2. cvičení z PST

10. října 2018

2.1 (srovnání výběru s opakováním a bez opakování)

V loterii je $k = 500$ výher a $n = 10^7$ účastníků. Jaká je pravděpodobnost, že Alice získá nějakou výhru, pokud

(a) každý může vyhrát nejvýše jednou,

(b) každý může vyhrát opakovaně.

Řešení:

Z hlediska výsledku je jedno budeme-li uvažovat uspořádaný nebo neuspořádaný výběr (protože i při neuspořádaném výběru musíme stejně vzít v úvahu všechny uspořádané možnosti, ze kterých daná neuspořádaná možnost vznikne - např. pro $k = 3$ odpovídají usp. trojice (A, A, B), (A, B, A) a (B, A, A) jedné neuspořádané možnosti, že A bylo vybráno $2\times$ a B bylo vybráno $1\times$).

Z hlediska výpočtu to ale není pokaždé stejné a bývá snadnější používat uspořádaný výběr.

(a) Uvažujme neuspořádaný výběr bez opakování. Pravděpodobnost spočítáme přes doplňkový jev. Počet všech možností je $\binom{n}{k}$. Počet nepříznivých možností je $\binom{n-1}{k}$ (vynecháme Alici). Pravděpodobnost p_1 je tedy

$$p_1 = 1 - \frac{\binom{n-1}{k}}{\binom{n}{k}} = 1 - \frac{n-k}{n} = \frac{k}{n} = \frac{500}{10^7} = 5 \cdot 10^{-5} \; .$$

(b) Uvažujme uspořádaný výběr s opakováním. Pravděpodobnost opět spočítáme přes doplňkový jev. Počet všech možností je n^k. Počet nepříznivých možností je $(n-1)^k$. Pravděpodobnost p_2 je tedy

$$p_2 = 1 - \frac{(n-1)^k}{n^k} = 1 - \left(1 - \frac{1}{n}\right)^k = 1 - (0.9999999)^{500} \doteq 1 - 0.999950001 = 4.9999 \cdot 10^{-5} \; .$$

Jak je vidět, pravděpodobnosti jsou téměř stejné. To můžeme vysvětlit buď pomocí binomického rozvoje

$$p_2 = 1 - \left(1 - \frac{1}{n}\right)^k = 1 - \left(1 - \frac{k}{n} + \binom{k}{2}\frac{1}{n^2} - \cdots\right) = \underbrace{\frac{k}{n}}_{p_1} - \binom{k}{2}\frac{1}{n^2} + \cdots$$

nebo pomocí limity (pro k pevné a n rostoucí nade všechny meze)

$$\lim_{n\to\infty} \frac{p_2}{p_1} = \lim_{n\to\infty} \frac{1 - \left(1 - \frac{1}{n}\right)^k}{\frac{k}{n}} = \left[x = -\frac{1}{n}\right] = \frac{1}{k} \cdot \lim_{x\to 0} \frac{(1+x)^k - 1}{x} = \frac{1}{k} \cdot \frac{\mathrm{d}}{\mathrm{d}x}(1+x)^k_{|x=0} = \frac{1}{k} \cdot k = 1 \; .$$

Současně tento výsledek odpovídá intuitivní představě, že v obrovském množství účastníků n se už v limitě (tj. pro $k << n$) ztratí to, jestli je losujeme opakovaně (tj. vlastně neměníme podmínky) nebo ne (tj. daného výherce vždy vyřadíme).

2.2 (podmíněná pravděpodobnost)

Pravděpodobnost toto, že napětí v elektrické síti překročí standardní hodnotu, je rovna $p_1 > 0$. Při přepětí je pravděpodobnost poruchy elektrického spotřebiče rovná p_2. K poruše přístroje může dojít jen při přepětí. Určete pravděpodobnost poruchy přístroje.

Řešení:
Označme si

$A =$ "*napětí překročí standardní hodnotu*",

$B =$ "*dojde k poruše přístroje*".

Ze zadání víme, že $P(A) = p_1$, $P(B|A) = p_2$ a $B \subseteq A$. Zajímá nás hodnota $P(B)$. Ze vztahu $B \subseteq A$ plyne, že $B \cap A = B$. Pomocí definice relativní pravděpodobnosti teď můžeme psát

$$P(B) = P(B \cap A) = \frac{P(B \cap A)}{P(A)} \cdot P(A) = P(B|A) \cdot P(A) = p_2 \cdot p_1 \ .$$

2.3 (nezávislé jevy)

Dva střelci střílí na terč po jedné ráně. Pravděpodobnost, že se první trefí je $p_1 = 0.7$. Pravděpodobnost, že se trefí druhý střelec je $p_2 = 0.8$. Jaká je pravděpodobnost, že

(a) alespoň jeden střelec zasáhne cíl?

