Střední hodnota náhodné veličiny

... zobecnění průměrného očekávání,

jde o vážený průměr, kde váhy odpovídají pravděpodobnostem.

- Kolik šestek průměrně padne při hodu 10 kostkami?
- Jak dlouho se obvykle čeká na tramvaj?
- Jaký je průměrný plat? I když většina zaměstnanců má podprůměrný plat, má zas nadprůměrně často narozeniny. ☺
- Jaká je průměrná škoda při havárii auta, přepočteno na všecha auta daného modelu a rok?
- Jaká je obvyklý počet uživatelů přihlášených na server?
- Kolik pevných bodů má náhodně zvolená permutace n prvků?

Průměrné očekávání závisí na:

... hodnotách, které náhodná veličina nabývá

... a jejím rozdělení pravděpodobnosti.

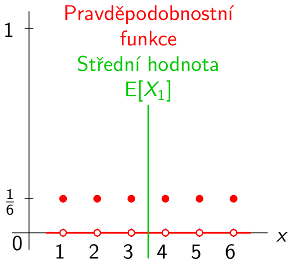
Střední hodnota náhodné veličiny

Definice: *Střední hodnota* diskrétní náhodné veličiny X na prostoru (Ω, P) je $\sum_{\omega \in \Omega} X(\omega)\mathrm{P}[\{\omega\}]$. Značí se $\mathrm{E}[X]$.

Na nekonečné Ω nemusí součet konvergovat; pak $\mathrm{E}[X]$ neexistuje.

Ukázka: X_1 – hod jednou kostkou

$\mathrm{E}[X_1] =$
$1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{7}{2}$



Sloučení jevů se stejnou hodnotou x náhodné veličiny X dává vztah $\mathrm{E}[X] = \sum_{x} x\mathrm{P}[X = x]$, přičemž stačí sčítat přes obor hodnot X.

Linearita střední hodnoty

Věta: Pro náhodné veličiny X a Y na Ω s konečnými středními hodnotami a libovolné $\alpha \in \mathbb{R}$ platí:

- $\mathrm{E}[X+Y] = \mathrm{E}[X] + \mathrm{E}[Y]$ (linearita vůči součtu) a
- $\mathrm{E}[\alpha X] = \alpha \mathrm{E}[X]$ (linearita vůči reálnému násobku).

Ukázka: X_2 — hod dvěma kostkami

X_1' — číslo na první kostce,

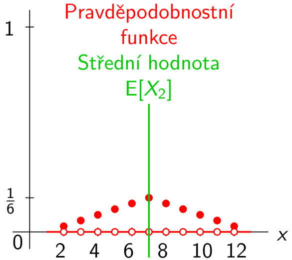
X_1'' — číslo na druhé kostce,

$X_2 = X_1' + X_1''$

X_1' a X_1'' mají stejné rozdělení jako X_1, a proto $\mathrm{E}[X_1'] = \mathrm{E}[X_1''] = \mathrm{E}[X_1]$.

$\mathrm{E}[X_2] = 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + \cdots + 12 \cdot \frac{1}{36}$

$= 7 = \frac{7}{2} + \frac{7}{2} = \mathrm{E}[X_1'] + \mathrm{E}[X_1'']$



Linearita střední hodnoty

Věta: Pro náhodné veličiny X a Y na Ω s konečnými středními hodnotami a libovolné $\alpha \in \mathbb{R}$ platí:

- $\mathsf{E}[X+Y] = \mathsf{E}[X] + \mathsf{E}[Y]$ (linearita vůči součtu) a
- $\mathsf{E}[\alpha X] = \alpha \mathsf{E}[X]$ (linearita vůči reálnému násobku).

... rozepsáním definic: $\mathsf{E}[\alpha X] = \sum_{\omega\in\Omega} (\alpha X)(\omega)\mathsf{P}[\{\omega\}]$
$= \sum_{\omega\in\Omega} \alpha X(\omega)\mathsf{P}[\{\omega\}] = \alpha \sum_{\omega\in\Omega} X(\omega)\mathsf{P}[\{\omega\}] = \alpha\mathsf{E}[X]$

$\mathsf{E}[X+Y] = \sum_{\omega\in\Omega} (X+Y)(\omega)\mathsf{P}[\{\omega\}] = \sum_{\omega\in\Omega} (X(\omega)+Y(\omega))\mathsf{P}[\{\omega\}]$
$= \sum_{\omega\in\Omega} X(\omega)\mathsf{P}[\{\omega\}] + \sum_{\omega\in\Omega} Y(\omega)\mathsf{P}[\{\omega\}] = \mathsf{E}[X] + \mathsf{E}[Y]$

Poznámka: Součet i součin lze definovat obecněji i pro náhodné veličiny na *různých* prostorech. Často při opakování pokusu, např. počet orlů při n hodech mincí je součet n veličin jednotlivých hodů. Věta o linearitě střední hodnoty platí i v tomto kontextu.

