

Elektronická podpora pre časť: Štatistické výpočty v MATLABe

Doplnenia a komentár textu skript, riešenia niektorých príkladov a úloh

Doplnenia textu skript sú pre lepšiu orientáciu označené symbolmi ♠(začiatok) ♣(koniec)

Kap. 3 Diskrétné pravdepodobnostné rozdelenia

- 3.1 Model náhodného pokusu
- 3.2 Diskrétné náhodné veličiny, nezávislosť
- 3.3 Číselné charakteristiky diskretného rozdelenia
- 3.4 Špeciálne diskretné rozdelenia
- 3.5 Doplnenia a úlohy

V predchádzajúcich kapitolách sme si uvedomili, že väčšina veličín, ktoré v živote stretávame, je náhodnej povahy. V tejto kapitole budeme hovoriť o tom, ako matematickými prostriedkami modelujeme situácie, v ktorých máme do činenia s náhodou. Oboznámime sa (v zjednodušenej podobe) s modelom, ktorý v tridsiatych rokoch minulého storočia navrhol A.N. Kolmogorov a ktorý sa odvtedy stal všeobecne akceptovaným.

3.1 Model náhodného pokusu

Pod náhodným pokusom sa obyčajne rozumie pokus, ktorého výsledok nie je možné predpovedať s absolútnou istotou. V takomto voľnom zmysle je náhodným pokusom prakticky každá situácia, a preto predpokladajme, že máme na mysli len také pokusy, pri ktorých

- vieme vymedziť podmienky, za ktorých sa pokus koná a vymenovať, resp. inak určiť všetky (logicky možné) výsledky pokusu,
- nie je možné s istotou predpovedať výsledok, ktorým pokus skončí,
- je možné (aspoň teoreticky) pokus veľa razy opakovať pri dodržaní stanovených podmienok,
- výsledkom pokusu, resp. náhodným udalostiam v ňom, je možné priradiť pravdepodobnosť.

Pre model náhodného pokusu sú podstatné tri skutočnosti:

- 1) množina výsledkov, ktorými môže pokus skončiť,
- 2) systém náhodných udalostí v pokuse,
- 3) pravdepodobnosť, ktorá je 'spojená' s náhodným pokusom.

Prvým naším cieľom je porozumieť, ako modelujeme tieto tri skutočnosti. Začneme s príkladmi veľmi jednoduchých náhodných pokusov:

- P1: hod mincou (napr. slovenskou korunovou mincou),
- P2: hod pripínáčikom (konkrétnym typom),
- P3: hod hracou kockou (bežnou kockou, ktorá má 6 stien, je symetrickou a z homogénneho materiálu),
- P4: hod sfalšovanou hracou kockou, keď sme do bežnej (pôvodnej) vmontovali olovený brok,
- P5: hod tromi mincami, ktoré sme označili tak, že ich vieme rozlíšiť,
- P6: trojnásobný hod jednou mincou,
- P7: hod dvomi hracími kockami: modrou a červenou (obe sú riadne, nie sú falšované),
- P8: hod dvomi hracími kockami: modrou a červenou (modrá je riadna, červená je sfalšovaná),
- P9: hod dvomi hracími kockami: modrou a červenou (obe sú sfalšované),
- P10: náhodné ťahanie dvoch výrobkov z krabice, v ktorej je 5 výrobkov prvej kvality, 3 druhej kvality a 2 výrobky tretej kvality.

Zatiaľ žiaden z pokusov nebol poriadne popísaný, lebo sme nepovedali jasne, čo chápeme pod *výsledkom* pokusu. Snáď pre prvé štyri je to jasné, pre určitosť povedzme, že:

- P1: výsledkom je 'Z', alebo 'Č', teda znak, alebo číslo; množina možných výsledkov má len dva prvky,
- P2: výsledkom je 'D', alebo 'H', (pripínáčik skončí hrotom Dole k podložke, alebo hrotom Hore),
- P3: čísla, počet bodov na vrchnej stene (po dokotúľaní kocky): 1, 2, 3, 4, 5 a 6
- P4: množina možných výsledkov je rovnaká ako v pokuse P3
- P5: množina všetkých usporiadaných trojíc: (Z, Z, Z) (Z, Z, Č) (Z, Č, Z) ... (Z, Č, Č) (Č, Č, Č)
- P6: počet padnutých znakov, t.j. množina možných výsledkov má 4 prvky: 0, 1, 2, 3
- P7: výsledkom nech je usporiadaná dvojica (i, j) , kde i je číslo na modrej, j číslo na červenej kocke
- P8: výsledky chápeme rovnako ako v pokuse P7
- P9: výsledky chápeme rovnako ako v pokuse P7
- P10: výsledkom je neusporiadaná dvojica $\{i, j\}$, $i, j \in \{1, 2, 3\}$, kde i, j sú kvality vytiahnutých výrobkov.

