Matematika III – 9. týden
Základní typy a vlastnosti náhodných veličin

Jan Slovák

Masarykova univerzita
Fakulta informatiky

10.-14. 11. 2014

Obsah přednášky

Plán přednášky

Kde je dobré číst?

- Karel Zvára, Josef Štěpán, Pravděpodobnost a matematická pravděpodobnost statistika, Matfyzpress, 2006, 230pp.
- J. Slovák, M. Panák, M. Bulant, Matematika drsně a svižně, Muni Press, Brno 2013, v+773 s., elektronická edice `www.math.muni.cz/Matematika_drsne_svizne`
- Marie Budíková, Štěpán Mikoláš, Pavel Osecký, Teorie pravděpodobnosti a matematická statistika (sbírka příkladů), Masarykova univerzita, 3. vydání, 2004, 117 stran, ISBN 80-210-3313-4.
- Marie Budíková, Tomáš Lerch, Štěpán Mikoláš, Základní statistické metody, Masarykova univerzita, 2005, 170 stran, ISBN 80-210-3886-1.
- Riley, K.F., Hobson, M.P., Bence, S.J. Mathematical Methods for Physics and Engineering, second edition, Cambridge University Press, Cambridge 2004, ISBN 0 521 89067 5, xxiii + 1232 pp.

Plán přednášky

Diskrétní náhodné veličiny

Jestliže náhodná veličina X na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ nabývá jen konečně nebo spočetně mnoha hodnot $x_1, x_2, \dots \in \mathbb{R}$, pak existuje **pravděpodobnostní funkce** $f(x)$ taková, že

$$f(x) = \begin{cases} P(X = x_i) & x = x_i \\ 0 & \text{jinak.} \end{cases}$$

Spojité náhodné veličiny

Hustota $f(x)$ **pravděpodobnosti** pro náhodnou veličinu X je funkce splňující pro $-\infty \leq a \leq b \leq \infty$

$$P(a < X < b) = \int_a^b f(x)dx. \qquad (*)$$

Náhodná veličina X, pro kterou existuje její **hustota pravděpodobnosti** splňující $(*)$, se nazývá **spojitá**.

Degenerované a alternativní rozdělení.

Degenerované rozdělení $\mathrm{D}(\mu)$ odpovídá konstantní hodnotě $X = \mu$. Distribuční funkce F_X a pravděpodobnostní funkce f_X jsou tedy rovny

$$F_X(t) = \begin{cases} 0 & t \leq \mu \\ 1 & t > \mu \end{cases} \qquad f_X(t) = \begin{cases} 1 & t = \mu \\ 0 & \text{jinak} \end{cases}.$$

Degenerované a alternativní rozdělení.

Degenerované rozdělení $\mathrm{D}(\mu)$ odpovídá konstantní hodnotě $X = \mu$. Distribuční funkce F_X a pravděpodobnostní funkce f_X jsou tedy rovny

$$F_X(t) = \begin{cases} 0 & t \leq \mu \\ 1 & t > \mu \end{cases} \qquad f_X(t) = \begin{cases} 1 & t = \mu \\ 0 & \text{jinak} \end{cases}.$$

Alternativní rozdělení $\mathrm{A}(p)$ popisuje pokus s pouze dvěma možnými výsledky, kterým budeme říkat zdar a nezdar. Náhodné veličině X pro určitost přiřadíme hodnotu 0 pro nezdar a 1 pro zdar. Pokud má zdar pravděpodobnost p, pak nezdar musí mít pravděpodobnost $1 - p$. Jsou tedy distribuční a pravděpodobnostní funkce tvaru:

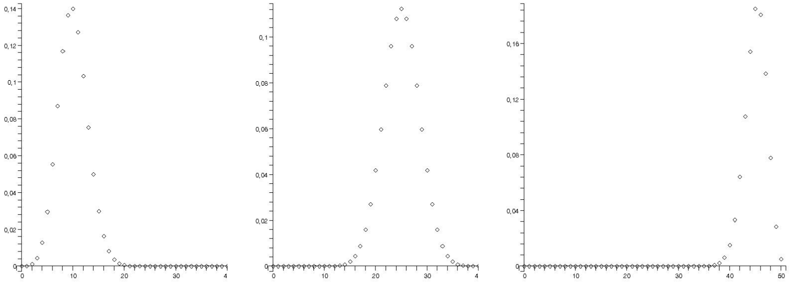
$$F_X(t) = \begin{cases} 0 & t \leq 0 \\ 1-p & 0 < t \leq 1 \\ 1 & t > 1 \end{cases} \qquad f_X(t) = \begin{cases} p & t = 1 \\ 1-p & t = 0 \\ 0 & \text{jinak} \end{cases}.$$

