1. cvičení z PST

23. září 2020

Uvažujme výběr z n různých předmětů, které vybíráme k-krát. Počet všech jednotlivých možností pro různé způsoby výběru uvádí následující tabulka:

Výběr	bez vracení (bez opakování)	s vracením (s opakováním)
uspořádaný (variace)	$n(n-1)\cdots(n-k+1)$	n^k
neuspořádaný (kombinace)	$\binom{n}{k}$	$\binom{n+k-1}{k}$

Odpovídající variace (příp. kombinace) se označují jako "k-té třídy z n prvků".

1.1 (kombinace s opakováním)

Ukažte, že počet všech kombinací s opakováním k-té třídy z n prvků je roven počtu všech nezáporných celočíselných řešení $(x_1, \ldots, x_n)$ rovnice

$$x_1 + x_2 + \cdots + x_n = k$$

a že tento počet je právě $\binom{n+k-1}{k}$.

Řešení:

Dané řešení $(x_1, \ldots, x_n)$ můžeme jednoznačně vyjádřit také tak, že pro k nerozlišitelných koulí a n různých přihrádek bude x_i znamenat počet koulí v i-té přihrádce.

Tato situace odpovídá kombinacím bez opakování (máme n druhů prvků v dostatečném množství a vybíráme z nich k prvků, při výběru přitom rozlišujeme vždy, jen kolik je od kterého druhů prvků).

Rozdělení koulí do přihrádek můžeme zakódovat pomocí $n-1$ nerozlišitelných přepážek a k nerozlišitelných koulí. To ale odpovídá permutacím s opakováním (pro dva druhy předmětů). Celkový počet řešení tak je kombinační číslo

$$\binom{n-1+k}{k} = \frac{(n-1+k)!}{(n-1)! \cdot k!} = \binom{n-1+k}{n-1} .$$

1.2 (variace s opakováním)

Kolik různých slov můžeme vytvořit přeskupením písmen slov:

(a) MISSISSIPPI,

(b) ANANAS,

(c) PROBLÉM (kde písmena B a R nestojí vedle sebe)?

Řešení:
Jde o permutace s opakováním. Pro n prvků, z nichž

- k_1 je 1. druhu, k_2 je 2. druhu atd. až k_ℓ je ℓ-tého druhu,
- přičemž $k_1 + \cdots + k_\ell = n$ a prvky stejného druhu jsou navzájem nerozlišitelné,

je počet všech uspořádaných n-tic roven $\frac{n!}{k_1! \cdots k_\ell!}$.

(a) Písmeno M se vyskytuje 1×, písmeno I pak 4×, písmeno S také 4× a písmeno P máme 2×. Počet znaků ve slově je $1 + 4 + 4 + 2 = 11$. Počet různých slov je tedy

$$\frac{11!}{1! \cdot 4! \cdot 4! \cdot 2!} = \frac{11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5}{4 \cdot 3 \cdot 2 \cdot 2} = 11 \cdot 10 \cdot 9 \cdot 7 \cdot 5 = 34650 \ .$$

(b) Podobně jako v (a): Písmeno A se vyskytuje 3×, písmeno N pak 2× a písmeno S máme 1×. Počet znaků ve slově je $3 + 2 + 1 = 6$. Počet různých slov je tedy

$$\frac{6!}{3! \cdot 2! \cdot 1!} = 5 \cdot 4 \cdot 3 = 60 \ .$$

(c) Jednodušší je zjistit počet zbylých slov, tj. těch, kde písmena B a R *stojí* vedle sebe, a odečíst je od počtu všech přesmyček původního slova. Skupinu slov B a R budeme považovat za jeden nedělitelný prvek, který se může vyskytovat ve 2! různých stavech: BR nebo RB. Budeme tedy permutovat 6 prvků: P, O, L, É, M a X (kde X={B, R}). Počet slov, kde písmena B a R *stojí* vedle sebe je tak $2! \cdot 6!$.

Počet slov, kde písmena B a R *nestojí* vedle sebe tudíž bude

$$7! - 2! \cdot 6! = 5 \cdot (6!) = 3600 \ .$$

1.3 (variace bez opakování)

Na jedné poličce je náhodně rozestavěno 10 knih. Jaká je pravděpodobnost, že 3 určité knihy jsou postaveny vedle sebe?

