# Teória pravdepodobnosti

## 7. Podmienené rozdelenia

Aleš Kozubík

Katedra Matematických metód a operačnej analýzy
Fakulta Riadenia a Informatiky
Žilinská Univerzita v Žiline

2. marca 2015

# Podmienené rozdelenie

## Motivácia

V predchádzajúcej časti sme videli, že pri znalosti zákona rozdelenia systému náhodných premenných $X_1, \ldots X_n$ vieme vždy určiť zákony rozdelenia pre jednotlivé náhodné premenné z tohto systému.

Budeme sa zaoberať otázkou riešiteľnosti opačnej úlohy, tj. otázkou rekonštrukcie združeného rozdelenia z jednotlivých marginálnych rozdelení.

Ukazuje sa, že táto úloha vo všeobecnosti riešiteľná nie je. Je potrebné poznať aj závislosti medzi jednotlivými náhodnými premennými.

Pre jednoduchosť sa budeme venovať len dvojici náhodných premenných.

# Podmienené rozdelenie

## Definícia

**Definícia (Podmienené rozdelenie)**

Nech $(X, Y)$ je dvojica diskrétnych náhodných premenných so združenou pravdepodobnostnou funkciou $p(x, y)$. *Podmienené rozdelenie náhodnej premennej $X$ pri podmienke $Y = y$* je určené *podmienenou pravdepodobnostnou funkciou*

$$p(x|y) = \mathbb{P}\left(X = x \middle| Y = y\right) = \frac{p(x, y)}{p_Y(y)} \quad \text{ak } p_Y(y) > 0.$$

# Podmienené rozdelenie
## Definícia

### Definícia (Podmienené rozdelenie)

Nech $(X, Y)$ je dvojica spojitých náhodných premenných so združenou hustotou pravdepodobnosti $f(x, y)$. *Podmienené rozdelenie náhodnej premennej $X$ pri podmienke $Y = y$* je určené *podmienenou hustotou pravdepodobnosti*

$$f_X(x|Y = y) = \frac{f_{X,Y}(x, y)}{f_Y(y)} \quad \text{ak } f_Y(y) > 0.$$

# Podmienená pravdepodobnostná funkcia

## Príklad

### Zadanie

Uvažujme trojnásobný hod mincou. Nech náhodná premenná $X$ označuje počet hláv, padnutých v prvých dvoch pokusoch a náhodná premenná $Y$ počet hláv padnutých v druhých dvoch pokusoch. Určíme združenú pravdepodobnostnú funkciu dvojice náhodných premenných $(X, Y)$ a podmienené pravdepodobnostné funkcie.

# Podmienená pravdepodobnostná funkcia
## Príklad

Videli sme, že združenú pravdepodobnostnú funkciu môžeme prehľadne zapísať vo forme tabuľky

| | | $y_1$ | $y_2$ | $y_3$ |
|---|---|---|---|---|
| | | 0 | 1 | 2 |
| $x_1$ | 0 | $\frac{1}{8}$ | $\frac{1}{8}$ | 0 |
| $x_2$ | 1 | $\frac{1}{8}$ | $\frac{1}{4}$ | $\frac{1}{8}$ |
| $x_3$ | 2 | 0 | $\frac{1}{8}$ | $\frac{1}{8}$ |

# Podmienená pravdepodobnostná funkcia
## Príklad

Z tejto združenej pravdepodobnostnej funkcie vidíme, že:

$$\mathbb{P}(X = 0 | Y = 0) = \frac{p(0,0)}{p_Y(0)} = \frac{\frac{1}{8}}{\frac{1}{4}} = \frac{1}{2},$$
$$\mathbb{P}(X = 1 | Y = 0) = \frac{p(1,0)}{p_Y(0)} = \frac{\frac{1}{8}}{\frac{1}{4}} = \frac{1}{2},$$
$$\mathbb{P}(X = 2 | Y = 0) = \frac{p(2,0)}{p_Y(0)} = \frac{0}{\frac{1}{4}} = 0.$$

Výsledkom je teda „dvojhodnotové“ rozdelenie, a to aj napriek tomu, že náhodná premenná $X$ môže nadobúdať až tri hodnoty.

