NSTP 097 Statistika Zima 2009[1]

Přehled pravděpodobnostních rozdělení

1 Diskrétní rozdělení

1.1 Alternativní (Bernoulliovo, nula-jedničkové) rozdělení

Značení: $X \sim \mathrm{Alt}(p)$

Parametry: $p \in (0,1)$

Nosič: $X \in \{0,1\}$

Hustota: $\mathrm{P}\,[X=j] = p^j(1-p)^{1-j}, \quad j \in \{0,1\}$

Střední hodnota: $\mathrm{E}\,X = p$

Rozptyl: $\mathrm{var}\,X = p(1-p)$

Poznámky: Rozdělení indikátoru náhodného jevu, úspěch vs. neúspěch.

1.2 Binomické rozdělení

Značení: $X \sim \mathrm{Bi}(n,p)$

Parametry: $p \in (0,1), n \in \mathbb{N}$

Nosič: $X \in \{0,1,\ldots,n\}$

Hustota: $\mathrm{P}\,[X=j] = \binom{n}{j} p^j(1-p)^{n-j}, \quad j \in \{0,1,\ldots,n\}$

Střední hodnota: $\mathrm{E}\,X = np$

Rozptyl: $\mathrm{var}\,X = np(1-p)$

Poznámky:

- Rozdělení součtu nezávislých alternativních veličin (počtu úspěchů mezi n pokusy). Přesněji, jsou-li $X_i \sim \mathrm{Alt}(p)$ nezávislé, $i=1,\ldots,n$, pak $\sum_{i=1}^n X_i \sim \mathrm{Bi}(n,p)$.
- $\mathrm{Bi}(1,p)$ jest $\mathrm{Alt}(p)$.
- Jestliže $n \to \infty$ a $np \to \lambda < \infty$, pak $\mathrm{Bi}(n,p)$ konverguje k $\mathrm{Po}(\lambda)$.

1.3 Geometrické rozdělení

Značení: $X \sim \mathrm{Geo}(p)$

Parametry: $p \in (0,1)$

[1] Michal Kulich, KPMS MFF UK, 29.9.2009

Nosič: $X \in \{0, 1, 2, \ldots\}$

Hustota: $\mathrm{P}[X = j] = p(1-p)^j, \quad j = 0, 1, 2, \ldots$

Střední hodnota: $\mathrm{E}\,X = \dfrac{1-p}{p}$

Rozptyl: $\operatorname{var} X = \dfrac{1-p}{p^2}$

Poznámky: Počet neúspěchů před prvním úspěchem v posloupnosti nezávislých pokusů.

1.4 Poissonovo rozdělení

Značení: $X \sim \mathrm{Po}(\lambda)$

Parametry: $\lambda > 0$

Nosič: $X \in \{0, 1, 2, \ldots\}$

Hustota: $\mathrm{P}[X = j] = \dfrac{\lambda^j}{j!}\mathrm{e}^{-\lambda}, \quad j = 0, 1, 2, \ldots$

Střední hodnota: $\mathrm{E}\,X = \lambda$

Rozptyl: $\operatorname{var} X = \lambda$

1.5 Negativně binomické rozdělení

Značení: $X \sim \mathrm{NB}(n, p)$

Parametry: $p \in (0, 1), n \in \mathbb{N}$

Nosič: $X \in \{0, 1, 2, \ldots\}$

Hustota: $\mathrm{P}[X = j] = \dbinom{n+j-1}{n-1} p^n (1-p)^j, \quad j = 0, 1, 2, \ldots$

Střední hodnota: $\mathrm{E}\,X = n\dfrac{1-p}{p}$

Rozptyl: $\operatorname{var} X = n\dfrac{1-p}{p^2}$

Poznámky: Rozdělení počtu neúspěchů předcházejících n-tému úspěchu v posloupnosti nezávislých pokusů. $\mathrm{NB}(1, p)$ jest $\mathrm{Geo}(p)$.

