METODICKÉ VARIÁCIE NA TÉMU PRAVDEPODOBNOSŤ I

BELOSLAV RIEČAN, Bratislava

Pravdepodobnosť a matematická štatistika si čoraz viac razia cestu do školského vyučovania. Je tomu tak v cudzine, ako aj u nás doma. Nezostáva nám preto nič iné, ako sa s touto skutočnosťou vyrovnať. Na nasledujúcich stránkach naznačíme jednu z možností, ako by sa mohli na našej základnej a strednej škole vyučovať základy pravdepodobnosti. Pravda, na základnej škole by mala do popredia vystúpiť hra a experiment. Naproti tomu u gymnazistov by malo ísť viac o logickú analýzu príkladov.

Pravdepodobnosť a početnosť

Vyjdeme z klasickej Laplaceovej schémy, v ktorej skúmaný experiment má konečný počet výsledkov (teda daná je konečná množina $\Omega = \{\omega_1, \ldots, \omega_n\}$), pričom o žiadnom z nich nemáme dôvod predpokladať, že by sa vyskytoval častejšie ako iné. Pravdepodobnosť množiny (= udalosti = javu) $A \subset \Omega$ sa definuje pomocou vzťahu

$$P(A) = \frac{|A|}{|\Omega|}$$

kde $|A|$ je počet prvkov množiny A, $|\Omega|$ je počet prvkov množiny Ω.

Príklad 1. Hádžeme pravidelnou kockou. Aká je pravdepodobnosť toho, že padne šestka? Aká je pravdepodobnosť toho, že padne párne číslo?

V tomto prípade $\Omega = \{\omega_1, \ldots, \omega_6\}$, $A = \{\omega_6\}$. Preto pravdepodobnosť toho, že padne šestka, je

$$P(A) = \frac{|A|}{|\Omega|} = \frac{1}{6}$$

V druhej úlohe $B = \{\omega_2, \omega_4, \omega_6\}$, teda

$$P(B) = \frac{|B|}{|\Omega|} = \frac{3}{6} = \frac{1}{2}$$

Uvedené elementárne výpočty je vhodné overiť experimentálne, a to napr. tak, že uskutočníme väčší počet hodov a výsledky zaznamenávame do *tab. 1.*

Tabuľka 1

1 5 3 4 3	1 4 6 5 6	2 6 5 5 2	4 4 6 6 4
2 3 6 6 3	4 2 6 4 3	3 1 1 1 2	4 1 3 1 2
4 3 2 2 5	5 1 6 4 3	6 3 1 6 4	6 1 2 3 6
4 6 4 3 4	4 6 1 2 3	4 3 4 2 3	4 6 1 3 3
6 4 2 4 6	1 6 6 1 1	2 4 1 5 4	6 3 1 3 5

Kvôli šetreniu miestom sme uviedli len prvých 100 čísel. Celkom sme urobili 200 bodov. Medzi nimi sa šestka vyskytla 34-krát. Preto (relatívna) početnosť padnutia šestky bola

$$\frac{34}{200} = 0{,}17$$

Porovnanie s pravdepodobnosťou

$$\frac{1}{6} = 0{,}1\overline{6}$$

vyzerá veľmi nádejne.

Keď si už žiaci tabuľku raz urobili, môžu ju použiť aj pri riešení iných príkladov. Napr. dvojnásobný hod kockou. Nemusia dvakrát hádzať. Tabuľku začnú odpisovať na náhodne vybratom mieste (napr. od 9. číslice v druhom riadku). Tieto čísla budú predstavovať výsledky prvého hodu. Potom opäť vyberieme náhodné miesto v tabuľke (napr. 4. číslicu vo 4. riadku). To budú výsledky druhého hodu. Dostaneme novú tabuľku usporiadaných dvojíc:

[4, 3] [3, 4] [3, 4] [1, 6] [1, 1] [1, 2] [2, 3]
[4, 4] [1, 3] [3, 4] atď.

Príklad 2. Aká je pravdepodobnosť toho, že pri dvojnásobnom hode kockou šestka padne aspoň raz?

