MATEMATIKA

pre **4.** ročník gymnázia
a **8.** ročník gymnázia
s osemročným štúdiom

SPN
SLOVENSKÉ PEDAGOGICKÉ NAKLADATEĽSTVO

SPN

Zbyněk Kubáček

MATEMATIKA

pre **4.** ročník gymnázia

a **8.** ročník gymnázia
s osemročným štúdiom

SPN

SLOVENSKÉ PEDAGOGICKÉ NAKLADATEĽSTVO

Lektori: Mgr. Jana Fraasová
PaedDr. Iveta Kohanová, PhD.

Schválilo Ministerstvo školstva, vedy, výskumu a športu Slovenskej republiky pod č. 2013-11435/34301:4-919 zo dňa 22. júla 2013 ako učebnicu matematiky pre 4. ročník gymnázia a 8. ročník gymnázia s osemročným štúdiom. Schvaľovacia doložka má platnosť 5 rokov.

Prvé vydanie, 2013

Zodpovedná redaktorka RNDr. Jana Belasová
Výtvarná redaktorka Mgr. Ľubica Suchalová
Technická redaktorka Ivana Bronišová

Vyšlo vo vydavateľstve
Slovenské pedagogické nakladateľstvo – Mladé letá, s. r. o.,
Sasinkova 5, 811 08 Bratislava

Vytlačil Polygraf print, spol. s r. o., Prešov

ISBN 978-80-10-02290-8

OBSAH

ÚVOD

Čomu ľudia nerozumejú, z toho si robia posmech. Preto existuje toľko bonmotov o štatistike. Na úvod sme si pripravili zopár najznámejších. Sú vhodné najmä pre tých, ktorí štatistike rozumieť nechcú a potrebujú nejaký pekný citát ako výhovorku. Pre tých ostatných je určená táto učebnica.

Sú tri typy lží: lži, svinské lži a štatistika.

O tomto výroku sa dodnes vedú spory, kto slávny ho vlastne povedal. Existuje ešte jeden podobný:

Sú tri druhy lží: obyčajné lži, štatistika a pravda.

Štatistikou možno dokázať čokoľvek – aj pravdu.

Verím iba štatistike, ktorú som sám sfalšoval.

Výrok pripisovaný Winstonovi Churchillovi, s veľkou pravdepodobnosťou ho však do obehu uviedla nacistická propaganda.

Kedysi nemali štatistiku, tak sa museli uchýliť ku klamstvám.

Stephen Leacock

Jeden z problémov pri používaní štatistických metód je správne pochopiť, čo vlastne „hovoria“ čísla, ktoré dostaneme po dosadení do rôznych štatistických vzorcov. **Matematika** totiž – hoci si to možno mnohí myslia – **nie je súbor vzorcov**, do ktorých sa dosadzuje. Oveľa bližšie k pravde je tvrdenie, že **matematika je cesta k týmto vzorcom** spojená so schopnosťou rozhodnúť, **kedy, čo a prečo do vzorca dosadiť** a čo znamená **výsledok**, ktorý sme tým dosadením dostali.

Takýto pohľad na matematiku ovplyvnil aj podobu tejto učebnice. Na niekoľkých desiatkach strán sa stretnete iba s tromi vzorcami – pozri tabuľku, zvyšok textu učebnice vypĺňa naša snaha ukázať, čo vyjadrujú, ako sa na ne prišlo a čo pomocou nich možno zistiť.

$\binom{n}{k} p^k (1-p)^{n-k}$	**pravdepodobnosť**, že pri n-násobnom opakovaní pokusu, v ktorom pravdepodobnosť úspechu je p (pokusom môže byť napr. hod mincou, pričom úspech je padnutie znaku, vtedy $p = 0{,}5$) dosiahneme presne k úspechov (teda znak padne k-krát)
$\left(np - 2\sqrt{np(1-p)},\ np + 2\sqrt{np(1-p)}\right)$	**interval**, v ktorom sa nachádza počet úspechov, ktorý *prakticky iste* dosiahneme pri n-násobnom opakovaní pokusu, v ktorom pravdepodobnosť úspechu je p
$p = \frac{k}{n} \pm \frac{1}{\sqrt{n}}$	číslo $\frac{1}{\sqrt{n}}$ určuje **presnosť**, s akou vieme zistiť pravdepodobnosť úspechu p v jednoduchom pokuse, ak sme pri n opakovaniach tohto pokusu dosiahli presne k úspechov

Ako vidno, vzorce nie sú zložité – iste nie zložitejšie ako vzorec na výpočet koreňov kvadratickej rovnice $x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ – a dosadiť napr. do výrazu $p = \frac{k}{n} \pm \frac{1}{\sqrt{n}}$ hodnoty k, n tiež nevyžaduje nejaké mimoriadne úsilie. Rovnako jednoducho sa dajú vymyslieť aj úlohy na precvičenie tohto vzťahu („z 1 200 náhodne vybraných dospelých občanov SR 214 odpovedalo, že nesúhlasí s povinnou školskou dochádzkou; ako presne vieme zistiť podiel takýchto občanov v celej dospelej populácii?“ a pod.). Preto ich vymýšľanie – alebo ešte radšej hľadanie vhodných námetov napr. v dennej tlači – prenecháme na vás. My sa namiesto precvičovania vzorcov sústredíme na cestu vedúcu k ich **pochopeniu**. Chceme, aby ste si ju prešli spolu s nami (nepožadujeme, aby ste ju boli schopní samostatne zopakovať, ale jednotlivé kroky by ste mali zvládnuť sami). Nepredstavujeme si to však tak, že prečítate zopár strán v učebnici, alebo si vypočujete výklad v škole. Aby naša spoločná práca mala zmysel, vyžaduje to vašu aktívnu spoluprácu – tou bude **riešenie úloh**. Na to, aby ste pochopili, čo naše tri vzorce „hovoria“, totiž stačí vyriešiť zopár príkladov, ktoré by ste mali vedieť vypočítať.

Na túto učebnicu sa teda môžete pozerať ako na zbierku úloh, ktorými si precvičujete poznatky a zručnosti, ktoré ste mali získať v predchádzajúcich ročníkoch (výpočet pravdepodobnosti, dosadenie do vzorca, použitie tabuľkového kalkulátora, odhad chyby približného čísla a pod.). Učebnica však má pridanú hodnotu – je zostavená tak, aby ju bolo možné využiť pri vysvetlení niektorých základných štatistických myšlienok. Stručne povedané: vy budete počítať a my vám vysvetlíme, čo z toho vyplýva. Oboznámite sa tak s niekoľkými úvahami, ktoré sa nachádzajú v základoch štatistiky. Veríme, že vďaka ich pochopeniu bude štatistika pre vás zrozumiteľnejšia.

V texte učebnice sa pri hľadaní pravdepodobností sústreďujeme na výpočet teoretickej hodnoty. Naliehavo však odporúčame všade, kde je to možné, doplniť výpočet aj experimentovaním. Experiment je dôležitý most medzi teoretickými úvahami a realitou, na ktorý by sa pri vyučovaní nemalo zabúdať.

Začali sme citátmi, preto – kvôli symetrii – citátmi tento úvod aj skončíme:

Je ľahké klamať so štatistikou, ešte ľahšie je klamať bez nej.

Frederick Mosteller, jeden z významných štatistikov 20. storočia

Používanie štatistických vzorcov nenahrádza pochopenie toho, čo človek vlastne robí.

Hubert Morse Blalock, Jr., sociológ známy svojimi prácami o štatistických metódach výskumu

Je mojou milou povinnosťou poďakovať všetkým, ktorí prispeli k výslednej podobe tejto učebnice: lektorkám PaedDr. Ivete Kohanovej, PhD. a Mgr. Jane Fraasovej, ktoré upozornili na viaceré nedostatky pôvodného rukopisu, PaedDr. Jánovi Žabkovi za sústavnú morálnu podporu, môjmu otcovi prof. RNDr. Ing. Lubomírovi Kubáčkovi, DrSc., Dr.h.c., za trpezlivosť pri nekonečných konzultáciách, ktoré mi ochotne poskytol a všetkým ďalším nemenovaným, bez ktorých by som nikdy nedokázal rukopis dokončiť.

Autor

1 JE TÁ MINCA FALOŠNÁ?

„Pekne o tom hovorí Winwood Reade," povedal Holmes.
„Tvrdí, že jednotlivec síce je neriešiteľná záhada, ale v množstve sa z neho stáva matematická istota. Človek napríklad nikdy nemôže predvídať, čo urobí jednotlivec, ale môže presne uhádnuť, ako sa zachová určité množstvo priemerných ľudí. Jedinci sa navzájom líšia, ale percentá sa nemenia. Štatistiky to dokazujú."

ARTHUR CONAN DOYLE: *Podpis štyroch*

V tejto kapitole sa vrátime k úvahám o pravdepodobnosti. V predchádzajúcich ročníkoch sme pravdepodobnosť opisovali ako ideálnu relatívnu početnosť (pripomenieme si to v článku 1.1). Takýto opis je v súlade s našimi skúsenosťami: napr. formulácia *pravdepodobnosť, že pri hode mincou padne znak je* ½ vyjadruje naše očakávanie, že pri veľkom počte hodov približne v polovici z nich padne znak. Neočakávame, že to bude *presne* polovica z celkového počtu hodov, očakávame však, že relatívna početnosť znakov (teda podiel *počet znakov* : *počet pokusov*) sa od ½ nebude veľmi líšiť.

Hlavná náplň tejto kapitoly súvisí so slovami *nebude sa veľmi líšiť* z predchádzajúcej vety. Bude nás zaujímať, ako veľké odchýlky od ideálnej hodnoty ½ ešte možno pokladať za normálne a ktoré už nie. Asi sa zhodneme na tom, že ak pri desiatich hodoch mincou nepadne znak ani raz, vzbudí to naše podozrenie, že minca nie je v poriadku – teda, že pravdepodobnosť padnutia znaku v jej prípade **nie je** ½. Otázkou je, či je takéto podozrenie namieste aj vtedy, keď nám z desiatich pokusov padol znak iba raz alebo dvakrát. Aby sme mohli podobné otázky zodpovedať, potrebujeme zistiť, ako pravdepodobné sú jednotlivé výsledky pri veľkom počte hodov – napr. ako pravdepodobné je, že pri 1 000 hodoch mincou padne znak 429-krát. Hľadanie odpovede nás privedie k *Pascalovmu trojuholníku* – trojuholníkovej tabuľke čísel s viacerými zaujímavými vlastnosťami, k *bernoulliovským pokusom* – nazvaným podľa švajčiarskeho matematika Jacoba Bernoulliho (1654 – 1705) a k *binomickému rozdeleniu pravdepodobnosti*.

1.1 Pravdepodobnosť ako ideálna relatívna početnosť

Ak šanca, že sa niečo pokazí, je 1 : 1, tak sa to pokazí v 9 prípadoch z 10.

Výsledok hodu mincou – teda to, či padne číslo alebo znak – nevieme odhadnúť vopred. Predpokladáme však, že pri štandardnej minci je jedna aj druhá možnosť rovnako pravdepodobná, t. j. že pri veľkom počte pokusov v ideálnom prípade v polovici týchto pokusov padne znak (a v zvyšnej polovici číslo).

Pripomíname základné pojmy:

HROMADNÝ NÁHODNÝ DEJ (tiež náhodný pokus alebo iba pokus) = dej, ktorého výsledok nie je vopred jednoznačne určený (nemožno ho vopred jednoznačne predpovedať) a ktorý sa vyskytuje hromadne (napr. životnosť veľkého počtu rovnakých žiaroviek) alebo opakovane (napr. žrebovanie čísel v lotérii).

NÁHODNÁ UDALOSŤ (náhodný jav) = jeden z možných výsledkov náhodného pokusu

RELATÍVNA POČETNOSŤ náhodnej udalosti (v danej skupine pokusov) je podiel

$$\frac{\textit{počet pokusov, v ktorých nastala sledovaná udalosť}}{\textit{celkový počet pokusov}}$$

To vyjadrujeme formuláciou *pravdepodobnosť, že padne znak, je jedna polovica*. Pripomeňme, že hodnotu pravdepodobnosti vyjadrujeme

	v našom prípade
ako číslo ležiace medzi 0 a 1	*pravdepodobnosť, že padne znak, je* $\frac{1}{2}$ *pravdepodobnosť, že padne znak, je* 0,5
ako počet percent	*pravdepodobnosť, že padne znak, je* 50 %

V skutočnosti ideálny prípad nenastáva často. Ak hodíme veľakrát mincou a budeme sledovať, koľkokrát padne znak, spravidla sa skutočná relatívna početnosť tejto náhodnej udalosti (padnutia znaku) bude od čísla ½ odlišovať. Teda znak len zriedkakedy padne presne v polovici z celkového počtu pokusov. Očakávame však, že s rastúcim počtom hodov sa rozdiel medzi skutočnou relatívnou početnosťou a ideálnou hodnotou ½ bude stále zmenšovať.

Pozor, zo zmenšovania rozdielu medzi relatívnou početnosťou a číslom $\frac{1}{2}$ **nevyplýva**, že sa musí zmenšovať aj rozdiel medzi počtom pokusov, v ktorých padol znak (nazvime ich priaznivé) a počtom tých, v ktorých padlo číslo (nepriaznivé).

Ak napr. znak padne

- pri 100 hodoch 52-krát, je relatívna početnosť 0,52 a rozdiel medzi počtom priaznivých a nepriaznivých výsledkov je 52 – 48 = 4
- pri 1 000 hodoch 510-krát, je relatívna početnosť 0,51 a rozdiel medzi počtom priaznivých a nepriaznivých výsledkov 510 – 490 = 20. Teda odchýlka relatívnej početnosti od ideálnej hodnoty $\frac{1}{2}$ je menšia ako v predchádzajúcom prípade, hoci rozdiel medzi počtom priaznivých a nepriaznivých hodov je väčší.

Inak povedané: očakávame, že pri veľmi veľkom počte hodov sa bude relatívna početnosť veľmi málo líšiť od čísla $\frac{1}{2}$.

Predstavou ideálneho prípadu si môžeme pomôcť aj pre iné hodnoty pravdepodobnosti. Formuláciu *pravdepodobnosť náhodnej udalosti je p %* alebo *pravdepodobnosť náhodnej udalosti je* $\frac{p}{100}$ môžeme interpretovať: v *ideálnom prípade* nastane táto udalosť v *p* % z veľkého počtu pokusov.

pravdepodobnosť náhodnej udalosti je 22,5 % (alebo 0,225) =
= v ideálnom prípade nastane táto udalosť v 22,5 % z veľkého počtu pokusov

Pravdepodobnosť teda môžeme chápať ako ideálnu relatívnu početnosť. Predpokladáme, že so zväčšujúcim sa počtom pokusov sa bude skutočná relatívna početnosť stále menej odlišovať od ideálnej hodnoty – pravdepodobnosti. Na predstave ideálneho prípadu budú založené viaceré výpočty pravdepodobnosti v nasledujúcich odsekoch.

1.2 Ako pravdepodobné sú jednotlivé výsledky?

- Prvé riešenie (postupne zväčšujeme počet opakovaní pokusu)
- Pascalov trojuholník
- Druhé riešenie (pomocou kombinačných čísel)
- Malá odbočka: Pascalov trojuholník a kombinačné čísla
- Bernoulliovské pokusy a binomické rozdelenie pravdepodobnosti

Vráťme sa k otázke z úvodu. Chceme diskutovať o tom, aký počet výsledkov *padol znak* pri veľkom počte hodov mincou ešte je a aký už nie je *normálny*. Preto potrebujeme zistiť, ako pravdepodobné sú jednotlivé možnosti. Ukážeme dva spôsoby riešenia tejto otázky. Prvý nás privedie k číselnej tabuľke, ktorá sa označuje ako *Pascalov trojuholník*, z druhého zistíme, že hodnoty v tejto tabuľke sú vlastne *kombinačné čísla*.

Predtým ešte malá terminologická poznámka. V tejto kapitole sa viackrát stane, že slovo pokus použijeme pri označení dvoch rôznych vecí. Napríklad pri skúmaní výsledkov desiatich hodov mincou bude predmetom nášho záujmu náhodný pokus, ktorý vznikne 10-násobným opakovaním jednoduchého náhodného pokusu (tým je jeden hod mincou). Pokusom pritom budeme označovať aj jeden hod mincou, aj jeho 10-násobné opakovanie. Veríme však, že z textu bude vždy jasné, čo máme na mysli.

PRVÉ RIEŠENIE (POSTUPNE ZVÄČŠUJEME POČET OPAKOVANÍ POKUSU)

Úvahy urobíme pre 10 hodov mincou, ktoré sme spomenuli v úvode. Budú nás zaujímať pravdepodobnosti 11 náhodných udalostí, ktoré môžu byť výsledkami tohto pokusu: *znak nepadol ani raz, znak padol* 1*-krát, znak padol* 2*-krát, ..., znak padol* 10*-krát*.

Postupne budeme počítať pravdepodobnosti pre výsledky po 2 hodoch, 3 hodoch atď. a dúfať, že vo výpočtoch sa objaví nejaká zákonitosť, ktorá nám pomôže nájsť odpoveď aj pre iné počty hodov.

Tento prístup je v matematike dosť častý: ak neviete vyriešiť komplikovanú úlohu, skúste začať riešením podobných, ale jednoduchších úloh.

ÚLOHA

1. a) Aká je pravdepodobnosť, že pri 2 hodoch mincou padne znak 0-krát (práve 1-krát, 2-krát)?
b) Aká je pravdepodobnosť jednotlivých možností pri 3 a 4 hodoch mincou?

RIEŠENIE

a) Pomôžeme si predstavou *ideálneho prípadu*. Pravdepodobnosť, že padne znak, je $\frac{1}{2}$, preto pri veľmi veľkom počte hodov mincou v ideálnom prípade padne v polovici hodov znak (a v zvyšnej polovici číslo). Predpokladajme, že sme urobili veľký počet pokusov, teda veľakrát sme 2-krát za sebou hodili mincou.

Pozrime sa na výsledky prvého hodu vo všetkých týchto pokusoch. V ideálnom prípade v prvom hode v polovici prípadov padne **číslo** a v druhej polovici **znak**. Výsledok po prvom hode teda bude:

$\frac{1}{2}$ pokusov **0 znakov**	$\frac{1}{2}$ pokusov **1 znak**

Výsledok 2. hodu nezávisí od toho, čo padlo v 1. hode. Preto výsledok 2. hodu bude rovnaký v tej skupine pokusov, kde v 1. hode padlo **číslo**, aj v tej skupine pokusov, kde v 1. hode padol **znak** (túto úvahu si dobre rozmyslite a diskutujte o nej).

To znamená, že v ideálnom prípade:

z pokusov,v ktorých v 1. hode padlo **číslo**		z pokusov, v ktorých v 1. hode padol **znak**	
padne v 2. hode		padne v 2. hode	
v polovici prípadov (teda v štvrtine z celkového počtu pokusov) **číslo**	v polovici prípadov **znak**	v polovici prípadov **číslo**	v polovici prípadov **znak**
$\frac{1}{4}$ pokusov	$\frac{1}{4}$ pokusov	$\frac{1}{4}$ pokusov	$\frac{1}{4}$ pokusov

Pozrime sa teraz na počet znakov, ktoré padli v prvých dvoch hodoch.
Ak v 2. hode padlo

- **číslo**, tak sa celkový počet znakov od predchádzajúceho pokusu nezmenil (v tabuľke to znázorňujeme zápisom **+ 0**),
- **znak**, tak sa celkový počet znakov od predchádzajúceho pokusu zväčšil o 1 (znázorňujeme to zápisom **+ 1**).

+ 0	**+ 1**	**+ 0**	**+ 1**

Preto v jednotlivých prípadoch znak padol celkom

0-krát (*ani v jednom hode*)	1-krát (*iba v 2. hode*)	1-krát (*iba v 1. hode*)	2-krát (*v obidvoch hodoch*)

a výsledok po 2 hodoch bude

$\frac{1}{4}$ pokusov **0 znakov**	$\frac{2}{4}$ pokusov **1 znak**	$\frac{1}{4}$ pokusov **2 znaky**

Teda v ideálnom prípade v 2 hodoch padne znak

- 0-krát v $\frac{1}{4}$ prípadov (teda pravdepodobnosť, že pri dvoch hodoch mincou nepadne znak ani raz, je $\frac{1}{4}$, t. j. 0,25 alebo 25 %),
- 1-krát v $\frac{2}{4}$, t. j. v $\frac{1}{2}$ prípadov (teda pravdepodobnosť, že ....),
- 2-krát v $\frac{1}{4}$ prípadov (teda ....).

DOPLŇTE CHÝBAJÚCI TEXT V DRUHEJ A TRETEJ ODRÁŽKE.

Naše úvahy súvisia s pojmami *podmienená pravdepodobnosť* a *nezávislé udalosti*. Pravdepodobnosti, ktoré sme počítali na tejto strane (pozri tiež tretí riadok v tab. 1 na s. 12, nad riadkom po 2. hode), sú podmienené, napr. v **červenej** bunke vľavo je to *pravdepodobnosť, že v 2. hode padne číslo, za predpokladu, že v 1. hode padlo číslo*, v nasledujúcej **modrej** bunke je to *pravdepodobnosť, že v 2. hode padne znak, za predpokladu, že v 1. hode padlo číslo*. Pri úvahách o týchto pravdepodobnostiach sa na dva za sebou nasledujúce hody mincou pozeráme ako na dvojicu náhodných pokusov, v ktorej výsledky druhého pokusu (druhý hod mincou) sú nezávislé od výsledkov prvého pokusu (prvý hod mincou) – v súvislosti s nezávislosťou pri hodoch mincou si spomeňte na heslo *Minca nemá pamäť*. Podobne sa možno pozerať aj na tri a viac hodov mincou.

Pripomeňme, že podobnými úvahami, aké sme použili pri určení pravdepodobností na s. 11, sme v učebnici pre 2. ročník zdôvodnili pravidlo na výpočet pravdepodobnosti v prípade nezávislých udalostí: pravdepodobnosť, že v jednom z dvojice pokusov nastane udalosť A a v druhom udalosť B, je súčin pravdepodobností $p(A)$ a $p(B)$. Napríklad pravdepodobnosť, že v prvých dvoch hodoch nepadol znak (= udalosť A) a v treťom padne znak (= udalosť B), je $\frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}$. Skontrolujte, že výsledky uvedené v tab. 1 sú v súlade s týmto pravidlom.

b) Ak v týchto úvahách budeme pokračovať, dostaneme pre počet znakov po 3 a 4 hodoch výsledky, ktoré znázorňuje spodná časť tab. 1 (v jej hornej časti sa opakujú úvahy pre 1. a 2. hod).

$\frac{1}{2}$ pokusov	$\frac{1}{2}$ pokusov	
$\frac{1}{2}$ pokusov **0 znakov**	$\frac{1}{2}$ pokusov **1 znak**	po 1. hode

$\frac{1}{4}$ pokusov **+ 0**	$\frac{1}{4}$ pokusov **+ 1**	$\frac{1}{4}$ pokusov **+ 0**	$\frac{1}{4}$ pokusov **+ 1**	
$\frac{1}{4}$ pokusov **0 znakov**	$\frac{2}{4}$ pokusov **1 znak**		$\frac{1}{4}$ pokusov **2 znaky**	po 2. hode

$\frac{1}{8}$ **+ 0**	$\frac{1}{8}$ **+ 1**	$\frac{2}{8}$ **+ 0**	$\frac{2}{8}$ **+ 1**	$\frac{1}{8}$ **+ 0**	$\frac{1}{8}$ **+ 1**	
$\frac{1}{8}$ **0 znakov**	$\frac{3}{8}$ **1 znak**		$\frac{3}{8}$ **2 znaky**		$\frac{1}{8}$ **3 znaky**	po 3. hode

$\frac{1}{16}$ **+ 0**	$\frac{1}{16}$ **+ 1**	$\frac{3}{16}$ **+ 0**	$\frac{3}{16}$ **+ 1**	$\frac{3}{16}$ **+ 0**	$\frac{3}{16}$ **+ 1**	$\frac{1}{16}$ **+ 0**	$\frac{1}{16}$ **+ 1**	
$\frac{1}{16}$ **0 znakov**	$\frac{4}{16}$ **1 znak**		$\frac{6}{16}$ **2 znaky**		$\frac{4}{16}$ **3 znaky**		$\frac{1}{16}$ **4 znaky**	po 4. hode

Tab. 1
Pravdepodobnosť padnutia znakov po 1., 2., 3., a 4. hode.

Naznačme úvahy pre 3. hod (pozri v tab. 1 riadok, ktorý je pod riadkom *po* 2. *hode*).

V ideálnom prípade v 3. hode

z tých pokusov, v ktorých v prvých dvoch hodoch nepadol ani jeden znak (tie tvoria $\frac{1}{4}$ z celkového počtu pokusov), padne

z tých pokusov, v ktorých v prvých dvoch hodoch padol znak iba raz (tie tvoria $\frac{2}{4}$ z celkového počtu pokusov), padne

z tých pokusov, v ktorých v prvých dvoch hodoch padli 2 znaky (tie tvoria $\frac{1}{4}$ z celkového počtu pokusov), padne

$\frac{1}{4}$ pokusov **0 znakov**		$\frac{2}{4}$ pokusov **1 znak**		$\frac{1}{4}$ pokusov **2 znaky**		po 2. hode
$\frac{1}{8}$ **+ 0**	$\frac{1}{8}$ **+ 1**	$\frac{2}{8}$ **+ 0**	$\frac{2}{8}$ **+ 1**	$\frac{1}{8}$ **+ 0**	$\frac{1}{8}$ **+ 1**	

v polovici prípadov **číslo** (to je $\frac{1}{2}$ z $\frac{1}{4}$ z celkového počtu pokusov, teda $\frac{1}{8}$ z celkového počtu pokusov), v týchto prípadoch padlo po 3 hodoch celkom **0 znakov**

a v polovici prípadov **znak** (to je tiež $\frac{1}{8}$ z celkového počtu pokusov), v týchto prípadoch padol po 3 hodoch celkom **1 znak**

v polovici prípadov **číslo** (to je $\frac{1}{2}$ z $\frac{1}{4}$ z celkového počtu pokusov, teda $\frac{2}{8}$ z celkového počtu pokusov), v týchto prípadoch padol po 3 hodoch celkom **1 znak**

a v polovici prípadov **znak** (to sú tiež $\frac{2}{8}$ z celkového počtu pokusov), v týchto prípadoch padli po 3 hodoch celkom **2 znaky**

v polovici prípadov **číslo** (to je $\frac{1}{8}$ z celkového počtu pokusov), v týchto prípadoch padli po 3 hodoch celkom **2 znaky**

a v polovici prípadov **znak** (to je tiež $\frac{1}{8}$ z celkového počtu pokusov), v týchto prípadoch padli po 3 hodoch celkom **3 znaky**

Preto po 3 hodoch padne (pozri riadok *po* 3. *hode* v tab. 1)

- 0 znakov v $\frac{1}{8}$ z celkového počtu pokusov,
- 1 znak v $\frac{1}{8} + \frac{2}{8}$, t. j. v $\frac{3}{8}$ z celkového počtu pokusov,
- 2 znaky v $\frac{2}{8} + \frac{1}{8} = \frac{3}{8}$ z celkového počtu pokusov,
- 3 znaky v $\frac{1}{8}$ z celkového počtu pokusov.

ÚLOHY

2. V diskusii sa uistite, že ste porozumeli úvahám, ktoré sme použili pri vypĺňaní tab. 1 a skontrolujte správnosť jednotlivých pravdepodobností po 4. hode.

3. Použite tieto úvahy na zistenie počtu znakov po 5. a 6. hode. Výsledky zapíšte do tabuľky a vyslovte ich „v reči pravdepodobnosti“ (napr. pravdepodobnosť, že pri 6 hodoch mincou padne znak práve 2-krát, je ... .)

PASCALOV TROJUHOLNÍK

Všimnime si teraz čísla v bielych riadkoch tab. 1 na s. 12, teda jednotlivé pravdepodobnosti. Vedome sme ich zapísali tak, aby všetky zlomky v jednom riadku mali rovnakého menovateľa (tým je v 1. riadku číslo 2, v druhom číslo $4 = 2^2$, v treťom číslo $8 = 2^3$, v štvrtom číslo $16 = 2^4$, pozri tab. 2).

<table>
<tr><td colspan="8">$\frac{1}{2}$</td><td colspan="8">$\frac{1}{2}$</td><td>1. hod</td></tr>
<tr><td colspan="4">$\frac{1}{4}$</td><td colspan="8">$\frac{2}{4}$</td><td colspan="4">$\frac{1}{4}$</td><td>2. hod</td></tr>
<tr><td colspan="2">$\frac{1}{8}$</td><td colspan="6">$\frac{3}{8}$</td><td colspan="6">$\frac{3}{8}$</td><td colspan="2">$\frac{1}{8}$</td><td>3. hod</td></tr>
<tr><td>$\frac{1}{16}$</td><td colspan="4">$\frac{4}{16}$</td><td colspan="6">$\frac{6}{16}$</td><td colspan="4">$\frac{4}{16}$</td><td>$\frac{1}{16}$</td><td>4. hod</td></tr>
</table>

Tab. 2
Jednotlivé pravdepodobnosti z tab. 1.

ÚLOHA

4. Aký je súčet čísel v jednotlivých riadkoch tab. 2? Výsledok možno zdôvodniť jednoduchými úvahami o celku (ktorý predstavujú všetky pokusy) a jeho častiach, urobte to! Potom svoje zdôvodnenie vyjadrite v reči pravdepodobnosti.

Úlohou 4 si tiež pripomíname jednoduché vlastnosti pravdepodobnosti, ktoré možno zdôvodniť úvahami o súbore všetkých pokusov a jeho častiach:

- pravdepodobnosť istej udalosti je 1 (istá udalosť nastane pri každom pokuse),
- ak A, B sú dva výsledky náhodného pokusu s pravdepodobnosťami $p(A)$, $p(B)$, ktoré **nemôžu** nastať súčasne, tak pravdepodobnosť, že nastane A alebo B, je $p(A \text{ alebo } B) = p(A) + p(B)$ (pripomeňte si v diskusii zdôvodnenie tejto vlastnosti).

Teraz nás budú zaujímať iba čitatele zlomkov z tab. 2, uvádzame ich v tab. 3.

<table>
<tr><td colspan="8">1</td><td colspan="8">1</td><td>1. hod</td></tr>
<tr><td colspan="4">1</td><td colspan="8">2</td><td colspan="4">1</td><td>2. hod</td></tr>
<tr><td colspan="2">1</td><td colspan="6">3</td><td colspan="6">3</td><td colspan="2">1</td><td>3. hod</td></tr>
<tr><td>1</td><td colspan="4">4</td><td colspan="6">6</td><td colspan="4">4</td><td>1</td><td>4. hod</td></tr>
</table>

Tab. 3
Čitatele zlomkov z tab. 2.

ÚLOHA

5. Aký je súčet čísel v jednotlivých riadkoch tab. 3? Výsledok možno odvodiť z výsledku úlohy 4, urobte to.

Súčtová vlastnosť

Ak ste správne vyriešili úlohu 3 na s. 13, iste ste objavili dôležitú vlastnosť čísel z tab. 3:

- na pravom aj ľavom kraji každého riadka je číslo 1,
- každé vnútorné číslo v nasledujúcom riadku je súčet dvoch čísel nad ním z predchádzajúceho riadka.

