

Střední hodnota n.v.

Definice střední hodnoty n.v. X : EX

Uvažujeme pravděpodobnostní prostor $(S, \mathcal{B}, P)$ a uvažujeme X je n.v. na tomto prostoru. Pak definujeme :

$$EX := \int_S X dP = \int_S X(e) dP(e). \quad (*)$$

Příklad !!! (i) Je-li X jednoduchá n.v., $X = \sum_{i=1}^n X_i \cdot \chi_{\{X=X_i\}}$ pak $EX = \sum_{i=1}^n X_i \cdot P(\{X=X_i\})$.

(ii) Je-li X diskrétní n.v. a $x_1, x_2, \dots, x_i, \dots$ jsou funkční hodnoty n.v. X , pak máme :

$$EX = \sum_i x_i \cdot P(\{X=x_i\}) = \sum_i x_i \cdot p_X(x_i).$$

(iii) Jestliže X je absolutně sp. n.v., f_X je hustota n.v. X , pak $EX = \int_{-\infty}^{\infty} x \cdot f_X(x) dx$.

Trváme uvažovat v předchozích příkladech nebudeme dožadovat. Zjednodušeně v příkladu (iii) vyžaduje hustotu f_X z jedné strany a integrálu.

Odpovědi definice :

definice čísla $\int_S X dP$ prohrává v několika krocích :

1. krok : X je jednoduchá n.v., $X = \sum_{i=1}^n x_i \cdot \chi_{A_i}$, kde

$$A_i = X^{-1}(\{x_i\}) = \{e \in S : X(e) = x_i\}.$$

Pak definujeme $EX = \sum_{i=1}^n x_i \cdot P(A_i) = \sum_{i=1}^n x_i \cdot p_X(x_i)$.

2. krok : H_X = obraz hodnot X = $\{x_1, \dots, x_n\}$.

3. krok : X je ustupková n.v. ($X \geq 0$ na S)

Sestrojíme $n \in \mathbb{N}$ tzv. mřížku aproximací $X_n : S \rightarrow \mathbb{R}^+$

takže : rozdělíme interval $\langle 0, a \rangle$ na 2^n intervalů :

$$I_k^n = \left[\frac{k-1}{2^n} a, \frac{k}{2^n} a \right) \text{ o délce } \frac{1}{2^n} a$$

kde $k = 1, \dots, n \cdot 2^n \Rightarrow \langle 0, a \rangle = \bigcup_{k=1}^{n \cdot 2^n} I_k^n$

Položíme-li $I_{n \cdot 2^n + 1}^n := \langle a, \infty \rangle$.

Potom dojdeme $\langle 0, a \rangle = \bigcup_{k=1}^{n \cdot 2^n} I_k^n$; označme jako I_k^n

její kanonický bod intervalu I_k^n .

Nyní definujeme $X_n : S \rightarrow \langle 0, a \rangle$ takto

$$X_n(e) = l(I_k^n), \text{ pokud } X(e) \in I_k^n.$$

FAKT 1 : $\forall e \in S : \{X_n(e)\}_{n=1}^{\infty}$ je neklesající posl.

FAKT 2 : $\forall e \in S : X_n(e) \rightarrow X(e)$. (X je bodová limitou posl. X_n).

FAKT 3 : $\forall m \in \mathbb{N}$ je X_n jednoduchou n.v. na S .

FAKT 4 : Posloupnost $\{EX_n\}_{n=1}^{\infty}$ je neklesající.

Definice EX : $EX := \lim_{n \rightarrow \infty} EX_n \in \langle 0, a \rangle \cup \{+\infty\}$.

Lemma : Je-li $X : S \rightarrow \langle 0, a \rangle$ n.v. a $\{X_n\}$ je

neklesající posloupnost ustupkových n.v. tak že platí :

(a) $\forall e \in S : \{X_n(e)\}$ je neklesající posl.

(b) $\forall e \in S : X_n(e) \rightarrow X(e)$ pro $n \rightarrow \infty$.

Pak $EX_n \rightarrow EX$ pro $n \rightarrow \infty$.

