NMSA202 – PST & MAT STAT: KAPITOLY 3–4

Dk V Výpočet střední hodnoty.

Aplikace věty o přenosu integrace. ■

Dk V Diskrétní a spojitá střední hodnota.

Aplikace Radon-Nikodýmovy věty. ■

Dk V Pravidlo Líného statistika.

Zobrazení $Y := h(X) : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ je náhodná veličina, protože je to složení dvou měřitelných zobrazení $(h \circ X)$.

Potom
$$EY = E[h(X)] \underset{\substack{\uparrow \\ \text{definice střední hodnoty}}}{=} \int_\Omega h[X(\omega)]\,dP(\omega) \underset{\substack{\uparrow \\ \text{věta o přenosu integrace}}}{=} \int_\mathbb{R} h(x)\,dP_X(x).$$
■

Dk V Vyšší momenty. $0 < \ell \le k \Rightarrow E|X|^\ell \le E|X|^k < +\infty$

Počítejme:
$$E|X|^\ell = \int_\mathbb{R} |x|^\ell dP_X(x) = \int_{|x|\le 1} |x|^\ell dP_X(x) + \int_{|x|>1} |x|^\ell dP_X(x) \le$$
$$\le \int_{|x|\le 1} dP_X(x) + \int_{|x|>1} |x|^k dP_X(x) \le 1 + E|X|^k < +\infty.$$
■

Dk V Linerita $E[\cdot]$.

Důsledek věty o přenosu integrace a linearity Lebesgueova integrálu. ■

Dk V Násobení při nezávislosti.

Uvažujme posloupnost fcí $\{g_n : \mathbb{R}^d \to \mathbb{R}\}_{n\in\mathbb{N}}$ definovaných $g_n(\mathbf{x}) = \prod_{\ell=1}^d |x_\ell| \mathbb{1}\{|x_\ell| \le n\}$.

Pak $g_n(\mathbf{X})$ je omezená a $E[g_n(\mathbf{X})] \in \mathbb{R}$. Navíc
$$E[g_n(\mathbf{X})] \underset{\substack{\uparrow \\ \text{nezávislost}}}{=} \int_{\mathbb{R}^d} \prod_{\ell=1}^d |x_\ell| \mathbb{1}\{|x_\ell|\le n\}\, d\Big(\bigotimes_{\ell=1}^d P_{X_\ell}\Big)(\mathbf{x}) \underset{\substack{\uparrow \\ \text{Fubini}}}{=} \underbrace{\int_\mathbb{R} \cdots \int_\mathbb{R}}_{d\times} \prod_{\ell=1}^d |x_\ell| \mathbb{1}\{|x_\ell|\le n\} \underbrace{\quad}_{dP_{X_1}(x_1)\dots dP_{X_d}(x_d)} =$$
$$\underset{\substack{\uparrow \\ \text{Linearita integrálu}}}{=} \prod_{\ell=1}^d E\big[|X_\ell| \mathbb{1}\{|X_\ell| \le n\}\big] \le \prod_{\ell=1}^d E|X_\ell|.$$

Navíc nezáporná fce $g_n(\mathbf{x}) \uparrow \prod_{\ell=1}^d |x_\ell|$ na $\mathbb{R}^d$.

Takže z Leviho věty $E[g_n(\mathbf{X})] \uparrow E\Big[\prod_{\ell=1}^d |X_\ell|\Big]$. Potom $E\Big|\prod_{\ell=1}^d X_\ell\Big| \le \prod_{\ell=1}^d E|X_\ell| < +\infty$.

Dále počítejme

$$E\left[\prod_{\ell=1}^{d} X_\ell\right] = \int_{\mathbb{R}^d} \prod_{\ell=1}^{d} x_\ell \, dP_X(x) \underset{\text{nezávislost}}{=} \int_{\mathbb{R}^d} \prod_{\ell=1}^{d} x_\ell \, d\left(\bigotimes_{\ell=1}^{d} P_{X_\ell}\right)(x) \underset{\text{Fubini}}{=} \underbrace{\int_{\mathbb{R}} \cdots \int_{\mathbb{R}}}_{dx} \prod_{\ell=1}^{d} x_\ell \, dP_{X_1}(x_1) \ldots dP_{X_d}(x_d) =$$

$$\underset{\text{Linearita integrálu}}{=} \prod_{\ell=1}^{d} \int_{\mathbb{R}} x_\ell \, dP_{X_\ell}(x_\ell) = \prod_{\ell=1}^{d} EX_\ell. \quad \blacksquare$$