Indikátor

… pomáhá rozložit složité náhodné veličiny na jednodušší.

Definice: *Indikátor* jevu $A \subseteq \Omega$ je náhodná veličina I_A taková, že

$$I_A(\omega) = \begin{cases} 1 & \text{pro } \omega \in A, \\ 0 & \text{pro } \omega \notin A. \end{cases}$$

Pozorování: $\mathrm{E}[I_A] = \sum_{\omega \in \Omega} I_A(\omega) \mathrm{P}[\{\omega\}] = \sum_{\omega \in A} \mathrm{P}[\{\omega\}] = \mathrm{P}[A]$.

Ukázka: Kolik průměrně padne šestek při hodu n kostkami?

Označme tuto náhodnou veličinu X. Pro $i \in \{1, \ldots, n\}$ zavedeme jev A_i označující, zda na i-té kostce padla šestka.
Platí $X = I_{A_1} + \cdots + I_{A_n}$ a také $\mathrm{E}[I_{A_i}] = \mathrm{P}[A_i] = \frac{1}{6}$.

Z linearity střední hodnoty dostáváme:

$$\begin{aligned} \mathrm{E}[X] &= \mathrm{E}[I_{A_1} + \cdots + I_{A_n}] \\ &= \mathrm{E}[I_{A_1}] + \cdots + \mathrm{E}[I_{A_n}] \\ &= \mathrm{P}[A_1] + \cdots + \mathrm{P}[A_n] = \frac{n}{6}. \end{aligned}$$

Ukázky indikátorové metody

Hrajeme-li mariáš, kolik srdcových karet
lze očekávat v prvních pěti rozdaných kartách?

Označme tuto náhodnou veličinu X a pro i od 1 do 5
zavedeme jevy A_i obsahující ta rozdání, kde i-tá karta je srdcová.

Platí: $X = I_{A_1} + \cdots + I_{A_5}$, ale také $\mathrm{E}[I_{A_i}] = \mathrm{P}[A_i] = \frac{8 \cdot 31!}{32!} = \frac{1}{4}$.

Z linearity dostáváme: $\mathrm{E}[X] = \mathrm{E}[I_{A_1}] + \cdots + \mathrm{E}[I_{A_5}] = \frac{5}{4}$.

Při výpočtu střední hodnoty nevadí, že jsou jevy závislé.

Zvolíme-li náhodně substituční šifru, kolik písmen zůstane
průměrně nezměněno? Neboli, jaká je střední hodnota veličiny X,
znamenající počet pevných bodů náhodné n-prvkové permutace?

Zavedeme jevy A_i obsahující permutace, kde i je pevným bodem.
Z $X = I_{A_1} + \cdots + I_{A_n}$ a $\mathrm{P}[A_i] = \frac{(n-1)!}{n!} = \frac{1}{n}$ plyne $\mathrm{E}[X] = n \cdot \frac{1}{n} = 1$.

Ukázky indikátorové metody

Rozdělujeme-li náhodně n úloh na n serverů, kolik serverů v průměru zůstane nevyužito?

Pravděpodobnostní prostor tvoří možná přiřazení úloh na servery, čili zobrazení z n-prvkové množiny úloh do n-prvkové množiny serverů. Odtud $|\Omega| = n^n$ a pro každé $\omega \in \Omega$ platí: $\mathrm{P}[\{\omega\}] = \frac{1}{n^n}$.

Označíme $X(\omega)$ počet nevyužitých serverů v zobrazení ω. (Jde o doplněk oboru hodnot zobrazení ω v množině serverů.)

Rozložíme X na součet indikátorů jevů A_i, obsahující zobrazení, kde i-tý server je neobsazený, neboli $X = I_{A_1} + \cdots + I_{A_n}$. Každé A_i obsahuje zobrazení do $(n-1)$-prvkové množiny. Odtud $|A_i| = (n-1)^n$ a $\mathrm{P}[A_i] = \frac{(n-1)^n}{n^n} = \left(1 - \frac{1}{n}\right)^n$.