Vo formálnom modeli je množina všetkých možných výsledkov náhodného pokusu označovaná symbolom Ω . Prvá úloha nás vyzýva stanoviť množiny možných výsledkov, t.j. množinu Ω , pre všetky náhodné pokusy P1 až P10.



Prvky množiny Ω označujeme ako ω_i , takže pre

P1: $\Omega = \{\omega_1, \omega_2\}$, $\omega_1 = 'Z'$, $\omega_2 = 'Č'$, (pokus má dva možné výsledky),

P2: $\Omega = \{\omega_1, \omega_2\}$, $\omega_1 = 'D'$, $\omega_2 = 'H'$, aj tento náhodný pokus má len dva výsledky,

P3: $\Omega = \{\omega_1, \omega_2, \dots, \omega_6\}$, $\omega_1 = '1'$, $\omega_2 = '2'$, $\omega_3 = '3'$, ..., $\omega_6 = '6'$, teraz máme 6 výsledkov

P4: $\Omega = \{\omega_1, \omega_2, \dots, \omega_6\}$, $\omega_1 = '1'$, $\omega_2 = '2'$, $\omega_3 = '3'$, ..., $\omega_6 = '6'$, opäť 6 (tých istých výsledkov).

To, čo odlišuje pokus P1 od P2, resp. pokus P3 od P4 je pravdepodobnosť, ktorá je s nimi spojená. Intuitívne je zrejmé, že kým v P1 oba možné výsledky majú rovnakú pravdepodobnosť, pri pokuse s pripínáčikom je to inak (analogicky je to s dvojicou P3 a P4). O pravdepodobnosti budeme hovoriť neskôr, najprv treba porozumieť tomu, ako modelujeme výsledky pokusu a náhodné udalosti v ňom.

P5: $\Omega = \{\omega_1, \omega_2, \dots, \omega_8\}$, $\omega_1 = (Z, Z, Z)$, $\omega_2 = (Z, Z, Č)$, $\omega_3 = (Z, Č, Z)$, $\omega_4 = (Č, Z, Z)$,
 $\omega_5 = (Z, Č, Č)$, $\omega_6 = (Č, Z, Č)$, $\omega_7 = (Č, Č, Z)$, $\omega_8 = (Č, Č, Č)$.

Zložky trojice znamenajú výsledky na prvej, na druhej a na tretej minci, napr. $\omega_6 = (Č, Z, Č)$ znamená: na prvej a na tretej padne číslo a pritom na druhej znak.

P6: $\Omega = \{\omega_0, \omega_1, \omega_2, \omega_3\}$, pričom ω_i znamená: počet padnutých znakov v pokuse sa rovná i .

P7: $\Omega = \{\omega_1, \omega_2, \dots, \omega_{36}\}$, $\omega_1 = (1, 1)$, $\omega_2 = (1, 2)$, $\omega_3 = (1, 3)$, ..., $\omega_{35} = (6, 5)$, $\omega_{36} = (6, 6)$. Štvorcovú mrežu 6×6 indexujeme po riadkoch zľava doprava a zhora dole.

P8 aj P9 majú rovnakú množinu Ω ako je v P7. Pre pokusy P7, P8, P9 sa dá povedať, že hoci majú rovnakú množinu možných výsledkov (a preto aj rovnakú množinu náhodných udalostí, čo hneď vysvetlíme), ide o tri rôzne pokusy, pretože pravdepodobnosť je v každom z nich iná (túto skutočnosť zatiaľ počítujeme len na intuitívnej úrovni).

P10 $\Omega = \{\omega_1, \omega_2, \dots, \omega_6\}$, $\omega_1: (1-1)$, $\omega_2: (1-2)$, $\omega_3: (1-3)$, $\omega_4: (2-2)$, $\omega_5: (2-3)$, $\omega_6: (3-3)$. Napr. výsledok ω_1 znamená, že oba vytiahnuté sú prvej kvality a ω_5 znamená, že vytiahnuté boli jeden druhej a jeden tretej kvality.

