PRAVDEPODOBNOSŤ a ŠTATISTIKA

Náhodné udalosti a pravdepodobnosť

doc. RNDr. Štefan Peško

Katedra matematických metód a operačnej analýzy, FRI ŽU

26. februára 2019

Náhodný pokus a náhodná udalosť

Náhodným pokusom rozumieme taký pokus, ktorý môžeme ľubovoľne veľa krát opakovať, pričom jeho výsledok nie je určený jednoznačne. Pri každom opakovaní pokusu je jeho výsledkom elementárna udalosť.

Neprázdnu množinu všetkých elementárnych udalostí označíme Ω a nazývame výberový priestor. Náhodnou udalosťou rozumieme ľubovoľnú množinu elementárnych udalostí t.j. náhodná udalosť $U \subset \Omega$.

Poznámka: Po realizácii náhodného pokusu musí nutne nastať práve jedna elementárna udalosť z množiny Ω pričom jej prvky musia byť navzájom nezlučiteľné (nemôžu nastať naraz viacero rôznych výsledky pokusu). O výsledku náhodného pokusu rozhoduje náhoda.

Príklad 1.1 (Skúška z algebry)

Výsledky študentov môžeme interpretovať ako realizácie náhodného pokusu - termínu skúšky. Elemetárnymi udalosťami sú tu získané známky študentov a tak $\Omega = \{A, B, C, D, E, FX\}$. Náhodnou udalosťou *úspech* rozumejú :-) študenti známky $\{A, \ldots, E\} \subset \Omega$.

	A	B	C	D	E	FX
1.	1	1	5	3	1	4
2.	0	1	1	4	3	9
3.	0	0	2	0	5	6
4.	0	1	2	5	16	15
5.	0	0	1	0	4	14
6.	0	0	0	1	10	9
7.	0	0	0	1	12	9
8.	0	0	0	0	9	11

Tabuľka: Výsledky z 8 termínov skúšky z algebry v ZS 2016

Majme náhodné udalosti $A, B, A^c \subset \Omega$, povieme že:

- $A^c = \Omega - A$ je doplnkovou udalosťou k udalosti A, t.j. udalosť A^c nastane práve vtedy keď nenastane udalosť A,
- $A \subset B$ ak udalosť B je dôsledkom udalosti A t.j. z nastatia udalosti A vyplýva nastatie udalosti B,
- $A = B$ ak udalosti A a B sú si rovné t.j. platí $A \subset B$ a súčasne $B \subset A$.
- $\emptyset$ je nemožná udalosť,
- Ω je istá udalosť,

Poznámka: Z definície náhodnej udalosti je zrejmé:

a) $\forall A \subset \Omega$ platí: $\emptyset \subset A, A \subset A, (A^c)^c = A$.

b) $\forall A, B, C \subset \Omega$ platí: ak $A \subset B, B \subset C$ potom $A \subset C$.

c) $\emptyset^c = \Omega, \Omega^c = \emptyset$.

Operácie s náhodnými udalosťami

Nech $A_1, A_2, \ldots, A_n$ sú náhodné udalosti. Potom

- zjednotením (súčtom) udalostí rozumieme takú udalosť A, ktorá nastane práve vtedy keď nastane aspoň jedna z udalosti $A_1, A_2, \ldots, A_n$; čo označujeme

$$A = \bigcup_{i=1}^{n} A_i.$$

- prienikom (súčinom) udalostí rozumieme takú udalosť A, ktorá nastane práve vtedy keď súčasne nastanú všetky udalosti $A_1, A_2, \ldots, A_n$; čo označujeme

$$A = \bigcap_{i=1}^{n} A_i.$$

Hovoríme, že udalosti A a B sú disjunktné (nezlučiteľné) ak to že nastanú súčasne je udalosť nemožná, $A \cap B = \emptyset$.

Príklad 1.2

Nech je udalosť A v tom, že náhodne vybrané prirodzené číslo je deliteľné 5 a udalosť B je v tom, že má na poslednom mieste 0.

Čo znamenajú udalosti:

$C = A \cap B$? $A \cap B = B$ a tak udalosť C nastane, keď bude vybrané číslo končiť nulou.

$D = A^c \cap B$? Udalosti A^c a B sú nezlučiteľné a tak je udalosť D nemožná, $D = \emptyset$.

