platí, že $\mathbb{E}\hat{\theta}(X_1, X_2, \dots, X_n) = \theta$, pak tento odhad nazýváme *nestranným*.

Definice

Jestliže pro bodový odhad $\hat{\theta}(X_1, X_2, \dots, X_n)$ parametru θ platí, že

- ① $\lim_{n \rightarrow \infty} \mathbb{E}\hat{\theta}(X_1, X_2, \dots, X_n) = \theta$ (je tzv. *asymptoticky nestranný*),
- ② $\lim_{n \rightarrow \infty} \text{var}\hat{\theta}(X_1, X_2, \dots, X_n) = 0$,

pak tento odhad nazýváme *konzistentním*.

Definice

Jestliže existuje více *nestranných bodových odhadů*, pak ten s *nejmenším rozptylem* $\text{var}\hat{\theta}(X_1, X_2, \dots, X_n)$ nazýváme *eficientním*.

Nechť $(x_1, x_2, \dots, x_n)^T$ je realizace náhodného výběru $(X_1, X_2, \dots, X_n)^T$ a rozdělení náhodných veličin X_1, \dots, X_n závisí na parametrech $\theta_1, \dots, \theta_k \in \Theta$, kde Θ je množina parametrů.

Předpoklad: $\mathbb{E}X_1^i < \infty \quad \forall i = 1, \dots, k$ a $\mathbb{E}X_1^i$ závisí na $\theta_1, \dots, \theta_k$.

Metoda: Položíme do rovnosti teoretické a odhadnuté momenty, tj.

$$\mathbb{E}X_1^i = m_i, \quad \text{kde} \quad m_i = \frac{1}{n} \sum_{j=1}^n x_j^i \quad \text{pro všechna} \quad i = 1, \dots, k.$$

Takto získáme soustavu k rovnic o k neznámých $\theta_1, \dots, \theta_k$, jejichž řešením jsou hledané odhady parametrů $\hat{\theta}_1, \dots, \hat{\theta}_k$.

Alternativa: Je-li $k = 2$, pak místo i -tých momentů, $i = 1, 2$, můžeme položit $\mathbb{E}X_1 = \bar{x}_n$ a $\text{var}X_1 = s_n^2$, kde \bar{x}_n a s_n^2 jsou hodnoty výběrového průměru, resp. výběrového rozptylu, získané z dat.

Výhoda: Jednoduchost.

Nevýhoda: Řešení nemusí existovat.

Nechť $(x_1, x_2, \dots, x_n)^T$ je realizace náhodného výběru $(X_1, X_2, \dots, X_n)^T$ a rozdělení (tj. $P_\theta(X_1 = \cdot)$) v diskrétním případě nebo hustota f_θ ve spojitém případě) náhodných veličin X_1, \dots, X_n závisí na parametru θ .

Definice

Hodnota $\hat{\theta}$ se nazývá maximálně věrohodným odhadem, jestliže

$$\prod_{i=1}^n P_{\hat{\theta}}(X_1 = x_i) = \max_{\theta \in \Theta} \prod_{i=1}^n P_{\theta}(X_1 = x_i),$$

resp.

$$\prod_{i=1}^n f_{\hat{\theta}}(x_i) = \max_{\theta \in \Theta} \prod_{i=1}^n f_{\theta}(x_i).$$

Poznámka

Obvykle je jednodušší pracovat s logaritmy těchto součinů, abychom při hledání extrému funkce derivovali součet, nikoliv součin.

Pro náhodný výběr z diskrétního rozdělení:

- 1 Zkonstruujeme věrohodnostní funkci $L(\theta) = \prod_{i=1}^n P_{\theta}(X_1 = x_i)$.
- 2 Zkonstruujeme logaritmicko-věrohodnostní funkci $l(\theta) = \log L(\theta) = \sum_{i=1}^n \log P_{\theta}(X_1 = x_i)$.
- 3 Položíme $\frac{\partial l(\theta)}{\partial \theta} = 0$.
- 4 Řešením rovnice $\frac{\partial l(\theta)}{\partial \theta} = 0$ je hledaný maximálně věrohodný odhad $\hat{\theta}$.

Pro náhodný výběr ze spojitého rozdělení:

- 1 Zkonstruujeme věrohodnostní funkci $L(\theta) = \prod_{i=1}^n f_{\theta}(x_i)$.
- 2 Zkonstruujeme logaritmicko-věrohodnostní funkci $l(\theta) = \log L(\theta) = \sum_{i=1}^n \log f_{\theta}(x_i)$.
- 3 Položíme $\frac{\partial l(\theta)}{\partial \theta} = 0$.
- 4 Řešením rovnice $\frac{\partial l(\theta)}{\partial \theta} = 0$ je hledaný maximálně věrohodný odhad $\hat{\theta}$.

Definice

Nechť $(X_1, X_2, \dots, X_n)^T$ je náhodný výběr a $\alpha \in (0, 1)$.

- 1 Dvojice $(\theta_L^*(X_1, \dots, X_n), \theta_U^*(X_1, \dots, X_n))$ se nazývá $(1 - \alpha) \cdot 100 \%$ intervalový odhad (nebo též interval spolehlivosti; označení $(1 - \alpha) \cdot 100 \%$ -CI) parametru θ , jestliže

$$P(\theta_L^*(X_1, \dots, X_n) < \theta < \theta_U^*(X_1, \dots, X_n)) = 1 - \alpha.$$

- 2 $(\theta_D^*(X_1, \dots, X_n))$ se nazývá dolní $(1 - \alpha) \cdot 100 \%$ intervalový odhad parametru θ , jestliže

$$P(\theta_D^*(X_1, \dots, X_n) < \theta) = 1 - \alpha.$$

- 3 $(\theta_H^*(X_1, \dots, X_n))$ se nazývá horní $(1 - \alpha) \cdot 100 \%$ intervalový odhad parametru θ , jestliže

$$P(\theta_H^*(X_1, \dots, X_n) > \theta) = 1 - \alpha.$$

Věta

Nechť $(X_1, X_2, \dots, X_n)^T$ je náhodný výběr z rozdělení $N(\mu, \sigma^2)$, $\mu \in \mathbb{R}$ je neznámý parametr a $\sigma^2 > 0$ je známá konstanta. Pak