Úloha (nová :-). Stanovte množinu Ω možných výsledkov pokusu P11, ktorý je takou modifikáciou P10, keď sa zo škatule náhodne ťahajú tri výrobky (inak všetko ostáva ako v P10).



Na strane 135 je vysvetlené, ako modelujeme náhodné udalosti v pokuse. Ďalšia úloha nás vyzýva, aby sme v pokuse P5 prirodzeným jazykom popísali udalosti, ktoré sú vyjadrené cez množinové operácie.

Úloha. Uvažujte náhodný pokus P5 (hod tromi označenými mincami). Nech A je udalosť spočívajúca v tom, že na prvej minci bude 'Z', nech B je udalosť, že na tretej minci bude 'Z'. Jazykom udalostí popíšte udalosť $C = A \cap B$, resp. udalosť $D = A \cup B$. Predstavte C, resp. D ako podmnožiny množiny Ω .



$C = A \cap B$ znamená udalosť: "na prvej minci bude znak a súčasne na tretej znak". Poznamenávame, že o druhej minci sa nič nehovorí, a preto $C = \{\omega_1, \omega_3\}$, $\omega_1 = (Z, Z, Z)$, $\omega_3 = (Z, Č, Z)$.

$D = A \cup B$ znamená udalosť: "alebo na prvej minci znak, alebo na tretej minci znak". A to znamená, že D pozostáva z výsledkov: $\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_7$. Pripomeňme si ich:

$\omega_1 = (Z, Z, Z)$, $\omega_2 = (Z, Z, Č)$, $\omega_3 = (Z, Č, Z)$, $\omega_4 = (Č, Z, Z)$, $\omega_5 = (Z, Č, Č)$, $\omega_7 = (Č, Č, Z)$.

Preto $D = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_7\}$.



Úloha. V pokusoch P1 až P10 zvolte náhodné udalosti, reprezentujte ich ako podmnožiny a vytvárajte z nich nové udalosti operáciami komplementu, prieniku a zjednotenia.

Túto úlohu tu riešiť nebudeme. Je naozaj ľahká. Upozorňujeme, že zvládnuť tvorbu nových udalostí pomocou množinových operácií je nevyhnutné pre pochopenie toho, čo nasleduje.

Záver článku 3.1 sa týka pravdepodobnosti. Na nasledujúcej strane sa pokúsime povedať inými slovami niečo z toho, čo objasňuje, približuje, pojem pravdepodobnostného priestoru. Predpokladáme, že

prvé dva symboly Ω a $\mathcal{U}$ sú už vysvetlené. Ω reprezentuje možné výsledky a $\mathcal{U}$ náhodné udalosti v pokuse. Po formálnej stránke:

Ω je konečná množina (samozrejme, neprázdna).

$\mathcal{U}$ je systém (sústava) všetkých podmnožín množiny Ω .

Tretím symbolom v pravdepodobnostnom priestore $(\Omega, \mathcal{U}, P)$ je pravdepodobnosť P . Intuitívne slovo pravdepodobnosť zväzujeme s náhodnou udalosťou – teda nehovoríme o "pravdepodobnosti" samej o sebe, ale o pravdepodobnosti nejakej udalosti. A preto pravdepodobnosť je pre bežného človeka *číslo*. Pre nás však pravdepodobnosť P bude *zobrazenie*, ktorého definičný obor je systém všetkých udalostí v pokuse a hodnotami sú tie "čísla". Teda pre nás (t.j. v Kolmogorovom modeli náhodného pokusu)

$$P: \mathcal{U} \rightarrow [0, 1]$$

$$A \mapsto P(A)$$

Ak A je náhodná udalosť, tak jej pravdepodobnosť $P(A)$ chápeme ako mieru toho, ako často udalosť A nastáva. Naša predstava o frekvencii nastávania udalosti sa opiera:

