8 Modus, medián, stredná hodnota, pravidlá počítania so strednou hodnotou.

8.1 Charakteristiky náhodnej premennej

Úplnú predstavu o náhodnej premennej dostaneme, keď uvedieme všetky hodnoty, ktoré môže nadobúdať a ich pravdepodobnosti. Niektoré vlastnosti náhodnej premennej však stačí popísať iba jedným číslom. Charakteristiky náhodnej premennej sú čísla, ktoré určujú niektoré vlastnosti jej pravdepodobnostnej funkcie, j

Stredná hodnota, medián a modus sú charakteristiky polohy, umiestnenia, alebo centra náhodnej premennej.

Modus (Mod_X)

Pre náhodnú premennú X s hodnotami $x_1, x_2, \ldots$ definujeme **modus** tejto náhodnej premennej, teda $Mod(X)$, ako tie hodnoty náhodnej premennej x_m, pre ktoré je ich PDF maximálne. Teda ak $Mod(x) = x_m$, potom platí:

$$PDF(x_m) = \max_{x_j \in X} PDF(X = x_j)$$

Zjednodušene povedané, modus je tá hodnota náhodnej premennej, ktorá má najväčšiu pravdepodobnosť.

Modus náhodnej premennej môže nadobudnúť viac hodnôt.

Medián (Med_X)

Pre náhodnú premennú X s hodnotami $x_1, x_2, \ldots$ definujeme **medián** tejto náhodnej premennej, teda $Med(X)$, ako takú hodnotu náhodnej premennej x_n, pre ktorú platí:

$$\Pr(X \geq Med_X) \geq \frac{1}{2}$$
$$\Pr(X \leq Med_X) \geq \frac{1}{2}$$

Zjednodušene povedané, medián je tá hodnota, po ktorú súčet pravdepodobností dosiahne polovicu.

Príklad 8.1 *Naprostá väčšina áut má podpriemerný počet kolies.*

Riešenie: Autá majú štyri, ale niekedy aj viac kolies. Naprostá väčšina áut však má 4 kolesá. Keďže existuje pár áut s väčším počtom kolies, je priemerný počet kolies áut o niečo málo viac ako 4. Preto všetky štvorkolesové autá, ktorých je väčšina, majú menej kolies ako je priemer. Teda podpriemerný počet kolies. □

V predchádzajúcom príklade je hodnota, ktorú si ľudia predstavia pod pojmom priemer rovná skôr mediánu, než priemeru. Medián počtu kolies auta je skutočne 4, lebo určite viac ako polovica áut má 4 kolesá.

Naopak v nasledujúcom príklade je rozumnejšie uvažovať o priemere, než o mediáne.

Príklad 8.2 *Meriame teplotu 5 prístrojov v miestnosti, ale ako výsledok dostaneme iba jeden údaj, medián.*

Z hodnôt

$$22;\ 25;\ 26;\ 30;\ 32$$

je medián hodnota 26.

Z hodnôt

$$22;\ 25;\ 26;\ 50;\ 82$$

je medián hodnota 26.

Aj z hodnôt

$$22;\ 25;\ 26;\ 50;\ 182$$

je medián hodnota 26 a to už v miestnosti horí.

Riešenie: Užitočnejšiu informáciu o teplote prístrojov by sme dostali, keby sme ako výsledok využili aritmetický priemer. □

Stredná hodnota náhodnej premennej (EX)

Stredná hodnota (*vážený priemer, mean, expected value, očakávaná hodnota*) náhodnej premennej X s hodnotami $x_1, x_2, \ldots$ je číslo, pre ktoré platí

$$E(X) = \sum_{x_i \in X} x_i \cdot \Pr(X = x_i) = \sum_{x_i \in X} x_i \cdot PDF(x_i)$$

Zjednodušene povedané, stredná hodnota náhodnej premennej je akýsi vážený priemer jej hodnôt, pričom váha každej hodnoty je jej pravdepodobnosť.