$E = A \cup B^c$? Udalosť E sa nastane ak je vybrané číslo deliteľné 5 alebo nemá na konci 0 čomu vyhovujú všetky prirodzené čísla takže $E = \Omega$.

$F = (A \cap B^c)^c$? Platí $(A \cap B^c)^c = A^c \cup B$ a tak udalosť F nastane keď vybrané číslo nebude deliteľné piatimi ale bude mať na konci nulu t.j. všetky prirodzené čísla okrem tých, ktoré majú na poslednom mieste 5.

Príklad 1.3 (Test súčiastky)

Dovážaná súčiastka je skúšaná troma testami. Udalosť $A_i, i = 1, 2, 3$ spočíva v tom, že súčiastka vyhovie v i-tom teste.

Ako vyjadríme, v množinovej symbolike, že súčiastka vyhovie:

a) len v prvom teste ? $A_1 \cap A_2^c \cap A_3^c$;

b) vo všetkých troch testoch ? $A_1 \cap A_2 \cap A_3$;

c) aspoň v dvoch testoch ? $(A_1 \cap A_2) \cup (A_1 \cap A_3)) \cup (A_2 \cap A_3)$;

d) práve v jednom teste ?

$$(A_1 \cap A_2^c \cap A_3^c) \cup (A_1^c \cap A_2 \cap A_3^c) \cup (A_1^c \cap A_2^c \cap A_3);$$

e) práve v dvoch testoch ?

$$(A_1 \cap A_2 \cap A_3^c) \cup (A_1 \cap A_2^c \cap A_3) \cup (A_1^c \cap A_2 \cap A_3);$$

f) maximálne v dvoch testoch? $(A_1 \cap A_2 \cap A_3)^c$.

Štatistická a klasická definícia pravdepodobnosti

Predpokladajme, že vykonáme sériu n pokusov v ktorej sa náhodná udalosť A vyskytne n_A krát. Potom štatistickou pravdepodobnosťou udalosti A rozumieme číslo

$$\mathcal{P}(A) = \lim_{n\to\infty} \frac{n_A}{n},$$

kde $\frac{n_A}{n}$ udáva relatívnu početnosť náhodnej udalosti A.

Nech je Ω konečná neprázdna množina rovnako možných elemetárnych udalostí. Potom klasickou pravdepodobnosťou náhodnej udalosti $A \subset \Omega$ rozumieme číslo

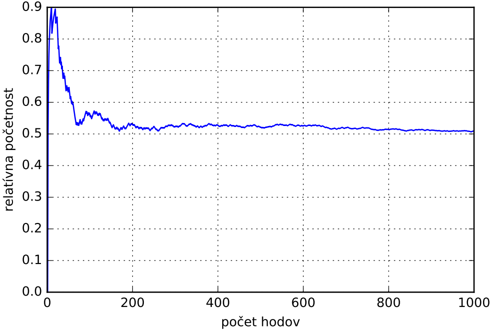
$$\mathcal{P}(A) = \frac{|A|}{|\Omega|}.$$

Poznámka: Zásadný rozdiel medzi štatistickou a klasickou definíciou pravdepodobnosti (pp.) je v tom, že pri použití klasickej definície pp. rozhodujeme o pp. pred pokusom zatiaľ čo pri štatistickej pp. až po ňom.

Príklad 1.4 (Štatistická pravdepodobnosť)

Hádžeme mincou 1000 krát a sledujeme udalosť A, že padne znak.

Možné výsledky hodu mincou sú $\Omega = \{\text{znak}, \text{čislo}\}$. Označme n_A počet udalostí keď v sérii n hodov padne znak.



Vidíme, že po 1000 hodoch je relatívna početnosť výskytu znaku $\frac{n_A}{n}$ blízka očakávanej štatistickej pravdepodobnosti $\frac{1}{2}$.