- 1) buď o dedukciu
- 2) alebo o empiriu

V prvom prípade ide o situácie, v ktorých výsledky považujeme za "rovnako možné", vďaka nejakým predpokladom o geometrických, či fyzikálnych vlastnostiach objektov, ktorými realizujeme pokus. Napr. ak ide o obyčajnú hraciu kocku, tak ide o pravidelnú kocku z homogénneho materiálu, a tak všetkých 6 výsledkov považujeme za rovnako možné. Činíme tak bez toho, aby sme pokus veľa razy opakovali. Jednoducho nevidíme dôvod, pre ktorý by sa nejaký výsledok mal objavovať častejšie ako iný. A ak ide o pravdepodobnosť A spočívajúcu v tom, že kocka ukáže aspoň 5 bodov, tak $P(A) = 2/6$ (čo sa komentuje slovami: všetkých možných výsledkov je 6 a dva z nich sú *priaznivé* pre udalosť A).

Niekedy ale takto postupovať nemôžeme. Ak napr. ide o pokus s pripínáčikom, tak tam si žiadne deduktívne usudzovanie nevieme predstaviť. Ak ale máme možnosť hod pripínáčikom veľa razy opakovať a empiricky zistiť "frekvenciu" objavovania sa výsledku "D" (t.j. hrotom k podložke), tak získame "predstavu" o tom, čo je $P(D)$. Napr. ak sme pokus realizovali 1000 razy a v 327 prípadoch nastala udalosť D , tak náš *odhad* čísla $P(D)$ sa rovná 0,327.

Ak sa musíme odvolávať na empiriu, tak je jasné, že pozorovaná frekvencia je závislá od konkrétnej postupnosti realizácií a tá je, samozrejme, vždy iná. Akceptujeme existenciu čísla $P(D)$, ktoré je limitou postupnosti pomerných početností (čo je pomerná početnosť je opísané na strane 136).

Predchádzajúce dva odstavce chceli priblížiť obsah pojmu pravdepodobnosť a tak – chtiac, nechtiac – hovorili sme o tom, ako pravdepodobnosť stanovujeme, ako hľadáme $P(A)$, či $P(D)$. Ale teraz je podstatné uvedomiť si, že prichádzajúca definícia pravdepodobnosti nám nedáva návod, ako pravdepodobnosť stanoviť, ako určiť jej hodnoty $P(A)$, či $P(D)$. Definícia pravdepodobnosti deklaruje tie vlastnosti, ktoré pravdepodobnosť má v *každom* náhodnom pokuse.

Trochu zjednodušene povedané, v prípade jednoduchých náhodných pokusov (majúcich len konečne veľa výsledkov) ide o vlastnosť *aditivity*. Aditivitu – ako vlastnosť – od pravdepodobnosti požadujeme preto, lebo relatívne frekvencie túto vlastnosť majú. To je vysvetlené na strane 136. Teraz uveďme riešenie úlohy (aspoň jej časti) sformulovanej na strane 136 (celkom dole).



Stanovme $P(A)$, $P(B)$, $P(C)$ a $P(D)$ pre udalosti z pokusu P5 (hod tromi označenými mincami).

Pripomeňme, že $\omega_1 = (Z, Z, Z)$, $\omega_2 = (Z, Z, \checkmark)$, $\omega_3 = (Z, \checkmark, Z)$, $\omega_4 = (\checkmark, Z, Z)$,
 $\omega_5 = (Z, \checkmark, \checkmark)$, $\omega_6 = (\checkmark, Z, \checkmark)$, $\omega_7 = (\checkmark, \checkmark, Z)$, $\omega_8 = (\checkmark, \checkmark, \checkmark)$.

A ... na prvej znak, $A = \{\omega_1, \omega_2, \omega_3, \omega_5\}$, $P(A) = 0.5$,

B ... na tretej znak, $B = \{\omega_1, \omega_3, \omega_4, \omega_7\}$, $P(B) = 0.5$,

$C = A \cap B$, $C = \{\omega_1, \omega_3\}$, $P(C) = 0.25$,

$D = A \cup B$, $D = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_7\}$ $P(D) = 0.75$

Všimnime si, že platí: $P(A) + P(B) = P(A \cup B) + P(A \cap B)$

Tento vzorec má všeobecnú platnosť a je dôsledkom aditivity pravdepodobnosti.

Na záver tohoto článku uveďme formálnu definíciu pravdepodobnostného priestoru. Pozorný čitateľ si všimol, že sme nekomentovali modelovanie pravdepodobnosti v pokusoch P8 a P9 (a tam je to naozaj zaujímavé!). Urobíme tak v nasledujúcom článku.