Tu máme celkom $36=6.6$ rovnako pravdepodobných výsledkov, teda Ω pozostáva z 36 prvkov — usporiadaných dvojíc čísel 1, ..., 6. (Ešte inak povedané $\Omega = \Omega_0 \times \Omega_0$, kde $\Omega_0 = \{1, 2, ..., 6\}$). Množina A, ktorej pravdepodobnosť hľadáme, pozostáva z tých dvojíc, v ktorých aspoň jeden člen je šestka. Takých dvojíc je celkom 11 (sú to dvojice [6, 6], [1, 6], ..., [5, 6], [6, 1], ..., [6, 5]). Preto

$$P(A)=\frac{11}{36}=0{,}305$$

Konfrontujme túto skutočnosť s experimentom. Napr. my sme urobili 100 dvojíc. Aspoň jedna šestka sa vyskytla v 29 dvojiciach, teda početnosť padnutia aspoň jednej šestky je

$$\frac{29}{100}=0{,}29$$

čo opäť pekne súhlasí s teoretickou pravdepodobnosťou 0,305.

Príklad 3. Aká je pravdepodobnosť toho, že pri hode mincou padne znak?

V tomto prípade $\Omega=\{\omega_1, \omega_2\}$, $A = \{\omega_1\}$. Preto

$$P(A)=\frac{1}{2}=0{,}5$$

Pri experimentovaní nemusíme skutočne hádzať mincou. Môžeme použiť opäť *tab. 1,* a to hneď dvojakým spôsobom.

Prvá metóda spočíva v uvedomení si skutočnosti, že sa v *tab. 1* rovnako často vyskytujú párne a nepárne čísla. Napíšeme preto napr. vždy 0, ak sa vyskytlo párne číslo, 1 v prípade nepárneho čísla. Dostávame

1 1 1 0 1	1 0 0 1 0	0 0 1 1 0	0 0 0 0 0
0 1 0 0 1	0 0 0 0 1	atď.	

Druhú metódu dostaneme, ak uvážime, že napr. 5 a 6 sa v *tab. 1* vyskytujú rovnako často. Preto v *tab. 1* budeme ignorovať ostatné číslice, zapisovať 5 v prípade jej vyskytnutia sa a 0 v prípade vyskytnutia sa šestky. Dostaneme *tab. 2.*

Tabuľka 2

5 0 5 0 0	5 5 0 0 0	0 0 5 5 0	0 0 0 0 0
0 0 0 0 0	0 5 0 5 atď.		

Takýmto spôsobom sme simulovali hod mincou a vyšlo nám, že z 200 pokusov znak padol 98-krát, teda početnosť padnutia znaku bola

$$\frac{98}{200} = 9{,}49$$

Keď sme už tejto tematike venovali toľko miesta, ukážeme ešte ako si žiaci sami môžu zostrojiť tabuľky náhodných čísel 0, 1, ..., 9. Vzhľadom na to, že sme naučení pracovať v dekadickej sústave, je to úloha veľmi užitočná.

Najprv v *tab. 1* budeme škrtať 6 a namiesto 5 budeme písať 0. Tak dostaneme tabuľku, v ktorej sa rovnako často budú vyskytovať 0, 1, ..., *5 (tab. 3).*

Tabuľka 3

1 0 3 4 3	1 4 0 2 0	0 2 4 4 4	2 3 3 4 2
4 3 3 1 1	1 2 4 1 3	atď.	

Teraz skombinujeme *tab. 2* a *3* tak, že zodpovedajúce čísla sčítame. Možné výsledky sú $5+0$, $5+1$, $5+2$, $5+3$, $5+4$, $0+0$, $0+1$, $0+2$, $0+3$, $0+4$, teda všetky celé čísla od 0 do 9. Z nasledujúceho diagramu vidieť, že sa vyskytujú s rovnakou pravdepodobnosťou

0		1		2		3		4	
/	\	/	\	/	\	/	\	/	\
(0,0)	(0,5)	(1,0)	(1,5)	(2,0)	(2,5)	(3,0)	(3,5)	(4,0)	(4,5)
0	5	1	6	2	7	3	8	4	9

Výsledky zaznamenáme do tabuľky náhodných čísel 0, 1, ..., 9:

6 0 8 4 3 6 9 0 2 0 0 2 9 9 4 2 3 3 4 2
4 3 3 1 1 1 7 4 6 atď.

Iná možnosť generovania náhodných čísel od 0 do 9 je hádzanie pravidelným desaťstenom s očíslovanými stenami. Ale k takému telesu sa ťažko dostať.

Pravdaže, najjednoduchšie je vziať knižku s hotovými tabuľkami náhodných čísel (napr. V. Dupač, J. Hájek; Pravděpodobnost ve vědě a technice, Praha 1962). Nazdávame sa však, že vzácnejšími a viac k jadru vedúcimi budú tabuľky, ktoré si žiaci urobia sami.

