19. ročník, úloha I . S ... *pravděpodobnost* *(5 bodů; průměr 3,46; řešilo 50 studentů)*

a) Z 32 karet se náhodně vyberou tři karty. Zjistěte pravděpodobnosti jevů, že mezi vybranými kartami bude právě jedno eso, alespoň jedno eso, ani jedno eso.

b) V nádobě se nachází N stejných částic. Určete pravděpodobnost, že v levé půlce bude o m částic více než v pravé půlce. Nakreslete graf závislosti pro $N = 10^{10}$. Rozsah m volte tak, aby pravděpodobnost na krajích intervalu byla desetinová oproti středu intervalu. Jak závisí šířka křivky (tj. rozdíl $m_2 - m_1$, kde $m_2 > 0$ a $m_1 < 0$ jsou hodnoty m, pro které je pravděpodobnost poloviční oproti maximu) na N?

c) Odhadněte velikost $\ln(n!)$ (bez použití Stirlingova vzorce).

Zadal autor seriálu Matouš Ringel.

První díl seriálu byl jednoduchý a spíše matematický, takže jej řešilo hodně řešitelů, kteří se s úlohami většinou dobře vypořádali. Museli překonávat i další obtíže; při výrobě první série u nás totiž řádil tiskařský šotek, který do seriálu zanesl hned dvě chyby, které vzápětí uvedeme na pravou míru.

a) V této podúloze nějaký záškodník změnil počet karet v balíčku na 36 místo správného počtu 32. Zde úlohu vyřešíme se správným počtem karet 32, kde jsou čtyři esa (samozřejmě uznáváme i řešení s 36 kartami).

Nejjednodušší je určit pravděpodobnost, že nevytáhneme ani jedno eso. Potom všechny vytažené karty musí být z množiny 28 „nees". Počet příznivých případů (výběrů) je $\binom{28}{3}$, celkový počet výběrů je $\binom{32}{3}$, takže pravděpodobnost vychází

$$p_0 = \frac{\binom{28}{3}}{\binom{32}{3}} = \frac{26 \cdot 27 \cdot 28}{30 \cdot 31 \cdot 32} \doteq 0{,}660\,.$$

Má-li být mezi kartami právě jedno eso, musíme jednu kartu vybrat z množiny es a zbylé dvě karty z množiny nees. Eso lze z es vybrat $4 = \binom{4}{1}$ způsoby, dvě karty z nees pak $\binom{28}{2}$ způsoby. Tyto dva výběry lze libovolně kombinovat, takže počet příznivých případů je $4 \cdot \binom{28}{2}$. Počet všech výběrů je stejný jako výše, čili pravděpodobnost, že mezi třemi kartami bude právě jedno eso, je

$$p_1 = \frac{\binom{4}{1}\binom{28}{2}}{\binom{32}{3}} = \frac{4 \cdot 3 \cdot 27 \cdot 28}{30 \cdot 31 \cdot 32} \doteq 0{,}304\,.$$

Jestliže požadujeme, aby ve výběru bylo alespoň jedno eso, můžeme postupovat oklikou. Je totiž zřejmé, že vždy nastává právě jeden z jevů „alespoň jedno eso" a „žádné eso". Pravděpodobnost jevu „alespoň jedno eso" pak bude doplňkem do jedničky pravděpodobnosti jevu „žádné eso", kterou již známe. Proto

$$p_{\geq 1} = 1 - p_0 = 1 - \frac{\binom{28}{3}}{\binom{32}{3}} \doteq 0{,}340\,.$$

b) Tato podúloha byla chybou zasažena na nejcitlivějším místě; z textu seriálu se při krácení omylem vytratil vzorec, pomocí kterého se měla vyřešit tato úloha, tzv. Stirlingův vzorec. Většina z řešitelů však Stirlingův vzorec buď znala, nebo si pomohla bodem c) úlohy, takže chyba neměla katastrofální důsledky.

Chceme určit pravděpodobnost, že z celkového počtu N částic bude $(N + m)/2$ v levé půlce a $(N - m)/2$ v pravé půlce (potom bude vlevo o m více částic než vpravo). Celkový počet možných uspořádání je 2^N (u každé částice můžeme říct, má-li být vpravo, nebo vlevo), počet příznivých případů je $\binom{N}{(N+m)/2}$ (ze všech částic vybíráme ty, které mají být vlevo. Počítaná pravděpodobnost tedy vychází

$$p = \frac{1}{2^N} \cdot \binom{N}{\frac{1}{2}(N+m)} .$$

Dále bylo úkolem namalovat graf této závislosti pro $N = 10^{10}$. Umocníme-li naň dvojku nebo počítáme-li z něj faktoriál, dostaneme obrovské číslo, které se do žádné kalkulačky ani PC nevejde. Navíc není na první pohled zřejmé, jak závisí kombinační číslo na svých argumentech. Budeme se jej proto snažit odhadnout, a to pomocí Stirlingova vzorce, který omylem nebyl uveden v seriálu (lze použít i odhad z bodu c). Stirlingův vzorec dává pro faktoriál čísla odhad

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n ,$$

platný pro dostatečně velká n (s chybou menší než $1\,\%$ se dá používat již od čísla $n = 9$).