Definícia. Diskrétne pravdepodobnostný priestor je trojica $(\Omega, \mathcal{U}, P)$, kde

Ω je konečná, resp. spočítateľná množina, predstavujúca všetky výsledky náhodného pokusu,

$\mathcal{U}$ je systém všetkých podmnožín množiny Ω , t.j. prvky $\mathcal{U}$ predstavujú náhodné udalosti v pokuse,

P je pravdepodobnosť, t.j. P je zobrazenie, ktoré udalostiam priradzuje ich pravdepodobnosť, pričom platí:

- $0 \leq P(\omega_i) \leq 1$ pre všetky i
- $\sum_{i=1}^n P(\omega_i) = 1$ ak Ω má n prvkov, resp. $\sum_{i=1}^{\infty} P(\omega_i) = 1$ (ak Ω je nekonečná)
- $P(A) = \sum_{i: \omega_i \in A} P(\omega_i)$ pre každú náhodnú udalosť A (aditivita pravdepodobnosti).

Poznámka. Zrejme $P(\emptyset) = 0$ a $P(\Omega) = 1$. Diskrétne pravdepodobnostný priestor jednoznačne určujú:

vektor výsledkov: $[\omega_1, \omega_2, \omega_3, \dots, \omega_n]$ (resp. $[\omega_1, \omega_2, \omega_3, \dots, \omega_n, \dots]$)
vektor ich pravdepodobností: $[p_1, p_2, p_3, \dots, p_n]$ (resp. $[p_1, p_2, p_3, \dots, p_n, \dots]$)

3.2 Diskrétne náhodné veličiny

Často v rámci náhodného pokusu nás zaujímajú také náhodné udalosti, ktoré majú, voľne povedané, číselné vyjadrenie. Napr. v rámci pokusu P5 nás môže zaujímať počet znakov, ktoré padli na minciach a nie aj to, na ktorých minciach znaky padli. Venujme tejto myšlienke vsuvku:



Uvažujme pokus P5, hod troch označenými mincami, o ktorom vieme, že má 8 možných, rovnako pravdepodobných výsledkov. Nech X znamená počet znakov, ktoré padli na minciach. Tabuľka pravdepodobnostnej funkcie má tvar:

x	0	1	2	3
$f(x)$	1/8	3/8	3/8	1/8

Táto tabuľka je prakticky zhodná s modelom pre pokus P6. Totiž pokus P6 môžeme reprezentovať (v zmysle poznámky za definíciou $(\Omega, \mathcal{U}, P)$) vektorom výsledkov a vektorom pravdepodobností:

$$\begin{aligned} &[\omega_0, \omega_1, \omega_2, \omega_3] \\ &[p_0, p_1, p_2, p_3] \end{aligned}$$

kde $p_0 = p_3 = 1/8$, $p_1 = p_2 = 3/8$. Vezmime na zreteľ, že je len na nás a na okolnostiach, kvôli ktorým pokus modelujeme, čo prehlásime za výsledky pokusu, alebo čo chápeme ako číselnú funkciu definovanú na bodoch nejakého pravdepodobnostného priestoru (na bodoch množiny Ω).

"Indikátor 6-ty" (v pokuse P3) je náhodná veličina X , ktorej pravdepodobnostná funkcia je daná tabuľkou:

x	0	1
$f(x)$	5/6	1/6

Pripomeňme, čo je indikátor. Je to náhodná veličina definovaná (teraz ako indikátor 6-ty) opisom:

X má hodnotu 1 práve vtedy, ak kocka ukáže 6-tku a X nadobudne hodnotu 0 práve vtedy, keď v pokuse nepadne číslo 6.

Vo všeobecnosti, indikátor sa môže týkať ľubovoľnej náhodnej udalosti! Indikátor udalosti A je náhodná veličina definovaná takto:

X má hodnotu 1 práve vtedy, ak udalosť A nastane (t.j. X nadobudne hodnotu 0 práve vtedy, keď A nenastane). Indikátor je teda náhodná veličina, ktorej hodnotová množina $H(X) = \{0, 1\}$.



Teraz na prenesieme na stranu 139 a budeme hovoriť o nezávislosti dvoch diskrétnych náhodných veličín. Táto diskusia nám pomôže pochopiť modelovanie fenoménu *nezávislosť*, ktorý pozorný čitateľ intuitívne cítil pri uvažovaní o pravdepodobnostnom modeli pokusov P8 a P9.