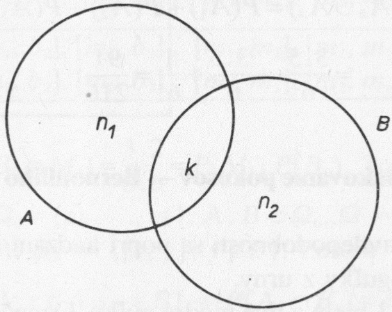
Tabuľky náhodných čísel majú pred experimentom ešte jednu výhodu.

Vyzerajú akademickejšie, starší študenti sa nebudú hanbiť nimi zaoberať. Veď hrať sa s kockou ako malé deti, to im už neprichodí.

Niektoré vlastnosti pravdepodobnosti

Pri logickej analýze mnohých úloh je užitočné poznať niektoré jednoduché pravidlá.

Veta 1. Pre ľubovoľné množiny A, $B \subset \Omega$ platí $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.



Obr. 1

Dôkaz. Ak má množina $A \cap B$ k prvkov, množina $A - B$ n_1 prvkov a množina $B - A$ n_2 prvkov, tak $A \cup B$ má celkom $n_1 + k + n_2$ prvkov, takže

$$|A \cup B| = n_1 + k + n_2 = (n_1 + k) + (n_2 + k) - k = \\ = |A| + |B| - |A \cap B|$$

odkiaľ

$$P(A \cup B) = \frac{|A \cup B|}{n} = \frac{|A|}{n} + \frac{|B|}{n} - \frac{|A \cap B|}{n} = P(A) + P(B) - P(A \cap B)$$

Dôsledok 1. Ak $A \cap B = \emptyset$, tak $P(A \cup B) = P(A) + P(B)$.

Dôsledok 2. $P(A') = 1 - P(A)$.

Príklad 4. Trikrát hádžeme kockou. Aká je pravdepodobnosť toho, že šestka padne a) nanajvýš jeden raz, b) aspoň jeden raz?

Základný priestor Ω pozostáva z usporiadaných trojíc. Nech A_i ($i=0$, 1, 2, 3) je množina tých trojíc, v ktorých je práve i šestiek. V úlohe a) máme pri tomto označení vypočítať $P(A_0 \cup A_1)$.

Pretože $A_0 \cap A_1 = \emptyset$, platí

$$P(A_0 \cup A_1) = P(A_0) + P(A_1) = \frac{5^3}{6^3} + \frac{3 \,.\, 5^3}{6^3} = \frac{200}{216}$$

V úlohe b) máme vypočítať

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) =$$

$$= \frac{3 \,.\, 5^2}{6^3} + \frac{3 \,.\, 5}{6^3} + \frac{1}{6^3} = \frac{91}{216}$$

Nezávislé opakovanie pokusov — Bernoulliho schéma

V učebniciach o pravdepodobnosti sa popri hádzaní kockou a mincou najčastejšie vyťahujú guľky z urny.

Majme teda v urne 3 biele a dve modré guľky. Označme $\Omega_0 = \{b_1, b_2, b_3, m_1, m_2\}$. Budeme z urny dvakrát ťahať. To sa môže uskutočniť dvojako:

1. po prvom ťahu guľku vrátime,
2. vytiahnutú guľku nevraciame.

Pri prvom spôsobe je pravdepodobnosť vytiahnutej bielej guľky v druhom ťahu $\frac{3}{5}$, a to bez ohľadu na to, aká guľka bola vytiahnutá v prvom ťahu.

Iná je situácia pri druhom spôsobe. Ak bola prvá vytiahnutá guľka biela, tak v urne zostali už len 2 biele a 2 modré, teda pravdepodobnosť vytiahnutia bielej v druhom ťahu je $\frac{2}{4}$. Ak bol vytiahnutá guľka modrá, tak pravdepodobnosť vytiahnutia bielej guľky v druhom ťahu sa bude rovnať $\frac{3}{4}$.

Vidíme, že pri prvom spôsobe sú prvý a druhý ťah „nezávislé“, pri druhom „závislé“. Rozoberme si podrobnejšie prvý spôsob, keďže nás zaujíma práve nezávislé opakovanie pokusov. V tomto prípade Ω pozostáva zo všetkých usporiadaných dvojíc prvkov z Ω_0, teda $\Omega = \Omega_0 \times \Omega_0$. (V druhom prípade je množina Ω menšia, chýba v nej uhlopriečka $[b_1, b_1]$, $[b_2, b_2]$, $[b_3, b_3]$, $[m_1, m_1]$, $[m_2, m_2]$.) Nech M_1 spočíva v tom, že prvá vytiahnutá guľka bude modrá, B_2 v tom, že druhá vytiahnutá guľka bude biela

