M3121 Pravděpodobnost a statistika I

12. Střední hodnota náhodné veličiny

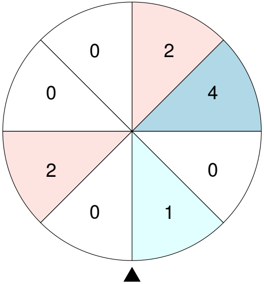
Jan Koláček (kolacek@math.muni.cz)

Ústav matematiky a statistiky, Přírodovědecká fakulta, Masarykova univerzita, Brno

Motivační příklad

Příklad 1

Uvažujme hru, kde účastník hry roztočí „kolo štěstí“. Každé pole tohoto kola definuje výhru (v Kč), která bude vyplacena hráči v případě, že na toto pole ukazuje šipka po zastavení kola. Za každou hru zaplatí hráč provozovateli 1 Kč. Budeme hrát? Tj. jaká je „očekávaná“ výhra?



Y ... **zisk** z jedné hry, X ... **částka**, kterou si vytočíme
Zřejmě $Y = X - 1$.

X	0	1	2	4
$p(x)$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$

„Očekávaná“ výhra

$$EY = EX - 1 = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{8} + 2 \cdot \frac{1}{4} + 4 \cdot \frac{1}{8} - 1$$
$$= \frac{1}{8} = \boxed{0{,}125} > 0.$$

Střední hodnota

Definice 1

Nechť X je náhodná veličina definovaná na $(\Omega, \mathcal{A}, P)$ a nechť existuje integrál $\int\limits_{\Omega} X(\omega)\, dP(\omega) < \infty$. Potom číslo

$$EX = \int\limits_{\Omega} X(\omega)\, dP(\omega)$$

nazýváme **střední hodnotou náhodné veličiny** X (**Expected Value, Mean**).

Značení: $\mathcal{L}_1(\Omega, \mathcal{A}, P)$ ... množina všech náhodných veličin definovaných na $(\Omega, \mathcal{A}, P)$, které mají **konečné střední hodnoty**.

Střední hodnota

Věta 2 (Vlastnosti)

Nechť X, X_1, X_2 *jsou náhodné veličiny definované na pravděpodobnostním prostoru* $(\Omega, \mathcal{A}, P)$, $a, a_1, a_2 \in \mathbb{R}$. *Potom*

- EX *existuje* $\quad \Leftrightarrow \quad E|X|$ *existuje.*
- *Jestliže* $P(X = a) = 1 \quad \Rightarrow \quad EX = a$.
- *Existují-li* EX_1, EX_2 $\quad \Rightarrow \quad E(a_1X_1 + a_2X_2) = a_1EX_1 + a_2EX_2$.
- *Nechť existují* EX_1, EX_2 *a platí* $X_1 \leq X_2 \quad \Rightarrow \quad EX_1 \leq EX_2$.
- *Nechť* $|X_1| \leq X_2$ *a* EX_2 *existuje* $\quad \Rightarrow \quad EX_1$ *existuje.*
- *Nechť* $P(X \geq 0) = 1 \quad \Rightarrow \quad EX \geq 0$.

Věta 3 (Věta o přenosu integrace)

Nechť $\mathbf{X} = (X_1, \ldots, X_n)'$ *je náhodný vektor definovaný na* $(\Omega, \mathcal{A}, P)$, g *je borelovsky měřitelná na* $(\mathbb{R}^n, \mathcal{B}^n)$, $P_{\mathbf{X}}$ *je rozdělení pstí náhodného vektoru* $\mathbf{X}$. *Pak*

$$\int\limits_{\Omega} g(\mathbf{X}(\omega))dP(\omega) = \int\limits_{\mathbb{R}^n} g(\mathbf{x})dP_{\mathbf{X}}(\mathbf{x})$$

Střední hodnota

Věta 4 (Výpočet)

Nechť X *je náhodná veličina na* $(\Omega, \mathcal{A}, P)$. *Pak platí*

$X \in \mathcal{L}_1(\Omega, \mathcal{A}, P) \Leftrightarrow \int\limits_{-\infty}^{\infty} x dF(x) < \infty$. *V tomto případě je* $\boxed{EX = \int\limits_{-\infty}^{\infty} x dF(x)}$.