Dk V Vlastnosti rozptylu

1) $\operatorname{var} X = E\underbrace{[X - EX]^2}_{\geq 0} \geq 0$ & $\operatorname{var} X = E[X - (EX)]^2 = E[X^2 - 2X(EX) + (EX)^2] =$

$$\underset{\text{Linearita } E(\cdot)}{=} E(X^2) - \underbrace{2(EX)EX}_{=2(EX)^2} + (EX)^2 = EX^2 - (EX)^2$$

2) $\operatorname{var}(aX+b) = E[aX + b - E(aX+b)]^2 = E[a\{X - EX\}]^2 = a^2 E[X - EX]^2 =$

$$= a^2 \operatorname{var} X$$

3) $\operatorname{var}\left(\sum_{\ell=1}^{d} a_\ell X_\ell\right) = E\left[\sum_{\ell=1}^{d} a_\ell X_\ell - E\left(\sum_{\ell=1}^{d} a_\ell X_\ell\right)\right]^2 = E\left[\sum_{\ell=1}^{d} a_\ell (X_\ell - EX_\ell)\right]^2 =$

$$= E\left[\sum_{\ell=1}^{d} a_\ell^2 (X_\ell - EX_\ell)^2 + 2 \sum\sum_{1 \leq j < \ell \leq d} a_j a_\ell (X_j - EX_j)(X_\ell - EX_\ell)\right] =$$

$$= \sum_{\ell=1}^{d} a_\ell^2 \underbrace{E(X_\ell - EX_\ell)^2}_{\operatorname{var} X_\ell} + 2 \sum\sum_{1 \leq j < \ell \leq d} a_j a_\ell \underbrace{E(X_j - EX_j)(X_\ell - EX_\ell)}$$

$$= E[X_j X_\ell - (EX_\ell)EX_j - (EX_j)EX_\ell + (EX_j)(EX_\ell)] =$$

$$= E(X_j X_\ell) - (EX_j)(EX_\ell) \underset{X_j \perp\!\!\!\perp X_\ell}{=} 0 \quad \blacksquare$$

Dk V Vlastnosti kovariance a korelace

1) $\operatorname{cov}(X,Y) = E[(X - EX)(Y - EY)] = E[XY - (EX)EY - (EY)EX + (EX)(EY)] =$

$$= E(XY) - (EX)(EY)$$

2) $\underset{\mathbb{R}}{g(a)} := E(aX - Y)^2 \Rightarrow 0 \leq E(aX - b)^2 = a^2 EX^2 - 2a E(XY) + EY^2$

$$\rightsquigarrow g'(a) = 2a EX^2 - 2E(XY) \Rightarrow \operatorname*{argmin}_{a \in \mathbb{R}} g(a) = \frac{E(XY)}{EX^2}$$

Dosadíme: $g\left(\frac{EXY}{EX^2}\right) = \frac{(EXY)^2}{EX^2} - 2\frac{(EXY)^2}{EX^2} + EY^2$ a to musí být ≥ 0.

Pak $(EXY)^2 \leq (EX^2)(EY^2)$ … Cauchy-Schwarzova nerovnost.

$$\Downarrow$$

$$[\operatorname{Corr}(X,Y)]^2 \leq 1$$

3) "$\Leftarrow$" $\mathrm{Cov}(X,Y) \underset{P[Y=aX+b]=1}{\overset{\uparrow}{=}} \mathrm{cov}(X, aX+b) = E[X(aX+b)] - (EX)[E(aX+b)] =$

$= a\,EX^2 + \cancel{b\,EX} - a(EX)^2 - \cancel{b\,EX} = a\,\mathrm{var}\,X$

$$|\mathrm{Corr}(X,Y)| = \frac{|\mathrm{Cov}(X,Y)|}{\sqrt{\mathrm{var}\,X\,\mathrm{var}\,Y}} = \frac{|a\,\mathrm{var}\,X|}{\sqrt{(\mathrm{var}\,X)\,a^2\,\mathrm{var}\,X}} = 1$$