$$
\begin{array}{cccccc}
 & [b_1, b_1], & [b_1, b_2], & [b_1, b_3], & [b_1, m_1], & [b_1, m_2] \\
B_2 & [b_2, b_1], & [b_2, b_2], & [b_2, b_3], & [b_2, m_1], & [b_2, m_2] \\
 & [b_3, b_1], & [b_3, b_2], & [b_3, b_3], & [b_3, m_1], & [b_3, m_2] \\
 & [m_1, b_1], & [m_1, b_2], & [m_1, b_3], & [m_1, m_1], & [m_1, m_2] \\
 & [m_2, b_1], & [m_2, b_2], & [m_2, b_3]. & [m_2, m_1], & [m_2, m_2]
\end{array}
\; M_1
$$

Vidíme, že $P(B_2 \cap M_1) = \dfrac{3.2}{6.6} = P(M_1)\, P(B_2)$. To nie je náhoda.

Veta 2. Nech $\Omega_0 = \{\omega_1, \ldots, \omega_i\}$, $A, B \subset \Omega_0$, $\Omega = \Omega_0 \times \Omega_0$. Nech $A_1 = \{[x, y];\, x \in A\}$, $B_2 = \{[x, y];\, y \in B\}$. Potom

$$P\{[x, y];\, x \in A \wedge y \in B\} = P(A_1 \cap B_2) = P(A_1)\, P(B_2)$$

Dôkaz. Nech A má m prvkov, B má n prvkov. Zrejme Ω má i^2 prvkov, A_1 má $m\,.\,i$ prvkov, B_2 má $n\,.\,i$ prvkov, $A_1 \cap B_2$ má $m\,.\,n$ prvkov. Preto

$$P(A_1) = \frac{m\,.\,i}{i^2} = \frac{m}{i}, \quad P(B_2) = \frac{n\,.\,i}{i^2} = \frac{n}{i}$$

$$P(A_1 \cap B_2) = \frac{m\,.\,n}{i^2} = \frac{m}{i} \cdot \frac{n}{i} = P(A_1)\, P(B_2)$$

Definícia. Pokusom rozumieme ľubovoľnú konečnú neprázdnu množinu Ω_0. Nezávislým (n-násobným) opakovaním pokusu Ω_0 rozumieme (n-násobný) karteziánsky súčin $\Omega = \Omega_0 \times \Omega_0 \times \ldots \times \Omega_0$

Zákonitosť, ktorú sme si všimli vo vete 2, sa zachová aj pri n-násobnom opakovaní.

Príklad 5. Hádžeme trikrát kockou. Aká je pravdepodobnosť toho, že prvýkrát padne párne číslo, druhýkrát nepárne a tretíkrát 1 alebo 2?

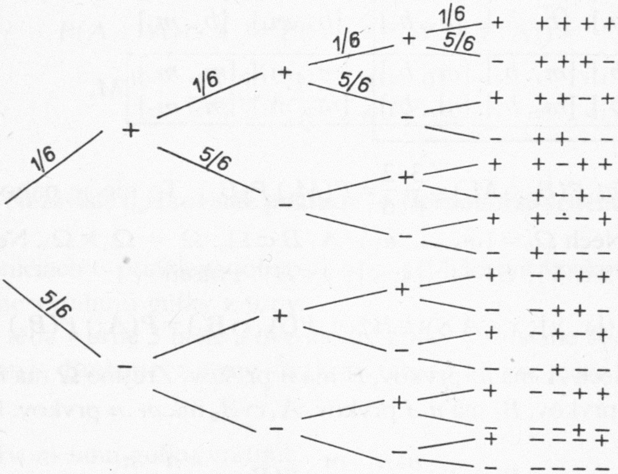
Pri zrejmom označení platí

$$P(A_1 \cap B_2 \cap C_3) = \frac{3.3.2}{6.6.6} = P(A_1)\, P(B_2)\, P(C_3)$$

Zdá sa, že vhodnou pomôckou pri riešení podobných príkladov sú stromy logických možností.

Príklad 6. Aká je pravdepodobnosť toho, že pri štvornásobnom hode kockou šestka padne práve dvakrát?

Zostavíme strom logických možností, pričom budeme písať $+$, ak padne 6, $-$ v prípade, že nepadne. Jednotlivé pravdepodobnosti sa (vďaka nezávislosti opakovania) vynásobia.