Nyní $\mathrm{E}[X] = \mathrm{E}[I_{A_1} + \cdots + I_{A_n}] = \mathrm{P}[A_1] + \cdots + \mathrm{P}[A_n] = n\left(1 - \frac{1}{n}\right)^n$.

Jinými slovy, podíl nevyužitých serverů se limitně blíží $\frac{1}{e}$.

Nezávislé náhodné veličiny

... zjišťování jedné veličiny neovlivní hodnotu jiné

Definice: Náhodné veličiny X a Y na diskrétním pravděpodobnostním prostoru Ω jsou *nezávislé*, pokud pro všechna reálná x a y jsou jevy $\{\omega \in \Omega\colon X(\omega) = x\}$ a $\{\omega \in \Omega\colon Y(\omega) = y\}$ nezávislé.

Formálně: $\forall x, y \in \mathbb{R}\colon \mathrm{P}[X = x]\mathrm{P}[Y = y] = \mathrm{P}[X = x \wedge Y = y]$.

Ukázky: Veličiny X_1' a X_1'' z ukázky o hodu dvěma kostkami (počet ok na první/druhé) jsou nezávislé. Libovolné $x, y \in \{1, \dots, 6\}$ totiž splňují: $\mathrm{P}[X_1' = x]\mathrm{P}[X_1'' = y] = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36} = \mathrm{P}[X_1' = x \wedge X_1'' = y]$.

Ze skupiny 2 děvčat a 1 chlapce vybereme dvojici. Veličina X s alternativním rozdělením udává, zdali byl chlapec vybrán v prvním kroku, a Y v druhém.

Image by Freepik.

Tyto veličiny jsou závislé, neboť pro $x = 1$ a $y = 1$ máme: $\mathrm{P}[X = 1]\mathrm{P}[Y = 1] = \frac{1}{3} \cdot \frac{1}{3} \neq 0 = \mathrm{P}[X = 1 \wedge Y = 1]$.

Nezávislé náhodné veličiny

... zjišťování jedné veličiny neovlivní hodnotu jiné

Definice: Náhodné veličiny X a Y na diskrétním pravděpodobnostním prostoru Ω jsou *nezávislé*, pokud pro všechna reálná x a y jsou jevy $\{\omega \in \Omega\colon X(\omega) = x\}$ a $\{\omega \in \Omega\colon Y(\omega) = y\}$ nezávislé.

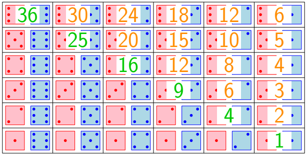
Formálně: $\forall x, y \in \mathbb{R}\colon \mathrm{P}[X = x]\mathrm{P}[Y = y] = \mathrm{P}[X = x \wedge Y = y]$.

Věta:

Pro *nezávislé* náhodné veličiny X a Y platí: $\mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$.

Ukázka: Střední hodnota *součinu* počtů ok padlých na dvou kostkách je podle definice:

$$\frac{30+24+18+12+6+20+15+10+5+12+8+4+6+3+2}{18} + \frac{36+25+16+9+4+1}{36} = \frac{91}{36} + \frac{175}{18} = \frac{441}{36} = \frac{49}{4}$$



Protože jde o součin dvou nezávislých veličin X_1' a X_1'',

platí podle věty: $\mathrm{E}[X_1' X_1''] = \mathrm{E}[X_1']\mathrm{E}[X_1''] = \frac{7}{2} \cdot \frac{7}{2} = \frac{49}{4}$

Nezávislé náhodné veličiny

... zjišťování jedné veličiny neovlivní hodnotu jiné

Definice: Náhodné veličiny X a Y na diskrétním pravděpodobnostním prostoru Ω jsou *nezávislé*, pokud pro všechna reálná x a y jsou jevy $\{\omega \in \Omega \colon X(\omega) = x\}$ a $\{\omega \in \Omega \colon Y(\omega) = y\}$ nezávislé.

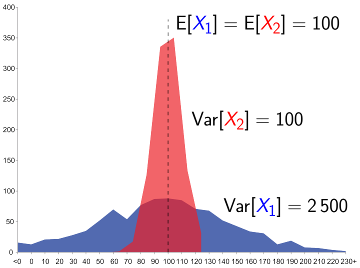
Formálně: $\forall x, y \in \mathbb{R} \colon \mathrm{P}[X = x]\mathrm{P}[Y = y] = \mathrm{P}[X = x \wedge Y = y]$.

