NMSA202 – PRAVDĚPODOBNOST A MATEMATICKÁ STATISTIKA

Dk L Pst sjednocení, $\forall A, B \in \mathcal{A}: P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Rozepišme $A \cup B = (A \cap B^c) \cup (A \cap B) \cup (A^c \cap B)$

disjunktní události

aditivita psti

$$P(A \cup B) = P(A \cap B^c) + P(A \cap B) + P(A^c \cap B)$$
$$= \underbrace{P(A \cap B^c) + P(A \cap B)} + P(A^c \cap B) + P(A \cap B) - P(A \cap B)$$
$$= P((A \cap B^c) \cup (A \cap B)) + P((A^c \cap B) \cup (A \cap B)) - P(A \cap B)$$
$$= P(A) + P(B) - P(A \cap B)$$

■

Dk V Spojitost psti, $A_n \uparrow A$ nebo $A_n \downarrow A \Rightarrow P(A_n) \xrightarrow[n \to \infty]{} P(A)$

a) $A_n \uparrow A \overset{def}{\Leftrightarrow} A_1 \subseteq A_2 \subseteq \ldots \;\&\; A = \bigcup_{i=1}^{\infty} A_i$; b) $A_n \downarrow A \overset{def}{\Leftrightarrow} A_1 \supseteq A_2 \supseteq \ldots \;\&\; A = \bigcap_{i=1}^{\infty} A_i$

a) definujme $B_1 := A_1$, $B_2 := \{\omega \in \Omega : \omega \in A_2, \omega \notin A_1\}$

$B_3 := \{\omega \in \Omega : \omega \in A_3, \omega \notin A_2, \omega \notin A_1\}$, $B_4 := \ldots$

$\Rightarrow B_1, B_2, \ldots$ jsou disjunktní a $A_n = \bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n} B_i \quad \forall n \in \mathbb{N}$

a $\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i$ (viz $\supseteq$)

Pak $P(A_n) = P\left(\bigcup_{i=1}^{n} B_i\right) = \sum_{i=1}^{n} P(B_i)$ a následně

$$\lim_{n \to \infty} P(A_n) = \lim_{n \to \infty} \sum_{i=1}^{n} P(B_i) = \sum_{i=1}^{\infty} P(B_i) = P\left(\bigcup_{i=1}^{\infty} B_i\right) = P(A)$$

■

b) analogicky

Dk V Zákon úplné psti, $\Omega = \bigcup_{i=1}^{\infty} A_i$, $A_i \cap A_j = \emptyset \; \forall i \neq j$, $P(A_i) > 0 \; \forall i$

$$\Rightarrow P(B) = \sum_{i=1}^{\infty} P(B|A_i) P(A_i)$$

Definujme $C_i := B \cap A_i \quad \forall i$

Nahlédneme, že $C_1, \ldots, C_k$ jsou disjunktní a $B = \bigcup_{i=1}^{k} C_i \quad \forall k$

Potom $P(B) = \sum_i P(C_i) = \sum_i P(B \cap A_i) = \sum_i P(B|A_i) P(A_i)$ z definice podm. psti ■

Dk Bayesovy věty: $P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{\infty} P(B|A_j)\,P(B_j)}$

Počítejme: $P(A_i|B) \underset{\uparrow}{=} \frac{P(A_i \cap B)}{\underbrace{P(B)}_{>0}} \underset{\uparrow}{=} \frac{P(B|A_i)P(A_i)}{P(B)} \underset{\uparrow}{=} \frac{P(B|A_i)P(A_i)}{\sum_j P(B|A_j)P(A_j)}$ ∎

def podm. psti ⟶ ; z o úplné psti

Dk O přenosu integrace: $(\Omega, \mathcal{A}, P) \xrightarrow{X} (M, \mathcal{M}) \xrightarrow{g}$

$P_X(M) = P[X^{-1} \in M],\ M \in \mathcal{M}$

$$\Rightarrow \int_\Omega g[X(\omega)]\,dP(\omega) = \int_M g(x)\,dP_X(x) \quad (*)$$

1) Nejprve uvažujme char. fci $g = 1_B,\ B \in \mathcal{M}$

tedy $1_B[X(\omega)] = 1$ pro $X(\omega) \in B$ a to jest pro $\omega \in X^{-1}(B)$

$LS(*)$: $\int_\Omega 1_B[X(\omega)]\,dP(\omega) = \int_{X^{-1}(B)} dP(\omega) = P[X^{-1}(B)]$

$\qquad\quad\ \equiv$

$PS(*)$: $\int_M 1_B(x)\,dP_X(x) = \int_B dP_X(x) = P_X(B) = P[X^{-1}(B)]$

Tím je $(*)$ pro měřitelnou charakteristickou fci dokázán.