- *Nechť* $X \sim (M, p)$ *je diskrétního typu, pak platí* $X \in \mathcal{L}_1(\Omega, \mathcal{A}, P) \Leftrightarrow \sum\limits_{x \in M} xp(x)$ *absolutně konverguje. V tomto případě* $\boxed{EX = \sum\limits_{x \in M} xp(x)}$.
- *Nechť* $X \sim f(x)$ *je absolutně spojitého typu, pak platí* $X \in \mathcal{L}_1(\Omega, \mathcal{A}, P) \Leftrightarrow xf(x)$ *je integrovatelná vzhledem k Lebesgueově míře.*

  V tomto případě je $\boxed{EX = \int\limits_{-\infty}^{\infty} xf(x)dx}$.

Střední hodnota transformované n. v.

Věta 5

Nechť $h(x)$ je borelovská funkce. Potom střední hodnota transformované náhodné veličiny $Y = h(X)$ existuje právě když existuje a je konečný integrál $\int\limits_{-\infty}^{\infty} h(x)dF(x) < \infty$. V tomto případě platí $\boxed{EY = Eh(X) = \int\limits_{-\infty}^{\infty} h(x)dF(x)}$.

- *Nechť $X \sim (M, p)$ je diskrétního typu, pak platí $Y \in \mathcal{L}_1(\Omega, \mathcal{A}, P) \Leftrightarrow \sum\limits_{x \in M} h(x)p(x)$ absolutně konverguje. V tomto případě $\boxed{EY = Eh(X) = \sum\limits_{x \in M} h(x)p(x)}$.*
- *Nechť $X \sim f(x)$ je absolutně spojitého typu. Potom EY existuje právě když je funkce $h(x)f(x)$ integrovatelná vzhledem k Lebesgueově míře a přitom platí $\boxed{EY = Eh(X) = \int\limits_{-\infty}^{\infty} h(x)f(x)dx}$, tj. $EY = Eh(X) \in \mathcal{L}_1(\Omega, \mathcal{A}, P) \Leftrightarrow h(x)f(x)$ je integrovatelná vzhledem k Lebesgueově míře.*

Př. $Y = X^2 \Rightarrow EY = EX^2 = \int x^2 f(x)dx$ nebo $EY = EX^2 = \sum x^2 p(x)$

Příklad

Příklad 2 (Střední hodnota Alternativního rozdělení)

Mějme náhodnou veličinu $X \sim A(\theta)$, $\theta \in (0,1)$ s pravděpodobnostní funkcí

$$p(x) = \begin{cases} 1-\theta & x = 0 \\ \theta & x = 1 \\ 0 & \textit{jinak.} \end{cases}$$

Vypočtěte střední hodnotu.

$$\boxed{EX} = \sum_{x=0}^{1} xp(x) = 0 \cdot (1-\theta) + 1 \cdot \theta = \boxed{\theta}.$$

Příklad

Příklad 3 (Střední hodnota Binomického rozdělení)

Mějme náhodnou veličinu $X \sim Bi(n,\theta)$, $n \in \mathbb{N}$, $\theta \in (0,1)$ *s pravděpodobnostní funkcí*

$$p(x) = \begin{cases} \binom{n}{x}\theta^x(1-\theta)^{n-x} & x \in M = \{0,1,\ldots,n\} \\ 0 & \textit{jinak.} \end{cases}$$

Vypočtěte střední hodnotu.

$$X = \sum_{i=1}^{n} Y_i, \qquad Y_i \sim A(\theta),\ i = 1,\ldots,n$$

$$\boxed{EX} = E\sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} EY_i = \boxed{n\theta}.$$

Nebo

$$\boxed{EX} = \sum_{x=0}^{n} xp(x) = \sum_{x=0}^{n} x\binom{n}{x}\theta^x(1-\theta)^{n-x} = \cdots = \boxed{n\theta}.$$

Příklad

Příklad 4

Biatlonista střílí nezávisle na sobě do terče, přičemž pravděpodobnost zásahu při každém výstřelu je $2/3$. Jaká je očekávaná hodnota počtu zasažených terčů ze 300 pokusů?

