# NMSA202 – PST & MAT STAT – KAPITOLY 5–6

## Dk V Implikace mezi typy konvergence

(a) Nechť $X_n \xrightarrow{a.s.} X$, $n\to\infty$. Pro $n\in\mathbb{N}$ definujme náhodné události

$$A_n := \{\omega\in\Omega : \exists m \geq n : |X_m(\omega)-X(\omega)| \geq \varepsilon\}$$

$$B_n := \{\omega\in\Omega : |X_n(\omega)-X(\omega)| \geq \varepsilon\}$$

Chceme ukázat, že $P(B_n)\to 0$. Víme, že $A_n \supseteq B_n$. Takže z monotonie psti stačí ukázat, že $P(A_n)\longrightarrow 0$.

- Reálná posloupnost $\{X_n(\omega)\}_n$ je konvergentní když $\exists N\in\mathbb{N}$ takové, že pro žádné $m\geq N$: $|X_m(\omega)-X(\omega)|\geq\varepsilon$. Potom $\lim\limits_{n\to\infty} A_n$ je událost, kdy $X_n$ diverguje. Dle konvergence skoro jistě dostáváme $P(\lim\limits_{n\to\infty} A_n)=0$.
- Platí, že $A_1\supseteq A_2\supseteq A_3\supseteq \dots$, protože když $|X_m(\omega)-X(\omega)|\geq\varepsilon$, potom $\omega$ patří do každé $A_n$, $n\leq m$, ale ne nutně $n>m$. Pak ze spojitosti psti máme $\lim\limits_{n\to\infty} P(A_n) = P(\lim\limits_{n\to\infty} A_n)$.

Pak ale dostaneme konvergenci v psti, protože

$$\lim_{n\to\infty} P(B_n) \leq \lim_{n\to\infty} (A_n) = P(\lim_{n\to\infty} A_n) = 0.$$

(b) Nechť $X_n \xrightarrow{L_p} X$, $n\to\infty$, $p\geq 1$. Dle Markovovy nerovnosti

$$P(|X_n-X|\geq\varepsilon) = P(|X_n-X|^p\geq\varepsilon^p) \leq \frac{E[|X_n-X|^p]}{\varepsilon^p}\longrightarrow 0.$$

(c) Nechť $X_n \xrightarrow{L_p} X$, $n\to\infty$, $p\geq q\geq 1$. Dle Jensenovy nerovnosti pro konvexní fci $g(x) := x^{p/q}$, $x\geq 0$ dostáváme

$$g(E[|X_n-X|^q]) \leq E[g(|X_n-X|^q)] = E[(|X_n-X|^q)^{p/q}] = E[|X_n-X|^p]\to 0.$$

(d) Nechť $X_n \xrightarrow{P} X$, $n\to\infty$. Zvolme libovolné pevné $\varepsilon>0$. Nechť $x\in\mathbb{R}$ je libovolný bod spojitosti $F$. Potom

$$\begin{aligned}F_n(x) = P(X_n\leq x) &= P(X_n\leq x, X\leq x+\varepsilon) + P(X_n\leq x, X>x+\varepsilon)\\ &\leq P(X\leq x+\varepsilon) + P(|X_n-X|>\varepsilon) = F(x+\varepsilon) + P(|X_n-X|>\varepsilon).\end{aligned}$$

Taktéž dostaneme

$$F(x-\varepsilon)=P(X\leq x-\varepsilon)=P(X\leq x-\varepsilon, X_n\leq x)+P(X\leq x-\varepsilon, X_n>\varepsilon)$$
$$\leq F_n(x)+P(|X_n-X|>\varepsilon).$$

Potom

$$F(x-\varepsilon)-P(|X_n-X|>\varepsilon)\leq F_n(x)\leq F(x+\varepsilon)+P(|X_n-X|>\varepsilon).$$

Pro limitní přechod $n\to\infty$ máme

$$F(x-\varepsilon)\leq \liminf_{n\to\infty} F_n(x)\leq \limsup_{n\to\infty} F_n(x)\leq F(x+\varepsilon).$$

Jelikož $\varepsilon>0$ bylo libovolné a $F$ je spojitá v $x$, pak dostáváme $\lim_{n\to\infty} F_n(x)=F(x)$.