Stredná hodnota nemusí byť hodnotou náhodnej premennej.

Príklad 8.3 *Pri pokladni v hypermarkete bude zákazník prvý v poradí s pravdepodobnosťou 0.4, druhý v poradí s pravdepodobnosťou 0.3, tretí v poradí s pravdepodobnosťou 0.2, štvrtý v poradí s pravdepodobnosťou 0.1. Ak by mal byť viac ako štvrtý v poradí, tak ku pokladni nepôjde.*
Na ktorom mieste v poradí bude zákazník v priemere?

Riešenie: Vypočítame vážený priemer z umiestnení v rade pri pokladni. Prvé miesto v rade má váhu 0.4, druhé 0.3, tretie 0.2 a štvrté 0.1. Iné miesta sa nevyskytujú. Vážený priemer týchto umiestnení je

$$1 \cdot 0.4 + 2 \cdot 0.3 + 3 \cdot 0.2 + 4 \cdot 0.1 = 2$$

Priemerné umiestnenie zákazníka v rade je druhé miesto. □

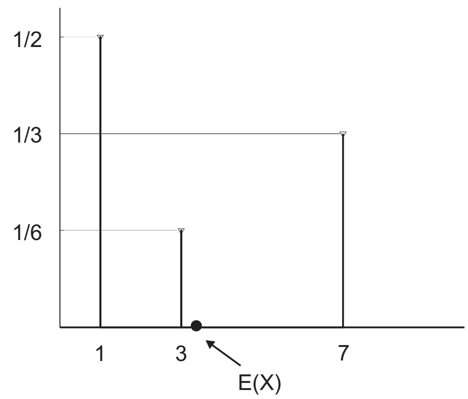
Príklad 8.4 *Náhodná premenná $\mathbb{X}$ popisuje hru, spočívajúcu v hode kockou a vyplatení príslušnej výhry. Ak na kocke padne číslo 1 alebo číslo 2, hráč dostane 20 Eur. Akú sumu má hráč zaplatiť ak padne niečo iné ako 1 a 2, aby hra bola spravodlivá?*

Riešenie: Hra bude spravodlivá, keď hráč zaplatí toľko, ako dostane. Teda keď stredná hodnota náhodnej premennej $\mathbb{X}$, ktorá popisuje hru, spočívajúcu v hode kockou a vyplatení príslušnej výhry, bude 0. Označme hľadanú sumu z.

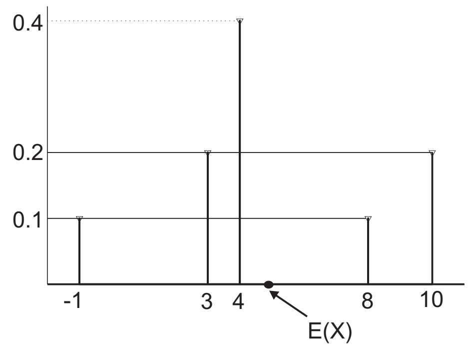
$$z = ?$$
$$E(\mathbb{X}) = z \cdot \frac{4}{6} + 20 \cdot \frac{2}{6} = 0$$
$$z = -10$$