# Podmienená pravdepodobnostná funkcia
## Príklad

Podobne určíme aj podmienené rozdelenie náhodnej premennej $Y$ pri podmienke $X = 0$:

$$\mathbb{P}(Y = 0|X = 0) = \frac{p(0,0)}{p_X(0)} = \frac{\frac{1}{8}}{\frac{1}{4}} = \frac{1}{2},$$

$$\mathbb{P}(Y = 1|X = 0) = \frac{p(0,1)}{p_X(0)} = \frac{\frac{1}{8}}{\frac{1}{4}} = \frac{1}{2},$$

$$\mathbb{P}(Y = 2|X = 0) = \frac{p(0,2)}{p_X(0)} = \frac{0}{\frac{1}{4}} = 0.$$

Výsledné rozdelenie náhodnej premennej $Y$ pri podmienke $X = 0$ je teda opäť „dvojhodnotové“ rozdelenie, ktoré je identické s rozdelením náhodnej premennej $X$ pri podmienke $Y = 0$.

# Podmienená pravdepodobnostná funkcia

## Príklad

Oproti tomu pre podmienené rozdelenie náhodnej premennej $X$ pri podmienke $Y = 1$ platí:

$$\mathbb{P}(X = 0|Y = 1) = \frac{p(0,1)}{p_Y(1)} = \frac{\frac{1}{8}}{\frac{1}{2}} = \frac{1}{4},$$

$$\mathbb{P}(X = 1|Y = 1) = \frac{p(1,1)}{p_Y(1)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2},$$

$$\mathbb{P}(X = 2|Y = 1) = \frac{p(2,1)}{p_Y(1)} = \frac{\frac{1}{8}}{\frac{1}{2}} = \frac{1}{4}.$$

Toto rozdelenie už od predchádzajúcich dvoch odlišné je.

# Podmienená hustota pravdepodobnosti

## Príklad

### Zadanie

Nech združená hustota náhodných premenných $(X, Y)$ má tvar

$$f_{X,Y}(x,y) = \begin{cases} k(x^2+y^2) & \text{pre } 0 \leq x \leq a, 0 \leq y \leq b, \\ 0 & \text{inde.} \end{cases}$$

a) Určme konštantu $k$.

b) Určme marginálne hustoty pre $X$ a $Y$.

c) Určme podmienené hustoty $f_X(x|Y=y), f_Y(y|X=x)$ a $f_Y(y|X<\frac{a}{2})$.

d) Určme distribučnú funkciu $F_{X,Y}(x,y)$.

# Podmienená hustota pravdepodobnosti
## Príklad

a) Z vlastností hustoty máme

$$\begin{aligned}1 &= \int_0^b \int_0^a k(x^2+y^2)\,\mathrm{d}x\,\mathrm{d}y = k\int_0^b \left[\frac{x^3}{3}+xy^2\right]_0^a \mathrm{d}y \\ &= k\int_0^b \left(\frac{a^3}{3}+ay^2\right)\mathrm{d}y = k\left[\frac{a^3y}{3}+\frac{ay^3}{3}\right]_0^b \\ &= k\left(\frac{a^3b}{3}+\frac{ab^3}{3}\right).\end{aligned}$$

Odtiaľ

$$k = \left(\frac{a^3b}{3}+\frac{ab^3}{3}\right)^{-1} = \frac{3}{(ab)(a^2+b^2)}.$$

# Podmienená hustota pravdepodobnosti
## Príklad

b) Pre marginálne hustoty môžeme písať:

$$\begin{aligned} f_X(x) &= \int_0^b k(x^2+y^2)\,\mathrm{d}y = k\left[yx^2+\tfrac{y^3}{3}\right]_0^b \\ &= k\left(bx^2+\tfrac{b^3}{3}\right) \quad \text{pre } 0 \le x \le a, \end{aligned}$$

pričom $k$ bolo určené v časti a).