- ❶ $(\bar{X}_n - u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{X}_n + u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}})$ je $(1 - \alpha) \cdot 100$ % intervalový odhad parametru μ ,
- ❷ $\bar{X}_n - u_{1-\alpha} \frac{\sigma}{\sqrt{n}}$ je dolní $(1 - \alpha) \cdot 100$ % intervalový odhad parametru μ ,
- ❸ $\bar{X}_n + u_{1-\alpha} \frac{\sigma}{\sqrt{n}}$ je horní $(1 - \alpha) \cdot 100$ % intervalový odhad parametru μ .

Věta

Nechť $(X_1, X_2, \dots, X_n)^T$ je náhodný výběr z rozdělení $N(\mu, \sigma^2)$ a oba parametry $\mu \in \mathbb{R}$, $\sigma^2 > 0$ jsou neznámé. Pak

- ❶ $(\bar{X}_n - t_{1-\alpha/2, n-1} \frac{S_n}{\sqrt{n}}, \bar{X}_n + t_{1-\alpha/2, n-1} \frac{S_n}{\sqrt{n}})$ je $(1 - \alpha) \cdot 100$ % intervalový odhad parametru μ ,
- ❷ $\bar{X}_n - t_{1-\alpha, n-1} \frac{S_n}{\sqrt{n}}$ je dolní $(1 - \alpha) \cdot 100$ % intervalový odhad parametru μ ,
- ❸ $\bar{X}_n + t_{1-\alpha, n-1} \frac{S_n}{\sqrt{n}}$ je horní $(1 - \alpha) \cdot 100$ % intervalový odhad parametru μ .
- ❹ $\left(\frac{(n-1)S_n^2}{\chi_{1-\alpha/2, n-1}^2}, \frac{(n-1)S_n^2}{\chi_{\alpha/2, n-1}^2} \right)$ je $(1 - \alpha) \cdot 100$ % intervalový odhad parametru σ^2 ,
- ❺ $\frac{(n-1)S_n^2}{\chi_{1-\alpha, n-1}^2}$ je dolní $(1 - \alpha) \cdot 100$ % intervalový odhad parametru σ^2 ,
- ❻ $\frac{(n-1)S_n^2}{\chi_{\alpha, n-1}^2}$ je horní $(1 - \alpha) \cdot 100$ % intervalový odhad parametru σ^2 .

Věta

Nechť $(X_1, X_2, \dots, X_n)^T$ je náhodný výběr z libovolného rozdělení s rozptylem $0 < \sigma^2 < \infty$. Pak asymptotický $(1 - \alpha) \cdot 100$ % intervalový odhad střední hodnoty $\mu = \mathbb{E}X$ je

$$(\bar{X}_n - u_{1-\alpha/2} \frac{S_n}{\sqrt{n}}, \bar{X}_n + u_{1-\alpha/2} \frac{S_n}{\sqrt{n}}).$$

Důkaz

Připomeňme, že pro velká n platí $\frac{S_n}{\sigma} \rightarrow 1$, tj. S_n je aproximace σ . Z CLV víme, že $\frac{\sum X_i - n\mu}{\sqrt{n\sigma^2}}$ má přibližně (asymptoticky) normální rozdělení, tj.

$$P(u_{\frac{\alpha}{2}} \leq \frac{\sum X_i - n\mu}{\sqrt{n\sigma^2}} \leq u_{1-\frac{\alpha}{2}}) = 1 - \alpha$$

$$P(u_{\frac{\alpha}{2}} \leq \frac{\sum X_i - n\mu}{\sqrt{n}S_n} \leq u_{1-\frac{\alpha}{2}}) = 1 - \alpha$$

$$P\left(\frac{\sum X_i}{n} + u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} \geq \mu \geq \frac{\sum X_i}{n} + u_{\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}}\right) = 1 - \alpha$$

$$P\left(\bar{X}_n + u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} \geq \mu \geq \bar{X}_n - u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}}\right) = 1 - \alpha,$$

což je definice $(1 - \alpha) \cdot 100\%$ intervalového odhadu parametru μ .

Věta

- ① *Nechť $(X_1, X_2, \dots, X_n)^T$ je náhodný výběr z rozdělení $Alt(p)$, $0 < p < 1$. Pak $(1 - \alpha) \cdot 100$ % intervalový odhad parametru p je*

$$(\bar{X}_n - u_{1-\alpha/2} \sqrt{\frac{\bar{X}_n(1 - \bar{X}_n)}{n}}, \bar{X}_n + u_{1-\alpha/2} \sqrt{\frac{\bar{X}_n(1 - \bar{X}_n)}{n}}).$$

- ② *Nechť $(X_1, X_2, \dots, X_n)^T$ je náhodný výběr z rozdělení $Po(\lambda)$, $\lambda > 0$. Pak $(1 - \alpha) \cdot 100$ % intervalový odhad parametru λ je*

$$(\bar{X}_n - u_{1-\alpha/2} \sqrt{\frac{\bar{X}_n}{n}}, \bar{X}_n + u_{1-\alpha/2} \sqrt{\frac{\bar{X}_n}{n}}).$$

Důkaz

Důkaz plyne z předešlé věty a faktu, že pro alternativní rozdělení je $\mathbb{E}X = p$, $\text{var}X = p(1 - p)$ a pro Poissonovo rozdělení je $\mathbb{E}X = \text{var}X = \lambda$.