Řešení:
Podobně jako v předchozím příkladu budeme dané 3 knihy považovat za jednu nedělitelnou skupinu. Dostaneme tak 8 prvků. V rámci nedělitelné skupiny máme 3! možností. Počet příznivých (tj. požadovaných) rozestavení je proto

$$3! \cdot 8! = 241\ 920 \ .$$

Počet všech (libovolných) rozestavení je 10!. Pravděpodobnost tedy je

$$p = \frac{\text{“počet příznivých možností”}}{\text{“počet všech možností”}} = \frac{3! \cdot 8!}{10!} = \frac{1}{5 \cdot 3} \doteq 0.067 \ .$$

1.4 (kombinace bez opakování)

Kolik různých volejbalových týmu lze složit ze skupiny 15 chlapců a 6 dívek, pokud v týmu vždy hrají 4 chlapci a 2 děvčata?

Jaká je pravděpodobnost, že v náhodně vybrané skupině 6 osob (z výše uvedené skupiny chlapců a děvčat) budou alespoň 4 chlapci?

Řešení:
Každý požadovaný tým je určený množinou chlapců, kterou lze zvolit $\binom{15}{4}$ způsoby, a množinou děvčat, kterou lze zvolit $\binom{6}{2}$ způsoby. Hledaný počet týmu je tedy

$$\binom{15}{4} \cdot \binom{6}{2} = \frac{15 \cdot 14 \cdot 13 \cdot 12}{4 \cdot 3 \cdot 2 \cdot 1} \cdot \frac{6 \cdot 5}{2 \cdot 1} = 20475 \ .$$

Dále, všech osob je $15 + 6 = 21$. Vybíráme 6 osob. Prostorem všech možných rovnocenných výsledků je tedy

$$\Omega = \text{“}\textit{všechny neuspořádané } 6\textit{-ce utvořené z } 21 \textit{ osob}\text{”}$$

tedy, kombinace bez opakování 6-té třídy z 21 prvků. Tudíž, $|\Omega| = \binom{21}{6}$. Hledáme pravděpodobnost jevu

$$A = \text{“ } 6\textit{-tice z } \Omega\textit{, které obsahují alespoň } 4 \textit{ chlapce}\text{”} \ .$$

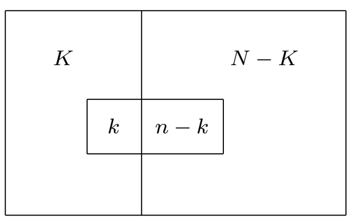
Velikost množiny A spočítáme rozdělením na 6-tice z Ω, které obsahují právě

- 4 chlapce: těch je $\binom{15}{4} \cdot \binom{6}{2}$,
- 5 chlapců: těch je $\binom{15}{5} \cdot \binom{6}{1}$,
- 6 chlapců: těch je $\binom{15}{6} \cdot \binom{6}{0}$.

Pravděpodobnost tedy je

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\binom{15}{4} \cdot \binom{6}{2} + \binom{15}{5} \cdot \binom{6}{1} + \binom{15}{6} \cdot \binom{6}{0}}{\binom{21}{6}} \doteq 0.8016 \ .$$

Poznámka: V tomto příkladu jsme setkali s tzv. *hypergeometrickým* rozdělením.



To se objevuje tehdy, když

- z množiny, která má N prvků (zde $N = 21$),
- z nichž právě K má nějakou vlastnost $\mathcal{V}$ (zde $K = 15$ a $\mathcal{V} =$"osoba je chlapec"),
- chceme vybrat právě n prvků (zde $n = 6$)

a ptáme se, s jakou pravděpodobností bude právě k z nich mít vlastnost $\mathcal{V}$.