Podobne

$$f_Y(y) = k\left(ay^2+\frac{a^3}{3}\right) \quad \text{pre } 0 \le y \le b.$$

# Podmienená hustota pravdepodobnosti
## Príklad

c) Vypočítame:

$$f_X(x|Y=y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{k(x^2+y^2)}{k\left(ay^2+\frac{a^3}{3}\right)} = \frac{3(x^2+y^2)}{a(a^2+3y^2)}$$

pre $0 \leq x \leq a$ a $0 \leq y \leq b$. Podobne

$$f_Y(y|X=x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{3(x^2+y^2)}{b(b^2+3x^2)}$$

pre $0 \leq x \leq a$ a $0 \leq y \leq b$.

# Podmienená hustota pravdepodobnosti

## Príklad

Konečne:

$$
\begin{aligned}
f_Y(y|X < \frac{a}{2}) &= \frac{\int_0^{\frac{a}{2}} k(x^2 + y^2)\,\mathrm{d}x}{\int_0^{\frac{a}{2}} k\left(bx^2 + \frac{b^3}{3}\right)\mathrm{d}x} \\
&= \frac{k\left[\frac{x^3}{3} + xy^2\right]_0^{\frac{a}{2}}}{k\left[b\frac{x^3}{3} + x\frac{b^3}{3}\right]_0^{\frac{a}{2}}} \\
&= \frac{\frac{a^3}{24} + \frac{ay^2}{2}}{\frac{ba^3}{24} + \frac{ab^3}{6}} \\
&= \frac{a^2 + 12y^2}{ba^2 + 4b^3} \quad \text{pre } 0 \le y \le b.
\end{aligned}
$$

# Podmienená hustota pravdepodobnosti

## Príklad

d) Podľa definície

$$\begin{aligned} F_{X,Y}(x,y) &= \int_0^y \int_0^x k(u^2+v^2)\,\mathrm{d}u\,\mathrm{d}v = \int_0^y k\left[\frac{u^3}{3}+uv^2\right]_0^x \\ &= k\int_0^y \left(\frac{x^3}{3}+xv^2\right)\mathrm{d}v = k\left(\frac{x^3y}{3}+\frac{xy^3}{3}\right) = \frac{xy(x^2+y^2)}{ab(a^2+b^2)}, \end{aligned}$$

pre $0 \le x \le a$ a $0 \le y \le b$. Teda celkom:

$$F_{X,Y}(x,y) = \begin{cases} 0 & \text{ak } x < 0 \text{ alebo } y < 0, \\ \frac{xy(x^2+y^2)}{ab(a^2+b^2)} & \text{ak } 0 \le x \le a \text{ a } 0 \le y \le b, \\ \frac{xb(x^2+b^2)}{ab(a^2+b^2)} & \text{ak } 0 \le x \le a \text{ a } y > b, \\ \frac{ay(a^2+y^2)}{ab(a^2+b^2)} & \text{ak } x > a \text{ a } 0 \le y \le b, \\ 1 & \text{ak } x > a \text{ a } y > b. \end{cases}$$

# Systém náhodných premenných

## Nezávislosť náhodných premenných

### Definícia (Nezávislosť)

O dvoch náhodných premenných $X$ a $Y$ povieme, že *sú nezávislé*, ak pre každé dva intervaly $A$ a $B$ platí

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B)\,. \tag{1}$$

### Poznámka

Porovnajte definíciu nezávislosti náhodných premenných s definíciou nezávislosti náhodných udalostí.

# Systém náhodných premenných

## Nezávislosť náhodných premenných

Definíciu nezávislosti môžeme ekvivalentne definovať pomocou distribučnej funkcie resp. funkcie hustoty.

**Definícia (Nezávislosť)**

Spojité náhodné premenné $X$ a $Y$ nazývame *nezávislé*, ak pre ich združenú funkciu hustoty platí

$$f_{X,Y}(x,y) = f_X(x) f_Y(y), \tag{2}$$

pre každý bod $(x, y)$, resp. ak pre ich združenú distribučnú funkciu platí

$$F_{X,Y}(x,y) = F_X(x) F_Y(y), \tag{3}$$

pre každý bod $(x, y)$.

# Systém náhodných premenných
## Nezávislosť náhodných premenných

Definíciu nezávislosti môžeme v diskrétnom prípade ekvivalentne definovať pomocou distribučnej funkcie resp. pravdepodobnostnej funkcie.