(e) Nechť $X_n\xrightarrow{\mathcal{D}}X$, $n\to\infty$ a $P(X=c)=1$ pro nějaké $c\in\mathbb{R}$.

Zvolme libovolné $\varepsilon>0$. Potom

$$P(|X_n-c|\geq\varepsilon)=P(X_n\leq c-\varepsilon)+P(X_n\geq c+\varepsilon)$$
$$\leq P(X_n\leq c-\varepsilon)+P(X_n>c+\tfrac{\varepsilon}{2})$$
$$=F_n(c-\varepsilon)+1-F_n(c+\tfrac{\varepsilon}{2})$$
$$\xrightarrow{n\to\infty}F(c-\varepsilon)+1-F(c+\tfrac{\varepsilon}{2})=0+1-1=0. \qquad \blacksquare$$

## Dk V O spojitém zobrazení (CMT)

- Nechť $X_n\xrightarrow{a.s.}X$, $n\to\infty$. Z definice spojitosti funkce $g$ platí $\lim_{n\to\infty}X_n(\omega)=X(\omega)\Rightarrow\lim_{n\to\infty}g(X_n(\omega))=g(X(\omega))$.

  Potom
  $$P[\omega\in\Omega:\lim_{n\to\infty}g(X_n(\omega))=g(X(\omega))]=P[\omega\in\Omega:\lim_{n\to\infty}X_n(\omega)=X(\omega)]=1.$$

- Nechť $X_n\xrightarrow{P}X$, $n\to\infty$. Zvolme libovolné pevné $\varepsilon>0$. Pak pro libovolné $\delta>0$ uvažujme množinu

  $$B_\delta:=\{x\in\mathbb{R}:\exists y\in\mathbb{R}: |x-y|<\delta \,\&\, |g(x)-g(y)|\geq\varepsilon\}.$$

  Zřejmě $B_\delta\downarrow\emptyset$, $\delta\downarrow 0+$. Potom

  $$P(|g(X_n)-g(X)|\geq\varepsilon)=P(|g(X_n)-g(X)|\geq\varepsilon\cap|X_n-X|\geq\delta)+$$
  $$+P(|g(X_n)-g(X)|\geq\varepsilon\cap|X_n-X|<\delta)\leq$$

$\leq P(|X_n - X| \geq \delta) + P(X \in B_\delta)$. Jelikož $\delta > 0$ bylo libovolné, potom $P(X \in B_\delta) \to 0$, když $\delta \downarrow 0+$. Navíc pro každé pevné $\delta > 0$ platí $P(|X_n - X| \geq \delta) \to 0$, když $n \to \infty$. ∎

- Bez dk: $X_n \xrightarrow{\mathcal{D}} X \Rightarrow g(X_n) \xrightarrow{\mathcal{D}} g(X)$.

## Dk Slutského věta

Důkaz provedeme pouze pro součet; pro součin je analogický. Vezmeme libovolné $x \in \mathbb{R}$, které je bodem spojitosti distribuční funkce náhodné veličiny $X + c$. Tedy bod $x - c$ je bodem spojitosti $F_X$. Zvolme libovolné $\eta > 0$. Potom existuje $\varepsilon_0 > 0$: $|F_X(x-c) - F_X(x-c-\varepsilon)| < \eta/3$, $\forall \varepsilon: |\varepsilon| < \varepsilon_0$. Jelikož $F_X$ je distribuční funkce, pak má spočetně mnoho bodů nespojitosti. (Nechť $D$ je mna bodů nespojitosti $F_X$. $\forall y \in D: F_X(y-) < F_X(y+)$. Potom $\exists q_y \in \mathbb{Q}: F_X(y-) < q_y < F_X(y+)$. Jelikož $F_X$ je neklesající, pak $y \neq x \Rightarrow q_y \neq q_x$. Takže zobrazení $y \mapsto q_y : D \xleftrightarrow{1-1} \mathbb{Q}$.) Tedy v každém okolí bodu $x - c$ najdeme vlevo i vpravo od $x - c$ bod spojitosti $F_X$. Zvolme tedy $0 < \varepsilon < \varepsilon_0$ tak, že $F_X$ je spojitá v bodech $x - c - \varepsilon$ i v $x - c + \varepsilon$. Potom

