5. cvičení

Pravděpodobnost a statistika 1, 15. 3. 2023 `https://iuuk.mff.cuni.cz/~chmel/2223/past/`

Na zahřátí

Úloha 1 (Náhodný řez)
Mějme neorientovaný graf $G = (V, E)$. Vybereme náhodně podmnožinu vrcholů $V' \subseteq V$ tak, že pro každý vrchol si nezávisle hodíme (spravedlivou) mincí, zda ho dáme do V'. Jaká je střední hodnota počtu hran, které vedou mezi V' a $V \setminus V'$?

Řešení
$\frac{|E|}{2}$, protože každá hrana tam padne s pravděpodobností 1/2 – střední hodnota pak půjde přes indikátory

Úloha 2 (Tři kostky)
Hodíme třemi šestistěnnými kostkami a výsledek přečteme jako trojciferné číslo (kostky jsou odlišitelné, a předem víme, že např. žlutá kostka bude na pozici stovek, červená na pozici desítek a modrá na pozici jednotek). Jaká je jeho střední hodnota?

Řešení
$350 + 35 + 3.5 = 388.5$

Úloha 3 (Dvě kostky)
Hodíme dvěma kostkami a výsledek přečteme jako dvojciferné číslo, přičemž kostku s větší hodnotou použijeme pro řád desítek. Jaká je teď střední hodnota?

Řešení
$\frac{2\cdot21+2\cdot31+2\cdot41+2\cdot51+2\cdot61+2\cdot32+2\cdot42+2\cdot52+2\cdot62+2\cdot43+2\cdot53+2\cdot63+2\cdot54+2\cdot64+2\cdot65+11+22+33+44+55+66}{36} = 189/4 = 47.25.$

Zacházení se střední hodnotou a rozptylem

Úloha 4 (Rozptyl a jeho (ne)linearita)
Nechť X, Y jsou diskrétní náhodné veličiny a $\alpha \in \mathbb{R}$.

a) Ukažte, že $\mathrm{var}(X + \alpha) = \mathrm{var}(X)$.

b) Vyjádřete $\mathrm{var}(\alpha X)$ pomocí $\mathrm{var}(X)$.

c) Ukažte, že $\mathrm{var}(X + Y) = \mathrm{var}(X) + \mathrm{var}(Y)$, pokud X, Y jsou nezávislé.

Řešení a) Evidentně $\mathbb{E}(((X + \alpha) - \mathbb{E}(X + \alpha))^2) = \sum_t ((X + \alpha) - \mathbb{E}(X + \alpha))^2 \cdot P[X + \alpha = t] = \sum_t (X + \alpha - \mathbb{E}(X) - \alpha)^2 \cdot P[X + \alpha = t] = \sum_t (X - \mathbb{E}(X))^2 \cdot P[X = t - \alpha] = \sum_t (X - \mathbb{E}(X))^2 \cdot P[X = t]$

b) $\mathbb{E}(((\alpha X) - \mathbb{E}(\alpha X))^2) = \mathbb{E}((\alpha(X - \mathbb{E}(X)))^2) = \mathbb{E}(\alpha^2((X - \mathbb{E}(X)))^2) = \alpha^2 \mathbb{E}((X - \mathbb{E}(X))^2)$.

c) $\mathbb{E}[((X+Y)-\mathbb{E}(X+Y))^2] = \mathbb{E}[((X-\mathbb{E}(X))+(Y-\mathbb{E}(Y)))^2] = \mathbb{E}[(X-\mathbb{E}(X))^2+2(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))+(Y-\mathbb{E}(Y))^2] = \mathbb{E}[(X-\mathbb{E}(X))^2]+\mathbb{E}[(Y-\mathbb{E}(Y))^2]+\mathbb{E}[2(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))]$, dále $\mathbb{E}[2(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))] = 2\mathbb{E}[XY - X\mathbb{E}(Y) - Y\mathbb{E}(X) + \mathbb{E}(X)\mathbb{E}(Y)] = \mathbb{E}[XY] - \mathbb{E}[X\mathbb{E}(Y)] - \mathbb{E}(Y\mathbb{E}(X)) + \mathbb{E}(X)\mathbb{E}(Y) = 0$ díky nezávislosti

Podmíněná střední hodnota

Úloha 5 (Testujeme)
V testu je 20 otázek s volbami $\{a, b, c, d\}$, vždy je správná právě jedna odpověď. Za správnou odpověď dostanete 1 bod, za špatnou $-1/4$ bodu, za nevyplněnou otázku nula. Každá otázka je s pravděpodobností p jednou z těch, co se Kvído naučil, a tedy zná správnou odpověď. Pokud správnou odpověď nezná, ví o tom a může se rozhodnout, zda tipovat.

a) Jaká je střední hodnota počtu bodů, které Kvído získá, pokud bude odpovídat jenom na otázky, u kterých zná odpověď?

b) A co když bude tipovat, když nezná správnou odpověď?

c) Jak by se musela změnit penalizace za chybnou odpověď, aby byly odpovědi v částech a, b stejné?

Řešení a) $\mathbb{E}[B_i|V_i] = 1 - \mathbb{E}[B_i] = 1 \cdot p + 0 \cdot (1-p) \rightsquigarrow \mathbb{E}[\sum_{i=1}^{20} B_i] = 20p$

b) $\mathbb{E}[B_i|V_i] = 1, \mathbb{E}[B_i|V_i^c] = 1/4 - 3/16 = 1/16 - \mathbb{E}[B_i] = 1 \cdot p + 1/16 \cdot (1-p) = \frac{15p+1}{16} \rightsquigarrow \mathbb{E}[\sum_{i=1}^{20} B_i] = \frac{5}{4} \cdot (15p+1)$

c) $\mathbb{E}[B_i|V_i] = 1, \mathbb{E}[B_i|V_i^c] = 1/4 - 3d/4 = 0 \rightsquigarrow d = 1/3$.

