
NMFM301 – Statistika pro finanční matematiky

Pravděpodobnostní rozdělení a pořádkové statistiky

Podrobné riešenie príkladov z 1. cvičenia

A Vzorové příklady s podrobným řešením

A1. Nechť X_i má distribuční funkci F pro všechna $i = 1, \dots, n$.

- Určete distribuční funkce $X_{(1)}$ a $X_{(n)}$.

Řešení:

Nech $F_{(1)}(x)$ je distribuční funkce $X_{(1)}$. Potom platí, že:

$$\begin{aligned} F_{(1)}(x) &= P[X_{(1)} \leq x] = 1 - P[X_{(1)} > x] = 1 - P[X_1 > x, \dots, X_n > x] \\ &\stackrel{(1)}{=} 1 - \prod_{i=1}^n P[X_i > x] \stackrel{(2)}{=} 1 - \prod_{i=1}^n [1 - F(x)] = 1 - (1 - F(x))^n, \end{aligned}$$

kde rovnost (1) plyne z nezávislosti náhodných veličin X_1 až X_n a rovnost (2) z definice distribuční funkce $F(x) = P[X_i \leq x]$ a faktu, že náhodné veličiny X_1, \dots, X_n sú rovnako rozdelené (pretože sa jedná o náhodný výber).

Nech $F_{(n)}(x)$ je distribuční funkce $X_{(n)}$. Potom analogicky platí, že:

$$F_{(n)}(x) = P[X_{(n)} \leq x] = P[X_1 \leq x, \dots, X_n \leq x] = \prod_{i=1}^n P[X_i \leq x] = [F(x)]^n,$$

kde sa v posledných dvoch rovnostiach opäť využila vlastnosť nezávislosti a rovnakého rozdelenia náhodných veličin X_1, \dots, X_n .

- Nechť X_i má hustotu f vzhledem k Lebesguově míře. Najděte hustoty $X_{(1)}$ a $X_{(n)}$.

Řešení:

Keďže X_i má hustotu f vzhledem k Lebesguově míře, tak platí, že $f(x) = F'(x)$ pre všetky $x \in \mathbb{R}$. Zároveň aj pre hustoty $f_{(1)}$ a $f_{(n)}$ náhodných veličin $X_{(1)}$ a $X_{(n)}$ platí, že $f_{(1)}(x) = F'_{(1)}(x)$ a $f_{(n)}(x) = F'_{(n)}(x)$. Preto:

$$\begin{aligned} f_{(1)}(x) &= \frac{\partial}{\partial x} F_{(1)}(x) = \frac{\partial}{\partial x} [1 - (1 - F(x))^n] = n(1 - F(x))^{n-1} \cdot \frac{\partial}{\partial x} F(x) \\ &= n f(x) [1 - F(x)]^{n-1} \quad \forall x \in \mathbb{R}; \\ f_{(n)}(x) &= \frac{\partial}{\partial x} F_{(n)}(x) = \frac{\partial}{\partial x} [F(x)]^n = n f(x) [F(x)]^{n-1} \quad \forall x \in \mathbb{R}; \end{aligned}$$

A2. Nechť X_i má distribuční funkci F pro všechna $i = 1, \dots, n$.

- Určete sdruženou distribuční funkci $X_{(1)}$ a $X_{(n)}$.

Řešení:

Nech $G : \mathbb{R}^2 \rightarrow [0, 1]$ označuje združenou distribuční funkci náhodného vektoru $(X_{(1)}, X_{(n)})^\top$. Potom (z definície združenej distribučnej funkcie) plynie, že

$$G(x, y) = P[X_{(1)} \leq x \wedge X_{(n)} \leq y] = P[\exists i : X_i \leq x \wedge X_1 \leq y, \dots, X_n \leq y]$$

$$= \begin{cases} P[X_1 \leq y, \dots, X_n \leq y] = [F(y)]^n & \text{pre } x \geq y; \\ P[\{X_1 \leq x \cup \dots \cup X_n \leq x\} \cap \{X_1 \leq y\} \cap \dots \cap \{X_n \leq y\}] & \text{pre } x < y; \end{cases}$$

Totíž, ak $x \geq y$ a chceme aby minimum z X_1, \dots, X_n bolo menšie ako x a zároveň maximum z X_1, \dots, X_n bolo menšie ako y , tak stačí, aby všetky náhodné veličiny X_1, \dots, X_n boli menšie, ako y . Na druhú stranu, ak $x < y$, tak minimálne jedna hodnota z X_1, \dots, X_n musí byť menšia, ako x a zároveň všetky hodnoty musia byť menšie, ako y .