<table>
<tr><td colspan="8">1</td><td colspan="8">1</td><td>1. hod</td></tr>
<tr><td colspan="4">1</td><td colspan="8">1 + 1 =
2</td><td colspan="4">1</td><td>2. hod</td></tr>
<tr><td colspan="2">1</td><td colspan="6">1 + 2 =
3</td><td colspan="6">2 + 1 =
3</td><td colspan="2">1</td><td>3. hod</td></tr>
<tr><td>1</td><td colspan="4">1 + 3 =
4</td><td colspan="6">3 + 3 =
6</td><td colspan="4">3 + 1 =
4</td><td>1</td><td>4. hod</td></tr>
</table>

Súčtová vlastnosť čísel z tab. 3.

ÚLOHY

6. Doplňte do tab. 3 ďalšie štyri riadky. Porovnajte čísla v piatom a šiestom riadku s vaším riešením úlohy 3.

7. Vysvetlite, ako možno riešenie úlohy 6 využiť pri hľadaní odpovedí na nasledujúce otázky a nájdite tieto odpovede:
a) Aká je pravdepodobnosť, že pri 7 hodoch mincou padnú práve 3 znaky?
b) Aká je pravdepodobnosť, že pri 8 hodoch mincou padnú najviac 2 znaky?
c) Aká je pravdepodobnosť, že pri 6 hodoch mincou padnú aspoň 4 znaky?

BLAISE PASCAL
(1623 – 1662)
VYNIKAJÚCI FRANCÚZSKY MATEMATIK A FILOZOF

Kvôli väčšej prehľadnosti sa tab. 3 väčšinou zapisuje v trojuholníkovom tvare (z neho je hneď zrejmé, súčtom ktorých dvojíc susedných čísel dostávame hodnoty v nasledujúcom riadku). Táto tabuľka sa nazýva **Pascalov trojuholník** na počesť Blaise Pascala, ktorý venoval číselnému trojuholníku celú knihu *Traité du triangle arithmétique (1654)*. V jej druhej časti opisuje použitie aritmetického trojuholníka pri riešení pravdepodobnostných úloh, predovšetkým slávnej úlohy o rozdelení výhry, ktorá stála pri zrode modernej teórie pravdepodobnosti (hovorili sme o nej v 2. ročníku).

riadok	
1.	1 1
2.	1 2 1
3.	1 3 3 1
4.	1 4 6 4 1
5.	1 5 10 10 5 1
6.	1 6 15 20 15 6 1
7.	1 7 21 35 35 21 7 1
8.	1 8 28 56 70 56 28 8 1
atď.	

Pascalov trojuholník

ÚLOHA

8. Teraz môžeme zodpovedať otázku z úvodu tohto článku: Ako pravdepodobné sú jednotlivé výsledky pri 10 hodoch mincou, teda ako pravdepodobné sú možnosti *padne* 0 *znakov*, *padne* 1 *znak*, ..., *padne* 10 *znakov*?

V ideálnom prípade padne šestka v $\frac{1}{6}$ z velkého počtu pokusov a v $\frac{5}{6}$ nepadne. Preto sa vo výsledku úlohy 9 budú vyskytovať súčiny týchto dvoch zlomkov.

Postup z úlohy 1 funguje aj pre iné pravdepodobnosti ako $\frac{1}{2}$

V nasledujúcej úlohe zistíme, že s číslami z tab. 3 – teda z Pascalovho trojuholníka – sa stretneme nielen pri hodoch mincou.

ÚLOHA

9. Úvahy, ktoré sme používali pri vyplňovaní tab. 1 na s. 12, použite na riešenie nasledujúcej úlohy: Ako pravdepodobné je, že pri 4 hodoch hracou kockou padne šestka 0-krát, 1-krát, 2-krát, 3-krát, 4-krát ?
Jednotlivé pravdepodobnosti v tabuľke píšte v tvare súčinu mocnín čísel $\frac{1}{6}$ a $\frac{5}{6}$, teda napr. $3 \cdot \left(\frac{5}{6}\right)^2 \cdot \frac{1}{6}$. Skontrolujte, že pri takomto zápise sa vo výsledkoch objavia čísla z prvých štyroch riadkov Pascalovho trojuholníka.

Ak ste správne vyriešili úlohu 9, dostali ste v poslednom riadku tabuľky tieto výsledky (**červeno** sme zvýraznili čísla zo 4. riadka Pascalovho trojuholníka):

pravdepodobnosť, že v 4 hodoch hracou kockou padne práve				
0 šestiek	1 šestka	2 šestky	3 šestky	4 šestky
$\mathbf{1} \cdot \left(\frac{5}{6}\right)^4$	$\mathbf{4} \cdot \left(\frac{5}{6}\right)^3 \cdot \frac{1}{6}$	$\mathbf{6} \cdot \left(\frac{5}{6}\right)^2 \cdot \left(\frac{1}{6}\right)^2$	$\mathbf{4} \cdot \frac{5}{6} \cdot \left(\frac{1}{6}\right)^3$	$\mathbf{1} \cdot \left(\frac{1}{6}\right)^4$
≈ 48,23 %	≈ 38,58 %	≈ 11,57 %	≈ 1,54 %	≈ 0,08 %

Tab. 4

Veríme, že pri riešení úlohy 9 ste objavili zákonitosť, ktorá vám umožní vyplniť ďalšie riadky tabuľky prakticky automaticky. Preverte si to v úlohe 10. (Ak si zistenou zákonitosťou nie ste istí, skúste vyplniť ešte niekoľko ďalších riadkov tabuľky z úlohy 9. Veríme, že pravidlo objavíte skôr, ako by ste sa pri takomto vypĺňaní dostali až do 12. riadka.)

ÚLOHA

10. Aká je pravdepodobnosť, že pri 12 hodoch hracou kockou
a) padnú práve 2 šestky,
b) padne práve 10 šestiek?

DRUHÉ RIEŠENIE (POMOCOU KOMBINAČNÝCH ČÍSEL)

Z úloh v predchádzajúcom článku vidno, aký význam majú čísla z Pascalovho trojuholníka pri výpočte pravdepodobností. Doteraz sme pri hľadaní týchto čísel využívali ich *súčtovú vlastnosť* (pozri s. 15). Pomocou nej síce dokážeme nájsť ľubovoľný prvok Pascalovho trojuholníka, je to však zdĺhavé (ak nás zaujímajú napr. prvky 50. riadka, musíme vyplniť všetkých predchádzajúcich 49 riadkov). V tomto odseku ukážeme, že jednotlivé čísla tohto trojuholníka vieme nájsť aj iným spôsobom – sú to totiž kombinačné čísla, na výpočet ktorých existuje jednoduchý vzorec.

Kombinačné číslo $\binom{n}{k}$ je počet k-prvkových kombinácií z n prvkov, teda počet možností, ako z n rôznych objektov vybrať skupinu obsahujúcu k objektov.
Na výpočet kombinačného čísla platí vzťah $\binom{n}{k} = \dfrac{n!}{(n-k)!\,k!}$, ktorý možno zapísať aj v tvare

$$\binom{n}{k} = \frac{\overbrace{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}^{k \text{ činiteľov}}}{\underbrace{k \cdot (k-1) \cdot \ldots \cdot 1}_{k \text{ činiteľov}}}$$

napr. $\binom{12}{3} = \dfrac{12!}{9!\,3!} = \dfrac{12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{(9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1) \cdot (3 \cdot 2 \cdot 1)} = \dfrac{12 \cdot 11 \cdot 10}{3 \cdot 2 \cdot 1}$.

(Pripomeňte si odvodenie tohto vzorca, napr. pomocou vzťahu medzi počtom k-prvkových variácií a kombinácií.)

O tom, že prvky Pascalovho trojuholníka sú kombinačné čísla, sa možno presvedčiť viacerými spôsobmi. My toto tvrdenie dokážeme vlastne mimochodom: budeme hľadať iný postup riešenia úlohy 9 a pritom objavíme vzťah medzi Pascalovým trojuholníkom a kombinačnými číslami. Ukážme si tento postup najprv na jednom konkrétnom čísle: V riešení úlohy 9 sme zistili, že pri **4** hodoch hracou kockou padnú práve **2** šestky s pravdepodobnosťou **6** $\cdot \left(\frac{5}{6}\right)^2 \cdot \left(\frac{1}{6}\right)^2$ (pozri tab. 4). Z riešenia úloh 11 až 13 vyplynie, že číslo **6** je kombinačné číslo $\binom{4}{2}$. Podobné úvahy budú platiť aj pre ostatné čísla v Pascalovom trojuholníku, uvedieme ich po vyriešení úloh 14 až 16.

ÚLOHA

11. Vypočítajte pravdepodobnosť, že v 4 hodoch hracou kockou hodíme šestku
a) iba v 1. a 4. hode,
b) iba v 2. a 3. hode.

RIEŠENIE

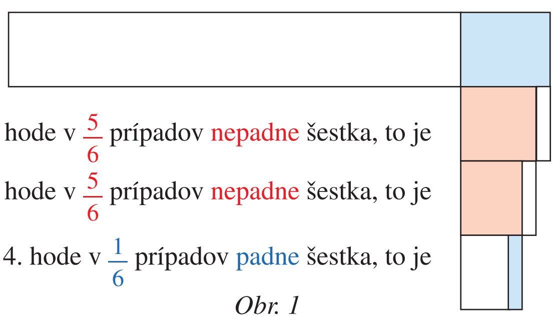
a) Použijeme rovnaký spôsob uvažovania ako pri vypĺňaní tab. 1. Stručne je celý postup znázornený na obr. 1.

V 1. hode v $\frac{1}{6}$ prípadov padne šestka

Z tejto $\frac{1}{6}$ pokusov v 2. hode v $\frac{5}{6}$ prípadov nepadne šestka, to je — $\frac{1}{6} \cdot \frac{5}{6}$ všetkých pokusov

Z týchto $\frac{1}{6} \cdot \frac{5}{6}$ pokusov v 3. hode v $\frac{5}{6}$ prípadov nepadne šestka, to je — $\frac{1}{6} \cdot \frac{5}{6} \cdot \frac{5}{6}$ všetkých pokusov

Z týchto $\frac{1}{6} \cdot \frac{5}{6} \cdot \frac{5}{6}$ pokusov v 4. hode v $\frac{1}{6}$ prípadov padne šestka, to je — $\frac{1}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{1}{6}$ všetkých pokusov



Obr. 1

Pravdepodobnosť, že pri hode hracou kockou padne šestka, je $\frac{1}{6}$. Pravdepodobnosť, že šestka nepadne, je $\frac{5}{6}$. Takéto dve pravdepodobnosti sa nazývajú navzájom *doplnkové.* Pripomeňme jednoduchý poznatok: ak p je pravdepodobnosť, že **nastane** nejaká náhodná udalosť, tak pravdepodobnosť, že táto náhodná udalosť **nenastane**, je $1 - p$. (Pripomeňte si zdôvodnenie tohto tvrdenia napr. úvahami o ideálnych pokusoch.)

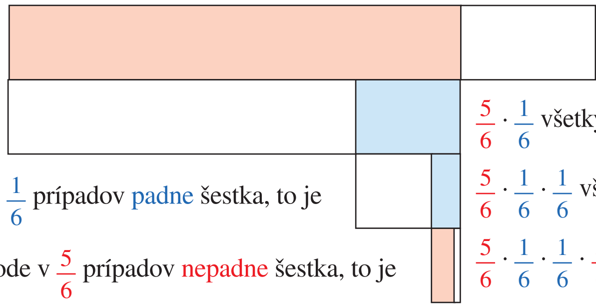
b) Riešenie úlohy 11 b) je podobné, postup sme znázornili na obr. 2. Vidno z neho, že aj v tomto prípade sa hľadaná pravdepodobnosť (že pri štyroch hodoch kockou padne šestka iba v 2. a 3. hode) rovná $\left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^2$.

V 1. hode v $\frac{5}{6}$ prípadov nepadne šestka

Z nich v 2. hode v $\frac{1}{6}$ prípadov padne šestka, to je $\frac{5}{6} \cdot \frac{1}{6}$ všetkých pokusov

Z týchto $\frac{5}{6} \cdot \frac{1}{6}$ pokusov v 3. hode v $\frac{1}{6}$ prípadov padne šestka, to je $\frac{5}{6} \cdot \frac{1}{6} \cdot \frac{1}{6}$ všetkých pokusov

Z týchto $\frac{5}{6} \cdot \frac{1}{6} \cdot \frac{1}{6}$ pokusov v 4. hode v $\frac{5}{6}$ prípadov nepadne šestka, to je $\frac{5}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{5}{6}$ všetkých pokusov



Obr. 2

ÚLOHY

12. Uveďte všetky zvyšné možnosti, ako môžu pri 4 hodoch kockou padnúť práve 2 šestky.

13. Pre každú možnosť z riešenia úlohy 12 vypočítajte pravdepodobnosť, že táto možnosť nastane.

V úlohe 11 sme uviedli dve rôzne možnosti, ako môže pri 4 hodoch hracou kockou padnúť šestka práve dvakrát (v 1. a 4. hode, v 2. a 3. hode).

Vo výsledku sa objavuje kombinačné číslo

Z riešenia úloh 11 až 13 vyplýva: Existuje celkom **6** rôznych možností, ako môžu padnúť práve 2 šestky pri štyroch hodoch hracou kockou:

v 1. a 2. hode, v 1. a 3., v 1. a 4., v 2. a 3. , v 2. a 4., v 3. a 4. hode, (*)

to ste zistili v úlohe 12.

Pravdepodobnosť každej z týchto možností je $\left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^2$, to ste mali zistiť v úlohe 13. Pravdepodobnosť, že v štyroch hodoch kockou padne šestka práve dvakrát, je súčtom týchto **6** rovnakých pravdepodobností, preto jej hodnota je $\mathbf{6} \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^2$.

Ako vidíme, dostali sme rovnaký výsledok ako v tab. 4 na s. 16. To samozrejme nie je nič prekvapujúce (podivné by bolo, keby sme dvomi spôsobmi výpočtu tej istej pravdepodobnosti dostali dva rôzne výsledky). Zaujímavé pre nás je, ako sme v tomto druhom postupe prišli k číslu **6**: šesť bol počet všetkých možností, ako zo **4** hodov vybrať tie **2**, v ktorých padne šestka – pozri riadok označený (*). Bol to teda počet **2**-prvkových kombinácií zo **4** prvkov (prvkami boli čísla 1, 2, 3, 4, ktoré označovali poradové číslo hodu). To znamená, že hodnotu **6** sme našli ako kombinačné číslo $\binom{4}{2}$, preto výslednú pravdepodobnosť môžeme zapísať ako $\binom{4}{2} \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^2$.

Pravdepodobnosť, že v **4** hodoch kockou padnú práve **2** šestky, je $\binom{4}{2} \cdot \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^{4-2}$.

Úvahy, ktoré sme v úlohách 11 až 13 urobili pre konkrétnu pravdepodobnosť úspechu $p = \frac{1}{6}$ (*úspechom je padnutie šestky*), konkrétny počet pokusov $n = 4$ (*pokusom je hod kockou*) a konkrétny počet úspechov $k = 2$, môžeme použiť aj pre iné hodnoty n, k a p.

Vyskúšame si to v nasledujúcich úlohách (opakovanie uvedených úvah je tiež príležitosťou na ich lepšie pochopenie). V prvých dvoch zostaneme pri hádzaní kockou, teda ponecháme hodnotu $p = \frac{1}{6}$:

- v úlohe 14 zapíšeme pomocou kombinačných čísel ďalšie dva z možných výsledkov pri hode 4 kockami – 1 šestka a 3 šestky (zvyšné dva prípady – 4 šestky a 0 šestiek prídu na rad v úlohách 16 a 17), zmeníme teda počet úspechov k, ale ponecháme počet pokusov $n = 4$,
- v úlohe 15 zmeníme počet pokusov n aj počet úspechov k.

Neskôr – v úlohách 18 a 19 zmeníme aj hodnotu pravdepodobnosti p.

ÚLOHY

14. Pomocou kombinačných čísel zapíšte pravdepodobnosti, že pri 4 hodoch kockou
a) padne práve 1 šestka, b) padnú práve 3 šestky.
Pri odvodzovaní výsledku použite postup z riešenia úloh 11 až 13.

15. Zopakujte úvahy z riešenia úloh 11 až 13 pri vyjadrení pravdepodobnosti, že
a) pri 7 hodoch kockou padnú práve 3 šestky,
b) pri 20 hodoch kockou padnú práve 4 šestky.

RIEŠENIE

Doterajšie úvahy o výpočte pravdepodobnosti počtu úspechov pri hode kockou môžeme zhrnúť: (skontrolujte, či je tento opis v súlade s vaším riešením úloh 14 a 15):

- Počet rôznych možností, v ktorých v k hodoch z celkového počtu n hodov padne šestka, vypočítame ako kombinačné číslo $\binom{n}{k}$.
- Všetky tieto možnosti majú rovnakú pravdepodobnosť: tou je súčin, v ktorom sa číslo $\frac{1}{6}$ objaví toľkokrát, koľkokrát padla šestka – teda k-krát a číslo $\frac{5}{6}$ toľkokrát, koľkokrát šestka nepadla – teda $(n - k)$-krát. Napríklad pravdepodobnosť, že šestka padne zo 7 hodov v 3., 4. a 7. hode, nájdeme ako súčin $\frac{5}{6} \cdot \frac{5}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{1}{6}$ (hodnota $\frac{1}{6}$ je na 3., 4. a 7. mieste), teda táto pravdepodobnosť je $\left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^4$.
- Pravdepodobnosť, že v n hodoch kockou padne šestka práve k-krát, je súčet týchto $\binom{n}{k}$ rovnakých pravdepodobností, preto jej hodnota je $\binom{n}{k} \cdot \left(\frac{1}{6}\right)^k \cdot \left(\frac{5}{6}\right)^{n-k}$.

Pravdepodobnosť, že v n hodoch kockou padne práve k šestiek, je $\binom{n}{k} \cdot \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{n-k}$.

ÚLOHA

16. Skontrolujte, že v tvare $\binom{n}{k} \cdot \left(\frac{1}{6}\right)^k \cdot \left(\frac{5}{6}\right)^{n-k}$, ktorý sme uviedli v závere predchádzajúceho textu, možno zapísať aj pravdepodobnosť, že
a) v 4 hodoch kockou padnú 4 šestky (v tomto prípade sa $n = 4$, $k = 4$),
b) v n hodoch kockou padne n šestiek.

Vzorec $\binom{n}{k} \cdot \left(\frac{1}{6}\right)^k \cdot \left(\frac{5}{6}\right)^{n-k}$ bude dôležitý v našich ďalších úvahách. Preto – hoci obidve pravdepodobnosti z úlohy 16 možno zapísať jednoduchším spôsobom – chceme skontrolovať, či všetky (aj tieto jednoduchšie zapísateľné) pravdepodobnosti možno vyjadriť jedným a tým istým vzorcom. Rovnaký cieľ bude mať aj úloha 17.

Malá odbočka

PASCALOV TROJUHOLNÍK A KOMBINAČNÉ ČÍSLA

Teraz sa môžeme vrátiť k vzťahu medzi Pascalovým trojuholníkom a kombinačnými číslami, ktorý sme naznačili v úvode tejto časti na s. 16. V tab. 5 porovnávame dve riešenia tej istej úlohy – nájsť pravdepodobnosti, že pri 4 hodoch kockou padne práve 0 šestiek, 1 šestka, 2 šestky, 3 šestky, 4 šestky:

- Výsledky v dolnom riadku tab. 5 sme našli pri riešení úloh 11 až 14 a 16 na s. 17 – 19.
- Výsledky uvedené o riadok vyššie sme našli v úlohe 9 (pozri tab. 4 na s. 16). Pripomeňme ešte, že čísla **1**, **4**, **6**, **4**, **1** v tomto riadku sú čísla v štvrtom riadku Pascalovho trojuholníka.

pravdepodobnosť, že v 4 hodoch hracou kockou padne práve				
0 šestiek	1 šestka	2 šestky	3 šestky	4 šestky
$\mathbf{1}\cdot\left(\frac{5}{6}\right)^4$	$\mathbf{4}\cdot\left(\frac{5}{6}\right)^3\cdot\frac{1}{6}$	$\mathbf{6}\cdot\left(\frac{5}{6}\right)^2\cdot\left(\frac{1}{6}\right)^2$	$\mathbf{4}\cdot\frac{5}{6}\cdot\left(\frac{1}{6}\right)^3$	$\mathbf{1}\cdot\left(\frac{1}{6}\right)^4$
	$\binom{\mathbf{4}}{\mathbf{1}}\cdot\left(\frac{5}{6}\right)^3\cdot\frac{1}{6}$	$\binom{\mathbf{4}}{\mathbf{2}}\cdot\left(\frac{5}{6}\right)^2\cdot\left(\frac{1}{6}\right)^2$	$\binom{\mathbf{4}}{\mathbf{3}}\cdot\frac{5}{6}\cdot\left(\frac{1}{6}\right)^3$	$\binom{\mathbf{4}}{\mathbf{4}}\cdot\left(\frac{1}{6}\right)^4$

Tab. 5

Z porovnania posledných dvoch riadkov tabuľky 5 vidno, že čísla **4, 6, 4** a **1** sa zhodujú s kombinančnými číslami $\binom{4}{1}$, $\binom{4}{2}$, $\binom{4}{3}$ a $\binom{4}{4}$. Zostáva ešte číslo **1** na začiatku horného riadka: aby sme aj to mohli zapísať v tvare kombinačného čísla, dohodneme sa, že symbol $\binom{4}{0}$ bude označovať číslo **1**.

Dohoda o zavedení symbolu $\binom{4}{0}$ nie je až taká formálna, ako by sa na prvý pohľad mohlo zdať. Kombinačné číslo $\binom{n}{k}$ je počet všetkých možností, ako z n rôznych prvkov vybrať skupinu obsahujúcu k prvkov. Podľa tohto opisu by číslo $\binom{4}{0}$ malo vyjadrovať počet možností, ako zo 4 rôznych prvkov vybrať 0 prvkov. Ak sa nad tým zamyslíme, tak uznáme, že sa to dá práve jedným spôsobom: všetky 4 prvky necháme na mieste. To je v súlade s ďalším poznatkom o kombinačných číslach: Každým výberom k-prvkovej skupiny z n prvkov je jednoznačne určená $(n-k)$-prvková skupina tých prvkov, ktoré sme nevybrali. Preto počet všetkých k-prvkových skupín musí byť rovnaký ako počet všetkých $(n-k)$-prvkových skupín. Zapísané pomocou kombinačných čísel: $\binom{n}{k}=\binom{n}{n-k}$. *Opticky* to znamená, že n-tý riadok Pascalovho trojuholníka je *súmerný podľa stredu* (skontroluj-te to). Dohoda $\binom{n}{0}=1$ zachováva túto súmernosť, pretože ako vieme, platí $\binom{n}{n}=1$.

Vďaka tejto dohode môžeme povedať, že štvrtý riadok Pascalovho trojuholníka tvoria kombinačné čísla $\binom{4}{0}$, $\binom{4}{1}$, $\binom{4}{2}$, $\binom{4}{3}$ a $\binom{4}{4}$. Rovnaké úvahy možno použiť pre ľubovoľný riadok v Pascalovom trojuholníku. Ak sa dohodneme, že prvú jednotku v n-tom riadku označíme symbolom $\binom{n}{0}$, tak n-tý riadok Pascalovho trojuholníka tvoria kombinačné čísla

$$\binom{n}{0},\ \binom{n}{1},\ \binom{n}{2},\ \ldots,\ \binom{n}{n-2},\ \binom{n}{n-1},\ \binom{n}{n}$$

Dohoda: $\binom{\mathbf{n}}{\mathbf{0}}=\mathbf{1}$

Aby sme na výpočet týchto nových kombinačných čísel $\binom{n}{0}$ mohli používať pôvodný vzorec $\binom{n}{k}=\frac{n!}{(n-k)!\,k!}$, musíme sa ešte dohodnúť, že symbol $0!$, ktorý sa v tomto vzorci pre $k=0$ vyskytne, budeme chápať ako číslo 1. Túto dohodu potrebujeme aj pri použití uvedeného vzorca na výpočet kombinačného čísla $\binom{n}{n}$, skontrolujte to!

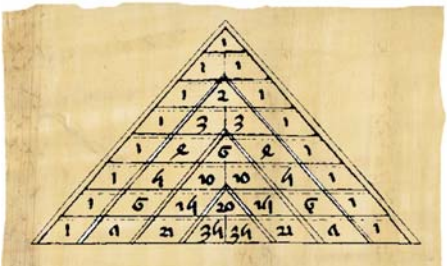
Dohoda: $\mathbf{0! = 1}$

Zvyčajne používaná podoba Pascalovho trojuholníka má aj nultý riadok, v ktorom je jedna jednotka (nasleduje prvý riadok s číslami 1, 1, druhý s číslami 1, 2, 1 atď.). Túto jednotku môžeme – v súlade so vzorcom $\binom{n}{k} = \dfrac{n!}{(n-k)!\,k!}$ a s dohodou, že $0! = 1$ – zapísať ako kombinačné číslo $\binom{0}{0}$ (skontrolujte to).

riadok

$$
\begin{array}{l}
0 \qquad \binom{0}{0} \\
1 \qquad \binom{1}{0} \quad \binom{1}{1} \\
2 \qquad \binom{2}{0} \quad \binom{2}{1} \quad \binom{2}{2} \\
3 \qquad \binom{3}{0} \quad \binom{3}{1} \quad \binom{3}{2} \quad \binom{3}{3} \\
4 \qquad \binom{4}{0} \quad \binom{4}{1} \quad \binom{4}{2} \quad \binom{4}{3} \quad \binom{4}{4} \\
\vdots \\
n \qquad \binom{n}{0} \quad \binom{n}{1} \quad \binom{n}{2} \quad \binom{n}{3} \quad \dots \quad \binom{n}{n-2} \quad \binom{n}{n-1} \quad \binom{n}{n} \\
\vdots
\end{array}
$$

Pascalov trojuholník zapísaný pomocou kombinačných čísel



„Pascalov“ trojuholník bol známy dávno pred Pascalom v Číne, Indii, Perzii aj Európe. Na obrázku je jeho vyobrazenie z r. 1407 v jednom z vydaní knihy Jordana de Nemore, európskeho matematika z 13. storočia. Za pozornosť stoja pre nás dnes nezvyčajné tvary arabských číslic.

To, že jednotlivé hodnoty v Pascalovom trojuholníku vyjadrujú počty kombinácií, bolo známe indickým matematikom už v 9. storočí. Nasledujúci kombinatorický dôkaz súčtovej vlastnosti Pascalovho trojuholníka však pochádza od Pascala.

Všetky $(k+1)$-prvkové kombinácie z $n+1$ prvkov – ktorých počet je $\binom{n+1}{k+1}$ – možno rozložiť na dve skupiny:

- Prvá obsahuje tie kombinácie, v ktorých sa vyskytuje nejaký vopred určený prvok. Týchto kombinácií je $\binom{n}{k}$, pretože k určenému prvku treba ešte pridať ďalších k prvkov (tie vyberáme celkovo z n prvkov, medzi ktorými zvolený prvok už nie je), aby sme dostali celkom $k+1$ prvkov.
- Druhá obsahuje kombinácie, ktoré tento daný prvok neobsahujú, tých je $\binom{n}{k+1}$, pretože všetkých $k+1$ prvkov treba vybrať spomedzi n prvkov, medzi ktorými nie je zvolený prvok.

Preto musí platiť $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$.

Čísla v Pascalovom trojuholníku majú ešte ďalšiu dôležitú vlastnosť. Ak sa pozrieme na pravé strany nasledujúcich rovností – dostali sme ich roznásobením výrazov na ich ľavej strane – zistíme, že koeficienty, ktoré sme vyznačili červeno, tvoria jednotlivé riadky Pascalovho trojuholníka:

$$(a+b)^2 = \mathbf{1} \cdot a^2 + \mathbf{2} \cdot ab + \mathbf{1} \cdot b^2$$
$$(a+b)^3 = \mathbf{1} \cdot a^3 + \mathbf{3} \cdot a^2b + \mathbf{3} \cdot ab^2 + \mathbf{1} \cdot b^3$$
$$(a+b)^4 = \mathbf{1} \cdot a^4 + \mathbf{4} \cdot a^3b + \mathbf{6} \cdot a^2b^2 + \mathbf{4} \cdot ab^3 + \mathbf{1} \cdot b^4$$

Čísla 1, 2, 1 predstavujú druhý riadok Pascalovho trojuholníka, čísla 1, 3, 3, 1 jeho tretí riadok, čísla 1, 4, 6, 4, 1 štvrtý riadok. Tento poznatok platí pre každé prirodzené číslo n: ak chceme zapísať výraz $(a+b)^n$ v roznásobenom tvare, použijeme ako koeficienty čísla z n-tého riadka Pascalovho trojuholníka. Napríklad jeho ôsmy riadok tvoria čísla

1 8 28 56 70 56 28 8 1

a výraz $(a+b)^8$ má po roznásobení podobu

$$(a+b)^8 = \mathbf{1} \cdot a^8 + \mathbf{8} \cdot a^7b + \mathbf{28} \cdot a^6b^2 + \mathbf{56} \cdot a^5b^3 + \mathbf{70} \cdot a^4b^4 + \mathbf{56} \cdot a^3b^5 + \mathbf{28} \cdot a^2b^6 + \mathbf{8} \cdot ab^7 + \mathbf{1} \cdot b^8$$

(všimnite si, že mocniny premennej a klesajú postupne od 8 k 0, naopak mocniny premennej b rastú od 0 po 8).

JACOB BERNOULLI
(1654 – 1705)
SA PRAVDEPODOBNOSŤOU POČTU ÚSPECHOV PRI VIACNÁSOBNOM OPAKOVANÍ POKUSU ZAOBERAL V SVOJOM DIELE *ARS CONJECTANDI (UMENIE USUDZOVAŤ)*. PRÁCA NADVÄZUJÚCA NA VÝSLEDKY ĎALŠÍCH VÝZNAČNÝCH MATEMATIKOV, KTORÍ STÁLI PRI ZRODE TEÓRIE PRAVDEPODOBNOSTI (HUYGENS, PASCAL, FERMAT) VYŠLA V R. 1713, OSEM ROKOV PO BERNOULLIHO SMRTI.

BERNOULLIOVSKÉ POKUSY A BINOMICKÉ ROZDELENIE PRAVDEPODOBNOSTI

Opakované hody mincou alebo hracou kockou, pri ktorých nás zaujíma počet úspechov (v našich príkladoch je úspechom v prípade mince padnutie znaku, v prípade kocky padnutie šestky) sú typické príklady **bernoulliovských pokusov**. Vo všeobecnosti bernoulliovským pokusom nazývame viacnásobné opakovanie jednoduchého pokusu, pričom v jednoduchom pokuse nás zaujíma iba jeden z jeho možných výsledkov, ktorý označujeme ako *úspech* (v príklade s hracou kockou nás z rôznych možných výsledkov, napr. *padne trojka*, *padne párne číslo*, *nepadne štvorka* a pod., zaujímal výsledok *padne šestka*, tento výsledok sme preto označili ako *úspech*).

V predchádzajúcich úlohách sme sa obmedzili na padnutie šestky pri hode kockou – teda na náhodný pokus, v ktorom pravdepodobnosť úspechu je $p = \frac{1}{6}$. Rovnaké úvahy však môžeme urobiť pre náhodný pokus s ľubovoľnou inou pravdepodobnosťou úspechu p. Vyskúšame si to v úlohe 18. Predtým dokončíme rozbor všetkých možných výsledkov pri opakovanom hode kockou.