$$P(X_n + Y_n \leq x) = P(X_n + Y_n \leq x \cap |Y_n - c| < \varepsilon) + P(X_n + Y_n \leq x \cap |Y_n - c| \geq \varepsilon) \leq$$

$$\leq P(X_n + c \leq x + \varepsilon \cap |Y_n - c| < \varepsilon) + P(|Y_n - c| \geq \varepsilon) \leq$$

$$\leq P(X_n \leq x - c + \varepsilon) + P(|Y_n - c| \geq \varepsilon).$$

Protože $F_X$ je spojitá v $x - c + \varepsilon$ a $Y_n \xrightarrow{P} c$, existuje $n_1 \in \mathbb{N}: \forall n \geq n_1$: $P(X_n \leq x - c + \varepsilon) \leq P(X \leq x - c + \varepsilon) + \eta/3$ & $P(|Y_n - c| \geq \varepsilon) \leq \eta/3$. Dohromady pro $n \geq n_1$ máme

$$P(X_n + Y_n \leq x) \leq P(X \leq x - c + \varepsilon) + 2\eta/3 \leq P(X + c \leq x) + \eta.$$

Opačnou nerovnost dokážeme obdobně. Nejdřív si uvědomíme

$$P(X_n + c \leq x - \varepsilon) = P(X_n + c \leq x - \varepsilon \cap |Y_n - c| < \varepsilon) + P(X_n + c \leq x - \varepsilon \cap |Y_n - c| \geq \varepsilon) \leq$$

$\leq P(X_n+Y_n\leq x)+P(|Y_n-c|\geq \varepsilon)$. Díky předpokladům věty a volbě $n_2\in\mathbb{N}: \forall n\geq n_2: P(X_n+c\leq x-\varepsilon)\geq P(X+c\leq x-\varepsilon)-\eta/3$ & $P(|Y_n-c|\geq\varepsilon)\leq \eta/3$. Pro $n\geq n_2$ tedy platí

$$P(X+c\leq x)\leq P(X+c\leq x-\varepsilon)+\eta/3\leq P(X_n+c\leq x-\varepsilon)+2\eta/3\leq$$
$$\leq P(X_n+Y_n\leq x)+P(|Y_n-c|\geq\varepsilon)+2\eta/3\leq P(X_n+Y_n\leq x)+\eta.$$

Pak pro $n\geq \max\{n_1,n_2\}$ dohromady dostáváme

$$P(X+c\leq x)-\eta\leq P(X_n+Y_n\leq x)\leq P(X+c\leq x)+\eta.$$

Protože $\eta>0$ bylo libovolné, je věta dokázána. ∎

Dk D Slutského věty

Jelikož $X_n\xrightarrow{P} a\in\mathbb{R}$, potom $X_n\xrightarrow{\mathcal{D}} a\in\mathbb{R}$. Potom dle Slutského věty $X_n+Y_n\xrightarrow{\mathcal{D}} a+b\in\mathbb{R}$. Ale věta o implikaci konvergencí (e) dává $X_n+Y_n\xrightarrow{P} a+b$. Důkaz pro součin je analogický. ∎

Dk Lévyho věta o spojitosti

Bez dk.