Tato pravděpodobnost je dána podílem příznivých možností ku všem. Příznivé jsou dány počtem způsobů jak vybrat k prvků (s $\mathcal{V}$) z K prvků (s $\mathcal{V}$) násobeno počtem způsobů jak vybrat zbytek, tj. $n - k$ prvků (bez $\mathcal{V}$) z $N - K$ prvků (bez $\mathcal{V}$). Celkem tedy

$$\frac{\binom{K}{k} \cdot \binom{N-K}{n-k}}{\binom{N}{n}}$$

přičemž rozsah proměnné k je určen pomocí

$$\max\{0, n + K - N\} \leq k \leq \min\{n, K\}$$

tedy

$$0 = \max\{0, 6 + 15 - 21\} \leq k \leq \min\{6, 15\} = 6 \ .$$

Tuto podmínku dostaneme ihned z nerovností

$$k \leq K, \quad n - k \leq N - K,$$

$$k \leq n, \quad n - k \leq n \ .$$

1.5 (kombinace bez opakování)

Na party se sešlo 14 studentů, z toho 8 chlapců a 6 děvčat. Jaká je pravděpodobnost, že v náhodně vybrané čtveřici osob

(a) bude právě jeden chlapec?

(b) budou alespoň dva chlapci?

Řešení:
Prostor všech možných rovnocenných výsledků je

$$\Omega = \text{“}\textit{všechny neuspořádané } 4\textit{-ce utvořené ze } 14 \textit{ osob”}$$

tedy všechny 4-prvkové podmnožiny 14-ti prvkové množiny (neboli *kombinace* 4*-té třídy ze* 14 *prvků bez opakování*). Jejich počet je $|\Omega| = \binom{14}{4} = \frac{14!}{4!(14-4)!}$.

(a) Každou čtveřici z jevu

$$A = \text{“}\textit{čtveřice z } \Omega\textit{, ve kterých je právě 1 chlapec”}$$

můžeme popsat pomocí trojice děvčat, kterých je $\binom{6}{3}$, a pomocí vybraného chlapce, kterých je $\binom{8}{1} = 8$. Tudíž hledaná pravděpodobnost je

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\binom{8}{1} \cdot \binom{6}{3}}{\binom{14}{4}} \doteq 0.16.$$

(b) Pravděpodobnost jevu

$$B = \text{“}\textit{čtveřice z } \Omega\textit{, ve kterých jsou alespoň 2 chlapci”}$$

je jednodušší spočítat pomocí doplňkového jevu

$$\overline{B} = \text{“}\textit{čtveřice z } \Omega\textit{, ve kterých je nejvýše 1 chlapec”}\ .$$

Velikost množiny $\overline{B}$ spočítáme rozdělením na čtveřice, které

- obsahují právě 1 chlapce: těch je $\binom{8}{1} \cdot \binom{6}{3}$,
- neobsahují žádného chlapce: těch je $\binom{8}{0} \cdot \binom{6}{4}$.

Pravděpodobnost tedy je

$$P(\overline{B}) = \frac{|\overline{B}|}{|\Omega|} = \frac{\binom{8}{0} \cdot \binom{6}{4} + \binom{8}{1} \cdot \binom{6}{3}}{\binom{14}{4}} = \frac{1 \cdot 15 + 8 \cdot 20}{1001} \doteq 0.1748$$

a

$$P(B) = 1 - P(\overline{B}) \doteq 1 - 0.1748 = 0.8252\ .$$

Opět se zde jedná o tzv. *hypergeometrické* rozdělení (viz příklad **1.4**).

1.6 (výběr bez opakování)

Jaká je pravděpodobnost, že ve skupině n lidí mají alespoň dva narozeniny ve stejný den? (Neuvažujte přestupné roky a předpokládejte, že se během celého roku děti rodí rovnoměrně.)

Řešení:
První část z předpokladů znamená, že počet všech dnů v roce pro nás bude vždy 365. Druhá část předpokladů pak znamená, že všechny dny v roce považujeme za rovnocenné. Tedy pravděpodobnost, že se daný člověk narodí v daný den v roce bude pro všechny dny stejná. To, že pracujeme se skupinou n lidí si také můžeme ekvivalentně představit tak, že máme urnu s lístky s čísly od 1 do 365 (představujícími očíslované dny v roce) a my z ní n-krát opakovaně budeme losovat čísla s tím, že lístky vždy vrátíme zpět.