V úlohe 16 sme overili, že vzorec $\binom{n}{k} \cdot \left(\frac{1}{6}\right)^k \cdot \left(\frac{5}{6}\right)^{n-k}$ možno použiť na zápis pravdepodobnosti jednej z extrémnych možností – *vo všetkých hodoch padne šestka*. Teraz nás bude zaujímať druhý extrém – *šestka nepadne ani raz*.

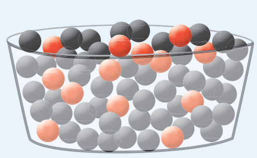
ÚLOHY

17. V článku o Pascalovom trojuholníku na s. 20 sme sa dohodli, že symbol $\binom{n}{0}$ bude označovať číslo 1. Skontrolujte, že vďaka tejto dohode možno pravdepodobnosť, že
a) v 4 hodoch kockou nepadne ani jedna šestka,
b) v n hodoch kockou nepadne ani jedna šestka
zapísať v tvare

$$\binom{n}{k} \cdot \left(\frac{1}{6}\right)^k \cdot \left(\frac{5}{6}\right)^{n-k}.$$

ZMYSEL TÝCHTO OTÁZOK SME VYSVETLILI V POZNÁMKE PRI ÚLOHE 16 NA S. 19: CHCEME SKONTROLOVAŤ, ČI VŠETKY PRAVDEPODOBNOSTI – AJ TIE, KTORÉ MOŽNO ZAPÍSAŤ JEDNODUCHŠÍM SPÔSOBOM – MOŽNO VYJADRIŤ JEDNÝM A TÝM ISTÝM VZORCOM.

18. V diskusii si ujasnite, že rovnako ako v úlohách 11 až 17 môžeme uvažovať aj o výsledkoch opakovania náhodného pokusu, v ktorom pravdepodobnosť úspechu je napr. $p = 0{,}237$. Riešte pre túto pravdepodobnosť p úlohy podobné predchádzajúcim: aká je pravdepodobnosť, že
a) zo 4 pokusov budú úspešné práve 2,
b) z 10 pokusov bude úspešných práve 7,
c) z n pokusov bude úspešných práve k,
d) z 20 pokusov bude úspešných 20,
e) z 8 pokusov nebude úspešný ani jeden,
f) z n pokusov nebude úspešný ani jeden.
Výsledky zapíšte pomocou kombinačných čísel, hodnoty v úlohách a), b), d), e) vypočítajte na kalkulačke.



Urnový model s vrátením

Pokus s pravdepodobnosťou $p = 0{,}237$ možno znázorniť pomocou tzv. urnového modelu. V urne je veľký počet guliek, z nich 23,7 % je červených, ostatné sú čierne (v urne je napr. 237 červených a 1 000 – 237 = 763 čiernych guliek).
Jednoduchý pokus je *vybratie jednej guľky*.

Ak chceme pokus opakovať, musíme vybratú guľku do urny vrátiť, aby sa pravdepodobnosť úspechu pri ďalšom pokuse nezmenila. Kvôli vráteniu guľky sa tento urnový model označuje ako *model s vrátením* (v poslednej kapitole sa stretneme aj s urnovými modelmi bez vrátenia).

Podobne si možno predstavovať aj pokusy s inou pravdepodobnosťou úspechu p.

ÚLOHA

19. Zapíšte pravdepodobnosti z úlohy 7 na s. 15 pomocou kombinačných čísel. Zapísané pravdepodobnosti vypočítajte a hodnoty porovnajte s výsledkami úlohy 7.

Úvahy, ktoré sme používali pri riešení úloh 11 až 19, sú stručne zhrnuté v nasledujúcej tabuľke – v prvom a druhom stĺpci pre porovnanie (a väčšiu názornosť) pre hod kockou, v treťom stĺpci vo všeobecnej podobe.

Hod kockou – pravdepodobnosť, že pri jednom hode padne šestka, je $\frac{1}{6}$		*Pokus, v ktorom pravdepodobnosť úspechu je* p
pravdepodobnosť, že pri **4** *hodoch kockou padnú práve* **2** *šestky*	*pravdepodobnosť, že pri* n *hodoch kockou padne práve* k *šestiek*	*pravdepodobnosť, že pri* n *opakovaniach pokusu bude práve* k *pokusov úspešných*
Existuje celkom $\binom{4}{2}$ možností, v ktorých **dvoch** zo **štyroch** hodov padne šestka.	Existuje celkom $\binom{n}{k}$ možností, v ktorých k z n hodov padne šestka.	Existuje celkom $\binom{n}{k}$ možností, ktorých k z n pokusov bude úspešných.
Pravdepodobnosť každej z týchto možností je $\left(\frac{1}{6}\right)^2 \cdot \left(1-\frac{1}{6}\right)^{4-2}$	Pravdepodobnosť každej z týchto možností je $\left(\frac{1}{6}\right)^k \cdot \left(1-\frac{1}{6}\right)^{n-k}$	Pravdepodobnosť každej z týchto možností je $p^k \cdot (1-p)^{n-k}$
Hľadaná pravdepodobnosť je súčet týchto $\binom{4}{2}$ rovnakých možností, preto sa rovná $\binom{4}{2} \cdot \left(\frac{1}{6}\right)^2 \cdot \left(1-\frac{1}{6}\right)^{4-2}$.	Hľadaná pravdepodobnosť je súčet týchto $\binom{n}{k}$ rovnakých možností, preto sa rovná $\binom{n}{k} \cdot \left(\frac{1}{6}\right)^k \cdot \left(1-\frac{1}{6}\right)^{n-k}$.	Hľadaná pravdepodobnosť je súčet týchto $\binom{n}{k}$ rovnakých možností, preto sa rovná $\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$.

Vyplýva z nich dôležitý poznatok, ktorý využijeme v našich nasledujúcich úvahách:

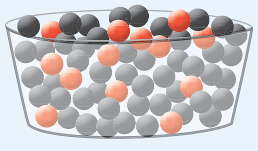
ak n-krát opakujeme pokus, v ktorom pravdepodobnosť úspechu je p, tak pravdepodobnosť, že z týchto n pokusov bude práve k úspešných, je $\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$. (*)

PRAVDEPODOBNOSŤ, ŽE PRI 19 HODOCH KOCKOU PADNE PRÁVE 8 JEDNOTIEK, JE $\binom{19}{8} \cdot \left(\frac{1}{6}\right)^8 \cdot \left(\frac{5}{6}\right)^{11}$ (PRAVDEPODOBNOSŤ, ŽE PRI JEDNOM HODE PADNE JEDNOTKA, JE $\frac{1}{6}$).

PRAVDEPODOBNOSŤ, ŽE PRI 12 HODOCH MINCOU PADNE ZNAK PRÁVE 7-KRÁT, JE $\binom{12}{7} \cdot 0{,}5^{12}$ (PRAVDEPODOBNOSŤ, ŽE PRI JEDNOM HODE PADNE ZNAK, JE 0,5).

Ak budeme 250-krát opakovať pokus, v ktorom pravdepodobnosť úspechu je $p = 0{,}4$, bude mať vzorec (*) tvar

$$\binom{250}{k} \cdot 0{,}4^k \cdot 0{,}6^{250-k}. \qquad (**)$$



Napríklad 250-krát za sebou vytiahneme guľku z urny, ktorá obsahuje 40 % červených a 60 % čiernych guliek, zistíme farbu vytiahnutej guľky a guľku vrátime do urny. Tento *zložený* náhodný pokus (vytvorený 250-násobným opakovaním *jednoduchého* pokusu, ktorým je vytiahnutie 1 guľky) môže mať 251 rôznych výsledkov – náhodných udalostí: *žiadna vytiahnutá guľka nebola červená, práve* 1 *vytiahnutá guľka bola červená, práve* 2 *vytiahnuté guľky boli červené, ..., práve* 249 *vytiahnutých guliek bolo červených, všetky vytiahnuté guľky boli červené.*

Ak do tohto vzorca dosadíme niektoré z čísel $k = 0, 1, 2, ..., 250$, dostaneme pravdepodobnosť, že pri 250 opakovaniach dosiahneme úspech práve k-krát. Na vzorec (**) sa teda môžeme pozerať ako na predpis funkcie, ktorá možným výsledkom náhodného pokusu (tie sú v tomto prípade opísané hodnotami $k = 0, 1, 2, ..., 250$) priraďuje ich pravdepodobnosti. Takéto funkcie sa v matematike nazývajú **rozdelenie pravdepodobnosti**. Predpis rozdelenia pravdepodobnosti (**) sme dostali tak, že v (*) sme zvolili konkrétne hodnoty $n = 250$ a $p = 0{,}4$. Každé rozdelenie pravdepodobnosti, ktoré takouto konkrétnou voľbou n a p dostaneme z predpisu $\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$, sa nazýva **binomické**. Teda pre každú konkrétnu dvojicu n a p (kde n je nejaké prirodzené číslo a p niektorá hodnota medzi 0 a 1) dostaneme jedno konkrétne binomické rozdelenie. Hodnoty n a p sa nazývajú *parametre* tohto rozdelenia. Napríklad rozdelenie určené vzorcom (**) je binomické rozdelenie pravdepodobnosti s parametrami $n = 250$, $p = 0{,}4$.

1.3 Ako to bolo s falošnou mincou

- Ktoré odchýlky od ideálneho výsledku sú pri 10 hodoch prakticky nemožné?
- Ako to bude pri iných počtoch hodov mincou
- Prechod k relatívnym početnostiam
- Hľadáme prakticky isté výsledky

Teraz sa môžeme vrátiť k otázke, ktorú sme vyslovili v úvode tejto kapitoly. Ak hádžeme 10-krát mincou, očakávame, že počet znakov, ktoré padnú, sa bude pohybovať okolo 5. Je nám pritom jasné, že nemusí padnúť presne 5 znakov. Ak sa napr. počet padnutých znakov od ideálnej hodnoty 5 odchýli nanajvýš o ± 1 (teda padne 4, 5 alebo 6 znakov), pokladáme to za normálne. Otázka je, aká veľká odchýlka od ideálnej hodnoty 5 už **nie je** normálna. Inak povedané: aká veľká odchýlka by už mala vzbudiť naše podozrenie, že minca je falošná (teda pravdepodobnosť padnutia znaku na nej **nie je** 0,5)?

ideálny výsledok = počet pokusov (n) × pravdepodobnosť úspechu (p)

KTORÉ ODCHÝLKY OD IDEÁLNEHO VÝSLEDKU SÚ PRI 10 HODOCH PRAKTICKY NEMOŽNÉ?

ÚLOHA

20. Pomocou tabuľkového kalkulátora
a) vytvorte tabuľku, v ktorej pre každú z hodnôt $k = 0, 1, ..., 10$ vypočítate pravdepodobnosť, že pri 10 hodoch štandardnou mincou padne znak práve k-krát,

b) potom v ďalšom stĺpci tabuľky vyjadrite v percentách pravdepodobnosti z bodu a),
c) zostrojte graf, znázorňujúci pre jednotlivé hodnoty $k = 0, 1, ..., 10$ príslušnú pravdepodobnosť (v %).
Vytvorenú tabuľku použijeme ešte v úlohách 21, 26.

TÁTO TABULKA OPISUJE BINOMICKÉ ROZDELENIE PRAVDEPODOBNOSTI PRE $n = 10$, $p = 0{,}4$ (POZRI DEFINÍCIU BINOMICKÉHO ROZDELENIA ZO S. 24).

RIEŠENIE

Z článku 1.2 vieme, že pravdepodobnosť, že pri 10 hodoch mincou padne práve k znakov, je $\binom{10}{k} \cdot 0{,}5^k \cdot (1 - 0{,}5)^{10-k}$, čo je to isté ako $\binom{10}{k} \cdot 0{,}5^{10}$ (skontrolujte, že obidva vzorce sú správne).
V tabuľkovom kalkulátore máme teda naprogramovať výpočet hodnôt funkcie $\binom{10}{k} \cdot 0{,}5^{10}$ pre $k = 0, 1, 2, ..., 10$.
Možný postup výpočtu sme znázornili na obrázku: Do bunky A1 vložíme hodnotu 0, do bunky A2 vzorec =A1 + 1 a rozšírime ho na bunky A3 až A11. Do bunky B1 vložíme vzorec $=\binom{10}{A1} \cdot 0{,}5^{10}$ a rozšírime ho na bunky B2 až B11. Do bunky C1 vložíme vzorec = 100 · B1 a rozšírime ho na bunky C2 až C11.
Hodnoty pravdepodobnosti z buniek C1 až C11 nájdete nad jednotlivými bodmi grafu na obr. 4 (s. 27). Skontrolujte, či sa zhodujú s vašimi výsledkami (porovnajte ich tiež s pravdepodobnosťami z riešenia úlohy 8 na s. 15).

<table>
<tr><th></th><th>A</th><th>B</th><th>C</th></tr>
<tr><td>1</td><td>0</td><td>$=\binom{10}{A1} \cdot 0{,}5^{10}$</td><td>= 100 · B1</td></tr>
<tr><td>2</td><td>= A1 + 1</td><td rowspan="7">vzorec rozšírime na bunky B2 – B11</td><td rowspan="7">vzorec rozšírime na bunky C2 – C11</td></tr>
<tr><td>3</td><td rowspan="5">vzorec rozšírime na bunky A3 – A11</td></tr>
<tr><td>4</td></tr>
<tr><td>⋮</td></tr>
<tr><td>10</td></tr>
<tr><td>11</td></tr>
</table>

ÚLOHA

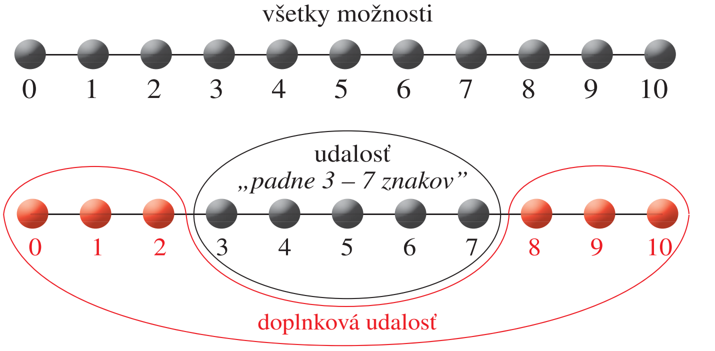
21. Použite tabuľku, ktorú ste vytvorili v úlohe 20 a) a pomocou tabuľkového kalkulátora vypočítajte pravdepodobnosť (v %), že pri 10 hodoch štandardnou mincou padne
a) 4 až 6 znakov (teda odchýlka od ideálneho stavu 5 znakov bude nanajvýš ± 1),
b) 3 až 7 znakov (odchýlka bude nanajvýš ± 2),
c) 2 až 8 znakov (odchýlka bude nanajvýš ± 3),
d) 1 až 9 znakov (odchýlka bude nanajvýš ± 4).
K jednotlivým pravdepodobnostiam vypočítajte doplnkové pravdepodobnosti v percentách (teda hodnotu 100 – pôvodná pravdepodobnosť v %). Opíšte udalosti, ktorým zodpovedajú tieto doplnkové pravdepodobnosti.

RIEŠENIE

Pravdepodobnosti a) až d) sú v prvých dvoch stĺpcoch tab. 6. Vyznačili sme ich tiež pod svorkami na obr. 4. Napríklad hodnotu **89,062 5 %** (vyjadrujúcu pravdepodobnosť, že pri 10 hodoch mincou padne 3 až 7 znakov, pozri vyznačený riadok v tab. 6) sme našli ako súčet pravdepodobností, že padne práve 3, práve 4, práve 5, práve 6 a práve 7 znakov:

11,718 75 + 20,507 812 5 + 24,609 375 + 20,507 812 5 + 11,718 75 = **89,062 5**,

teda v tabuľke z úlohy 20 sme sčítali pravdepodobnosti zodpovedajúce hodnotám $k = 3$ až $k = 7$.



Obr. 3
Pomocou ideálnych pokusov pripomeňme vzťah pravdepodobnosti a doplnkovej pravdepodobnosti:
3 – 7 *znakov padne v ideálnom prípade v* 89,0625 %
z veľkého počtu pokusov (jeden pokus = 10 *hodov mincou).*
V zvyšných pokusoch – tých je 100 – 89,0625 = 10,937 5 *percent – táto možnosť* ***nenastane****, teda padne iný počet znakov ako* 3 – 7.
Preto pravdepodobnosť doplnkovej udalosti (padne iný počet znakov ako 3 – 7*) je* 10,937 5 %.

Doplnková pravdepodobnosť je v tomto prípade 100 – 89,062 5 = 10,937 5 (%). Doplnková pravdepodobnosť je pravdepodobnosť toho, že **nenastane** daná možnosť. V našom prípade je to teda pravdepodobnosť, že pri 10 hodoch mincou bude počet znakov 0, 1, 2, 8, 9 alebo 10, t. j. odchýlka od ideálneho počtu 5 znakov bude väčšia ako ± 2.

pravdepodobnosť, že pri 10 *hodoch mincou padne*	%	*doplnková pravdepodobnosť*	= *pravdepodobnosť, že odchýlka od ideálnej hodnoty 5 bude väčšia ako*
4 až 6 znakov	65,625	34,375	± 1
3 až 7 znakov	89,062 5	10,937 5	± 2
2 až 8 znakov	97,851 562 5	2,148 437 5	± 3
1 až 9 znakov	99,804 687 5	0,195 312 5	± 4

Tab. 6

Všimnime si posledný riadok tab. 6. Pravdepodobnosť, že pri 10 hodoch štandardnou mincou padne 1 – 9 znakov, je 99,804 687 5 %. Udalosť s takouto pravdepodobnosťou – tak málo sa líšiacu od 100 % – už môžeme pokladať za prakticky istú. Je teda prakticky isté, že pri 10 hodoch štandardnou mincou bude počet znakov medzi 1 a 9. Inak povedané, je prakticky nemožné, aby pri 10 hodoch takouto mincou padlo 0 alebo 10 znakov (t. j. odchýlka väčšia ako ± 4 od ideálneho stavu 5 znakov je prakticky nemožná).

Možnosť *pri* 10 *hodoch mincou padne* 0 *alebo* 10 *znakov* je doplnková k možnosti *padne* 1 *až* 9 *znakov*, jej pravdepodobnosť (v %) je preto 100 – pravdepodobnosť možnosti *padne* 1 *až* 9 *znakov*, t. j.
100 – 99,804 687 5 = 0,195 312 5 (počítali sme ju v treťom stĺpci tab. 6), teda približne 0, 2 %.

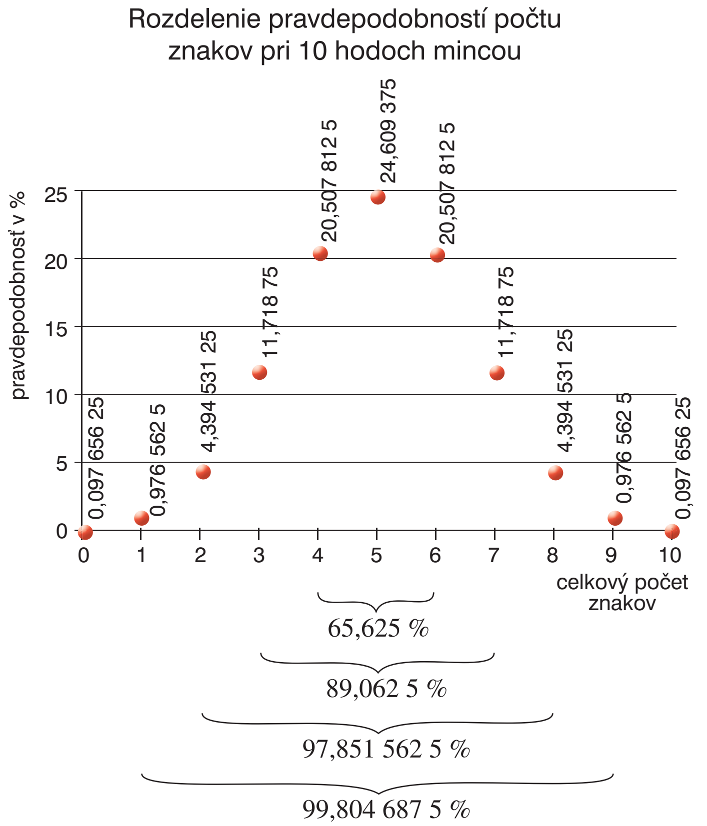
Ak pri 10 hodoch mincou padne 0 znakov, nastala situácia, ktorá je v prípade štandardnej mince prakticky nemožná (odchýlka od ideálnej hodnoty 5 je väčšia ako ± 4). Môžeme preto odôvodnene predpokladať, že naša minca je falošná (teda pravdepodobnosť padnutia znaku na nej **nie** je 0,5). Rovnako môžeme uvažovať, ak pri 10 hodoch mincou padlo 10 znakov.

Pri 10 hodoch mincou padlo 10 znakov. Taká veľká odchýlka od ideálnej hodnoty 5 znakov je v prípade štandardnej mince prakticky nemožná (jej pravdepodobnosť pokladáme za zanedbateľnú). Môžeme preto odôvodnene predpokladať, že minca je falošná.

ÚLOHA

22. Podrobne diskutujte o úvahách, ktorými sme dospeli k poznatku, že pri 10 hodoch mincou je prakticky nemožné, aby odchýlka od ideálneho počtu 5 znakov bola väčšia ako ± 4. Rovnako diskutujte o úvahe, na základe ktorej považujeme mincu za falošnú, ak v 10 hodoch padne 10 znakov.

Všimnite si, ako táto úvaha pripomína dôkaz nepravdivosti daného tvrdenia.
(Pripomeňme, že takýto dôkaz je súčasťou dôkazu sporom: ak dokazujeme sporom výrok C, tak tvrdenie, ktorého nepravdivosť dokazujeme je výrok $\neg C$).



Obr. 4
Graf binomického rozdelenia pravdepodobnosti pre $n = 10$, $p = 0,5$. *Pod svorkami sú uvedené pravdepodobnosti, že pri* 10 *hodoch mincou padne* 4 – 6, 3 – 7, 2 – 8, 1 – 9 *znakov.*

Schéma dôkazu nepravdivosti tvrdenia A	Naše úvahy pred úlohou 22
Keby platilo A,	Keby platilo A (teda minca je štandardná),
tak iste (teda so 100%-nou pravdepodobnosťou) nastane B.	tak skoro iste (teda s pravdepodobnosťou málo sa líšiacou od 100 %) nastane B (z 10 hodov padne znak 1- až 9-krát).
Ak B nenastalo, odvodzujeme z toho, že A nemôže byť pravdivé.	Ak B nenastalo (teda znak padol 0 alebo 10-krát), môžeme odôvodnene predpokladať, že A nebolo pravdivé (minca teda nie je štandardná).

Keby pri 10 hodoch mincou padlo napr. 8 znakov, odchýlka od ideálnej hodnoty je väčšia ako ± 2. Pravdepodobnosť takejto odchýlky je 100 – 89,062 5 = 10,937 5 (%), pozri vyznačený riadok v tab. 6. Táto pravdepodobnosť je príliš veľká na to, aby sme ju mohli pokladať za zanedbateľnú. Preto 8 znakov z 10 pokusov nie je pre nás dostatočný dôvod na to, aby sme mincu mohli považovať za falošnú.

ÚLOHA

23. Venujte dostatočný čas diskusii o tom, prečo 8 znakov pri 10 hodoch mincou nepokladáme ešte za dostatočný dôvod na to, aby sme mincu mohli považovať za falošnú.

V úvahe pred úlohou 22 sme za prakticky istú pokladali udalosť, ktorej pravdepodobnosť bola približne 99,8 %. Spravidla sa pri takýchto úvahách za prakticky isté pokladajú buď udalosti, ktorých pravdepodobnosť je väčšia ako 99 %, alebo – oveľa častejšie –

udalosti, ktorých pravdepodobnosť je väčšia ako 95 %. V našich ďalších úvahách budeme používať hodnotu 95 %. To znamená, že pravdepodobnosti menšie ako 5 % budeme už pokladať za zanedbateľné a udalosti, ktoré nastanú s pravdepodobnosťou menšou ako 5 %, za prakticky nemožné (toto si dobre rozmyslite).

Dohoda:
prakticky istá udalosť – s pravdepodobnosťou väčšou ako 95 %
prakticky nemožná udalosť – s pravdepodobnosťou menšou ako 5 %

ÚLOHA

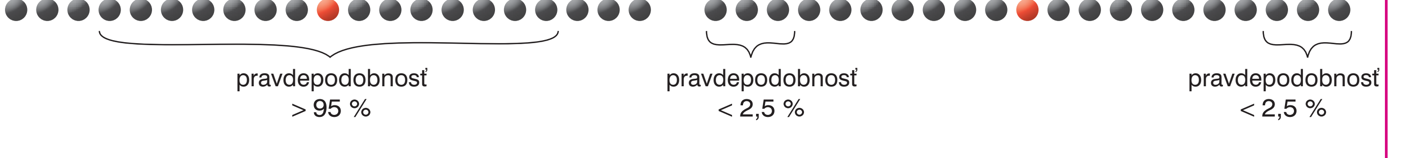
24. V predchádzajúcich úvahách sme za prakticky nemožnú považovali odchýlku väčšiu ako ± 4. Akú odchýlku budeme pokladať za prakticky nemožnú, ak použijeme našu novú dohodu a za prakticky istú budeme pokladať udalosť s pravdepodobnosťou väčšou ako 95% ?

AKO TO BUDE PRI INÝCH POČTOCH HODOV MINCOU

ÚLOHA

25. Zopakujte predchádzajúce úvahy pre a) 50, b) 200, c) 1 000 hodov mincov.
Teda: pomocou tabuľkového kalkulátora
- vytvorte tabuľku, v ktorej vypočítate pravdepodobnosti jednotlivých výsledkov hodov mincou (formálnejšie vyjadrené, máte vytvoriť tabuľku hodnôt binomického rozdelenia s parametrami a) $n = 50$ a $p = 0{,}5$, b) $n = 200$ a $p = 0{,}5$, c) $n = 1\,000$ a $p = 0{,}5$),
- zostrojte graf, ktorý podobne ako graf na obr. 4 znázorňuje rozdelenie pravdepodobností pri 50, 200 a 1 000 hodoch mincou,
- rozhodnite, ktoré odchýlky od ideálneho počtu znakov budeme pokladať za prakticky nemožné (ak uplatníme našu dohodu, že za prakticky istú pokladáme udalosť s pravdepodobnosťou väčšou ako 95 %).

Pri riešení zadania z poslednej odrážky odporúčame použiť postup opísaný v texte k obr. 5 b).
Vytvorené tabuľky použijeme ešte v úlohe 26.



Obr. 5

a) Postup, ktorý sme doteraz používali pri hľadaní najväčšej „ešte normálnej“ odchýlky od ideálnej hodnoty (napr. v úlohe 24) bol: sčítali sme „úseky pravdepodobností“ zodpovedajúcich postupne odchýlke nanajvýš ± 1, nanajvýš ± 2 atď. (pozri obr. 4) a zistili sme, kedy vypočítaný súčet pravdepodobností prvýkrát prevýši hodnotu 95 %. Ak je počet pokusov (teda číslo n) veľmi veľký, je tento postup – aj na tabuľkovom kalkulátore – zdĺhavý a je rozumné nahradiť ho postupom znázorneným na obr. b).

b) Hľadáme čo najdlhší „úsek pravdepodobností“ začínajúci na ľavom konci všetkých možností (teda pravdepodobnosťou, že padne 0 znakov) tak, aby jeho súčet bol ešte menší ako 2,5 %. Podobne hľadáme čo najdlhší úsek s celkovou pravdepodobnosťou menšou ako 2,5 % na pravom konci všetkých možností. Overte si správnosť tohto postupu najprv na obr. 4.

VÝPOČET SÚČTOV PRAVDEPODOBNOSTÍ V TAKÝCHTO „KRAJNÝCH" ÚSEKOCH MOŽNO NA TABUĽKOVOM KALKULÁTORE POMERNE ĽAHKO NAPROGRAMOVAŤ. HĽADANIE SA EŠTE ZJEDNODUŠÍ, AK TABUĽKOVÝ KALKULÁTOR AUTOMATICKY POČÍTA SÚČET HODNÔT VO VYZNAČENÝCH BUNKÁCH. VTEDY STAČÍ LEN ZVÄČŠOVAŤ DĹŽKU VYZNAČENÉHO ÚSEKU (NAPR. ŤAHOM MYŠOU) A SLEDOVAŤ, KEDY SÚČET PRAVDEPODOBNOSTÍ PRVÝKRÁT PREVÝŠI HODNOTU 2,5 %.

POZNAMENAJME EŠTE, ŽE V PRÍPADE HODOV MINCOU STAČÍ HĽADAŤ IBA ĽAVÝ Z TÝCHTO DVOCH ÚSEKOV, PRETOŽE PRAVDEPODOBNOSTI SÚ V TOMTO PRÍPADE „SÚMERNÉ" PODĽA STREDU (SKONTROLUJTE TO NA OBR. 4 A DISKUTUJTE O TOM, AKO TÁTO VLASTNOSŤ SÚVISÍ SO SÚMERNOSŤOU n-TÉHO RIADKA PASCALOVHO TROJUHOLNÍKA PODĽA STREDU, O KTOREJ SME HOVORILI V POZNÁMKE NA S. 20). PRETO PRAVÝ A ĽAVÝ ÚSEK MAJÚ ROVNAKÚ DĹŽKU. PRE NÁHODNÉ POKUSY S INOU PRAVDEPODOBNOSŤOU ÚSPECHU AKO $p = 0{,}5$ TO UŽ NEBUDE PLATIŤ A DĹŽKY NÁJDENÉHO PRAVÉHO A ĽAVÉHO KRAJNÉHO ÚSEKU BUDÚ ROZDIELNE.

RIEŠENIE

Výsledné grafy sú na obr. 6 až 8 na s. 30. V prípade **50 hodov** najdlhší ľavý úsek s pravdepodobnosťou ešte menšou ako 2,5 % je úsek od $k = 0$ po $k = 17$. Súčet pravdepodobností v tomto úseku – teda pravdepodobnosť, že pri 50 hodoch padne 0 – 17 znakov – je 1,64 ... % (súčet pravdepodobností pre hodnoty $k = 0$ až $k = 18$ je už 3,24... %). Najdlhší pravý úsek s pravdepodobnosťou ešte menšou ako 2,5 % je od $k = 33$ po $k = 50$. Preto pravdepodobnosť, že padne nanajvýš 17 alebo aspoň 33 znakov, je menšia ako 5 %, alebo – to isté inými slovami – pravdepodobnosť, že padne 18 – 32 znakov, je väčšia ako 95 %.

Teda: je prakticky isté, že pri 50 hodoch štandardnou mincou padne 18 – 32 znakov, t. j. odchýlka od ideálnej hodnoty (ktorá je v tomto prípade 25 znakov) bude nanajvýš ± 7.