Pri sume 10 Eur, ktoré platí hráč, keď prehrá, bude popísaná hra spravodlivá. □

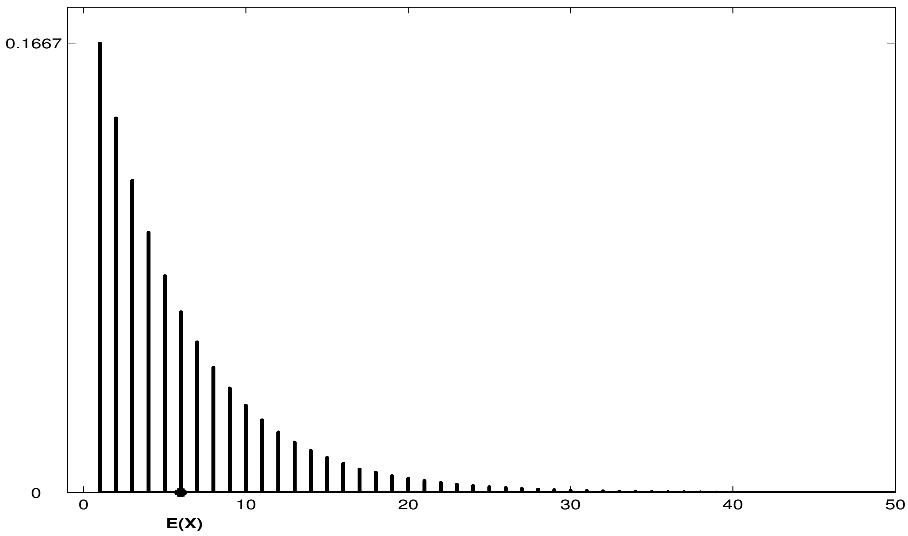
Na obrázkoch 1, 2 a 3 sú znázornené rôzne náhodné premenné so svojimi rozdeleniami pravdepodobnosti a vyznačenou strednou hodnotou. Vidíme, že stredná hodnota nie je v strede medzi hodnotami, ale ak by sme si predstavili jednotlivé pravdepodobnosti ako závažia na prevažovacej hojdačke, tak stredná hodnota je to miesto, kde treba umiestniť stred, aby bola hojdačka v rovnováhe.



Obr. 1: Rozdelenie pravdepodobnosti náhodnej premennej s hodnotami 1, 3, 7 a s pravdepodobnosťami 1/2, 1/6, 1/3. Stredná hodnota tejto náhodnej premennej je $E(\mathbb{X}) = 3.3333$.



Obr. 2: Rozdelenie pravdepodobnosti náhodnej premennej s hodnotami -1, 3, 4, 8, 10 a s pravdepodobnosťami 0.1, 0.2, 0.4, 0.1, 0.2 a jej stredná hodnota je $E(\mathbb{X}) = 6.9000$



Obr. 3: Rozdelenie pravdepodobnosti náhodnej premennej a jej stredná hodnota $E(\mathbb{X}) = 6$

Pravidlá počítania so strednou hodnotou

- Nech pre náhodnú premennú Y a nenáhodnú hodnotu $a \in R$ platí $Y = X + a$. Potom:

$$E(Y) = E(X + a) = E(X) + a$$

Dôkaz:

$$E(X + a) = E(Y) = \sum_{y_i \in Y} y_i \cdot PDF(y_i) = \sum_{x_i \in X} (x_i + a) \cdot PDF(x_i) =$$

$$= \sum_{x_i \in X} x_i \cdot PDF(x_i) + \sum_{x_i \in X} a \cdot PDF(x_i) = E(X) + a \cdot \sum_{x_i \in X} PDF(x_i) = E(X) + a \cdot 1$$

- Nech pre náhodnú premennú Y a nenáhodnú hodnotu $c \in R$ platí $Y = c \cdot X$. Potom:

$$E(Y) = c \cdot E(X)$$

- Pre náhodnú premennú $Z = X + Y$ platí

$$E(Z) = E(X) + E(Y)$$

Tento vzťah platí pre nezávislé aj závislé náhodné premenné X a Y. $\square$

- Predchádzajúce tvrdenia môžeme zovšeobecniť:

$$E(c_1 \cdot R_1 + c_2 \cdot R_2 + + \cdots + c_k \cdot R_k + a) = a_1 \cdot E(R_1) + a_2 \cdot E(R_2) + \cdots + a_k \cdot E(R_k) + a$$