Výsledky pokusu jsou tudíž uspořádané n-tice s hodnotami od 1 do 365. Tedy $\Omega = \{1, \ldots, 365\}^n$ (tj. kartézský součin množiny $\{1, \ldots, 365\}$ a to celkem n-krát, podobně jako třeba zapisujeme $\mathbb{R}^n$).

Nechť A je jev, že ve skupině n lidí mají alespoň dva narozeniny ve stejný den. Pak jev $\overline{A}$ znamená, že ve skupině n lidí má každý člověk narozeniny v jiný den. Jde tedy o variace bez opakování třídy n z 365 prvků.

$$|\Omega| = 365^n$$

$$|\overline{A}| = 365 \cdot 364 \cdot \cdots \cdot (365 - n + 1)$$

a tedy

$$P(\overline{A}) = \frac{365 \cdot 364 \cdot \cdots \cdot (365 - n + 1)}{365^n},$$

a tudíž

$$P(A) = 1 - P(\overline{A}) = 1 - \frac{365 \cdot 364 \cdot \cdots \cdot (365 - n + 1)}{365^n}.$$

Pro zajímavost se ještě podíváme na přibližnou hodnotu této pravděpodobnosti. Pro jednoduchost označme $H = 365$. Pak máme

$$P(A) = 1 - \frac{H \cdot (H-1) \cdot \cdots \cdot (H - (n-1))}{H^n} = 1 - \left(1 - \frac{1}{H}\right) \ldots \left(1 - \frac{n-1}{H}\right) =$$

$$= 1 - e^{\ln\left(1-\frac{1}{H}\right)\ldots\left(1-\frac{n-1}{H}\right)} = 1 - e^{\sum\limits_{k=1}^{n-1} \ln\left(1-\frac{k}{H}\right)}$$

Použijeme teď lineární aproximaci funkce

$$f(x) = \sum_{k=1}^{n-1} \ln(1 - kx)$$

v bodě $x_0 = 0$, kterou pak vyčíslíme pro $x = \frac{1}{H}$, která je blízká nule, a to jako $f(x) \approx f(0) + f'(0) \cdot (x - 0)$. Tedy

$$f'(x) = \left(\sum_{k=1}^{n-1} \ln(1 - kx)\right)' = -\sum_{k=1}^{n-1} \frac{k}{1 - kx}$$

$$f'(0) = -\sum_{k=1}^{n-1} k = -\frac{n(n-1)}{2}$$

a

$$\sum_{k=1}^{n-1} \ln\left(1 - \frac{k}{H}\right) = f\left(\frac{1}{H}\right) \approx 0 + f'(0) \cdot \frac{1}{H} = -\frac{n(n-1)}{2H}\ .$$

Odsud máme, že

$$P(A) \doteq 1 - e^{-\frac{n(n-1)}{2 \cdot 365}}$$

a ukázku některých hodnot v tabulce:

n	23	24	25	26	27	50
$P(A)$	50%	53.05%	56.04%	58.95%	61.77%	96.51%

1.7 (srovnání výběru s opakováním a bez opakování)
V balíčku máme 32 karet, z toho 4 esa. Dvakrát za sebou vytáhneme náhodně jednu kartu. Stanovte

pravděpodobnost jevu

$$A = \text{“}\textit{alespoň jedna z vytažených karet je eso}\text{”},$$

jestliže po prvním tahu kartu

(1) vrátíme,

(2) nevrátíme

zpět do balíčku.

Řešení:
(1) Výsledky pokusu jsou opět uspořádané dvojice (množina Ω_1). První člen dvojice odpovídá kartě vytažené v prvním tahu a druhý člen kartě vytažené v druhém tahu. V prvním tahu můžeme kartu vytáhnout 32 způsoby. Protože vytaženou kartu vracíme zpět do urny, i v druhém tahu máme 32 možností. Počet všech možných případů je tedy $|\Omega_1| = 32^2$. Příznivým případům odpovídají tahy (libovolná karta - eso), (eso - libovolná karta), (eso - eso). Počet příznivých případů je $|A_1| = 28 \cdot 4 + 4 \cdot 28 + 4 \cdot 4$. Hledaná pravděpodobnost je rovna

$$P(A_1) = \frac{|A_1|}{|\Omega_1|} = \frac{28 \cdot 4 + 4 \cdot 28 + 4 \cdot 4}{32^2} = \frac{15}{64} \doteq 0.2344.$$