Binomické rozdělení $\mathrm{Bi}(n, p)$

odpovídá n–krát nezávisle opakovanému pokusu popsanému alternativním rozdělením, přičemž naše náhodná veličina měří počet zdarů. Je tedy zjevné, že pravděpodobnostní funkce bude mít nenulové hodnoty právě v celých číslech $0, \dots, n$ odpovídajícím celkovému počtu úspěchů v pokusech (a nezáleží nám na pořadí). Je tedy

$$f_X(t) = \begin{cases} \binom{n}{t} p^t (1-p)^{1-t} & t \in \{0, 1, \dots, n\} \\ 0 & \text{jinak} \end{cases} .$$

Na obrázku jsou pravděpodobnostní funkce pro Bi$(50, 0.2)$, Bi$(50, 0.5)$ a Bi$(50, 0.9)$. Rozdělení pravděpodobnosti dobře odpovídá intuici, že nejvíce výsledků bude blízko u hodnoty np:



S binomickým rozdělením se potkáváme velice často v praktických úlohách. Jednou z nich je popis náhodné veličiny, která popisuje počet X předmětů v jedné zvolené přihrádek z n možných, do nichž jsme náhodně rozdělili r předmětů.
Umístění kteréhokoliv předmětu do pevně zvolené přihrádky má pravděpodobnost $1/n$ (každá z nich je stejně pravděpodobná).
Zjevně tedy bude pro jakýkoliv počet $k = 0, \dots, r$

$$P(X = k) = \binom{r}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{r-k} = \binom{r}{k} \frac{(n-1)^{r-k}}{n^r},$$

jde proto o rozložení X typu $\mathrm{Bi}(r, 1/n)$.

Jestliže nám bude vzrůstat počet přihrádek n společně s počtem předmětů r_n tak, že v průměru nám na každou přihrádku bude připadat (přibližně) stejný počet prvků λ, můžeme dobře vyjádřit chování našeho rozdělení veličin X_n při limitním přechodu $n \to \infty$:

Jestliže nám bude vzrůstat počet přihrádek n společně s počtem předmětů r_n tak, že v průměru nám na každou přihrádku bude připadat (přibližně) stejný počet prvků λ, můžeme dobře vyjádřit chování našeho rozdělení veličin X_n při limitním přechodu $n \to \infty$:

$$
\begin{aligned}
\lim_{n\to\infty} P(X_n = k) &= \lim_{n\to\infty} \binom{r_n}{k} \frac{(n-1)^{r_n-k}}{n^{r_n}} \\
&= \lim_{n\to\infty} \frac{r_n(r_n-1)\dots(r_n-k+1)}{(n-1)^k} \frac{1}{k!} \left(1-\frac{1}{n}\right)^{r_n} \\
&= \frac{\lambda^k}{k!} \lim_{n\to\infty} \left(1+\frac{-\frac{r_n}{n}}{r_n}\right)^{r_n} \\
&= \frac{\lambda^k}{k!}\, \mathrm{e}^{-\lambda}
\end{aligned}
$$

protože obecně funkce $(1+x/n)^n$ konvergují stejnoměrně k funkci e^x na každém omezeném intervalu v $\mathbb{R}$.
To dává **Poissonovo rozdělení** $\mathrm{Po}(\lambda)$.

Poissonovo rozdělení

Poissonovo rozdělení Po(λ) popisuje např. události, které se vyskytují náhodně v čase a přitom pravděpodobnost výskytu v následujícím časovém intervalu o jednotkové délce nezávisí na předchozí historii a je rovna stále stejné hodnotě λ.
V praxi jsou takové procesy spojeny např. s poruchovostí strojů a zařízení.