2 Spojitá rozdělení

2.1 Rovnoměrné rozdělení

Značení: $X \sim \mathrm{R}(a, b)$

Parametry: $a, b \in \mathbb{R}, a < b$

Nosič: $X \in (a, b)$

Hustota: $f(x) = \frac{1}{b-a}\mathbb{I}_{(a,b)}(x), \quad x \in \mathbb{R}$

Distribuční funkce:

$$F(x) = \begin{cases} 0 & x < a, \\ \frac{x-a}{b-a} & a < x < b, \\ 1 & x > b \end{cases}$$

Střední hodnota: $\mathrm{E}\, X = \frac{a+b}{2}$

Rozptyl: $\operatorname{var} X = \frac{(b-a)^2}{12}$

Poznámky: Rozdělení s konstantní hustotou.

2.2 Normální rozdělení

Značení: $X \sim \mathrm{N}(\mu, \sigma^2)$

Parametry: $\mu \in \mathbb{R}, \sigma^2 > 0$

Nosič: $X \in \mathbb{R}$

Hustota:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-(x-\mu)^2/(2\sigma^2)}, \quad x \in \mathbb{R}$$

Distribuční funkce: $F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$, kde

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{e}^{-u^2/2} du$$

Střední hodnota: $\mathrm{E}\, X = \mu$

Rozptyl: $\operatorname{var} X = \sigma^2$

Vyšší momenty: $\mu_k = \begin{cases} 0 & k \text{ liché} \\ \{(k-1)(k-3)\cdots 3 \cdot 1\}\sigma^k & k \text{ sudé} \end{cases}$

Špičatost: $\gamma_4 = 3$

2.3 Cauchyovo rozdělení

Značení: $X \sim \mathrm{C}(a, b)$

Parametry: $a \in \mathbb{R}$, $b > 0$

Nosič: $X \in \mathbb{R}$

Hustota: $f(x) = \dfrac{1}{b\pi}\left[1 + \left(\dfrac{x-a}{b}\right)^2\right]^{-1}$

Distribuční funkce: $F(x) = \dfrac{1}{2} + \dfrac{1}{\pi}\operatorname{arctg}\dfrac{x-a}{b}$

Střední hodnota: $\mathrm{E}\,X$ neexistuje

Rozptyl: $\operatorname{var} X$ neexistuje

Poznámky: Hustota je symetrická kolem a, momenty neexistují, $\mathrm{E}\,|X| = +\infty$.

2.4 Exponenciální rozdělení

Značení: $X \sim \mathrm{Exp}(\lambda)$

Parametry: $\lambda > 0$

Nosič: $X \in (0, \infty)$

Hustota: $f(x) = \lambda \mathrm{e}^{-\lambda x}\mathbb{I}_{(0,\infty)}(x)$

Distribuční funkce:

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - \mathrm{e}^{-\lambda x} & x > 0 \end{cases}$$

Střední hodnota: $\mathrm{E}\,X = \dfrac{1}{\lambda}$

Rozptyl: $\operatorname{var} X = \dfrac{1}{\lambda^2}$

2.5 Gama rozdělení

Značení: $X \sim \Gamma(a, p)$

Parametry: $a > 0$, $p > 0$

Nosič: $X \in (0, \infty)$

Hustota: $f(x) = \dfrac{a^p}{\Gamma(p)} x^{p-1}\mathrm{e}^{-ax}\mathbb{I}_{(0,\infty)}(x), \quad x \in \mathbb{R}$

Střední hodnota: $\mathrm{E}\,X = \frac{p}{a}$

Rozptyl: $\operatorname{var} X = \frac{p}{a^2}$

Poznámky:

- $\Gamma(a, 1)$ jest $\mathrm{Exp}(a)$
- $\Gamma(a, k)$, $k \in \mathbb{N}$ jest součtem k nezávislých veličin s rozdělením $\mathrm{Exp}(a)$; toto rozdělení se též nazývá Erlangovo.
- $\Gamma\left(\frac{1}{2}, \frac{r}{2}\right)$ jest χ^2_r-rozdělení.

2.6 Beta rozdělení

Značení: $X \sim \mathrm{B}(\alpha, \beta)$

Parametry: $\alpha, \beta > 0$

Nosič: $X \in (0, 1)$

Hustota: $f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}\mathbb{I}_{(0,1)}(x)$

Střední hodnota: $\mathrm{E}\,X = \frac{\alpha}{\alpha + \beta}$

Rozptyl: $\operatorname{var} X = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

Poznámky: Systém konstantních, lineárních a kvadratických hustot na omezeném intervalu $(0, 1)$. $\mathrm{B}(1, 1)$ jest $\mathrm{R}(0, 1)$.