2) Nechť g je jednoduchá měřitelná fce, tj. $g(\cdot) = \sum_{k=1}^{m} c_k 1_{B_k}(\cdot)$, $c_k \in \mathbb{R}$, $B_k \in \mathcal{M},\ \forall k$

z linearity integrálu plyne:

$LS(*)$: $\int_\Omega \sum_{k=1}^{m} c_k 1_{B_k}[X(\omega)]\,dP(\omega) = \sum_{k=1}^{m} c_k \int_{X^{-1}(B_k)} dP(\omega) = \sum_{k=1}^{m} c_k P[X^{-1}(B_k)]$

$\qquad\equiv$

$PS(*)$: $\int_M \sum_{k=1}^{m} c_k 1_{B_k}(x)\,dP_X(x) = \sum_{k=1}^{m} c_k \int_{B_k} dP_X(x) = \sum_{k=1}^{m} c_k P_X(B_k) = \sum_{k=1}^{m} c_k P[X^{-1}(B_k)]$

3) Teď vezmeme $g \geq 0$ měřitelnou. Pak existuje posloupnost $\{g_n\}_{n\in\mathbb{N}}$ nezáporných jednoduchých měřitelných fcí takové, že $g_n(x) \underset{n\to\infty}{\uparrow} g(x)\ \forall x \in M$

Dle Leviho věty o monotónní konvergenci:

$LS(*)$: $\int_\Omega g[X(\omega)]\,dP(\omega) \xleftarrow[n\to\infty]{\text{Levi}} \int_\Omega g_n[X(\omega)]\,dP(\omega) \overset{2)}{=} \int_M g_n(x)\,dP_X(x) \xrightarrow[n\to\infty]{\text{Levi}} \int_M g(x)\,dP(x) = PS(*)$

4) Je-li g reálná měřitelná fce na $(M, \mathcal{M})$, užijeme rozklad $g = g^+ - g^-$, kde g^+, g^- jsou nezáporné měřitelné. Pak z 3) plyne $\int_\Omega g^+[X(\omega)]dP(\omega) = \int_M g^+(x)dP_X(x)$ a analogicky pro g^-. Linearita integrálu potom dává $(*)$. ∎

__Dk Výpočet psti__ (a) $P_X(B) = P[X \in B]$, (b) $f_X = \frac{dP_X}{d\mu}$, μ ... σ-konečná $\Rightarrow P[X \in B] = \int_B f_X d\mu$

a) $P[X \in B] = P[\{\omega \in \Omega : X(\omega) \in B\}] = \int_\Omega 1_B[X(\omega)]dP(\omega) \underset{\uparrow}{=} \int_{\mathbb{R}} 1_B(x) dP_X(x) = P_X(B)$

V o přenosu integrace, kde $(M, \mathcal{M}) \equiv (\mathbb{R}, \mathcal{B}(\mathbb{R}))$

b) $P[X \in B] = \int_\Omega 1_B[X(\omega)]dP(\omega) = \int_{\mathbb{R}} 1_B(x) dP_X(x) \underset{\uparrow}{=} \int_{\mathbb{R}} 1_B(x) f_X(x) d\mu(x) = \int_B f_X(x) d\mu(x)$

Radon–Nikodým ∎

__Dk Tři základní vlastnosti CDF__, (i) neklesající, (ii) normovaná, (iii) zprava spojitá.

(i) Zvolme $x_1 < x_2$. Pak

$$F(x_1) = P[X \le x_1] = P[X \in (-\infty, x_1]] \underset{\uparrow}{\overset{\text{monotonie míry}}{\le}} P[(-\infty, x_2]] = F(x_2)$$

