Výfučtení: Kombinatorika a pravděpodobnost

Milí řešitelé, vítejte u již pátého letošního Výfučtení. V této sérii se trochu odkloníme od klasických fyzikálních témat, nicméně nemusíte být zklamaní. Ukážeme si základy kombinatoriky a pravděpodobnosti, které mají nejen ve fyzice, ale i v logice a především matematice velice široké a praktické využití. Jistě se vám už někdy stalo, že jste s nevěřícím výrazem sledovali své okolí a kladli si otázku: „Jaká je pravděpodobnost, že…?“. V tomto Výfučtení se dozvíte, jak najít správnou odpověď.

Základy kombinatoriky

Nejdříve začneme rychlým přehledem základních pojmů. Samotná kombinatorika je matematická disciplína z oblasti diskrétní matematiky (což znamená, že pracuje pouze s přirozenými čísly). Kombinatorika nám říká, jak a kolika způsoby můžeme za stanovených podmínek z daného počtu různých prvků (množiny), vybrat určité (ne nutně různé) počty prvků s nějakou společnou vlastností nebo třeba v konkrétním pořadí. Obecně jsou tyto metody výběru celkem tři.

Permutace

Permutace využijeme ve chvíli, kdy vybíráme vždy všechny prvky. Jednotlivé varianty výběru se tedy budou lišit pouze pořadím (typické využití je u otázky kolik n-ciferných čísel dokážeme sestavit, pokud máme k dispozici n číslic). Takové varianty nazýváme *permutace* a jejich celkový počet spočítáme tak, že si pro n různých prvků představíme přesně n pozic, na které je chceme umístit, a položíme si otázku: „Kolik mám možností, jak obsadit první místo?“ Odpověď je n. Nyní se ptáme na druhé místo, ale tam již můžeme dát jen $n-1$ prvků, protože jeden už obsadil první pozici. A takhle pokračujeme až do konce. Není těžké si rozmyslet, že se jednotlivé počty možností mezi sebou násobí, protože pro každou jednu volbu na dané pozici se počty výběru větví do různých možností na pozicích dalších. Tím získáme tzv. *faktoriál*, který značíme $n!$ a jde o zkrácený zápis výrazu

$$P(n) = n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 1\,.$$

Výpočet počtu permutací nám však může zkomplikovat opakování. Může nastat situace, kdy vybíráme z prvků, které nejsou všechny různé – některé se opakují. Náš výpočet se bude lišit tím, že počet permutací všech prvků vydělíme počtem permutací každého z opakujících se prvků (tedy pokud se nám jeden prvek z výběru opakuje 3krát a druhý prvek 2krát, budeme dělit $3! \cdot 2!$). To si opět můžeme představit tak, že mezi sebou prohazujeme stejné prvky. Jenže tím, že se prvky od sebe neliší, vlastně na výsledném pořadí nic neměníme. Tímto dělením tedy eliminujeme všechna duplicitní uspořádání.

Variace

Variace jsou podobné permutacím, ovšem s tím rozdílem, že vybíráme vždy menší počet prvků, než kolik jich máme celkem k dispozici. Rovněž však záleží na pořadí prvků v daném výběru.

Počet variací n různých prvků je tedy opět součinem, ale nikoliv čísel od 1 do n, jako je tomu u permutace, ale pouze od $n-k+1$ do n, kde k je počet prvků, které vybíráme. Máme totiž pouze k pozic, u nichž se postupně ptáme, kolik prvků do nich můžeme dosadit. Počet variací k-té třídy z n prvků (vybíráme k prvků z n a záleží na jejich pořadí) zapíšeme jako:

$$V(k,n) = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}\,.$$

I u variace se můžeme setkat s opakujícími se prvky. V tomto případě opět vydělíme součinem faktoriálů počtu, kolikrát je ve výčtu opakující se prvek obsažen (zcela stejně jako u permutací).

Kombinace

Poslední metodou výběru jsou kombinace, kde z n prvků vybíráme k prvků, na jejichž pořadí nám vůbec nezáleží. Jde tedy vlastně o variace bez permutací každé vybrané k-tice. Toto reprezentují tzv. *kombinační čísla*.

Kombinační čísla využijeme nejen pro výpočet kombinací, jejich využití přesahuje i do výpočtů pravděpodobnosti, které si ukážeme dále. Kombinační číslo čteme „n nad k“ a vypočítáme jej jako

$$K(k,n) = \binom{n}{k} = \frac{n!}{k!(n-k)!}\,.$$

Je jasné, že nezáleží na tom, zda se prvky, z nichž vybíráme, opakují, či jsou různé. Jejich pořadí totiž nebereme v potaz.