Nech B je udalosť, ktorej pravdepodobnosť hľadáme. Zrejme

$$P(B) = P(+\ +\ -\ -) + P(+\ -\ +\ -) + P(+\ -\ -\ +) +$$
$$+ P(-\ +\ +\ -) + P(-\ +\ -\ +) + P(-\ -\ +\ +) =$$
$$= \frac{1}{6}\cdot\frac{1}{6}\cdot\frac{5}{6}\cdot\frac{5}{6} + \frac{1}{6}\cdot\frac{5}{6}\cdot\frac{1}{6}\cdot\frac{5}{6} + \ldots = 6\cdot\left(\frac{1}{6}\right)^2\cdot\left(\frac{5}{6}\right)^2$$

To, že všetky vyskytujúce sa pravdepodobnosti budú rovnaké, to sme tušili. Prečo je ich práve 6? Je to počet takých štvoríc, v ktorých je $+$ práve na dvoch miestach. Teda je to počet kombinácií 2. triedy zo

4 prvkov (t. j. dvojprvkových podmnožín štvorprvkovej množiny). A to je

$$\binom{4}{2}=\frac{4\,.\,3}{2\,.\,1}=6$$

teda

$$P(B)=\binom{4}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^{4-2}$$

Zovšeobecnenie sa nám priamo núka. Pre pravdepodobnosť $P(B)$ toho, že z n nezávislých pokusov nastane udalosť A (ktorej pravdepodobnosť je p) práve k-krát, platí

$$P(B)=\binom{n}{k}p^k(1-p)^{n-k}$$

Pre náročnejšieho čitateľa to aj dokážeme.

Veta 3. Nech $A\subset\Omega_0$, $p=\dfrac{|A|}{|\Omega_0|}$, $\Omega = \Omega_0 \times \Omega_0 \times \ldots \times \Omega_0$ (n-krát). Nech $B\subset\Omega$ je množina tých n-tíc, v ktorých práve k súradníc patrí do A. Potom

$$P(B)=\binom{n}{k}p^k(1-p)^{n-k}$$

Dôkaz. Najprv poriadne dokážeme to, o čom sme čitateľa presviedčali v príkladoch 5 a 6. Nech μ je nejaká k-prvková podmnožina množiny $\{1, 2, \ldots, n\}$. Ak označíme

$$B_\mu=\{[x_1,\ldots,x_n];\ x_i\in A \text{ pre } i\in\mu,\ x_i\in\Omega_0-A \text{ pre } i\notin\mu\}$$

tak

$$P(B_\mu)=\frac{|B_\mu|}{|\Omega|}=\frac{|B_\mu|}{|\Omega_0|^n}=p^k(1-p)^{n-k}$$

Ak totiž Ω_0 pozostáva z m prvkov a množina A z j prvkov (teda $p=\dfrac{j}{m}$), tak počet n-tíc patriacich do B_μ je $j^k(m-j)^{n-k}$, počet prvkov množiny Ω je m^n. Preto

$$P(B_\mu)=\frac{j^k(m-j)^{n-k}}{m^n}=\left(\frac{j}{m}\right)^k\left(1-\frac{j}{m}\right)^{n-k}=p^k(\mathrm{I}-p)^{n-k}$$

Nech C je množina všetkých k-prvkových podmnožín množiny $\{1, \ldots, n\}$; počet jej prvkov je $\binom{n}{k}$. Pretože

$$B=\bigcup_{\mu\in C} B_\mu$$

a B_μ sú navzájom disjunktné, platí podľa I. časti dôkazu

$$P(B)=\sum_{\mu\in C} P(B_\mu)=\sum_{\mu\in C} p^k(1-p)^{n-k}=\binom{n}{k} p^k(1-p)^{n-k}$$

Početnosť a pravdepodobnosť

Na začiatku sme radili robiť experimenty, výsledky zaznamenávať a príslušné početnosti porovnávať s teoretickými pravdepodobnosťami. Pritom sme dosiahli často výbornú zhodu. Prečo?

Aby sme mohli dať fundovanejšiu odpoveď na túto otázku, musíme sa zamyslieť nad tým, čo to vlastne (relatívne) početnosť je.