- Pre dve nezávislé náhodné premenné X a Y platí

$$E(X \cdot Y) = E(X) \cdot E(Y)$$

- Nech náhodná premenná X nadobúda hodnoty $k = 0, 1, 2, 3, ...$ Potom platí

$$E(X) = \sum_{k=0}^{\infty} \Pr(X > k) = \sum_{k=0}^{\infty} TDF_X(k)$$

Dôkaz:

$$\sum_{k=0}^{\infty} TDF_X(k) = TDF_X(0) + TDF_X(1) + TDF_X(2) + TDF_X(3) + ... =$$
$$= (p_1 + p_2 + p_3 + ...) + (p_2 + p_3 + p_4 + ...) + (p_3 + p_4 + ...) + (p_4 + ...) + ... =$$
$$= p_1 + 2p_2 + 3p_3 + 4p_4 + ... = EX$$

□

Dôkaz existencie nejakého objektu pomocou strednej hodnoty

Pomocou strednej hodnoty náhodnej premennej môžeme dokázať existenciu objektov s danou vlastnosťou.

Príklad 8.5 *Nech sú všetky body kružnice zafarbené buď modrou, alebo červenou farbou. Ako sú na kružnici tieto farby rozdelené, to nevieme. Vieme iba, že* $1/5$ *bodov je červená a* $4/5$ *bodov sú modré. Dokážte, že pri ľubovoľnom rozložení modrých a červených bodov existuje rovnostranný trojuholník* ABC *vpísaný tejto kružnici, ktorého všetky tri vrcholy sú modré body.*

Riešenie: Úlohu nebudeme riešiť tak, že vykreslíme všetky možnosti zafarbenia kružnice a ku každej takejto možnosti, nájdeme trojuholník danej vlastnosti. Takéto riešenie by síce bolo korektné, ale trvalo by príliš dlho.

Namiesto toho ukážeme, že pre ľubovoľné rozloženie farieb (kde $1/5$ bodov je červená a $4/5$ bodov sú modré), musí existovať trojuholník požadovanej vlastnosti (= vpísaný, rovnostranný, s 3 modrými vrcholmi).

Uvažujme ľubovoľné rozloženie farieb a všetky možné vpísané trojuholníky ABC. Aký bude stredný počet červených vrcholov? Pravdepodobnosť, že vrchol A bude červený je $1/5$, rovnako to platí pre vrcholy B a C. Stredná hodnota náhodnej premennej $\mathbb{A}$ popisujúcej počet červených vrcholov A je $E(\mathbb{A}) = 1/5$, stredná hodnota náhodnej premennej $\mathbb{B}$ popisujúcej počet červených vrcholov B je $E(\mathbb{B}) = 1/5$ a stredná hodnota náhodnej premennej $\mathbb{C}$ popisujúcej počet červených vrcholov C je $E(\mathbb{C}) = 1/5$.

Stredná hodnota počtu červených vrcholov je $3/5$.

$$E(\mathbb{A} + \mathbb{B} + \mathbb{C}) = E(\mathbb{A}) + E(\mathbb{B}) + E(\mathbb{C}) = \frac{1}{5} + \frac{1}{5} + \frac{1}{5} = \frac{3}{5}$$

Počet červených vrcholov v jednom trojuholníku je jedno z čísel 0, 1, 2 a 3. Stredná hodnota z týchto čísel je $3/5$, teda číslo menšie ako 1. Preto spomedzi všetkých trojuholníkov musí byť niekoľko takých, ktoré majú 0 červených vrcholov. Keby to tak nebolo, priemer by bol iba z čísel 1, 2 a 3. To by muselo byť číslo väčšie alebo rovné 1. Teda existujú trojuholníky s 0 červenými vrcholmi. To znamená, že všetky ich vrcholy sú modré.

Trojuholník s požadovanou vlastnosťou sme nenašli, ale zaručili sme jeho existenciu.

□

Postup použitý v úlohe 8.5 sa nazýva **pravdepodobnostná metóda** a rozvinul ju vo svojich prácach Paul Erdős (1913–1996).