

Matematika III – 10. týden
Číselné charakteristiky, momentová funkce a
centrální limitní věta

Jan Slovák

Masarykova univerzita
Fakulta informatiky

28. 11 – 2. 12. 2016

Obsah přednášky

- 1 Literatura
- 2 Střední hodnota a rozptyl
- 3 Kovariance
- 4 Momentová funkce
- 5 Centrální limitní věta

Kde je dobré číst?

- Karel Zvára, Josef Štěpán, Pravděpodobnost a matematická pravděpodobnost statistika, Matfyzpress, 2006, 230pp.
- J. Slovák, M. Panák, M. Bulant, Matematika drsně a svižně, Muni Press, Brno 2013, v+773 s., elektronická edice www.math.muni.cz/Matematika_drsne_svizne
- Marie Budíková, Štěpán Mikoláš, Pavel Osecký, Teorie pravděpodobnosti a matematická statistika (sbírka příkladů), Masarykova univerzita, 3. vydání, 2004, 117 stran, ISBN 80-210-3313-4.
- Marie Budíková, Tomáš Lerch, Štěpán Mikoláš, Základní statistické metody, Masarykova univerzita, 2005, 170 stran, ISBN 80-210-3886-1.
- Riley, K.F., Hobson, M.P., Bence, S.J. Mathematical Methods for Physics and Engineering, second edition, Cambridge University Press, Cambridge 2004, ISBN 0 521 89067 5, xxiii + 1232 pp.

Střední hodnota

Nechť X je náhodná veličina s diskretním rozdělením. Jestliže řada $\sum_{k=1}^{\infty} x_i P(X = x_i)$ konverguje absolutně (zejména tedy pro všechny X s konečně mnoha možnými hodnotami x_i), pak její součet EX nazýváme **střední hodnotou** X .

Je-li X náhodná veličina se spojitým rozdělením s hustotou $f(x)$ a nevlastní integrál $\int_{-\infty}^{\infty} xf(x)dx$ konverguje absolutně, pak jeho hodnota EX se nazývá **střední hodnota** X .

Vlastnosti střední hodnoty

Theorem

Uvažme náhodné veličiny X , Y , skaláry a , $b \in \mathbb{R}$, náhodný vektor $W = (X_1, \dots, X_n)$, vektor c a čtvercovou skalární matici B s n řádky.

- *Pro konstantní náhodnou veličinu $X = a \in \mathbb{R}$ je $E a = a$.*
- *$E(a + bX) = a + b E X$.*
- *$E(X + Y) = E X + E Y$.*
- *$E(c + BW) = c + B(E W)$.*

Theorem

Jsou-li veličiny X a Y nezávislé, pak $E(XY) = E X E Y$.

Rozptyl

Další charakteristika popisuje, jak moc se dá čekat, že se hodnoty náhodné veličiny „hemží“ kolem nějaké hodnoty.

Definition

Nechť X je náhodná veličina s konečnou střední hodnotou. Pak definujeme **rozptyl** veličiny X výrazem

$$\text{var } X = E(X - E X)^2,$$

pokud taková konečná hodnota existuje.

Odmocnina z rozptylu $\sqrt{\text{var } X}$ se nazývá **směrodatná odchylna** náhodné veličiny X .

Jde o zjevnou obdobu definice kvadrátu vzdálenosti vektorů nebo funkcí. Zachycujeme tak „očekávanou vzdálenost“ hodnot X od její střední hodnoty.

Theorem

Jestliže má náhodná veličina X konečný rozptyl, pro libovolné skaláry $a, b \in \mathbb{R}$ platí

- $\text{var } X = E X^2 - (E X)^2$
- $\text{var}(a + bX) = b^2 \text{var } X$
- $\sqrt{\text{var}(a + bX)} = |b| \sqrt{\text{var } X}$.

Občas přiřazujeme k X **normovanou** veličinu Z ,

$$Z = \frac{X - E X}{\sqrt{\text{var } X}},$$

kteřá má zjevně nulovou střední hodnotu a jednotkový rozptyl.

kvantilová funkce

Je-li $F(x)$ distribuční funkce náhodné veličiny X , pak

$$F^{-1}(u) = \inf\{x \in \mathbb{R}; F(x) \geq u\}, \quad 0 < u < 1$$

je kvantilová funkce náhodné veličiny X .

Hodnota $F^{-1}(\alpha)$ se nazývá α -kvantil.

Tzv. kritické hodnoty pro veličinu X jsou pak $F^{-1}(1 - \alpha)$.

Čebyševova nerovnost

Theorem

Má-li X rozptyl a $\epsilon > 0$ je libovolné, pak platí

$$P(|X - EX| \geq \epsilon) \leq \frac{\text{var } X}{\epsilon^2}.$$

Kovariance veličin

Jsou-li X a Y dvě náhodné veličiny, pro které existují jejich konečné rozptyly, pak definujeme jejich **kovarianci** vztahem

$$\text{cov}(X, Y) = E(X - E X)(Y - E Y).$$

Evidentně je $\text{cov}(X, X) = \text{var } X$ a $\text{cov}(X, Y) = \text{cov}(Y, X)$.

