Přehled PST

Jaroslav Šmolík

15. ledna 2016

Abstrakt

Stručný přehled vzorců a vztahů z předmětu PST. Neobsahuje všechny *triviální* věci (pravděpodobnost jevu) a některé *komplikovanější*, které se nedají snadno stručně shrnout (CLT, Markova nerovnost etc.). Veškerou většinu ostatní potřebné výbavy na zkoušku ale ano.

1 Pravděpodobnost

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

2 Náhodné veličiny

Střední hodnota

$$\mathrm{E}\, X = \sum_{\forall x_i} x_i p_X(x_i)$$

$$\mathrm{E}\, X = \int_{-\infty}^{\infty} x f_X(x)$$

Střední hodnota, rozptyl

$$\sigma^2 = \operatorname{var} X = \mathrm{E}(X - \mathrm{E}\, X)^2 = \mathrm{E}\, X^2 - (\mathrm{E}\, X)^2$$

$$\sigma = \operatorname{sd} X = \sqrt{\operatorname{var} X}$$

Moment

$$\mu_k = \mathrm{E}\, X^k$$

Centrální moment

$$\sigma_k = \mathrm{E}(X - \mathrm{E}\, X)^k$$

Koeficient šimosti

$$\gamma_1 = \frac{\sigma^3}{\sigma_3}$$

Koeficient je kladný, záporný, podle toho, kam se vrchol vychyluje. (kladný doleva, záporný doprava)

Koeficient špičatosti

$$\gamma_2 = \frac{\sigma^4}{\sigma_4} - 3$$

Koeficient je kladný, záporný, podle toho, jestli je špičatější než normální rozdělení (kladný špičatější, záporný méně špičatý).

Kvantil

$$x_\alpha = Q_X(\alpha) = \inf\{x : F_X \geq \alpha\}$$

Tedy bod, který dělí veličinu tak, že před ním je $(\alpha \cdot 100)\%$ hodnot.

Kritická hodnota

$$z_\alpha = Q_X(1-\alpha)$$

Moment generující funkce

$$M_X(s) = \mathrm{E}(e^{sX})$$

k-tý moment je pak

$$m_k = \frac{d^k}{ds^k} M_X(s)$$

3 Náhodné vektory

Marginální hustota

$$p_X(x) = \sum_{\forall y_i} p_{X,Y}(x, y_i)$$

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

Podmíněná hustota

$$p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

Kovariance

$$\mathrm{cov}_{X,Y} = \mathrm{E}(X - \mathrm{E}\,X)(Y - \mathrm{E}\,Y)$$

Korelační koeficient

$$\rho_{X,Y} = \frac{\mathrm{cov}_{X,Y}}{\sqrt{\mathrm{var}\,X}\sqrt{\mathrm{var}\,Y}}$$

4 Příklady rozdělení

4.1 Příklady diskrétních rozdělení

Tabulka 1: Příklady diskrétních rozdělení

	Be(p)	Binom(p,n)	Geom(p)	Poisson(λ)
p_X	p	$\binom{n}{k}p^k(1-p)^{n-k}$	$(1-p)^{k-1}p$	$\frac{\lambda^k}{k!}e^{-\lambda}$
$\mathrm{E}\,X$	p	$n \cdot p$	$\frac{1}{p}$	λ
$\mathrm{var}\,X$	$pq = p(1-p)$	$n \cdot pq = n \cdot p(1-p)$	$\frac{q}{p^2} = \frac{1-p}{p^2}$	λ

4.2 Příklady spojitých rozdělení

Tabulka 2: Příklady spojitých rozdělení

	Unif(a,b)	Exp(λ)	Norm(μ, σ^2)
p_X	$\begin{cases} \frac{1}{b-a}, & x \in (a,b) \\ 0, & \text{else} \end{cases}$	$\begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{else} \end{cases}$	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
$\mathrm{E}\,X$	$\frac{a+b}{2}$	$\frac{1}{\lambda}$	μ
$\mathrm{var}\,X$	$\frac{(b-a)^2}{12}$	$\frac{1}{\lambda^2}$	σ^2

5 Bodové odhady

Výběrový průměr

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

Výběrový rozptyl

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{X})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{X}^2\right)$$

Výběrová odchylka

$$S_n = \sqrt{S_n^2}$$

Výběrový moment

$$m_k = \frac{i=1}{n}\sum_{1}^{n} x_i^k$$

Výběrová kovarinace

$$s_{X,Y} = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y_n})(x_i - \bar{x_n}) = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i y_i - n\bar{x}_i\bar{y}_i\right)$$

Výběrový korelační koeficient

$$r = \frac{s_{X,Y}}{s_X s_Y}$$

6 Konfidenční $(1-\alpha)100\%$ intervaly

6.1 Střední hodnota

Známý rozptyl

$$\left(\bar{X}_n - z_{\frac{\alpha}{2},n-1}\frac{\sqrt{\sigma^2}}{\sqrt{n}}, \bar{X}_n + z_{\frac{\alpha}{2},n-1}\frac{\sqrt{\sigma^2}}{\sqrt{n}}\right)$$

$$\left(\bar{X}_n - z_{\alpha,n-1}\frac{\sqrt{\sigma^2}}{\sqrt{n}}, \infty\right), \left(-\infty, \bar{X}_n + z_{\alpha,n-1}\frac{\sqrt{\sigma^2}}{\sqrt{n}}\right)$$

Nezmámý rozptyl

$$\left(\bar{X}_n - t_{\frac{\alpha}{2},n-1}\frac{\sqrt{s_n^2}}{\sqrt{n}}, \bar{X}_n + t_{\frac{\alpha}{2},n-1}\frac{\sqrt{s_n^2}}{\sqrt{n}}\right)$$

$$\left(\bar{X}_n - t_{\alpha,n-1}\frac{\sqrt{s_n^2}}{\sqrt{n}}, \infty\right), \left(-\infty, \bar{X}_n + t_{\alpha,n-1}\frac{\sqrt{s_n^2}}{\sqrt{n}}\right)$$

6.2 Rozptyl

Pouze normální rozdělení. Pro obecná nelze určit KI rozptylu.

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

$$\left(\frac{(n-1)s_n^2}{\chi^2_{\alpha,n-1}}, \infty\right), \left(0, \frac{(n-1)s_n^2}{\chi^2_{1-\alpha,n-1}}\right)$$

7 Lineární regrese

Přímka

$$b_0 = \bar{Y_n} - b_1\bar{x_n}$$
$$b_1 = \frac{s_{X,Y}}{s_x^2} = r\frac{s_Y}{s_X}$$

RSS (residual sum of squares)

$$S_e = \sum_{i=1}^{n} (y_i - (b_0 + b_1 x_i))^2$$

Přesnost

$$R^2 = 1 - \frac{S_e}{(n-1)s_Y^2}$$

Čím více je koeficient blíže 1, tím přesněji odpovídá model datům.