Príklad 1.5 (Klasická pravdepodobnosť)

Hádžeme hracou kockou. Aká je pravdepodobnosť, že

- a) pri jednom hode padne 6 ?
  Možné výsledky hodu sú elementárne udalosti z $\Omega = \{1, 2, 3, 4, 5, 6\}$, pp. udalosti A, že padne 6 je $\mathcal{P}(A) = \frac{|A|}{|\Omega|} = \frac{1}{6}$;
- b) pri jednom hode padne párne číslo ?
  Možné výsledky hodu sú náhodné udalosti z $B = \{2, 4, 6\}$, pp. udalosti B je $\mathcal{P}(B) = \frac{|B|}{|\Omega|} = \frac{3}{6} = \frac{1}{2}$;
- c) po dvoch hodoch padne 2 krát 6?
  Výberový priestor je tvorený usporiadanými dvojicami $\Omega = \{(1, 1), (1, 2), \dots, (6, 5), (6, 6)\}$. Priaznivým výsledkom hodu je len udalosť $C = \{(6, 6)\}$, a tak pp. udalosti C je $\mathcal{P}(C) = \frac{|C|}{|\Omega|} = \frac{1}{36}$;

Príklad 1.6

Aká je pravdepodobnosť, že pri jednom hode troma kockami bude súčet bodov 11, a aká je pravdepodobnosť, že tento súčet bude 12?

Chevalier de Méré (1607–1684) pozoroval pri hre s kockami, že častejšie padá súčet 11 než súčet 12 aj keď podľa jeho názoru majú oba súčty rovnakú pravdepodobnosť. Vychádzal z chybnej úvahy, že súčty 11 aj 12 majú priaznivých 6 výsledkov:

$$11 = 4+4+3 = 4+5+2 = 4+6+1 = 5+3+3 = 5+5+1 = 6+3+2$$
$$12 = 4+4+4 = 4+5+3 = 4+6+2 = 5+5+2 = 6+5+1 = 6+3+3.$$

Obrátil sa preto na veľmi slavného matematika Blaisa Pascala (1623–1662), ktorý ho upozornil na to, že uvedené možnosti nie sú rovnako pravdepodobné. Napr. vo výsledku $4+4+3$ sú zahrnuté aj výsledky $4+3+4$ a $3+4+4$. Možných výsledkov je $6 \cdot 6 \cdot 6 = 216$ a tak jeho pp. je $\frac{3}{216}$. Overte že pp. hodu súčtu 11 je $\frac{27}{216}$ a hodu súčtu 12 je $\frac{25}{216}$, čo je už v zhode s de Mérého pozorovaním.

Nech Ω je podmnožina n-rozmerného Euklidovského priestoru s konečnou mierou μ a $G \subset \Omega$. Potom pravdepodobnosť, že náhodne zvolený bod bodovej množiny Ω je bodom množiny G rozumieme číslo

$$\mathcal{P}(A) = \frac{\mu(G)}{\mu(\Omega)},$$

kde $\mu(G)$ je miera množiny G a $\mu(\Omega)$ je miera množiny Ω.

Príklad 1.7 (Prerušenie telefónneho spojenia)

Medzi miestami K a L vzdialenými 10 km bolo prerušené telefónne spojenie. Aká je pravdepodobnosť, že miesto poruchy bolo vzdialené od miesta K najviac 500 m?

Označme A udalosť, že miesto poruchy je nanajvýš 500 m vzdialené od miesta K. Množina všetkých možných prerušení medzi miestami K a L je tvorená bodmi množiny Ω. Množina vyhovujúcich miest poruchy G je tvorená množinou bodov, ktoré sú od miesta K vzdialené najviac 500 m.

Mierou množín je dĺžka resp. vzdialenosť. Potom

$$\mathcal{P}(A) = \frac{\mu(G)}{\mu(\Omega)} = \frac{l(G)}{l(\Omega)} = \frac{500}{10000} = \frac{1}{20}.$$

Príklad 1.8 (Kancelária)

Kolegovia Karol, Ľuboš a Martin prídu do spoločnej kancelárie medzi ôsmou a štrnástou hodinou nezávisle na sebe. Vieme, že každý z týchto troch kolegov sa zdrží v kancelárii presne hodinu. Vypočítajte pravdepodobnosť, že: a) Karol sa stretne s Ľubošom. b) Všetci traja sa stretnú. c) Aspoň jedna dvojica kolegov sa stretne.