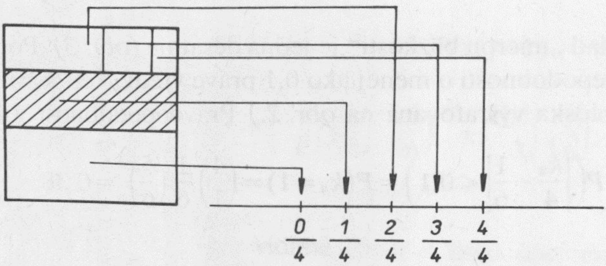
Hádžeme, povedzme, 4-krát kockou a sledujme relatívnu početnosť padnutia šestky. V prípade zaznamenania výsledkov [6, 6, 1, 6] je tá početnosť $\frac{3}{4}$, v prípade štvorice [1, 1, 6, 5] je $\frac{1}{4}$. Relatívna početnosť (v prípade štvornásobného hodu kockou) je teda funkcia $\frac{k}{4}$, definovaná na množine všetkých usporiadaných štvoríc prvkov množiny $\Omega_0=\{1, 2, \ldots, 6\}$, teda na množine $\Omega_0 \times \Omega_0 \times \Omega_0 \times \Omega_0$. Pritom $k(x_1, x_2, x_3, x_4)$ je počet šestiek vo štvorici $[x_1, x_2, x_3, x_4]$. Pravda, funkcia k závisí od počtu n uskutočnených pokusov (v našom prípade to bolo $n=4$). Preto budeme relatívnu početnosť označovať znakmi $\frac{k_4}{4}$, $\frac{k_5}{5}$, $\frac{k_6}{6}$ a pod.

Relatívna početnosť (množiny A) je teda reálna funkcia $\frac{k_n}{n}$ definovaná na množine $\Omega_0 \times \Omega_0 \times \ldots \times \Omega_0$, a to tak, že

$$k_n([x_1, \ldots, x_n]) = k$$

ak práve k súradníc x_i patrí do A.

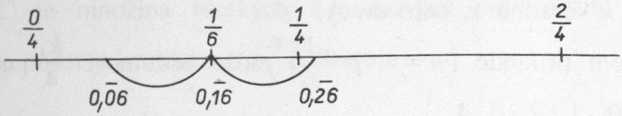
V našom príklade $\left(n=4,\ p=\frac{1}{6}\right)$ môže početnosť $\frac{k_4}{4}$ nadobúdať hodnoty $\frac{0}{4}, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, \frac{4}{4}$, pravda, nie rovnako často *(obr. 2)*



Obr. 2

absolútna početnosť	relatívna početnosť	počet štvoríc	pravdepodobnosť
0	0	5^4	$\left(\frac{5}{6}\right)^4 = 0{,}4808$
1	$\frac{1}{4}$	4.5^3	$\binom{4}{3}\left(\frac{5}{6}\right)^3 \frac{1}{6} = 0{,}3873$
2	$\frac{2}{4}$	6.5^2	$\binom{4}{2}\left(\frac{5}{5}\right)^2 \left(\frac{1}{6}\right)^2 = 0{,}1157$
3	$\frac{3}{4}$	4.5	$\binom{4}{1}\frac{5}{6}\left(\frac{1}{6}\right)^3 = 0{,}0154$
4	1	1	$\left(\frac{1}{6}\right)^4 = 0{,}0008$

Aká je pravdepodobnosť toho, že sa početnosť $\frac{k_4}{4}$ nebude líšiť veľmi od $p=\frac{1}{6}$? Závisí od toho, čo budeme rozumieť pod „veľmi" resp. „málo".

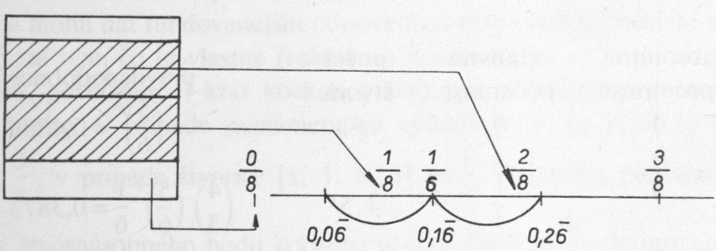


Obr. 3

Nech napríklad „mierou blízkosti" je jedna desatina *(obr. 3)* Početnosť sa líši od pravdepodobnosti o menej ako 0,1 práve vtedy, keď $k_4=1$. (Tomu zodpovedá plôška vyšrafovaná na *obr. 2.)* Pravdepodobnosť toho je

$$P\left(\left|\frac{k_4}{4}-\frac{1}{6}\right|<0{,}1\right)=P(k_4=1)=\binom{4}{1}\frac{1}{6}\left(\frac{5}{6}\right)^3=0{,}38$$

To nie je veľa, ale zodpovedá to malému počtu uskutočnených pokusov.



