Pravděpodobnost a statistika - cvičení

Simona Domesová

simona.domesova@vsb.cz

místnost: RA310 (budova CPIT)

web: http://homel.vsb.cz/~dom0015

Cíle předmětu

- vyhodnocování dat pomocí statistických metod
- porozumění výsledkům statistických analýz
- použití statistického softwaru

Podmínky získání zápočtu

- testy na cvičeních: 10 testů po 2 bodech = 20 bodů (minimum 6 bodů)
- domácí úkoly: 4 úkoly po 5 bodech = 20 bodů (minimum 10 bodů)
- aktivní účast na cvičeních 80%

$\longrightarrow$ celkem za zápočet: 40 bodů (minimum 20 bodů)

Zkouška

- praktická část: 50 bodů
- teoretická část: 10 bodů

Materiály k předmětu

- http://k470.vsb.cz/litschmannova/vyuka/statistika/

Cvičení

- pracovní listy
  - http://homel.vsb.cz/~lit40/STA1/Materialy/Biostatistika_PL.pdf
- vzorce a tabulky
  - http://homel.vsb.cz/~lit40/STA1/Materialy/Vzorce-a-tabulky.pdf
- software: RkWard

Pravděpodobnost a statistika - cvičení 1

Kombinatorika

Simona Domesová
http://homel.vsb.cz/~dom0015

Variace

= uspořádaná k-tice vybraná z n prvků, každý prvek může být vybrán jen jednou

- Postupně vybereme k prvků z n prvkové množiny, přičemž záleží na pořadí výběru.
- Počet všech variací:

$$V(n, k) = \frac{n!}{(n-k)!}$$

- $n = 3$, $k = 2$:

$$\{\mathrm{A},\ \mathrm{B},\ \mathrm{C}\} \longrightarrow (\mathrm{A},\ \mathrm{B}),\ (\mathrm{A},\ \mathrm{C}),\ (\mathrm{B},\ \mathrm{A}),\ (\mathrm{B},\ \mathrm{C}),\ (\mathrm{C},\ \mathrm{A}),\ (\mathrm{C},\ \mathrm{B})$$

- např. kolik uspořádaných trojic karet můžu sestavit, pokud mám k dispozici balíček 52 různých karet?

Permutace

$=$ libovolné uspořádání n různých prvků do posloupnosti

- Počet všech permutací n prvkové množiny:

$$P(n) = V(n, n) = n!$$

- $n = 3$:

$$\{\text{A, B, C}\} \longrightarrow \text{(A, B, C), (A, C, B), (B, A, C), (B, C, A),}$$
$$\text{(C, A, B), (C, B, A)}$$

- např. kolika způsoby může 5 osob vytvořit jednu řadu?

Kombinace

$= k$ prvková podmnožina n prvkové množiny, každý prvek může být vybrán jen jednou

- Vybíráme k prvků z n prvkové množiny, přičemž nezáleží na pořadí výběru.
- Počet všech kombinací:

$$C(n,k) = \left(\begin{array}{c} n \\ k \end{array} \right) = \frac{n!}{k! \cdot (n-k)!}$$

- $n = 3$, $k = 2$:

$$\{\text{A, B, C}\} \longrightarrow \{\text{A, B}\}, \{\text{A, C}\}, \{\text{B, C}\}$$

- např. počet všech možných trojic studentů vybraných z dvacetičlenné skupiny

Variace s opakováním

= uspořádaná k-tice sestavená z n prvků, jednotlivé prvky se mohou opakovat

- Postupně vybíráme k prvků z n prvkové množiny, vybraný prvek vždy vracíme zpět do množiny. Záleží na pořadí výběru.
- Počet všech variací s opakováním:

$$V^*(n,k) = n^k$$

- $n = 3$, $k = 2$:

$$\{\mathrm{A, B, C}\} \longrightarrow \mathrm{(A, A), (A, B), (A, C), (B, A), (B, B), (B, C),} \\ \mathrm{(C, A), (C, B), (C, C)}$$

- např. počet všech možných devítimístných telefonních čísel sestavených z cifer 0-9

Permutace s opakováním

- Sestavíme uspořádanou n-tici z prvků $(\alpha_1, \alpha_2, \ldots, \alpha_k)$ tak, aby prvek α_1 byl zastoupen n_1-krát, prvek α_2 zastoupen α_2-krát, atd.
  (Platí $n_1 + n_2 + \cdots + n_k = n$.)