Př $X\sim N(0,1)\Rightarrow \varphi_X(t)=\exp\{-t^2/2\},\ t\in\mathbb{R}$

$$\varphi_X(t)=E[\exp\{itX\}]=E[\cos(tX)]+i\,E[\sin(tX)]=$$
$$=\int_{-\infty}^{+\infty}\cos(tx)\,dP_X(x)+\underbrace{i\int_{-\infty}^{+\infty}\sin(tx)\,dP_X(x)}_{=0}=E[\cos(tX)]$$

↑ lichá

majoranta

$$\frac{d}{dt}\varphi_X(t)=-\int_{-\infty}^{+\infty}x\sin(tx)\,dP_X(x)=-\int_{-\infty}^{+\infty}x\sin(tx)f_X(x)\,dx$$

$\left\{ f_X(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\ x\in\mathbb{R}\right.$

$$\overset{P.P.}{=}\underbrace{\Big[\sin(tx)f_X(x)\Big]_{-\infty}^{+\infty}}_{=0}-\int_{-\infty}^{+\infty}t\cos(tx)f_X(x)\,dx=-t\,E[\cos(tX)]$$

$$\Rightarrow \frac{d}{dt}\varphi_X(t)=-t\varphi_X(t)\ \&\ \varphi_X(0)=E[\cos(0)]=1\Rightarrow \varphi_X(t)=\exp\{-\tfrac{t^2}{2}\}.$$

Př $X\sim N(\mu,\sigma^2)\Rightarrow\varphi_X(t)=\exp\{i\mu t-\frac{1}{2}\sigma^2t^2\},\ t\in\mathbb{R}$

$Z:=\frac{X-\mu}{\sigma}\sim N(0,1)\Rightarrow\varphi_Z(t)=\exp\{-\frac{t^2}{2}\}$

$$\varphi_X(t)=E[\exp\{itX\}]=E[\exp\{it(\mu+\sigma Z)\}]=\exp\{it\mu\}E[\exp\{i(t\sigma)Z\}]=$$
$$=\exp\{it\mu\}\varphi_Z(t\sigma)=\exp\{i\mu t-\tfrac{1}{2}\sigma^2t^2\}.$$

## Dk Slabý zákon velkých čísel

Rozvineme komplexní exponenciálu pomocí Taylora 1. řádu

$$\exp\{i t y\} = 1 + i t y + o(t), \quad t \to 0, \quad \forall y \in \mathbb{R}.$$

Potom $E[\exp\{i t Y\}] = 1 + i t EY + o(t), \; t \to 0.$

Pak počítejme

$$\varphi_{\overline{X}_n}(t) = \varphi_{\frac{1}{n}\sum X_i}(t) = \varphi_{\sum \frac{X_i}{n}}(t) \underset{\substack{\uparrow \\ CF(viii)}}{\overset{IID}{=}} \left\{ \varphi_{\frac{X_1}{n}}(t) \right\}^n \underset{\substack{\uparrow \\ CF(v)}}{=} \left\{ \varphi_{X_1}\left(\frac{t}{n}\right) \right\}^n =$$

$$= \left\{ E\left[\exp\{i t X_1 / n\}\right] \right\}^n \underset{\substack{\uparrow \\ \text{Taylor 1.ř.}}}{=} \left\{ 1 + \frac{it}{n} EX_1 + o\left(\frac{t}{n}\right) \right\}^n, \quad \forall t \in \mathbb{R}, \; \frac{t}{n} \to 0, n \to \infty.$$

$$\lim_{n\to\infty} \varphi_{\overline{X}_n}(t) = \lim_{n\to\infty} \left\{ 1 + \frac{it}{n} EX_1 + o\left(\frac{t}{n}\right) \right\}^n = \left\{ E\left[\exp\{i t EX_1 / n\}\right] \right\}^n =$$

$$= E\left[\exp\{i t EX_1\}\right] = \varphi_{EX_1}(t), \quad \forall t \in \mathbb{R} \; (\text{bodově}).$$

Dle Lévyho věty o spojitosti dostáváme $\overline{X}_n \xrightarrow[n\to\infty]{\mathcal{D}} EX_1 \in \mathbb{R}$.