Theorem

Nechť existují konečné rozptyly veličin X a Y . Pak

- $\text{cov}(X, Y) = E(XY) - (E X)(E Y)$
- *pro jakékoliv skaláry a, b, c, d platí*
 $\text{cov}(a + bX, c + dY) = bd \text{cov}(X, Y)$
- $\text{var}(X + Y) = \text{var } X + \text{var } Y + 2 \text{cov}(X, Y)$.

Od kovariance snadno odvodíme tzv. **korelační koeficient** dvou náhodných veličin X a Y . Definujeme jej jako kovarianci příslušných normovaných veličin:

$$\rho_{X,Y} = \text{cov} \left(\frac{X - EX}{\sqrt{\text{var } X}}, \frac{Y - EY}{\sqrt{\text{var } Y}} \right) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var } X \text{ var } Y}}.$$

Theorem

- $\rho_{a+bX, c+dY} = \text{sign}(bd)\rho_{X,Y}$, pro $bd \neq 0$
- $\rho_{X,X} = 1$
- $\rho_{X,Y} = 0$, pokud jsou veličiny X a Y nezávislé.
- pokud je $\rho_{X,Y}$ definován, pak je roven jedné právě, když existují konstanty a, b, c tak, že $P(aX + bY = c) = 1$.

Varianční matice

Uvažme náhodný vektor $W = (X_1, \dots, X_n)$ takový, že pro všechny jeho komponenty existuje rozptyl. Pak **varianční matice** $\text{var } W$ je dána

$$\text{var } W = \begin{pmatrix} \text{var } X_1 & \text{cov}(X_1, X_2) & \dots & \text{cov}(X_1, X_n) \\ \text{cov}(X_2, X_1) & \text{var } X_2 & \dots & \text{cov}(X_2, X_n) \\ & & \dots & \\ \text{cov}(X_n, X_1) & \text{cov}(X_n, X_2) & \dots & \text{var } X_n \end{pmatrix}.$$

Theorem

Pro náhodný vektor X , skaláry a , matice skalárů B platí

$$\text{var}(a + BX) = B \text{var } XB^T.$$

Momenty

Podobně jako rozptyl můžeme uvažovat výrazy vyšších řádů:

$$\mu'_k = E X^k, \quad \mu_k = E(X - E X)^k.$$

Nazýváme je k -tý moment a k -tý centrální moment náhodné veličiny X . Momenty lze všechny dostat jako koeficienty v mocninné řadě následujícím způsobem.

Pro volný reálný parametr t definujeme **momentovou vytvořující funkci** pro náhodnou veličinu X vztahem

$$M_X(t) = E e^{tX}.$$

Tato funkce (za docela rozumných předpokladů následující věty) zcela určuje náhodné veličiny a má řadu užitečných vlastností (tj. *stejná momentová funkce na nějakém netriviálním intervalu \implies stejná distribuční funkce*).

Theorem

Nechť X je náhodná veličina pro kterou na intervalu $(-a, a)$ existuje její analytická momentová vytvořující funkce. Pak na tomto intervalu je $M_X(t)$ dána absolutně konvergující řadou

$$M_t(X) = \sum_{k=0}^{\infty} \frac{t^k}{k!} E X^k.$$

Theorem

Pro součet náhodných veličin platí:

$$M_{X+Y}(t) = M_X(t)M_Y(t).$$

Momentová vytvořující funkce pro $X \sim \text{Bi}(0, 1)$

Často je jednodušší počítat momenty z jejich vytvořující funkce než přímo.

Pro alternativní rozdělení náhodné veličiny $Y \sim A(p)$ spočteme snadno

$$M_Y(t) = E e^{tY} = e^0(1 - p) + e^t p = p(e^t - 1) + 1.$$

Protože je binomické rozdělení $X \sim \text{Bi}(n, p)$ dáno jako součet n alternativních rozdělení $Y_i \sim A(p)$, je zjevně v tomto případě

$$M(t) = M_X(t) = (p(e^t - 1) + 1)^n.$$

Obecně platí $\mu'_k = \frac{d^k}{dt^k} M_X(t)|_{t=0}$. Je tedy např. první moment binomického rozdělení skutečně np (první derivace $M(t)$ v nule), což je střední hodnota. Druhý moment je $np(1 - p)$, čímž jsme ověřili výsledek pro rozptyl.

Momentová vytvořující funkce pro $Z \sim N(0, 1)$

$$\begin{aligned}
 M_Z(t) &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}} \exp(-x^2/2) dx \\
 &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2 - 2tx + t^2 - t^2}{2}\right) dx \\
 &= \exp(t^2/2) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-t)^2}{2}\right) dx \\
 &= \exp(t^2/2).
 \end{aligned}$$

(V předposledním řádku je integrálem dána pravděpodobnost jakékoliv hodnoty pro normální rozdělení, proto je to jednička.)

Derivováním: $(M_Z)'(0) = 0$ a $(M_Z)''(0) = (te^{t^2/2})'(0) = 1$. Je tedy skutečně

$$E Z = 0, \quad \text{var } Z = 1.$$

Uvažme nezávislé náhodné veličiny Y_1, Y_2, \dots , které mají všechny stejné rozdělení se střední hodnotou 0 a rozptylem 1.