Nezávislosť dvojice diskretných náhodných veličín

Intuitívne sme o závislosti, resp. nezávislosti hovorili v článku 1 Kap. 2 (väzba znakov farba očí a farba vlasov). Teraz, na pôde teórie pravdepodobnosti, môžeme hovoriť konkrétnejšie. Ale, aby bolo jasné – romieť budú iba tí, ktorí porozumeli argumentácii v článku 2.1 kapitoly 2.



Uvažujme dve diskretné náhodné veličiny V a O . Nech $H(V) = \{1, 2\}$, $H(O) = \{1, 2, 3\}$. Nech vektor $p = [p_1, p_2]$ predstavuje rozdelenie veličiny V a vektor $q = [q_1, q_2, q_3]$ predstavuje rozdelenie veličiny O . Tabuľka spoločného rozdelenia nech je (zatiaľ) neznáma a má tvar:

$V \setminus O$	1	2	3	
1				p_1
2				p_2
	q_1	q_2	q_3	

Položme si otázku: Aké pravdepodobnosti majú vystupovať v poli 2×3 (ktoré je zatiaľ prázdne), aby sme mohli povedať, že veličiny sú nezávislé?

V zmysle toho, čo bolo povedané v článku 2.1, kap.2 vieme, že nezávislosť znamená, že pravdepodobnosti v riadku pre $V = 1$ majú byť v rovnakom pomere, v akom sú v poslednom riadku, (v ňom sú čísla: q_1, q_2, q_3). Teda v riadku pre $V = 1$ majú vystupovať čísla: $a \cdot q_1, a \cdot q_2, a \cdot q_3$ (koeficient úmernosti a je zatiaľ neznámy).

Z aditivity pravdepodobnosti vieme, že platí:

$$p_1 = P(V = 1) = P(V = 1, O = 1) + P(V = 1, O = 2) + P(V = 1, O = 3),$$

a preto:

$$p_1 = P(V = 1) = P(V = 1, O = 1) + P(V = 1, O = 2) + P(V = 1, O = 3) = a \cdot q_1 + a \cdot q_2 + a \cdot q_3 = a \cdot (q_1 + q_2 + q_3) = a$$

(lebo zrejme $q_1 + q_2 + q_3 = 1$).

Vďaka aditivite sme získali a , $a = p_1$. Ak majú byť veličiny V a O nezávislé, tak v riadku pre $V = 1$ musia vystupovať hodnoty $p_1 \cdot q_1, p_1 \cdot q_2, p_1 \cdot q_3$. Teda tabuľka musí vyzeráť takto:

$V \setminus O$	1	2	3	
1	$p_1 q_1$	$p_1 q_2$	$p_1 q_3$	p_1
2				p_2
	q_1	q_2	q_3	

Rovnako

$$p_2 = P(V = 2) = P(V = 2, O = 1) + P(V = 2, O = 2) + P(V = 2, O = 3) = b \cdot q_1 + b \cdot q_2 + b \cdot q_3 = b \cdot (q_1 + q_2 + q_3) = b$$

čo znamená, že konštanta úmernosti pre druhý riadok, $b = p_2$. Ak majú byť veličiny V a O nezávislé (v zmysle, v akom sme o ňom hovorili v čl.1, kap.2), tak tabuľka spoločného rozdelenia je

$V \setminus O$	1	2	3	
1	$p_1 q_1$	$p_1 q_2$	$p_1 q_3$	p_1
2	$p_2 q_1$	$p_2 q_2$	$p_2 q_3$	p_2
	q_1	q_2	q_3	

Teraz už zrejme akceptujeme nasledujúcu definíciu nezávislosti dvojice diskretných veličín. Stačí si uvedomiť, že len označenie je iné, napr.:

$$[p_1, p_2] = [f_1(x_1), f_1(x_2)] \quad \text{a} \quad [q_1, q_2, q_3] = [f_2(y_1), f_2(y_2), f_2(y_3)].$$



Definícia. Diskretné náhodné veličiny sú *stochasticky nezávislé* práve vtedy, ak pre ich pravdepodobnostné funkcie platí:

$$f(x, y) = f_1(x) \cdot f_2(y) \quad \text{pre všetky reálne } x, y.$$

(definícia je formulovaná tak, aby obsiahla aj prípady, keď veličiny majú nekonečne veľa hodnôt).