3. krok : Někdy $X : S \rightarrow \mathbb{R}$ je lib. n.v. Položíme :

$$\forall e \in S : X^+(e) = \max(X(e), 0), \quad X^-(e) = \max(-X(e), 0).$$

Takto jsme definovali fce $X^+ : S \rightarrow \mathbb{R}^+$ a $X^- : S \rightarrow \mathbb{R}^+$.

Jako cv. dokažte :

a) $\forall e \in S : X(e) = X^+(e) - X^-(e) ; |X(e)| = X^+(e) + X^-(e)$.

b) Funkce X^+ a X^- jsou ustupkové n.v. na S .

Definice EX : předpokládáme, že není pravda :

$$EX^+ = +\infty \wedge EX^- = +\infty$$

Potom je to pravda, $EX^+ - EX^-$ je ve tvaru $a - \infty$, pak říkáme, že EX neexistuje !

V opačném případě definujeme :

$$EX := \int_S X dP := EX^+ - EX^-.$$

Lemma : $EX \in \mathbb{R} \Leftrightarrow EX^+ \in \mathbb{R} \wedge EX^- \in \mathbb{R}$;

$$EX \in \mathbb{R} \Leftrightarrow E(|X|) < \infty.$$

D. cv. □

Lemma : Jestliže X, Y n.v. na S , pak jejich součet

$$Z = X + Y \text{ je n.v. na } S \text{ a}$$

$$EZ = EX + EY.$$

Poznámka. V tvrzení předchozí lemmatu předpokládáme, že

má práva strana $EX + EY$ smysl. Tj., že EX, EY

existují a součet není ve tvaru $\infty - \infty$.

VĚTA : Jestliže $X_1, \dots, X_n$ n.v. na S a $Z = X_1 + \dots + X_n$.

Potom je Z n.v. na S a platí :

$$EZ = EX_1 + \dots + EX_n,$$

ne-li, pravá strana smysl.

Důkaz. Někdy-li řečeno jinak, definujeme $0 \cdot \infty = \infty \cdot 0 =$

$= 0$.

VĚTA : $\forall a \in \mathbb{R} \wedge \forall$ n.v. X na S platí, že ex-kl. EX ,

potom :

$$E(aX) = a \cdot EX.$$

Nuž uvažujeme další vlastnost střední hodnoty, definujeme

pojmy (stochastické) nezávislosti náhodných veličin.

Definice : Někdy $X_1, \dots, X_n$ jsou n.v. na $(S, \mathcal{B}, P)$. Pak

říkáme, že jsou tyto n.v. nezávislé, pokud pro každé

Borelovské množiny $B_1, \dots, B_n$ v $\mathbb{R}$ platí :

$$\begin{aligned} P(\{X_1 \in B_1, \dots, X_n \in B_n\}) &= \\ &= P(\{e \in S : e \in \bigcap_{i=1}^n X_i^{-1}(B_i)\}) = \\ &= P(\{X_1 \in B_1\}) \cdot P(\{X_2 \in B_2\}) \cdot \dots \cdot P(\{X_n \in B_n\}) \\ &= \prod_{i=1}^n P(\{X_i \in B_i\}). \end{aligned}$$

VĚTA : Mějme danou n.v. $X_1, \dots, X_n$ na pravd. prostoru $(S, \mathcal{B}, P)$,

kde jsou nezávislé.

a) Jestliže $X_1 \geq 0, \dots, X_n \geq 0$, pak

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

b) Jestliže $\forall i=1, \dots, n$ jsou střední hodnoty EX_i konečné,

pak platí :

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

b)

D. cv. Pověz : jednoduché vel. X, Y .

Někdy $X = \sum_{i=1}^n x_i \cdot \chi_{\{X=x_i\}}, Y = \sum_{j=1}^m y_j \cdot \chi_{\{Y=y_j\}}$

$$\Rightarrow X \cdot Y = \left(\sum_{i=1}^n x_i \cdot \chi_{\{X=x_i\}} \right) \left(\sum_{j=1}^m y_j \cdot \chi_{\{Y=y_j\}} \right) = \begin{cases} x_i \cdot y_j & \text{pr } e \in A \cap B \\ 0 & \text{pr } e \notin A \cap B \end{cases} \Rightarrow X \cdot Y = \chi_{A \cap B}$$