(b) první střelec se trefí a druhý ne?

Řešení:
Uvažujme jevy

$S_1 =$ "*první střelec se trefí*",

$S_2 =$ "*druhý střelec se trefí*".

Tyto jevy jsou nezávislé, $P(S_1) = 0.7$ a $P(S_2) = 0.8$.

(a) Pro jev

$A =$ "*alespoň jeden střelec zasáhne cíl*"

máme $A = S_1 \cup S_2$ a tedy

$$P(A) = P(S_1 \cup S_2) = P(S_1) + P(S_2) - P(S_1 \cap S_2) \overset{\text{(nezav.)}}{=}$$
$$\overset{\text{(nezav.)}}{=} P(S_1) + P(S_2) - P(S_1) \cdot P(S_2) = 0.7 + 0.8 - 0.7 \cdot 0.8 = 0.94 \ .$$

(b) Pro jev

$B =$ "*první střelec se trefí a druhý ne*"

máme $B = S_1 \cap \overline{S_2}$ a tedy

$$P(B) = P(S_1 \cap \overline{S_2}) \overset{\text{(nezav.)}}{=} P(S_1) \cdot P(\overline{S_2}) = 0.7 \cdot (1 - 0.8) = 0.14 \ .$$

2.4 ((ne)závislost jevů)

Automat vyrábí podložky ve tvaru obdélníka. Tolerance v šířce není dodržena v 8%, tolerance v délce v 7% a v obou rozměrech ve 3% případů.

(a) Rozhodněte, zda jsou porušení tolerance v délce a v šířce závislé nebo nezávislé jevy.

(b) Vypočtěte pravděpodobnost toho, že náhodně vybraný kus má oba rozměry v toleranci.

Řešení:

Uvažujme jevy

$A =$ "*není dodržena tolerance v šířce*",

$B =$ "*není dodržena tolerance v délce*".

Ze zadání máme, že $P(A) = 0.08$, $P(B) = 0.07$ a $P(A \cap B) = 0.03$.

(a) Protože máme $P(A \cap B) = 0.03 \neq 0.08 \cdot 0.07 = P(A) \cdot P(B)$, jsou jevy A a B závislé.

(b) Pro jev

$C =$ "*oba rozměry jsou v toleranci*"

máme $C = \overline{A} \cap \overline{B} = \overline{A \cup B}$. Takže dostáváme

$$P(C) = P(\overline{A \cup B}) = 1 - P(A \cup B)$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.08 + 0.07 - 0.03 = 0.12$$

a tudíž

$$P(C) = 1 - P(A \cup B) = 1 - 0.12 = 0.88\ .$$

2.5 (geometrická pravděpodobnost)

Tyč délky ℓ se náhodně rozpadne na 3 části. Jaká je pravděpodobnost, že z částí lze sestavit trojúhelník?

Řešení:

V rámci geometrické pravděpodobnosti pracujeme vždy v $\mathbb{R}^n$, kde máme obvyklý n-rozměrný objem $\mathrm{vol}(\cdot)$ (a tudíž pracujeme s množinami, kterým nějaký objem přiřadit lze - tzv. borelovské). Pojmu σ-algebra (a dalším definicím spojeným s pravděpodobností) se tak prostě nelze vyhnout, pokud máme používat velikost plochy nebo objem množiny (a později i integrovat funkce).

Kolmogorovým modelem v tomto případě pak bude trojice $(\Omega, \mathcal{B}, P)$, kde $\Omega \subseteq \mathbb{R}^n$ je borelovská množina taková, že $\mathrm{vol}(\Omega) < \infty$, $\mathcal{B}$ je σ-algebra tvořena všemi borelovskými množinami obsaženými v Ω a $P(A) = \frac{\mathrm{vol}(A)}{\mathrm{vol}(\Omega)}$ pro $A \in \mathcal{B}$.

V našem případě se tyč se rozpadne na části o délkách a, b a c, kde $0 < a, b, c$ a $a + b + c = \ell$. Za jevové pole si zvolíme