- Necht' $(X_1, X_2, \dots, X_n)^T$ je náhodný výběr z rozdělení, které závisí na parametru $\theta \in \Theta$.
- Tvzení, že θ patří do nějaké množiny Θ_0 , se nazývá nulová hypotéza (značíme $H_0 : \theta \in \Theta_0$).
- Na základě náhodného výběru $(X_1, X_2, \dots, X_n)^T$ testujeme nulovou hypotézu vůči alternativní hypotéze $H_A : \theta \in \Theta \setminus \Theta_0$. K tomu stanovíme množinu W (tzv. kritický obor) tak, že H_0 zamítáme ve prospěch H_A , jestliže $\mathbb{X} \in W$, v opačném případě H_0 ve prospěch H_A nezamítáme.

Poznámka

Většinou testujeme $H_0 : \theta = \theta_0$, kde θ_0 je konkrétní hodnota, takže přirozenou alternativou je $H_A : \theta \neq \theta_0$. Občas však dává větší smysl testovat H_0 vůči $H_A : \theta > \theta_0$ nebo $H_A : \theta < \theta_0$, jelikož opačná situace nedává v tu chvíli praktický smysl.

Při testování mohou nastat následující situace:

- H_0 platí a test ji nezamítá ✓
- H_0 neplatí a test ji zamítá ✓
- H_0 platí a test ji zamítá → chyba prvního druhu
- H_0 neplatí a test ji nezamítá → chyba druhého druhu

Testovací hladina:

Zvolíme hodnotu α (obvykle 0.05, někdy 0.01 nebo 0.1) a kritický obor W konstruueme tak, aby chyba prvního druhu nebyla větší než (obvykle byla rovna) α . Takové α se nazývá testovací hladina.

Testování střední hodnoty normálního rozdělení:

Jednovýběrový t -test

99

- Nechť $(X_1, X_2, \dots, X_n)^T$ je náhodný výběr z rozdělení $N(\mu, \sigma^2)$, kde $\sigma^2 > 0$ a ani jeden parametr není známý. Víme, že náhodná veličina

$$T = \frac{\bar{X}_n - \mu}{S_n} \sqrt{n} \sim t_{n-1}.$$

- Testování $H_0 : \mu = \mu_0$ vůči $H_A : \mu \neq \mu_0$ tedy probíhá následovně:
 - 1 Spočítáme tzv. testovou statistiku (hodnotu) $T_0 = \frac{\bar{X}_n - \mu_0}{S_n} \sqrt{n}$.
 - 2 Jestliže $|T_0| \geq t_{1-\alpha/2, n-1}$, zamítáme H_0 ve prospěch H_A , v opačném případě H_0 ve prospěch H_A nezamítáme.
- Testování $H_0 : \mu = \mu_0$ vůči $H_A : \mu > \mu_0$ je analogické:
 - 1 Spočítáme hodnotu $T_0 = \frac{\bar{X}_n - \mu_0}{S_n} \sqrt{n}$.
 - 2 Je-li $T_0 \geq t_{1-\alpha, n-1}$, zamítáme H_0 ve prospěch H_A , v opačném případě H_0 ve prospěch H_A nezamítáme.
- Nulová hypotéza $H_0 : \mu = \mu_0$ vůči $H_A : \mu < \mu_0$ je pak zamítnutá v případě, že $T_0 \leq t_{\alpha, n-1} = -t_{1-\alpha, n-1}$.

- Používá se tehdy, když pozorujeme párový znak na jednom objektu (např. dioptrie na levém a pravém oku, dobu zpracování stejných dat jednou a druhou metodou atd.).
- Máme náhodný výběr $(Y_1, Z_1), (Y_2, Z_2), \dots, (Y_n, Z_n)^T$ a testujeme $H_0 : \mathbb{E}Y_i - \mathbb{E}Z_i = \mu_0$ (většinou $\mu_0 = 0$, tj. shodu středních hodnot) vůči některé z alternativních hypotéz zmíněných výše.
- Položíme

$$X_1 = Y_1 - Z_1, \dots, X_n = Y_n - Z_n$$

a jestliže X_1, \dots, X_n pochází z normálního rozdělení, použijeme jednovýběrový t -test popsany výše.

- Uvažujme dva nezávislé výběry, a to $(X_1, X_2, \dots, X_m)^T$ z rozdělení $N(\mu_1, \sigma^2)$ a $(Y_1, Y_2, \dots, Y_n)^T$ z $N(\mu_2, \sigma^2)$, kde $\sigma^2 > 0$.
- Označme \bar{X} výběrový průměr náhodného výběru $(X_1, X_2, \dots, X_m)^T$, \bar{Y} výběrový průměr $(Y_1, Y_2, \dots, Y_n)^T$, S_X^2 výběrový rozptyl $(X_1, X_2, \dots, X_m)^T$ a S_Y^2 výběrový rozptyl $(Y_1, Y_2, \dots, Y_n)^T$.
- Náhodná veličina

$$T = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{(m-1)S_X^2 + (n-1)S_Y^2}} \sqrt{\frac{mn(m+n-2)}{m+n}} \sim t_{m+n-2}.$$

- Testování $H_0 : \mu_1 - \mu_2 = \mu_0$ vůči $H_A : \mu_1 - \mu_2 \neq \mu_0$ tedy probíhá následovně:
 - 1 Spočteme $T_0 = \frac{\bar{X} - \bar{Y} - \mu_0}{\sqrt{(m-1)S_X^2 + (n-1)S_Y^2}} \sqrt{\frac{mn(m+n-2)}{m+n}}$.
 - 2 Je-li $|T_0| \geq t_{1-\alpha/2, m+n-2}$, zamítáme H_0 ve prospěch H_A , v opačném případě H_0 ve prospěch H_A nezamítáme.