Ak ide o dvojicu nezávislých veličín X, Y tak je možné stanoviť spoločné rozdelenie, t.j. funkciu $f(x, y)$, na základe znalostí "jednorozmerných" funkcií $f_1(\cdot)$ a $f_2(\cdot)$. V prípade, že veličiny sú nezávislé, jednoducho kladieme

$$f(x, y) = f_1(x) \cdot f_2(y), \text{ pre všetky } x, y.$$

Niekedy totiž nezávislosť je evidentná zo samotnej réžie náhodného pokusu, ako je to v pokusoch P8 a P9, ktoré teraz rozoberieme.



P8: hod dvomi hracími kockami: modrou a červenou (modrá je riadna, červená je sfalšovaná).

Na rozdiel od prístupu z úvodu tejto kapitoly, teraz sa pozerať na náhodný pokus tak, že hodnoty na modrej kocke chápeme ako hodnoty veličiny X a hodnoty na červenej kocke ako hodnoty veličiny Y . Hodnoty oboch veličín sú rovnaké: 1, 2, ..., 6.

Podľa popisu náhodného pokusu, pravdepodobnostné rozdelenie veličiny X určuje vektor

$$p = [1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6]$$

a povedzme, že rozdelenie veličiny Y je dané vektorom q :

$$q = [1/12 \quad 1/12 \quad 1/12 \quad 3/12 \quad 3/12 \quad 3/12]$$

vektor q predstavuje (reprezentuje) spôsob, akým bola červená kocka sfalšovaná. Pokus spočíva v hode dvomi kockami a je zrejmé, že pri bežnom vykonávaní pokusu sa kocky kotúľajú každá "po svojom" a teda veríme, že nezávisle. Preto tabuľka pravdepodobnostnej funkcie vektora (X, Y) je

$X \setminus Y$	1	2	3	4	5	6	
1	1/72	1/72	1/72	3/72	3/72	3/72	12/72
2	1/72	1/72	1/72	3/72	3/72	3/72	12/72
3	1/72	1/72	1/72	3/72	3/72	3/72	12/72
4	1/72	1/72	1/72	3/72	3/72	3/72	12/72
5	1/72	1/72	1/72	3/72	3/72	3/72	12/72
6	1/72	1/72	1/72	3/72	3/72	3/72	12/72
	6/72	6/72	6/72	18/72	18/72	18/72	



P9: hod dvomi hracími kockami: modrou a červenou (obe sú sfalšované).

Teraz použijeme MATLAB. Falošnou je teraz aj modrá, a preto nech pravdepodobnostné vektory sú dané napr. takto:

$$p = [1/12 \quad 3/12 \quad 1/12 \quad 1/12 \quad 3/12 \quad 3/12]$$

$$q = [1/12 \quad 1/12 \quad 1/12 \quad 3/12 \quad 3/12 \quad 3/12]$$

$$p =$$

0.0833	0.2500	0.0833	0.0833	0.2500	0.2500
--------	--------	--------	--------	--------	--------

$$q =$$

0.0833	0.0833	0.0833	0.2500	0.2500	0.2500
--------	--------	--------	--------	--------	--------

Maticu spoločného rozdelenia dostaneme operáciou násobenia stĺpca p' s riadkom q :

$$\text{mat} = p' * q$$

$$\text{mat} =$$

0.0069	0.0069	0.0069	0.0208	0.0208	0.0208
0.0208	0.0208	0.0208	0.0625	0.0625	0.0625
0.0069	0.0069	0.0069	0.0208	0.0208	0.0208
0.0069	0.0069	0.0069	0.0208	0.0208	0.0208
0.0208	0.0208	0.0208	0.0625	0.0625	0.0625
0.0208	0.0208	0.0208	0.0625	0.0625	0.0625

Keď už máme spoločné rozdelenie, ľahko vieme zodpovedať všetky možné otázky, napríklad:

Aká je pravdepodobnosť, že súčet bodov na kockách bude aspoň 9?

$$P = P((3, 6), (4, 5), (4, 6), (5, 4), (5, 5), (5, 6), (6, 3), (6, 4), (6, 5), (6, 6)) = 0.0208 + 0.0208 + 0.0208 + \\ + .0625 + 0.0625 + 0.0625 + 0.0208 + 0.0625 + 0.0625 + 0.0625 = 0.4582$$



Úlohy

Úlohy **3.5.1, 3.5.2** sú jednoduché. Venujme sa ďalším úlohám.

