Teória pravdepodobnosti

9. Číselné charakteristiky systému náhodných premenných

Aleš Kozubík

Katedra Matematických metód
Fakulta Riadenia a Informatiky
Žilinská Univerzita v Žiline

18. novembra 2015

Stredná hodnota

Definícia (Stredná hodnota)

Nech náhodné premenné $(X_1, \dots, X_n)$ majú združenú hustotu $f(x_1, \dots, x_n)$ a nech $g(x_1, \dots, x_n)$ je reálna funkcia premenných $x_1, \dots, x_n$. *Strednou hodnotou* funkcie $g(X_1, \dots X_n)$ nazývame integrál

$$\mathbb{E}\left(g(X_1, \dots X_n)\right) = \int_{\mathbb{R}^n} g(x_1, \dots, x_n) f(x_1, \dots, x_n) \, \mathrm{d}x_1 \dots \mathrm{d}x_n, \quad (1)$$

pokiaľ daný integrál konverguje.

Stredná hodnota

V prípade diskrétnej náhodnej premennej daný integrál prepíšeme do tvaru sumy

Definícia (Stredná hodnota)

Nech diskrétne náhodné premenné $(X_1, \ldots, X_n)$ majú združenú pravdepodobnostnú funkciu $p(x_1, \ldots, x_n)$, nech náhodná premenná X_i nadobúda hodnoty z množiny H_i, $i = 1, \ldots, n$ a nech $g(x_1, \ldots, x_n)$ je reálna funkcia premenných $x_1, \ldots, x_n$. *Strednou hodnotou* funkcie $g(X_1, \ldots X_n)$ nazývame súčet

$$\mathbb{E}\left(g(X_1, \ldots X_n)\right) = \sum_{x_1 \in H_1} \cdots \sum_{x_n \in H_n} g(x_1, \ldots, x_n) p(x_1, \ldots, x_n), \quad (2)$$

pokiaľ daný rad konverguje.

Stredná hodnota

Poznámka

Pre stredné hodnoty jednotlivých náhodných premenných zo systému dostávame vyjadrenie

$$\mathbb{E}\left(X_i\right) = \int_{\mathbb{R}^n} x_i f\left(x_1, \ldots, x_n\right) \mathrm{d}x_1 \ldots \mathrm{d}x_n = \int_{-\infty}^{\infty} x f_i(x) \, \mathrm{d}x,$$

resp.

$$\mathbb{E}\left(X_i\right) = \sum_{x_1 \in H_1} \cdots \sum_{x_n \in H_n} x_i p(x_1, \ldots, x_n) = \sum_{x \in H_i} x p_i(x).$$

Stredná hodnota
Príklad

Zadanie

Náhodné premenné X a Y sú nezávislé a majú rovnomerné rozdelenie na intervale $\langle 0; 1\rangle$. Určme strednú hodnotu ich vzdialenosti t.j. $\mathbb{E}\left(|X - Y|\right)$.

Stredná hodnota
Riešenie príkladu

Podľa (1) platí

$$\begin{aligned}
\mathbb{E}\left(|X-Y|\right) &= \int_0^1 \int_0^1 |x-y| \, \mathrm{d}x \, \mathrm{d}y \\
&= \int_0^1 \left[\int_0^y (y-x) \, \mathrm{d}x + \int_y^1 (x-y) \, \mathrm{d}x \right] \mathrm{d}y \\
&= \int_0^1 \left[\left(y^2 - \frac{y^2}{2} \right) + \left(\frac{1-y^2}{2} - y(1-y) \right) \right] \mathrm{d}y \\
&= \int_0^1 \left[\frac{1}{2} - y + y^2 \right] \mathrm{d}y \\
&= \frac{1}{2} - \frac{1}{2} + \frac{1}{3} = \frac{1}{3}.
\end{aligned}$$

Momenty

S využitím (1) resp. (2) je možné definovať:

Definícia (Začiatočný moment)