- Uvažujme dva nezávislé výběry, a to $(X_1, X_2, \dots, X_m)^T$ z rozdělení $N(\mu_1, \sigma_1^2)$ a $(Y_1, Y_2, \dots, Y_n)^T$ z $N(\mu_2, \sigma_2^2)$, kde $\sigma_1^2 > 0$ a $\sigma_2^2 > 0$.
- Označme S_X^2 výběrový rozptyl $(X_1, X_2, \dots, X_m)^T$ a S_Y^2 výběrový rozptyl $(Y_1, Y_2, \dots, Y_n)^T$.
- Náhodná veličina

$$F = \frac{S_X^2}{S_Y^2} \sim F_{m-1, n-1}.$$

- Testování $H_0 : \sigma_1^2 = \sigma_2^2$ vůči $H_A : \sigma_1^2 \neq \sigma_2^2$ tedy probíhá následovně:
 - 1 Spočteme $F_0 = \frac{S_X^2}{S_Y^2}$.
 - 2 Je-li $F_0 \leq F_{\alpha/2, m-1, n-1}$ nebo $F_0 \geq F_{1-\alpha/2, m-1, n-1}$, zamítáme H_0 ve prospěch H_A , v opačném případě H_0 ve prospěch H_A nezamítáme.

Multinomické rozdělení

- Je zobecněním rozdělení binomického ve smyslu, že uvažujeme n -krát opakovaný náhodný pokus, který může pokaždé skončit nějakým z výsledků A_1, A_2, \dots, A_k (nikoliv pouze jedním ze dvou výsledků "úspěch" nebo "neúspěch").
- Pro $i = 1, \dots, k$ označme $p_i = P(A_i)$ (kde zřejmě $\sum_{i=1}^k p_i = 1$) a X_i počet výsledků A_i ve výše zmíněných n pokusech. Pak

$$P(X_1 = x_1, \dots, X_k = x_k) = \frac{n!}{x_1! \dots x_k!} p_1^{x_1} \dots p_k^{x_k}, \quad \text{kde } \sum_{i=1}^k x_i = n.$$

- Rozdělení náhodného vektoru $(X_1, X_2, \dots, X_k)^T$ se nazývá multinomickým.

- Testujeme nulovou hypotézu

H_0 : "marginální pravděpodobnosti jsou rovny hodnotám p_1, \dots, p_k "
proti alternativní hypotéze

H_A : "alespoň jedno p_i je jiné".

- Test probíhá následovně:

① Spočteme hodnotu $\chi_0^2 = \sum_{i=1}^k \frac{(X_i - np_i)^2}{np_i}$.

② Jestliže $\chi_0^2 > \chi_{1-\alpha, k-1}^2$, zamítáme H_0 ve prospěch H_A , v opačném případě H_0 ve prospěch H_A nezamítáme.

- Mějme náhodný výběr $(Y_1, Z_1), (Y_2, Z_2), \dots, (Y_n, Z_n)$, kde pro $k = 1, \dots, n$ nabývají Y_k a Z_k hodnot $1, \dots, r$, resp. $1, \dots, c$.
- Testujeme nulovou hypotézu H_0 : " Y a Z jsou vzájemně nezávislé" vůči alternativní hypotéze H_A : " Y a Z nejsou nezávislé".
- Označme n_{ij} počet dvojic $(Y_k = i, Z_k = j)$. Pak matici o rozměrech $r \times c$ s prvky n_{ij} nazýváme kontingenční tabulkou a prvkům n_{ij} říkáme sdružené četnosti.
- Marginalní četnosti jsou

$$n_{i.} = \sum_j n_{ij}, \quad n_{.j} = \sum_i n_{ij}.$$

- Test nezávislosti probíhá následovně:

- 1 Spočteme hodnotu

$$\chi_0^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - \frac{n_{i.} n_{.j}}{n})^2}{\frac{n_{i.} n_{.j}}{n}}.$$

- 2 Jestliže $\chi_0^2 \geq \chi_{1-\alpha, (r-1)(c-1)}^2$, zamítáme H_0 ve prospěch H_A , v opačném případě H_0 ve prospěch H_A nezamítáme.

Definice

Nechť (Ω, \mathcal{A}, P) je pravděpodobnostní prostor a $T \subset \mathbb{R}$. Rodina reálných náhodných veličin $\{X_t, t \in T\}$ definovaných na (Ω, \mathcal{A}, P) se nazývá náhodný (nebo také stochastický) proces.

Definice

Je-li $T = \mathbb{Z}$ nebo $T = \mathbb{N}$, mluvíme o náhodném procesu s diskrétním časem. Je-li $T = [a, b]$, kde $-\infty \leq a < b \leq \infty$, mluvíme o náhodném procesu se spojitým časem.

Definice

Dvojice (S, \mathcal{E}) , kde S je množina hodnot náhodných veličin X_t a \mathcal{E} je σ -algebra na množině S , se nazývá stavový prostor.

Definice

Pokud náhodné veličiny X_t nabývají pouze diskrétních hodnot, mluvíme o náhodném procesu s diskrétními stavy. Pokud náhodné veličiny X_t nabývají spojitých hodnot, mluvíme o náhodném procesu se spojitými stavy.

Náhodný proces $\{X_t, t \in T\}$ můžeme chápat jako funkci dvou proměnných ω a t . Pro pevné t je tato funkce náhodnou veličinou, pro pevné ω se jedná o funkci jedné reálné proměnné t .

Definice

Mějme pevné $\omega \in \Omega$. Pak funkce $t \rightarrow X_t$ se nazývá trajektorie procesu $\{X_t, t \in T\}$.

Definice

Proces se nazývá spojitý, jsou-li všechny jeho trajektorie spojité.