Rovnakým postupom zistíme, že

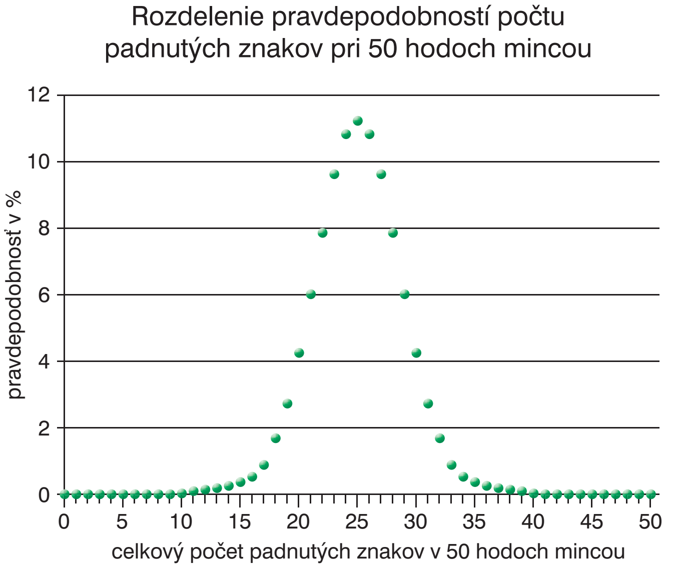
pri **200 hodoch** (ideálna hodnota je 100 znakov)	pri **1 000 hodoch** (ideálna hodnota je 500 znakov)
je prakticky isté, že padne	
86 – 114 znakov	469 – 531 znakov
t. j. odchýlka od ideálnej hodnoty je	
nanajvýš ± 14	nanajvýš ± 31
(súčet pravdepodobností v úseku $k = 0$ až $k = 85$ je 2,001... %, kým súčet pre $k = 0$ až $k = 86$ je už 2,798... %).	(súčet pravdepodobností v úseku $k = 0$ až $k = 468$ je 2,31... %, kým súčet pre $k = 0$ až $k = 469$ je už 2,68... %).

Ak teda napr. pri 200 hodoch mincou bude počet padnutých znakov mimo intervalu 86 až 114 znakov, možno odôvodnene predpokladať, že minca nie je štandardná. Keby pri 200 hodoch mincou padlo napr. 89 znakov, nebol by to dostatočný dôvod na tvrdenie, že minca nie je štandardná.

Poznámka.
Slovné spojenia prakticky istý *a* odôvodnene predpokladať *by ste v odbornej literatúre hľadali márne. Zvolili sme ich v snahe o čo najväčšiu názornosť. Štatistik by tvrdenia z predchádzajúceho odseku sformuloval takto:*

- *pri* 200 *hodoch mincou sa počet znakov nachádza mimo intervalu* 86 – 114, *preto* hypotézu minca je štandardná zamietame na hladine významnosti 5 %,
- *pri* 200 *hodoch mincou padlo* 89 *znakov, preto* hypotézu minca je štandardná nezamietame na hladine významnosti 5 %.

AKO OSTATNÉ OKULTNÉ TECHNIKY VEŠTENIA, AJ ŠTATISTIKA MÁ VLASTNÝ ŽARGÓN VYMYSLENÝ TAK, ABY ZAKRYL JEJ METÓDY PRED NEPROFESIONÁLOM.
(NEZNÁMY ŠTATISTIK)



Obr. 6

Binomické rozdelenie pravdepodobnosti pre n = 50 a p = 0,5.

Obrázok môže vzbudiť dojem, že pre k menšie ako 10 alebo väčšie ako 40 sú už hodnoty pravdepodobnosti nulové, nie je to však pravda. V skutočnosti znázornené hodnoty postupne rastú od

$$P(0) = 100 \cdot \binom{50}{0} \cdot 0{,}5^{50} = 8{,}881\,784 \ldots \cdot 10^{-16}\ \%$$

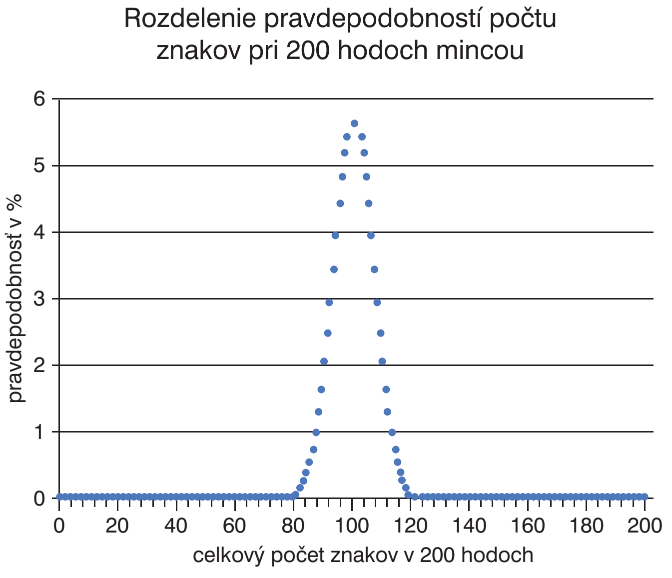
(teda nula, desatinná čiarka, za ňou 15 núl a až potom nasleduje prvá nenulová číslica 8) až po najväčšiu hodnotu

$$P(25) = 100 \cdot \binom{50}{25} \cdot 0{,}5^{50} = 11{,}227\,517 \ldots\ \%,$$

a potom zasa klesajú až po

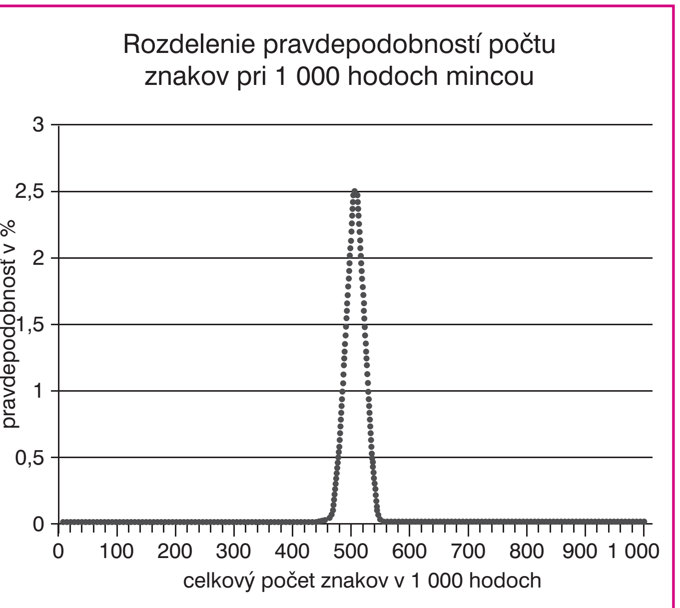
$$P(50) = 8{,}881\,784 \ldots \cdot 10^{-16}\ \%.$$

Rovnaká poznámka sa vzťahuje aj na grafy z obr. 7 a 8. (Činiteľ 100, ktorý sme vo vzorcoch pre $P(0)$ a $P(25)$ uviedli ako prvý v poradí, je tam preto, lebo pravdepodobnosti vyjadrujeme v percentách.)



Obr. 7

Binomické rozdelenie pravdepodobnosti pre n = 200 a p = 0,5.



Obr. 8

Binomické rozdelenie pravdepodobnosti pre n = 1 000 a p = 0,5.

Galton a quincunx (Galtonova doska)

Galton navrhol zariadenie znázorňujúce, ako opakovaním náhodného pokusu vzniká binomické rozdelenie pravdepodobnosti (pozri obrázok). Ide o dosku s niekoľkými radmi kolíkov. Tie sú rozmiestnené tak, aby guľka spustená zhora mala pri náraze do ľubovoľného kolíka rovnakú pravdepodobnosť pokračovať vpravo alebo vľavo od neho. Každý náraz guľky je teda náhodný pokus s pravdepodobnosťou úspechu ½ a počet radov kolíkov zodpovedá počtu opakovaní tohto pokusu. Po prejdení celej trasy guľka spadne do jedného z priečinkov umiestnených pod doskou. Ich šírka je zvolená tak, aby v nich guľky ležali jedna na druhej. Vďaka tomu zodpovedá výška guliek v jednotlivých priečinkoch počtu guliek, ktoré do priečinka spadli.

Ak je počet spustených guliek dostatočne veľký, budú výšky guliek v jednotlivých priečinkoch vytvárať približný histogram binomického rozdelenia pravdepodobnosti. Na internete možno nájsť množstvo simulácií tohto zariadenia, oveľa lepšie však je vyskúšať si ho v skutočnosti, nie iba virtuálne.

Názov quincunx označuje päť predmetov usporiadaných tak, ako sú usporiadané bodky v čísle 5 na hracej kocke. Túto podobu vytvára aj pätica susedných kolíkov Galtonovej dosky: jeden v strede, dva v rade nad ním a dva v rade pod ním.

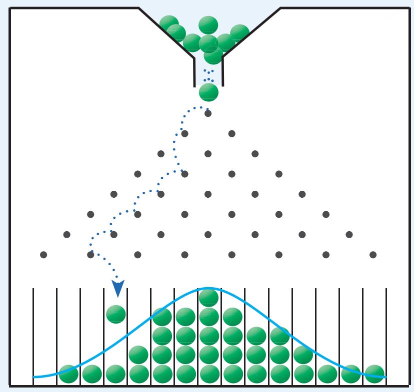


Schéma quincunxu (Galtonovej dosky).

PRECHOD K RELATÍVNYM POČETNOSTIAM

Ak sa pozrieme na obr. 4, 6, 7 a 8, všimneme si, že so zväčšovaním celkového počtu pokusov z $n = 10$ (obr. 4) cez $n = 50, 200$ (obr. 6, 7) až po $n = 1\,000$ (obr. 8) sa grafy „zužujú" (vidíte to aj vy?). Podobné pozorovanie o „zužovaní" platí aj pre interval, v ktorom ležia výsledky, ktoré ešte pokladáme za normálne (pozri riešenie úlohy 25 na s. 29): 2 – 8 znakov pri 10 hodoch, 18 – 32 pri 50 hodoch, 86 – 114 pri 200 hodoch, 469 – 531 pri 1 000 hodoch. Pre porovnanie:

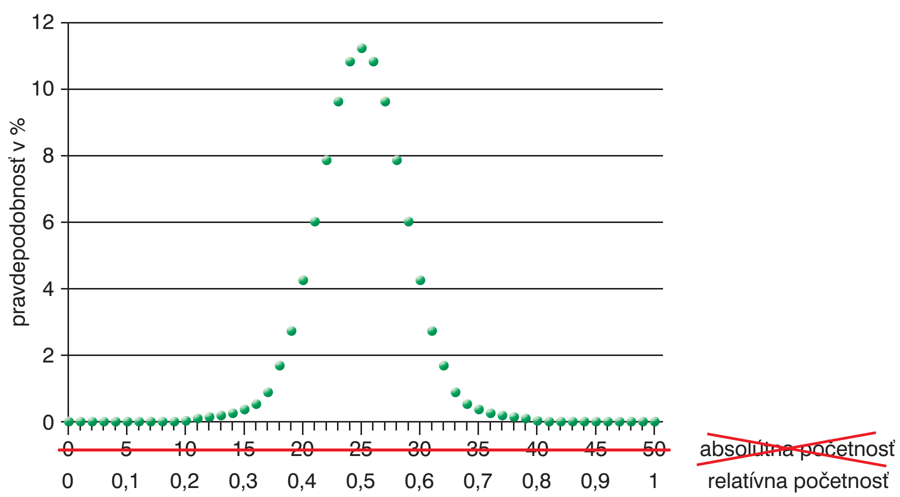
- na obr. 4 interval 2 – 8 predstavuje $\frac{6}{10} = 0{,}6$ dĺžky intervalu 0 – 10,
- na obr. 8 interval 469 – 531 predstavuje už iba $\frac{62}{1\,000} = 0{,}062$ dĺžky intervalu 0 – 1 000, teda skoro 10-krát menšiu časť ako na obr. 4.

Úvahy o „zužovaní" budú názornejšie, ak v grafoch nahradíme na vodorovnej osi absolútnu početnosť relatívnou (t. j. podielom $\frac{\text{počet znakov}}{\text{počet pokusov}}$, tento postup sme znázornili na obr. 9). Teda namiesto *počtu* padnutých znakov uvedieme, akú *časť* z celkového počtu pokusov tvoria tie pokusy, v ktorých padol znak (napr. v grafe na obr. 7 na vodorovnej osi hodnotu 140 nahradí hodnota $\frac{140}{200} = 0{,}7$, ideálny počet znakov vo všetkých štyroch

Sir Francis Galton

(1822 – 1911),

význačný anglický učenec, zaslúžil sa o rozvoj viacerých vedných disciplín (napr. psychológie, štatistiky alebo meteorológie). Svojimi výskumami položil, okrem iného, základ pre používanie odtlačkov prstov v kriminalistike.



Obr. 9

Graf rozdelenia pravdepodobnosti absolútnych *početností z obr. 6 prerábame na graf rozdelenia pravdepodobností* relatívnych *početností: na vodorovnej osi hodnotu „*počet znakov*" nahrádzame hodnotou* $\frac{\text{počet znakov}}{\text{počet pokusov}}$, *v tomto prípade je to* $\frac{\text{počet znakov}}{50}$.

grafoch nahradí hodnota 0,5, rozmyslite si to, až potom čítajte ďalej). Dosiahneme tak, že vo všetkých grafoch budú na vodorovnej osi čísla z intervalu od 0 do 1. Vďaka tomu môžeme takto získané grafy rozdelenia pravdepodobností relatívnych početností ľahko porovnávať.

ÚLOHA

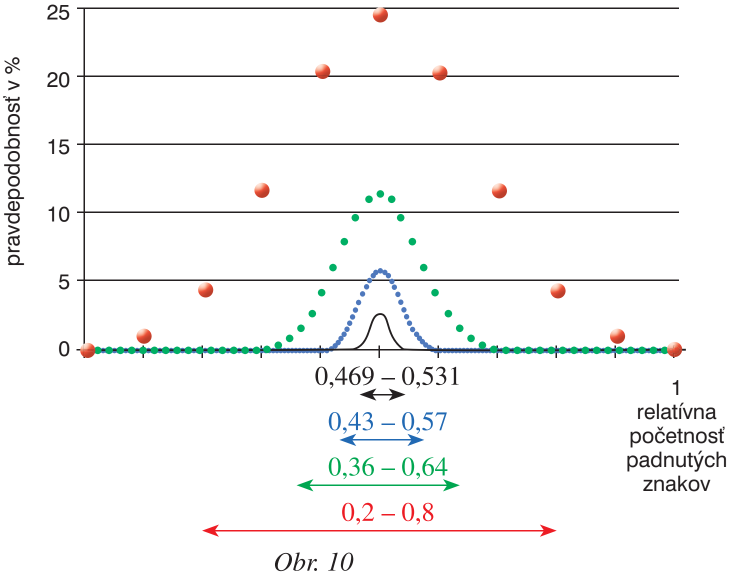
26. Vráťte sa k tabuľkám, ktoré ste pomocou tabuľkového kalkulátora vytvorili v úlohách 20 a 25.

a) Vytvorte pomocou nich grafy rozdelenia pravdepodobností (pre 10, 50, 200 a 1 000 hodov mincou), v ktorých na vodorovnej osi bude *relatívna* početnosť počtu znakov. (Teda: doplňte do tabuľky stĺpec *relatívna početnosť* a potom zostrojte graf, ktorý znázorní závislosť medzi stĺpcami *relatívna početnosť* a *pravdepodobnosť v %*.)

b) Pre každý z týchto štyroch prípadov (10, 50, 200 a 1000 hodov mincou) opíšte výsledok, ktorý pokladáme za prakticky istý, pomocou relatívnych početností.

RIEŠENIE

Na obr. 10 sme všetky štyri grafy nakreslili do jedného obrázka. Intervaly *prakticky istých výsledkov* vyjadrené pomocou relatívnych početností uvádzame v tabuľke na s. 33 aj pod obr. 10 (absolútne početnosti v intervaloch 2 – 8, 18 – 32, 86 – 114, 469 – 531 sme prepočítali na relatívne, teda vydelili sme ich celkovým počtom pokusov, napr. intervalu absolútnych početností 86 – 114 pri 200 pokusoch zodpovedá interval relatívnych početností $\frac{86}{200}$ – $\frac{114}{200}$, t. j. 0,43 – 0,57).



Obr. 10

*Porovnanie rozdelenia pravdepodobností z obr. 4, 6, 7 a 8 – na vodorovnej osi je relatívna početnosť počtu znakov. Farebné šípky znázorňujú interval, v ktorom pri danom počte pokusov prakticky iste leží relatívna početnosť znakov. Všimnite si, že s rastúcim počtom pokusov (od **n = 10** cez **n = 50** a **n = 200** po **n = 1 000**) sa krajné body tohto intervalu stále menej líšia od ideálnej relatívnej početnosti* 0,5 *(teda* 50 %*).*

pri 10 hodoch	pri 50 hodoch	pri 200 hodoch	pri 1 000 hodoch
štandardnou mincou je prakticky isté, že relatívna početnosť znakov bude v rozmedzí			
0,2 – 0,8	0,36 – 0,64	0,43 – 0,57	0,469 – 0,531
teda v percentách			
20 – 80 %	36 – 64 %	43 – 57 %	46,9 – 53,1 %

Pokladáme teda napr. za prakticky isté, že pri 200 hodoch štandardnou mincou padne znak v 43 – 57 percentách hodov. Pri relatívnej početnosti menšej ako 43 % alebo väčšej ako 57 % by sme už mohli odôvodnene predpokladať, že minca **nie** je štandardná (teda pravdepodobnosť padnutia znaku na nej **nie** je 0,5).

HĽADÁME PRAKTICKY ISTÉ VÝSLEDKY

Doteraz sme uvažovali o hádzaní mincou, teda o pokuse s pravdepodobnosťou $p = 0{,}5$. V nasledujúcich úlohách si podobné úvahy vyskúšame pre pokusy s inou pravdepodobnosťou úspechu.

Koľko správnych odpovedí sa dá „natipovať" v teste?

Test sa skladá z 30 otázok s výberom odpovede – v každej otázke si možno vybrať jednu zo 4 odpovedí, pričom iba jedna z nich je správna. Zaujíma nás, koľko správnych odpovedí by sme dosiahli, keby sme v každej otázke vybrali jednu zo 4 ponúkaných možností úplne náhodne. Takáto náhodná odpoveď nie je nič iné ako pokus s pravdepodobnosťou úspechu $p = 0{,}25$ (prečo?) a náhodné odpovedanie na 30 otázok je 30-násobné opakovanie tohto pokusu (ujasnite si to v diskusii).

ÚLOHA

27. a) Pomocou tabuľkového kalkulátora zostavte tabuľku rozdelenia pravdepodobností počtu správnych odpovedí. Zistite, aký počet správnych odpovedí možno pokladať za prakticky istý u študenta, ktorý na všetkých 30 otázok v teste odpovedal náhodne (použite postup opísaný na obr. 5 b na s. 28).
b) Ako sa zmení odpoveď, ak test bude mať 75 otázok (každú opäť s výberom zo 4 možností)?

Na základe výsledkov diskutujte o tom, aký počet správnych odpovedí by ste stanovili za minimálnu hranicu na úspešné absolvovanie testu. (Predpokladáme, že jediné kritérium na hodnotenie testu je počet správnych odpovedí, teda žiak pri hodnotení nie je napr. penalizovaný za počet nezodpovedaných alebo zle zodpovedaných otázok).

GRAFY ROZDELENÍ PRAVDEPODOBNOSTÍ SÚVISIACICH S RIEŠENÍM ÚLOHY 27 SÚ NA OBR. 11, S. 34.

Koľko šestiek by malo padnúť pri hode hracou kockou?

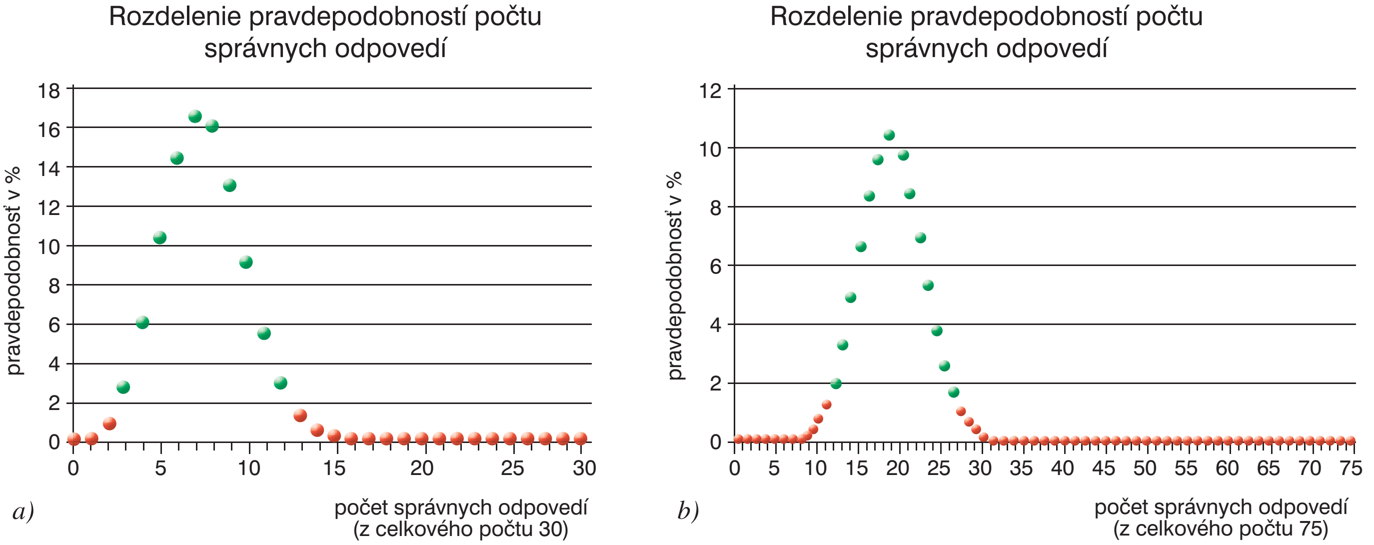
ÚLOHA

28. Zaujíma nás, aký počet šestiek možno pokladať za prakticky istý pri
a) $n = 60$, b) $n = 120$, c) $n = 600$
hodoch štandardnou hracou kockou.

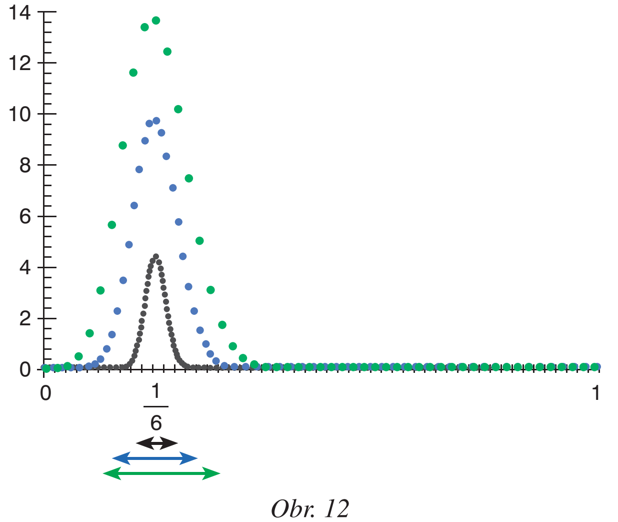
Pre uvedené počty hodov

- zostrojte pomocou tabuľkového kalkulátora tabuľku, v ktorej vypočítate pravdepodobnosť každého z možných výsledkov *padne* 0 *šestiek*, *padne* 1 *šestka*, ..., *padne n šestiek*,
- zostrojte graf rozdelenia pravdepodobností pre absolútnu aj pre relatívnu početnosť šestiek (v prvom budú na vodorovnej osi hodnoty $k = 0, 1, 2, ..., n$, v druhom relatívne početnosti $0, \frac{1}{n}, \frac{2}{n}, ..., 1$),
- pomocou postupu znázorneného na obr. 5 b) na s. 28 zistite, aký počet šestiek možno pri danom počte hodov pokladať za prakticky istý. Tento výsledok vyjadrite aj pomocou relatívnych početností.

Grafy rozdelení relatívnych početností súvisiacich s riešením úlohy 28 sú na obr. 12.



Obr. 11
Graf binomického rozdelenia pre $p = 0{,}25$ *a a)* $n = 30$, *b)* $n = 75$.
Zelenou sme vyznačili hodnoty nad intervalom, do ktorého prakticky iste padne počet správnych odpovedí pri náhodnom tipovaní.



Obr. 12
Rozdelenie pravdepodobností relatívnych početností šestiek pri 60, 120 *a* **600** *hodoch štandardnou hracou kockou. Farebné šípky označujú interval, v ktorom sa prakticky iste bude nachádzať relatívna početnosť šestiek pri danom počte hodov.*

Keby sme v každom z týchto troch grafov nahradili na vodorovnej osi relatívne početnosti absolútnymi (je to opačný postup k postupu znázornenému na obr. 9, hodnotu 1 *na vodorovnej osi by v zelenom grafe nahradilo číslo* 60, *v modrom* 120 *a v čiernom* 600*), dostali by sme grafy binomického rozdelenia pravdepodobnosti s parametrami*

$n = 60$ *a* $p = \frac{1}{6}$,

$n = 120$ *a* $p = \frac{1}{6}$,

$n = 600$ *a* $p = \frac{1}{6}$.

3. Po 5. hode:

0 znakov	1 znak	2 znaky	3 znaky	4 znaky	5 znakov
$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

Tab. 7

Po 6. hode:

0 znakov	1 znak	2 znaky	3 znaky	4 znaky	5 znakov	6 znakov
$\frac{1}{64}$	$\frac{6}{64}$	$\frac{15}{64}$	$\frac{20}{64}$	$\frac{15}{64}$	$\frac{6}{64}$	$\frac{1}{64}$

Tab. 8

Napríklad pravdepodobnosť, že pri 6 hodoch mincou padne znak práve dvakrát, je $\frac{15}{64} = 0{,}234\,375$, t. j. približne 23,44 % (3. stĺpec v tab. 8). Pravdepodobnosť, že pri 5 hodoch padne znak práve 3-krát, je $\frac{10}{32} = 0{,}312\,5$, t. j. 31,25 % (4. stĺpec v tab. 7).

4. Súčet čísel v každom riadku je 1.

Zdôvodnenie	
„v reči počtu pokusov“	*„v reči pravdepodobnosti“*
V každom kroku rozdeľujeme celok (predstavujúci všetky pokusy) na časti, ktoré sa navzájom neprekrývajú (napr. pri úvahách o 2. hode sú to 3 časti: „pokusy, v ktorých v prvých 2 hodoch nepadol ani 1 znak“, „pokusy, v ktorých v prvých 2 hodoch padol práve 1 znak“, „pokusy, v ktorých v prvých 2 hodoch *padli 2 znaky*“).	V každom riadku uvažujeme o náhodných udalostiach (napr. pri úvahách o 2. hode sú to 3 náhodné udalosti: *v prvých* 2 *hodoch nepadol ani* 1 *znak, v prvých* 2 *hodoch padol práve* 1 *znak, v prvých* 2 *hodoch padli* 2 *znaky*), ktorých súhrn je istá udalosť (pri úvahách o 2. hode je to udalosť v prvých 2 hodoch padne 0, alebo 1, alebo 2 znaky), pričom žiadne dve z týchto náhodných udalostí nemôžu nastať súčasne.
Čísla v príslušnom riadku tabuľky 2 určujú, akú časť celku (teda všetkých pokusov) tvoria tieto časti.	Čísla v príslušnom riadku tabuľky sú pravdepodobnosti týchto náhodných udalostí. Ak označíme náhodné udalosti v n-tom riadku A_0 (*v prvých n hodoch padlo* 0 *znakov*), A_1 (*v prvých n hodoch padol* 1 *znak*), A_2, ..., A_n, tak čísla v n-tom riadku sú hodnoty $p(A_0), p(A_1), \ldots, p(A_n)$ (*)
Vieme, že • časti sa navzájom neprekrývajú • a spolu tvoria celý celok (každý pokus patrí práve do jednej časti), preto sa súčet čísel v riadku rovná 1.	Vieme, že • žiadne dve z udalostí $A_0, A_1, A_2, \ldots, A_n$ nemôžu nastať súčasne, preto $p(A_0 \text{ alebo } A_1 \ldots \text{ alebo } A_n) = p(A_0) + p(A_1) + \ldots + p(A_n)$ (**) • A_0 alebo A_1 alebo ... alebo A_n (*po n hodoch padne* 0, *alebo* 1, *alebo* 2, ..., *alebo n znakov*) je istá udalosť, preto $p(A_0 \text{ alebo } A_1 \ldots \text{ alebo } A_n) = 1$ (***) Z (*), (**) a (***) vyplýva, že súčet čísel v n-tom riadku je 1.

5. V prvom riadku 2, v druhom 4 (t. j. 2^2), v treťom 8 (t. j. 2^3), v štvrtom 16 (t. j. 2^4), vo všeobecnosti v n-tom riadku to bude 2^n. Čísla v n-tom riadku tab. 2 majú všetky menovateľa 2^n a ich súčet je 1 (úloha 4), preto sa súčet čitateľov týchto čísel rovná 2^n.

6.

<table>
<tr><td>1</td><td>1</td></tr>
<tr><td>1</td><td>2</td><td>1</td></tr>
<tr><td>1</td><td>3</td><td>3</td><td>1</td></tr>
<tr><td>1</td><td>4</td><td>6</td><td>4</td><td>1</td></tr>
<tr><td>1</td><td>5</td><td>10</td><td>10</td><td>5</td><td>1</td></tr>
<tr><td>1</td><td>6</td><td>15</td><td>20</td><td>15</td><td>6</td><td>1</td></tr>
<tr><td>1</td><td>7</td><td>21</td><td>35</td><td>35</td><td>21</td><td>7</td><td>1</td></tr>
<tr><td>1</td><td>8</td><td>28</td><td>56</td><td>70</td><td>56</td><td>28</td><td>8</td><td>1</td></tr>
</table>

Tab. 9

Tento zápis s rastúcim počtom riadkov stráca prehľadnosť (najmä, keď chceme, aby bolo zrejmé, ktoré dve čísla z predchádzajúceho riadka máme sčítať). Preto ho nahradíme zápisom v tvare trojuholníka, pozri obrázok na s. 15 pred úlohou 8. Čísla v 5. riadku tab. 9 (1, 5, 10, 10, 5, 1) sú čitatele zlomkov z tab. 7 na s. 35 (menovateľ týchto zlomkov je $32 = 2^5$). Podobne čísla v 6. riadku tab. 9 sú čitatele zlomkov z tab. 8 (ich menovateľ je $64 = 2^6$).

7. Z riešenia úloh 1 a 2 vyplýva, že ak čísla v 6. (7., 8.) riadku tabuľky vydelíme číslom $64 = 2^6$ (číslom $128 = 2^7$, číslom $256 = 2^8$), dostaneme zlomky vyjadrujúce pravdepodobnosti, že pri 6 (7, 8) hodoch mincou padne práve 0 znakov, 1 znak atď.

a) $\frac{35}{128}$ (4. číslo v 7. riadku tabuľky sme vydelili číslom $128 = 2^7$).