"$\Rightarrow$" Rovnost nastává, když $|\mathrm{Cov}(X,Y)| = \sqrt{\mathrm{var}X\,\mathrm{var}Y}$. A to je právě když $[E\underbrace{(X-EX)}_{=:\tilde{X}}\underbrace{(Y-EY)}_{=:\tilde{Y}}]^2 = \underbrace{[E(X-EX)^2]}_{=E\tilde{X}^2}\underbrace{[E(Y-EY)^2]}_{=E\tilde{Y}^2}$.

Dosadíme $a = \frac{E\tilde{X}\tilde{Y}}{E\tilde{X}^2}$ do $g(a)$ z 2) a dostáváme

$0 = E\left[\frac{E\tilde{X}\tilde{Y}}{E\tilde{X}^2}\tilde{X} - \tilde{Y}\right]^2$ a tedy musí platit $P[a\tilde{X}-\tilde{Y}=0]=1$.

Pak s pstí 1 platí $aX - aEX + EY = Y$.

4) $\mathrm{Cov}(X,Y) = EXY - (EX)(EY) \underset{X\perp\!\!\!\perp Y}{\overset{\uparrow}{=}} (EX)(EY) - (EX)(EY) = 0$ ∎

Dk D Rozptyl součtu

Viz Dk V Vlastnosti rozptylu 3). ∎

Dk V Vlastnosti varianční-kovarianční matice

Přímý výpočet. ∎

Platí pro $h>0$: $|x|^m \le \frac{m!\,e^{|x|h}}{h^m}$, $\forall x\in\mathbb{R}, \forall m\in\mathbb{N}$

Dk V Vlastnosti MGF

- Definujme fci $g(x) := (m!)\left(\frac{6}{\epsilon}\right)^m\left[\exp\{-\frac{\epsilon}{2}x\} + \exp\{+\frac{\epsilon}{2}x\}\right]$, $x\in\mathbb{R}$.

Potom platí $|x|^m\exp\{tx\} \le (m!)\left(\frac{6}{\epsilon}\right)^m \exp\{|x|\frac{\epsilon}{6}\}\exp\{|x|\frac{\epsilon}{3}\} \le g(x)$, $\forall x\in\mathbb{R}$, $\forall t\in[-\frac{\epsilon}{3},\frac{\epsilon}{3}]$

Podle předpokladů věty máme

$$\int g(x)\,dP_X(x) = (m!)\left(\frac{6}{\epsilon}\right)^m\left[\psi_X(-\tfrac{\epsilon}{2}) + \psi_X(\tfrac{\epsilon}{2})\right] < \infty$$ a tedy

dostáváme konvergentní majorantu. Pak můžeme zaměnit derivaci a integrál

$$\frac{d}{dt}\psi_X(t) = \frac{d}{dt}\int \exp\{tx\}\,dP_X(x) = \int x\exp\{tx\}\,dP_X(x)$$

a dosadíme $\psi_X'(0) = \int x e^0\,dP_X(x) = \int x\,dP_X(x) = EX$. Zbytek indukce.

- Počítejme
$$\psi_Y(t) = E[\exp\{t(aX+b)\}] = \exp\{tb\}E[\exp\{(at)X\}] = e^{tb}\psi_X(at)$$
- Z nezávislosti $\{X_1,\dots,X_d\}$ plyne $\perp\!\!\!\perp \{e^{tX_1},\dots,e^{tX_d}\}\ \forall t$ (viz V Alternativní charakterizace nezávislosti)
$$\psi_Y(t) = E[\exp\{t\sum_{\ell=1}^{d} X_\ell\}] = E[\prod_{\ell=1}^{d}\exp\{tX_\ell\}] \overset{\downarrow}{=} \prod_{\ell=1}^{d} E\exp\{tX_\ell\} = $$
$$= \prod_{\ell=1}^{d}\psi_{X_\ell}(t) \quad \blacksquare$$

Dk V Vlastnosti CF

(i) Víme, že $\forall x\ \forall t: |e^{itx}|^2 = \sin^2(tx)+\cos^2(tx) = 1$.
Pak $E|e^{itX}|^2 = 1$. Potom $E|e^{itX}| \underset{\text{Jensen}}{\le} \sqrt{E|e^{itX}|^2} = 1$ a tedy e^{itx} je integrovatelná.