X ... počet zásahů, $X \sim Bi(300, 2/3)$

$$\boxed{EX} = n\theta = 300 \cdot 2/3 = \boxed{\mathbf{200}}.$$

Příklad

Příklad 5 (Střední hodnota Poissonova rozdělení)

Mějme náhodnou veličinu $X \sim Po(\lambda)$, $\lambda > 0$ *s pravděpodobnostní funkcí*

$$p(x) = \begin{cases} e^{-\lambda}\frac{\lambda^x}{x!} & x \in M = \{0, 1, \ldots\} \\ 0 & \textit{jinak.} \end{cases}$$

Vypočtěte střední hodnotu.

$$\boxed{EX} = \sum_{x=0}^{\infty} xp(x) = \sum_{x=0}^{\infty} xe^{-\lambda}\frac{\lambda^x}{x!} = e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^x}{(x-1)!} = \lambda e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!}$$

$$= \left|subst.\ y = x-1\right| = \lambda \underbrace{\sum_{y=0}^{\infty} e^{-\lambda}\frac{\lambda^y}{y!}}_{1=\sum\limits_{y\in M} p(y)} = \boxed{\lambda}.$$

Příklad

Příklad 6 (Střední hodnota Rovnoměrného rozdělení)

Mějme náhodnou veličinu $X \sim Rs(a,b)$, $a < b$, $a,b \in \mathbb{R}$ *s hustotou*

$$f(x) = \begin{cases} \frac{1}{b-a} & x \in (a,b) \\ 0 & \textit{jinak.} \end{cases}$$

Vypočtěte střední hodnotu.

$$\boxed{EX} = \int_{-\infty}^{\infty} xf(x)dx = \int_a^b x\frac{1}{b-a}dx = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b = \frac{1}{b-a}\cdot\frac{b^2-a^2}{2}$$

$$= \frac{1}{b-a}\frac{(b-a)(b+a)}{2} = \boxed{\frac{a+b}{2}}.$$

Příklad

Příklad 7 (Střední hodnota Normálního rozdělení)

Mějme náhodnou veličinu s normálním rozdělením $X \sim N(\mu, \sigma^2)$ *s hustotou*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}.$$

$$EX = \int_{-\infty}^{\infty} x f(x) dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} x e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx.$$

Položíme-li $y = \frac{x-\mu}{\sigma}$, tj. $x = \sigma y + \mu$ a $dx = \sigma dy$, pak

$$\boxed{EX} = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} x e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma y + \mu) e^{-\frac{1}{2}y^2} dy$$

$$= \frac{\sigma}{\sqrt{2\pi}} \underbrace{\int_{-\infty}^{\infty} y e^{-\frac{1}{2}y^2} dy}_{=0 \text{ (lichá funkce)}} + \mu \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy}_{=1 \text{ (hustota } Y \sim N(0,1))} = \boxed{\mu}.$$

Příklad

Příklad 8

Náhodná veličina X má binomické rozdělení $X \sim Bi(n, \theta)$. Vypočtěte střední hodnotu náhodné veličiny $Y = e^{2X}$.

$$
\begin{aligned}
EY = E(e^{2X}) &= \sum_{x=0}^{n} e^{2x} \binom{n}{x} \theta^x (1-\theta)^{n-x} \\
&= \sum_{x=0}^{n} \binom{n}{x} \left(\theta e^2\right)^x (1-\theta)^{n-x} \\
&\overset{\text{binom. věta}}{=} \left(\theta e^2 + 1 - \theta\right)^n
\end{aligned}
$$

Střední hodnota

Věta 6 (Střední hodnota součinu nezávislých náhodných veličin)

Nechť $X_1, \ldots, X_n$ *jsou* **nezávislé** *náhodné veličiny definované na* $(\Omega, \mathcal{A}, P)$ *a nechť existují střední hodnoty* $EX_1, \ldots, EX_n$*. Pak platí*

$$E\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} EX_i.$$

Podmíněná střední hodnota

$\mathbf{Z} = (Y, X)'$ ... sdružená distribuční funkce $F(y, x)$
X a Y ... marginální distribuční funkce $F_X(x)$, $F_Y(y)$

Definice 7

Nechť pro $\forall B \in \mathcal{B}$ a $\forall y \in \mathbb{R}$ existuje funkce $F_{Y|X}(y|x)$ tak, že platí

$$P(Y \leq y, X \in B) = \int_B F_{Y|X}(y|x) dF_X(x).$$

Pak funkci $F_{Y|X}(y|x)$ nazveme **podmíněnou distribuční funkcí** náhodné veličiny Y při daném $X = x$.

Podmíněná střední hodnota

Definice 8

Nechť pro $\forall x \in \mathbb{R}$ existuje konečný integrál

$$h(x) = E(Y|X = x) = \int_{\mathbb{R}} y dF_{Y|X}(y|x).$$

Pak funkci $h(x)$ nazveme **podmíněnou střední hodnotou náhodné veličiny** Y **za podmínky** $X = x$.
Symbolem

$$E(Y|X) = h(X)$$

definujeme náhodnou veličinu, kterou nazýváme **podmíněnou střední hodnotou náhodné veličiny** Y **při daném** X.