Jednodušeji se k výsledku můžeme dostat přes doplňkový jev

$$\overline{A_1} = \Omega_1 \setminus A_1 = \text{“žádná z vytažených karet není eso”},$$

Pro pravděpodobnost pak platí, že

$$P(\overline{A_1}) = \frac{|\Omega_1 \setminus A_1|}{|\Omega_1|} = \frac{|\Omega_1| - |A_1|}{|\Omega_1|} = 1 - P(A_1)$$

Počet prvků $\overline{A_1}$ je pak (analogicky jako u Ω_1) roven 28^2, takže

$$P(A_1) = 1 - P(\overline{A_1}) = 1 - \frac{28 \cdot 28}{32 \cdot 32} = 1 - \frac{7 \cdot 7}{8 \cdot 8} = \frac{15}{64} \doteq 0.2344.$$

(2) Tentokrát z uspořádaných dvojic karet musíme vyloučit ty, kde první i druhá karta jsou stejné (dostaneme tak množinu Ω_2). Počet možných případů je vzhledem k tomu, že po prvním tahu kartu nevrátíme, $|\Omega_2| = 32 \cdot 31$. Příznivým případům odpovídají opět tahy (libovolná karta - eso), (eso - libovolná karta), (eso - eso). Počet příznivých případů je nyní $|A_2| = 28 \cdot 4 + 4 \cdot 28 + 4 \cdot 3$. Hledaná pravděpodobnost je rovna

$$P(A_2) = \frac{|A_2|}{|\Omega_2|} = \frac{2 \cdot 4 \cdot 28 + 4 \cdot 3}{32 \cdot 31} = \frac{59}{248} \doteq 0.2379.$$

Přes doplňkový jev je to opět jednodušší:

$$\overline{A_2} = \text{“žádná z vytažených karet není eso”},$$

$$|\overline{A_2}| = 28 \cdot 27$$

$$P(A_2) = 1 - P(\overline{A_2}) = 1 - \frac{28 \cdot 27}{32 \cdot 31} = \frac{59}{248} \doteq 0.2379.$$

Kromě toho je vidět, že pravděpodobnosti se v případech vracení (0.2344) i nevracení (0.2379) příliš neliší. Je to proto, že pokud máme velké množství N (zde $N = 32$), ze kterého taháme pouze k-krát (zde $k = 2$), tj. ”několikrát”v porovnání s tím, jak velké množství máme, neboli $k << N$, tak pravděpodobnost, že bychom opakovaně vytáhli znovu tentýž předmět je zanedbatelná. Tudíž obě pravděpodobnosti se budou téměř shodovat.

1.8 Los za 60 korun obsahuje deset políček, která postupně v námi zvoleném pořadí odkrýváme. Na dvou políčkách je znak V (výhra), na dvou políčkách je znak P (prohra) a na zbývajících šesti není nic. Pokud odkryjeme obě políčka se znakem V a žádné políčko se znakem P, vyhráváme 300 korun.

1. Jaká je pravděpodobnost výhry?
2. Je cena losu férová?

Řešení:
Nejdříve si musíme ujasnit, jak situaci popsat. I zde můžeme využít Laplaceovu pravděpodobnost, tj. takové výsledky, které budeme považovat za rovnocenné. Mohli bychom uvažovat nejdříve všechny možné losy, tj. všechny možnosti, jak mohou být uspořádána políčka se symboly a ke každému tomuto losu pak ještě všechny možné způsoby, jak políčka postupně odkrývat. Protože ale každý los je jen permutací jiného losu, bude počet všech možných odkrývání pro každý los stejný a stejný bude i počet výherních kombinací. Abychom si tedy ušetřili práci, vezmeme si pevně jeden z losů a k němu spočítáme všechny možné způsoby odkrývání a počet výherních možností.