Začiatočným momentom $\nu_{k_1,\ldots,k_n}$, $(k_1,\ldots,k_n)$–tého rádu systému náhodných premenných $(X_1,\ldots,X_n)$ sa nazýva stredná hodnota funkcie $g(x_1,\ldots,x_n) = x_1^{k_1}\cdots x_n^{k_n}$, t.j.

$$\nu_{k_1,\ldots,k_n} = \mathbb{E}\left(X_1^{k_1}X_2^{k_2}\cdots X_n^{k_n}\right).$$

Momenty

S využitím (1) resp. (2) je možné definovať:

Definícia (Centrálny moment)

Centrálnym momentom $\mu_{k_1,\dots,k_n}$, $(k_1,\dots,k_n)$–tého rádu systému náhodných premenných $(X_1,\dots,X_n)$ nazývame strednú hodnotu funkcie $g(x_1,\dots,x_n) = (x_1 - \mathbb{E}(X_1))^{k_1}\cdots(x_n - \mathbb{E}(X_n))^{k_n}$, t.j.

$$\mu_{k_1,\dots,k_n} = \mathbb{E}\left((X_1 - \mathbb{E}(X_1))^{k_1}\cdots(X_n - \mathbb{E}(X_n))^{k_n}\right).$$

Významné momenty

Stredné hodnoty

Zrejme počiatočné momenty, v ktorých je jeden index rovný 1 a ostatné sú nulové zodpovedajú stredným hodnotám, jednotlivých náhodných premenných, teda:

$$\begin{aligned} \nu_{1,0,0,\ldots,0} &= \mathbb{E}\left(X_1\right) \\ \nu_{0,1,0,\ldots,0} &= \mathbb{E}\left(X_2\right) \\ &\vdots \\ \nu_{0,0,0,\ldots,1} &= \mathbb{E}\left(X_n\right) \end{aligned}$$

Významné momenty

Rozptyl

Zrejme centrálne momenty, v ktorých je jeden index rovný 2 a ostatné sú nulové zodpovedajú rozptylom, jednotlivých náhodných premenných, teda:

$$\mu_{2,0,0,\dots,0} = \mathbb{D}\left(X_1\right)$$
$$\mu_{0,2,0,\dots,0} = \mathbb{D}\left(X_2\right)$$
$$\vdots$$
$$\mu_{0,0,0,\dots,2} = \mathbb{D}\left(X_n\right)$$

Významné momenty

Definícia (Kovariancia)

Pod *kovarianciou* dvoch náhodných premenných X_i a X_j, $i \neq j$ rozumieme centrálny moment, ktorého i-ty a j-ty index sú rovné 1 a ostatné sú nulové. Kovarianciu náhodných premenných X_i a X_j označujeme symbolom $\operatorname{cov}(X_i; X_j)$. Platí teda

$$\begin{aligned}
\mu_{1,1,0,\ldots,0} &= \operatorname{cov}(X_1; X_2) \\
\mu_{1,0,1,0\ldots,0} &= \operatorname{cov}(X_1; X_3) \\
&\vdots \\
\mu_{0,0,\ldots,0,1,1} &= \operatorname{cov}(X_{n-1}; X_n)
\end{aligned}$$

Významné momenty

Definícia (Korelačný koeficient)

Korelačným koeficientom $\rho_{X,Y}$ dvoch náhodných premenných X a Y rozumieme podiel

$$\rho_{X,Y} = \frac{\operatorname{cov}(X;Y)}{\sqrt{\mathbb{D}(X)\,\mathbb{D}(Y)}}. \tag{3}$$

Významné momenty
Alternatívne vyjadrenie kovariancie

Veta (Alternatívne vyjadrenie kovariancie)

Nech X a Y sú náhodné premenné také, že $\mathbb{E}(X)$, $\mathbb{E}(Y)$ a $\mathbb{E}(XY)$ existujú. Potom platí

$$\operatorname{cov}(X;Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y)\,. \tag{4}$$

