# 4. cvičení

Pravděpodobnost a statistika 1, 12. 3. 2026 https://iuuk.mff.cuni.cz/~chmel/2526/past/

## Ještě troška náhodných veličin

| Název | Značení | Pravděpodobnostní funkce | Rozsah (Im($X$)) | Střední hodnota |
|---|---|---|---|---|
| Bernoulliho | $X \sim \mathrm{Ber}(p)$ | $p_X(1) = p$, $p_X(0) = 1 - p$ | $\{0, 1\}$ | $p$ |
| Binomické | $X \sim \mathrm{Bin}(n, p)$ | $p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$ | $\{0, 1, \ldots, n\}$ | $np$ |
| Geometrické | $X \sim \mathrm{Geo}(p)$ | $p_X(k) = (1-p)^{k-1} p$ | $\{1, 2, \ldots\}$ | $1/p$ |
| Poissonovo | $X \sim \mathrm{Poi}(\lambda)$ | $p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}$ | $\{0, 1, \ldots\}$ | $\lambda$ |
| Uniformní | $X \sim \mathrm{Uni}(a, b)$ | $p_X(k) = \frac{1}{b-a+1}$ | $\{a, a+1, \ldots, b\}$ | $\frac{a+b}{2}$ |
| Hypergeometrické | $X \sim \mathrm{Hyp}(N, K, n)$ | $p_X(k) =$ | $\{0, 1, 2, \ldots, \min(n, K)\}$ | $n\frac{K}{N}$ |

**Úloha 1** (Která je tohle distribuce?)
Uvažme $m + n$ hodů spravedlivou šestistěnnou kostkou. Označme $X$ počet šestek z prvních $m$ hodů, $Y$ počet šestek z posledních $n$ hodů. Jaká je distribuce $X$, $Y$ a $X + Y$?

**Řešení**
Po řadě $\mathrm{Bin}(m, 1/6), \mathrm{Bin}(n, 1/6), \mathrm{Bin}(m + n, 1/6)$.

**Úloha 2** (Odpověď zní: Jigglypuff, pohled shora)
V pytlíku je $N$ bonbónů, z nichž $K$ je dobrých. Náhodně vytáhneme $n$ z nich, označíme $X$ počet dobrých vytažených bonbónů.

a) Jak se jmenuje rozdělení n.v. $X$?

b) Jaká je $P(X = k)$?

c) S pomocí tabulky určete $\mathbb{E}[X]$, pro $n = 1$ si rozmyslete, že to je jasné.

**Řešení** a) Hypergeometrické, mělo zaznít na přednášce

b) $P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$

c) Dle tabulky $n\frac{K}{N}$.

## Střední hodnota

Pro diskrétní náhodnou veličinu $X$ definujeme její střední hodnotu $\mathbb{E}[X] = \sum_{x \in \mathrm{Im}(X)} x \cdot P[X = x]$.

**Úloha 3** (Kostky jsou vrženy)
Po hodu kostkou dostaneme za šestku 10 korun, za pětku 7 korun, za ostatní zaplatíme 5 korun. Jaká je střední hodnota výhry?

**Řešení**
Podle definice $\mathbb{E}(X) = 10 \cdot \frac{1}{6} + 7 \cdot \frac{1}{6} + (-5) \cdot \frac{4}{6} = \frac{10+7-20}{6} = -\frac{1}{2}$.
Střední hodnota výhry je tedy $-1/2$ koruny.

**Úloha 4** (Lovci ponožek)
Máme neomezený počet černých a červených ponožek. Ponožky vytahujeme poslepu, obě barvy jsou stejně pravděpodobné.

a) Kolik ponožek ve střední hodnotě vytáhneme, než budeme mít dvě stejné barvy?

b) Řešte totéž pro tři různé barvy.

**Řešení** a) S pravděpodobností 1/2 nám stačí 2 tahy, a za 3 tahy máme dvě barvy určitě (tj. pravděpodobnost je zbývající polovina), tedy $\mathbb{E}[X] = 2.5$

b) $2 \cdot 3 \cdot (1/3)^2 + 3 \cdot 6 \cdot (1/3)^2 \cdot (2/3) + 4 \cdot (1 - 1/3 - 4/9) = 2 \cdot 1/3 + 3 \cdot 4/9 + 4 \cdot 2/9 = 26/9$

**Úloha 5** (Know when to walk away)
Nové kasino nabízí následující hru: vsadíme $x$ korun, s pravděpodobností 1/2 o ně přijdeme, ale s pravděpodobností 1/2 vyhrajeme $2x$ (navíc k našim $x$ korunám).

a) Jaká je střední hodnota výhry?

b) Začínáme s $k$ korunami. Pokud chceme maximalizovat střední hodnotu našich zisků po $n$ kolech, jak to udělat (a kolik ta střední hodnota je)?

c) Jaká je pravděpodobnost, že s takovou strategií přijdeme o všechny peníze?

d) Jakou strategii byste zvolili?

**Řešení** a) Čistá výhra má střední hodnotu $\mathbb{E}[X] = (-x) \cdot 1/2 + 2x \cdot 1/2 = x/2$, kde $x$ je sázka. Tedy ve střední hodnotě po jednom kole máme $3x/2$ mincí.

b) Indukcí z předchozího pak máme, že po $\ell$ kolech bychom měli $(\frac{3}{2})^\ell x$, a tedy je nejlepší vždycky sázet celou částku pro $(\frac{3}{2})^n k$

c) Musíme aspoň jednou prohrát, což se stane s pravděpodobností $1 - 2^{-n}$.

d) Pokud maximalizujeme pouze střední hodnotu konečného bohatství, vychází „vsaď všechno“. To je ale extrémně riskantní.