3.5.3 Uvažujme pokus P8, keď červená kocka je zhodná s tou falošnou, ktorú ste zvolili v úlohe 3.5.1. Nech X je počet bodov na modrej, Y je počet bodov na červenej. Je zrejmé, že X, Y sú nezávislé (to vyplýva z režie pokusu, keď každá veličina sa týka inej kocky). Nech $Z = X + Y$. Určte pravdepodobnostnú funkciu g veličiny Z . Nájdite $E(Z)$, $var(Z)$.

Riešenie. Teraz môžeme nadviazať na to, čo sme napísali na predchádzajúcej strane. Modrej kocke odpovedá pravdepodobnostný vektor

$$p = [1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6 \quad 1/6]$$

a predpokladajme, že červenej vektor

$$q = [1/12 \quad 1/12 \quad 1/12 \quad 3/12 \quad 3/12 \quad 3/12]$$

Ako sme už uviedli, tabuľka spoločného rozdelenia je takáto:

$X \setminus Y$	1	2	3	4	5	6	
1	1/72	1/72	1/72	3/72	3/72	3/72	12/72
2	1/72	1/72	1/72	3/72	3/72	3/72	12/72
3	1/72	1/72	1/72	3/72	3/72	3/72	12/72
4	1/72	1/72	1/72	3/72	3/72	3/72	12/72
5	1/72	1/72	1/72	3/72	3/72	3/72	12/72
6	1/72	1/72	1/72	3/72	3/72	3/72	12/72
	6/72	6/72	6/72	18/72	18/72	18/72	

Rozdelenie náhodnej veličiny $Z = X + Y$ získame sčítaním po vedľajších diagonálach. Objasníme to na prípade určenia $P(Z = 5)$. Zrejme

$$\begin{aligned} P(Z = 5) &= P(\{(X, Y) = (1, 4)\} \cup \{(X, Y) = (2, 3)\} \cup \{(X, Y) = (3, 2)\} \cup \{(X, Y) = (4, 1)\}) = \\ &= P((X, Y) = (1, 4)) + P((X, Y) = (2, 3)) + P((X, Y) = (3, 2)) + P((X, Y) = (4, 1)) = \\ &= 3/72 + 1/72 + 1/72 + 1/72 = 0.0833 \end{aligned}$$

Tabuľka pravdepodobnostnej funkcie g náhodnej veličiny Z má tvar:

z	2	3	4	5	6	7	8	9	10	11	12
$g(z)$	1/72	2/72	3/72	6/72	9/72	12/72	11/72	10/72	9/72	6/72	3/72

(overte, že súčet v druhom riadku sa rovná 1).

3.5.4 Uvažujme pokus P9, keď obe kocky sú falošné a nech sú zhodné s tou, ktorú ste zvolili v úlohe 3.5.1. Nájdite odpovede na otázky, ktoré sme položili v predchádzajúcej úlohe!

Návrh: Postupujte za pomoci MATLABu tak, ako sme to ukázali na predchádzajúcej strane. Pravdepodobnostné vektory môžete voliť ľubovoľne, avšak, musí ísť o pravdepodobnostné vektory. Teda o vektory kladných čísel, pričom súčet zložiek musí byť rovný 1.

3.5.5 Uvažujte o náhodnom vektore (X, Y) , ktorého tabuľka pravdepodobnostnej funkcie má tvar:

$X \setminus Y$	-1	0	1
-1	0	0.25	0
0	0.25	0	0.25
1	0	0.25	0

Nájdite pravdepodobnostné rozdelenia zložiek. Ukážte, že X, Y sú *nekorelované*, t.j. $cov(X, Y) = 0$. Je ľahké vidieť, že pritom X, Y nie sú nezávislé! Vezmime na vedomie, že implikácia:

Ak X, Y sú nezávislé, potom sú nekorelované, lebo $E(X \cdot Y) = E(X) \cdot E(Y)$, a preto $cov(X, Y) = 0$

platí vždy, ale naša situácia ilustruje, že opačná implikácia (vo všeobecnosti) neplatí!