Obr. 4

Urobme teraz 8 pokusov. Aká je pravdepodobnosť toho, že početnosť

$\frac{k_4}{4}$ sa bude od $\frac{1}{6}$ líšiť o menej ako 0,1?

$$P\left(\left|\frac{k_8}{8}-\frac{1}{6}\right|<0,1\right)=P(k_8=1)+P(k_8=2)=$$

$$=\binom{8}{1}\frac{1}{6}\left(\frac{5}{6}\right)^7+\binom{8}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^6=0,63$$

Pre $n=12$ sú už tri hodnoty, pre ktoré $\left|\frac{k_{12}}{12}-\frac{1}{6}\right|<0,1$. Sú to $k_{12}=1, 2, 3$. Preto

$$P\left(\left|\frac{k_{12}}{12}-\frac{1}{6}\right|<0,1\right)=P(k_{12}=1)+P(k_{12}=2)+P(k_{12}=3)=$$

$$=\binom{12}{1}\frac{1}{6}\left(\frac{5}{6}\right)^{11}+\binom{12}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^{10}+\binom{12}{3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^9=$$

$$=0,77$$

Už na týchto príkladoch vidíme (hoci n bolo dosť malé), že spejeme pomaly, ale iste k „istote", že odchýlka medzi početnosťou a pravdepodobnosťou nepresiahne prípustnú toleranciu ($\varepsilon=0,1$).

Na rozdiel od iných zákonov matematiky táto „istota" nie je stopercentná. Môže sa stať aj to, že pri nejakom opakovaní sa tolerancia poruší. Ale to je málo pravdepodobné. (Napr. pri dvanásťnásobnom opakovaní sa to môže stať tak v 23 prípadoch zo 100.)

Zhoda medzi pravdepodobnosťou a početnosťou nie je náhodná. Vyplýva napr. z nerovnosti

$$P\left(\left|\frac{k_n}{n}-p\right|<\varepsilon\right)\geqq 1-\frac{1}{n\varepsilon^2} \qquad (*)$$

ktorá platí pre každé $n\in N$ a každé $\varepsilon>0$ a ktorú dokážeme (opäť len pre veľmi trpezlivých čitateľov) v dodatku.

Posledne uvedená nerovnosť nám umožňuje odhadovať neznámu pravdepodobnosť. Predstavme si napr. obrovské množstvo výrobkov, z ktorých je určité percento chybných. Toto percento nepoznáme, to by sme museli preskúšať každý výrobok. Namiesto toho preskúšame len určité množstvo z nich, povedzme n. Z nich bude chybných k_n. Na posúdenie percenta nepodarkovosti budeme mať k dispozícii podiel k_n/n namiesto presného p.

Lenže (*) nám dáva aj odhad čísla p

$$P\left(\frac{k_n}{n}-\varepsilon<p<\frac{k_n}{n}+\varepsilon\right)\geqq 1-\frac{1}{n\varepsilon^2}$$

Číslo n poznáme, dokonca sme ho volili, to je počet opakovaných pokusov, k_n sme namerali, ε je nami zvolená tolerancia. Dostávame interval $\left(\frac{k_n}{n}-\varepsilon,\ \frac{k_n}{n}+\varepsilon\right)$, v ktorom sa p nachádza s určitou pravdepodobnosťou. Tá je dosť veľká, blíži sa k istote, ak je dosť veľký súčin $n\varepsilon^2$. To môžeme dosiahnuť dvojako: buď zväčšovaním n, alebo (do určitej miery) zväčšovaním ε, teda zväčšením tolerancie.

Treba priznať, že je to odhad dosť mizerný. Napr. pri $\varepsilon=0{,}1$ by sme museli urobiť $n=10\,000$ pokusov, aby sme dospeli k 99 % istote. Vtedy je totiž

$$n\varepsilon^2=10\,000\,.(0{,}1)^2=100$$

$$1-\frac{1}{n\varepsilon^2}=1-0{,}01=0{,}99$$

Pravda, chyba je v našej nešikovnosti, nie v podstate veci.

Zručný štatistik by pri takom veľkom počte pokusov vedel s 99 % istotou povedať, že neznáma pravdepodobnosť sa od danej početnosti bude líšiť o menej ako $\varepsilon=0{,}005$. Na toleranciu $\varepsilon=0{,}1$ by stačilo už 100 pokusov pri 99% istote. (O týchto odhadoch sa možno dočítať napr. v ods. 11.12 knihy J. Hátleho a J. Likeša: Základy počtu pravdepodobnosti a matematické statistiky, Praha 1972.)