(ii) $\varphi_X(0) = E[\exp\{i\cdot 0\cdot X\}] = E[1] = 1$.

(iii) Plyne z (i).

(iv) Položme $h := t-s$. Potom
$$|\varphi_X(t)-\varphi_X(s)| = |E[\exp\{isX\}(\exp\{ihX\}-1)]| \le$$
$$\le E[|\exp\{isX\}(\exp\{ihX\}-1)|] \le$$
$$\le E[\underbrace{|\exp\{isX\}|}_{\le 1 \text{ dle (iii)}}|\exp\{ihX\}-1|] \le E|\exp\{ihX\}-1|.$$
Víme, že $\exp\{ihX\}-1 \to 0$ když $h\to 0$ a $|\exp\{ihX\}-1| \le 2$, což poslouží jako integrovatelná majoranta. Pak podle Lebesgueovy věty dostáváme $\lim_{h\to 0} E|\exp\{ihX\}-1| = 0$.

Takže pro $\forall \varepsilon>0\ \exists \delta>0\ \forall t,s\in\mathbb{R}: |t-s|\le\delta : |\varphi_X(t)-\varphi_X(s)| \le \varepsilon$.

(v) Z definice máme
$$\varphi_{a+bX}(t) = E[\exp\{it(a+bX)\}] = \exp\{ita\}E[\exp\{itbX\}] = e^{iat}\varphi_X(bt).$$

(vi) Počítejme
$$\varphi_{-X}(t) = E[\exp\{it(-X)\}] = E[\cos(-tX)+i\sin(-tX)] = E[\cos(+tX)-i\sin(+tX)] =$$
$$= \overline{E[\cos(tX)+i\sin(tX)]} = \overline{\varphi_X}(t).$$

(vii) X má symetrické rozdělení kolem $0 \iff X \overset{d}{=} -X \overset{\text{následující věta}}{\iff} \varphi_{-X}(t) = \varphi_X(t)\ \forall t$
(ale dle (vi) $\varphi_{-X}(t) = \overline{\varphi_X}(t)$) $\iff \varphi_X(t) = \overline{\varphi_X}(t) \iff \varphi_X(t)\in\mathbb{R}\ \forall t\in\mathbb{R}$

(viii) Z nezávislosti $X \perp\!\!\!\perp Y$ máme $e^{itX} \perp\!\!\!\perp e^{itY}$ $\forall t$

$$\varphi_{X+Y}(t) = E[\exp\{it(X+Y)\}] = E[\exp\{itX\}\exp\{itY\}] = \varphi_X(t)\varphi_Y(t) \; \forall t \quad \blacksquare$$

Dk V Inverzní formule

(Δ): omezení (majoranta) $|e^{z}-1| \leq \frac{3}{2}|z|$ $\forall z \in \mathbb{C}: |z| \leq \frac{1}{2}$

Uvažujme integrál $\forall a<b$

$$\frac{1}{2\pi}\int_{-T}^{T} \frac{e^{-ita}-e^{-itb}}{it}\varphi_X(t)\,dt = \frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}\int_{\mathbb{R}} e^{itx}\,dP_X(x) =$$

$$= \int_{-T}^{T}\int_{-\infty}^{+\infty}\frac{e^{it(x-a)}-e^{it(x-b)}}{2\pi i t}\,dP_X(x)\,dt \overset{\text{Fubini}}{=} \int_{-\infty}^{+\infty}\int_{-T}^{T}\frac{e^{it(x-a)}-e^{it(x-b)}}{2\pi i t}\,dt\,dP_X(x).$$

Všimneme si, že $\forall c \in \mathbb{R}$: $\int_{-T}^{T}\frac{e^{itc}}{2it}\,dt = \int_{0}^{T}\frac{\sin(tc)}{t}\,dt.$