**Definícia (Nezávislosť)**

Diskrétne náhodné premenné $X$ a $Y$ nazývame *nezávislé*, ak pre ich združenú pravdepodobnostnú funkciu platí

$$p_{X,Y}(x,y) = p_X(x)p_Y(y), \tag{4}$$

pre každý bod $(x,y)$, resp. ak pre ich združenú distribučnú funkciu platí

$$F_{X,Y}(x,y) = F_X(x)F_Y(y), \tag{5}$$

pre každý bod $(x,y)$.

# Nezávislosť náhodných premenných

## Príklad

### Zadanie

Združená funkcia hustoty dvoch náhodných premenných $X$ a $Y$ je daná predpisom

$$f_{X,Y}(x,y) = cxy\,\mathrm{e}^{-x^2-y^2}, \quad \text{pre } x \geq 0, y \geq 0,$$

kde $c$ je konštanta. Rozhodnime, či náhodné premenné $X$ a $Y$ sú nezávislé.

# Nezávislosť náhodných premenných
## Príklad

Podľa vlastností združenej funkcie hustoty platí

$$\begin{aligned}\int_0^\infty \int_0^\infty cxy\, \mathrm{e}^{-x^2-y^2}\,\mathrm{d}x\,\mathrm{d}y &= c\int_0^\infty x\,\mathrm{e}^{-x^2}\,\mathrm{d}x \int_0^\infty y\,\mathrm{e}^{-y^2}\,\mathrm{d}y \\ &= c\left[-\frac{\mathrm{e}^{-x^2}}{2}\right]_0^\infty \left[-\frac{\mathrm{e}^{-y^2}}{2}\right]_0^\infty \\ &= c\frac{1}{2}\frac{1}{2} = \frac{c}{4} = 1\end{aligned}$$

odkiaľ dostávame, že $c = 4$.

# Nezávislosť náhodných premenných
## Príklad

Pre marginálne hustoty dostávame

$$f_X(x) = \int_0^{\infty} 4xy\,\mathrm{e}^{-x^2-y^2}\,\mathrm{d}y = 2x\,\mathrm{e}^{-x^2}\int_0^{\infty} 2y\,\mathrm{e}^{-y^2}\,\mathrm{d}y$$
$$= 2x\,\mathrm{e}^{-x^2}\left[-\,\mathrm{e}^{-y^2}\right]_0^{\infty} = 2x\,\mathrm{e}^{-x^2},$$

pre $x \geq 0$. Analogicky zistíme

$$f_Y(z) = 2y\,\mathrm{e}^{-y^2},$$

pre $y \geq 0$.

# Nezávislosť náhodných premenných
## Príklad

Teraz už stačí len dosadiť do podmienky (2).

Tak dostávame

$$f_X(x) \cdot f_Y(y) = 2x\,\mathrm{e}^{-x^2} \cdot 2y\,\mathrm{e}^{-y^2} = 4xy\,\mathrm{e}^{-x^2-y^2} = f_{X,Y}(x,y),$$

čo potvrdzuje, že náhodné premenné $X$ a $Y$ sú nezávislé.

# Nezávislosť náhodných premenných
## Príklad

### Rovnomerné rozdelenie na jednotkovom kruhu

Združená funkcia hustoty dvoch náhodných premenných $X$ a $Y$ rovnomerne rozdelených na jednotkovom kruhu $D = \{(x, y) : x^2 + y^2 < 1\}$ je daná predpisom

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{\pi} & \text{ak } (x, y) \in D \\ 0 & \text{inak.} \end{cases}$$

Rozhodnime, či náhodné premenné $X$ a $Y$ sú nezávislé.

# Nezávislosť náhodných premenných
## Príklad

Už sme zistili, že

$$f_X(x) = \begin{cases} \frac{2}{\pi}\sqrt{1-x^2} & \text{ak } x \in (-1;1) \\ 0 & \text{inak,} \end{cases}$$

a

$$f_Y(y) = \begin{cases} \frac{2}{\pi}\sqrt{1-y^2} & \text{ak } y \in (-1;1) \\ 0 & \text{inak.} \end{cases}$$

# Nezávislosť náhodných premenných
## Príklad

V tomto prípade je zrejmé, že

$$f_X(x) \cdot f_Y(y) \neq f_{X,Y}(x,y),$$

takže náhodné premenné $X$ a $Y$ sú v tomto prípade závislé.

