Matematická statistika 2

Mirko Navara
Centrum strojového vnímání
katedra kybernetiky FEL ČVUT
Karlovo náměstí, budova G, místnost 104a
http://cmp.felk.cvut.cz/~navara/psi

21. září 2012

1 Podmíněná rozdělení pravděpodobnosti

1.1 Motivační úloha

Příklad (počet es)**:**

1. Hráč má 10 z balíčku 32 karet, v němž jsou 4 esa. Náhodná veličina Y je počet es v jeho ruce. Pravděpodobnost, že má y es, je

$$p_Y(y) = \frac{\binom{4}{y}\binom{32-4}{10-y}}{\binom{32}{10}}, \qquad y \in \{0, 1, 2, 3, 4\},$$

(hypergeometrické rozdělení),

y	0	1	2	3	4
$p_Y(y)$	0.203	0.428	0.289	0.073	0.006

$$\mathsf{E}Y = \frac{10}{32}4 = \frac{5}{4}.$$

2. Podívám se na svých 10 karet (ze stejného balíčku) a vidím, že nastal jev B: mám 2 esa.

Tím se pravděpodobnost počtu es prvního hráče mění na

$$p_{Y|B}(y) = \frac{\binom{2}{y}\binom{22-2}{10-y}}{\binom{22}{10}}\,, \qquad y \in \{0,1,2\}\,,$$

(vybíráme z 22 karet, mezi nimiž jsou 2 esa),

y	0	1	2
$p_Y(y\|B)$	$\frac{2}{7} \doteq 0.286$	$\frac{40}{77} \doteq 0.519$	$\frac{15}{77} \doteq 0.195$

$$\mathsf{E}\,(Y|B) = \frac{10}{22}\,2 = \frac{10}{11}\,.$$

3. Obecněji: Náhodná veličina X je počet es v mé ruce a x její realizace („skutečná hodnota“), tj. nastal jev $X = x$.

   Pravděpodobnost počtu es prvního hráče se mění na

$$p_{Y|X=x}(y) = p_{Y|X}(y|x) = \frac{\binom{4-x}{y}\binom{22-4+x}{10-y}}{\binom{22}{10}}, \quad x \in \{0,\ldots,4\}, \ y \in \{0,\ldots,4-x\}$$

   (vybíráme z $22 - 4 + x$ karet, mezi nimiž je $4 - x$ es),

$$\mathsf{E}\,(Y|X = x) = \frac{10}{22}\,(4 - x) = \frac{5}{11}\,(4 - x)\,, \qquad x \in \{0,\ldots,4\}\,.$$

Tento výsledek popisuje rozdělení náhodné veličiny

$$\mathsf{E}\,(Y|X) = \frac{5}{11}\,(4 - X)\ ,$$

která nabývá hodnot

$$\mathsf{E}\,(Y|X = x) = \frac{5}{11}\,(4 - x) \in \left\{\frac{20}{11}, \frac{15}{11}, \frac{10}{11}, \frac{5}{11}, 0\right\}$$

s pravděpodobnostmi

$$p_X(x) = p_Y(x) = \frac{\binom{4}{x}\binom{32-4}{10-x}}{\binom{32}{10}}\,, \qquad x \in \{0, 1, 2, 3, 4\}\,.$$

x	0	1	2	3	4
$\mathsf{E}\,(Y\|X = x)$	$\frac{20}{11}$	$\frac{15}{11}$	$\frac{10}{11}$	$\frac{5}{11}$	0

Náhodná veličina $\mathsf{E}\,(Y|X)$ má střední hodnotu

$$\begin{aligned}\mathsf{E}\,(\mathsf{E}\,(Y|X)) &= \frac{20}{11}\,p_X(0) + \frac{15}{11}\,p_X(1) + \frac{10}{11}\,p_X(2) + \frac{5}{11}\,p_X(3) \\ &= \mathsf{E}\left(\frac{5}{11}\,(4 - X)\right) = \left(\frac{5}{11}\,(4 - \mathsf{E}X)\right) = \left(\frac{5}{11}\,\left(4 - \frac{5}{4}\right)\right) = \frac{5}{4} = \mathsf{E}Y\,.\end{aligned}$$

1.2 Opakování podmíněné pravděpodobnosti

Pravděpodobnost (=pravděpodobnostní míra) P

Značení: $P(\,.\,)$ pravděpodobnostní míra,

$P[\,.\,]$ pravděpodobnost jevu popsaného v závorce

(je-li argument jev, obě značení se shodují)

Jev B, $P(B) \neq 0 \neq P(\bar{B})$

$\forall$ jev $A : P(A) = c\, P(A|B) + c'\, P(A|\bar{B})$,

kde $c = P(B)$, $c' = P(\bar{B})$, tj. $c + c' = 1$.

