Princip inkluze a exkluze

Kombinatorika pro bioinformatiky

21.10.2020

1 Princip inkluze a exkluze

Jednoduchý příklad $|A \cup B| = |A| + |B| - |A \cap B|$

Jaká je velikost sjednocení více množin, známe-li jejich průniky?

Příklad: Kolik z čísel $1, 2, \ldots, 30$ je dělitelných některým číslem z $2, 3, 5$?

Řešení: Čísel dělitelných 2 bude $\lfloor 30/2 \rfloor = 15$, dělitelných 3 $\lfloor 30/3 \rfloor = 10$ a dělitelných 5 $\lfloor 30/5 \rfloor = 6$. Čísla dělitelná dvěmi z $2, 3, 5$ jsme ale počítali dvakrát, je tedy třeba je odečíst. To se týká čísel dělitelných $2 \cdot 3 = 6$, těch je $\lfloor 30/6 \rfloor = 5$, dále je třeba odečíst čísla dělitelná $2 \cdot 5 = 10$, kterých je $\lfloor 30/10 \rfloor = 3$, a $3 \cdot 5 = 15$, kterých je $\lfloor 30/15 \rfloor = 2$. Nakonec čísla dělitelná všemi třemi byla třikrát přičtena a třikrát odečtena. Přičteme je tedy ještě jednou. Ve skutečnosti jde pouze o jedno číslo, neboť $\lfloor 30/(2 \cdot 3 \cdot 5) \rfloor = 1$. Celkem nám tedy vyjde, že čísly $2, 3, 5$ je dělitelných $15 + 10 + 6 - 5 - 3 - 2 + 1 = 22$ z čísel $1, 2, \ldots, 30$.

$\square$

Věta 1 (PIE) *Nechť $A_1, A_2, \ldots, A_n$ jsou konečné množiny. Potom*

$$\begin{aligned}
\left|\bigcup_{i=1}^{n} A_i\right| &= |A_1| + |A_2| + \ldots |A_n| - |A_1 \cap A_2| - \ldots - |A_{n-1} \cap A_n| + \\
&\quad |A_1 \cap A_2 \cap A_3| + \ldots + (-1)^{n+1} |A_1 \cap A_2 \cap \ldots \cap A_n| = \\
&= \sum_{k=1}^{n} (-1)^{k+1} \sum_{I \in \binom{[n]}{k}} \left|\bigcap_{i \in I} A_i\right| = \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \left|\bigcap_{i \in I} A_i\right| .
\end{aligned}$$

Důkaz: Tvrzení dokážeme matematickou indukcí. Snadno nahlédneme platnost tvrzení pro $n = 1, 2, 3$. Budeme tedy dokazovat indukční krok.

$$\left|\bigcup_{i=1}^{n+1} A_i\right| = \left|\left(\bigcup_{i=1}^{n} A_i\right) \cup A_{n+1}\right| = \left|\bigcup_{i=1}^{n} A_i\right| + |A_{n+1}| - \left|\left(\bigcup_{i=1}^{n} A_i\right) \cap A_{n+1}\right|$$

První člen na pravé straně rozvedeme z indukčního předpokladu. Poslední člen nejprve upravíme $(\bigcup_{i=1}^n A_i) \cap A_{n+1} = \bigcup_{i=1}^n (A_i \cap A_{n+1})$ a opět použijeme indukční předpoklad. Tím dostáváme

$$\begin{aligned}\left|\bigcup_{i=1}^{n}(A_i \cap A_{n+1})\right| &= \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \left| \left(\bigcap_{i \in I} A_i\right) \cap A_{n+1} \right| = \\ &= \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \left| \bigcap_{i \in I \cap \{n+1\}} A_i \right| .\end{aligned}$$

Můžeme dosadit

$$\begin{aligned}\left|\bigcup_{i=1}^{n+1} A_i\right| &= \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \left|\bigcap_{i \in I} A_i\right| + |A_{n+1}| - \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \left| \bigcap_{i \in I \cap \{n+1\}} A_i \right| = \\ &= \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \left|\bigcap_{i \in I} A_i\right| + \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|} \left| \bigcap_{i \in I \cap \{n+1\}} A_i \right| + |A_{n+1}| = \\ &= \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \left|\bigcap_{i \in I} A_i\right| + \sum_{I \subseteq [n]} (-1)^{|I \cap \{n+1\}|+1} \left| \bigcap_{i \in I \cap \{n+1\}} A_i \right| = \\ &= \sum_{\emptyset \neq I \subseteq [n+1]} (-1)^{|I|+1} \left|\bigcap_{i \in I} A_i\right| .\end{aligned}$$