Dôkaz:

$$\begin{aligned}
\operatorname{cov}(X;Y) &= \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right)\\
&= \mathbb{E}\left(XY - \mathbb{E}(X)\,Y - \mathbb{E}(Y)\,X + \mathbb{E}(X)\,\mathbb{E}(Y)\right)\\
&= \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y) - \mathbb{E}(Y)\,\mathbb{E}(X) + \mathbb{E}(X)\,\mathbb{E}(Y)\\
&= \mathbb{E}(XY) - \mathbb{E}(Y)\,\mathbb{E}(X)\,.
\end{aligned}$$

Významné momenty
Príklad

Zadanie

Nech náhodný vektor (X, Y) má rovnomerné rozdelenie na jednotkovom kruhu $D = \{(x, y) : x^2 + y^2 < 1\}$. Určme kovarianciu $\operatorname{cov}(X; Y)$.

Riešenie:
Už sme zistili, že $\mathbb{E}(X)=\mathbb{E}(Y)=0$. Preto:

$$\operatorname{cov}(X; Y) = \int_{-1}^{1} \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} xy \,\mathrm{d}y \,\mathrm{d}x = \int_{-1}^{1} \frac{2x}{\pi} \sqrt{1 - x^2} \,\mathrm{d}x = 0.$$

Významné momenty
Kovariancia a nezávislosť

Veta

Nech náhodné premenné X a Y sú nezávislé. Potom $\operatorname{cov}(X;Y)=0$.

Dôkaz:
Ak X a Y sú nezávislé, tak $\mathbb{E}(XY)=\mathbb{E}(X)\,\mathbb{E}(Y)$.

Poznámka

Predchádzajúci príklad ukazuje, že opak neplatí. Teda ak $\operatorname{cov}(X;Y)=0$, tak náhodné premenné X a Y nemusia byť nezávislé. Ak $\operatorname{cov}(X;Y)=0$, tak náhodné premenné X a Y nazývame *nekorelované*.

Významné momenty
Aditivita rozptylu

Určme rozptyl náhodnej premennej $Z = X + Y$.

$$\begin{aligned}
\mathbb{D}(Z) &= \mathbb{E}\left((X+Y)^2\right) - \left[\mathbb{E}(X+Y)\right]^2 \\
&= \mathbb{E}\left(X^2 + Y^2 + 2XY\right) - \left[\mathbb{E}(X+Y)\right]^2 \\
&= \mathbb{E}\left(X^2\right) + \mathbb{E}\left(Y^2\right) + 2\mathbb{E}(XY) \\
&\quad - \left[\mathbb{E}(X)\right]^2 - \left[\mathbb{E}(Y)\right]^2 - 2\mathbb{E}(X)\,\mathbb{E}(Y) \\
&= \mathbb{D}(X) + \mathbb{D}(Y) + 2\operatorname{cov}(X;Y).
\end{aligned}$$

Podmienená stredná hodnota

Definícia (Podmienená stredná hodnota)

Nech A je náhodná udalosť taká, že $\mathbb{P}(A) \neq 0$ a X nech je náhodná premenná. Ak je X diskrétna náhodná premenná tak definujeme *podmienenú strednú hodnotu X za podmienky A* vzťahom:

$$\mathbb{E}_A(X) = \sum_{x: p_{X|A}(x) > 0} x p_{X|A}(x). \tag{5}$$

Ak je X spojitá náhodná premenná taká, že existuje hustota $f_{X|A}(x)$, tak definujeme *podmienenú strednú hodnotu X za podmienky A* vzťahom:

$$\mathbb{E}_A(X) = \int_{-\infty}^{\infty} x f_{X|A}(x)\,\mathrm{d}x. \tag{6}$$

Podmienená stredná hodnota

Ak náhodnú udalosť A spojíme s náhodnou premennou Y a jej nadobudnutím hodnoty y, tak môžeme definovať:

Definícia (Podmienená stredná hodnota)