Věta:
Pro *nezávislé* náhodné veličiny X a Y platí: $\mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$.

... z definic plyne: $\mathrm{E}[X]\mathrm{E}[Y] = \left(\sum_x x\mathrm{P}[X = x]\right)\left(\sum_y y\mathrm{P}[Y = y]\right) = \sum_x \sum_y xy\mathrm{P}[X = x]\mathrm{P}[Y = y] = \sum_x \sum_y xy\mathrm{P}[X = x \wedge Y = y]$

Seskupením x a y splňujících $xy = z$ dostáváme:
$\sum_x \sum_y xy\mathrm{P}[X = x \wedge Y = y] = \sum_z z\mathrm{P}[XY = z] = \mathrm{E}[XY]$.

Rozptyl náhodné veličiny



. . . popisuje, jak se hodnoty náhodné veličiny liší od průměrného očekávání.

Definice: *Rozptyl* náhodné veličiny X je $\mathrm{E}\left[(X - \mathrm{E}[X])^2\right]$. Značí se $\mathrm{Var}[X]$.

Pozorování: Pro rozptyl platí: $\mathrm{Var}[X] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2$.

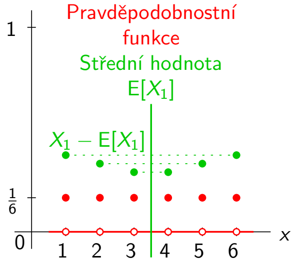
Ukázka: X_1 – hod jednou kostkou

$\mathrm{Var}[X_1] =$

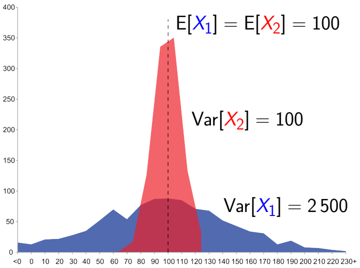
$$\frac{1}{6}\left[\frac{25}{4} + \frac{9}{4} + \frac{1}{4} + \frac{1}{4} + \frac{9}{4} + \frac{25}{4}\right] = \frac{35}{12}$$

$\mathrm{Var}[X_1] = \mathrm{E}[X_1^2] - (\mathrm{E}[X_1])^2 =$

$$\frac{1+4+9+16+25+36}{6} - \frac{49}{4} = \frac{182-147}{12} = \frac{35}{12}$$



Rozptyl náhodné veličiny



... popisuje, jak se hodnoty náhodné veličiny liší od průměrného očekávání.

Definice: *Rozptyl* náhodné veličiny X je $E\left[(X - E[X])^2\right]$. Značí se $Var[X]$.

Pozorování: Pro rozptyl platí: $Var[X] = E[X^2] - (E[X])^2$.

... $Var[X] = E\left[(X - E[X])^2\right] = E\left[X^2 - 2XE[X] + (E[X])^2\right] =$
$E[X^2] - 2E[X]E[X] + E\left[(E[X])^2\right] = E[X^2] - (E[X])^2$

Odhady pravděpodobnosti výjimečných jevů

Markovova nerovnost: Pro *nezápornou* náhodnou veličinu X a všechna $a \in \mathbb{R}$ platí: $\mathrm{P}[X \geq a] \leq \frac{\mathrm{E}[X]}{a}$.

Důsledek: $\mathrm{P}[X \geq t\mathrm{E}[X]] \leq \frac{1}{t}$ ... dosazením $a = t\mathrm{E}[X]$.

Ukázky: Pravděpodobnost, že při hodu 10 kostkami padnou alespoň 3 šestky je nejvýše $\frac{5}{9}$: $\mathrm{P}[X \geq 3] \leq \frac{10/6}{3}$.

Server zpracuje v průměru 1 000 dotazů za jednotku času.

Jak lze omezit pravděpodobnost, že bude překročen limit 1 500 dotazů?

$$\mathrm{P}[X \geq 1\,500] \leq \frac{1\,000}{1\,500} = \frac{2}{3}$$

Čebyševova nerovnost: Pro *každou* náhodnou veličinu X a všechna $a \geq 0$ platí: $\mathrm{P}[|X - \mathrm{E}[X]| \geq a] \leq \frac{\mathrm{Var}[X]}{a^2}$.