Poissonovo rozdělení

Poissonovo rozdělení $\mathrm{Po}(\lambda)$ popisuje např. události, které se vyskytují náhodně v čase a přitom pravděpodobnost výskytu v následujícím časovém intervalu o jednotkové délce nezávisí na předchozí historii a je rovna stále stejné hodnotě λ.
V praxi jsou takové procesy spojeny např. s poruchovostí strojů a zařízení.

Theorem (Poissonova věta)

Jsou-li $X_n \sim \mathrm{Bi}(n, p_n)$ *a* $\lim_{n\to\infty} n \cdot p_n = \lambda$ *je konečná, pak*

$$\lim_{n\to\infty} P(X_n = k) = P(X = k),$$

kde $X \sim \mathrm{Po}(\lambda)$.

Příklady spojitých rozdělení

Nejjednodušší je tzv. **rovnoměrné rozdělení**. Jestliže chceme, aby pravděpodobnost každé hodnoty v předem daném intervalu $(a, b) \subset \mathbb{R}$ byla stejná, pak hustota f_X našeho rozdělení náhodné veličiny X má být konstantní. Pak ovšem jsou pro libovolná reálná čísla $-\infty < a < b < \infty$ jen jediné možné hodnoty

$$f_X(t) = \begin{cases} 0 & t \leq a \\ \frac{1}{b-a} & t \in (a, b) \\ 0 & t \geq b, \end{cases} \qquad F_X(t) = \begin{cases} 0 & t \leq a \\ \frac{t-a}{b-a} & t \in (a, b) \\ 1 & t \geq b. \end{cases}$$

Exponenciální rozdělení $\mathrm{ex}(\lambda)$

je dalším rozdělením, které je snadno určeno požadovanými vlastnostmi náhodné veličiny. Předpokládejme, že sledujeme výskyt náhodného jevu tak, že výskyty v nepřekrývajících se intervalech jsou nezávislé. Je-li tedy $P(t)$ pravděpodobnost, že jev **nenastane** během intervalu délky t, pak nutně $P(t+s) = P(t)P(s)$ pro všechna $t, s > 0$. Předpokládejme navíc diferencovatelnost funkce P a $P(0) = 1$. Pak $\ln P(t+s) = \ln P(t) + \ln P(s)$.

Exponenciální rozdělení $\mathrm{ex}(\lambda)$

je dalším rozdělením, které je snadno určeno požadovanými vlastnostmi náhodné veličiny. Předpokládejme, že sledujeme výskyt náhodného jevu tak, že výskyty v nepřekrývajících se intervalech jsou nezávislé. Je-li tedy $P(t)$ pravděpodobnost, že jev **nenastane** během intervalu délky t, pak nutně $P(t+s) = P(t)P(s)$ pro všechna $t, s > 0$. Předpokládejme navíc diferencovatelnost funkce P a $P(0) = 1$. Pak $\ln P(t+s) = \ln P(t) + \ln P(s)$. Limitním přechodem:

$$\lim_{s \to 0_+} \frac{\ln P(t+s) - \ln P(t)}{s} = (\ln P)'(0) = -\lambda.$$

Odtud vyplývá diferenciální rovnice

$$(\ln P(t))' = -\lambda.$$

Odtud dostáváme $\ln P(t) = -\lambda t + C$ a počáteční podmínka dává jediné řešení

$$P(t) = e^{-\lambda t}.$$

Všimněme si, že z definice našich objektů vyplývá, že $\lambda > 0$.

Odtud dostáváme $\ln P(t) = -\lambda t + C$ a počáteční podmínka dává jediné řešení

$$P(t) = e^{-\lambda t}.$$

Všimněme si, že z definice našich objektů vyplývá, že $\lambda > 0$.
Uvažme náhodnou veličinu X udávající okamžik, kdy náš jev poprvé **nastane**. Zřejmě tedy je distribuční funkce rozdělení pro X dána

$$F_X(t) = 1 - P(t) = \begin{cases} 1 - e^{-\lambda t} & t > 0 \\ 0 & t \le 0. \end{cases}$$

Je vidět, že je to rostoucí funkce s hodnotami mezi nulou a jedničkou a správnými limitami v $\pm\infty$.
Hustotu tohoto rozdělení dostaneme derivováním distribuční funkce, tj.