Nech X je diskrétna a Y ľubovoľná náhodná premenná, taká, že $p_{X|Y}$ existuje. *Podmienenú strednú hodnotu X za podmienky $Y = y$* definujeme vzťahom:

$$\mathbb{E}_y(X) = \sum_{x: p_{X|A}(x,y)>0} x p_{X|Y}(x, y). \tag{7}$$

Ak je X spojitá a Y ľubovoľná náhodná premenná taká, že existuje hustota $f_{X|Y}$, tak definujeme *podmienenú strednú hodnotu X za podmienky $Y = y$* vzťahom:

$$\mathbb{E}_y(X) = \int_{-\infty}^{\infty} x f_{X|Y}(x, y)\, \mathrm{d}x. \tag{8}$$

Podmienená stredná hodnota

Podmienenú strednú hodnotu môžme vyjadriť pomocou podmienených rozdelení.

Podmienená stredná hodnota

Nech X a Y majú združenú pravdepodobnostnú funkciu $p(x, y)$. Podmienenú strednú hodnotu X pri podmienke $Y = y$ potom môžeme definovať ako

$$\mathbb{E}\,(X|Y = y) = \sum_x xp(x|y) = \frac{\sum_x xp(x, y)}{\sum_x p(x, y)}. \tag{9}$$

resp.

$$\mathbb{E}\,(X|Y = y) = \int_{-\infty}^{\infty} xf(x|y)\,\mathrm{d}x = \frac{\int_{-\infty}^{\infty} xf(x, y)\,\mathrm{d}x}{\int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}x}. \tag{10}$$

Podmienená stredná hodnota

Pre podmienenú strednú hodnotu ostávajú v platnosti všetky tvrdenia o strednej hodnote, najmä:

Vlastnosti podmienenej strednej hodnoty

$$
\begin{aligned}
\mathbb{E}_y(aX+b) &= a\mathbb{E}_y(X)+b,\\
\mathbb{E}_y(X_1+X_2) &= \mathbb{E}_y(X_1)+\mathbb{E}_y(X_2),\\
\mathbb{E}_y(g(X)) &= \sum_{x:p_{X|Y}>0} g(x)p_{X|Y}(x,y)\\
\text{resp.}&\\
\mathbb{E}_y(g(X)) &= \int_{-\infty}^{\infty} g(x)f_{X|Y}(x,y)\,\mathrm{d}x.
\end{aligned}
$$

Podmienená stredná hodnota

Nasledujúca veta je obdobou vety o úplnej pravdepodobnosti, a ukazuje, ako je možné využiť podmieňovanie

Veta (O úplnej strednej hodnote)

Nech X a Y sú náhodné premenné definované na rovnakom pravdepodobnostnom priestore. Predpokladajme, že $\mathbb{E}(X)$ a $\mathbb{E}(X|Y=y)$ existujú pre každé y. Potom

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Y=y)),$$

t.j.

$$\mathbb{E}(X) = \sum_y \mathbb{E}(X|Y=y)\, p_Y(y), \tag{11}$$

resp.

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} \mathbb{E}(X|Y=y)\, f_Y(y)\, \mathrm{d}y. \tag{12}$$

Podmienená stredná hodnota

Príklad

Zadanie

Nech náhodný vektor (X, Y) má rovnomerné rozdelenie na pravom jednotkovom polkruhu $D = \{(x, y) : x^2 + y^2 < 1, 0 < x\}$. Určte podmienenú strednú hodnotu $\mathbb{E}_y(X)$ a $\mathbb{E}\,(X)$.