V riešení úloh 7 b, c využijeme tvrdenie, ktoré sme si pripomenuli v riešení úlohy 4:

Ak A, B, C sú výsledky náhodného pokusu, z ktorých žiadne dva nemôžu nastať súčasne, tak

$$p(A \text{ alebo } B \text{ alebo } C) = p(A) + p(B) + p(C). \qquad (*)$$

b) $\frac{1}{256} + \frac{8}{256} + \frac{28}{256} = \frac{37}{256}$ (v tomto prípade sú *A, B, C* v tomto poradí udalosti *pri* 8 *hodoch mincou padne* 0 *znakov*, *pri* 8 *hodoch mincou padne* 1 *znak*, *pri* 8 *hodoch mincou padnú* 2 *znaky*),

c) $\frac{15}{64} + \frac{6}{64} + \frac{1}{64} = \frac{22}{64}$ (v tomto prípade sú *A, B, C* v tomto poradí udalosti *pri* 6 *hodoch mincou padnú* 4 *znaky*, *pri* 6 *hodoch mincou padne* 5 znakov, *pri* 6 *hodoch mincou padne* 6 *znakov*).

8. Jednotlivé pravdepodobnosti dostaneme, ak čísla v 10. riadku Pascalovho trojuholníka vydelíme hodnotou $2^{10} = 1\,024$:

pravdepodobnosť, že pri 10 hodoch štandardnou mincou padne										
0	1	2	3	4	5	6	7	8	9	10
znakov, je										
$\frac{1}{1\,024}$	$\frac{10}{1\,024}$	$\frac{45}{1\,024}$	$\frac{120}{1\,024}$	$\frac{210}{1\,024}$	$\frac{252}{1\,024}$	$\frac{210}{1\,024}$	$\frac{120}{1\,024}$	$\frac{45}{1\,024}$	$\frac{10}{1\,024}$	$\frac{1}{1\,024}$
= 0,000 976 562 5	= 0,009 765 625	= 0,043 945 312 5	= 0,117 187 5	= 0,205 078 125	= 0,246 093 75	= 0,205 078 125	= 0,117 187 5	= 0,043 945 312 5	= 0,009 765 625	= 0,000 976 562 5

(pravdepodobnosti sú zapísané ako čísla z intervalu $\langle 0,\ 1\rangle$, pravdepodobnosti vyjadrené v percentách dostaneme, ak výsledky uvedené v tabuľke vynásobíme 100).

9. V tab. 10 obsahujúcej výpočty sme – kvôli veľkosti vzorcov – nemohli dodržať správne pomery medzi šírkami jednotlivých obdĺžnikov. Čísla z prvých 4 riadkov Pascalovho trojuholníka sme zvýraznili červene. Výpočty medzi riadkami *po* 3. *hode* a *po* 4. *hode* sme vynechali (prenechávame ich na čitateľa).

<table>
<tr><td colspan="3">$\mathbf{1} \cdot \frac{5}{6}$ pokusov
0 šestiek</td><td colspan="3">$\mathbf{1} \cdot \frac{1}{6}$ pokusov
1 šestka</td><td>*po 1. hode*</td></tr>
<tr><td colspan="2">$\frac{5}{6} \cdot \frac{5}{6} = \left(\frac{5}{6}\right)^2$
+ 0</td><td>$\frac{5}{6} \cdot \frac{1}{6}$
+ 1</td><td>$\frac{1}{6} \cdot \frac{5}{6}$
+ 0</td><td colspan="2">$\frac{1}{6} \cdot \frac{1}{6} = \left(\frac{1}{6}\right)^2$
+ 1</td><td></td></tr>
<tr><td colspan="2">$\mathbf{1} \cdot \left(\frac{5}{6}\right)^2$
0 šestiek</td><td colspan="2">$\mathbf{2} \cdot \frac{5}{6} \cdot \frac{1}{6}$
1 šestka</td><td colspan="2">$\mathbf{1} \cdot \left(\frac{1}{6}\right)^2$
2 šestky</td><td>*po 2. hode*</td></tr>
<tr><td>$\left(\frac{5}{6}\right)^2 \cdot \frac{5}{6} = \left(\frac{5}{6}\right)^3$
+ 0</td><td>$\left(\frac{5}{6}\right)^2 \cdot \frac{1}{6}$
+ 1</td><td>$\left(2 \cdot \frac{5}{6} \cdot \frac{1}{6}\right) \cdot \frac{5}{6} =$
$= 2 \cdot \left(\frac{5}{6}\right)^2 \cdot \frac{1}{6}$
+ 0</td><td>$\left(2 \cdot \frac{5}{6} \cdot \frac{1}{6}\right) \cdot \frac{1}{6} =$
$= 2 \cdot \frac{5}{6} \cdot \left(\frac{1}{6}\right)^2$
+ 1</td><td>$\left(\frac{1}{6}\right)^2 \cdot \frac{5}{6}$
+ 0</td><td>$\left(\frac{1}{6}\right)^2 \cdot \frac{1}{6} = \left(\frac{1}{6}\right)^3$
+ 1</td><td></td></tr>
<tr><td>$\mathbf{1} \cdot \left(\frac{5}{6}\right)^3$
0 šestiek</td><td colspan="2">$\mathbf{3} \cdot \left(\frac{5}{6}\right)^2 \cdot \frac{1}{6}$
1 šestka</td><td colspan="2">$\mathbf{3} \cdot \frac{5}{6} \cdot \left(\frac{1}{6}\right)^2$
2 šestky</td><td>$\mathbf{1} \cdot \left(\frac{1}{6}\right)^3$
3 šestky</td><td>*po 3. hode*</td></tr>
<tr><td>$\mathbf{1} \cdot \left(\frac{5}{6}\right)^4$
0 šestiek</td><td>$\mathbf{4} \cdot \left(\frac{5}{6}\right)^3 \cdot \frac{1}{6}$
1 šestka</td><td colspan="2">$\mathbf{6} \cdot \left(\frac{5}{6}\right)^2 \cdot \left(\frac{1}{6}\right)^2$
2 šestky</td><td>$\mathbf{4} \cdot \frac{5}{6} \cdot \left(\frac{1}{6}\right)^3$
3 šestky</td><td>$\mathbf{1} \cdot \left(\frac{1}{6}\right)^4$
4 šestky</td><td>*po 4. hode*</td></tr>
</table>

Tab. 10

10. Spomínanú zákonitosť možno sformulovať takto: n-tý riadok tabuľky z riešenia úlohy 9 (teda riadok označený *po n-tom hode*) dostaneme, ak jednotlivé čísla v n-tom riadku Pascalovho trojuholníka násobíme postupne číslami $\left(\frac{5}{6}\right)^n$ (dostaneme pravdepodobnosť, že pri n hodoch nepadne žiadny znak), $\left(\frac{5}{6}\right)^{n-1} \cdot \left(\frac{1}{6}\right)$ (dostaneme pravdepodobnosť, že pri n hodoch padne práve 1 znak), $\left(\frac{5}{6}\right)^{n-2} \cdot \left(\frac{1}{6}\right)^2$ (dostaneme pravdepodobnosť, že pri n hodoch padnú práve 2 znaky) atď.

a) $66 \cdot \left(\frac{5}{6}\right)^{10} \cdot \left(\frac{1}{6}\right)^2 = 0{,}296\,09\ ...$, t. j. približne 29,61 %,

b) $66 \cdot \left(\frac{5}{6}\right)^{2} \cdot \left(\frac{1}{6}\right)^{10} = 0{,}000\,000\,757\ ...$, t. j. približne 8 stotisícin percenta (udalosť s pravdepodobnosťou 0,000 08 %, t. j. 0,000 000 8 by pri 10 miliónoch pokusov nastala v ideálnom prípade v 8 prípadoch, to je pri jednom z 1 250 000 pokusov).

12. v 1. a 2. hode, v 1. a 3. hode, v 2. a 4. hode, v 3. a 4. hode.

13. Pravdepodobnosť je pre každý z uvedených prípadov $\left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^2$.

14. a) $\binom{4}{1} \cdot \frac{1}{6} \cdot \left(\frac{5}{6}\right)^3$. V 4 hodoch kockou môže 1 šestka padnúť celkom $\binom{4}{1} = 4$ rôznymi spôsobmi (v 1. hode, v 2. hode, v 3. hode, v 4. hode – teda vyberáme 1 hod zo štyroch). Každý z nich má rovnakú pravdepodobnosť $\frac{1}{6} \cdot \left(\frac{5}{6}\right)^3$ (a žiadne dva z týchto spôsobov nemôžu nastať súčasne) . Preto pravdepodobnosť, že v 4 hodoch kockou padne práve 1 šestka, je $\binom{4}{1} \cdot \frac{1}{6} \cdot \left(\frac{5}{6}\right)^3$. **b)** $\underbrace{\binom{4}{3}}_{=4} \cdot \left(\frac{1}{6}\right)^3 \cdot \frac{5}{6}$.

15. a) $\binom{7}{3} \cdot \left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^4 = 0{,}078\,142\ ...$, t. j. približne 7,8 %. Tie 3 zo 7 hodov, v ktorých padne šestka, môžeme vybrať celkom $\binom{7}{3} = \frac{7 \cdot 6 \cdot 5}{1 \cdot 2 \cdot 3} = 35$ spôsobmi. Každá z týchto 35 možností má rovnakú pravdepodobnosť $\left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^4$.

b) $\binom{20}{4} \cdot \left(\frac{1}{6}\right)^4 \cdot \left(\frac{5}{6}\right)^{16}$.

16. a) Z riešenia úlohy 9 vieme, že táto pravdepodobnosť je $\left(\frac{1}{6}\right)^4$, pozri tiež tab. 4 na s. 16. Môžeme ju nájsť aj použitím úvah sformulovaných pred úlohou 16, tým zodpovedá zápis $\binom{4}{4} \cdot \left(\frac{1}{6}\right)^4$, kde kombinačné číslo $\binom{4}{4} = 1$ určuje počet všetkých rôznych možností, ako zo 4 hodov vybrať tie, v ktorých padne šestka – v tomto prípade je taká možnosť zrejme jediná, číslo $\left(\frac{1}{6}\right)^4$ je pravdepodobnosť tejto jedinej možnosti. Máme skontrolovať, či tento výsledok možno zapísať v tvare $\binom{4}{4} \cdot \left(\frac{1}{6}\right)^4 \cdot \left(\frac{5}{6}\right)^0$. To je pravda, pretože symbol $\left(\frac{5}{6}\right)^0$ označuje číslo 1 (dohodu, že pre každé kladné číslo q platí $q^0 = 1$, sme urobili pri zavádzaní mocnín, skúste si spomenúť, prečo). Teda: $\left(\frac{1}{6}\right)^4 = \underbrace{\binom{4}{4}}_{=1} \cdot \left(\frac{1}{6}\right)^4 \cdot \underbrace{\left(\frac{5}{6}\right)^0}_{=1}$.

b) Podobne ako v riešení úlohy 16a) platí $\left(\frac{1}{6}\right)^n = \underbrace{\binom{n}{n}}_{=1} \cdot \left(\frac{1}{6}\right)^n \cdot \underbrace{\left(\frac{5}{6}\right)^0}_{=1}$.

17. a) Táto pravdepodobnosť je $\left(\frac{5}{6}\right)^4$ (skontrolujte to). Máme overiť rovnosť $\left(\frac{5}{6}\right)^4 = \binom{4}{0} \cdot \left(\frac{1}{6}\right)^0 \cdot \left(\frac{5}{6}\right)^4$. Tá platí, pretože $\binom{4}{0} = 1$ a $\left(\frac{1}{6}\right)^0 = 1$.

b) $\left(\frac{5}{6}\right)^n = \underbrace{\binom{n}{0}}_{=1} \cdot \underbrace{\left(\frac{1}{6}\right)^0}_{=1} \cdot \left(\frac{5}{6}\right)^n$.

18. a) $\binom{4}{2} \cdot 0{,}237^2 \cdot 0{,}763^2 = 0{,}196\,199\,103\,366$,

b) $\binom{10}{7} \cdot 0{,}237^7 \cdot 0{,}763^3 = 0{,}002\,238\ ...$,

c) $\binom{n}{k} \cdot 0{,}237^k \cdot 0{,}763^{n-k}$,

d) $\binom{20}{20} \cdot 0{,}237^{20} \cdot 0{,}763^0 = 0{,}237^{20} = 3{,}125\ ... \cdot 10^{-13} = 0{,}000\,000\,000\,000\,312\,5\ ...$,

e) $\binom{8}{0} \cdot 0{,}237^0 \cdot 0{,}763^8 = 0{,}763^8 = 0{,}114\,867\ ...$,

f) $\binom{n}{0} \cdot 0{,}237^0 \cdot 0{,}763^n = 0{,}763^n$.

19. a) $\binom{7}{3} \cdot \left(\frac{1}{2}\right)^3 \cdot \left(1 - \frac{1}{2}\right)^4 = \underbrace{\binom{7}{3}}_{=35} \cdot \underbrace{\left(\frac{1}{2}\right)^7}_{=\frac{1}{128}} = 0{,}273\,437\,5$,

t. j. približne 27,3 % (hodnoty pod svorkami uvádzame kvôli porovnaniu s riešením úlohy 7 a).

b) $\binom{8}{0} \cdot \left(\frac{1}{2}\right)^0 \cdot \left(1 - \frac{1}{2}\right)^8 + \binom{8}{1} \cdot \left(\frac{1}{2}\right)^1 \cdot \left(1 - \frac{1}{2}\right)^7 + \binom{8}{0} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(1 - \frac{1}{2}\right)^6 = \binom{8}{0} \cdot \left(\frac{1}{2}\right)^8 + \binom{8}{1} \cdot \left(\frac{1}{2}\right)^8 + \binom{8}{2} \cdot \left(\frac{1}{2}\right)^8 = \underbrace{\left(\frac{1}{2}\right)^8}_{=\frac{1}{256}} \left[\underbrace{\binom{8}{0}}_{=1} + \underbrace{\binom{8}{1}}_{=8} + \underbrace{\binom{8}{2}}_{=28}\right] =$
$= 0{,}144\,531\,25$, t. j. približne 14,5 %,

c) $\binom{6}{4} \cdot \left(\frac{1}{2}\right)^6 + \binom{6}{5} \cdot \left(\frac{1}{2}\right)^6 + \binom{6}{6} \cdot \left(\frac{1}{2}\right)^6 = \underbrace{\left(\frac{1}{2}\right)^6}_{=\frac{1}{64}} \cdot \left[\underbrace{\binom{6}{4}}_{=15} + \underbrace{\binom{6}{5}}_{=6} + \underbrace{\binom{6}{6}}_{=1}\right] = 0{,}343\,75$, t. j. približne 34,4 %.

24. Odchýlku väčšiu ako ± 3 znaky. Z riešenia úlohy 21 vieme, že pravdepodobnosť odchýlky od ideálnej hodnoty 5 prekročí hodnotu 95 % prvýkrát pri odchýlke do ± 3 (pozri tab. 6 na s. 26 a obr. 4 na s. 27). Teda pravdepodobnosť, že odchýlka neprekročí ± 2 znaky, je ešte menšia ako 95 %, ale pravdepodobnosť, že odchýlka neprekročí ± 3 znaky (t. j. padne 2 – 8 znakov), je už viac ako 95 %.

27. a) Tabuľku vytvoríme podobne ako v riešení úlohy 20, pozri obrázok.

<table>
<tr><th></th><th>A
(počet správnych odpovedí)</th><th>B
(pravdepodobnosť ako hodnota medzi 0 a 1)</th><th>C
(pravdepodobnosť v %)</th></tr>
<tr><td>1</td><td>0</td><td>$= \binom{30}{A1} \cdot 0{,}25^{A1} \cdot 0{,}75^{(30 - A1)}$</td><td>= 100 · B1</td></tr>
<tr><td>2</td><td>= A1 + 1</td><td rowspan="6">vzorec rozšírime na bunky B2 – B31</td><td rowspan="6">vzorec rozšírime na bunky C2 – C31</td></tr>
<tr><td>3</td><td rowspan="5">vzorec rozšírime na bunky A3 – A31</td></tr>
<tr><td>...</td></tr>
<tr><td>29</td></tr>
<tr><td>30</td></tr>
<tr><td>31</td></tr>
</table>

Ak študent odpovie na všetky otázky náhodne, tak prakticky iste dosiahne 3 – 12 správnych odpovedí (teda 10 až 40 percent).

b) Ak študent odpovie na všetkých 75 otázok náhodne, tak prakticky iste dosiahne 12 – 26 správnych odpovedí (teda 16 až približne 35 percent).

28. Je prakticky isté, že padne **a)** 5 – 16 šestiek ($8\frac{1}{3}$ až $26\frac{2}{3}$ %, t. j. približne 8,33 až 26,67 %), **b)** 12 – 28 šestiek (10 až $23\frac{1}{3}$ %, t. j. približne 10 až 23,33 %), **c)** 82 – 118 šestiek ($13\frac{2}{3}$ až $19\frac{2}{3}$ %, t. j. približne 13,67 – 19,67 %). *Ideálna* relatívna frekvencia vyjadrená v percentách je $16\frac{2}{3}$ %, t. j. približne 16,67 %.

2 DVA UŽITOČNÉ VZORCE

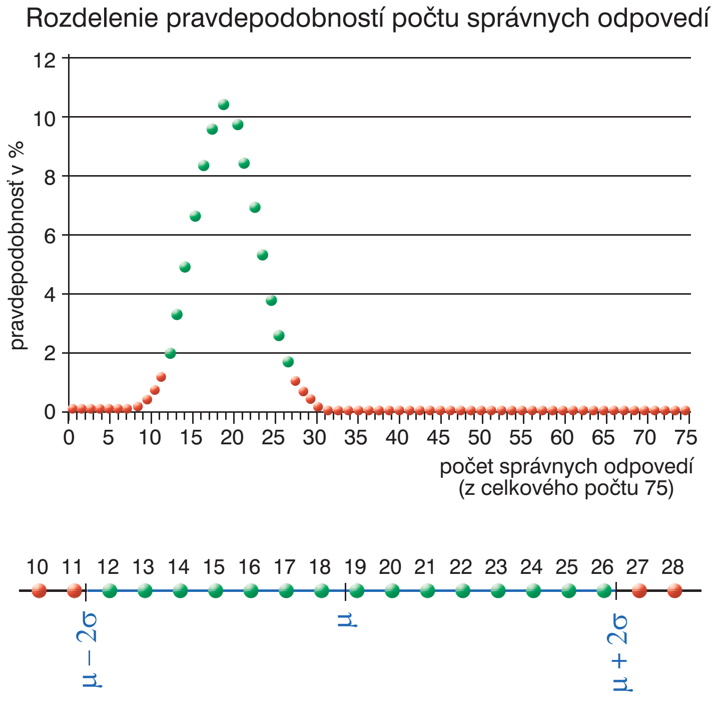
Vo viacerých úlohách z predchádzajúcej kapitoly bolo dôležité zistiť, aký výsledok (t. j. počet úspechov v mnohokrát opakovanom pokuse) možno pokladať za *prakticky istý*. Tabuľkový kalkulátor, ktorý sme používali, dokáže takéto výpočty urýchliť, aj ten má však svoje obmedzenia – nájsť postupom z predchádzajúcej kapitoly interval, v ktorom sa nachádzajú prakticky isté výsledky napr. pri 1 000 000 hodoch kockou, bude asi nad jeho sily. Našťastie (keďže úvahy o prakticky istých výsledkoch budú pre nás dôležité) na určenie intervalu, v ktorom sa nachádzajú takéto výsledky, existuje vzorec. Ten bude hlavnou náplňou prvej časti tejto kapitoly.

Nasledovať bude krátka nepovinná vsuvka – rozprávanie o krivkách normálneho rozdelenia, s ktorými binomické rozdelenie pravdepodobnosti úzko súvisí.

Potom – v druhej časti kapitoly – zmeníme uhol pohľadu. V predchádzajúcich úvahách sme za známu pokladali pravdepodobnosť úspechu pri jednom pokuse a zaujímalo nás, koľko úspechov môžeme *prakticky iste* pri viacnásobnom opakovaní tohto pokusu dosiahnuť. Teraz otázku obrátime: pravdepodobnosť úspechu pri jednom pokuse **ne**budeme poznať, budeme vedieť len počet úspechov, ktoré sme dosiahli pri n-násobnom opakovaní tohto pokusu. Otázka bude, ako presne možno na základe toho určiť neznámu pravdepodobnosť úspechu. Najprv si ujasníme, čo chápeme pod *presnosťou odhadu* pravdepodobnosti. Potom už príde na rad druhý zo vzorcov spomínaný v názve kapitoly – jednoduchý vzorec na odhad tejto presnosti.

2.1 Vzorec na určenie *prakticky istých* výsledkov

Všeobecne	*Napríklad*
Už vieme, že výsledky $\boldsymbol{n}$-krát opakovaného pokusu, ktorý má pravdepodobnosť úspechu $\boldsymbol{p}$, opisuje binomické rozdelenie pravdepodobnosti. To každému z čísel $k = 0$, $k = 1$, $k = 2$, ..., $k = n$ priradí pravdepodobnosť p_k, že pri n opakovaniach pokusu dosiahneme úspech presne k-krát. Ak chceme zistiť, v akom rozmedzí sa pri n opakovaniach pokusu bude *prakticky iste* nachádzať počet úspechov, môžeme postupovať takto:	V úlohe 27 b) na s. 33 sme sa stretli s binomickým rozdelením pravdepodobnosti s parametrami $n = \mathbf{75}$ a $p = \mathbf{0{,}25}$ (p bola pravdepodobnosť uhádnuť správnu odpoveď na jednu otázku, n bol počet otázok). Jeho predpis je $p_k = \binom{75}{k} \cdot 0{,}25^k \cdot (1 - 0{,}25)^{75-k}$ (p_k je pravdepodobnosť – vyjadrená ako číslo z intervalu $\langle 0, 1\rangle$ – že zo 75 otázok uhádneme práve k). Graf tohto rozdelenia je na obr. 1 a).
Vypočítame hodnoty	Pre toto rozdelenie pravdepodobnosti sa
$\mu = \boldsymbol{n} \cdot \boldsymbol{p}$	$\mu = \mathbf{75} \cdot \mathbf{0{,}25} = 18{,}75$,
$\sigma = \sqrt{\boldsymbol{n} \cdot \boldsymbol{p} \cdot (1 - \boldsymbol{p})}$	$\sigma = \sqrt{\mathbf{75} \cdot \mathbf{0{,}25} \cdot (1 - \mathbf{0{,}25})} = \sqrt{14{,}062\,5} = 3{,}75$
Potom – ak n je veľké číslo (teda, ak pokus opakujeme veľakrát) – platí: pravdepodobnosť, že počet úspešných pokusov je z intervalu $(\mu - 2\sigma, \mu + 2\sigma)$, je väčšia ako 95 %. To znamená, že ***prakticky isté* výsledky sú v intervale** $(\mu - 2\sigma, \mu + 2\sigma)$.	Pozrime sa na interval $(\mu - 2\sigma; \mu + 2\sigma) = (18{,}75 - 2 \cdot 3{,}75; 18{,}75 + 2 \cdot 3{,}75) = (11{,}25; 26{,}25)$. Z riešenia úlohy 27 b) vieme, že v tomto prípade počet uhádnutých odpovedí je prakticky iste 12 – 26. Ako vidno, intervalom $(11{,}25; 26{,}25)$ sú *prakticky isté* výsledky určené úplne presne. Sú to všetky celočíselné hodnoty z intervalu $(11{,}25; 26{,}25)$, pozri obr. 1 b).



Obr. 1

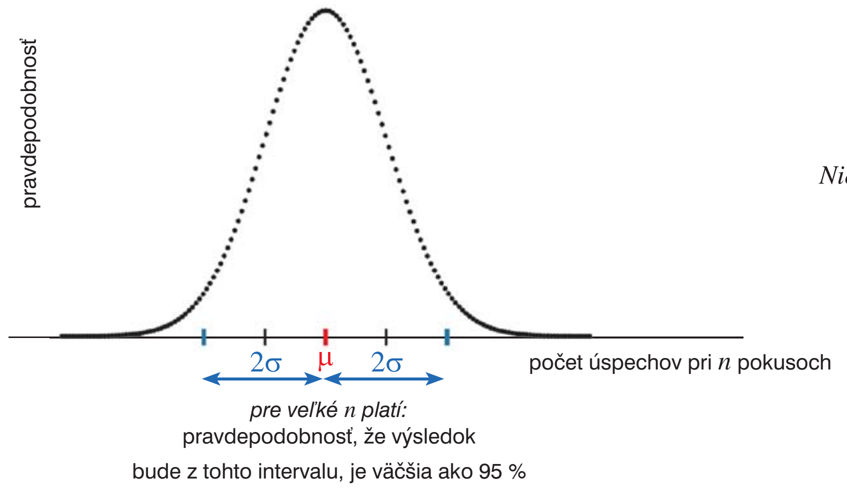
a) Binomické rozdelenie pravdepodobnosti s parametrami $n = 75$, $p = 0{,}25$ (pripomeňme, že v úlohe 27 b) z kapitoly 1.3 bolo číslo n počet otázok v teste a p bola pravdepodobnosť uhádnuť správnu odpoveď na jednu otázku). Zelenou sme vyznačili hodnoty nad intervalom, v ktorom ležia prakticky isté *výsledky (tie sme našli v riešení úlohy 27 b) na s. 39).*

b) Interval $(\mu - 2\sigma,\ \mu + 2\sigma)$ *(modrá úsečka) a* prakticky isté *výsledky (zelené bodky) pre binomické rozdelenie z obr. 1a).*

Môžeme tiež povedať: je prakticky isté, že pri veľkom počte pokusov n sa počet úspešných pokusov neodchýli od hodnoty μ o viac ako ± 2 σ.

Ak n (počet opakovaní pokusu) je veľké číslo, tak *prakticky isté* výsledky sú z intervalu $(\mu - 2\sigma,\ \mu + 2\sigma)$

$$\mu = n \cdot p, \quad \sigma = \sqrt{n \cdot p \cdot (1 - p)}$$



Obr. 2

Niektoré vlastnosti binomického rozdelenia pravdepodobnosti.

$\mu = n \cdot p$
$\sigma = \sqrt{n \cdot p \cdot (1 - p)}$
n – celkový počet pokusov
p – pravdepodobnosť úspechu pri 1 pokuse (číslo z intervalu $\langle 0, 1 \rangle$)

1. Všimnite si, že hodnota $\mu = n \cdot p$ je totožná s *ideálnou hodnotou*, ktorú sme už viackrát použili v našich úvahách (pozri napr. zvýraznený text pred úlohou 20 na s. 24).
2. Čísla μ a σ možno opísať pomocou štatistických pojmov, s ktorými sme sa stretli v učebnici pre 3. ročník. Na binomické rozdelenie pravdepodobnosti sa môžeme pozerať ako na súbor, v ktorom pravdepodobnosť p_k chápeme ako relatívnu početnosť hodnoty k. Dá sa ukázať, že tento súbor má strednú hodnotu (teda aritmetický priemer) $\mu = n \cdot p$ a rozptyl $\sigma = \sqrt{n \cdot p \cdot (1 - p)}$.

3. DÔKAZ TVRDENIA O INTERVALE $(\mu - 2\sigma,\ \mu + 2\sigma)$ AKO INTERVALE *PRAKTICKY ISTÝCH VÝSLEDKOV* JE UŽ ZA HRANICAMI STREDOŠKOLSKEJ MATEMATIKY. VYUŽÍVA SA V ŇOM POZNATOK, ŽE PRE VEĽKÉ HODNOTY n – TEDA PRE VEĽKÝ POČET POKUSOV – MOŽNO GRAF BINOMICKÉHO ROZDELENIA DOSTATOČNE PRESNE NAHRADIŤ NIEKTOROU Z KRIVIEK NORMÁLNEHO ROZDELENIA (PODROBNEJŠIE SA O TOM ZMIENIME V NASLEDUJÚCOM NEPOVINNOM DODATKU, POZRI OBR. 6 A OBR. 7 NA S. 45). VĎAKA TOMU SA NIEKTORÉ VLASTNOSTI KRIVIEK NORMÁLNEHO ROZDELENIA – JEDNOU Z NICH JE PRÁVE POLOHA INTERVALU *PRAKTICKY ISTÝCH VÝSLEDKOV* – PRENESÚ NA BINOMICKÉ ROZDELENIE.

Nebudeme dokazovať tvrdenie o intervale $(\mu - 2\sigma,\ \mu + 2\sigma)$ ako intervale *prakticky istých výsledkov* (pozri tiež text predchádzajúcej poznámky). Obmedzíme sa na to, že sa presvedčíme o jeho správnosti na binomických rozdeleniach z predchádzajúcej kapitoly.

ÚLOHY

1. Pre binomické rozdelenia pravdepodobností z úloh 20, 25, 27 a 28 z predchádzajúcej kapitoly vypočítajte hodnoty $\mu = n \cdot p$, $\sigma = \sqrt{n \cdot p \cdot (1 - p)}$ a určte hranice intervalu $(\mu - 2\sigma,\ \mu + 2\sigma)$. Potom skontrolujte, nakoľko čísla $\mu - 2\sigma$, $\mu + 2\sigma$ vystihujú polohu *intervalu prakticky istých výsledkov* (tieto intervaly ste našli v riešeniach úloh 24, 25, 27 a 28 z predchádzajúcej kapitoly). Pre binomické rozdelenie z úlohy 27 b) sme takéto porovnanie naznačili v pravom stĺpci tabuľky na s. 40.

RIEŠENIE

Ako dobre interval $(\mu - 2\sigma;\ \mu + 2\sigma)$ vystihuje polohu *intervalu prakticky istých výsledkov*, vidno z nasledujúcej tabuľky. (Naliehavo odporúčame, aby ste výpočty urobili sami a výsledky porovnali s číslami v tabuľke, pomôže vám to zvyknúť si na poznatok o súvise čísel μ a σ s intervalom *prakticky istých* výsledkov.)

binomické rozdelenie s parametrami n	p	$\mu = n \cdot p$	$\sigma = \sqrt{n \cdot p \cdot (1-p)}$	$(\mu - 2\sigma,\ \mu + 2\sigma)$	interval prakticky istých výsledkov	číslo úlohy z predchádzajúcej kapitoly
10	0,5	5	1,58...	(1,83...; 8,16...)	2 – 8	24
50	0,5	25	3,53...	(17,92...; 32,07...)	18 – 32	25
200	0,5	100	7,07...	(85,85...; 114,14...)	86 – 114	25
1 000	0,5	500	15,81...	(468,37...; 531,62...)	469 – 531	25
30	0,25	7,5	2,37...	(2,75...; 12,24...)	3 – 12	27
75	0,25	18,75	3,75	(11,25...; 26,25...)	12 – 26	27
60	$\frac{1}{6}$	10	2,88...	(4,22...; 15,77...)	5 – 16	28
120	$\frac{1}{6}$	20	4,08...	(11,83...; 28,16...)	12 – 28	28
600	$\frac{1}{6}$	100	9,12...	(81,74...; 118,25...)	82 – 118	28

Vo všetkých prípadoch okrem jednej výnimky (podfarbené bunky) je situácia rovnaká ako na obr. 1 b): *interval prakticky istých výsledkov* tvoria všetky prirodzené čísla ležiace v intervale $(\mu - 2\sigma,\ \mu + 2\sigma)$. V prípade podfarbených buniek je rozdiel iba v hodnote 16, tá už leží mimo intervalu $(\mu - 2\sigma,\ \mu + 2\sigma)$. Všimnite si, že pre rovnakú pravdepodobnosť $\left(p = \frac{1}{6}\right)$ a väčšie hodnoty n (nasledujúce dva riadky v tabuľke) už určenie *prakticky istých výsledkov* pomocou intervalu $(\mu - 2\sigma,\ \mu + 2\sigma)$ funguje bezchybne. To je v súlade s tým, že uvedené tvrdenie platí pre veľké hodnoty n.