- Počet všech permutací n prvkové množiny:

$$P^* (n_1, n_2, \ldots, n_k) = \frac{P(n)}{P(n_1) \cdot P(n_2) \cdot \cdots \cdot P(n_k)} = \frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_k!}$$

- $(n_1, n_2) = (1, 2)$, $(\alpha_1, \alpha_2) = (A, B)$:

$$(\text{A, B}) \longrightarrow (\text{A, B, B}), (\text{B, A, B}), (\text{B, B, A})$$

- např. počet anagramů slova ANAGRAM

Kombinace s opakováním

= libovolná k prvková skupina sestavená z prvků n prvkové množiny

- Vybíráme k prvků, každý prvek vybereme z n prvkové množiny. Na pořadí výběru nezáleží.

- Počet všech kombinací s opakováním:

$$C^*(n,k) = C(n+k-1,k) = \binom{n+k-1}{k} = \frac{(n+k-1)!}{(n-1)! \cdot k!}$$

- $n = 2$, $k = 3$:

$$\{\mathrm{A},\ \mathrm{B}\} \longrightarrow [\mathrm{A},\ \mathrm{A},\ \mathrm{A}],\ [\mathrm{A},\ \mathrm{A},\ \mathrm{B}],\ [\mathrm{A},\ \mathrm{B},\ \mathrm{B}],\ [\mathrm{B},\ \mathrm{B},\ \mathrm{B}]$$

- např. v obchodě prodávají 4 druhy piva, kolika způsoby můžeme nakoupit deset lahví?

Uspořádané výběry	Bez opakování	Variace	$V\left(n,k\right)=\frac{n!}{(n-k)!}$
		Permutace	$P\left(n\right)=n!$
	S opakováním	Variace s opakováním	$V^{*}\left(n,k\right)=n^{k}$
		Permutace s opakováním	$P^{*}\left(n_{1},n_{2},\ldots,n_{k}\right)=\frac{n!}{n_{1}!\cdot n_{2}!\cdot\ldots\cdot n_{k}!}$
Neuspořádané výběry	Bez opakování	Kombinace	$C\left(n,k\right)=\frac{n!}{k!\cdot(n-k)!}$
	S opakováním	Kombinace s opakováním	$C^{*}\left(n,k\right)=\frac{(n+k-1)!}{(n-1)!\cdot k!}$

Příklad 1

Které heslo je bezpečnější?

- **Heslo o délce osm znaků složené pouze z číslic.**
- **Heslo o délce pět znaků složené pouze z písmen anglické abecedy.**

Příklad 1

Které heslo je bezpečnější?

- **Heslo o délce osm znaků složené pouze z číslic.**
- **Heslo o délce pět znaků složené pouze z písmen anglické abecedy.**

Počet hesel, která lze sestavit z osmi číslic:

$$V^*(n = 10, k = 8) = 10^8 = 100\,000\,000$$

Počet hesel, která lze sestavit z pěti písmen, jestliže **nerozlišujeme** malá a velká písmena:

$$V^*(n = 26, k = 5) = 26^5 = 11\,881\,376$$

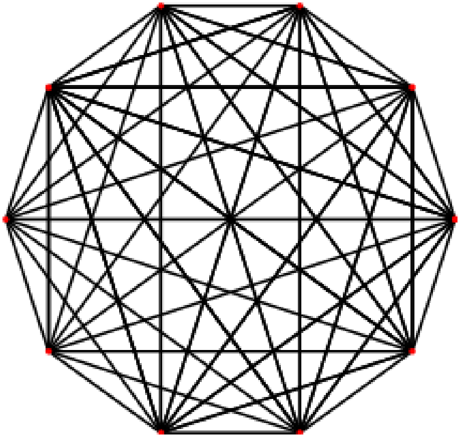
Počet hesel, která lze sestavit z pěti písmen, jestliže **rozlišujeme** malá a velká písmena:

$$V^*(n = 52, k = 5) = 52^5 = 380\,204\,032$$

Pokud bychom nerozlišovali malá a velká písmena, bylo by bezpečnější heslo složené z čísel. Pokud velká a malá písmena rozlišíme, bezpečnější bude heslo složené z písmen.

Příklad 2

Jak dlouho by trvalo vyřešení problému obchodního cestujícího pro $n = 10$ měst hrubou silou, jestliže vyhodnocení délky každé z možných cest trvá $1\ \mu$s?