Definice

Nechť $\{X_t, t \in T\}$ je náhodný proces takový, že pro každé $t \in T$ existuje střední hodnota $\mathbb{E}X_t$. Potom funkce $\mu_t = \mathbb{E}X_t$ definovaná na T se nazývá střední hodnota procesu $\{X_t\}$. Jestliže platí $\mathbb{E}|X_t|^2 < \infty$ pro všechna $t \in T$, potom funkce dvou proměnných definovaná na $T \times T$ předpisem $R(s, t) = \mathbb{E}(X_s - \mu_s)(X_t - \mu_t)$ se nazývá autokovarianční funkce procesu $\{X_t\}$. Hodnota $R(t, t)$ se nazývá rozptyl procesu $\{X_t\}$ v čase t .

Definice

Řekneme, že náhodný proces $\{X_t, t \in T\}$ je slabě stacionární, jestliže $R(s, t)$ je funkcí pouze rozdílu $s - t$, tj.

$$R(s, t) = \tilde{R}(s - t)$$

Důsledek:

$$R(s, t) = R(s + h, t + h)$$

pro každé $h \in \mathbb{R}$ takové, že $s + h \in T$ a $t + h \in T$.

Označme

$$F_{t_1, \dots, t_n}(x_1, \dots, x_n) = P(X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n).$$

Definice

Řekneme, že náhodný proces $\{X_t, t \in T\}$ je striktně stacionární, jestliže pro libovolné $n \in \mathbb{N}$, pro libovolná reálná x_1, \dots, x_n a pro libovolná reálná t_1, \dots, t_n a h taková, že $t_k \in T$, $t_k + h \in T$, $1 \leq k \leq n$, platí

$$F_{t_1, \dots, t_n}(x_1, \dots, x_n) = F_{t_1+h, \dots, t_n+h}(x_1, \dots, x_n). \quad (2)$$

Poznámka

Pro procesy s diskrétními stavy je vztah (2) je ekvivalentní vztahu

$$P(X_{t_1} = x_1, \dots, X_{t_n} = x_n) = P(X_{t_1+h} = x_1, \dots, X_{t_n+h} = x_n).$$

Definice

Nechť náhodné procesy $\{X_t, t \in T\}$ a $\{Y_t, t \in T\}$ jsou definované na stejném pravděpodobnostním prostoru s hodnotami ve stejném stavovém prostoru. Pak

- ❶ $\{X_t\}$ a $\{Y_t\}$ jsou stochasticky ekvivalentní, jestliže

$$P(X_t = Y_t) = P(\omega : X_t(\omega) = Y_t(\omega)) = 1, \quad \forall t \in T.$$

Říkáme také, že proces $\{X_t\}$ je stochastickou verzí, popř. modifikací, procesu $\{Y_t\}$.

- ❷ $\{X_t\}$ a $\{Y_t\}$ jsou nerozlišitelné, jestliže

$$P(X_t = Y_t, \forall t \in T) = P(\omega : X_t(\omega) = Y_t(\omega), \forall t \in T) = 1.$$

Mějme

- pravděpodobnostní prostor (Ω, \mathcal{A}, P) ,
- na něm posloupnost náhodných veličin $\{X_n, n \in \mathbb{N}\}$,
- stavový prostor (S, \mathcal{E}) , kde množina S je konečná nebo spočetná, bez újmy na obecnosti předpokládejme $S = \{0, 1, \dots, N\}$, resp. $S = \{0, 1, \dots\}$.

Definice

Posloupnost náhodných veličin $\{X_n, n \in \mathbb{N}\}$ nazveme Markovský řetězec s diskrétním časem, jestliže

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P(X_{n+1} = j | X_n = i)$$

pro všechna $n = 0, 1, \dots$ a všechna $i, j, i_{n-1}, \dots, i_0 \in S$ taková, že $P(X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) > 0$.

Nechť Y_1, Y_2, \dots jsou nezávislé stejně rozdělené náhodné veličiny nabývající hodnot ± 1 s pravděpodobnostmi $1/2$.

Definujme

$$X_0 = 0$$
$$X_n = \sum_{i=1}^n Y_i.$$

Pak posloupnost (proces, řetězec) $\{X_n, n \in \mathbb{N}\}$ se nazývá *náhodná procházka*.

Definice

Podmíněné pravděpodobnosti

- 1 $P(X_{n+1} = j | X_n = i) = p_{ij}(n, n+1)$ nazveme *pravděpodobnostmi přechodu ze stavu i v čase n do stavu j v čase $n+1$ nebo také pravděpodobnostmi přechodu 1.řádu*;
- 2 $P(X_{n+m} = j | X_n = i) = p_{ij}(n, n+m)$ nazveme *pravděpodobnostmi přechodu ze stavu i v čase n do stavu j v čase $n+m$ nebo také pravděpodobnostmi přechodu m -tého řádu*.

Definice

Jestliže pravděpodobnosti přechodu $p_{ij}(n, n+m)$ nezávisí na časových okamžicích n a $n+m$, ale pouze na rozdílu m , nazývá se příslušný Markovský řetězec homogenní.

- Uvažujme homogenní řetězec a označme $p_{ij} := p_{ij}(n, n+1)$.
- Tyto prvky lze seřadit do čtvercové matice $P = \{p_{ij}, i, j \in S\}$, pro niž zřejmě platí

$$p_{ij} \geq 0, \forall i, j \in S \quad \text{a} \quad \sum_{j \in S} p_{ij} = 1, \forall i \in S.$$

Definice

Matice $P = \{p_{ij}, i, j \in S\}$ se nazývá matice pravděpodobností přechodu.

Označme dále

$$p_i = P(X_0 = i), \quad \forall i \in S,$$

pro které zřejmě platí

$$p_i \geq 0, \forall i \in S \quad \text{a} \quad \sum_{i \in S} p_i = 1.$$

Definice

Vektor $p = \{p_i, i \in S\}$ se nazývá počáteční rozdělení Markovského řetězce.