Pravděpodobnost si rozepíšeme pomocí faktoriálů a odhadneme Stirlingovým vzorcem

$$\begin{aligned} p &= \frac{1}{2^N} \frac{N!}{\left(\frac{N+m}{2}\right)!\left(\frac{N-m}{2}\right)!} \approx \\ &\approx \frac{1}{2^N} \frac{\sqrt{2\pi N}\left(\frac{N}{\mathrm{e}}\right)^N}{\sqrt{\pi(N+m)}\left(\frac{N+m}{2\mathrm{e}}\right)^{(N+m)/2} \cdot \sqrt{\pi(N-m)}\left(\frac{N-m}{2\mathrm{e}}\right)^{(N-m)/2}} . \end{aligned}$$

Tento výraz se po provedení elementárních operací (např. typu $N^N = \exp(N \ln N)$) dostane do tvaru

$$p \approx \sqrt{\frac{2}{N\pi}} \frac{1}{2^N} \frac{\exp\left(N \ln N - \frac{N+m}{2} \ln\left(\frac{N+m}{2}\right) - \frac{N-m}{2} \ln\left(\frac{N-m}{2}\right)\right)}{\sqrt{1 - \frac{m^2}{N^2}}} ,$$

který nám však stále mnoho neříká. My však očekáváme, že při růstu m bude pravděpodobnost rychle klesat; důležité budou tudíž jen oblasti s $m \ll N$. Můžeme tedy zanedbat odmocninu ve jmenovateli, která je téměř rovna jedné, a rozvinout logaritmy v exponenciále do Taylorovy řady (stačí do druhého řádu, tj. použijeme vztah $\ln(1+x) \approx x + x^2/2$). Tímto způsobem nakonec získáme vzorec

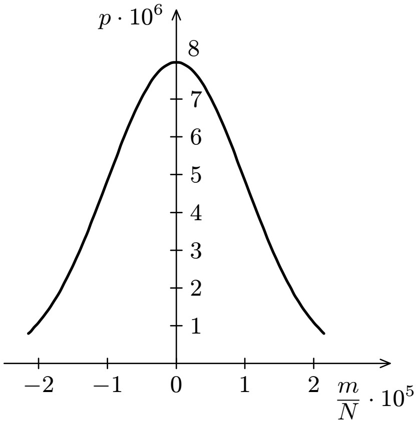
$$p \approx \sqrt{\frac{2}{\pi N}} \exp\left(-\frac{m^2}{2N}\right) ,$$

který nám odůvodněně připomíná známou Gaussovu křivku (viz teorie pravděpodobnosti, Gaussovo rozdělení je totiž limitním případem tzv. binomického rozdělení). Všimněme si, že je normovaná (neb m běží pouze přes sudá čísla).

Nyní máme nakreslit graf v takovém měřítku, aby na krajích byla desetinová pravděpodobnost proti středu, tj.

$$\exp\left(\frac{m^2}{2N}\right) = \frac{p(0)}{p(m)} = 10 ,$$

pročež $m^2/N = 2\ln 10 \approx 4{,}61$ a $m/N \approx 2{,}15 \cdot 10^{-5}$, takže naše přiblížení bylo jistě v pořádku. Vidíme, že šířka křivky Δ (definovaná v zadání úlohy) je dána vzorcem $\Delta = \sqrt{N}\sqrt{2\ln 2}$. Pokud by nás ale zajímala relativní šířka, tj. $\delta = \Delta/N$, dostali bychom $\delta = \sqrt{2\ln 2}/\sqrt{N}$. Pravděpodobnost určité relativní fluktuace počtu částic v obou polovinách klesá s celkovým počtem částic N, což je dosti všeobecná vlastnost a důvod, proč funguje statistická fyzika. V systému obsahujícím řekněme 1000 částic se bude například tlak molekul na stěnu pohybovat řádově v 1% okolí své průměrné (rovnovážné) hodnoty. Pokud však místo tisíce vezmeme částic 10^{23}, tlak bude mít odchylku od průměrné hodnoty typicky $3 \cdot 10^{-12}\,\%$, což je naprosto zanedbatelné oproti jiným vlivům.



Obr. 1

c) Nakonec poněkud ospravedlníme Stirlingův vzorec. Odhadneme totiž $\ln(n!)$. Celý odhad je založen na následujícím povšimnutí

$$\ln(n!) = \ln(1 \cdot 2 \cdot \dots \cdot n) = \ln 1 + \ln 2 + \dots + \ln n = \sum_{k=1}^{n} \ln k\,.$$

Pro velká n se logaritmus mění velice pozvolně, a je proto možné nahradit sumu integrálem. Takto dostáváme odhad

$$\ln(n!) = \sum_{k=1}^{n} \ln k \approx \int_1^n \ln x \,\mathrm{d}x \approx n \ln n - n\,,$$

čili

$$n! \approx \mathrm{e}^{n\ln n - n} = \left(\frac{n}{\mathrm{e}}\right)^n .$$

Matouš Ringel
matous@fykos.mff.cuni.cz