Podmienená stredná hodnota
Riešenie príkladu

Pre podmienenú strednú hodnotu platí

$$\begin{aligned}
\mathbb{E}_y(X) &= \int_{-\infty}^{\infty} x f_{X|Y}(x,y)\,\mathrm{d}x \\
&= \int_0^{\sqrt{1-y^2}} \frac{x}{\sqrt{1-y^2}}\,\mathrm{d}x \\
&= \frac{\sqrt{1-y^2}}{2}, \quad \text{pre } y \in (-1;1)
\end{aligned}$$

Podmienená stredná hodnota

Riešenie príkladu

Pre strednú hodnotu náhodnej premennej X potom máme:

$$
\begin{aligned}
\mathbb{E}\left(X\right) = \mathbb{E}\left(\mathbb{E}_y(X)\right) &= \int_{-\infty}^{\infty} \mathbb{E}_y(X) f_Y(y)\,\mathrm{d}y \\
&= \int_{-1}^{1} \frac{\sqrt{1-y^2}}{2} \cdot \frac{2}{\pi}\sqrt{1-y^2}\,\mathrm{d}y \\
&= \int_{-1}^{1} \frac{1-y^2}{\pi}\,\mathrm{d}y = \frac{4}{3\pi}.
\end{aligned}
$$

Podmienený rozptyl

Definícia (Podmienený rozptyl)

Nech náhodný vektor (X, Y) má združenú funkciu pravdepodobnosti $p(x, y)$ a nech $\mu_X = \mathbb{E}(X|Y = y)$. *Podmieneným rozptylom X za podmienky $Y = y$* rozumieme hodnotu:

$$\mathbb{D}(X|Y = y) = \mathbb{E}\left((X - \mu_X)^2 | Y = y\right) = \sum_x (x - \mu_X)^2 p(x|y). \tag{13}$$

Ak (X, Y) má združenú hustotu $f(x, y)$, tak podmienený rozptyl definujeme ako:

$$\mathbb{D}(X|Y = y) = \frac{\int_{-\infty}^{\infty} (x - \mu_X)^2 f(x, y)\, dx}{\int_{-\infty}^{\infty} f(x, y)\, dx}. \tag{14}$$

Najlepší prediktor

Veta (Najlepší prediktor X pri danom Y)

Nech X a Y sú náhodné premenné a nech $g(Y)$ je ľubovoľná funkcia. Potom stredná kvadratická odchýlka $\mathbb{E}\left([X - g(Y)]^2\right)$ je minimálna, ak $g(y) = \mathbb{E}\left(X|Y = y\right)$

Najlepší prediktor
Dôkaz

$$
\begin{aligned}
\mathbb{E}\left([X-g(Y)]^2\right) &= \mathbb{E}\left([X-\mathbb{E}(X|Y=y)+\mathbb{E}(X|Y=y)-g(y)]^2\right)\\
&= \mathbb{E}\left([X-\mathbb{E}(X|Y=y)]^2\right)+\mathbb{E}\left([\mathbb{E}(X|Y=y)-g(y)]^2\right.\\
&\quad +2\mathbb{E}\left([X-\mathbb{E}(X|Y=y)][\mathbb{E}(X|Y=y)-g(y)]\right)
\end{aligned}
$$

Podľa vlastností strednej hodnoty môžeme posledný člen zapísať v tvare

$$2\mathbb{E}\left(\mathbb{E}\left([X-\mathbb{E}(X|Y=y)][\mathbb{E}(X|Y=y)-g(y)]|Y=y\right)\right)$$

Najlepší prediktor
Dôkaz

Všimnime si, že platí

$$\begin{aligned}
\mathbb{E}_y(k(X)h(Y)) &= \int_{-\infty}^{\infty} k(x)h(y)f_{X|Y}(x,y)\,\mathrm{d}x \\
&= h(y)\int_{-\infty}^{\infty} k(x)f_{X|Y}(x,y)\,\mathrm{d}x \\
&= h(y)\mathbb{E}_y(k(x)).
\end{aligned}$$