Združená distribučná funcia pre prípad $x \geq y$ je preto zrejماً, a platí, že

$$G(x, y) = [F(y)]^n \quad \text{pre } x \geq y.$$

Pre prípad $x < y$ zavedieme náhodné javy $A_i = \{X_i \leq x \wedge X_1 \leq y \wedge \dots \wedge X_n \leq y\}$, t.j. náhodný jav A_i značí skutočnosť, že náhodná veličina X_i je menšia než x a všetky náhodné veličiny (samozrejme vrátane X_i) sú menšie, než y . Premyslieť, že náhodný jav A_i môže znamenať aj to, že všetky náhodné veličiny X_1, \dots, X_n sú menšie, než x (t.j. náhodné javy A_i nejsou neslúčitelne).

Následne môžeme vyjadriť pravdepodobnosti:

$$P(A_i) = F(x) \cdot [F(y)]^{n-1}$$

$$P(A_i \cap A_j) = [F(x)]^2 \cdot [F(y)]^{n-2} \quad \text{pre } i \neq j$$

$$P(A_i \cap A_j \cap A_k) = [F(x)]^3 \cdot [F(y)]^{n-3} \quad \text{pre } i \neq j \neq k$$

$$\dots$$

$$P(A_1 \cap \dots \cap A_n) = [F(x)]^n$$

Opäť premyslieť, že náhodný jav, napr. $\{A_i \cap A_j \cap A_k\}$, znamená, že náhodné veličiny X_i, X_j a X_k sú menšie, než x a zároveň všetky veličiny X_1, \dots, X_n sú menšie, než y . Samozrejme to opäť zahŕňa aj situáciu, že všetky náhodné veličiny X_1, \dots, X_n sú menšie, než x .

Pre prípad $x < y$ môžeme preto pomocou princípu inkluze a exkluze vyjadriť distribuční funkciu $G(x, y)$ ako

$$G(x, y) = P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i) - \sum_{i \neq j} P(A_i \cap A_j) + \sum_{i \neq j \neq k} P(A_i \cap A_j \cap A_k) + \dots$$

$$\dots + (-1)^{n-1} P(A_1 \cap \dots \cap A_n)$$

$$= nF(x)[F(y)]^{n-1} - \binom{n}{2}[F(x)]^2 \cdot [F(y)]^{n-2} + \dots + (-1)^{n-1} \binom{n}{n}[F(x)]^n.$$

Do výrazu na poslednom riadku vhodne pripočítame nulu, a to v nasledujúcom tvare:

$$0 = [F(y)]^n - [F(y)]^n = [F(y)]^n - \binom{n}{0}[F(x)]^0[F(y)]^n$$

Binomickou vetou a jednoduchou úpravou potom dostaneme

$$G(x, y) = [F(y)]^n - \sum_{j=0}^n \binom{n}{j} [-F(x)]^j [F(y)]^{n-j} = [F(y)]^n - [F(y) - F(x)]^n \quad \text{pre } x < y.$$

Dohromady teda, združená distribučná funkcia náhodného vektoru $(X_{(1)}, X_{(n)})^\top$ je pre všetky $(x, y) \in \mathbb{R}^2$ daná predpisom

$$G(x, y) = \begin{cases} [F(y)]^n & \text{pre } x \geq y; \\ [F(y)]^n - [F(y) - F(x)]^n & \text{pre } x < y. \end{cases}$$

- Nechť X_i má hustotu f vzhľadom k Lebesguovej mieri. Najdte združenou hustotu $X_{(1)}$ a $X_{(n)}$.

Řešení:

Združenou hustotu náhodného vektoru $(X_{(1)}, X_{(n)})^\top$ získame parciálnym derivovaním združenej distribučnej funkcie, podľa oboch premenných:

$$g(x, y) = \frac{\partial^2}{\partial x \partial y} G(x, y) = \begin{cases} 0 & \text{pre } x \geq y; \\ n(n-1)f(x)f(y)[F(y) - F(x)]^{n-2} & \text{pre } x < y; \end{cases}$$

Pamätať, že združená hustota je takto definovaná pre všetky body $(x, y) \in \mathbb{R}^2$.

A3. Nechť X_i má spojité rozdelenie s distribuční funkcií F a hustotou f .

- Určete združenou hustotu rozmezí W_n a minima $X_{(1)}$.