Poznamenajme ešte, že to, ktoré hodnoty n sú už *dostatočne veľké*, závisí od hodnoty pravdepodobnosti p. Vo všeobecnosti platí: čím viac sa p odlišuje od hodnoty 0,5, tým väčšie musí byť n, aby interval $(\mu - 2\sigma,\ \mu + 2\sigma)$ dobre vystihol polohu *prakticky istých* výsledkov.

V PRAXI SA NAMIESTO INTERVALU $(\mu - 2\sigma,\ \mu + 2\sigma)$ POUŽÍVA *PRESNEJŠÍ* INTERVAL $(\mu - 1{,}96 \cdot \sigma,\ \mu + 1{,}96 \cdot \sigma)$. INTERVAL $(\mu - 2\sigma,\ \mu + 2\sigma)$ SA VŠAK ĽAHŠIE PAMÄTÁ A PRE NAŠE ÚVAHY JE POSTAČUJÚCI.

ÚLOHY

2. Z urny, ktorá obsahuje 45 % bielych a 55 % čiernych guliek, sme 200-krát náhodne vybrali jednu guľku, zistili sme jej farbu a guľku sme vrátili do urny. Koľko bielych guliek možno *prakticky iste* očakávať medzi týmito 200 vytiahnutými guľkami?

3. Z urny, v ktorej malo byť 70 % bielych a 30 % čiernych guliek, sme 150-krát náhodne vybrali jednu guľku, zistili sme jej farbu a opäť ju vrátili do urny. Z týchto 150 guliek bolo 32 čiernych. Je tento výsledok v súlade s informáciou o počte čiernych guliek v urne?

V riešení úloh 26 b), 27 a 28 z kapitoly 1.3 sme informáciu o *prakticky istých* výsledkoch sformulovali aj pomocou *relatívnych* početností, napr.

formulácia o počte výsledkov	*formulácia o* relatívnom *počte výsledkov*
Pri 200 hodoch mincou je prakticky isté, že	
znak padne 86- až 114-krát (pozri tretí riadok v tabuľke na s. 42).	relatívna početnosť znakov bude od 0,43 do 0,57 (teda znak padne v 43 % – 57 % všetkých hodov).

ÚLOHA

4. a) Pripomeňte si, ako sme vypočítali relatívne početnosti 0,43 a 0,57.

b) Sformulujte takto pomocou relatívnych početností nasledujúce tvrdenie o polohe *prakticky istých* výsledkov z obr. 2 (s. 41)

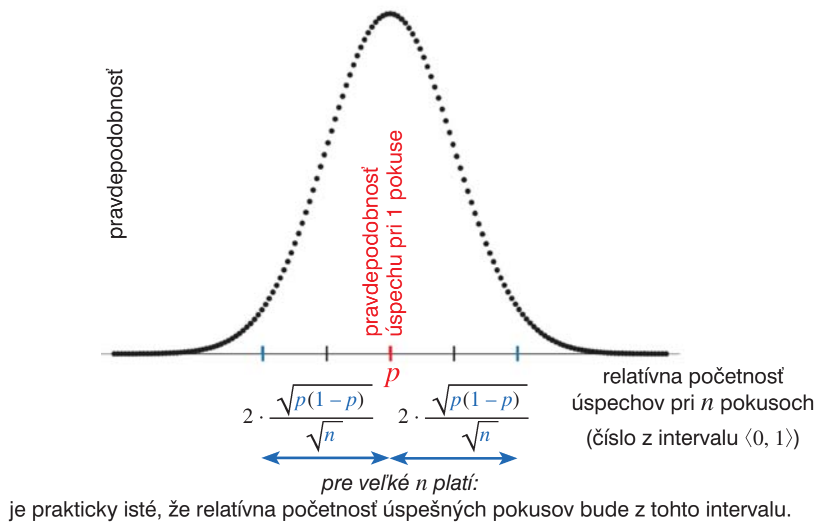
Ak n je veľké číslo, tak *prakticky isté* výsledky ležia v intervale

$(\mu - 2\sigma,\ \mu + 2\sigma)$, t. j. $\left(np - 2\sqrt{n}\ \sqrt{p(1-p)},\ np + 2\sqrt{n}\ \sqrt{p(1-p)}\right)$.

c) Doplňte opis výsledku úlohy b):

Ak n je veľké číslo, tak je *prakticky isté*, že relatívna početnosť úspešných pokusov sa od pravdepodobnosti p líši o menej ako ....

Relatívna početnosť *prakticky istých* výsledkov sa od pravdepodobnosti p líši o menej ako $\pm 2 \cdot \dfrac{\sqrt{p(1-p)}}{\sqrt{n}}$.



Obr. 3

Graf rozdelenia pravdepodobností relatívnych početností dostaneme z obr. 2 postupom, ktorý sme znázornili na obr. 9 (s. 31) – každú hodnotu na vodorovnej osi vydelíme celkovým počtom pokusov, teda číslom n.

V našich ďalších úvahách už budeme za prakticky isté výsledky automaticky pokladať všetky celočíselné hodnoty ležiace v intervale $(\mu - 2\sigma,\ \mu + 2\sigma)$.

Dohoda:

prakticky isté výsledky pri n-násobnom opakovaní pokusu (pre veľkú hodnotu n) = = všetky celočíselné hodnoty z intervalu $(\mu - 2\sigma,\ \mu + 2\sigma)$

ÚLOHA

5. Keby všetci obyvatelia Slovenska – vrátane dojčiat a starcov – naraz hodili hracou kockou (každý svojou a každý štandardnou), akú relatívnu početnosť šestiek v tomto celoslovenskom pokuse možno pokladať za prakticky istú? Ako by takýto pokus dopadol v Číne? Výsledky uvádzajte ako čísla z intervalu $\langle 0, 1 \rangle$ zaokrúhlené na štyri desatinné miesta, potom ich prerátajte na percentá.

Nepovinný dodatok

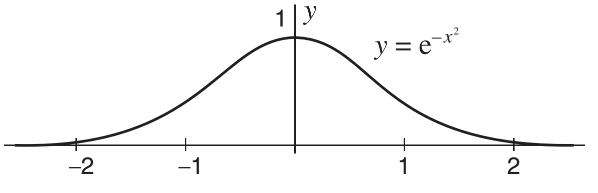
OD BERNOULLIOVSKÝCH POKUSOV KU KRIVKÁM NORMÁLNEHO ROZDELENIA

Každý verí normálnemu rozdeleniu: experimentátori si myslia, že ho dokázali matematici a matematici veria, že je výsledkom pozorovaní.

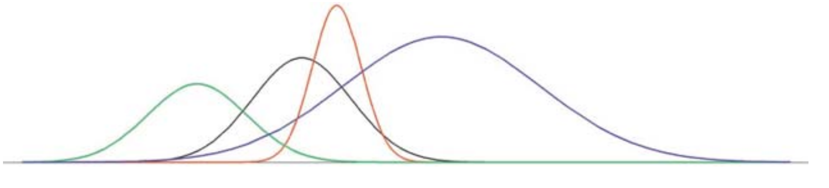
E. T. WHITTAKER, G. ROBINSON

Pri skúmaní bernoulliovských pokusov matematici narazili na krivky, ktoré mali neskôr hrať v štatistike významnú úlohu – krivky normálneho rozdelenia. Ich modelovým predstaviteľom je graf funkcie $y = e^{-x^2}$, pozri obr. 4.

Ľubovoľnú z kriviek normálneho rozdelenia dostaneme, ak tento graf stlačíme alebo roztiahneme v smere niektorej z osí x alebo y a prípadne ho ešte posunieme pozdĺž osi x. Príklady takýchto kriviek sú na obr. 5. Všeobecná podoba predpisu takto porozťahovanej a poposúvanej funkcie je $y = A \cdot e^{-(Bx + C)^2}$, kde konštanty A a B určujú veľkosť roztiahnutia alebo stlačenia grafu v smere osí y a x a číslo C súvisí s veľkosťou posunutia grafu pozdĺž osi x.



Obr. 4
Graf funkcie $y = e^{-x^2}$.

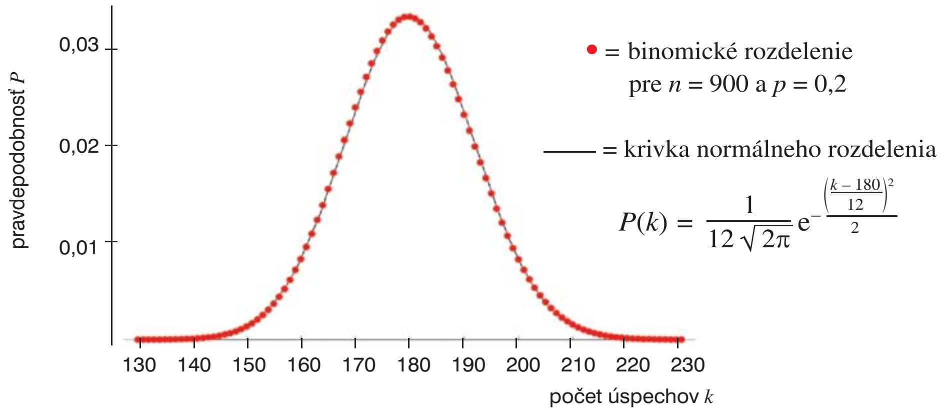


Obr. 5
Rôzne krivky normálneho rozdelenia. Podľa ich tvaru sa pre ne – najmä v anglicky písanej literatúre – vžil názov bell curve alebo bell-shaped curve (krivka tvaru zvona).

Pre veľké n možno hodnoty binomického rozdelenia dostatočne presne nahradiť hodnotami funkcie, ktorej grafom je niektorá z kriviek normálneho rozdelenia. Napríklad hodnoty binomického rozdelenia s parametrami $n = 900$ a $p = 0{,}2$ (to opisuje pravdepodobnosť výsledkov 900-násobného opakovania pokusu, v ktorom pravdepodobnosť úspechu je $p = 0{,}2$) možno veľmi dobre nahradiť hodnotami funkcie

$$P(k) = \frac{1}{12\sqrt{2\pi}} e^{-\frac{\left(\frac{k-180}{12}\right)^2}{2}}$$

pozri obr. 6.



Obr. 6
Porovnanie grafu binomického rozdelenia a zodpovedajúcej krivky normálneho rozdelenia.

De Moivre nahrádza binomické rozdelenie krivkou normálneho rozdelenia

Súvis medzi binomickým rozdelením pravdepodobnosti a krivkami normálneho rozdelenia objavil francúzsky matematik Abraham de Moivre, ktorý pred náboženským prenasledovaním utiekol do Anglicka. Svoj objav publikoval v roku 1733, neskôr ho včlenil do druhého vydania jedného zo svojich hlavných diel *The Doctrine of Chances*. De Moivre hľadal funkciu, ktorej hodnoty by boli dostatočne presnou náhradou hodnôt binomického rozdelenia. Súčty viacerých pravdepodobností z nejakého úseku tohto rozdelenia (teda výpočty, ktoré sme pre malú hodnotu n robili v úlohe 21 na s. 25) potom nahradil výpočtom plochy pod grafom nájdenej funkcie, čo bola (hoci to na prvý pohľad tak nevyzerá) jednoduchšia úloha.

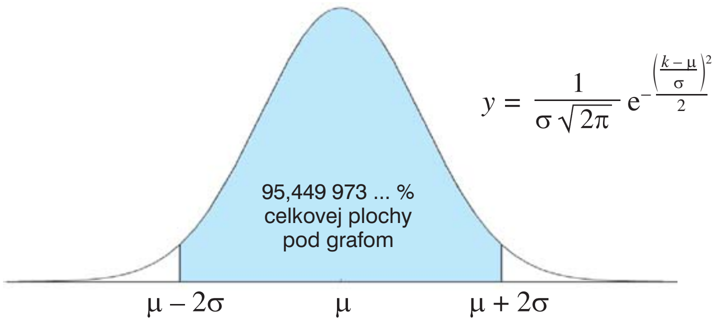
Vo všeobecnosti platí: pre veľké hodnoty n možno hodnoty binomického rozdelenia s parametrami n a p dostatočne presne nahradiť hodnotami funkcie

$$f_{\mu,\sigma}(k) = \frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{\left(\frac{k-\mu}{\sigma}\right)^2}{2}}$$

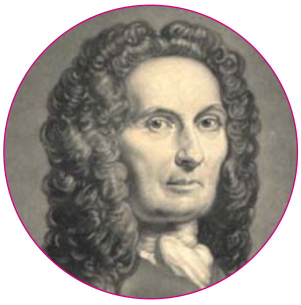
kde $\mu = n \cdot p$ je stredná hodnota a $\sigma = \sqrt{n \cdot p \cdot (1-p)}$ je štandardná odchýlka tohto binomického rozdelenia. Napríklad, pre rozdelenie z obr. 6 ($n = 900$, $p = 0{,}2$) sa $\mu = 180$, $\sigma = 12$ (skontrolujte to), predpis nahrádzajúcej funkcie je

$$f_{180,\,12}(k) = \frac{1}{12\sqrt{2\pi}}\,e^{-\frac{\left(\frac{k-180}{12}\right)^2}{2}}$$

(porovnaj s obr. 6).



Obr. 7
Jednu z dôležitých vlastností kriviek normálneho rozdelenia, ktorú sme – spolu s poznatkom, že hodnoty binomického rozdelenia možno pre veľké n s dostatočnou presnosťou nahradiť hodnotami funkcie $f_{\mu,\sigma}$ – využili v tejto kapitole, pozri obr. 2 na s. 41.



ABRAHAM DE MOIVRE
(1667 – 1754)
PREŠTUDOVAL NEWTONOVO DIELO *PRINCIPIA* (ZÁKLAD KLASICKEJ MECHANIKY) TAK DÔKLADNE, ŽE NEWTON NAŇ ODKAZOVAL TÝCH, KTORÍ NEROZUMELI NIEČOMU Z TEJTO PRÁCE, SO SLOVAMI: „CHOĎTE ZA PÁNOM DE MOIVRE, ON VIE TIETO VECI LEPŠIE AKO JA.“

PIERRE-SIMON LAPLACE
(1749 – 1827)
VÝZNAMNÝ FRANCÚZSKY MATEMATIK A FYZIK, PODOBNE AKO C. F. GAUSS, SVOJIMI PRÁCAMI POTVRDZOVAL VÝZNAM KRIVIEK NORMÁLNEHO ROZDELENIA PRI SKÚMANÍ CHÝB MERANÍ V ASTRONÓMII A ZEMEMERAČSTVE.

ADOLPHE QUETELET
(1796 – 1874)
AKO PRVÝ POUKÁZAL NA MOŽNOSŤ POUŽITIA KRIVIEK NORMÁLNEHO ROZDELENIA V SOCIÁLNYCH VEDÁCH.

2.2 Ako presne sa dá určiť pravdepodobnosť z 200 pokusov?

- Úvodný príklad
- Jednoduchý vzorec na odhad absolútnej chyby

V tejto časti nadviažeme na úvahy z kapitoly 1. V nej sme zistili, že ak pri veľkom počte hodov mincou padne znak podozrivo málo (alebo podozrivo veľa) ráz, je veľmi pravdepodobné, že minca nie je štandarná – teda pravdepodobnosť padnutia znaku **nie** je ½. Otázka potom je: ak pravdepodobnosť nie je ½, aká teda je?

ÚVODNÝ PRÍKLAD

V úlohe 25 na s. 29 sme zistili, že pri 200 hodoch štandardnou mincou je prakticky isté, že padne 86 – 114 znakov. Z toho sme odvodili, že keby pri 200 hodoch mincou padlo napr. 81 znakov, môžeme opodstatnene predpokladať, že minca nie je štandardná.

ÚLOHA

6. Pripomeňte si
a) ako sme našli *interval prakticky istých výsledkov* 86 – 114 znakov v úlohe 25 na s. 29 a ako ho možno nájsť pomocou vzorca zo s. 41,
b) úvahu, ktorou sme zdôvodnili, že pri padnutí 81 znakov z 200 hodov už máme príčinu domnievať sa, že minca nie je štandardná.

Ak je dôvod domnievať sa, že minca **nie** je štandardná, tak prirodzene vzniká otázka: Aká je pre túto mincu pravdepodobnosť, že padne znak? A ako presne vieme túto pravdepodobnosť zistiť z 200 hodov?

Z 200 hodov mincou padol znak 81-krát.
S akou pravdepodobnosťou p padá na tejto minci znak?
Ako presne vieme túto pravdepodobnosť odhadnúť na základe 200 hodov?

Odhad...
Prirodzené je za odhad – teda približnú hodnotu – hľadanej pravdepodobnosti p pokladať relatívnu početnosť $\frac{81}{200} = 0{,}405$, t. j. 40,5 %.

$$p \approx \frac{81}{200} = 0{,}405$$

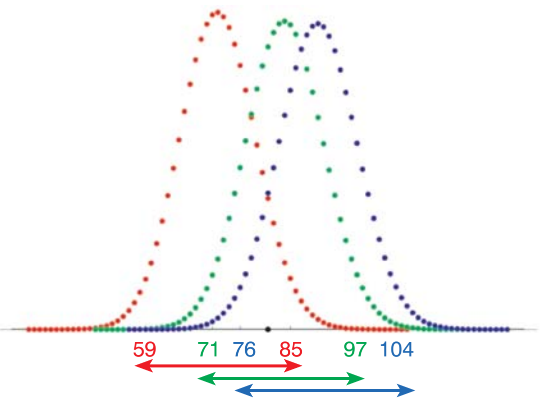
To je v súlade s myšlienkou empirickej pravdepodobnosti, s ktorou sme sa stretli v učebnici pre 2. ročník (namiesto označenia *empirická* sa často používa názov *štatistická* pravdepodobnosť): ak skutočnú hodnotu pravdepodobnosti nepoznáme, tak ju nahrádzame relatívnou početnosťou úspešných pokusov.

Toto číslo je skutočne iba odhadom skutočnej pravdepodobnosti. Prečo? Tvrdiť, že hľadaná pravdepodobnosť je *presne* 0,405, by znamenalo tvrdiť, že 81 je *ideálny* výsledok 200 hodov (skontrolujte, že je to v súlade s opisom nášho pojmu ideálny výsledok zo s. 24). To však odporuje skúsenosti, že nie pri každom pokuse sa podarí dosiahnuť ideálny výsledok. Prirodzenejšie je preto predpokladať menej: 81 nemusí byť *ideálny* výsledok pri 200 hodoch, ale je to jeden z *prakticky istých* výsledkov (o tomto dostatočne dlho diskutujte, až potom čítajte ďalej). Práve s týmto predpokladom súvisí odpoveď na našu druhú otázku: ako presný je odhad $p \approx 0{,}405$?

Výsledok 81 úspechov z 200 pokusov pokladáme za jeden z prakticky istých výsledkov.

... a jeho presnosť
Existuje viac pravdepodobností úspechu p, pre ktoré 81 úspechov z 200 pokusov patrí medzi prakticky isté výsledky, pozri obr. 8.

V našom prípade sa dá ukázať, že takýmito pravdepodobnosťami p sú – ak sa obmedzíme na čísla s 3 desatinnými miestami – všetky hodnoty medzi 0,339 a 0,475 , t. j. od 33,9 % do 47,5 % (podrobnosti nájde záujemca v poznámke **Len pre záujemcov** na tejto strane).



Obr. 8

81 znakov pri 200 hodoch patrí medzi prakticky isté výsledky napr. pri

- *pravdepodobnosti padnutia znaku p =* **36 %** *(prakticky iste padne 59 – 85 znakov)*
- *pravdepodobnosti padnutia znaku p =* **42 %** *(prakticky iste padne 71 – 97 znakov)*
- *pravdepodobnosti padnutia znaku p =* **45 %** *(prakticky iste padne 76 – 104 znakov)*

Prakticky isté výsledky sme určovali pomocou intervalu $(\mu - 2\sigma, \mu + 2\sigma)$. *Hodnotu 81 sme na vodorovnej osi označili čiernou bodkou.*

ÚLOHA

7. Skontrolujte, že ak 200-krát zopakujeme pokus s pravdepodobnosťou úspechu
- $p = 0{,}339$
- $p = 0{,}475$

tak medzi prakticky isté výsledky bude patriť hodnota 81. Na určenie prakticky istých výsledkov použite interval $(\mu - 2\sigma,\ \mu + 2\sigma)$, pozri obr. 2 na s. 41.

Len pre záujemcov

PRE KTORÉ PRAVDEPODOBNOSTI ÚSPECHU p JE 81 ÚSPECHOV Z 200 POKUSOV PRAKTICKY ISTÝ VÝSLEDOK?

Vo všeobecnosti, hodnota k patrí medzi prakticky isté výsledky pri n pokusoch (v našom konkrétnom prípade $k = 81$, $n = 200$), ak k leží v intervale $(\mu - 2\sigma, \mu + 2\sigma)$, teda ak

vzdialenosť medzi k a μ je menšia ako 2σ, (A)

kde $\mu = n \cdot p$, $\sigma = \sqrt{n \cdot p \cdot (1-p)}$, pozri obr. 2 na s. 41.

Túto podmienku možno vysloviť aj v podobe (ak všetky čísla v (A) vydelíme počtom pokusov n)

vzdialenosť medzi $\frac{k}{n}$ a $\frac{\mu}{n}$ je menšia ako $\frac{2\sigma}{n}$,

teda, keďže $\frac{\mu}{n} = p$, $\frac{2\sigma}{n} = 2\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}$,

vzdialenosť medzi $\frac{k}{n}$ a p je menšia ako $2\frac{\sqrt{p\cdot(1-p)}}{\sqrt{n}}$, (B)

pozri obr. 3 a riešenie úlohy 4 c) na s. 43. Formulácia (B) nám vyhovuje viac, pretože sa v nej vyskytuje relatívna početnosť $\frac{k}{n}$, ktorú pokladáme za približnú hodnotu hľadanej pravdepodobnosti.

V našom prípade ($k = 81$, $n = 200$) sa $\frac{k}{n} = \frac{81}{200} = 0{,}405$, a podmienka (B) má tvar

vzdialenosť medzi 0,405 a p je menšia ako $2\frac{\sqrt{p\cdot(1-p)}}{\sqrt{200}}$. (C)

Hľadáme hodnoty p, ktoré spĺňajú podmienku (C)

Podmienku (C) možno zapísať v tvare $|p - 0{,}405| < 2\frac{\sqrt{p\cdot(1-p)}}{\sqrt{200}}$. (*)

Postupnými ekvivalentnými úpravami dostaneme $200(p^2 - 0{,}81p + 0{,}164\,025) < 4p - 4p^2$ (obidve strany nerovnosti sme najprv umocnili na druhú, potom vynásobili číslom 200)

$$204p^2 - 166p + 32{,}805 < 0.$$

Riešením kvadratickej nerovnice $ap^2 + bp + c < 0$ s kladným diskriminantom $D = b^2 - 4ac$ (to je práve náš prípad) sú všetky čísla ležiace medzi koreňmi rovnice $ap^2 + bp + c = 0$, ktorými sú $p_{1,2} = \frac{-b \pm \sqrt{D}}{2a}$. V našom prípade

$$p_1 = \frac{166 - \sqrt{166^2 - 4\cdot 204\cdot 32{,}805}}{408} = \frac{166 - \sqrt{787{,}12}}{408} = 0{,}338\,098...,$$

$$p_2 = \frac{166 + \sqrt{787{,}12}}{408} = 0{,}475\,626...$$

Podmienku (C) preto spĺňajú všetky čísla z intervalu (p_1, p_2).

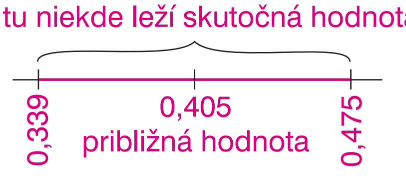
Hľadanou pravdepodobnosťou p je teda niektoré číslo medzi 0,339 a 0,475.

Tým, že uvažujeme iba o *prakticky istých* výsledkoch, vylučujeme z našich úvah situácie, ktoré pokladáme za veľmi málo pravdepodobné. Nemôžeme však tvrdiť, že vylúčené situácie *iste* nenastanú. Preto ani tvrdenie, že hľadaná pravdepodobnosť p je z intervalu 0,339 – 0,475, neplatí so 100-percentnou istotou.

Na druhej strane si treba uvedomiť, že keby sme z našich úvah nechceli vylúčiť žiadne (ani veľmi málo pravdepodobné) možnosti, tak o pravdepodobnosti p nevieme povedať vôbec nič. Nech je totiž pravdepodobnosť úspechu akékoľvek číslo medzi 0 a 1, *môže* sa stať, že pri 200 pokusoch uspejeme 81-krát.

Z tejto úvahy vidno, že bez vylúčenia možností, ktoré pokladáme za málo pravdepodobné, sa ďalej nepohneme. Nevyhnutnou cenou za to, že o hľadanej pravdepodobnosti p vieme povedať niečo zmysluplné, je teda fakt, že naša informácia nebude platiť so 100-percentnou istotou.

Na základe toho môžeme určiť, ako dobre približná hodnota $\frac{81}{200} = 0{,}405$ nahrádza skutočnú (pre nás neznámu) pravdepodobnosť p, teda o koľko sa môže p líšiť od čísla 0,405.



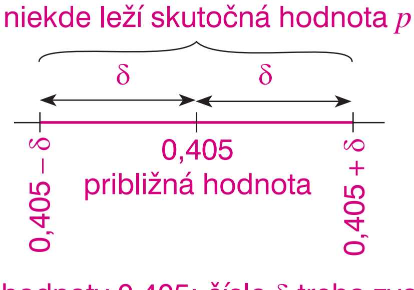
Výsledok 81 úspechov z 200 pokusov patrí medzi prakticky isté výsledky, ak pravdepodobnosť úspechu p je z intervalu $\langle 0{,}339;\ 0{,}475\rangle$.

Obr. 9

S otázkou, ako veľmi sa odlišujú približná a skutočná hodnota sme sa už stretli v kapitole o približných číslach (v učebnici pre 2. ročník). Informácia o veľkosti tejto odchýlky sa nazýva *odhad absolútnej chyby približného čísla*. Pripomeňme, že odhad absolútnej chyby približnej hodnoty 0,405 je kladné číslo δ, pre ktoré platí

vzdialenosť medzi 0,405 a p nie je väčšia ako δ. (D)

Inak povedané, skutočná hodnota p leží niekde medzi $0{,}405 - \delta$ a $0{,}405 + \delta$ (toto si rozmyslite, až potom čítajte ďalej). Zapisujeme to $p = 0{,}405 \pm \delta$.



Odhad δ absolútnej chyby približnej hodnoty 0,405: číslo δ treba zvoliť tak, aby sa skutočná hodnota p nachádzala v intervale $\langle 0{,}405 - \delta;\ 0{,}405 + \delta\rangle$.

Obr. 10

ÚLOHA

8. Skontrolujte, že neznáma hodnota p sa od približnej hodnoty $\frac{81}{200} = 0{,}405$ nelíši o viac ako $\pm\, 0{,}070$. Sformulujte tento výsledok pomocou pojmu *odhad absolútnej chyby*.

$$p = 0{,}405 \pm 0{,}070$$

Pripomeňme ešte všeobecnú zásadu pri práci s približnými číslami: zápis sa upraví tak, aby v odhade chyby bola jedna, nanajvýš dve platné cifry. Zápis $p = 0{,}405 \pm 0{,}070$ tejto podmienke vyhovuje – odhad chyby má dve platné cifry. Keby sme chceli dostať zápis s jednou platnou cifrou v odhade chyby, zaokrúhlime odhad na jednu platnú cifru nahor – v našom prípade nahor na stotiny, a s rovnakou presnosťou – teda na stotiny (tentokrát aritmeticky) zaokrúhlime aj približnú hodnotu (diskutujte o tom, prečo asi je jedno zaokrúhlenie nahor a druhé aritmetické). Dostaneme tak zápis

$$p = 0{,}41 \pm 0{,}07$$

JEDNODUCHÝ VZOREC NA ODHAD ABSOLÚTNEJ CHYBY

Hlavným cieľom úvodného príkladu bolo ujasniť si, čo rozumieme pod odhadom absolútnej chyby. Pre približnú hodnotu $\frac{k}{n} = \frac{81}{200}$ sme hľadali číslo δ – odhad chyby – tak, aby v intervale $\left\langle \frac{81}{200} - \delta;\ \frac{81}{200} + \delta \right\rangle$ ležali všetky pravdepodobnosti, pre ktoré relatívna početnosť $\frac{81}{200}$ – t. j. 81 úspechov z 200 pokusov – patrí medzi *prakticky isté* pri $n = 200$ pokusoch (pozri obr. 9, 10). Teda, aby platilo: ak sa pravdepodobnosť p od čísla $\frac{k}{n} = \frac{81}{200}$ líši o viac ako δ, tak je už prakticky nemožné, aby sme pri $n = 200$ opakovaniach pokusu, v ktorom pravdepodobnosť úspechu je p, dosiahli $k = 81$ úspechov (toto si dobre rozmyslite, až potom čítajte ďalej).

Jedna z možností, ako nájsť takúto hodnotu δ, je založená na výpočtoch, ktoré sme opísali v poznámke **Len pre záujemcov** na s. 47. Tento postup možno použiť pre ľubovoľné hodnoty k a n. Z hľadiska našich potrieb je však pomerne zdĺhavý, preto budeme používať iný odhad, ktorý je síce menej presný, ale ľahko sa pamätá.

Dá sa totiž ukázať nasledujúce tvrdenie (podrobnosti uvádzame v poznámke **Len pre záujemcov** na s. 50): ak pri n pokusoch – pričom n je dostatočne veľké číslo – dosiahneme k úspechov, tak za odhad absolútnej chyby približnej hodnoty $\frac{k}{n}$ možno zvoliť číslo $\delta = \frac{1}{\sqrt{n}}$.
Inak povedané: ak sme pri n pokusoch dosiahli k úspechov (pričom tento výsledok pokladáme za jeden z *prakticky istých*), tak sa pravdepodobnosť úspechu p v jednom pokuse nachádza v intervale $\left\langle \frac{k}{n} - \frac{1}{\sqrt{n}};\ \frac{k}{n} + \frac{1}{\sqrt{n}} \right\rangle$.

Ešte inak: ak sa pravdepodobnosť p od čísla $\frac{k}{n}$ líši o viac ako $\frac{1}{\sqrt{n}}$, tak je už prakticky nemožné, aby sme pri n opakovaniach pokusu, v ktorom pravdepodobnosť úspechu je p, dosiahli k úspechov.

Ak n-krát (pričom n je dostatočne veľké číslo) zopakujeme pokus a dosiahneme pritom k úspechov, tak pravdepodobnosť p úspechu v tomto pokuse je

$$p = \frac{k}{n} \pm \frac{1}{\sqrt{n}}$$

(za predpokladu, že dosiahnutý výsledok k úspechov z n pokusov pokladáme za prakticky istý).

Nasledujúce úlohy by nám mali pomôcť vytvoriť si predstavu, akú presnosť je reálne očakávať, ak pravdepodobnosť úspechu zisťujeme z veľkého počtu opakovaní náhodného pokusu. Na odhad tejto pravdepodobnosti použijeme vzťah $p = \frac{k}{n} \pm \frac{1}{\sqrt{n}}$.

ÚLOHY

9. Náhodný pokus s neznámou pravdepodobnosťou úspechu sme zopakovali 100-krát. S akou presnosťou vieme z týchto 100 opakovaní zistiť neznámu hodnotu pravdepodobnosti úspechu?