Potom

$$\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}\varphi_X(t)\,dt = \int_{-\infty}^{+\infty}\frac{1}{\pi}\left\{\int_0^T\frac{\sin[t(x-a)]}{t}\,dt - \int_0^T\frac{\sin[t(x-b)]}{t}\,dt\right\}dP_X(x). \quad (*)$$

Když $T\to\infty$, pak víme "integrál sinu"

$$\frac{1}{\pi}\int_0^T\frac{\sin[t(x-a)]}{t}\,dt \longrightarrow \begin{cases}-\frac{1}{2}, & x<a\\ +\frac{1}{2}, & x>a\\ 0, & x=a\end{cases}$$

Pak

$$\frac{1}{\pi}\left\{\int_0^T\frac{\sin[t(x-a)]}{t}\,dt - \int_0^T\frac{\sin[t(x-b)]}{t}\,dt\right\} \xrightarrow{T\to\infty} \begin{cases}\frac{1}{2}, & x=a\\ \frac{1}{2}, & x=b\\ 1, & a<x<b\\ 0, & \text{jinak}\end{cases}$$

Substitujeme do $(*)$ a užijeme Lebesgueovu větu o majorantě (Δ):

$$\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}\varphi_X(t)\,dt = P[a<X<b] + \frac{P[X=a]+P[X=b]}{2} \quad \blacksquare$$

Dk D Jednoznačná charakterizace rozdělení

Jelikož pravá strana předchozí věty platí $\forall a<b$, pak φ_X musí existovat jednoznačně pro X.

Dk V Markovova nerovnost

$X \geq 0 \Rightarrow P[X\geq 0]=1 \Rightarrow \int_{\{\omega\in\Omega: X(\omega)<0\}} dP(\omega) = 0$. Proto $\forall \varepsilon>0$ máme

$$P[X\geq\varepsilon] = \int_{\{\omega\in\Omega: X(\omega)\geq\varepsilon\}} dP(\omega) \leq \int_{\{\omega: X(\omega)\geq\varepsilon\}}\frac{X(\omega)}{\varepsilon}\,dP(\omega) \leq \int_{\{\omega: X(\omega)\geq 0\}}\frac{X(\omega)}{\varepsilon}\,dP(\omega) =$$

$$= \frac{1}{\varepsilon} \int_\Omega X(\omega)\, dP(\omega) = \frac{EX}{\varepsilon} \quad \blacksquare$$

Dk D zobecněná Markovova nerovnost

Uvažujme Markovovu nerovnost pro $\{X^r > \varepsilon^r\}$. $\blacksquare$

Dk V Čebyševova nerovnost

Aplikujeme předchozí důsledek na $|X-EX|$ a $r=2$. $\blacksquare$

Dk V Cauchy-Schwarzova nerovnost

Viz V Vlastnosti kovariance a korelace 2).

Dk V Jensenova nerovnost

Nechť je $l(x) = a + bx$ tečna k fci $g(x)$ v bodě EX.

Nechť je $l(x)=a+bx$ tečna k fci $g(x)$ v bodě EX. Když g je konvexní, pak $l(x) \leq g(x)\ \forall x \in \mathbb{R}$ a $l(EX) = g(EX)$.

Potom $E[g(X)] \geq E[l(X)] = E[a+bX] = a + bEX = l(EX) = g(EX)$.

Analogicky pro g konkávní. $\blacksquare$

Pozn. $x \in \mathbb{C}: |x| \leq \frac{1}{2} \Rightarrow \frac{1}{2}|x| \leq |\exp\{x\} - 1| \leq \frac{3}{2}|x|$

Dk. $\exp\{x\} = \sum_{n=1}^{\infty} \frac{x^n}{n!} \Rightarrow |e^x - 1 - x| = \left|\sum_{n=2}^{\infty} \frac{x^n}{n!}\right| \leq \sum_{n=2}^{\infty} \left|\frac{x^n}{n!}\right| \leq$

$$\leq \sum_{n=2}^{\infty} \frac{|x|^n}{2} = \frac{1}{2} \frac{|x|^2}{1-|x|} \leq \frac{1}{2}|x| \quad (\Leftarrow |x| \leq \tfrac{1}{2}).$$

Δ-ner. $\Rightarrow \frac{1}{2}|x| \leq |e^x - 1| \leq \frac{3}{2}|x| \quad \blacksquare$