$(-\infty, x_1] \subseteq (-\infty, x_2]$

(ii) $$\lim_{x \downarrow -\infty} F(x) = \lim_{x \downarrow -\infty} P[X \in (-\infty, x]] \underset{\Uparrow}{\overset{\text{spojitost psti}}{=}} P[X \in \emptyset] = 0$$

$(-\infty, x] \downarrow \emptyset$

$$\lim_{x \uparrow +\infty} F(x) = \lim_{x \uparrow +\infty} P[X \in (-\infty, x]] \underset{\Uparrow}{\overset{\text{spojitost psti}}{=}} P[X \in \mathbb{R}] = 1$$

$(-\infty, x] \uparrow \mathbb{R}$

(iii) $$\lim_{y \downarrow x+} F(y) = \lim_{y \downarrow x+} P[X \in (-\infty, y]] \overset{\text{spojitost psti}}{=} P[X \in (-\infty, x]] = F(x)$$

$(-\infty, y] \underset{y \downarrow x+}{\downarrow} (-\infty, x]$

__Dk L Důsledky definice CDF__

1) $P[X = x] = P[X \in \{x\}] = P[X \in (-\infty, x] \setminus (-\infty, x)] = P[X \in (-\infty, x]] - P[X \in (-\infty, x)] =$
$= F(x) - P[X \in \bigcup_{n \in \mathbb{N}} (-\infty, x - \frac{1}{a_n}]] \overset{\text{Heine}}{=} F(x) - \lim_{y \uparrow x-} P[X \in (-\infty, y]] = F(x) - F(x^-)$

lib. $a_n \uparrow \infty$

2) $P[x < X \le y] = P[X \in (-\infty, y] \setminus (-\infty, x]] = P[X \in (-\infty, y]] - P[X \in (-\infty, x]] =$
$= F(y) - F(x)$

3) $P[X > x] = P[X \in (x, +\infty)] = 1 - P[X \in (-\infty, x]] = 1 - F(x)$

4) X spojitá $\Rightarrow P[X = x] = \int_{\{x\}} f(\lambda) d\lambda(\lambda) = 0 \quad \forall x \in \mathbb{R} \Rightarrow P[X \le x] = P[X < x]$ ∎

Lebesgue

Dk V Vlastnosti sdružené CDF

(i) F je po složkách neklesající & zprava spojité

(ii) $\lim_{x_\ell \downarrow -\infty} F(x) = 0 \quad \forall \ell = 1,\dots,d$ & $\lim_{x_\ell \uparrow +\infty \, \forall \ell} F(x) = 1$

(i) Nechť $x_1,\dots,x_{\ell-1},x_{\ell+1},\dots,x_d \in \mathbb{R}$ jsou pevné. Definujme fci

$$G(x) := F_X(x_1,\dots,x_{\ell-1},x,x_{\ell+1},\dots,x_d), \quad x \in \mathbb{R}.$$

Z monotonie psti je G fce neklesající ✓ & nezáporná.

Jelikož je G neklesající, musí existovat $\lim_{y\downarrow x+} G(y) \geq G(x)$.

Potřebujeme ale ukázat rovnost.

Víme, že $\lim_{y\downarrow x+} G(y) = \lim_{n\to\infty} G(x+\frac{1}{n})$.

Označme $B_n := (-\infty,x_1]\times\dots\times(-\infty,x_{\ell-1}]\times(-\infty,x+\frac{1}{n})\times(-\infty,x_{\ell+1}]\times\dots\times(-\infty,x_d]$.

Pak $B_n \downarrow B = (-\infty,x_1]\times\dots\times(-\infty,x_{\ell-1}]\times(-\infty,x]\times(-\infty,x_{\ell+1}]\times\dots\times(-\infty,x_d]$.

Potom ze spojitosti míry P_X dostáváme

$$\lim_{y\downarrow x+} G(y) = \lim_{n\to\infty} G(x+\tfrac{1}{n}) = \lim_{n\to\infty} P_X(B_n) = P_X\Big(\bigcap_{n=1}^{\infty} B_n\Big) = G(x),$$

což je zprava spojitost.