10. Koľkokrát by sme mali opakovať náhodný pokus, ak chceme neznámu pravdepodobnosť úspechu p zistiť s presnosťou
a) na celé percentá (teda chyba by mala byť menšia ako 0,5 percentuálneho bodu),
b) na desatiny percenta,
c) na stotiny percenta?

Len pre záujemcov

Zjednodušujeme odhad absolútnej chyby

Vráťme sa k odhadu absolútnej chyby približnej hodnoty $\frac{81}{200} = 0{,}405$. Hľadáme kladné číslo δ tak, aby platilo

vzdialenosť medzi 0,405 a p nie je väčšia ako δ (*)

pozri podmienku (D) na s. 48. O skutočnej – nám však neznámej – hodnote pravdepodobnosti p predpokladáme, že pre ňu 81 úspechov z 200 pokusov patrí medzi prakticky isté výsledky. Vyslovené pomocou relatívnych početností: $\frac{81}{200} = 0{,}405$ patrí pri 200 pokusoch medzi relatívne početnosti prakticky istých výsledkov.

To znamená, že

vzdialenosť medzi 0,405 a p je menšia ako $2 \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{200}}$ (**)

(pozri obr. 3 na s. 43 a podmienku (C) na s. 47).

Stručne zapísané: *hľadáme* δ tak, aby platilo (*), pričom *vieme*, že pre p platí (**).

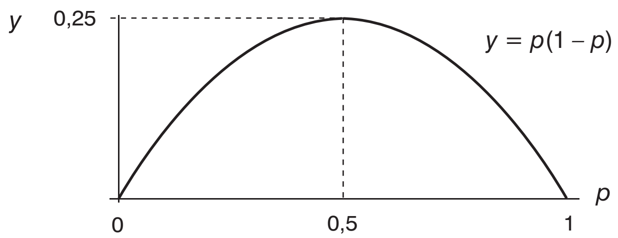
Všimnite si, ako sa formulácie (*) a (**) podobajú. Práve to nám umožní odhadnúť absolútnu chybu. Využijeme pritom tento poznatok (pozri obr. 11):

pre $p \in \langle 0, 1\rangle$ súčin $p \cdot (1-p)$ nikdy nie je väčší ako 0,25.

Z toho vyplýva, že číslo $2 \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{200}}$ z podmienky (**) nemôže byť väčšie ako $2 \cdot \frac{\sqrt{0{,}25}}{\sqrt{200}} = \frac{2 \cdot 0{,}5}{\sqrt{200}} = \frac{1}{\sqrt{200}}$:

$$2 \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{200}} \leq \frac{1}{\sqrt{200}} = 0{,}070\,7... \qquad (***)$$

Teda: hoci hodnotu p nepoznáme, vieme, že vzdialenosť $2 \cdot \frac{\sqrt{p \cdot (1-p)}}{\sqrt{200}}$ z (**) neprevýši hodnotu $\frac{1}{\sqrt{200}}$.



Obr. 11
Grafom kvadratickej funkcie $y = p \cdot (1 - p)$ pre p z intervalu $\langle 0, 1\rangle$ je časť paraboly, ktorá pretína vodorovnú os v hodnotách $p = 0$ a $p = 1$. Najväčšiu hodnotu nadobúda táto funkcia pre $p = 0{,}5$, je ňou číslo $y(0{,}5) = 0{,}5 \cdot 0{,}5 = 0{,}25$.

Toto tvrdenie možno zdôvodniť viacerými spôsobmi, napr. odvolaním sa na súmernosť paraboly, ktorá je grafom kvadratickej funkcie, alebo úpravou predpisu funkcie $y = p \cdot (1 - p)$ na „úplný štvorec“, t. j. na tvar $y = 0{,}25 - (p - 0{,}5)^2$. Venujte dostatočný čas diskusii o týchto zdôvodneniach.

ÚLOHY

11. Vysvetlite, ako z (*), (**) a (***) vyplýva toto tvrdenie:
Vzdialenosť medzi skutočnou hodnotou pravdepodobnosti p a približnou hodnotou 0,405 *nie je väčšia ako* $\frac{1}{\sqrt{200}}$. *Preto za odhad absolútnej chyby možno zvoliť číslo* $\delta = \frac{1}{\sqrt{200}} = 0{,}070\,7\ldots$.

12. Rovnaké úvahy možno použiť aj vo všeobecnom prípade, teda ak pri n opakovaniach pokusu s neznámou pravdepodobnosťou úspechu p uspejeme celkom k-krát. Diskutujte o tom a sformulujte tvrdenie o odhade presnosti približnej hodnoty $\frac{k}{n}$.

2.3 Ďalšie úlohy

ÚLOHY

13. V optimálnych podmienkach sme nechali klíčiť 500 semien, ktorých klíčivosť je 75 %. Na vyklíčenie sa môžeme pozerať ako na 500-krát opakovaný pokus s pravdepodobnosťou úspechu 75 % (diskutujte o tom najprv v triede.)
a) V akom rozmedzí by sa mal pohybovať počet vyklíčených semien?
b) Výsledky z časti a) uveďte v percentách (s presnosťou na celé percentá).
c) Pri akom počte vyklíčených semien by sme mali dôvod oprávnene pochybovať o správnosti uvedenej hodnoty klíčivosti?

14. Ukážte, že na základe uvedených údajov možno odôvodnene predpokladať, že pravdepodobnosť narodenia chlapca **nie** je $\frac{1}{2}$. Zistite, ako presne možno z týchto údajov odhadnúť pravdepodobnosť narodenia chlapca.

Klíčivosť

Klíčivosť – teda schopnosť semien rastliny začať svoj rast – sa vyjadruje v percentách. Napríklad, klíčivosť 80 % znamená, že v laboratórnych podmienkach pri optimálnej vlhkosti a teplote vyklíčilo 80 % z celkového počtu skúmaných semien.

Aká je pravdepodobnosť narodenia chlapca?

V roku 2011 sa v SR narodilo 60 813 živých novorodencov. Dominovalo mužské pohlavie, chlapcov sa narodilo 31 114. Vyplýva to z údajov, ktoré poskytol Štatistický úrad.

DVA UŽITOČNÉ VZORCE

Testovanie liekov

Jednou z hlavných úloh pri testovaní liekov je preukázať, že liek skutočne účinkuje. Opíšme stručne jeden z najjednoduchších používaných postupov. Vyberieme dve skupiny chorých. Prvej skupine podáme liek a zistíme, v koľkých prípadoch liek zaúčinkoval. Druhej skupine podáme neškodnú látku – placebo – a tiež zistíme, v koľkých prípadoch sa stav zlepšil. Pacientov do obidvoch skupín vyberáme náhodne a nikto z nich nevie, či dostal skutočný liek alebo placebo.
Na podanie lieku aj podanie placeba sa môžeme pozerať ako na náhodný pokus. V obidvoch prípadoch nás zaujíma pravdepodobnosť úspechu (tým je zlepšenie pacientovho stavu).
Ak máme liek považovať za účinný, musí byť pravdepodobnosť úspechu prvého pokusu (zlepšenie nastalo po podaní lieku) väčšia ako druhého (zlepšenie nastalo samovoľne, bez podania lieku).

V tabuľke uvádzame výsledky testovaní dvoch liekov na bolesti spôsobené diabetickou neuropatiou:

	Prvý liek	*Druhý liek*
Počet zlepšení / počet pacientov, ktorí dostali liek	105/144	81/165
Počet zlepšení / počet pacientov, ktorí dostali placebo	180/260	73/205

ÚLOHA

15. a) Pre obidva lieky nájdite interval, v ktorom leží pravdepodobnosť, že zlepšenie nastane po podaní lieku a interval pre pravdepodobnosť, že zlepšenie nastane bez podania lieku. Na základe toho rozhodnite, či testovaný liek možno pokladať za účinný.
b) Aký by musel byť v prípade prvého z uvedených liekov počet samovoľných zlepšení stavu, aby sme prvý liek **ne**mohli na základe opísaného postupu pokladať za účinný?

DVA UŽITOČNÉ VZORCE – VÝSLEDKY

2. 76 – 104 bielych guliek; v tomto prípade sa $n = 200$, $p = 0{,}45$, $\mu = 200 \cdot 0{,}45 = 90$, $\sigma = \sqrt{200 \cdot 0{,}45 \cdot 0{,}55} = 7{,}03...$, $(\mu - 2\sigma;\ \mu + 2\sigma) = (75{,}92...;\ 104{,}07...)$.
3. Spomedzi 150 vytiahnutých guliek by malo byť *prakticky iste* 34 – 56 čiernych. Hodnota 32 leží mimo tohto intervalu, môžeme preto odôvodnene predpokladať, že údaj o počte čiernych guliek nebol správny. (V tomto prípade sa $n = 150$, $p = 0{,}3$, $\mu = 150 \cdot 0{,}3 = 45$, $\sigma = \sqrt{150 \cdot 0{,}3 \cdot 0{,}7} = 5{,}61...$, $(\mu - 2\sigma;\ \mu + 2\sigma) = (33{,}7...;\ 56{,}2...)$.)

4. a) $\frac{86}{200} = 0{,}43$ (t. j. 43 %), $\frac{114}{200} = 0{,}57$ (t. j. 57 %), vo všeobecnosti

$$\text{relatívna početnosť} = \frac{\text{absolútna početnosť}}{\text{celkový počet pokusov}}\ \left(\text{v tomto prípade } \frac{\text{absolútna početnosť}}{200}\right),$$

$$\text{relatívna početnosť v \%} = \frac{\text{absolútna početnosť}}{\text{celkový počet pokusov}} \cdot 100.$$

b) Ak absolútna početnosť k leží medzi číslami $\mu - 2\sigma$ a $\mu + 2\sigma$, tak relatívna početnosť $\frac{k}{n}$ leží medzi číslami $\frac{\mu}{n} - \frac{2\sigma}{n}$ a $\frac{\mu}{n} + \frac{2\sigma}{n}$, pritom

$$\frac{\mu}{n} = \frac{np}{n} = p, \qquad \frac{2\sigma}{n} = 2 \cdot \underbrace{\frac{\sqrt{n}}{n}}_{=\frac{1}{\sqrt{n}}} \cdot \sqrt{p(1-p)} = 2 \cdot \frac{\sqrt{p\,(1-p)}}{\sqrt{n}},$$

(pri úprave $\frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}}$ sme využili, že kladné číslo n možno zapísať v tvare $n = \sqrt{n} \cdot \sqrt{n}$, preto $\frac{\sqrt{n}}{n} = \frac{\sqrt{n}}{\sqrt{n} \cdot \sqrt{n}} = \frac{1}{\sqrt{n}}$),

t. j. $$\frac{\mu}{n} - \frac{2\sigma}{n} = p - 2 \cdot \frac{\sqrt{p\,(1-p)}}{\sqrt{n}}, \qquad \frac{\mu}{n} + \frac{2\sigma}{n} = p + 2 \cdot \frac{\sqrt{p\,(1-p)}}{\sqrt{n}}.$$

Teda: ak je n veľké číslo, tak relatívna početnosť úspechov v n pokusoch leží prakticky iste medzi

$$p - 2 \cdot \frac{\sqrt{p\,(1-p)}}{\sqrt{n}} \ \text{ a } \ p + 2 \cdot \frac{\sqrt{p\,(1-p)}}{\sqrt{n}}.$$

c) (o menej ako) $\pm 2 \cdot \frac{\sqrt{p\,(1-p)}}{\sqrt{n}}$.

5. Podľa obr. 3 relatívna početnosť šestiek leží prakticky iste v intervale $\left(\frac{1}{6} - 2 \cdot \frac{\sqrt{\frac{1}{6} \cdot \frac{5}{6}}}{\sqrt{n}},\ \frac{1}{6} + 2 \cdot \frac{\sqrt{\frac{1}{6} \cdot \frac{5}{6}}}{\sqrt{n}}\right)$.
Za n sme dosadili $n = 5\,404\,555$ (počet obyvateľov SR k 31. 3. 2012), resp. $n = 1\,347\,350\,000$ (odhad počtu obyvateľov Číny k 31. 12. 2011). Výsledky po zaokrúhlení: SR (0,166 3; 0,167 0), t. j. 16,63 % až 16,70 %, Čína (0,166 6; 0,166 7), t. j. 16,66 % až 16,67 %.
7. Pre $p = 0{,}339$ je pravý krajný bod intervalu $(\mu - 2\sigma;\ \mu + 2\sigma)$ číslo $\mu + 2\sigma = np + 2\sqrt{np(1-p)} = 200 \cdot 0{,}339 + 2\sqrt{200 \cdot 0{,}339 \cdot 0{,}661} = 81{,}188...$, preto hodnota 81 leží v tomto intervale. Pre $p = 0{,}475$ je ľavý krajný bod intervalu $(\mu - 2\sigma,\ \mu + 2\sigma)$ číslo $\mu - 2\sigma = np - 2\sqrt{np(1-p)} = 200 \cdot 0{,}475 - 2\sqrt{200 \cdot 0{,}475 \cdot 0{,}525} = 80{,}875...$.
8. Väčší z rozdielov 0,475 – 0,405 a 0,405 – 0,339 má hodnotu 0,070, teda žiadne číslo z intervalu $\langle 0{,}339; 0{,}475\rangle$, v ktorom leží neznáma skutočná hodnota p, sa od 0,405 nelíši o viac ako ± 0,070. Preto možno za odhad absolútnej chyby približnej hodnoty 0,405 zvoliť $\delta = 0{,}070$.
9. S presnosťou $\pm \frac{1}{\sqrt{100}} = \pm \frac{1}{10}$ (teda – ak pravdepodobnosť vyjadrujeme v percentách – je to presnosť ± 10 percentuálnych bodov).
10. a) Chyba približnej hodnoty p je nanajvýš $\frac{1}{\sqrt{n}}$, ak je pravdepodobnosť číslo z intervalu $\langle 0, 1\rangle$. Ak pravdepodobnosť vyjadríme v percentách, bude chyba tejto hodnoty nanajvýš $100 \cdot \frac{1}{\sqrt{n}}$. Rovnosť $\frac{100}{\sqrt{n}} = 0{,}5$ platí pre $n = 40\,000$. Preto by počet pokusov mal byť 40 000 a viac.
b) Aspoň 4 milióny-krát. Chyba približnej hodnoty vyjadrenej v percentách by mala byť nanajvýš 0,05.
Rovnosť $\frac{100}{\sqrt{n}} = 0{,}05$ platí pre $n = \left(\frac{100}{0{,}05}\right)^2 = 4\,000\,000$. **c)** Aspoň 400 000 000-krát.
12. Môžeme použiť rovnakú schému uvažovania ako v poznámke **Len pre záujemcov** na s. 50:
- hľadáme δ tak, aby platilo

vzdialenosť medzi približnou hodnotou $\frac{k}{n}$ a skutočnou pravdepodobnosťou p nie je väčšia ako δ,

pritom
- pre každé p, ktoré by mohlo byť skutočnou hodnotou pravdepodobnosti, platí

vzdialenosť medzi $\frac{k}{n}$ a p je menšia ako $2 \cdot \frac{\sqrt{p(1-p)}}{\sqrt{n}}$

(pretože podľa nášho predpokladu patrí $\frac{k}{n}$ medzi relatívne početnosti prakticky istých výsledkov pri n pokusoch, pozri obr. 3 na s. 43)
- súčin $p \cdot (1-p)$ nie je väčší ako 0,25, teda platia nerovnosti $\sqrt{p \cdot (1-p)} \le \sqrt{0{,}25} = 0{,}5$ a $2 \cdot \frac{\sqrt{p(1-p)}}{\sqrt{n}} \le 2 \cdot \frac{0{,}5}{\sqrt{n}} = \frac{1}{\sqrt{n}}$.

Preto vzdialenosť medzi $\frac{k}{n}$ a každým p, ktoré by mohlo byť skutočnou hodnotou pravdepodobnosti, je menšia ako $\frac{1}{\sqrt{n}}$.
Z toho vyplýva, že za δ možno zvoliť číslo $\frac{1}{\sqrt{n}}$.
13. a) 356 – 394 (t. j. 375 ± 24) semien. V tomto prípade sa $n = 500$, $p = 0{,}75$, $\mu = 500 \cdot 0{,}75 = 375$, $\sigma = \sqrt{500 \cdot 0{,}75 \cdot 0{,}25} = 9{,}682...$, za prakticky isté výsledky (počet vyklíčených semien) možno pokladať celé čísla z intervalu $(\mu - 2\sigma,\ \mu + 2\sigma) = (355{,}63...;\ 394{,}36...)$.
b) 71 – 79 % (t. j. 75 ± 4 %), pretože $\frac{355{,}63...}{500} = 0{,}711...$, $\frac{394{,}36...}{500} = 0{,}788...$.
c) Ak počet vyklíčených semien nebude v rozmedzí 356 – 394 semien (pozri výsledok časti a) tejto úlohy).
14. Na údaj o počte narodených chlapcov sa môžeme pozerať ako na výsledok náhodného pokusu, v ktorom pravdepodobnosť úspechu (v tomto prípade narodenia chlapca) je p, pričom pokus sme opakovali 60 813-krát.
Keby pravdepodobnosť bola $p = 0{,}5$, tak pre $n = 60\,813$ sa $\mu = np = 60\,813 \cdot 0{,}5 = 30\,406{,}5$,
$2\sigma = 2 \cdot \sqrt{np(1-p)} = 2 \cdot \sqrt{60\,813 \cdot 0{,}5 \cdot 0{,}5} = 246{,}602...$ a za prakticky isté výsledky (počet narodených chlapcov) možno pokladať celé čísla z intervalu $(\mu - 2\sigma,\ \mu + 2\sigma) = (30\,159{,}897...;\ 30\,653{,}102...)$. Hodnota 31 114 leží mimo tohto intervalu, preto možno odôvodnene predpokladať, že **ne**platí $p = 0{,}5$.

Pravdepodobnosť narodenia chlapca je

$$p = \frac{31\,114}{60\,813} \pm \frac{1}{\sqrt{60\,813}} = 0{,}511\,6... \pm 0{,}004\,055... = 0{,}512 \pm 0{,}005$$

(približnú hodnotu pravdepodobnosti aj odhad chyby sme zaokrúhlili tak, ako je to opísané za úlohou 8 na s. 49).

15. a) prvý liek: Liek možno pokladať za účinný, pretože

- pravdepodobnosť, že zlepšenie nastane po podaní lieku, je

$$\frac{105}{144} \pm \frac{1}{\sqrt{144}}, \text{ teda aspoň } \frac{105}{144} - \frac{1}{\sqrt{144}} \approx 0{,}646,$$

- pravdepodobnosť, že zlepšenie nastane bez podania lieku, je

$$\frac{81}{165} \pm \frac{1}{\sqrt{165}}, \text{ teda nanajvýš } \frac{81}{165} + \frac{1}{\sqrt{165}} \approx 0{,}569\,.$$

druhý liek:

Liek možno pokladať za účinný:

- pravdepodobnosť, že zlepšenie nastane po podaní lieku:

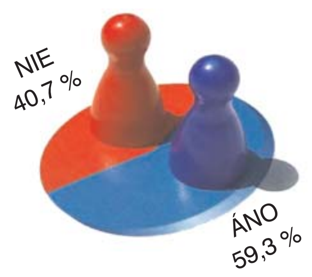
$$\frac{180}{260} \pm \frac{1}{\sqrt{260}}, \text{ teda aspoň } \frac{180}{260} - \frac{1}{\sqrt{260}} \approx 0{,}630,$$

- pravdepodobnosť, že zlepšenie nastane bez podania lieku:

$$\frac{73}{205} \pm \frac{1}{\sqrt{205}}, \text{ teda nanajvýš } \frac{73}{205} + \frac{1}{\sqrt{205}} \approx 0{,}426.$$

b) Viac ako 93. Ak počet samovoľných zlepšení je x, tak pravdepodobnosť, že zlepšenie nastalo bez podania lieku, je $\frac{x}{165} \pm \frac{1}{\sqrt{165}}$, teda nanajvýš $\frac{x}{165} + \frac{1}{\sqrt{165}}$. Ak je táto hodnota väčšia ako $\frac{105}{144} - \frac{1}{\sqrt{144}} \approx 0{,}646$ (dolný odhad pravdepodobnosti, že zlepšenie nastalo po podaní lieku z časti a) tejto úlohy), nemôžeme liek na základe tohto postupu pokladať za účinný. Rovnosť $\frac{x}{165} + \frac{1}{\sqrt{165}} = \frac{105}{144} - \frac{1}{\sqrt{144}}$ nastane pre $x = 165 \cdot \left(\frac{105}{144} - \frac{1}{\sqrt{144}} - \frac{1}{\sqrt{165}}\right) = 93{,}7...$.

3 *Pars pro toto* – KEĎ ČASŤ ZASTUPUJE CELOK



Štatistika tvrdí, že jeden zo štyroch ľudí je duševne chorý.
Tak sa pozri na troch svojich priateľov: ak sú oni v poriadku, tak si to ty.

S prieskumom verejnej mienky – napr. pred voľbami – sa už asi stretol každý z nás. Hlavná myšlienka je jednoduchá: namiesto toho, aby sme sa pýtali všetkých ľudí, vyberieme vhodnú menšiu skupinu a zistíme názory iba v nej. Matematicky vyjadrené: namiesto celého základného súboru vyberieme vhodný výberový súbor a pracujeme iba s ním. V tomto prípade *vhodný* je taký, o ktorom máme dôvod predpokladať, že sa správa podobne ako celý základný súbor.

Otázka je, ako nájsť takýto vhodný výberový súbor. V tejto kapitole ukážeme, že jednou z možností je náhodný výber – zo základného súboru *náhodne* vyberieme primerane veľkú skupinu prvkov, pričom spôsob vyberania je zvolený tak, aby každý prvok základného súboru mal rovnakú šancu byť vybratý.

Náhodne vybrať znamená, že každý prvok základného súboru má rovnakú šancu byť vybratý.

3.1 „Urnový“ model

Obmedzíme sa na najjednoduchší prípad použitia náhodného výberu: zisťovanie, aká časť základného súboru má nejakú vlastnosť – napr., koľko percent populácie Slovenska má krvnú skupinu A (teda, aká je relatívna početnosť tejto krvnej skupiny). Takéto úlohy možno preformulovať do „urnovej“ podoby:

v urne je 5 miliónov guliek	tie predstavujú populáciu Slovenska (o ktorej kvôli jednoduchosti predpokladáme, že je presne 5-miliónová)
z nich časť je biela	biele guľky predstavujú ľudí s krvnou skupinou A
ostatné sú čierne	to sú ľudia s inými krvnými skupinami

Nevieme, aká časť guliek v urne je biela a chceme to zistiť. Zo všetkých 5 miliónov guliek náhodne vyberieme nejakú časť – napr. 1 250 guliek – a zistíme, koľko z vytiahnutých guliek je bielych. Otázkou je, čo na základe toho vieme povedať o percente bielych guliek v urne.

V urne je 5 miliónov guliek. Náhodne vyberieme 1 250 guliek a zistíme, koľko z nich je bielych.
Čo na základe toho vieme povedať o percente bielych guliek v urne?

Zdravý rozum napovedá dve veci:
- na jednej strane je dosť málo pravdepodobné, že relatívna početnosť bielych guliek vo vybraných 1 250 guľkách (nech je to p percent) bude *presne* rovnaká ako ich relatívna početnosť v urne,
- na druhej strane však **ne**očakávame, že odchýlka čísla p od skutočnej relatívnej početnosti bielych guliek v urne bude príliš veľká.

Či zdravý rozum napovedá správne, o tom sa presvedčíme v nasledujúcich častiach tejto kapitoly.

3.2 Je jedno, či guľky vyberáme naraz alebo postupne

Sled našich úvah bude podobný ako v predchádzajúcich kapitolách. Začneme jednoduchšou úlohou. Budeme predpokladať, že poznáme zastúpenie – teda relatívnu početnosť – bielych guliek v urne (nech je to napr. 40 %) a bude nás zaujímať, ako môže dopadnúť výber 1 250 guliek z tejto urny, t. j. ako pravdepodobné je, že z 1 250 vybraných guliek bude práve

• 0 bielych, • 1 biela • 2 biele, ..., • 1 249 bielych, • 1 250 bielych (*)

V urne je 5 miliónov guliek: 2 000 000 bielych a 3 000 000 čiernych. Náhodne vyberieme 1 250 guliek. Aká je pravdepodobnosť, že práve k z nich je bielych?

Výpočet týchto pravdepodobností je pomerne jednoduchý, ak predpokladáme, že 1 250 guliek vyberieme z urny *naraz*. Na tento výpočet sa pozrieme v prvom riešení úlohy 1. V druhom riešení tejto úlohy zistíme, že rovnaký výsledok dostaneme aj vtedy, keď predpokladáme, že 1 250 guliek vyberáme *postupne jednu po druhej*.

S urnovými modelmi sme sa už stretli v poznámke pri úlohe 18 na s. 22. V nej sme avizovali, že okrem modelov s vrátením – o ktorých sme hovorili v 1. kapitole – sa stretneme aj s urnovými modelmi bez vrátenia.
Vyberanie 1 250 guliek jednu po druhej je príkladom urnového modelu bez vrátenia.

Aby sme nemuseli pracovať s veľmi veľkými číslami (2 milióny bielych guliek, 3 milióny čiernych, 1 250 vybraných guliek), použijeme vo výpočtoch menšie hodnoty. Myšlienka výpočtu sa tým nijak nezmení.

ÚLOHA

1. V urne je 200 bielych a 300 čiernych guliek. Náhodne vyberieme 4 z nich. Aká je pravdepodobnosť, že 2 budú biele a 2 čierne?

PRVÝ POSTUP – GUĽKY VYBERIEME NARAZ

V tomto riešení predpokladáme, že všetky 4 guľky vyberieme naraz – teda nerozlišujeme, ktorá z nich je prvá, druhá, tretia, štvrtá.

Pri výpočte pravdepodobnosti použijeme Laplaceovu schému. Lepšie sa nám bude uvažovať, ak si predstavíme, že biele guľky sú očíslované od 1 do 200 a čierne majú čísla od 201 do 500 (uistite sa v diskusii, že hľadaná pravdepodobnosť nezávisí od toho, či guľky sú alebo nie sú očíslované).

Pripomíname:
Laplaceovu schému používame, ak výsledky náhodného pokusu možno opísať pomocou možností, ktoré sú rovnako pravdepodobné (rozmyslite si, čo sú tieto rovnako pravdepodobné možnosti v našom prípade). Ak týchto možností je celkom n, každá z nich nastane v ideálnom prípade v $\frac{1}{n}$ (jednej n-tine) z celkového počtu pokusov.

Ak z týchto n možností je k priaznivých pre jav, ktorého pravdepodobnosť hľadáme, tak priaznivé možnosti predstavujú $\frac{k}{n}$ (ká entín) z celkového počtu pokusov. Preto pravdepodobnosť je $\frac{k}{n}$, teda

$$\text{pravdepodobnosť javu} = \frac{\text{počet priaznivých možností}}{\text{počet všetkých možností}}$$

Začnime ***počtom všetkých možností***. Štyri guľky z 500 – teda štyri rôzne čísla z čísel 1 až 500 – možno vybrať celkom $\binom{500}{4}$ spôsobmi (toto si rozmyslite).

Nájsť ***počet všetkých priaznivých možností*** znamená, vypočítať počet všetkých štvoríc navzájom rôznych čísel, pre ktoré platí: dve čísla sú od 1 do 200, zvyšné dve čísla sú od 201 do 500. Dve rôzne čísla z čísel 1 až 200 (teda dve biele guľky) môžeme vybrať $\binom{200}{2}$ spôsobmi, dve rôzne čísla z 300 čísel od 201 do 500 (dve čierne guľky) $\binom{300}{2}$ spôsobmi. Pritom každý výber bielych guliek môžeme skombinovať s ľubovoľným výberom čiernych guliek, pozri tab. 1 (nečítajte ďalej, kým si to nerozmyslíte). Preto celkový počet možností, ako vybrať štvoricu čísel požadovaných vlastnostostí – teda počet všetkých priaznivých možností – je $\binom{200}{2} \cdot \binom{300}{2}$.

ÚVAHA, KTORÚ SME POUŽILI, SA NAZÝVA *KOMBINATORICKÉ PRAVIDLO SÚČINU*.

		možnosti výberu dvoch rôznych čísel spomedzi čísel 1 – 200 (dvoch bielych guliek)											
		1,2	1,3	…	1,200	2,3	2,4	…	2,200	…	199,200
možnosti výberu dvoch rôznych čísel z čísel 201 – 500 (dvoch čiernych guliek)	201,202	1,2 201,202	1,3 201,202	…	1,200 201,202	2,3 201,202	2,4 201,202	…	2,200 201,202	…	199,200 201,202
	201,203	1,2 201,203	1,3 201,203	…	1,200 201,203	2,3 201,203	2,4 201,203	…	2,200 201,203	…	199,200 201,203
	201,204	1,2 201,204	1,3 201,204	…	1,200 201,204	2,3 201,204	2,4 201,204	…	2,200 201,204	…	199,200 201,204
	⋮	⋮	⋮		⋮	⋮	⋮		⋮		⋮
	201,500	1,2 201,500	1,3 201,500	…	1,200 201,500	2,3 201,500	2,4 201,500	…	2,200 201,500	…	199,200 201,500
	202,203	1,2 202,203	1,3 202,203	…	1,200 202,203	2,3 202,203	2,4 202,203	…	2,200 202,203	…	199,200 202,203
	202,204	1,2 202,204	1,3 202,204	…	1,200 202,204	2,3 202,204	2,4 202,204	…	2,200 202,204	…	199,200 202,204
	⋮	⋮	⋮		⋮	⋮	⋮		⋮		⋮
	202,500	1,2 202,500	1,3 202,500	…	1,200 202,500	2,3 202,500	2,4 202,500	…	2,200 202,500	…	199,200 202,500
	⋮	⋮	⋮		⋮	⋮	⋮		⋮		⋮
	499,500	1,2 499,500	1,3 499,500	…	1,200 499,500	2,3 499,500	2,4 499,500	…	2,200 499,500	…	199,200 499,500

Tab. 1

V hornom riadku sú všetky možnosti výberu dvoch bielych guliek, v ľavom stĺpci všetky možnosti výberu dvoch čiernych guliek (viete zistiť, podľa akého pravidla sme tieto možnosti zoraďovali?). Každá bunka v hrubo orámovanej tabuľke predstavuje jednu priaznivú možnosť.

Celkový počet buniek je počet stĺpcov $\left(= \text{počet výberu bielych guliek, ten je } \binom{200}{2} = 19\,900\right) \times$ *počet riadkov* $\left(= \text{počet výberov čiernych guliek, ten je } \binom{300}{2} = 44\,850\right)$*, t. j.* $\binom{200}{2} \cdot \binom{300}{2}$

Podľa Laplaceovej schémy hľadaná pravdepodobnosť je

$$\frac{\text{počet priaznivých výberov 4 guliek}}{\text{počet všetkých výberov 4 guliek}} = \frac{\binom{200}{2}\cdot\binom{300}{2}}{\binom{500}{4}} = \frac{\frac{200\cdot 199}{2\cdot 1}\cdot\frac{300\cdot 299}{2\cdot 1}}{\frac{500\cdot 499\cdot 498\cdot 497}{4\cdot 3\cdot 2\cdot 1}} = 0{,}346\ 87... \quad \text{(P1)}$$

teda približne 34,7 %.