# Veta o úplnej pravdepodobnosti

## Veta (Veta o úplnej pravdepodobnosti)

Nech $X$ je náhodná premenná a $A$ náhodná udalosť. Ak hustoty $f_{X|A}$ a $f_X$ existujú pre každé $x$, tak platí

$$\mathbb{P}(A) = \int_{-\infty}^{\infty} \mathbb{P}(A|X = x)\, f_X(x)\, \mathrm{d}x, \tag{6}$$

a ak $X$ a $Y$ sú spojité náhodné premenné a $f_{X|Y}$ a $f_Y$ existujú pre každé $x$ a $y$, tak

$$f_X(x) = \int_{-\infty}^{\infty} f_{X|Y}(x, y) f_Y(y)\, \mathrm{d}y. \tag{7}$$

# Bayesova veta

**Veta (Bayesova veta pre podmienené hustoty)**

Nech $(X, Y)$ je náhodný vektor so združenou hustotou $f(x, y)$. Potom pre každé $x$ a $y$ také, že $f_X(x) > 0$, $f_Y(y) > 0$ platí

$$f(y|x) = \frac{f(x|y)f_Y(y)}{f_X(x)}. \tag{8}$$

**Poznámka**

S vyžitím tejto Bayesovej vety je možné jednu podmienenú hustotu previesť na druhú.

# Bayesova veta

S použitím združenej hustoty a podmienenej hustoty môžeme tvrdenie Bayesovej vety prepísať do tvaru:

## Poznámka

Ak $X$ a $Y$ sú spojité náhodné premenné a existujú hustoty $f_{X|Y}$ a $f_Y$, tak pre každé $x, y$, také že $f_X(x) \neq 0$ platí

$$f_{Y|X}(y,x) = \frac{f_{x|y}(x,y) f_Y(y)}{\int_{-\infty}^{\infty} f_{x|y}(x,y) f_Y(y)\,\mathrm{d}y}.$$

# Bayesova veta
## Dôkaz

Postupne dostávame

$$\frac{f(x|y)f_Y(y)}{f_X(x)} = \frac{\frac{f(x,y)}{f_Y(y)}f_Y(y)}{f_X(x)} = \frac{f(x,y)}{f_X(x)} = f(y|x).$$

# Bayesova veta
## Príklad

### Zadanie – dvojstupňový experiment

Predpokladajme, že $X$ je náhodná premenná, ktorá nadobúda kladné hodnoty a jej hustota nech je $f(x)$. Pri danej hodnote $X = x$ sa náhodne vyberie číslo $Y$ z intervalu $(0; x)$. Predpokladajme, že sa dozvieme len hodnotu náhodnej premennej $Y$. Ako by sme mohli odhadnúť $x$?

# Bayesova veta
## Riešenie príkladu

Zo zadania máme $X \sim f(x)$, $Y|X = x \sim \mathrm{Ro}\,(0, x)$. Chceme určiť $\mathbb{E}\,(X|Y = y)$.
V prvom rade musíme určiť podmienenú hustotu $f(x|y)$. Podľa definície máme

$$f(x|y) = \frac{f(x, y)}{f_Y(y)} = \frac{f(y|x)f(x)}{f_Y(y)} = \frac{\frac{1}{x}I_{\{x \geq y\}}f(x)}{\int_y^\infty \frac{1}{x}f(x)\,\mathrm{d}x}$$

preto

$$\mathbb{E}\,(X|Y = y) = \int_y^\infty xf(x|y)\,\mathrm{d}x = \frac{\int_y^\infty x\frac{1}{x}f(x)\,\mathrm{d}x}{\int_y^\infty \frac{1}{x}f(x)\,\mathrm{d}x} = \frac{1 - F(y)}{\int_y^\infty \frac{1}{x}f(x)\,\mathrm{d}x}.$$