2.7 χ^2 rozdělení

Značení: $X \sim \chi^2_r$

Parametry: $r \in \mathbb{N}$, stupně volnosti

Nosič: $X \in (0, \infty)$

Hustota: $f(x) = \frac{1}{2^{r/2}\Gamma(r/2)} x^{(r-2)/2} \mathrm{e}^{-x/2} \mathbb{I}_{(0,\infty)}(x)$

Střední hodnota: $\mathrm{E}\,X = r$

Rozptyl: $\operatorname{var} X = 2r$

Poznámky:

- Speciální případ Γ-rozdělení s parametry $\frac{1}{2}$, $\frac{r}{2}$.
- Rozdělení součtu kvadrátů r nezávislých normovaných normálních veličin.

2.8 Studentovo t-rozdělení

Značení: $X \sim \mathrm{t}_k$

Parametry: $k \in \mathbb{N}$, počet stupňů volnosti

Nosič: $X \in \mathbb{R}$

Hustota: $f(x) = \dfrac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{k\pi}} \left(1 + \dfrac{x^2}{k}\right)^{-(k+1)/2}$

Střední hodnota: $\mathrm{E}\, X = 0$ pro $k \geq 2$, neexistuje pro $k = 1$.

Rozptyl: $\operatorname{var} X = \dfrac{k}{k-2}$ pro $k \geq 3$, $\operatorname{var} X = \infty$ pro $k = 1, 2$.

Poznámky:

- t_1 jest $\mathrm{C}(0,1)$
- Jsou-li $X \sim \mathrm{N}(0,1)$ a $Z \sim \chi^2_k$ nezávislé, pak

$$\frac{X}{\sqrt{Z/k}} \sim \mathrm{t}_k$$

2.9 (Fisherovo) F-rozdělení

Značení: $X \sim \mathrm{F}_{m,n}$

Parametry: $m, n \in \mathbb{N}$, stupně volnosti

Nosič: $X \in (0, \infty)$

Hustota: $f(x) = \dfrac{1}{B\left(\frac{m}{2}, \frac{n}{2}\right)} \left(\dfrac{m}{n}\right)^{m/2} x^{m/2-1} \left(1 + \dfrac{m}{n}x\right)^{-(m+n)/2} \mathbb{I}_{(0,\infty)}(x)$

Střední hodnota: $\mathrm{E}\, X = \dfrac{n}{n-2}$ pro $n \geq 3$, neexistuje pro $n = 1, 2$.

Rozptyl: $\operatorname{var} X = \dfrac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$ pro $n \geq 5$, $\operatorname{var} X = \infty$ pro $n \leq 4$.

Poznámky:

- Jsou-li $Z_1 \sim \chi^2_m$ a $Z_2 \sim \chi^2_n$ nezávislé, pak $\dfrac{Z_1/m}{Z_2/n} \sim \mathrm{F}_{m,n}$
- Je-li $Y \sim B(m/2, n/2)$, pak $\dfrac{n}{m}\dfrac{Y}{1-Y} \sim \mathrm{F}_{m,n}$

3 Mnohorozměrná diskrétní rozdělení

3.1 Multinomické rozdělení

Značení: $\boldsymbol{X} \sim \mathrm{Mult}_k(n; \boldsymbol{p})$

Parametry: $k \geq 2$ počet přihrádek, $n \in \mathbb{N}$ počet pokusů, $\boldsymbol{p} \in \{(0,1)^k : \sum_{j=1}^k p_j = 1\}$ pravděpodobnosti přihrádek.