Kombinace s opakováním řešíme v případě, kdy vybíráme z n různých prvků nějakých k prvků a ve výběru se mohou jednotlivé prvky libovolně opakovat (k tedy může být větší než n). Představme si, že máme v obchodě na výběr z $n = 4$ druhů ovoce (například banány, jablka, meruňky a ananasy) a chceme nakoupit celkem $k = 8$ kusů (je nám jedno, z kolika druhů ovoce se náš nákup bude sestávat, může to být z jednoho, ale i ze všech čtyř). Tuto úlohu budeme řešit pomocí tzv. oddělovačů – tedy od sebe pomyslně oddělíme jednotlivé druhy ovoce. Např. v situaci $BBB|JJ|M|AA$ vidíme, že abychom od sebe oddělili jednotlivé druhy ovoce, museli jsme použít 3 oddělovače. Tyto oddělovače budou měnit svou pozici v závislosti na množství jednotlivého ovoce (jiné složení nákupu by vypadalo např. $|JJJJJJ||AA$). Celkový počet kombinací nákupu proto zjistíme jako $\binom{8+3}{3} = 165$ možností. Tento vzorec si snadno můžeme zobecnit – vybíráme pozice $n-1$ oddělovačů (neboť oddělovačů bude vždy o 1 méně, než je počet variant, ze kterých máme na výběr) z $n+k-1$ možných umístění. To je tedy

$$K'(k,n) = \binom{n+k-1}{n-1} = \binom{n+k-1}{k}\,.$$

Pascalův trojúhelník

Kombinatorické schéma, které shrnuje všechna kombinační čísla a dává nám do nich jakýsi vhled, se nazývá Pascalův trojúhelník. Ten si můžeme představit třeba tak, že se jedná o trojúhelník, který má nahoře 1 a v každém dalším patře má součet čísel nad ním. Po stranách horní 1 si tedy představíme 0 a můžeme začít sčítat další a další patra. To je asi nejintuitivnější způsob konstrukce Pascalova trojúhelníku.

Pro nás je však tento matematický konstrukt zajímavý z úplně jiného důvodu, a tím jsou právě kombinační čísla. Na první pohled se to může zdát trochu neintuitivní, nicméně Pascalův

$$
\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & 1 & & 1 & & & & & \\
 & & & & 1 & & 2 & & 1 & & & & \\
 & & & 1 & & 3 & & 3 & & 1 & & & \\
 & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
 & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1
\end{array}
$$

Obrázek 1: Pascalův trojúhelník

trojúhelník je vlastně přehled všech kombinačních čísel. Jeho jednotlivé členy, které jsme původně konstruovali součtem předcházejících členů, totiž můžeme bez jakékoliv znalosti předchozích pater vypočítat právě z příslušných kombinačních čísel.

Každé patro je totiž kombinační číslo (počet prvků, ze kterých vybíráme, je počet čísel v předchozím patře a poloha v řádku je, kolik čísel z nich chceme vybrat). Zmíněná 1 na samotném vrcholu je totiž kombinační číslo $\binom{0}{0}$. Tedy kolika způsoby můžeme vybrat právě žádný prvek z celkem nula prvků – konkrétně jedním (ano, formálně je tento způsob to, že vlastně nic neuděláme).

Z vlastností Pascalova trojúhelníku můžeme odpozorovat několik esenciálních vlastností kombinačních čísel. První z nich je evidentní ze samotného sčítacího mechanismu vzniku trojúhelníku. Tedy to, že když sečteme dvě kombinační čísla, dostaneme číslo „pod nimi“. V řeči formálního zápisu kombinačních čísel to vyjádříme

$$\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}.$$

Zápis si již můžete sami pomocí faktoriálů rozepsat a ověřit, že rovnost skutečně platí.

Další užitečná vlastnost vychází ze symetrie trojúhelníku podél výšky procházející číslem v nejvyšším řádku. Říká nám, že některé kombinace jsou shodné, konkrétně to platí pro

$$\binom{n}{k} = \binom{n}{n-k}.$$

Tvrzení si můžeme rovněž ověřit úvahou, že pokaždé, když vybíráme k věcí z n možností, je to jako z n vybírat $n-k$ věcí, které vlastně nevybereme. V prvním případě tedy vybíráme samotné prvky množiny a v druhém vybíráme prvky, které v množině „necháme“. Je zřejmé, že tato čísla musí vyjít stejná, protože pro každou jednu kombinaci prvků, které v množině ponecháme, existuje právě jedna kombinace prvků, již si vybereme. Tedy každá jedna varianta musí být obsažena v obou číslech.

Samotné využití Pascalova trojúhelníku v realitě není moc rozšířené. Může fungovat jako tabulka, z níž lze číst kombinační čísla, pokud by se nám nechtělo počítat faktoriály. Nicméně jako demonstrace několika základních principů kombinačních čísel a toho, jak v praxi fungují, slouží opravdu dobře.