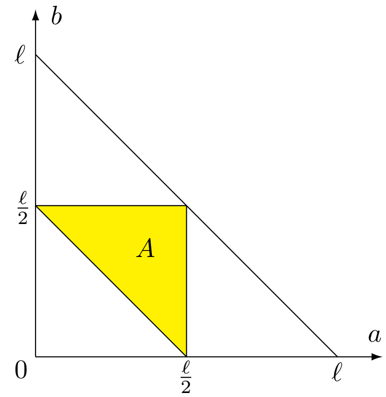
$$\Omega = \left\{(a, b) \in \mathbb{R}^2 \mid 0 < a, b \ \ \& \ \ a + b < \ell\right\},$$

kde původní elementární jev (a, b, c) popíšeme pouze prvními dvěma složkami (a, b) a třetí je jednoznačně určena jako $c = \ell - (a+b)$. Množina Ω je rovnoramenný pravoúhlý trojúhelník s odvěsnami délky ℓ. Jeho plocha tak je $\mathrm{vol}(\Omega) = \frac{\ell^2}{2}$. Abychom mohli sestavit trojúhelník, musí platit trojúhelníková nerovnost. Zajímá nás tedy tudíž jev

$$A = \{(a, b) \in \Omega \mid a + b > \ell - (a + b) \ \ \wedge \ \ b + \ell - (a + b) > a \ \ \wedge \ \ a + \ell - (a + b) > b\} =$$

$$= \left\{(a, b) \in \Omega \mid a + b > \tfrac{\ell}{2} \ \ \wedge \ \ a, b < \tfrac{\ell}{2}\right\}.$$

Množina A vytváří v množině Ω trojúhelník, jehož vrcholy jsou středy stran trojúhelníku Ω.



Velikost plochy A tak zřejmě je $\text{vol}(A) = \frac{1}{4}\text{vol}(\Omega)$. Proto máme

$$P(A) = \frac{\text{vol}(A)}{\text{vol}(\Omega)} = \frac{1}{4} = 0.25 \ .$$

K výpočtu lze použít také jevové pole

$$\Omega' = \big\{(a,b,c) \in \mathbb{R}^3 \mid 0 < a,b,c \quad \wedge \quad a+b+c = \ell\big\} \ .$$

To je rovnostranný trojúhelník o straně délky ℓ. Jeho plocha tak je $\text{vol}(\Omega') = \frac{\sqrt{3}}{4}\ell^2$. Odpovídající jev sestrojení trojúhelníku pak je

$$A' = \big\{(a,b,c) \in \Omega' \mid a+b > c \quad \wedge \quad b+c > a \quad \wedge \quad a+c > b\big\} \ ,$$

který opět v množině Ω' vytváří trojúhelník, jehož vrcholy jsou středy stran trojúhelníku Ω'. Velikost plochy A' je opět $\text{vol}(A') = \frac{1}{4}\text{vol}(\Omega')$. Proto znovu máme $P(A') = \frac{\text{vol}(A')}{\text{vol}(\Omega')} = \frac{1}{4}$.

Co není u podobných příkladů ihned zřejmé, je to, zda a proč budou různé přístupy dávat stejný výsledek. Obecně tomu tak být nemusí. Zde ale zřejmě ano. Důvod je obecněji ten, že máme zobrazení $\varphi : \Omega \to \Omega'$,

$$\varphi(a,b) = (a, b, \ell - a - b)$$

které parametrizuje množinu Ω' pomocí původní množiny Ω a přitom platí $\varphi(A) = A'$. Toto zobrazení je vlastně "natažení"trojúhelníku Ω do podoby trojúhelníku Ω'. Množina A se přitom natáhne stejným způsobem (do množiny A') a proto poměry velikostí zůstanou zachovány. Tedy pravděpodobnost vyjde stejně.

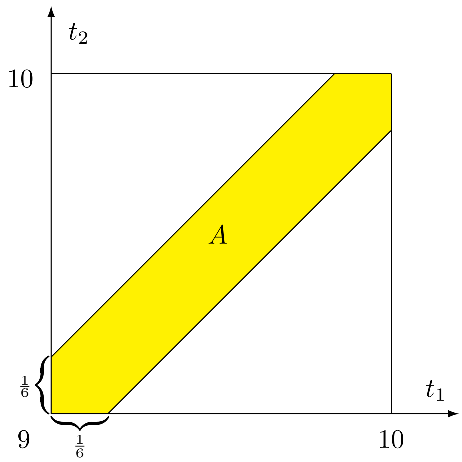
2.6 (geometrická pravděpodobnost)

Dva přátelé A a B si domluví schůzku mezi 9.00 a 10.00. Jejich příchody na dané místo jsou náhodné v rámci smluveného časového intervalu. Každý bude čekat 10 minut a pak odchází. Jaká je pravděpodobnost, že dojde k setkání?