$\square$

Důkaz jiným způsobem: Vezměme si nějaké $x \in \cup_{i=1}^n A_i$ a položme $J = \{i \in [n], x \in A_i\}$ a označme $j = |J|$. Je jasné, že na levé straně je x započteno jednou. Na pravé straně se objeví právě v těch průnicích $\cap_{i \in I} A_i$ kde $I \subseteq J$. Postupujme dle mohutnosti I: existuje právě j 1-prvkových indexových podmnožin obsažených v J, $\binom{j}{2}$ 2-prvkových atd.. Obecně se x objeví v $\binom{j}{i}$ (pro $i \leq j$) průnicích na pravé straně. Jelikož průniky se sudými indexovými množinami se započítávají se znaménkem minus, je na pravé straně celkem $\binom{j}{1} - \binom{j}{2} + \ldots + (-1)^{j+1}\binom{j}{j} = 1 - \left[\sum_{i=0}^{j} \binom{j}{i}(-1)^i\right] = 1 - (1-1)^j = 1$ krát.

$\square$

2 Aplikace – problém šatnářky

Do klubu přicházejí večer pánové a každý z nich si v šatně odloží klobouk. Šatnářka je však roztržitá a nepamatuje si, který klobouk komu patří. Rozdává je tedy pánům při odchodu zcela náhodně. Jaká je pravděpodobnost, že žádný z pánů nedostane zpět svůj klobouk?

Rozdání klobouků je vlastně permutace, ptáme se tedy na to, jaká je pravděpodobnost, že náhodně zvolená permutace nemá pevný bod (tj. i takové, že $\pi(i) = i$). Označme

$$\check{s}(n) = \text{ počet permutací bez pevného bodu na množině } \{1, 2, \ldots, n\}$$

Potom hledaná pravděpodobnost bude $\check{s}(n)/n!$.

Tvrzení 2 *Pro $n \geq 1$ je $\check{s}(n) = n!(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \ldots + (-1)^n \frac{1}{n!})$.*

Důkaz: Označme $F_i = \{\pi \in S_n, \pi(i) = i\}$, potom $\check{s}(n) = |S_n \setminus \cup_{i=1}^n F_i| = n! - |\cup_{i=1}^n F_i|$. Pro výpočet $|\cup_{i=1}^n F_i|$ použijeme PIE.

Nejprve si určíme $|\cap_{i\in I} F_i|$. Snadno nahlédneme, že prvky $\cap_{i\in I} F_i$ jsou právě ty permutace$\pi \in S_n$, pro které $\pi(i) = i$ pro $\forall i \in I$. Proto $|\cap_{i\in I} F_i| = (n - |I|)!$. Lze tedy počítat

$$\begin{aligned} |\cup_{i=1}^n F_i| &= \sum_{k=1}^{n} (-1)^{k+1} \sum_{I \in \binom{[n]}{k}} (n - |I|)! = \\ &= \sum_{k=1}^{n} (-1)^{k+1} \binom{n}{k} (n-k)! = \sum_{k=1}^{n} (-1)^{k+1} \frac{n!}{k!}\,. \end{aligned}$$

Tedy

$$\check{s}(n) = n! - \sum_{k=1}^{n} (-1)^{k+1} \frac{n!}{k!} = n! \left(1 + \sum_{k=1}^{n} (-1)^k \frac{1}{k!}\right) = n! \sum_{k=0}^{n} (-1)^k \frac{1}{k!}\,.$$

□

Důsledek 3 *Pravděpodobnost, že žádný z pánů nedostane zpět svůj klobouk se blíží $1/e \doteq 0.368787\ldots$.*

Důkaz: $p = \frac{\check{S}(n)}{n!} = \sum_{k=0}^{n} \frac{(-1)^k}{k!} \to e^{-1}$, neboť $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$.

□

3 Aplikace jiná – počet surjekcí

Věta 4 *Nechť X, Y jsou konečné množiny, $|X| = n, |Y| = m$ a $n \geq m > 0$. Potom počet zobrazení f množiny X na množinu Y je $\sum_{k=0}^{m}(-1)^k\binom{m}{k}(m-k)^n$.*

Důkaz: Počet všech funkcí z X do Y je m^n.

Dále si označíme prvky množiny Y jako $y_1, y_2, \ldots, y_m$ a definujeme množiny $F_i = \{f : X \mapsto Y | \nexists x \in X : f(x) = y_i\}$ pro $i = 1, 2, \ldots, m$.