Lze ukázat (Věta o násobení pravděpodobnosti), že

$$P(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = p_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}.$$

Označme dále $p_{ij}^{(1)} = p_{ij}$ a definujme pro přirozené $n \geq 1$ postupně

$$p_{ij}^{(n+1)} = \sum_{k \in S} p_{ik}^{(n)} p_{kj}. \quad (3)$$

Lze ukázat, že $p_{ij}^{(n)} \leq 1$ a navíc pro matice pravděpodobností přechodů platí

$$P^{(2)} = P \cdot P = P^2 \text{ a obecně } P^{(n+1)} = P^{(n)} \cdot P = P \cdot P^{(n)} = P^{n+1}.$$

Věta

Nechť $\{X_n, n \in \mathbb{N}\}$ je homogenní Markovský řetězec s maticí přechodu P . Potom pro pravděpodobnosti přechodu n -tého řádu platí

$$P(X_{m+n} = j | X_m = i) = p_{ij}^{(n)}, \quad \forall i, j \in S$$

pro všechna přirozená m a n a pro $P(X_m = i) > 0$.

Vztah (3) lze zobecnit. Toto zobecnění se nazývá

Chapman-Kolmogorova rovnost

definována jako

$$p_{ij}^{(m+n)} = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)},$$

zapsáno maticově

$$P^{(m+n)} = P^{(m)} \cdot P^{(n)}.$$

- Vychází-li řetězec $\{X_n, n \in \mathbb{N}\}$ ze stavu j , tj. $P(X_0 = j) = 1$, pak označíme

$$P(\cdot | X_0 = j) = P_j(\cdot).$$

- Definujme náhodnou veličinu

$$\tau_j = \inf\{n > 0 : X_n = j\}$$

čas prvního návratu řetězce do stavu j .

- Střední hodnotu doby prvního návratu označíme $\mu_j = \mathbb{E}[\tau_j | X_0 = j]$.
- Největší společný dělitel čísel $n \geq 1$, pro které $p_{jj}^{(n)} > 0$, označíme d_j .

Definice

Stav j Markovského řetězce se nazývá trvalý, jestliže

$$P_j(\tau_j < \infty) = 1.$$

Stav j Markovského řetězce se nazývá přechodný, jestliže

$$P_j(\tau_j = \infty) > 0.$$

Definice

Trvalý stav j Markovského řetězce se nazývá nenulový, jestliže $\mu_j < \infty$ a nulový, jestliže $\mu_j = \infty$.

Definice

Je-li $d_j > 1$, stav j Markovského řetězce se nazývá periodický s periodou d_j , je-li $d_j = 1$, stav j Markovského řetězce se nazývá neperiodický.

Věta

- a) *Nechť j je přechodný stav. Potom $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0, \forall i \in S$.*
- b) *Nechť j je trvalý nulový stav. Potom $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0, \forall i \in S$.*
- c) *Nechť j je trvalý nenulový a neperiodický stav. Potom $\lim_{n \rightarrow \infty} p_{jj}^{(n)} = \frac{1}{\mu_j}$.*
- d) *Nechť j je trvalý nenulový stav s periodou d_j . Potom $\lim_{n \rightarrow \infty} p_{jj}^{(nd_j)} = \frac{d_j}{\mu_j}$.*

Věta

Trvalý stav j je nulový právě tehdy, když $\lim_{n \rightarrow \infty} p_{jj}^{(n)} = 0$.

Uvažujme nyní řetězec s množinou přechodných stavů R a definujme náhodnou veličinu

$$\tau = \inf\{n \geq 0 : X_n \notin R\}$$

značící čas výstupu z množiny přechodných stavů.

Věta

V řetězci s konečně mnoha stavy je

$$P_j(\tau = \infty) = 0, \quad j \in R.$$

Definice

Řekneme, že stav j je dosažitelný ze stavu i , jestliže existuje $n \in \mathbb{N}$ takové, že $p_{ij}^{(n)} > 0$. Jestliže $p_{ij}^{(n)} = 0$, pro všechna $n \in \mathbb{N}$, pak říkáme, že stav j je nedosažitelný ze stavu i .

Definice

Množina stavů C se nazývá uzavřená, jestliže žádný stav vně C není dosažitelný z žádného stavu uvnitř C .

Definice

Množina stavů C se nazývá komponentou, jestliže žádný stav vně C není dosažitelný z žádného stavu uvnitř C a opačně, a přitom všechny stavy uvnitř komponenty jsou vzájemně dosažitelné.

Věta

Množina stavů je uzavřená právě tehdy, je-li $p_{ij} = 0$ pro všechna $i \in C, j \notin C$.

Definice

Markovský řetězec se nazývá nerozložitelný, jestliže každý jeho stav je dosažitelný z každého jiného stavu. V opačném případě je řetězec rozložitelný.

Definice

Je-li jednobodová množina stavů $\{j\}$ uzavřená, tj. je-li $p_{jj} = 1$, pak se stav j nazývá absorpční.

Definice

Řetězec s konečně mnoha stavy, jehož všechny trvalé stavy jsou absorpční, se nazývá absorpční řetězec.

Definice

Nechť $\{X_n, n \in \mathbb{N}\}$ je homogenní řetězec s množinou stavů S a maticí pravděpodobností přechodu P . Nechť $\pi = \{\pi_j, j \in S\}$ je nějaké pravděpodobnostní rozdělení na množině S , tj.

$\pi_j \geq 0, j \in S, \sum_{j \in S} \pi_j = 1$. Potom π se nazývá stacionární rozdělení daného řetězce, jestliže platí

$$\pi^T = \pi^T P,$$

neboli

$$\pi_j = \sum_{k \in S} \pi_k p_{kj}, j \in S.$$

Věta

Nechť počáteční rozdělení homogenního Markovského řetězce je stacionární. Pak je tento řetězec striktně stacionární a pro všechna $n \in \mathbb{N}$ platí

$$p_j(n) = P(X_n = j) = \pi_j, \quad j \in S,$$

kde π_j jsou počáteční stacionární pravděpodobnosti.