P je lineární (dokonce **konvexní**) kombinace podmíněných pravděpodobností $P(\,.\, |B)$, resp. $P(\,.\, |\bar{B})$,

což jsou obyčejné pravděpodobnosti popisující jiné modely (kde B je jev jistý, resp. nemožný).

Nadále předpoklad $P(B) \neq 0$ budeme potřebovat,

předpoklad $P(\bar{B}) \neq 0$ nikoli.

1.3 Rozdělení podmíněné jevem

Podmíněná pravděpodobnostní funkce $p_{Y|B}$ *diskrétní* náhodné veličiny Y za podmínky

$$p_{Y|B}(y) = P[Y = y|B] = \frac{P[Y = y, B]}{P[B]} \ ,$$

kde $P[Y = y, B] = P\left(\{\omega \in \Omega \mid Y(\omega) = y, \omega \in B\}\right) = P\left(\{\omega \in B \mid Y(\omega) = y\}\right)$.

1.3 Rozdělení podmíněné jevem

Podmíněná pravděpodobnostní funkce $p_{Y|B}$ *diskrétní* náhodné veličiny Y za podmínky

$$p_{Y|B}(y) = P[Y = y|B] = \frac{P[Y = y, B]}{P[B]} \ ,$$

kde $P[Y = y, B] = P\left(\{\omega \in \Omega \mid Y(\omega) = y, \omega \in B\}\right) = P\left(\{\omega \in B \mid Y(\omega) = y\}\right)$.

Podmíněná distribuční funkce $F_{Y|B}$ náhodné veličiny Y za podmínky B:

$$F_{Y|B}(y) = P[Y \leq y|B] = \frac{P[Y \leq y, B]}{P[B]} \ ,$$

kde $P[Y \leq y, B] = P\left(\{\omega \in B \mid Y(\omega) \leq y\}\right)$.

1.3 Rozdělení podmíněné jevem

Podmíněná pravděpodobnostní funkce $p_{Y|B}$ *diskrétní* náhodné veličiny Y za podmínky

$$p_{Y|B}(y) = P[Y = y|B] = \frac{P[Y = y, B]}{P[B]} ,$$

kde $P[Y = y, B] = P(\{\omega \in \Omega \mid Y(\omega) = y, \omega \in B\}) = P(\{\omega \in B \mid Y(\omega) = y\})$.

Podmíněná distribuční funkce $F_{Y|B}$ náhodné veličiny Y za podmínky B:

$$F_{Y|B}(y) = P[Y \leq y|B] = \frac{P[Y \leq y, B]}{P[B]} ,$$

kde $P[Y \leq y, B] = P(\{\omega \in B \mid Y(\omega) \leq y\})$.

Pro daný jev B má všechny vlastnosti distribuční funkce.

Definována, i když Y není diskrétní

1.3 Rozdělení podmíněné jevem

Podmíněná pravděpodobnostní funkce $p_{Y|B}$ *diskrétní* náhodné veličiny Y za podmínky

$$p_{Y|B}(y) = P[Y = y|B] = \frac{P[Y = y, B]}{P[B]} \; ,$$

kde $P[Y = y, B] = P\left(\{\omega \in \Omega \mid Y(\omega) = y, \omega \in B\}\right) = P\left(\{\omega \in B \mid Y(\omega) = y\}\right)$.

Podmíněná distribuční funkce $F_{Y|B}$ náhodné veličiny Y za podmínky B:

$$F_{Y|B}(y) = P[Y \leq y|B] = \frac{P[Y \leq y, B]}{P[B]} \; ,$$

kde $P[Y \leq y, B] = P\left(\{\omega \in B \mid Y(\omega) \leq y\}\right)$.

Pro daný jev B má všechny vlastnosti distribuční funkce.

Definována, i když Y není diskrétní

$\Rightarrow$ **podmíněná kvantilová funkce** $q_{Y|B}$

$\Rightarrow$ **podmíněná hustota** $f_{Y|B}$ *spojité* náhodné veličiny

$$F_{Y|B}(y) = \int_0^y f_{Y|B}(t)\,\mathrm{d}t$$

$\Rightarrow$ **podmíněná hustota** $f_{Y|B}$ *spojité* náhodné veličiny

$$F_{Y|B}(y) = \int_0^y f_{Y|B}(t)\,\mathrm{d}t$$

$\Rightarrow$ **podmíněná střední hodnota** $\mathsf{E}(Y|B)$

$\Rightarrow$ **podmíněný rozptyl** $\mathsf{D}(Y|B)$

1.4 Příklady podmíněných rozdělení

Příklad: První hráč hodí čtyřmi stejnými mincemi; ty, na kterých padne líc, si ponechá jako výhru (náhodná veličina X). Poté druhý hráč hodí zbylými mincemi; ty, na kterých padne líc, tvoří jeho výhru (náhodná veličina Y). Popište sdružené rozdělení X, Y, marginální rozdělení X a Y, podmíněná rozdělení $Y|X$ a $X|Y$. U všech rozdělení stanovte (podmíněnou) střední hodnotu a (podmíněný) rozptyl.