Nyní je potřeba si ujasnit, které možnosti považujeme za rovnocenné, tj. jak má vypadat množina všech možných výsledků. Místo postupného odkrývání políček si můžeme představovat, že pořadí políček si losujeme - a ještě lépe, že si symboly, které jsou na losu zakryté, prostě postupně náhodně vytahujeme z urny. Symboly považujeme pochopitelně za rozlišitelné (protože jsou v urně fyzicky každý samostatně a protože každý odpovídá jinému místu na losu). Stejně jako v jiných příkladech, i zde můžeme považovat všechny možné posloupnosti vytažení všech políček za rovnocenné. Abychom dodrželi podmínky, za kterých se vyhrává, řekneme, že výherní posloupnosti jsou ty, kdy vytáhneme dva symboly V dříve, než nějaký ze symbolů P. Neúplné posloupnosti délky menší než 10 (a takové, kdy jsme odkryli jen část symbolů a výsledek je z nich už zřejmý) není vhodné použít, protože ty nejsou z hlediska výskytu rovnocenné - mají totiž obecně různý počet doplnění na “úplné” posloupnosti a tedy jejich pravděpodobnost by byla úměrná počtu těchto doplnění. Tím bychom si jen zkomplikovali výpočet.

No a konečně, poslední zjednodušení: To, jestli posloupnost deseti znaků bude vyhrávající, určuje pouze pořadí znaků P a V mezi sebou. Jestliže si tedy tyto znaky označíme pro rozlišení jako V_1, V_2, P_1, P_2 a prázdná políčka jako $\circ_1, \ldots, \circ_6$, pak ke každému zvolenému uspořádání (např. (V_2, P_1, P_2, V_1), což je zrovna nevyhrávající) budeme mít vždy právě N posloupností o 10 členech, které v sobě toto pořadí znaků obsahují (takováto posloupnost je např. $(\circ_3, V_2, P_1, \circ_5, \circ_2, P_2, \circ_6, \circ_1, \circ_4, V_1)$). Toto N pochopitelně nezávisí na zvoleném pořadí výše uvedených 4 znaků a nemusíme ho ani počítat, protože v rámci pravděpodobnosti, o kterou nám jde, se pak ve zlomku zkrátí.

Nakonec tedy místo s posloupostmi délky 10 můžeme počítat jen s posloupnostmi délky 4 složenými ze znaků V_1, V_2, P_1, P_2. Tyto posloupnosti nyní můžeme považovat za rovnocenné výsledky (právě díky společnému N).

(1) Prostor Ω všech posloupností délky 4 má zřejmě $4! = 24$ prvků. Jakou velikost má nyní jev

$$A = \text{"}\textit{posloupnost je výherní}\text{"?}$$

Výherní jsou ty, kde nejdříve jsou pouze znaky pro výhru V a pak teprve následují znaky P. Těch je tedy $2! \cdot 2! = 4$, tj. $|A| = 4$.

Máme tedy $P(A) = \frac{|A|}{|\Omega|} = \frac{4}{24} = \frac{1}{6} \doteq 0.167$.

(2) Férova cena losu bude taková, která bude odpovídat střední ceně hodnoty, kterou díky losu získáme. Pro posloupnost $\omega \in \Omega$ si označme $H(\omega)$ hodnotu, kterou díky losu získáme, tj.

$$H(\omega) = \begin{cases} 300 \text{ Kč}, & \omega \textit{ je výherní} \\ 0 \text{ Kč}, & \omega \textit{ je nevýherní.} \end{cases}$$

Střední hodnotu pak určíme jako

$$E(H) = \frac{\sum\limits_{\omega \in \Omega} H(\omega)}{|\Omega|} = \frac{\left(\sum\limits_{\omega \in A} 300\right) + \left(\sum\limits_{\omega' \in \overline{A}} 0\right)}{|\Omega|} = 300 \cdot \frac{|A|}{|\Omega|} + 0 \cdot \frac{|\overline{A}|}{|\Omega|} =$$

$$= 300 \cdot P(A) + 0 \cdot P(\overline{A}) = 300 \cdot \frac{1}{6} + 0 \cdot \frac{5}{6} = 50 \text{ Kč}$$

Cena losu tak neni férová vůči tomu, kdo si los kupuje, protože zaplatí 60 Kč, ale vyhraje průměrně jen 50 Kč.