Rozumnější strategie závisí na cíli:

- chceme-li maximalizovat střední hodnotu peněz, pak *vždy všechno*;
- chceme-li výrazně snížit riziko bankrotu, pak sázet jen část kapitálu;
- při maximalizaci očekávaného logaritmu bohatství (Kellyho kritérium) bychom volili takový podíl $f$, který maximalizuje
$$\frac{1}{2}\log(1-f) + \frac{1}{2}\log(1+2f).$$

Derivací vyjde optimum $\boxed{f = \frac{1}{4}}$, tedy sázet čtvrtinu aktuálního kapitálu. Zdůvodnění, proč to dává smysl, půjde až po zákonech velkých čísel. Ale intuitivně: tvářím se, že přesně v polovině případů vyhraju, v polovině prohraju, na pořadí nezáleží, maximalizuju výhru pro takovéhle pořadí.

## Linearita střední hodnoty

Máme-li náhodné veličiny $X_1, \ldots, X_n$ a $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$, pak $\mathbb{E}[\alpha_1 X_1 + \cdots + \alpha_n X_n] = \alpha_1 \mathbb{E}[X_1] + \cdots + \alpha_n \mathbb{E}[X_n]$.

**Úloha 6** (Hody mincí)
Hodíme $n$-krát korunou, která má pravděpodobnost, že padne Panna, rovnou $p$.

a) Označme $X$ počet po sobě jdoucích hodů PO. (Např. pokud $n = 6$ a padlo postupně POOPOO, tak $X = 2$.) Určete $\mathbb{E}(X)$.

b) Rozmyslete si, proč se nejedná o binomické rozdělení.

c) Označme teď $Y$ počet opakování hodů POP, jaká je $\mathbb{E}(Y)$?

**Řešení** a) $\mathbb{E}(X) = (n-1) \cdot p \cdot (1-p)$

b) Indikátory nejsou navzájem nezávislé, $X_i$ a $X_{i+1}$ nemohou být zároveň 1.

c) $\mathbb{E}(Y) = (n-2) \cdot p^2 \cdot (1-p)$

**Úloha 7** ($G(n,p)$)
Hodíme $\binom{n}{2}$-krát korunou, na které padne Panna s pravděpodobností $p$. Přitom tvoříme graf s vrcholy $V = \{1, 2, \ldots, n\}$. Postupně pro všechny dvojice $\{i, j\} \in \binom{V}{2}$ určíme, jestli jsou spojené hranou – to bude tehdy, když příslušným hodem padla Panna. Vzniklému grafu se říká (Erdősův-Rényiho) náhodný graf $G(n,p)$.
Ukažte, že střední hodnota počtu hran v grafu je $p\binom{n}{2}$ a střední hodnota počtu trojúhelníků v grafu je $p^3\binom{n}{3}$.

**Řešení**
Standardní indikátory: máme $\binom{n}{2}$ hran, každá s pravděpodobností $p$. Máme $\binom{n}{3}$ trojúhelníků, každý se objeví s pravděpodobností $p^3$ (všechny hrany musí být vybrané).

**Úloha 8** (Za dva dny je PI(E) day!)
Připomeňte si definici indikátorové náhodné veličiny $I_A$.

a) Jaká je $\mathbb{E}(I_A)$?

b) Nechť $A = A_1 \cup A_2 \cup \cdots \cup A_n$. Ověřte rovnost $1 - I_A = \prod_{i=1}^{n}(1 - I_{A_i})$.

c) Roznásobte a použijte větu o linearitě střední hodnoty. Dostanete princip inkluze a exkluze, který znáte z diskrétní matematiky.

**Řešení**
1. $P(I_A)$
2. Pokud $I_A = 1$, pak nějaké $I_{A_i} = 1$ a naopak.
3. TODO

---

**Bonusové úlohy**

**Úloha 9** (Sběratel Hot Wheels)
Ke každému nákupu dostaneme jako dárek autíčko – náhodně vybrané z $n$ typů. Kolik průměrně musíme udělat nákupů, než dostaneme všechny typy autíček?

**Řešení**
Nemůžeme udělat náhodné veličiny typu "počet jiných autíček do prvního správného autíčka", takže uděláme následující: celkový počet nákupů budiž $T = T_1 + \ldots + T_n$, kde $T_i$ je počet nákupů mezi získáním $(i-1)$-ního a $i$-tého nového autíčka. Jistě vidíme, že $T_i \sim \text{Geo}(\frac{n-(i-1)}{n})$, což má střední hodnotu $\mathbb{E}[T_i] = \frac{n}{n-(i-1)}$. Pak $\mathbb{E}[T] = \sum \mathbb{E}[T_i] = \sum_{i=1}^{n} \frac{n}{n-(i-1)} = n \cdot H_n = n \log n + \gamma n + \frac{1}{2} + \mathcal{O}(\frac{1}{n})$, kde $\gamma$ je Eulerova-Mascheroniho konstanta, nebo obecně slaběji $\mathcal{O}(n \log n)$.