Nosič: $\boldsymbol{X} \in \{\{0,1,\ldots,n\}^k : \sum_{j=1}^k X_i = n\}$

Hustota: $\mathrm{P}[X_1 = r_1, \ldots, X_k = r_k] = \dfrac{n!}{r_1! \cdots r_k!} p_1^{r_1} \cdots p_k^{r_k} \mathbb{I}_A(\boldsymbol{r})$, kde $A = \{\boldsymbol{r} \in \{0,\ldots,n\}^k : \sum_{j=1}^k r_j = n\}$

Střední hodnota: $\mathrm{E}\,\boldsymbol{X} = n\boldsymbol{p}$

Rozptyl: $\mathrm{var}\,\boldsymbol{X} = n[\mathrm{diag}\,(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}^\mathsf{T}]$

Poznámky:

- Rozdělení počtu koulí padlých do každé z k přihrádek při n nezávislých pokusech.
- Marginální rozdělení X_j jest $\mathrm{Bi}(n, p_j)$, $j = 1, \ldots, k$

4 Mnohorozměrná spojitá rozdělení

4.1 Mnohorozměrné normální rozdělení

Značení: $\boldsymbol{X} \sim \mathrm{N}_k(\boldsymbol{\mu}, \Sigma)$

Parametry: $\boldsymbol{\mu} \in \mathbb{R}^k$ střední hodnota, $\Sigma > 0$ positivně definitní rozptylová matice

Nosič: $\boldsymbol{X} \in \mathbb{R}^k$

Hustota: $$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{k/2}\sqrt{\det \Sigma}} \exp\left\{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^\mathsf{T} \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu})\right\},\ \boldsymbol{x} \in \mathbb{R}^k$$

Střední hodnota: $\mathrm{E}\,\boldsymbol{X} = \boldsymbol{\mu}$

Rozptyl: $\mathrm{var}\,\boldsymbol{X} = \Sigma$

Poznámky: Je-li $k = 2$, můžeme vyjádřit $\Sigma = \begin{pmatrix} \sigma_1^2 & \varrho\sigma_1\sigma_2 \\ \varrho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$, kde $\varrho = \mathrm{cor}\,(X_1, X_2)$. Pak

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{1-\varrho^2}\sigma_1\sigma_2} \exp\left\{ -\frac{1}{2(1-\varrho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\varrho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

5 Appendix

5.1 Gama funkce

Definice: $\Gamma(p) = \int_0^\infty x^{p-1} \mathrm{e}^{-x}\, dx, \quad p > 0$ se nazývá **gama funkce**.

Vlastnosti:
- $\Gamma(k) = (k-1)!, \quad k = 1, 2, 3, \dots$
- $\Gamma(1/2) = \sqrt{\pi}$
- $\Gamma(p+1) = p\Gamma(p),\ p > 0$

5.2 Beta funkce

Definice: $B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\, dx, \quad \alpha, \beta > 0$ se nazývá **beta funkce**.

Vlastnosti:
- $B(\alpha, \beta) = B(\beta, \alpha) = \dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$

5.3 Normované normální rozdělení

Distribuční funkce: $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-u^2/2} du$

Vlastnosti: $\Phi(-x) = 1 - \Phi(x)$

Hodnoty:

x	0	0.5	1	1.5	2	3
$\Phi(x)$	0.5	0.6915	0.8413	0.9332	0.9772	0.9987

Kvantily:

α	0.5	0.8	0.9	0.95	0.975	0.99	0.995
$\Phi^{-1}(\alpha)$	0	0.8416	1.282	1.645	1.96	2.326	2.576

5.4 Kvantily χ^2-rozdělení

Definice: Pro dané $\alpha \in (0,1)$ a $r \geq 1$ hledáme x takové, že

$$\frac{1}{2^{r/2}\Gamma(r/2)} \int_0^x u^{(r-2)/2} \mathrm{e}^{-u/2} du = \alpha$$

Tabulka:

	α			
r	0.9	0.95	0.99	0.999
1	2.706	3.841	6.635	10.828
2	4.605	5.991	9.210	13.816
3	6.251	7.815	11.345	16.266
4	7.779	9.488	13.277	18.467
5	9.236	11.070	15.086	20.515
6	10.645	12.592	16.812	22.458
7	12.017	14.067	18.475	24.322
8	13.362	15.507	20.090	26.124
9	14.684	16.919	21.666	27.877
10	15.987	18.307	23.209	29.588
15	22.307	24.996	30.578	37.697
20	28.412	31.410	37.566	45.315
30	40.256	43.773	50.892	59.703
40	51.805	55.758	63.691	73.402
50	63.167	67.505	76.154	86.661

Obsah