Jelikož je limita v distribuci konstantou n.j., potom

$\overline{X}_n \xrightarrow[n\to\infty]{P} EX_1$ $\left( \xrightarrow{\mathcal{D}} \text{ \& konst. det. lim.} \Rightarrow \xrightarrow{P} \right)$. $\blacksquare$

## Dk Centrální limitní věta

Rozvineme komplexní exponenciálu pomocí Taylora 2. řádu

$$\exp\{i t y\} = 1 + i t y + (i t y)^2/2 + o(t^2), \quad t \to 0, \quad \forall y \in \mathbb{R}.$$

Potom $E[\exp\{i t Y\}] = 1 + i t EY - \frac{t^2}{2} EY^2 + o(t^2), \; t \to 0.$

Definujme $Y_n := (X_n - EX_1) / \sqrt{\operatorname{var} X_1}$. Pak $EY_n = 0$, $\operatorname{var} Y_n = 1$,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i = \sqrt{n} \underbrace{\left( \frac{1}{n} \sum_{i=1}^n Y_i \right)}_{=: \overline{Y}_n} = \sqrt{n}\, \overline{Y}_n = \sqrt{n} \frac{\overline{X}_n - EX_1}{\sqrt{\operatorname{var} X_1}} = Z_n, \; \forall n.$$

Počítejme

$$\varphi_{Z_n}(t) = \varphi_{\sum_{i=1}^n \frac{1}{\sqrt{n}} Y_i} \underset{\substack{\uparrow \\ CF(viii)}}{\overset{IID}{=}} \left\{ \varphi_{\frac{Y_1}{\sqrt{n}}}(t) \right\}^n \underset{\substack{\uparrow \\ CF(v)}}{=} \left\{ \varphi_{Y_1}\left( \frac{t}{\sqrt{n}} \right) \right\}^n = \left\{ E\left[ \exp\left\{ i \frac{t}{\sqrt{n}} Y_1 \right\} \right] \right\}^n =$$

$$\underset{\substack{\uparrow \\ \text{Taylor 2.ř.}}}{=} \left\{ 1 + \frac{it}{\sqrt{n}} \underbrace{EY_1}_{=0} - \frac{t^2}{2n} \underbrace{EY_1^2}_{=\operatorname{var} Y_1 = 1} + o\left( \frac{t^2}{n} \right) \right\}^n, \quad \forall t \in \mathbb{R}, \; \frac{t^2}{n} \to 0, n \to \infty.$$

Jelikož $e^y = \lim_{n\to\infty} \left(1+\frac{y}{n}\right)^n \; \forall y \in \mathbb{R}$, dostáváme

$$\lim_{n\to\infty} \varphi_{Z_n}(t) = \lim_{n\to\infty} \left\{1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right\}^n = \exp\left\{-\frac{1}{2}t^2\right\} = \varphi_Z(t), \forall t \in \mathbb{R}$$
$Z \sim N(0,1)$.

Dle Lévyho věty o spojitosti dostáváme $Z_n \xrightarrow[n\to\infty]{\mathcal{D}} Z$, $n \to \infty$. ∎

## Dk Cramér-Woldova věta

Důsledek Lévyho věty o spojitosti.

## Dk Mnohorozměrné CLT

Důsledek Cramér-Woldovy věty.

## Dk Delta metody

Podle věty o střední hodnotě (tj. aproximace Taylorovým rozvojem do 1. řádu) dostáváme

(*) $g(Y_n) = g(\mu) + g'(\tilde{\mu})(Y_n - \mu)$, kde $\tilde{\mu}$ leží mezi $Y_n$ a $\mu$.

Když $\sqrt{n}(Y_n - \mu) \xrightarrow[n\to\infty]{\mathcal{D}} N(0,\sigma^2)$, potom $Y_n - \mu \xrightarrow[n\to 0]{\mathcal{D}} 0$. Pak tedy $Y_n - \mu \xrightarrow[n\to\infty]{P} 0$. Jelikož $|\tilde{\mu} - \mu| \leq |Y_n - \mu|$, musí platit, že $\tilde{\mu} - \mu \xrightarrow[n\to\infty]{P} 0$. Dle věty o spojitém zobrazení dostáváme $g'(\tilde{\mu}) \xrightarrow[n\to\infty]{P} g(\mu)$ ($g'(\cdot)$ je totiž spojité dle předpokladu).