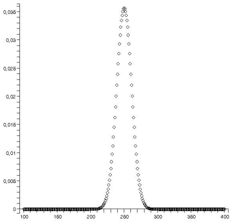
$$f_X = \begin{cases} \lambda e^{-\lambda t} & t > 0 \\ 0 & t \le 0. \end{cases}$$

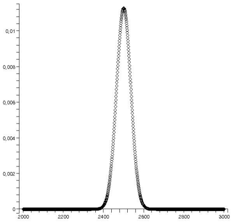
Normální rozdělení

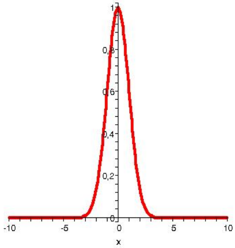
je ze všech nejdůležitější.

Normální rozdělení

je ze všech nejdůležitější.
Jestliže v binomiálním rozdělení zachováme konstatní úspěšnost p, ale budeme přidávat počet pokusů n, bude pravděpodobnostní funkce kupodivu pořád mít podobný tvar (i když jiné rozměry). Na obrázku při rostoucím n se budou vynesené bodové hodnoty slévat do křivky, pro hodnoty $\mathrm{Bi}(500, 0.5)$ a $\mathrm{Bi}(5000, 0.5)$ je výsledek vidět na obrázku níže. Třetí křivka na obrázku je grafem funkce $f(x) = \mathrm{e}^{-x^2/2}$.







Hledáme-li podobné spojité rozdělení, potřebovali bychom spočíst $\int_a^b e^{-x^2/2}\,dx$ což není pomocí elementárních funkcí možné. Je však možné (i když ne úplně snadné) ověřit, že příslušný nevlastní integrál konverguje k hodnotě

$$\int_{-\infty}^{\infty} e^{-x^2/2}\,dx = \sqrt{2\pi}.$$

Odtud vyplývá, že možná hustota rozdělení náhodného rozdělení může být

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^x.$$

Hledáme-li podobné spojité rozdělení, potřebovali bychom spočíst $\int_a^b \mathrm{e}^{-x^2/2}\,dx$ což není pomocí elementárních funkcí možné. Je však možné (i když ne úplně snadné) ověřit, že příslušný nevlastní integrál konverguje k hodnotě

$$\int_{-\infty}^{\infty} \mathrm{e}^{-x^2/2}\,dx = \sqrt{2\pi}.$$

Odtud vyplývá, že možná hustota rozdělení náhodného rozdělení může být

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^x.$$

Rozdělení s touto hustotou se nazývá **normální rozdělení** $\mathrm{N}(0,1)$. Příslušnou distribuční funkci

$$F_X(x) = \int_{-\infty}^{x} \mathrm{e}^{-x^2/2}\,dx$$

nelze vyjádřit pomocí elementárních funkcí, přesto se s ní numericky běžně počítá (pomocí tabulek nebo softwarových aplikací).

Hledáme-li podobné spojité rozdělení, potřebovali bychom spočíst $\int_a^b \mathrm{e}^{-x^2/2}\,dx$ což není pomocí elementárních funkcí možné. Je však možné (i když ne úplně snadné) ověřit, že příslušný nevlastní integrál konverguje k hodnotě

$$\int_{-\infty}^{\infty} \mathrm{e}^{-x^2/2}\,dx = \sqrt{2\pi}.$$

Odtud vyplývá, že možná hustota rozdělení náhodného rozdělení může být

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^x.$$

Rozdělení s touto hustotou se nazývá **normální rozdělení** $\mathrm{N}(0,1)$. Příslušnou distribuční funkci

$$F_X(x) = \int_{-\infty}^{x} \mathrm{e}^{-x^2/2}\,dx$$

nelze vyjádřit pomocí elementárních funkcí, přesto se s ní numericky běžně počítá (pomocí tabulek nebo softwarových aplikací). Hustotě f_X se také často říká **Gaussova křivka**.