Řešení:

Využijeme výsledok predchádzajúceho príkladu—t.j., združené rozdelenie (distribučná funkcia, resp. hustota) náhodného vektoru $(X_{(1)}, X_{(n)})^\top$ je dané funkciou $G(x, y)$, resp. $g(x, y)$ (pre ľubovoľné $(x, y) \in \mathbb{R}^2$). Keďže v zadaní sa požaduje pouze združená hustota, postačí vyšetrovať prípad pre $x < y$, pretože $g(x, y) = 0$ pre $x \geq y$. Následne použijeme vetu o transformácii náhodného vektoru s transformačnou funkciou $t: \mathbb{R}^2 \rightarrow [0, \infty) \times \mathbb{R}$ definovanou nasledovne:

$$t(x, y) = \begin{pmatrix} y - x \\ x \end{pmatrix} = \begin{pmatrix} w \\ z \end{pmatrix}.$$

Príslušné inverzné zobrazenie (keďže $w = y - x$ a $z = x$) je

$$t^{-1}(w, z) = \begin{pmatrix} z \\ w + z \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Keďže $x < y$, tak máme $z < z + w$ a teda $w > 0$. Pre Jakobián inverzného zobrazenia t^{-1} sa ľahko overí, že platí $|J_{t^{-1}}| = 1$ a preto pre združenú hustotu náhodného vektoru $(W_n, X_{(1)})^\top$ dostaneme z vety o transformácii:

$$h(w, z) = g(z, z + w) = n(n-1)f(z)f(z+w)[F(z+w) - F(z)]^{n-2}, \quad \text{pre } z \in \mathbb{R} \text{ a } w > 0,$$

a $h(w, z) = 0$ jinak (pozor na tento 'nulový' prípad, hustota je totiž definovaná pre všetky body $(w, z) \in \mathbb{R}^2$, takže nestačí pouze zmieniť prípady, pre ktoré je hustota nenulová).

- Nechť X_i má exponenciální rozdělení. Ukažte, že W_n a $X_{(1)}$ jsou nezávislé. Určete rozdělení náhodné veličiny $\exp\{-\lambda W_n\}$.

Řešení:

Nechť $X_i \sim \text{Exp}(\lambda)$ pre nejaký parameter $\lambda > 0$, tak, že príslušná hustota je $f(x) = \lambda \exp\{-\lambda x\}$ pre $x \geq 0$ a $f(x) = 0$ pre $x < 0$ (t.j. $EX = 1/\lambda$). Príslušná distribučná funkcia je

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{pre } x \geq 0; \\ 0 & \text{inak.} \end{cases}$$

Hustotu $f(x)$ je možné zapísať aj v tvare $f(x) = \lambda e^{-\lambda x} \cdot \mathbb{I}_{\{x \geq 0\}}$, kde $\mathbb{I}_{\{\cdot\}}$ je identifikátorová funkcia, ktorá vráti hodnotu jedna, ak je podmienka v argumente splnená a vráti hodnotu nula, ak podmienka splnená nie je. Analogicky aj pre distribučnú funkciu môžeme použiť kompaktnější zápis $F(x) = (1 - e^{-\lambda x}) \cdot \mathbb{I}_{\{x \geq 0\}}$.

Priamým dosadením do hustoty $h(w, z)$ dostaneme, že

$$\begin{aligned} h(w, z) &= n(n-1)f(z)f(z+w)[F(z+w) - F(z)]^{n-2} \\ &= n(n-1) \cdot \lambda e^{-\lambda z} \mathbb{I}_{\{z \geq 0\}} \cdot \lambda e^{-\lambda(z+w)} \mathbb{I}_{\{z+w \geq 0\}} \cdot [e^{-\lambda z} - e^{-\lambda(z+w)}]^{n-2} \cdot \mathbb{I}_{\{w > 0\}} \\ &= n(n-1)\lambda^2 e^{-n\lambda z} e^{-\lambda w} [1 - e^{-\lambda w}]^{n-2} \mathbb{I}_{\{z \geq 0\}} \mathbb{I}_{\{w > 0\}} \\ &= \underbrace{n\lambda e^{-n\lambda z} \mathbb{I}_{\{z \geq 0\}}}_{\text{Exp}(n\lambda)} \cdot \underbrace{(n-1)\lambda e^{-\lambda w} [1 - e^{-\lambda w}]^{n-2} \mathbb{I}_{\{w > 0\}}}_{f_{W_n}(w)}. \end{aligned}$$

Hustotu $h(w, z)$ je teda možné faktorizovať na súčin dvoch členov, pričom oba členy závisia iba na jednej premennej. Prvý člen je evidentne hustota náhodnej veličiny s exponenciálnym rozdelením s parametrom $n\lambda > 0$ (t.j. hustota náhodnej veličiny $X_{(1)}$) a druhý člen nezávisí na $z \in \mathbb{R}$ a jedná sa teda o hustotu náhodnej veličiny W_n . Vďaka tejto faktorizácii vieme, že náhodné veličiny $X_{(1)}$ a W_n sú vzájomne nezávislé (pripomeňte si ekvivalenčný vzťah, ktorý dáva do súvislosti práve takúto faktorizáciu hustoty s nezávislosťou príslušných náhodných veličín).