Pravděpodobnost

Nyní si ještě uděláme krátký vhled do základů pravděpodobnosti. Pravděpodobnost nějakého obecného jevu spočítáme jako podíl případů, ve kterých k jevu došlo, a všech případů, kdy k němu dojít mohlo. Z toho lze vyvodit důležitý poznatek, že žádná pravděpodobnost není větší než 1.

Nejdůležitější vlastností, které mohou dva jevy z hlediska pravděpodobnosti mít, je nezávislost. Znamená to, že spolu dvě veličiny vůbec nesouvisí a jedna s druhou se nijak neovlivňují. Pokud jsou naše dva jevy nezávislé, jejich pravděpodobnosti jsou naprosto separátní hodnoty, tedy fakt, že nastala jedna z nich, nám vůbec nic neřekne o věci druhé. Nezávislé veličiny mohou být například čísla padlá na kostce při dvou po sobě následujících hodech.

S pravděpodobností nezávislých jevů se tedy poté počítá tak, že pokud chceme zjistit, s jakou pravděpodobností nastanou oba zároveň, vynásobíme spolu dílčí pravděpodobnosti jednotlivých jevů.

Binomické rozdělení pravděpodobnosti

To, k čemu se kombinatorika, kombinační čísla a pravděpodobnost opravdu hojně používají (nejen ve fyzice), je binomické rozdělení pravděpodobnosti.

Slovo „rozdělení“ nám popisuje škálu nebo míru jednotlivých událostí. Lze si to představit tak, že hodněkrát uděláte nějakou věc, která může mít jen omezený počet výsledků (typicky se udává hod kostkou nebo mincí), a postupně si zapisujete, jak každý jeden pokus dopadl. Poté se podíváte, v jaké míře vycházely které výsledky, a – voilà – máte rozdělení.

Binomické rozdělení pravděpodobnosti nám poté udává pravděpodobnost daného výsledku. Slovo *binomické* nám popisuje jakousi dualitu, kterou v sobě toto rozdělení skrývá. Binomické rozdělení se totiž zaměřuje pouze na pokusy, které mohou skončit buď úspěchem, nebo neúspěchem. Pravděpodobnost úspěchu, resp. neúspěchu, přitom pro všechny pokusy zůstává stejná. Jak si tato dvě kritéria zadefinujete už je na vás, nicméně musí být vzájemně jednoznačně rozlišitelná a musí dohromady pokrýt všechny možné výsledky (u hodu kostkou to může být například úspěch sudé číslo a neúspěch liché).

Vzorec pro pravděpodobnost, že v binomickém rozdělení, v němž pravděpodobnost úspěchu pokusu je p, nastane z celkových n pokusů právě k úspěchů, je

$$\binom{n}{k} p^k (1-p)^{n-k} \,.$$

Takto můžeme například vypočítat, jaká je pravděpodobnost, že se střelec s úspěšností $p = 70\,\% = 0{,}7$ trefí z $n = 5$ pokusů právě $k = 4$krát.

První činitel, jímž je kombinační číslo, odpovídá počtu všech možných pořadí, v nichž mohly 4 trefy nastat. Druhý a třetím činitel, tj. p^k s $(1-p)^{n-k}$, pak vyjadřují pravděpodobnost, že nastane přesně k úspěchů a k tomu přesně $n-k$ neúspěchů v jednom daném pořadí. Dáme-li vše dohromady, dostaneme pravděpodobnost jevu (druhý s třetím činitelem) v určité konfiguraci přenásobenou počtem těchto konfigurací, tj. možností, jak k jevu mohlo dojít (první činitel).

Závěr

To by od nás bylo o kombinatorice a pravděpodobnosti vše. V matematice, fyzice, ale především statistice se jedná o důležité nástroje, které nám mohou mnohé prozradit o pravděpodobnosti našeho úspěchu v nějaké deterministické situaci, hazardních hrách, či kolik a jakých možností máme, abychom mohli vybrat tu optimální (tento princip se hojně využívá v tzv. *diskrétní matematice*). Takže příště, až budete koukat na test, který jste úspěšně natipovali, můžete si spočítat, jaké jste vlastně měli štěstí, neboli s jakou pravděpodobností se vám podaří trefit minimálně ten váš počet správných odpovědí.

Monika Drexlerová
`monika.drexlerova@vyfuk.org`

Anežka Čechová
`anezka.cechova@vyfuk.org`

Korespondenční seminář Výfuk je organizován studenty a přáteli MFF UK. Je zastřešen Oddělením propagace a mediální komunikace MFF UK a podporován Katedrou didaktiky fyziky MFF UK, jejími zaměstnanci a Jednotou českých matematiků a fyziků. Realizace projektu byla podpořena Ministerstvem školství, mládeže a tělovýchovy.

Toto dílo je šířeno pod licencí Creative Commons Attribution-Share Alike 3.0 Unported.
Pro zobrazení kopie této licence navštivte `https://creativecommons.org/licenses/by-sa/3.0/`.