(ii) Opět nechť $x_1,\dots,x_{\ell-1},x_{\ell+1},\dots,x_d \in \mathbb{R}$ jsou pevné. Uvažujme opět fci G z (i), která je neklesající a nezáporná. Proto musí existovat nezáporná $\lim_{x\downarrow -\infty} G(x) = \lim_{n\to+\infty} G(-n)$. Vezmeme

$C_n := (-\infty,x_1]\times\dots\times(-\infty,x_{\ell-1}]\times(-\infty,-n]\times(-\infty,x_{\ell+1}]\times\dots\times(-\infty,x_d]$. Platí $C_n \downarrow \emptyset$.

Ze spojitosti míry P_X máme

$$\lim_{x_\ell\downarrow -\infty} F(x) = \lim_{x\downarrow -\infty} G(x) = \lim_{n\to+\infty} G(-n) = \lim_{n\to\infty} P_X\Big(\bigcap_{n=1}^{+\infty} C_n\Big) = P_X(\emptyset) = 0.$$

- Pro druhou část tvrzení (ii) si uvědomíme: $x_\ell \uparrow +\infty \; \forall \ell \iff \min_\ell x_\ell \uparrow +\infty$.

Z monotonie psti máme $1 \geq F_X(x) \geq F_X\big(\min_{1\le \ell\le d}\{x_\ell\}[1,\dots,1]^T\big)$.

Položme $H(x) := F_X(x[1,\dots,1]^T)$, $x \in \mathbb{R}^1$. Z monotonie psti plyne, že H je neklesající. Z normovanosti psti dostáváme, že $H(x) \leq 1$, $\forall x \in \mathbb{R}$.

Pak tedy existuje $\lim_{x\uparrow+\infty} H(x) \leq 1$ a $\lim_{x\uparrow+\infty} H(x) = \lim_{n\to\infty} H(n)$.

Označme $D_n := (-\infty,n]^d$. Potom $D_n \uparrow \mathbb{R}^d$. Ze spojitosti míry P_X máme

$$\lim_{x_\ell\uparrow+\infty\,\forall \ell} F_X(x) = \lim_{n\to\infty} H(n) = \lim_{n\to\infty} P_X(D_n) = P_X\Big(\bigcup_{n=1}^{+\infty} D_n\Big) = P_X(\mathbb{R}^d) = 1. \blacksquare$$

Dk V Marginální distribuční fce

Buď $\{x_n\}_{n=1}^{+\infty}$ libovolná posloupnost taková, že $\lim_{n\to\infty} x_n = +\infty$.

Označme $B := \bigcap_{\ell=1}^{d-1} \{X_\ell \le x_\ell\}$, $B_n := \left(\bigcap_{\ell=1}^{d-1} \{X_\ell \le x_\ell\}\right) \cap \{X_d \le x_n\}$,

$D_n := \left(\bigcup_{m=n}^{\infty} B_m^C\right)^C$ $\forall n \in \mathbb{N}$.

Platí $D_n \subseteq B_n \subseteq B = \bigcup_{n=1}^{\infty} B_n$ a $D_n \uparrow B$, $n \to \infty$.

Ze spojitosti míry P tedy máme $\lim_{n\to\infty} P(D_n) = P(B)$.

Z monotonie míry P dostáváme $P(D_n) \le P(B_n) \le P(B)$.

Potom $\lim_{n\to\infty} P(B_n) = P(B)$. Pak z Heineho věty o spojitosti dostáváme $\lim_{x_d \to \infty} P(B \cap \{X_d \le x_d\}) = P(B)$. ■

Dk V Hustota vzhledem k součinové referenční míře

- Z Radonovy-Nikodýmovy věty dostáváme existenci a tvrzení o f_X.
- Z Fubiniovy věty a předchozí věty o marginální distribuční funkci dostáváme existenci a vlastnosti f_{X_ℓ}. ■

Dk D Marginální rozdělení pro dvourozměrné náhodné vektory

Plyne z předchozí věty, kde položíme $d=2$, $[X_1,\dots,X_d]^T = [X,Y]$ a $\mu_\ell \equiv \delta_{(\cdot)}$ $\forall \ell$ pro diskrétní případ nebo $\mu_\ell \equiv \lambda$ $\forall \ell$ pro spojitý případ. ■