Celkom teda

$$\mathbb{E}_y(k(X)h(Y)) = h(Y)\mathbb{E}_Y(k(x)).$$

Najlepší prediktor
Dôkaz

Ak v predchádzajúcom výsledku položíme $k(X) = X - \mathbb{E}_Y(X)$ a $h(Y) = \mathbb{E}_Y(X) - g(Y)$, tak dostávame

$$\begin{aligned}
&\mathbb{E}_Y[(X - \mathbb{E}_Y(X))(\mathbb{E}_Y(X) - g(Y))] \\
&\qquad = (\mathbb{E}_Y(X) - g(Y))\mathbb{E}_Y[(X - \mathbb{E}_Y(X))] \\
&\qquad = (\mathbb{E}_Y(X) - g(Y))[\mathbb{E}_Y(X) - \mathbb{E}_Y(\mathbb{E}_Y(X))] = 0,
\end{aligned}$$

lebo $\mathbb{E}_Y(\mathbb{E}_Y(X)) = \mathbb{E}_Y(X)$.

Najlepší prediktor
Dôkaz

Celkom teda máme

$$\mathbb{E}\left([X-g(y)]^2\right)=\mathbb{E}\left([X-\mathbb{E}_Y(X)]^2\right)+\mathbb{E}\left([\mathbb{E}_Y(X)-g(Y)]^2\right).$$

Obidva členy na pravej strane sú nezáporné a preto ich súčet dosiahne minimálnu hodnotu vtedy, ak $g(Y)=\mathbb{E}_Y(X)$.

Zákon podmieneného rozptylu

Nasledujúca veta ukazuje prepojenie medzi podmieneným a nepodmieneným rozptylom.

Veta (Zákon podmieneného rozptylu)

Pre náhodné premenné X a Y platí

$$\mathbb{D}(X) = \mathbb{E}(\mathbb{D}(X|Y)) + \mathbb{D}([\mathbb{E}(X|Y)) .$$

Zákon podmieneného rozptylu
Dôkaz

Z definície rozptylu postupne dostávame

$$\begin{aligned}
\mathbb{D}(X) &= \mathbb{E}\left((X-\mathbb{E}(X))^2\right)\\
&= \mathbb{E}\left((X-\mathbb{E}(X|Y)+\mathbb{E}(X|Y)-\mathbb{E}(X))^2\right)\\
&= \mathbb{E}\left((X-\mathbb{E}(X|Y))^2\right)+\mathbb{E}\left((\mathbb{E}(X|Y)-\mathbb{E}(X))^2\right)\\
&\quad +2\mathbb{E}\left((X-\mathbb{E}(X|Y))(\mathbb{E}(X|Y)-\mathbb{E}(X))\right)
\end{aligned}$$

Zákon podmieneného rozptylu
Dôkaz

Podľa vety o úplnej strednej hodnote pre prvý člen súčtu platí

$$\mathbb{E}\left((X - \mathbb{E}(X|Y))^2\right) = \mathbb{E}\left(\mathbb{E}\left((X - \mathbb{E}(X|Y))^2 | Y\right)\right) = \mathbb{E}(\mathbb{D}(X|Y)).$$

Druhý sčítanec je zrejme rovný

$$\mathbb{E}\left((\mathbb{E}(X|Y) - \mathbb{E}(X))^2\right) = \mathbb{D}(\mathbb{E}(X|Y))$$

pretože $\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Y))$.

Zákon podmieneného rozptylu
Dôkaz

Ostáva ukázať, že posledný člen je rovný nule. Označme $h(Y) = 2(\mathbb{E}(X|Y) - \mathbb{E}(X))$. Potom pre tretí člen dostávame:

$$\begin{aligned}
\mathbb{E}\left((\mathbb{E}(X|Y) - \mathbb{E}(X))^2\right) &= \mathbb{E}(Xh(Y)) - \mathbb{E}(\mathbb{E}(X|Y)\, h(Y)) \\
&= \mathbb{E}(Xh(Y)) - \mathbb{E}(\mathbb{E}(Xh(Y)|Y)) \\
&= \mathbb{E}(Xh(Y)) - \mathbb{E}(Xh(Y)) \\
&= 0.
\end{aligned}$$