Pak množina $\bigcup_{i=1}^{m} F_i$ je právě množina funkcí které nejsou surjektivní a její mohutnost stačí odečíst od počtu všech funkcí. Mohutnost této množiny spočteme s použitím principu inkluze a exkluze

$$
\begin{aligned}
|\bigcup_{i=1}^{m} F_i| &= \sum_{k=1}^{m}(-1)^{k+1} \sum_{I \in \binom{\{1,\ldots,m\}}{k}} |\bigcap_{i \in I} F_i| \\
&= \sum_{k=1}^{m}(-1)^{k+1}\binom{m}{k}(m-k)^n
\end{aligned}
$$

Počet surjekcí je pak

$$m^n - \sum_{k=1}^{m}(-1)^{k+1}\binom{m}{k}(m-k)^n$$

což lze ještě upravit

$$m^n + \sum_{k=1}^{m}(-1)^{k}\binom{m}{k}(m-k)^n = \sum_{k=0}^{m}(-1)^{k}\binom{m}{k}(m-k)^n$$

□

Výslednou sumu stačí počítat jen do $k = m-1$, pro $k = m$ je sčítanec nulový. Na rozdíl od počítání prostých funkcí je zde třeba mít podmínku na mohutnost množin, neboť pro $|Y| > |X|$ neexistuje surjektivní funkce, ale výraz může být nenulový.

4 Aplikace poslední – Eulerova funkce

Definice 5 $\varphi(n) =$ *počet čísel z* $1, 2, \ldots, n$ *nesoudělných s* n

Tvrzení 6 *Ať* $n = p_1^{e_1} \cdot p_2^{e_2} \cdot \ldots \cdot p_k^{e_k}$ *je prvočíselný rozklad (tj.* $p_1, p_2, \ldots, p_k$ *různá prvočísla,* $e_i \geq 1$*). Potom* $\varphi(n) = n(1 - \frac{1}{p_1})(1 - \frac{1}{p_2}) \ldots (1 - \frac{1}{p_k})$*.*

Důkaz: Pro $i = 1, 2, \ldots, k$ definujme $A_i = \{a \in \{1, 2, \ldots, n\}. p_i \backslash a\}$. Potom zřejmě $\bigcup_{i=1}^{k} A_i$ je množina všech čísel z $1, 2, \ldots, n$ soudělných s n. Tedy $\varphi(n) = n - |\cup_{i=1}^{k} A_i|$.

Počítejme

$$\begin{aligned} |A_i| &= \frac{n}{p_i} \\ |A_i \cap A_j| &= \frac{n}{p_i p_j}, \; i \neq j \\ |\cap_{i \in I} A_i| &= \frac{n}{\prod_{i \in I} p_i}. \end{aligned}$$

Takže dle PIE

$$\begin{aligned} |\bigcup_{i=1}^{k} A_i| &= \sum_{I \in 2^{[n]} \backslash \emptyset} (-1)^{|I|+1} \frac{n}{\prod_{i \in I} p_i} \\ &= n \left(\frac{1}{p_1} + \ldots + \frac{1}{p_k} - \frac{1}{p_1 p_2} - \ldots + \frac{1}{p_1 p_2 p_3} + \ldots + (-1)^{k+1} \frac{1}{p_1 \ldots p_k} \right). \end{aligned}$$

a

$$\begin{aligned} \varphi(n) &= n \left(1 - \frac{1}{p_1} - \ldots - \frac{1}{p_k} + \frac{1}{p_1 p_2} + \ldots + (-1)^k \frac{1}{p_1 \ldots p_k} \right) \\ &= n \left(1 - \frac{1}{p_1} \right) \left(1 - \frac{1}{p_2} \right) \ldots \left(1 - \frac{1}{p_k} \right). \end{aligned}$$

□

Poznámka 7

$$\varphi(n) = p_1^{e_1} \ldots p_k^{e_k} (1 - \frac{1}{p_1}) \ldots (1 - \frac{1}{p_k}) = \prod_{i=1}^{k} p_i^{e_i} (1 - \frac{1}{p_i}) = \prod_{i=1}^{k} (p_i^{e_i} - p_i^{e_i - 1})$$

Alternativní důkaz pak může jít následujícím způsobem.

1. $k = 1$: $\varphi(p^e) = p^e - p^{e-1}$
2. $k \geq 2$: $\varphi(r \cdot s) = \varphi(r) \cdot \varphi(s)$ *pro* r, s *nesoudělné*