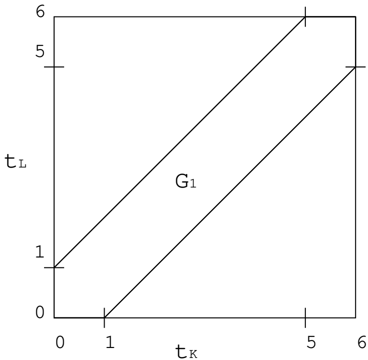
Oznanačme postupne t_K, t_L, t_M príchody Karola, Ľuboša a Martina do kancelárie po 8:00.

a) Označme A_1 náhodnú udalosť, že Karol sa stretne s Ľubošom. Potom množina všetkých možných situácií je

$$\Omega_1 = \{(t_K, t_L) \in \Re^2 : t_K, t_L \in (0, 6)\}.$$

Aby sa tretli musia byť rozdiely medzi časmi príchodu v absolútnej hodnote menšie ako 1 t.j.

$$G_1 = \{(t_K, t_L) \in \Omega_1 : |t_K - t_L| < 1\}.$$



Z obrázku vidíme, že plocha G_1 je rovná ploche štvorca $S(\Omega_1)$ od ktorej odčítame plochu dvoch pravouhlých trojuholníkov t.j.

$S(G_1) = 36 - 2 \cdot \frac{5 \cdot 5}{2} = 11$. A tak $\mathcal{P}(A_1) = \frac{S(G_1)}{S(\Omega_1)} = \frac{11}{36}$.

b) Označme A_2 náhodnú udalosť, že traja kolegovia sa stretnú. Potom množina všetkých možných situácií je

$$\Omega_2 = \{(t_K, t_L, t_M) \in \Re^3 : t_K, t_L, t_M \in (0, 6)\}.$$

Aby sa všetci traja stretli, ak pobudnú v kancelárii hodinu, musí platiť, že rozdiely medzi príchodmi každej dvojice je nanajvýš 1 t.j.

$$G_2 = \{(t_K, t_L, t_M) \in \Omega_2 : |t_K - t_L| < 1, |t_K - t_M| < 1, |t_L - t_M| < 1\}.$$

Mierou je objem a tak $\mathcal{P}(A_2) = \frac{V(G_2)}{V(\Omega_2)} = \frac{???}{6^3}$.

Objem množiny G_2 už nevypočítame pomocou obrázku ale pomocou trojného integrálu. Nerovnice z definície G_2 si rozdelíme do 6-tich objemovo zhodných oblastí podľa usporiadania príchodov t_K, t_L, t_M: $t_K < t_L < t_M$, $t_K < t_M < t_L$, $t_L < t_K < t_M$, $t_L < t_M < t_K$, $t_M < t_K < t_L$, $t_M < t_L < t_K$.

Ak je usporiadanie $t_K < t_L < t_M$ potom $t_M - t_K < 1$ teda $t_M < t_K + 1$. Ale $t_M < 6$ a takže $t_M < min(6, t_K + 1)$ a tak $t_M \in (t_K, min(6, t_K + 1))$.
Teda

$$V = \int_0^6 \int_{t_k}^{min(6, t_K+1)} \int_{t_K}^{t_M} \mathrm{d}t_L \mathrm{d}t_M \mathrm{d}t_K = V_1 + V_2,$$

$$\begin{aligned} V_1 &= \int_0^5 \int_{t_K}^{t_K+1} \int_{t_K}^{t_M} \mathrm{d}t_L \mathrm{d}t_M \mathrm{d}t_K = \int_0^5 \int_{t_K}^{t_K+1} (t_M - t_K) \mathrm{d}t_M \mathrm{d}t_K \\ &= \int_0^5 \left[\frac{t_M^2}{2} - t_K t_M\right]_{t_K}^{t_K+1} \mathrm{d}t_K = \\ &= \int_0^5 \left[\frac{(t_K + 1)^2}{2} - t_K(t_K + 1) - \left(\frac{t_K^2}{2} - t_K^2\right)\right] \mathrm{d}t_K = \cdots = \frac{5}{2}. \end{aligned}$$

$$V_2 = \int_5^6 \int_{t_K}^6 \int_{t_K}^{t_M} \mathrm{d}t_L \mathrm{d}t_M \mathrm{d}t_K = \int_5^6 \int_{t_K}^6 (t_M - t_K) \mathrm{d}t_M \mathrm{d}t_K$$

$$= \int_5^6 \left[\frac{t_M^2}{2} - t_K t_M\right]_{t_K}^6 \mathrm{d}t_K = \int_0^6 \left[\frac{6^2}{2} - 6t_K - \left(\frac{t_K^2}{2} - t_K^2\right)\right] \mathrm{d}t_K$$

$$= \int_5^6 \left[\frac{t_K^2}{2} - 6t_K + \frac{6^2}{2}\right] \mathrm{d}t_K = \left[\frac{t_K^3}{6} - 6\frac{t_K^2}{2} + \frac{6^2 t_K}{2}\right]_5^6 = \cdots = \frac{1}{6}.$$