Pokračování ukázky: Jak se změní odhad pravděpodobnosti zahlcení serveru, pokud *navíc víme*, že X má rozptyl 200?

$$\begin{aligned} \mathrm{P}[X \geq 1\,500] &= \mathrm{P}[X - 1\,000 \geq 500] \\ &\leq \mathrm{P}[|X - 1\,000| \geq 500] \leq \frac{200}{500^2} = 0{,}08\% \end{aligned}$$

Odhady pravděpodobnosti výjimečných jevů

Markovova nerovnost: Pro *nezápornou* náhodnou veličinu X a všechna $a \in \mathbb{R}$ platí: $\mathrm{P}[X \geq a] \leq \frac{\mathrm{E}[X]}{a}$.

Důsledek: $\mathrm{P}[X \geq t\mathrm{E}[X]] \leq \frac{1}{t}$ ... dosazením $a = t\mathrm{E}[X]$.

... rozepsáním a odhady: $\mathrm{E}[X] = \sum_x x\mathrm{P}[X = x] \geq$
$\sum_{x \geq a} x\mathrm{P}[X = x] \geq \sum_{x \geq a} a\mathrm{P}[X = x] = a \sum_{x \geq a} \mathrm{P}[X = x] = a\mathrm{P}[X \geq a]$

Čebyševova nerovnost: Pro *každou* náhodnou veličinu X a všechna $a \geq 0$ platí: $\mathrm{P}[|X - \mathrm{E}[X]| \geq a] \leq \frac{\mathrm{Var}[X]}{a^2}$.

... do Markovovy nerovnosti ve tvaru $\mathrm{P}[Y \geq b] \leq \frac{\mathrm{E}[Y]}{b}$
dosadíme $Y = (X - \mathrm{E}[X])^2$ a $b = a^2$, čímž dostaneme:
$\mathrm{P}[(X - \mathrm{E}[X])^2 \geq a^2] \leq \frac{\mathrm{E}[(X - \mathrm{E}[X])^2]}{a^2} \Leftrightarrow \mathrm{P}[|X - \mathrm{E}[X]| \geq a] \leq \frac{\mathrm{Var}[X]}{a^2}$

Kvíz — řešení

Je-li u některých otázek více možností správných, vyberte všechny.

1. Pravda nebo lež? Rozptyl veličin X a $X + 5$ je stejný.
2. Střední hodnota binomického rozdělení s parametry n a p
   a) np, b) p^n, c) $p^n(1-p)^n$, d) $np(1-p)$, e) $p^n + (1-p)^n$.
3. Jakou hodnotu bychom dostali, pokud bychom odchylku náhodné veličiny X od průměru měřili výrazem $\mathrm{E}[X - \mathrm{E}[X]]$?
   a) 0, b) $\frac{1}{2}\mathrm{E}[X]$, c) $\mathrm{E}[X]$, d) $\sqrt{\mathrm{Var}[X]}$.
4. Má-li náhodné veličina X rozptyl a, potom $2X$ má rozptyl
   a) $\sqrt{a}$, b) $\sqrt{2}a$, c) a, d) $2a$, e) a^2, f) $4a$, g) $4a^2$.
5. Pravda nebo lež?
   Jsou-li náhodné veličiny závislé, potom $\mathrm{E}[XY] \neq \mathrm{E}[X]\mathrm{E}[Y]$.

Otázky k porozumění tématu přednášky

- Jak lze spočítat *přesnou* pravděpodobnost, že při hodu 10 kostkami padnou alespoň 3 šestky? Vyjde $\frac{13\,671\,875}{30\,233\,088} \doteq 45\%$.
- Lze k Markovově a Čebyševově nerovnosti nalézt i nějaký netriviální *dolní* odhad na pravděpodobnost?

Poznámky k pojmosloví a značení

Příbuzný pojem ke střední hodnotě je *medián*.
Zjednodušeně řečeno jde o m takové, že $\mathrm{P}[X \leq m] = \mathrm{P}[X \geq m]$,
čili m je „uprostřed" hodnot nabývaných X.

Směrodatná odchylka je odmocnina z rozptylu.

Střední hodnota se někdy značí μ (angl. *mean*), směrodatná odchylka σ (*standard deviation*) a rozptyl pak σ^2 (*variance*).

Nezávislost náhodných veličin se často definuje pomocí nerovností:
$\forall x, y \in \mathbb{R}$: $\mathrm{P}[X \leq x]\mathrm{P}[Y \leq y] = \mathrm{P}[X \leq x \wedge Y \leq y]$,
protože tento vztah má smysl i pro spojité náhodné veličiny.
Pro diskrétní náhodné veličiny je ekvivalentní vztahu s rovnostmi.