$$= \sum_{x_i, y_j} x_i \cdot y_j \cdot \chi_{\{X=x_i\}} \cdot \chi_{\{Y=y_j\}} = \sum_{i,j} x_i \cdot y_j \cdot \chi_{\{X=x_i \cap Y=y_j\}} \Rightarrow$$

$$\Rightarrow E(X \cdot Y) = \sum_{i,j} x_i \cdot y_j \cdot P(\{X=x_i \cap Y=y_j\}) = \sum_{i,j} x_i \cdot y_j \cdot P(\{X=x_i\}) \cdot P(\{Y=y_j\}) = \left(\sum_i x_i \cdot P(\{X=x_i\}) \right) \cdot \left(\sum_j y_j \cdot P(\{Y=y_j\}) \right) = EX \cdot EY. \quad \checkmark$$

Definice rozptylu : Předp. je X je n.v. na $(S, \mathcal{B}, P)$.

Označme $\mu = EX$. Dále pak hodnotu $DX = E((X-\mu)^2)$

nazýváme rozptylem (dispert) n.v. X . Označme $\sigma^2 = \sigma^2(X) =$

$$= \sigma_X^2 = DX. \text{ Znaménko } \sigma \text{ disperze } \sqrt{DX} = \sigma \text{ a tedy}$$

nazýváme standardní nebo směrodatnou odchylkou n.v. X .

Lemma $DX = EX^2 - (EX)^2$.

$$\text{D. cv. } DX = E(X^2 - 2\mu X + \mu^2) = EX^2 - 2\mu EX + E\mu^2 =$$

$$= EX^2 - 2(EX)^2 + \mu^2 = EX^2 - 2(EX)^2 + (EX)^2 =$$

$$= EX^2 - (EX)^2. \quad \square$$

Cvičení. Předp. je X n.v. mající konečnou mřížku pravdep.

Tj. X můžeme chápat jako veličinu označující počet

výsledků v n opakováních n.h. nezávislých pokusů, přičemž

předpokládáme, že je rovná p .

$$\text{tedy } p_X(k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}.$$

Ukažte, že pro střední nebo očekávanou hodnotu EX platí :

$$EX = n \cdot p.$$

Součteme dle k dokazu tzv. slabého zákona velkých čísel :

VĚTA Mějme danou n.v. $X_1, X_2, \dots, X_n$ na $(S, \mathcal{B}, P)$ a

předp. je jsou tyto n.v. nezávislé a mají každá konečnou

střední hodnotu a rozptyl. Dále předpokládáme, že každá

konstanta $M > 0$ taková, že $\forall i : \sigma_i = \sigma(X_i) \leq M$ a

a $\forall n$ platíme :

$$Y_n = \sum_{i=1}^n X_i.$$

Pak $\forall \varepsilon > 0$ platí :

$$P\left(\left|\frac{Y_n - E(Y_n)}{n}\right| \geq \varepsilon\right) \rightarrow 0 \text{ pro } n \rightarrow \infty.$$

VĚTA (Čebyšev) a) Někdy X je ustupková n.v. na $(S, \mathcal{B}, P)$

a $\varepsilon > 0$. Potom platí nerovnost :

$$P(\{X \geq \varepsilon\}) \leq \frac{EX}{\varepsilon}.$$

b) Někdy X je libovolná n.v. na $(S, \mathcal{B}, P)$, $c \in \mathbb{R}$, $m > 0$, $\varepsilon > 0$.

Pak

$$P(\{|X-c| \geq \varepsilon\}) \leq \frac{E(X-c)^2}{\varepsilon^2}.$$

c) Někdy X je lib. n.v. na $(S, \mathcal{B}, P)$ a uvaž. má X konečnou

střední hodnotu μ a rozptyl $\sigma^2 > 0$ a $k > 0$. Potom platí :

$$P(\{|X-\mu| \geq k \cdot \sigma\}) \leq \frac{1}{k^2}.$$

Lemma. Někdy X a Y jsou dvě n.v. na $(S, \mathcal{B}, P)$ a uvaž.