Příklad 2

Jak dlouho by trvalo vyřešení problému obchodního cestujícího pro $n = 10$ měst hrubou silou, jestliže vyhodnocení délky každé z možných cest trvá $1\ \mu$s?

Graf problému můžeme považovat za úplný. Počet permutací n měst:

$$P(n) = n!$$

První město považujeme za fixní:

$$P(n-1) = (n-1)!$$

Každá cesta je nyní obsažena dvakrát. Počet všech různých cest přes všechna města (tzv. hamiltonovských kružnic):

$$\frac{P(n-1)}{2} = \frac{(n-1)!}{2}$$

$n = 10 \longrightarrow \frac{P(9)}{2} = 181\,440 \longrightarrow 0.181$ sekund

Příklad 2

Jak dlouho by trvalo vyřešení problému obchodního cestujícího pro $n = 10$ měst hrubou silou, jestliže vyhodnocení délky každé z možných cest trvá $1\ \mu$s?

Graf problému můžeme považovat za úplný. Počet permutací n měst:

$$P(n) = n!$$

První město považujeme za fixní:

$$P(n-1) = (n-1)!$$

Každá cesta je nyní obsažena dvakrát. Počet všech různých cest přes všechna města (tzv. hamiltonovských kružnic):

$$\frac{P(n-1)}{2} = \frac{(n-1)!}{2}$$

$n = 10 \longrightarrow \frac{P(9)}{2} = 181\,440 \longrightarrow 0.181$ sekund
$n = 15 \longrightarrow \frac{P(14)}{2} = 43,6 \cdot 10^9 \longrightarrow$ cca 12 hodin
$n = 20 \longrightarrow \frac{P(19)}{2} = 6,1 \cdot 10^{16} \longrightarrow$ cca 1929 let

Příklad 3 (Partition problem)

Problém dvou loupežníků je NP-úplný problém, jak rozdělit kořist mezi 2 loupežníky, aby dostali oba věci ve stejné hodnotě. Tj. lze rozdělit N zadaných čísel do dvou skupin tak, aby součet čísel v obou skupinách byl stejný?

$\longrightarrow$ **Kolik možností by bylo třeba vyzkoušet, pokud bychom úlohu řešili hrubou silou?**

Příklad 3 (Partition problem)

Problém dvou loupežníků je NP-úplný problém, jak rozdělit kořist mezi 2 loupežníky, aby dostali oba věci ve stejné hodnotě. Tj. lze rozdělit N zadaných čísel do dvou skupin tak, aby součet čísel v obou skupinách byl stejný?

$\longrightarrow$ Kolik možností by bylo třeba vyzkoušet, pokud bychom úlohu řešili hrubou silou?

Každou z N věcí přiřadíme jednomu ze dvou loupežníků.

$$V^*(n=2, k=N) = 2^N$$

Dostaneme složitější úlohu, pokud zdvojnásobíme počet věcí, nebo pokud přidáme jednoho loupežníka?

Příklad 3 (Partition problem)

Problém dvou loupežníků je NP-úplný problém, jak rozdělit kořist mezi 2 loupežníky, aby dostali oba věci ve stejné hodnotě. Tj. lze rozdělit N zadaných čísel do dvou skupin tak, aby součet čísel v obou skupinách byl stejný?

$\longrightarrow$ Kolik možností by bylo třeba vyzkoušet, pokud bychom úlohu řešili hrubou silou?

Každou z N věcí přiřadíme jednomu ze dvou loupežníků.

$$V^* (n = 2, k = N) = 2^N$$

Dostaneme složitější úlohu, pokud zdvojnásobíme počet věcí, nebo pokud přidáme jednoho loupežníka?

- zdvojnásobení počtu věcí:

$$V^* (n = 2, k = 2N) = 2^{2N} = 4^N$$

- tři loupežníci místo dvou:

$$V^* (n = 3, k = N) = 3^N$$

Příklad 3 (Partition problem)

Problém K loupežníků je NP-úplný problém, jak rozdělit kořist mezi loupežníky, aby všichni dostali věci ve stejné hodnotě. Tj. lze rozdělit N zadaných čísel do K skupin tak, aby součet čísel ve všech skupinách byl stejný?

Podobnou úlohu řešíme, pokud chceme rozdělit N nezávislých úloh trvajících různou dobu mezi K paralelních procesů. (Nezkoumáme, zda je takto lze rozdělit, ale hledáme optimální rozdělení úloh, aby paralelní procesy dokončily výpočet přibližně ve stejnou dobu.)