Předpokládejme, že třetí absolutní moment $E|Y_i|^3$ je konečný.

Pro náhodnou veličinu $S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i$ spočtěme momentovou funkci (koeficient $n^{-1/2}$ je volen tak, aby rozptyl S_n byl stále 1)

$$M_{S_n} = \prod_{i=1}^n E e^{(t/\sqrt{n})Y_i} = (M_Y(t/\sqrt{n}))^n,$$

kde M_Y je společná momentová funkce všech veličin Y_i . Nyní

$$M_Y(t/\sqrt{n}) = 1 + 0 \frac{t}{\sqrt{n}} + 1 \frac{t^2}{2n} + o(t^2/n)$$

a v limitě proto dostáváme

$$\lim_{n \rightarrow \infty} M_{S_n}(t) = \lim_{n \rightarrow \infty} \left(1 + \frac{t^2}{2n} + o(1/n) \right)^n = e^{t^2/2}.$$

To je právě momentová funkce pro rozdělení $N(0, 1)$!

Tím jsme skoro dokázali:

Theorem (Centrální limitní věta)

Nechť Y_1, Y_2, \dots jsou nezávislé náhodné veličiny se společnou střední hodnotou $E Y_i = \mu$, rozptylem $\text{var } Y_i = \sigma^2 > 0$ a konečným třetím absolutním momentem $E|Y_i|^3$. Pro distribuční funkce náhodných veličin

$$S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{1}{\sigma} (Y_i - \mu)$$

platí

$$\lim_{n \rightarrow \infty} P(S_n < x) = \Phi(x),$$

kde $\Phi(x)$ je distribuční funkce normálního rozdělení $N(0, 1)$.

Všimněme si: součty $X_n = \sum_{i=1}^n Y_i$ mají střední hodnotu $n\mu$ a rozptyl $n\sigma^2$. Veličiny S_n jsou tedy právě normované veličiny X_n .

Pokud jsou $Y_i \sim A(p)$ nezávislé, pak $E(Y_i)^3 = p < \infty$ a všechny podmínky centrální limitní věty jsou splněny, $\mu = p$, $\sigma^2 = p(1 - p)$. Součtové veličiny $X_n = \sum_{i=1}^n Y_i$ pak představují právě binomická rozdělení $\text{Bi}(n, p)$ a příslušné normované veličiny jsou

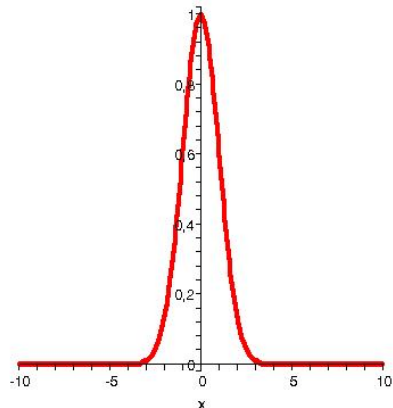
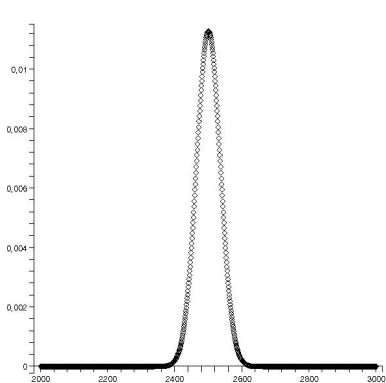
$$S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\frac{Y_i - p}{\sqrt{p(1-p)}} \right) = \frac{X_n - np}{\sqrt{np(1-p)}}.$$

Podle centrální limitní věty má tato veličina pro velká n rozdělení velmi podobné rozdělení $N(0, 1)$.

Jinými slovy, rozdělení $\text{Bi}(n, p)$ je velice blízké rozdělení $N(np, np(1 - p))$ pro velká n . To je obsahem tzv.

Laplaceovy–Moivreovy věty. To jsme už viděli minule na obrázcích:

Pro hodnoty $Bi(5000, 0.5)$ je výsledek vidět na obrázku níže. Druhá křivka na obrázku je grafem funkce $f(x) = e^{-x^2/2}$.



Aproximace binomického rozdělení normálním se často považuje v praxi za dostatečnou, jestliže $np(1 - p) > 9$

Při praktických průzkumech zpravidla věříme „zákonu velkých čísel“. Potřebujeme přitom rozhodnout, jak velký vzorek už postačuje.

Typickým příkladem je např. tato úloha: Chceme zjistit poměr p osob s danou krevní skupinou A v populaci. U kolika osob je třeba krevní skupinu skutečně zjistit, abychom měli 90% pravděpodobnost, že naše zjištění se nebude lišit o více než 5%. Propočítáním zjistíme, že (nezávisle na p) vždy stačí odhadnout $p = X/n$, kde X je náhodná veličina udávající počet osob majících požadovanou skupinu, pro vzorek 270 lidí.