$\forall e \in S : X(e) \leq Y(e)$. Pak

$$EX \leq EY.$$

Důkaz. Cvičení. □

Důkaz Čebyševovy věty

a) Nyní přejdeme k vlastnostem důkaz Čebyševovy nerovnosti :

položme $A_\varepsilon = \{e \in S : X(e) \geq \varepsilon\}$. Pak $X \geq X \cdot \chi_{A_\varepsilon}$, kde

χ_{A_ε} je charakteristická funkce množiny A_ε . Dále platí :

$$X \cdot \chi_{A_\varepsilon} \geq \varepsilon \cdot \chi_{A_\varepsilon}.$$

Takto vyplývá aplikací lemmatu o sčítání nerovností :

$$EX \geq E(X \cdot \chi_{A_\varepsilon}) \geq E(\varepsilon \cdot \chi_{A_\varepsilon}) = \varepsilon \cdot E(\chi_{A_\varepsilon}) =$$

$$= \varepsilon \cdot P(A_\varepsilon) = \varepsilon \cdot P(\{X \geq \varepsilon\}) \Rightarrow$$

$$\Rightarrow P(\{X \geq \varepsilon\}) \leq \frac{1}{\varepsilon} \cdot EX. \quad \square$$

b) $P(\{|X-c| \geq \varepsilon\}) = P(\{|X-c| \geq \varepsilon^m\}) \leq \frac{E(X-c)^2}{\varepsilon^{2m}}$,

kde poslední nerovnost je důsledkem tvrzení části a). □

c) Tvrzení plyne z b) kde položíme $c = \mu$, $\varepsilon = k \cdot \sigma$ a $m = 2$. □

VĚTA (Vlastnosti rozptylu n.v.)

a) Bud' X n.v. na $(S, \mathcal{B}, P)$ a uvaž. má rozptyl σ^2 (připouštíme

i. $\sigma = \infty$), $a, b \in \mathbb{R}$. Potom platí :

$$D(aX+b) = a^2 \sigma^2.$$

b) Někdy jsou $X_1, \dots, X_n$ nezávislé n.v. na $(S, \mathcal{B}, P)$ a

uvaž. $\forall i=1, \dots, n$ je $EX_i < \infty$. Potom :

$$D\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n D(X_i).$$

Důkaz slabého zákona velkých čísel

2. předpokládáme nyní například, že $\forall n \in \mathbb{N}$:

$$D(Y_n) = \sum_{i=1}^n D(X_i) \leq n \cdot M.$$

Dále $\forall n \in \mathbb{N}$ uvažujeme n.v. $Z = \frac{Y_n - E(Y_n)}{n}$. Nyní v důsledku

Čebyševovy věty (b) máme :

$$P\left(\left|\frac{Z - 0}{1} \right| \geq \varepsilon\right) = P\left(\left|\frac{Y_n - E(Y_n)}{n}\right| \geq \varepsilon\right) \leq$$

$$\leq \frac{1}{\varepsilon^2} E\left[\left(\frac{Y_n - E(Y_n)}{n}\right)^2\right] = \frac{1}{n^2 \varepsilon^2} E[(Y_n - E(Y_n))^2] =$$

$$= \frac{D(Y_n)}{n^2 \varepsilon^2} \leq \frac{n \cdot M}{n^2 \varepsilon^2} = \frac{M}{\varepsilon^2} \cdot \frac{1}{n} \rightarrow 0 \text{ pro } n \rightarrow \infty. \quad \square$$

Poznámka. Předpokládáme hypotézu, že každou nekonstantní fci

opakovat nezávislých pokusů n krát a p je pravděpodobnost, že

padne panna. Potom po n krocích panna :

$$f(n) = \frac{\text{počet pokusů kdy padla panna}}{n},$$

pak očekáváme, že $f(n) \rightarrow p$ pro $n \rightarrow \infty$.

Tento předpoklad je použitelný silněji tvrzení než by se dolo-

žovat zslabého zákona velkých čísel. Důsledkem tohoto

tvrzení je pouze to, že pro dostatečně velké $n \in \mathbb{N}$ je téměř

jisté, že relativní frekvence $f(n)$ je blízká k ex-kl. p .

Je-li $X_i = \begin{cases} 1 & \text{pokud padla panna} \\ 0 & \text{v opačném případě} \end{cases}$

$$E(Y_n) = E(X_1 + \dots + X_n) = \sum_{i=1}^n E(X_i) = n \cdot p, \text{ kde } \forall i : E(X_i) = p.$$

$$\frac{Y_n - E(Y_n)}{n} = \frac{X_1 + \dots + X_n}{n} - p.$$