Podmíněná střední hodnota

Příklad 9 (Diskrétní případ)

$\mathbf{Z} = (Y, X)' \sim (M; p(y,x))$, $X \sim (M_X; p_X(x))$, $Y \sim (M_Y; p_Y(y))$

Podmíněná distribuční funkce

$$F(y|x) = \begin{cases} \sum\limits_{t \leq y} \frac{p(t,x)}{p_X(x)} & \text{pro } x \in M_X, \\ 0 & \text{pro } x \in \mathbb{R} \setminus M_X. \end{cases}$$

Podmíněná pravděpodobnostní funkce

$$p(y|x) = \begin{cases} \frac{p(y,x)}{p_X(x)} & \text{pro } x \in M_X, \\ 0 & \text{pro } x \in \mathbb{R} \setminus M_X. \end{cases}$$

Podmíněná střední hodnota

$$E(Y|X=x) = \sum_{y \in M_Y} y\, p(y|x) = \sum_{y \in M_Y} y \frac{p(y,x)}{p_X(x)}, \quad \text{pro } \forall x \in M_X.$$

Podmíněná střední hodnota

Příklad 10 (Spojitý případ)

Označme $M_X = \{x \in \mathbb{R} : f_X(x) > 0\}$, $M_Y = \{y \in \mathbb{R} : f_Y(y) > 0\}$.

Podmíněná distribuční funkce

$$F(y|x) = \begin{cases} \int\limits_{-\infty}^{y} \frac{f(t,x)}{f_X(x)} dt & \text{pro } x \in M_X, \\ 0 & \text{pro } x \in \mathbb{R} \setminus M_X \end{cases}$$

Podmíněná hustota

$$f(y|x) = \begin{cases} \frac{f(y,x)}{f_X(x)} & \text{pro } x \in M_X, \\ 0 & \text{pro } x \in \mathbb{R} \setminus M_X. \end{cases}$$

Podmíněná střední hodnota

$$E(Y|X = x) = \int_{\mathbb{R}} y f(y|x) dy = \int_{\mathbb{R}} y \frac{f(y,x)}{f_X(x)} dy, \quad \text{pro } \forall x \in M_X.$$

Podmíněná střední hodnota

Vlastnosti

- Nechť Y_1, Y_2, X jsou náhodné veličiny a a_0, a_1, a_2 jsou reálné konstanty, pak pokud střední hodnoty EY_1, EY_2 existují, platí

$$E(a_0 + a_1Y_1 + a_2Y_2|X) = a_0 + a_1E(Y_1|X) + a_2E(Y_2|X).$$

- Nechť X, Y jsou náhodné veličiny a střední hodnota EY existuje, pak

$$E\left[E(Y|X)\right] = EY.$$

Podmíněná střední hodnota

Příklad 11

Počet vajíček X nakladených samičkou Octomilky obecné se řídí Poissonovým rozdělením s parametrem λ. Počet vylíhnutých jedinců Y v případě, že bylo nakladeno x vajíček, tj. $Y|X = x$ se řídí binomickým rozdělením $Bi(x, \theta)$. Jakým rozdělením se řídí náhodná veličina Y?

$$
\begin{aligned}
p_Y(y) &= \sum_{x=-\infty}^{\infty} p(x,y) = \sum_{x=-\infty}^{\infty} p(y|x) p_X(x) = \sum_{x=y}^{\infty} \binom{x}{y} \theta^y (1-\theta)^{x-y} e^{-\lambda} \frac{\lambda^x}{x!} \\
&= e^{-\lambda} \frac{(\theta\lambda)^y}{y!} \sum_{x=y}^{\infty} \frac{x!}{(x-y)!} (1-\theta)^{x-y} \frac{\lambda^{x-y}}{x!} \\
&= e^{-\lambda} \frac{(\theta\lambda)^y}{y!} \sum_{x=0}^{\infty} \frac{1}{x!} (1-\theta)^x \lambda^x \\
&= e^{-\lambda} \frac{(\theta\lambda)^y}{y!} e^{(1-\theta)\lambda} = e^{-\theta\lambda} \frac{(\theta\lambda)^y}{y!} \Rightarrow \boxed{Y \sim Po(\theta\lambda)}
\end{aligned}
$$