Věta

Pro nerozložitelný Markovský řetězec platí

- ❶ *Jsou-li všechny jeho stavy přechodné nebo všechny trvalé nulové, stacionární rozdělení neexistuje.*
- ❷ *Jsou-li všechny jeho stavy trvalé nenulové, stacionární rozdělení existuje a je jednoznačné.*

- ❶ *Jsou-li všechny stavy neperiodické, potom pro všechna $i, j \in S$*

$$\pi_j = \lim_{n \rightarrow \infty} p_{ij}^{(n)} > 0 \quad a \quad \pi_j = \lim_{n \rightarrow \infty} p_j(n) > 0.$$

- ❷ *Jsou-li všechny stavy periodické, potom pro všechna $i, j \in S$*

$$\pi_j = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n p_{ij}^{(k)} > 0 \quad a \quad \pi_j = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n p_j(k) > 0.$$

- ❸ *V nerozložitelném řetězci s konečně mnoha stavy stacionární rozdělení existuje.*

- Uvažujme rozložitelný Markovský řetězec s konečně mnoha stavy $j \in S$, který lze rozdělit na K komponent.
- Existuje permutace těchto stavů $\tilde{j} = \text{perm}(j)$ pro všechna $j \in S$ taková, že matice pravděpodobností přechodu pro stavy \tilde{j} je tvaru

$$P = \begin{pmatrix} P_1 & 0 & \dots & 0 \\ 0 & P_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & P_K \end{pmatrix},$$

kde všechny matice P_k , $k = 1, \dots, K$, jsou stochastické.

- Označme $\pi^{(k)}$ stacionární rozdělení pro řetězec s maticí pravděpodobností přechodu P_k , $k = 1, \dots, K$.
- Pak Markovský řetězec s maticí pravděpodobností přechodu P má stacionární rozdělení

$$\pi = (c_1 \pi^{(1)}, \dots, c_K \pi^{(K)}),$$

kde $0 \leq c_k \leq 1$ pro všechna $k = 1, \dots, K$ a $\sum_{k=1}^K c_k = 1$.

- Uvažujme rozložitelný Markovský řetězec se stavy $j \in S$, z nichž právě m je absorpčních.
- Existuje permutace těchto stavů $\tilde{j} = \text{perm}(j)$ pro všechna $j \in S$ taková, že matice pravděpodobností přechodu pro stavy \tilde{j} je tvaru

$$P = \begin{pmatrix} \mathbb{I}_m & 0 \\ R & Q \end{pmatrix},$$

kde \mathbb{I}_m je jednotková matice typu $m \times m$.

- Pak

$$P^\infty = \lim_{n \rightarrow \infty} P^n = \begin{pmatrix} \mathbb{I}_m & 0 \\ M & 0 \end{pmatrix},$$

kde $M = FR$ a $F = (\mathbb{I}_{|S|-m} + Q + Q^2 + Q^3 + \dots) = (\mathbb{I}_{|S|-m} - Q)^{-1}$ je tzv. fundamentální matice.

- Prvky p_{ij} , $i = m+1, \dots, |S|$, $j = 1, \dots, m$ (tj. prvky matice M) vyjadřují pravděpodobnosti, že řetězec, který vyšel ze stavu i , skončí v absorpčním stavu j .

Definice

Systém celočíselných náhodných veličin $\{X_t, t \geq 0\}$ definovaných na pravděpodobnostním prostoru (Ω, \mathcal{A}, P) nazveme Markovský řetězec se spojitým časem, jestliže

$$P(X_t = j | X_s = i, X_{t_n} = i_n, \dots, X_{t_1} = i_1) = P(X_t = j | X_s = i)$$

pro všechna $0 \leq t_1 < \dots < t_n < s < t$ a všechna $i, j, i_n, \dots, i_1 \in S$ taková, že $P(X_s = i, X_{t_n} = i_n, \dots, X_{t_1} = i_1) > 0$.

- $P(X_t = j | X_s = i) = p_{ij}(s, t)$ nazveme pravděpodobnostmi přechodu ze stavu i v čase s do stavu j v čase t ;
- pro homogenní řetězec, kde pravděpodobnosti přechodu závisí pouze na rozdílech časů, budeme značit $p_{ij}(s, s + t)$ jako $p_{ij}(t)$;
- absolutní pravděpodobnosti budeme značit $p_j(t) = P(X_t = j), j \in S$ a $p_j(0) = P(X_0 = j), j \in S$ pak budou počáteční pravděpodobnosti.

- Pro každé t je $P(t) = \{p_{ij}(t), i, j \in S\}$ čtvercová matice \rightarrow systém matic pravděpodobností přechodu $\{P(t), t \geq 0\}$.
- Zřejmě platí: $\{P(0) = I\}$, kde I je jednotková matice.
- Dále dostáváme

$$p_j(t) = \sum_{i \in S} p_i(0) p_{ij}(t) \quad \forall j \in S,$$

zapsáno maticově

$$p(t)^T = p(0)^T \cdot P(t).$$

Ten lze zobecnit na

$$p_{ij}(s+t) = \sum_{k \in S} p_{ik}(s) p_{kj}(t) \quad \forall i, j \in S,$$

zapsáno maticově

$$P(s+t) = P(s) \cdot P(t),$$

což je **Chapman-Kolmogorova rovnost**.

V dalším textu budeme předpokládat, že

$$\lim_{t \rightarrow 0+} p_{ij}(t) = \delta_{ij}, \quad i, j \in S,$$

kde δ_{ij} značí Dirackovu funkci, tj. $\delta_{ij} = 1$ pro $i = j$ a $\delta_{ij} = 0$ jinak. Tento předpoklad společně se skutečností, že $p_{ij}(0) = \delta_{ij}$ znamená, že řetězec je zprava spojitý v 0.