Plán přednášky

Obdobně definujeme distribuční funkce a hustotu a pravděpodobnostní funkci pro spojité a diskrétní náhodné vektory. Hovoříme také o **simultánních (sdružených) pravděpodobnostních funkcích a hustotách**.
Pro dvě proměnné (vektor (X, Y) náhodných veličin):

$$f(x,y) = \begin{cases} P(X = x_i \wedge Y = y_j) & x = x_i \wedge y = y_j \\ 0 & \text{jinak.} \end{cases}$$

u diskrétních a pro všechny $a, b \in \mathbb{R}$ pro spojité:

$$F(b,a) = P(-\infty < X < b, \infty < Y < a) = \int_{-\infty}^{a} \int_{-\infty}^{b} f(x,y)dxdy.$$

Marginální rozložení

pro jednu z proměnných obdržíme tak, že přes ostatní posčítáme nebo zintegrujeme. Např. u diskrétních vektorových veličin (X, Y) tvoří jevy $(X = x_i, Y = y_j)$ pro všechny možné hodnoty x_i a y_j s nenulovými pravděpodobnostmi pro X a Y úplný systém jevů pro vektor (X, Y) a dostáváme vztah:

$$P(X = x_i) = \sum_{j=1}^{\infty} P(X = x_i, Y = y_j)$$

mezi **marginálním rozdělením pravděpodobnosti** náhodné veličiny X a **sdruženým rozdělením pravděpodobnosti** náhodného vektoru (X, Y).

Náhodné veličiny X a Y jsou **stochasticky nezávislé**, jestliže jejich sdružená distribuční funkce splňuje

$$F(x, y) = G(x) \cdot H(y),$$

kde G a H jsou distribuční funkce veličin X a Y.

Plán přednášky

Definition

Pro danou spojitou funkci $\psi : \mathbb{R} \to \mathbb{R}$ a náhodnou veličinu X máme dánu také náhodnou veličinu $Y = \psi(X)$. Nazýváme ji **funkcí náhodné veličiny** X.
V případě náhodného vektoru $(X_1, \ldots, X_n)$ a funkce $\psi : \mathbb{R}^n \to \mathbb{R}$ hovoříme o funkci $Y = \psi(X_1, \ldots, X_n)$ náhodného vektoru.

Požadavek spojitosti ψ zaručuje, že je Y opět náhodnou veličinou podle naší definice, protože vzor borelovské množiny ve spojitém zobrazení je opět borelovská množina.
Obecněji můžeme právě tento požadavek na ψ vztáhnout pro každý speciální případ veličiny či vektoru a definovat tak pojem funkce z náhodné veličiny či vektoru obecněji.

Nejjednodušší funkcí po konstantách je afinní závislost

$$\psi(X) = a + bX$$

s konstantami $a, b \in \mathbb{R}$, $b \neq 0$.
Je-li $f_X(x)$ pravděpodobnostní funkce náhodné veličiny s diskrétním rozdělením, snadno se vypočte

$$f_{\psi(X)}(y) = P(\psi(X) = y) = \sum_{\psi(x_i)=y} f(x_i).$$

V případě afinní závislosti $Y = a + bX$ je proto pravděpodobnostní funkce nenulová právě v bodech $y_i = ax_i + b$.
Např. součet n nezávislých náhodných veličin s alternativním rozdělením $\mathrm{A}(p)$ je veličina s binomiální rozdělení $\mathrm{Bi}(n, p)$.

Podobně můžeme přepočíst distribuční funkci rozdělení funkce ze spojité náhodné veličiny, či vektoru.
Např. má-li Z s normální rozdělení $\mathrm{N}(0,1)$, pak veličiny $Y = \mu + \sigma Z$ budou mít normální rozdělení $\mathrm{N}(\mu, \sigma)$.

Podobně můžeme přepočíst distribuční funkci rozdělení funkce ze spojité náhodné veličiny, či vektoru.
Např. má-li Z s normální rozdělení $\mathrm{N}(0,1)$, pak veličiny $Y = \mu + \sigma Z$ budou mít normální rozdělení $\mathrm{N}(\mu,\sigma)$.
Se součty nezávislých spojitých veličin X a Y s hustotami f_X a f_Y je to složitější. Přímým výpočtem spočteme distribuční funkci náhodné promnné $V = X + Y$.

$$\begin{aligned} F_V(u) &= \int_{x+y<u} f_X(x) f_Y(y)\, dxdy \\ &= \int_{-\infty}^{u} \left(\int_{-\infty}^{\infty} f_X(x) f_Y(v-x)\, dx \right) dv. \end{aligned}$$

Je tedy sdruženou hustotou součtu dvou nezávislých veličin právě konvoluce jejich hustot

$$f_V = f_X * f_Y.$$