Dodatok — dôkaz zákona veľkých čísel

Lema. Pre ľubovoľné $n\in N$, $p\in R$ platí

$$\sum_{k=0}^{n}\left(\frac{k}{n}-p\right)^2\binom{n}{k}p^k(1-p)^{n-k}=\frac{p(1-p)}{n}$$

Dôkaz. Dvojčlen $\left(\frac{k}{n}-p\right)^2$ umocníme, a potom zvlášť vypočítame súčty

$$\sum_{k=0}^{n} \frac{k^2}{n^2}\binom{n}{k} p^k (1-p)^{n-k} = p^2 + \frac{p(1-p)}{n} \tag{1}$$

$$\sum_{k=0}^{n} \frac{-2kp}{n}\binom{n}{k} p^k (1-p)^{n-k} = -2p^2 \tag{2}$$

$$\sum_{k=0}^{n} p^2 \binom{n}{k} p^k (1-p)^{n-k} = p^2 \tag{3}$$

Zo vzťahov (1)—(3) už ľahko dostaneme, že

$$\sum_{k=0}^{n} \left(\frac{k}{n}-p\right)^2 \binom{n}{k} p^k (1-p)^{n-k} =$$

$$= p^2 + \frac{p(1-p)}{n} - 2p^2 + p^2 = \frac{p(1-p)}{n}$$

Rovnosť (3) dostaneme priamo z binomickej vety

$$\sum_{k=0}^{n} p^2 \binom{n}{k} p^k (1-p)^{n-k} = p^2(p+(1-p))^n = p^2$$

Aby sme dokázali (2), použijeme najprv vzťah $k\binom{n}{k} = n\binom{n-1}{k-1}$

$$\sum_{k=0}^{n} \frac{-2kp}{n}\binom{n}{k} p^k (1-p)^{n-k} =$$

$$= -\frac{2p}{n}\sum_{k=1}^{n} k\binom{n}{k} p^k (1-p)^{n-k} =$$

$$= -\frac{2p}{n} np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{n-1-k+1} =$$

$$= -2p^2(1+(1-p))^{n-1} = -2p^2$$

V prípade (1) musíme predošlú procedúru urobiť dvakrát:

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

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

$$=\frac{1}{n}\sum_{k=2}^{n}(k-1)\binom{n-1}{k-1}p^k(1-p)^{n-k}+\frac{p}{n}\sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1}(1-p)^{n-k}=$$

$$=\frac{n-1}{n}\sum_{k=2}^{n}\binom{n-2}{k-2}p^k(1-p)^{n-k}+\frac{p}{n}(p+(1-p))^{n-1}=$$

$$=\frac{n-1}{n}p^2(p+(1-p))^{n-2}+\frac{p}{n}=p^2+\frac{p(1-p)}{n}$$

Veta 4 (zákon veľkých čísel). Nech A je nejaká udalosť, ktorej pravdepodobnosť je p. Nech k_n je počet tých pokusov (pri nezávislom n-násobnom opakovaní), v ktorých udalosť A nastane (t. j. počet súradníc, ktoré patria do A). Potom pre každé $\varepsilon>0$ platí

$$\lim_{n\to\infty}P\left(\left|\frac{k_n}{n}-p\right|<\varepsilon\right)=1$$

Dôkaz. Nech α je množina tých k spomedzi $0, 1, \ldots, n$, pre ktoré platí

$$\left|\frac{k}{n}-p\right|\geqq\varepsilon$$

Potom

$$P\left(\left|\frac{k_n}{n}-p\right|\geqq\varepsilon\right)=\sum_{k\in\alpha}P(k_n=k)=$$

$$=\frac{1}{\varepsilon^2}\sum_{k\in\alpha}\varepsilon^2P(k_n=k)\leqq$$

$$\leqq\frac{1}{\varepsilon^2}\sum_{k\in\alpha}\left(\frac{k}{n}-p\right)^2\binom{n}{k}p^k(1-p)^{n-k}\leqq$$

$$\leqq\frac{1}{\varepsilon^2}\sum_{k=0}^{n}\left(\frac{k}{n}-p\right)^2\binom{n}{k}p^k(1-p)^{n-k}=$$

$$=\frac{p(1-p)}{n\varepsilon^2}\leqq\frac{1}{n\varepsilon^2}$$

Odtiaľ zase

$$P\left(\left|\frac{k_n}{n}-p\right|<\varepsilon\right)=1-P\left(\left|\frac{k_n}{n}-p\right|\geqq\varepsilon\right)\geqq$$

$$\geqq 1-\frac{1}{n\varepsilon^2}$$