Dk V Ekvivalentní charakterizace nezávislosti

"$\Rightarrow$" Z definice nezávislosti $\{X_\ell\}_{\ell=1}^{d}$, linearity integrálu a Fubiniovy věty plyne, $\forall x = [x_1,\dots,x_d]^T \in \mathbb{R}^d$,

$$F_X(x) \overset{\text{def II}}{=} \prod_{\ell=1}^{d} F_{X_\ell}(x_\ell) \underbrace{=}_{\text{V o hustotě}} \prod_{\ell=1}^{d} \int_{-\infty}^{x_\ell} f_{X_\ell}(t_\ell)\, d\mu_\ell(t_\ell) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_d} \prod_{\ell=1}^{d} f_{X_\ell}(t_\ell)\, d\mu_d(t_d) \dots d\mu_1(t_1) =$$

(pro $\mu_1 \otimes \dots \otimes \mu_d = \lambda^d$ nebo ν (= čítací míra na $\mathbb{R}^d$))

$$= \int_{(-\infty, x]} \prod_{\ell=1}^{d} f_{X_\ell}(t_\ell)\, d(\mu_1 \otimes \dots \otimes \mu_d)(t_1,\dots,t_d).$$

Pak nutně $f_X(t_1,\dots,t_d) = \prod_{\ell=1}^{d} f_{X_\ell}(t_\ell)$ pro $\mu_1 \otimes \dots \otimes \mu_d$-s.v. $[t_1,\dots,t_d] \in \mathbb{R}^d$.

"$\Leftarrow$" Opačnou implikaci dokážeme obráceným postupem. ■

Dk D Kartézský součin a nezávislost,

Definujme $\bar{g}(x) := \dfrac{g(x)}{\int g(t)\,d\mu(t)}$. Analogicky $\bar{h}(x)$. Pak lehce nahlédneme, že $\bar{g}$ a $\bar{h}$ jsou hustoty/pstní fce z předchozí věty. ∎

Dk V Alternativní charakterizace nezávislosti,

$\perp\!\!\!\perp \{X_1,\dots,X_d\} \iff P_{\mathbf{X}} = \bigotimes_{\ell=1}^{d} P_{X_\ell}$

"$\Leftarrow$" Pokud pro $\forall$ mny $B_\ell \in \mathcal{B}(\mathbb{R})$, $\ell \in \{1,\dots,d\}$ platí $P_{\mathbf{X}}\left(\bigtimes_{\ell=1}^{d} B_\ell\right) = \prod_{\ell=1}^{d} P_{X_\ell}(B_\ell)$, pak vezmeme $B_\ell = (-\infty, x_\ell]$ a počítáme

$$F_{\mathbf{X}}(\mathbf{x}) = P_{\mathbf{X}}\left(\bigtimes_{\ell=1}^{d}(-\infty, x_\ell]\right) = \prod_{\ell=1}^{d} P_{X_\ell}((-\infty, x_\ell]) = \prod_{\ell=1}^{d} F_{X_\ell}(x_\ell) \quad \forall \mathbf{x} = [x_1,\dots,x_d]^T \in \mathbb{R}$$

"$\Rightarrow$" Pokud platí $F_{\mathbf{X}}(\mathbf{x}) = \prod_{\ell=1}^{d} F_{X_\ell}(x_\ell)$, pak $P_{\mathbf{X}} = \bigotimes_{\ell=1}^{d} P_{X_\ell}$ na množinách ze systému $\mathcal{D} = \left\{\bigtimes_{\ell=1}^{d}(-\infty, x_\ell] : x_\ell \in \mathbb{R}\ \forall \ell = \{1,\dots,d\}\right\}$. Ale systém $\mathcal{D}$ je uzavřený na průniky a generuje celou $\mathcal{B}(\mathbb{R}^d)$. Protože je $P_{\mathbf{X}}(\mathbb{R}^d) = 1$ a tedy konečná, dostáváme z věty o jednoznačnosti míry rovnost obou měr (tj. $P_{\mathbf{X}}$ a $\bigotimes_{\ell=1}^{d} P_{X_\ell}$) na celém $\mathcal{B}(\mathbb{R}^d)$. ∎