Řešení:
Jako elementární jev si zvolíme dvojici (t_1, t_2), která znamená příchody jednotlivých osob v jednotkách *hodin*. Tedy $\Omega = \langle 9, 10\rangle \times \langle 9, 10\rangle$. Jev $A \subseteq \Omega$ setkání obou přátel bude pak

$$A = \{(t_1, t_2) \in \Omega \mid \ |t_1 - t_2| \leq 1/6\}$$

(za jednotku jsme si zvolili hodinu, takže 10 min $= \frac{1}{6}$ hod). Z grafického znázornění množin v $\mathbb{R}^2$



snadno zjistíme, že $\mathrm{vol}(A) = 1 - (\frac{5}{6})^2 = \frac{11}{36}$ a $\mathrm{vol}(\Omega) = 1$, takže

$$P(A) = \frac{11}{36} \ .$$

2.7 (Kolmogorův model)

Zjistěte, zda $(\Omega, \mathcal{A}, P)$ je Kolmogorův model pravděpodobnosti, je-li dáno:

- $\Omega = \{1, 2, 3\}$,
- $\mathcal{A} = \big\{\emptyset, \{1, 2\}, \{1, 3\}, \{2\}, \{3\}, \{1, 2, 3\}\big\}$,
- $P(A) = \frac{1}{3}|A|$, $A \in \mathcal{A}$ (kde $|A|$ je počet prvků množiny A).

Řešení:

Pro Kolmogorův model je potřeba ověřit, že $\mathcal{A}$ je σ-algebra:

- $\emptyset \in \mathcal{A}$,
- $A \in \mathcal{A} \ \Rightarrow \overline{A} \in \mathcal{A}$,
- $\{A_n \mid n \in \mathbb{N}\} \subseteq \mathcal{A} \ \Rightarrow \ \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$,

a že $P : \mathcal{A} \to \langle 0, 1\rangle$ je pravděpodobnost:

- $P(\Omega) = 1$,
- $\{A_n \mid n \in \mathbb{N}\} \subseteq \mathcal{A}$ & A_n jsou navzájem disjunktní $\Rightarrow \ P\big(\bigcup_{n\in\mathbb{N}} A_n\big) = \sum_{n\in\mathbb{N}} P(A_n)$,

První dvě podmínky pro σ-algebru jsou zřejmě splněny, poslední ne, protože

$$\{2\}, \{3\} \in \mathcal{A}, \quad \text{ale} \quad \{2\} \cup \{3\} \notin \mathcal{A} \ .$$

Množina $\mathcal{A}$ tedy *není* σ-algebra a $(\Omega, \mathcal{A}, P)$ proto *není* Kolmogorův model.

Tuto nedokonalost, ale můžeme spravit tak, že k $\mathcal{A}$ přidáme prvky, které chybí: tedy prvek $\{2\}\cup\{3\} = \{2, 3\}$ a $\overline{\{2, 3\}} = \{1\}$. Dostaneme tak celou potenční množinu $\mathcal{A}' = \exp(\Omega)$, která σ-algebrou určitě je.

Teď ještě ukážeme, že $P : \mathcal{A}' \to \langle 0, 1\rangle$ (stále uvažujeme stejný předpis) je v tomto případě pravděpodobnost. Zřejmě

$$P(\Omega) = \frac{3}{3} = 1$$

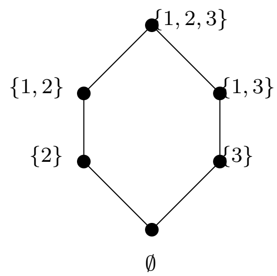
a dále pro $\{A_n \mid n \in \mathbb{N}\} \subseteq \exp(\Omega)$ navzájem disjunktní máme

$$P\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \frac{1}{3}\left|\bigcup_{n\in\mathbb{N}} A_n\right| = \frac{1}{3}\sum_{n\in\mathbb{N}} |A_n| = \sum_{n\in\mathbb{N}} P(A_n)\ .$$

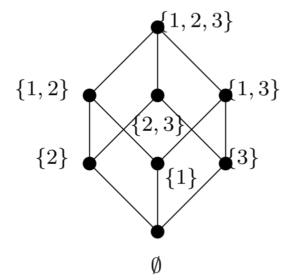
To, že jsme ukázali, že P je pravděpodobnost nakonec není žádné překvapení, protože to celé je prostě Laplaceův model pravděpodobnosti (tj. počet příznivých případů ku počtu všech.)

Uspořádané množiny (v našem případě je uspořádání dáno inkluzí) můžeme ještě zakreslit tzv. Hasseovým diagramem (větší prvky se zakreslují nad menší a spojují se čárkou, pokud už mezi nimi žádné další prvky nejsou). Dostáváme tak:

$\mathcal{A}$:



exp(Ω) :



To, že jsme ve druhém případě dostali obrázek, který vypadá jako krychle, není náhoda. Konečné σ-algebry budou mít vždy Hasseův diagram ve tvaru vícerozměrné krychle.