Na záver, rozdelenie náhodnej veličiny $Y = \exp\{-\lambda W_n\}$ získame opäť pomocou vety o transformácii (tentokrát postačí len jednorozmerná varianta).

Transformácia: $t: w \rightarrow e^{-\lambda w}$, t.j. náhodná veličina $Y = t(W_n) = e^{-\lambda W_n}$;

Inv. transformácia: $t^{-1}: y \rightarrow -\frac{1}{\lambda} \log y$, t.j. náhodná veličina $W_n = t^{-1}(Y) = -\frac{1}{\lambda} \log(e^{-\lambda W_n})$; (zároveň pre $w > 0$ dostávame, že $y \in (0, 1)$)

Podľa vety o transformácii náhodnej veličiny máme pre hustotu náhodnej veličiny Y nasledujúci vzťah:

$$\begin{aligned} f_Y(y) &= f_{W_n} \left(-\frac{1}{\lambda} \log y \right) \cdot \left| \frac{\partial t^{-1}(y)}{\partial y} \right| = (n-1)\lambda e^{\log y} [1 - e^{\log y}]^{n-2} \cdot \frac{1}{\lambda y} \\ &= (n-1)(1-y)^{n-2}, \quad \text{pre } y \in (0, 1), \end{aligned}$$

a samozrejme $f_Y(y) = 0$ jinak. Zároveň sa jedná o hustotu náhodnej veličiny s beta rozdelením s parametrami 1 a $n-1$ (t.j. $Y \sim \text{Beta}(1, n-1)$).

A4. Nechť X_1, \dots, X_n je náhodný výběr z rozdělení $R(0, 1)$. Spočítejte hustotu W_n , $E W_n$ a $\text{var } W_n$.

Řešení:

Pre združenú hustotu náhodných veličín W_n a $X_{(1)}$ už máme hustotu v tvare:

$$h(w, z) = n(n-1)f(z)f(z+w)[F(z+w) - F(z)]^{n-2} \quad \text{pre } z \in \mathbb{R} \text{ a } w > 0,$$

a $h(w, z) = 0$ jinak. Keďže náhodné veličiny X_1, \dots, X_n majú rovnomerné rozdelenie na intervale $(0, 1)$, tak poznáme tvar hustoty $f(x)$ aj distribučnej funkcie $F(x)$. Platí, že:

$$\begin{aligned} f(x) &= \mathbb{I}_{\{x \in (0,1)\}} \\ \text{a tiež} \\ F(x) &= x \quad \text{pre } x \in (0, 1), \end{aligned}$$

pričom $F(x) = 0$ pre $x < 0$ a $F(x) = 1$ pre $x > 1$. Po dosadení preto dostaneme združenú hustotu

$$\begin{aligned} h(w, z) &= n(n-1)\mathbb{I}_{\{x \in (0,1)\}}\mathbb{I}_{\{z+w \in (0,1)\}}[z+w-z]^{n-2}\mathbb{I}_{\{w>0\}} \\ &= n(n-1)\mathbb{I}_{\{z \in (0,1-w)\}}w^{n-2}\mathbb{I}_{\{w \in (0,1)\}}. \end{aligned}$$

Potrebuje totíž, aby

$$\begin{aligned} 0 &< z + w < 1 \\ -w &< z < 1 - w \end{aligned}$$

ale zároveň aj $z > 0$ a $0 < w < 1$. Preto identifikátorové funkcie v danom tvare. Pre hustotu náhodnej veličiny W_n (a príslušné charakteristiky strednej hodnoty a rozptylu) stačí integrovať združenú hustotu vzhľadom k premennej $z \in \mathbb{R}$. Dostaneme:

$$\begin{aligned} f_{w_n}(w) &= \int_{-\infty}^{\infty} h(w, z) dz = n(n-1)w^{n-2} \int_0^{1-w} 1 dz \\ &= n(n-1)w^{n-2}(1-w) \quad \text{pre } w \in (0, 1) \end{aligned}$$

a $f_{w_n}(w) = 0$ inak. Teraz je užitočné uvedomiť si, že sa jedná o hustotu náhodnej veličiny, ktorá má Beta rozdelenie, s parametrami $\alpha = n-1$ a $\beta = 2$. To znamená, že

$$W_n \sim \text{Beta}(\alpha, \beta) \equiv \text{Beta}(n-1, 2),$$

pričom z vlastnosti Beta rozdelenia vieme, že pre strednú hodnotu a rozptyl platia nasledujúce vzťahy:

$$E W_n = \frac{\alpha}{\alpha + \beta} = \frac{n-1}{n+1} \quad \text{Var}(W_n) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} = \frac{2(n-1)}{(n+2)(n+1)^2}.$$