Z (*) máme $\sqrt{n}\left[g(Y_n) - g(\mu)\right] = g'(\tilde{\mu})\sqrt{n}(Y_n - \mu)$.

Jelikož (dle předpokladu) $\sqrt{n}(Y_n - \mu) \xrightarrow[n\to\infty]{\mathcal{D}} N(0,\sigma^2)$, potom s využitím Slutského věty máme

$$\sqrt{n}\left[g(Y_n) - g(\mu)\right] \xrightarrow[n\to\infty]{\mathcal{D}} N\left(0, [g'(\mu)]^2\sigma^2\right). \blacksquare$$

## Dk Rozklad MSE = bias² + var

Z definice počítáme

$$\mathrm{MSE}(\hat{\theta}_n) = E_\theta(\hat{\theta}_n - \theta)^2 = E_\theta\left[\hat{\theta}_n - E_\theta(\hat{\theta}_n) + E_\theta(\hat{\theta}_n) - \theta\right]^2 =$$
$$= \underbrace{E_\theta\left[\hat{\theta}_n - E_\theta(\hat{\theta}_n)\right]^2}_{=\mathrm{var}_\theta \hat{\theta}_n} + \underbrace{\left[E_\theta(\hat{\theta}_n) - \theta\right]}_{\text{deterministické}}\underbrace{\left\{E_\theta\left[\hat{\theta}_n - E_\theta(\hat{\theta}_n)\right]\right\}}_{=0} + \underbrace{\left[E_\theta(\hat{\theta}_n) - \theta\right]^2}_{=\mathrm{bias}^2 \hat{\theta}_n}.$$

∎

## Dk V Postačující podmínka pro konzistenci,

Podle předchozí věty máme

$$\underset{\substack{\text{def}\ \| \\ E_\theta[\hat{\theta}_n - \theta]^2 \\ \downarrow n\to\infty \\ 0}}{MSE(\hat{\theta}_n)} = \underbrace{\text{bias}^2\theta_n}_{\to 0} + \underbrace{\text{var}\,\hat{\theta}_n}_{\to 0}$$

To je ale definice konvergence v $L_2$, tedy

$$\hat{\theta}_n \xrightarrow[n\to\infty]{L_2} \theta \implies \hat{\theta}_n \xrightarrow[n\to\infty]{P} \theta$$ (tj. konzistence $\hat{\theta}_n$). ∎

(implikace mezi typy konvergence)

## Dk CI založený na normalitě,

Podle předpokladů věty máme:

- $\dfrac{\hat{\theta}_n - \theta}{se(\hat{\theta}_n)} \xrightarrow[n\to\infty]{\mathcal{D}} N(0,1)$ (asymptoticky standardně normální),
- $\hat{se}(\hat{\theta}_n) - se(\hat{\theta}_n) \xrightarrow[n\to\infty]{P} 0$ (konzistentní odhad se).

Ze Sluckého věty plyne

$$\frac{\hat{\theta}_n - \theta}{\hat{se}(\hat{\theta}_n)} \xrightarrow[n\to\infty]{\mathcal{D}} N(0,1),$$ což implikuje

$$1-\alpha = \lim_{n\to\infty} P_\theta\left[-u_{1-\frac{\alpha}{2}} \leq \frac{\hat{\theta}_n - \theta}{\hat{se}(\hat{\theta}_n)} \leq u_{1-\frac{\alpha}{2}}\right] =$$

$$= \lim_{n\to\infty} P_\theta\left[\hat{\theta}_n - u_{1-\frac{\alpha}{2}}\hat{se}(\hat{\theta}_n) \leq \theta \leq \hat{\theta}_n + u_{1-\frac{\alpha}{2}}\hat{se}(\hat{\theta}_n)\right] =$$

$$= \lim_{n\to\infty} P_\theta[\theta \in C_n].$$ ∎