1.4 Příklady podmíněných rozdělení

Příklad: První hráč hodí čtyřmi stejnými mincemi; ty, na kterých padne líc, si ponechá jako výhru (náhodná veličina X). Poté druhý hráč hodí zbylými mincemi; ty, na kterých padne líc, tvoří jeho výhru (náhodná veličina Y). Popište sdružené rozdělení X, Y, marginální rozdělení X a Y, podmíněná rozdělení $Y|X$ a $X|Y$. U všech rozdělení stanovte (podmíněnou) střední hodnotu a (podmíněný) rozptyl.

Řešení: (Čísla za tabulkami jsou pravděpodobnostní funkce marginálních rozdělení, což v případě sdruženého rozdělení jsou řádkové, resp. sloupcové součty.)

$p_{X,Y}$ $x \backslash y$	0	1	2	3	4	$p_X(x)$
0	$\frac{1}{256}$	$\frac{1}{64}$	$\frac{3}{128}$	$\frac{1}{64}$	$\frac{1}{256}$	$\frac{1}{16}$
1	$\frac{1}{32}$	$\frac{3}{32}$	$\frac{3}{32}$	$\frac{1}{32}$	0	$\frac{1}{4}$
2	$\frac{3}{32}$	$\frac{3}{16}$	$\frac{3}{32}$	0	0	$\frac{3}{8}$
3	$\frac{1}{8}$	$\frac{1}{8}$	0	0	0	$\frac{1}{4}$
4	$\frac{1}{16}$	0	0	0	0	$\frac{1}{16}$
$p_Y(y)$	$\frac{81}{256}$	$\frac{27}{64}$	$\frac{27}{128}$	$\frac{3}{64}$	$\frac{1}{256}$	

$$\mathsf{E}X = 2\,, \qquad \mathsf{E}Y = \frac{27}{64}\,1 + \frac{27}{128}\,2 + \frac{3}{64}\,3 + \frac{1}{256}\,4 = 1$$

$p_{Y\|X}$ $\quad y$ / x	0	1	2	3	4	$p_X(x)$	$\mathsf{E}(Y\|X=x)$
0	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$	$\frac{1}{16}$	2
1	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	0	$\frac{1}{4}$	$\frac{3}{2}$
2	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	0	0	$\frac{3}{8}$	1
3	$\frac{1}{2}$	$\frac{1}{2}$	0	0	0	$\frac{1}{4}$	$\frac{1}{2}$
4	1	0	0	0	0	$\frac{1}{16}$	0
$p_Y(y)$	$\frac{81}{256}$	$\frac{27}{64}$	$\frac{27}{128}$	$\frac{3}{64}$	$\frac{1}{256}$		

$$\mathsf{E}(\mathsf{E}(Y|X)) = \frac{1}{16}\,2 + \frac{1}{4}\,\frac{3}{2} + \frac{3}{8}\,1 + \frac{1}{4}\,\frac{1}{2} = 1 = \mathsf{E}Y$$

$p_{X\|Y}$ y / x	0	1	2	3	4	$p_X(x)$
0	$\frac{1}{81}$	$\frac{1}{27}$	$\frac{1}{9}$	$\frac{1}{3}$	1	$\frac{1}{16}$
1	$\frac{8}{81}$	$\frac{2}{9}$	$\frac{4}{9}$	$\frac{2}{3}$	0	$\frac{1}{4}$
2	$\frac{8}{27}$	$\frac{4}{9}$	$\frac{4}{9}$	0	0	$\frac{3}{8}$
3	$\frac{32}{81}$	$\frac{8}{27}$	0	0	0	$\frac{1}{4}$
4	$\frac{16}{81}$	0	0	0	0	$\frac{1}{16}$
$p_Y(y)$	$\frac{81}{256}$	$\frac{27}{64}$	$\frac{27}{128}$	$\frac{3}{64}$	$\frac{1}{256}$	
$\mathsf{E}(X\|Y=y)$	$\frac{8}{3}$	2	$\frac{4}{3}$	$\frac{2}{3}$	0	

$$\mathsf{E}(\mathsf{E}(X|Y)) = \frac{81}{256}\,\frac{8}{3} + \frac{27}{64}\,2 + \frac{27}{128}\,\frac{4}{3} + \frac{3}{64}\,\frac{2}{3} = 2 = \mathsf{E}X$$