Dále si označme

$$\lim_{h \rightarrow 0+} \frac{1 - p_{ii}(h)}{h} := \lambda_i \quad \text{a} \quad \lim_{h \rightarrow 0+} \frac{p_{ij}(h)}{h} := \lambda_{ij}. \quad (4)$$

Definice

Nezáporná čísla λ_{ij} z (4) se nazývají intenzity přechodu, číslo λ_i z (4) je pak celková intenzita. Matice $\Lambda = \{\lambda_{ij}, i, j \in S\}$, kde $\lambda_{ii} = -\lambda_i$ se nazývá matice intenzit přechodu.

Věta

Pro homogenní Markovský řetězec platí

$$P(X_t = i \text{ pro } t \in (s, s + h) | X_s = i) = e^{-\lambda_i h} \quad \forall s \geq 0, h \geq 0.$$

Věta

Je-li $\lambda_i = 0$, pak $p_{ii}(t) = 1$ pro všechna $t \geq 0$. Je-li $0 < \lambda_i < \infty$, má doba, po kterou řetězec setrvává ve stavu i , exponenciální rozdělení s parametrem λ_i .

Věta

Nechť $0 < \lambda_i < \infty$. Potom pravděpodobnost, že řetězec z počátečního stavu i přejde nejdříve do stavu j , je rovna $\frac{\lambda_{ij}}{\lambda_i}$ pro všechna $j \neq i$.

Nechť $P(X_0 = j) = 1$, J je čas, kdy řetězec poprvé opustí stav j , a

$$\tau_j = \inf\{t \geq J : X_t = j\}.$$

Definice

Stav j Markovského řetězce se nazývá trvalý, jestliže buď $q_j = 0$ nebo $q_j > 0$ a zároveň

$$P_j(\tau_j < \infty) = 1.$$

Stav j Markovského řetězce se nazývá přechodný, jestliže $q_j > 0$ a zároveň

$$P_j(\tau_j = \infty) > 0.$$

Definice

Trvalý stav j Markovského řetězce se nazývá nenulový, jestliže buď $q_j = 0$ nebo $\mathbb{E}[\tau_j] < \infty$. V opačném případě se řetězec nazývá nulový.

Definice

Stav $j \in S$ se nazývá absorpční, jestliže $q_j = 0$. Jestliže $q_j > 0$, pak se stav j nazývá stabilní, pokud $q_j < \infty$, a nestabilní, pokud $q_j = \infty$.

Definice

Řekneme, že stav j je dosažitelný ze stavu i , jestliže existuje $t > 0$ takové, že $p_{ij}(t) > 0$.

Poznámka

Analogicky jako pro řetězce s diskrétním časem můžeme definovat také nerozložitelnost řetězce se spojitým časem.

Definice

Nechť $\{X_t, t \geq 0\}$ je homogenní řetězec se spojitým časem, množinou stavů S a maticemi pravděpodobností přechodu $P(t), t \geq 0$. Potom π se nazývá stacionární rozdělení daného řetězce, jestliže platí

$$\pi^T = \pi^T P(t), \quad \forall t \geq 0.$$

Věta

Nechť počáteční rozdělení homogenního Markovského řetězce $\{X_t, t \geq 0\}$ je stacionární. Pak je tento řetězec striktně stacionární a pro všechna $t \geq 0$ platí

$$p_j(t) = P(X_t = j) = \pi_j, \quad j \in S,$$

kde π_j jsou počáteční stacionární pravděpodobnosti.

Poissonův proces $\{N_t, t \geq 0\}$ popisuje počet událostí do času t .

Předpoklady:

- počty událostí v disjunktních časových intervalech jsou nezávislé náhodné veličiny (proces s nezávislými přírůstky),
- počty událostí v časovém intervalu $(t, t + h)$ závisí pouze na h ,
- pro počty událostí v časovém intervalu $(t, t + h)$ platí

$$P(N_{t+h} - N_t = 0) = 1 - \lambda h + o(h),$$

$$P(N_{t+h} - N_t = 1) = \lambda h + o(h),$$

$$P(N_{t+h} - N_t \geq 2) = o(h),$$

kde symbol $o(h)$ značí, že $o(h)/h \rightarrow 0$ při $h \rightarrow 0+$, a λ je konstanta, která se nazývá intenzita Poissonova procesu.

- Z předpokladu nezávislosti přírůstků plyne Markovská vlastnost
- Pro pravděpodobnosti přechodu platí

$$\begin{aligned}
 P(N_{t+h} = j | N_t = i) &= \lambda h + o(h) & j = i + 1 \\
 &= 1 - \lambda h + o(h) & j = i \\
 &= o(h) & j > i + 1 \\
 &= 0 & j < i.
 \end{aligned}$$

- Intenzity přechodu jsou

$$q_{i,i+1} = \lambda, \quad q_i = -q_{ii} = \lambda, \quad q_{ij} = 0 \text{ jinak.}$$

- Navíc se dá ukázat, že pro tento proces platí

$$P(N_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, \dots,$$

což je Poissonovo rozdělení.

- Není to Markovský řetězec, pouze Markovský proces (spojitá množina stavů)!
- Wienerův proces (někdy také nazýván *Brownův pohyb*) $\{W_t, t \geq 0\}$ je definován následujícími vlastnostmi:
 - $\{W_t, t \geq 0\}$ má spojité trajektorie,
 - $W_0 = 0$,
 - $\{W_t, t \geq 0\}$ má nezávislé přírůstky,
 - přírůstek hodnoty v časovém intervalu (s, t) má normální rozdělení s nulovou střední hodnotou a rozptylem $(t - s)$.

Poznámka

Občas se místo $(t - s)$ jako rozptyl uvádí $\sigma^2(t - s)$, kde σ^2 je kladná konstanta.