Úloha 6 (Člověče, nezlob se)
Házíme běžnou kostkou, při šestce házíme znovu, a to i opakovaně. Spočítejte střední hodnotu součtu všech hozených čísel (nemusíte ručně vyčíslovat).

Řešení
Jevy H_i - přesně i-krát jsem hodil šestku $\rightsquigarrow \mathbb{E}[X|H_i] = 6i + 3$. $\mathbb{E}[X] = \sum_{i=0}^{\infty} \mathbb{E}[X|H_i] \cdot \frac{5}{6^{i+1}} = \frac{(6i+3)\cdot 5}{6^{i+1}} = 4.2$

Úloha 7 (Nezávislost jevů a indikátorů)
Ukažte, že jevy A, B jsou nezávislé, právě když jsou nezávislé jejich indikátorové veličiny.

Řešení
Stačí rozepsat.

Bonusové úlohy

Úloha 8 (Velký lov)
Velký hon: n myslivců se snaží ulovit n lišek. Míří náhodně: každý myslivec strefí rovnoměrně náhodně vybranou lišku. Spočítejte střední počet lišek, které hon přežijí (nikdo je netrefil).

Řešení
P[jedna konkrétní liška přežije] $= (\frac{n-1}{n})^n$. $\mathbb{E}$[počet lišek] $= n \cdot (\frac{n-1}{n})^n$.

Úloha 9 (Počty levých maxim)
Mějme náhodnou permutaci π na množině $\{1, \ldots, n\}$. Prvek i je *levým maximem*, pokud $\pi(j) < \pi(i)$ pro všechna $j < i$ (tj. všechny hodnoty nalevo jsou menší). Jaká je střední hodnota počtu levých maxim? Analogie: různě vysocí lidé stojí ve frontě, kolik z nich vidí na začátek fronty?

Řešení
TODO

Úloha 10 (Indikátory a PIE)
Připomeňte si definici indikátorové náhodné veličiny I_A.

a) Nechť $A = A_1 \cup A_2 \cup \cdots \cup A_n$. Ověřte rovnost

$$1 - I_A = \prod_{i=1}^{n} (1 - I_{A_i}).$$

b) Roznásobte a použijte větu o linearitě střední hodnoty, abyste získali princip inkluze a exkluze.

Řešení
TODO

Úloha 11 (Střední hodnota a rozptyl podruhé)
Dokažte:

a) Pokud $\mathbb{E}(X^2) = 0$, tak $P(X = 0) = 1$.

b) Předpokládejme, že $\mathrm{var}(X) = 0$, dále že $\mathbb{E}(X)$ existuje a je konečná.
Pak $X = \mathbb{E}(X)$ skoro jistě, neboli $P[X = \mathbb{E}(X)] = 1$.

Řešení
TODO

Úloha 12 (Nezávislost pro jevy $\leq k$)
Ukažte, že pro diskrétní nezávislé n.v. X, Y platí

$$P[X \leq x \wedge Y \leq y] = P[X \leq x] \cdot P[Y \leq y].$$

Pro jednoduchost můžete předpokládat, že $\mathrm{Supp}(X) = \mathrm{Supp}(Y) = \{1, 2, \ldots, n\}$ pro nějaké n.

Řešení
TODO

Tahák

- *Střední hodnota* veličiny X: $\mathbb{E}(X) = \sum_t t \cdot P(X = t)$.
- *Linearita střední hodnoty:* $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$, $\mathbb{E}(\alpha X) = \alpha\mathbb{E}(X)$.
- *Podmíněná střední hodnota* veličiny X za podmínky jevu B: $\mathbb{E}(X|B) = \sum_t t \cdot P(X = t|B)$.
- *Rozbor případů:* $\mathbb{E}(X) = \mathbb{E}(X|A) \cdot P(A) + \mathbb{E}(X|A^c) \cdot P(A^c)$.
- *LOTUS (pravidlo naivního statistika):* $\mathbb{E}(g(X)) = \sum_a g(a) \cdot P(X = a)$.
- *Indikátorová náhodná veličina* I_A jevu A nabývá hodnot 0 a 1, přičemž $I_A = 1 \Leftrightarrow A$ nastane. Platí, že $\mathbb{E}(I_A) = P(A)$.
- *Rozptyl* $\mathrm{var}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$.
- Veličiny X a Y jsou *nezávislé* $\Leftrightarrow P(X = a \wedge Y = b) = P(X = a) \cdot P(Y = b)$.
- Pro X, Y nezávislé platí $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

Domácí úkol 5

`https://iuuk.mff.cuni.cz/~chmel/2223/past/` Zadáno 15. 3. 2023
Pravděpodobnost a statistika 1 Odevzdejte do 22. 3. 2023 9:00 přes Poštovní sovu

---

Úloha 1 (Výzkum)
Na přednášce „Jak vystudovat matfyz“ se přednášející poněkud oddálil od tématu, protože začal mluvit o počtu gumovacích per, které průměrný student během studia potřebuje. Zároveň prohlásil, že empiricky se zdá, že tento počet lze modelovat náhodnou veličinou X, která mimo jiné splňuje $\mathbb{E}[X] = 5, \mathbb{E}[X^2] = 20$. V ten okamžik jste jako studenti obeznámení s pravděpodobností a statistikou zbystřili – co je na jeho tvrzení podezřelé?