DRUHÝ POSTUP – GUĽKY VYBERÁME JEDNU PO DRUHEJ

Teraz budeme guľky vyberať postupne – teda vyberieme prvú, druhú, tretiu a napokon štvrtú guľku. Aj v tomto prípade si pomôžeme predstavou, že guľky sú očíslované (biele od 1 do 200, čierne od 201 do 500) a použijeme Laplaceovu schému.

Vybrať postupne 4 guľky z 500 je to isté ako vybrať postupne 4 rôzne čísla z čísel 1 až 500 (pričom zohľadňujeme aj poradie, v akom sme čísla vybrali). Preto počet ***všetkých možností*** je počet všetkých usporiadaných štvoríc navzájom rôznych čísel z 1 až 500. Tých je **500 · 499 · 498 · 497**.

Pripomeňme, že ide o 4-prvkové variácie z 500 prvkov a že výpočet ich počtu sa zakladá na opakovanom použití kombinatorického pravidla súčinu:

- na prvom mieste môže byť ľubovoľný z 500 prvkov,
- akokoľvek zvolíme prvok na 1. mieste, vždy máme 499 možností výberu prvku na druhom mieste, teda pre výber prvku na 1. a 2. mieste máme celkom 500 · 499 možností,
- pre každú z týchto 500 · 499 možností máme 498 možností výberu prvku na 3. mieste ...
  (zvyšok zdôvodnenia už prenechávame vám).

Priaznivé možnosti sú tie usporiadané štvorice, v ktorých na dvoch miestach sú čísla od 1 do 200 a na zvyšných dvoch čísla od 201 do 500. Všetky priaznivé možnosti môžeme rozdeliť na skupiny podľa toho, na ktorom mieste v nich stoja čísla od 1 do 200:

na 1. a 2. mieste, na 1. a 3. mieste, na 1. a 4. mieste,
na 2. a 3. mieste, na 2. a 4. mieste, na 3. a 4. mieste (*)

TÁTO ÚVAHA SA NAZÝVA *KOMBINATORICKÉ PRAVIDLO SÚČTU.*

Každá priaznivá možnosť patrí iba do jednej z týchto skupín (skontrolujte to), preto počet všetkých priaznivých možností je súčet prvkov jednotlivých skupín.

V koľkých prípadoch vytiahneme bielu guľku iba ako prvú a štvrtú?

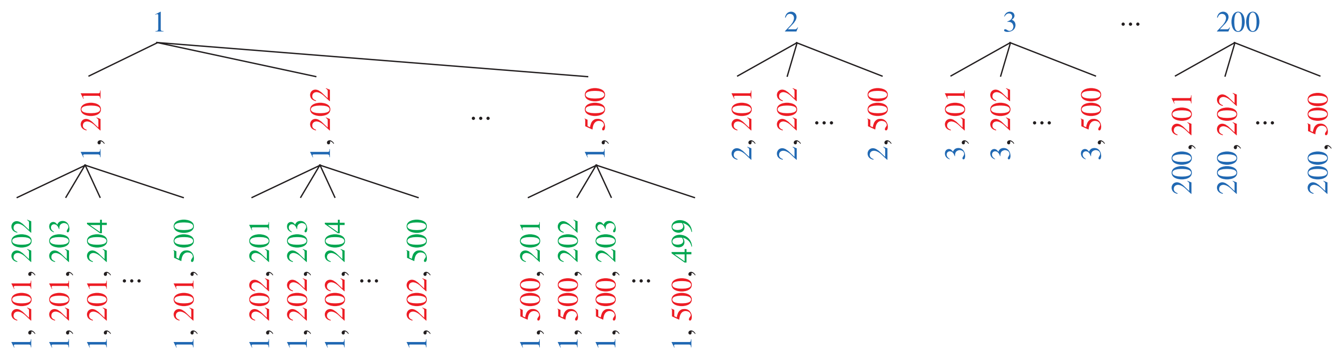
Vypočítajme na ukážku počet tých priaznivých možností, v ktorých sú čísla od 1 do 200 na prvom a štvrtom mieste (vo výpočte opäť použijeme kombinatorické pravidlo súčinu):

ÚVAHY OPÍSANÉ V PRVÝCH TROCH ODRÁŽKACH SCHEMATICKY ZNÁZORŇUJE OBR. 1.

- na **prvom** mieste môže byť ktorékoľvek z čísel 1 – 200, to je 200 možností,
- na **druhom** mieste môže byť ktorékoľvek z 300 čísel 201 – 500, preto počet možností výberu dvoch prvých čísel je 200 · 300,
- pre každú z týchto 200 · 300 možností máme na výber 299 čísel na **treťom** mieste (všetky čísla 201 – 500 okrem čísla, ktoré sme vybrali na druhé miesto), preto počet možností výberu prvých troch čísel je 200 · 300 · 299,
- pre každú z týchto 200 · 300 · 299 možností máme na výber 199 čísel na štvrté miesto (ktorékoľvek z čísel 1 – 200 okrem čísla, ktoré sme vybrali na prvé miesto).
  Preto celkový počet možností je **200 · 300 · 299 · 199**.

Výsledná pravdepodobnosť

Rovnaký výsledok – len s iným poradím činiteľov 200, 300, 299, 199 – by sme dostali pre každú zo **6** skupín uvedených v (*), rozmyslite si to. Preto počet všetkých



Obr. 1

*Číslo na prvom mieste je **modré**, na druhom mieste **červené**, na treťom mieste **zelené***

(štvrté miesto už kvôli prehľadnosti obrázku neznázorňujeme, z rovnakého dôvodu sme výber čísla na treťom mieste naznačili len pre možnosti, v ktorých je na prvom mieste číslo 1).

priaznivých možností je $\mathbf{6 \cdot 200 \cdot 300 \cdot 299 \cdot 199}$. Všimnime si ešte, že číslo 6 bol počet všetkých možností, ako vybrať zo 4 miest tie 2, na ktorých sú čísla od 1 do 200. Teda 6 bol počet všetkých 2-prvkových kombinácií zo 4 prvkov: $\mathbf{6} = \binom{4}{2}$.

Hľadaná pravdepodobnosť je podľa Laplaceovej schémy

$$\frac{\mathbf{6 \cdot 200 \cdot 199 \cdot 300 \cdot 299}}{\mathbf{500 \cdot 499 \cdot 498 \cdot 497}}, \text{ t. j. } \binom{4}{2} \cdot \frac{\mathbf{200 \cdot 199 \cdot 300 \cdot 299}}{\mathbf{500 \cdot 499 \cdot 498 \cdot 497}}. \qquad \text{(P2)}$$

Dostali sme rovnaký výsledok ako v prvom riešení

Nie je ťažké skontrolovať, že sme dostali rovnaký výsledok ako v prvom riešení (v ktorom sme guľky vyberali naraz) – stačí trochu inak zapísať zlomok v (P1)

$$\frac{\frac{200 \cdot 199}{2 \cdot 1} \cdot \frac{300 \cdot 299}{2 \cdot 1}}{\frac{500 \cdot 499 \cdot 498 \cdot 497}{4 \cdot 3 \cdot 2 \cdot 1}} = \frac{\frac{1}{2 \cdot 1} \cdot \frac{1}{2 \cdot 1}}{\frac{1}{4 \cdot 3 \cdot 2 \cdot 1}} \cdot \frac{\mathbf{200 \cdot 199 \cdot 300 \cdot 299}}{\mathbf{500 \cdot 499 \cdot 498 \cdot 497}}$$

a vhodne upraviť výraz $\dfrac{\frac{1}{2 \cdot 1} \cdot \frac{1}{2 \cdot 1}}{\frac{1}{4 \cdot 3 \cdot 2 \cdot 1}}$

$$\frac{\frac{1}{2 \cdot 1} \cdot \frac{1}{2 \cdot 1}}{\frac{1}{4 \cdot 3 \cdot 2 \cdot 1}} = \frac{4 \cdot 3 \cdot 2 \cdot 1}{(2 \cdot 1) \cdot (2 \cdot 1)} = \frac{4!}{(2!) \cdot (2!)} = \binom{4}{2}$$

(skontrolujte, že z týchto úvah už vyplýva, že (P1) a (P2) sú rovnaké hodnoty).

Pripomeňme vzorec na výpočet kombinačného čísla „en nad ká“:

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

ÚLOHA

2. Zapíšte pomocou kombinačných čísel – teda tak ako v (P1) – aj pomocou súčinov – ako v (P2) – pravdepodobnosti ostatných možností, ktoré môžu nastať pri výbere 4 guliek z 500 (z ktorých je 200 bielych a 300 čiernych), teda pravdepodobnosti, že zo 4 vybraných guliek

a) je 0 bielych (t. j. všetky sú čierne)
b) je práve 1 biela,
c) sú práve 3 biele,
d) sú všetky 4 biele.

Skontrolujte – podobne, ako sme to urobili na záver riešenia úlohy 1 – že obidva zápisy vyjadrujú tú istú hodnotu. Obidve zapísané hodnoty potom vypočítajte na kalkulačke alebo tabuľkovom kalkulátore.

NA CHVÍĽU SA VRÁTIME K PÔVODNÝM VEĽKÝM ČÍSLAM Z OTÁZKY O KRVNÝCH SKUPINÁCH, POTOM SKÚSIME ZAPÍSAŤ VÝSLEDOK VO VŠEOBECNOM PRÍPADE.

ÚLOHY

3. V urne je 5 miliónov guliek: 2 milióny bielych a 3 milióny čiernych. Zapíšte pomocou kombinančných čísel pravdepodobnosť, že z 1 250 vybraných guliek bude práve
a) 490 bielych,
b) k bielych (k je niektoré z čísel 0, 1, 2, ..., 1 250).

VÝSLEDKOM ÚLOHY 3 a) JE ČÍSLO $\dfrac{\binom{2\,000\,000}{490}\cdot\binom{3\,000\,000}{760}}{\binom{5\,000\,000}{1\,250}}$. KOMBINAČNÉ ČÍSLA V ČITATELI AJ MENOVATELI TOHTO ZLOMKU SÚ OBROVSKÉ, VÝPOČET HODNOTY TOHTO PODIELU POSTUPOM *VYPOČÍTAME HODNOTU ČITATEĽA, POTOM MENOVATEĽA A VYDELÍME ICH* NA TABUĽKOVOM KALKULÁTORE JE PRAKTICKY NEMOŽNÝ. AK BY SME CHCELI UVEDENÚ PRAVDEPODOBNOSŤ VYPOČÍTAŤ, MUSÍME ZVOLIŤ INÝ POSTUP. VÝSLEDOK NAPÍŠEME TAK AKO (P2) V RIEŠENÍ ÚLOHY 1:

$$\binom{1\,250}{490}\cdot\overbrace{\frac{2\,000\,000}{5\,000\,000}\cdot\frac{1\,999\,999}{4\,999\,999}\cdot\frac{1\,999\,998}{4\,999\,998}\cdot\ldots\cdot\frac{1\,999\,511}{4\,999\,511}}^{B\,=\,490\text{ činiteľov}}\cdot\overbrace{\frac{3\,000\,000}{4\,999\,510}\cdot\frac{2\,999\,999}{4\,999\,509}\cdot\ldots\cdot\frac{2\,999\,241}{4\,998\,750}}^{C\,=\,1\,250\,-\,490\,=\,760\text{ činiteľov}}\qquad(*)$$

A AKO SÚČIN ZAPÍŠEME AJ KOMBINAČNÉ ČÍSLO

$$\binom{1\,250}{490}=\overbrace{\frac{1\,250}{490}\cdot\frac{1\,249}{489}\cdot\ldots\cdot\frac{762}{2}\cdot\frac{761}{1}}^{A\,=\,490\text{ činiteľov}}$$

POTOM NA TABUĽKOVOM KALKULÁTORE NAPROGRAMUJEME VÝPOČET SÚČINU V PORADÍ:

PRVÝ ČINITEĽ Z A × PRVÝ ČINITEĽ Z B × PRVÝ ČINITEĽ Z C ×
× DRUHÝ ČINITEĽ Z A × DRUHÝ ČINITEĽ Z B × DRUHÝ ČINITEĽ Z C × ... ×
× 490-TY ČINITEĽ Z A × 490-TY ČINITEĽ Z B × 490-TY ČINITEĽ Z C ×
× 491-VÝ ČINITEĽ Z C × 492-HÝ ČINITEĽ Z C × ... × 760-TY ČINITEĽ Z C

PODSTATNÉ JE, ŽE POČAS VÝPOČTU NIKDY NEPOČÍTAME S PRÍLIŠ VEĽKÝMI ČÍSLAMI (KTORÉ BY MOHLI BYŤ PRE KALKULÁTOR PROBLÉMOM). DOSIAHNEME TO VĎAKA TOMU, ŽE VEĽKÉ ČINITELE ZO SÚČINU A „POPREKLADÁME" MALÝMI HODNOTAMI ZO SÚČINOV B A C.

4. Z urny, v ktorej je A bielych a B čiernych guliek (iné guľky tam už nie sú), vyberáme n guliek. Zapíšte pravdepodobnosť, že práve k z vybraných guliek je bielych.

3.3 Pomôžu nám bernoulliovské pokusy

- URNOVÝ MODEL BEZ VRÁTENIA ...
- ... A URNOVÝ MODEL S VRÁTENÍM ...
- ... DÁVAJÚ ZA ISTÝCH PREDPOKLADOV PRAKTICKY ROVNAKÉ VÝSLEDKY

V riešení úlohy 1 sme zistili, že výsledky (P1) a (P2) sú rovnaké (pre ďalšie prípady sme to overili v riešení úlohy 2). To znamená, že hľadaná pravdepodobnosť nezávisí od toho, či guľky vyberáme naraz – výsledok (P1) alebo postupne – výsledok (P2).

URNOVÝ MODEL BEZ VRÁTENIA ...

Druhý – postupný – spôsob vyberania nás zaujíma preto, že od neho vedie cesta k zjednodušeniu výpočtov aj celého riešenia. Tento postup sa označuje ako *urnový model bez vrátenia*: z urny obsahujúcej biele a čierne guľky vyberieme guľku, zistíme jej farbu, ale guľku do urny už nevrátime. Potom vyberieme ďalšiu guľku atď. Celkovo takto vyberieme n guliek, zaujíma nás, koľko z nich je bielych.

... A URNOVÝ MODEL S VRÁTENÍM ...

S vyberaním guliek z urny sme sa už stretli pri bernoulliovských pokusoch (stručne si ich pripomeňte, potom čítajte ďalej). Tie možno opísať ako *urnový model s vrátením.*

V ňom – na rozdiel od modelu bez vrátenia – vybranú guľku po zistení jej farby vrátime do urny, a až potom vyberáme ďalšiu guľku.

... DÁVAJÚ ZA ISTÝCH PREDPOKLADOV PRAKTICKY ROVNAKÉ VÝSLEDKY

Tento rozdiel – či guľku do urny vrátime alebo nevrátime – sa prejaví na pravdepodobnosti vybratia bielej guľky. Platí totiž

pravdepodobnosť, že z urny vyberieme bielu guľku = relatívna početnosť bielych guliek v urne

(toto si dobre rozmyslite).

Čo z toho vyplýva? Pri bernoulliovských pokusoch – teda pri urnovom modeli *s vrátením* – sa relatívna početnosť bielych guliek v urne nemení, preto pravdepodobnosť vybrať z urny bielu guľku (t. j. pravdepodobnosť úspechu) je pri každom výbere rovnaká.
Pri urnovom modeli *bez vrátenia* – ktorý nás zaujíma v tejto kapitole – sa relatívna početnosť bielych guliek v urne mení v každom kroku, teda s každou vybranou guľkou sa mení aj pravdepodobnosť výberu bielej guľky v nasledujúcom kroku.

Ak z urny obsahujúcej 200 bielych a 300 čiernych guliek vyberieme postupne bielu, čiernu, čiernu a bielu guľku (túto možnosť sme opisovali v riešení úlohy 1), tak

- v prvom ťahu je pravdepodobnosť vybratia bielej guľky $\frac{200}{500}$ (v urne je 500 guliek, z toho 200 bielych),
- v druhom ťahu je táto pravdepodobnosť $\frac{199}{499}$ (po vybratí 1 bielej guľky v prvom ťahu zostane v urne 499 guliek, z toho 199 bielych),
- v treťom ťahu to je $\frac{199}{498}$ (po vybratí 1 bielej a 1 čiernej guľky zostane v urne 498 guliek, z toho 199 bielych),
- a napokon v štvrtom ťahu je táto pravdepodobnosť $\frac{199}{497}$ (prečo?).

Uvedené štyri pravdepodobnosti sú skutočne rôzne, všimnite si však, že všetky sa len veľmi málo líšia od hodnoty $\frac{200}{500} = 0{,}4$, t. j. od počiatočnej relatívnej početnosti bielych guliek v urne.

Ak je počet vyberaných guliek v porovnaní s počtom guliek v urne veľmi malý, je veľmi malá aj zmena početnosti bielych guliek. V takom prípade je pravdepodobnosť vybrať bielu guľku prakticky rovnaká pri každom výbere. Preto sa nedopustíme veľkej chyby, ak túto pravdepodobnosť budeme pokladať – tak ako v bernoulliovských pokusoch – za nemennú, jej hodnotou p bude počiatočná relatívna početnosť bielych guliek v urne (nečítajte ďalej, kým si to nerozmyslíte).

	napríklad
Na výber n guliek z urny sa potom môžeme pozerať ako na n-násobné opakovanie pokusu, v ktorom pravdepodobnosť úspechu – vybratia bielej guľky – je p. Inými slovami: urnový model bez vrátenia môžeme nahradiť urnovým modelom s vrátením (teda bernoulliovským pokusom).	Vo výpočtoch z úloh 1 a 2 sa nedopustíme veľkej chyby, ak budeme predpokladať, že *z urny vyberáme postupne 4 guľky, pričom pravdepodobnosť vybrať bielu guľku je v každom z týchto 4 krokov rovnaká* (v našom prípade 0,4, t. j. 40 %).

ÚLOHA

5. Vyťahovanie 4 guliek z urny, v ktorej je 500 guliek (200 bielych a 300 čiernych) nahraďte zodpovedajúcim bernoulliovským pokusom. Vypočítajte pravdepodobnosti pre jednotlivé počty úspechov v tomto bernoulliovskom pokuse. Výsledky porovnajte s pravdepodobnosťami, ktoré sme vypočítali v úlohách 1 a 2.

Skontrolujme uvedenú úvahu na výsledkoch úloh 1 a 2.

Ak je počet vybraných guliek v porovnaní s celkovým počtom guliek v urne malý, tak pravdepodobnosti pre urnový model bez vrátenia sú prakticky rovnaké ako pre model s vrátením (t. j. pre bernoulliovský pokus).

Preto:

	napríklad
ak je v základnom súbore veľký počet guliek, z toho p % bielych,	ak je v základnom súbore 5 miliónov guliek, z toho 40 % bielych,
tak pravdepodobnosť, že z n vybraných guliek (bez vrátenia) bude práve k bielych, je približne	tak pravdepodobnosť, že z 1 250 vybraných guliek (bez vrátenia) bude práve 490 bielych, je približne
$\binom{n}{k} \cdot \left(\frac{p}{100}\right)^k \cdot \left(1 - \frac{p}{100}\right)^{n-k}$	$\binom{1\,250}{490} \cdot 0{,}4^{490} \cdot 0{,}6^{760}$

ÚLOHA

6. a) Skontrolujte zápis hodnoty pravdepodobnosti v zodpovedajúcom bernoulliovskom pokuse v pravom stĺpci predchádzajúcej tabuľky.
Presnú hodnotu pravdepodobnosti, že z 1 250 vybraných guliek – bez vrátenia – bude práve 490 bielych, sme zapísali v riešení úlohy 3a). Jej ďalší zápis je (*) v poznámke za úlohou 3 na s. 60.
b) To, že približná hodnota tejto pravdepodobnosti je $\binom{1\,250}{490} \cdot 0{,}4^{490} \cdot 0{,}6^{760}$, možno zdôvodniť aj pomocou zápisu (*). Diskutujte o tom, ako.

3.4 Odpoveď na otázku z úvodu

V predchádzajúcom článku sme videli, že urnové modely bez vrátenia môžeme, ak je počet vyberaných guliek veľmi malý v porovnaní s počtom všetkých guliek v urne, nahradiť modelmi s vrátením (teda bernoulliovskými pokusmi). To nám umožňuje využiť výsledky, ku ktorým sme dospeli v predchádzajúcich kapitolách. Vďaka tomu teraz môžeme zodpovedať otázku z úvodu (úloha 8 b). Zaradili sme pred ňu ešte niekoľko otázok, ktorých riešenie rovnako využíva poznatky o bernoulliovských pokusoch.

Pred riešením úlohy 7 odporúčame najprv pripomenúť si základné vlastnosti binomického rozdelenia pravdepodobnosti znázornené na obr. 2 a 3 na s. 41 a 43. Podobne pred riešením úlohy 8 je rozumné vrátiť sa k textu na s. 49, v ktorom uvádzame jednoduchý vzťah na určenie presnosti, s akou vieme z veľkého počtu opakovaní pokusu zistiť pravdepodobnosť úspechu v tomto pokuse.

ÚLOHY

7. Z urny, v ktorej je 5 miliónov guliek (2 milióny bielych a 3 milióny čiernych) náhodne vyberieme 1 250 guliek.
a) Aký počet bielych guliek môžeme prakticky isto očakávať medzi vybranými guľkami? Pri riešení využite interval $(\mu - 2\sigma,\ \mu + 2\sigma)$.
b) Akú početnosť bielych guliek môžeme prakticky isto očakávať v tomto výbere 1 250 guliek?

8. Z urny obsahujúcej 5 miliónov guliek (časť sú biele, zvyšok čierne) sme vybrali 1 250 guliek. Z nich bolo 431 bielych.
a) Keby sme predpokladali, že v urne je 40 % bielych guliek, zodpovedal by tento výsledok nášmu očakávaniu?
b) Ako presne vieme na základe výsledku 431 *bielych z* 1 250 *vybratých guliek* odhadnúť relatívnu početnosť bielych guliek v urne?
c) Diskutujte o tom, či sa odpoveď na otázky a) a b) zmení, ak celkový počet guliek v urne nepoznáme (a vieme iba, že je veľmi veľký).

d) Koľko guliek z urny by sme museli vybrať, aby sme relatívnu početnosť bielych guliek v nej určili s chybou menšou ako 0,01 (teda relatívna početnosť vyjadrená v percentách by mala chybu nanajvýš $\pm$ 1)?

Ak z urny obsahujúcej veľký počet bielych a čiernych guliek vyberieme 1 250 guliek a zistíme, že 431 z nich je bielych, tak relatívna početnosť bielych guliek v urne je

$$p = \frac{431}{1\,250} \pm \frac{1}{\sqrt{1\,250}} = 0{,}34 \pm 0{,}03$$

(za predpokladu, že výsledok 431 *bielych z 1 250 vybratých* pokladáme za jeden z prakticky istých výsledkov).

Pars pro toto – KEĎ ČASŤ ZASTUPUJE CELOK – VÝSLEDKY

2. a) $\dfrac{\binom{300}{4}}{\binom{500}{4}} = 0{,}128\,560\,891...,$ to môžeme – aby mal náš zápis rovnakú štruktúru ako ostatné výsledky – zapísať v podobe $\dfrac{\binom{200}{0}\cdot\binom{300}{4}}{\binom{500}{4}}$ (pripomeňme dohodu, podľa ktorej sa $\binom{200}{0} = 1$), $\dfrac{300\cdot 299\cdot 298\cdot 297}{500\cdot 499\cdot 498\cdot 497}$, aj tento zápis môžeme – kvôli jeho zjednoteniu s ostatnými výsledkami – písať v tvare $\binom{4}{0}\cdot\dfrac{300\cdot 299\cdot 298\cdot 297}{500\cdot 499\cdot 498\cdot 497}$.

b) $\dfrac{\binom{200}{1}\cdot\binom{300}{3}}{\binom{500}{4}} = \binom{4}{1}\cdot\dfrac{200\cdot 300\cdot 299\cdot 298}{500\cdot 499\cdot 498\cdot 497} = 0{,}346\,291\,963...,$ **c)** $\dfrac{\binom{200}{3}\cdot\binom{300}{1}}{\binom{500}{4}} = \binom{4}{3}\cdot\dfrac{200\cdot 199\cdot 198\cdot 300}{500\cdot 499\cdot 498\cdot 497} = 0{,}153\,134\,564...,$

d) $\dfrac{\binom{200}{4}\cdot\binom{300}{0}}{\binom{500}{4}} = \binom{4}{4}\cdot\dfrac{200\cdot 199\cdot 198\cdot 197}{500\cdot 499\cdot 498\cdot 497} = 0{,}025\,139\,590...,$ na použitie kombinačných čísel $\binom{300}{0} = 1$ a $\binom{4}{4} = 1$ sa vzťahuje rovnaká poznámka ako v riešení úlohy 2 a).

3. a) $\dfrac{\binom{2\,000\,00}{490}\cdot\binom{3\,000\,000}{760}}{\binom{5\,000\,000}{1\,250}}$ (pravdepodobnosť je približne 0,019 6, t. j. približne 1,96 %). **b)** $\dfrac{\binom{2\,000\,000}{k}\cdot\binom{3\,000\,000}{1\,250-k}}{\binom{5\,000\,000}{1\,250}}$.

4. $\dfrac{\binom{A}{k}\cdot\binom{B}{n-k}}{\binom{A+B}{n}}$.

5. Pravdepodobnosť úspechu v zodpovedajúcom bernoulliovskom pokuse je $p = \dfrac{200}{500} = 0{,}4$.

pravdepodobnosť, že zo 4 vytiahnutých guliek	*urnový model s vrátením (bernoulliovské pokusy)*	*urnový model bez vrátenia (riešenia úloh 1 a 2)*
je 0 bielych	$0{,}6^4 = 0{,}129\,6 \approx 0{,}130$	$0{,}128\,56 ... \approx 0{,}129$
je práve 1 biela	$\binom{4}{1}\cdot 0{,}4\cdot 0{,}6^3 = 0{,}345\,6 \approx 0{,}346$	$0{,}346\,29 ... \approx 0{,}346$
sú práve 2 biele	$\binom{4}{2}\cdot 0{,}4^2\cdot 0{,}6^2 = 0{,}345\,6 \approx 0{,}346$	$0{,}346\,87 ... \approx 0{,}347$
sú práve 3 biele	$\binom{4}{3}\cdot 0{,}4^3\cdot 0{,}6 = 0{,}153\,6 \approx 0{,}154$	$0{,}153\,13 ... \approx 0{,}153$
sú práve 4 biele	$0{,}4^4 = 0{,}025\,6 \approx 0{,}026$	$0{,}025\,13 ... \approx 0{,}025$

6. b) V zápise (*) každý zo 490 činiteľov v súčine B má hodnotu približne 0,4 a každý zo 760 činiteľov v súčine C má hodnotu približne 0,6 (porovnaj tiež s poznámkou pred úlohou 5 na s. 61).

7. Na výber 1 250 guliek sa môžeme pozerať ako na 1 250-násobné opakovanie pokusu, v ktorom pravdepodobnosť úspechu je $p = 0{,}4$.

a) Počet úspešných pokusov – teda počet bielych guliek – prakticky iste leží v intervale $(\mu - 2\sigma,\ \mu + 2\sigma)$, kde $\mu = n \cdot p = 1\,250 \cdot 0{,}4 = 500$, $\sigma = \sqrt{n \cdot p(1-p)} = \sqrt{1\,250 \cdot 0{,}4 \cdot 0{,}6} = \sqrt{300}$, pozri obr. 2 na s. 41, $\mu - 2\sigma = 465{,}358\ldots$, $\mu + 2\sigma = 534{,}641\ldots$, preto počet vytiahnutých bielych guliek bude prakticky isto medzi 466 a 534.

b) Očakávanú relatívnu početnosť môžeme

- vypočítať z výsledku v časti a): bude to medzi $\frac{466}{1\,250} = 0{,}372\,8$ a $\frac{534}{1\,250} = 0{,}427\,2$,
- alebo použiť odhad z obr. 3 na s. 43 – relatívna početnosť bielych guliek sa prakticky isto od hodnoty $p = 0{,}4$ líši o menej ako $\pm 2 \cdot \frac{\sqrt{p(1-p)}}{\sqrt{n}} = \pm 2 \cdot \frac{\sqrt{0{,}4 \cdot 0{,}6}}{\sqrt{1\,250}} = 0{,}027\,7\ldots$.

8. a) Keby početnosť bielych guliek v urne bola 40 %, tak 431 *bielych z* 1250 *vybraných guliek* **ne**patrí medzi prakticky isté výsledky, vieme to z riešenia úlohy 7a). Odchýlka výsledku 431 od očakávanej ideálnej hodnoty $1\,250 \cdot 0{,}4 = 500$ je teda tak veľká, že už ju pokladáme za prakticky nemožnú. Preto môžeme odôvodnene predpokladať, že relatívna početnosť bielych guliek v urne **nie** je 40 %.

b) Relatívna početnosť bielych guliek v urne sa rovná pravdepodobnosti úspechu v zodpovedajúcom bernoulliovskom pokuse. Prirodzené je za približnú hodnotu tejto pravdepodobnosti – teda za približnú hodnotu relatívnej početnosti bielych guliek v urne – považovať relatívnu početnosť vo výbere, teda číslo $\frac{431}{1\,250} = 0{,}344\,8$. Z textu na s. 49 vieme, že skutočná hodnota hľadanej početnosti (za predpokladu, že výsledok 431 *bielych z* 1250 *vybratých guliek* pokladáme za prakticky istý) je

$$\frac{431}{1\,250} \pm \frac{1}{\sqrt{1\,250}} = 0{,}344\,8 \pm 0{,}028\,284\ldots = 0{,}34 \pm 0{,}03$$

teda medzi 31 % a 37 % (približnú hodnotu pravdepodobnosti aj odhad chyby sme zaokrúhlili tak, ako je to opísané za úlohou 8 na s. 49).

c) Nezmení, odpoveď závisí iba od počtu vyberaných guliek, ktorý zodpovedá počtu opakovaní v príslušných bernoulliovských pokusoch.

d) 10 000 alebo viacej. Chceme, aby platilo $\frac{1}{\sqrt{n}} \leq 0{,}01$. Rovnosť $\frac{1}{\sqrt{n}} = 0{,}01$, t. j. $\frac{1}{\sqrt{n}} = \frac{1}{100}$ nastáva pre $n = 100^2 = 10\,000$.

LITERATÚRA

[1] Edwards, A. W. F.: Pascal‘s Arithmetical Triangle: The Story of a Mathematical Idea. John Hopkins University Press : Baltimore, 2002. ISBN 0801869463, 9780801869464.

[2] Hald, A.: A History of Probability and Statistics and Their Applications Before 1750. John Wiley & Sons : Hoboken, 2005. ISBN 047172517X, 9780471725176.

[3] Janko, J.: Jak vytváří statistika obrazy světa a života, I. díl. Jednota československých matematiků a fysiků : Praha, 1947.

[4] McQuay, H. J., Moore, R. A.: Using Numeric Results from Systematic Reviews in Clinical Practice. *Annals of Internal Medicine*, 1 May 1997. 126:712-720.

[5] Stigler, S. M.: The history of statistics: the measurement of uncertainty before 1900. Harvard University Press, 1986. ISBN 067440341X.

[6] Stigler, S. M.: Statistics on the table: the history of statistical concepts and methods. Harvard University Press, 2002. ISBN 0674009797.