Potom $V = V_1 + V_2 = \frac{5}{2} + \frac{1}{6} = \frac{8}{3}$ a tak

$$\mathcal{P}(A_2) = \frac{V(G_2)}{V(\Omega_2)} = \frac{6\frac{8}{3}}{6^3} = \frac{2}{27}.$$

c) Označme A_3 náhodnú udalosť, že aspoň jedna dvojica sa stretne. Jej pp. vypočítame pomocou už vypočítaných pp. Každá dvojica je rovnako pravdepodobná s pp. $\mathcal{P}(A_1) = \frac{11}{36}$ no pripúšťa aj možnosť stretnutia všetkých kolegov. Preto

$\mathcal{P}(A_3) = 3\mathcal{P}(A_1) - 2\mathcal{P}(A_2) = 3\frac{11}{36} - 2\frac{2}{3^3} = \frac{83}{108}.$

Pole udalostí

Nech je daný neprázdny výberový priestor Ω a systém jeho podmnožín $\mathcal{F}$. Tento systém podmnožín nazývame pole udalostí ak sú splnené tieto podmienky:

- 1. $\emptyset \in \mathcal{F}, \Omega \in \mathcal{F}$.
- 2. Ak $A \in \mathcal{F}$ potom tiež $A^c \in \mathcal{F}$.
- 3. Ak $A_1, A_2, A_3, \dots \in \mathcal{F}$ potom tiež $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Majme výberový priestor Ω a jeho pole udalostí $\mathcal{F}$. Potom pravdepodobnosťou rozumieme ľubovolnú reálnu funkcia $\mathcal{P} : \mathcal{F} \to \Re$, ktoré má tieto vlastnosti:

Axióma 1. Pre každé $A \in \mathcal{F}$ platí $0 \leq \mathcal{P}(A) \leq 1$.

Axióma 2. $\mathcal{P}(\Omega) = 1$.

Axióma 3. Pre ľubovoľné $A_1, A_2, A_3, \dots \in \mathcal{F}$ pre ktoré platí $A_i \cap A_j = \emptyset, 1 \leq i < j$ je

$$\mathcal{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathcal{P}(A_i).$$

Usporiadanú trojicu $(\Omega, \mathcal{F}, \mathcal{P})$ potom nazývame (Kolmogorov) pravdepodobnostný priestor.

Vlastnosti pravdepodobnosti

Tvrdenie 1

Nech $(\Omega, \mathcal{F}, \mathcal{P})$ je pravdepodobnostný prietor. Potom platí:

a) $\mathcal{P}(\emptyset) = 0$.

b) Pre každú $A \in \mathcal{F}$ je $\mathcal{P}(A) \leq 1, \mathcal{P}(A^c) = 1 - \mathcal{P}(A)$.

c) Ak $A, B \in \mathcal{F}$ také že $A \subseteq B$, tak $\mathcal{P}(A) \leq \mathcal{P}(B)$.

d) Ak $A, B \in \mathcal{F}$ potom platí
$\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A \cap B)$.

e) Nech $A_i \in \mathcal{F}$ pre $i = 1, 2, \ldots$ Potom platí
$$\mathcal{P}\Big(\bigcup_{i=1}^{\infty} A_i\Big) \leq \sum_{i=1}^{\infty} \mathcal{P}(A_i).$$

f) Nech $A_i \in \mathcal{F}$ pre $i = 1, 2, \ldots, n$. Potom platí
$$\mathcal{P}\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{i=1}^{n} \mathcal{P}(A_i) - \sum_{1 \leq i < j \leq n} \mathcal{P}(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq n} \mathcal{P}(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-1} \mathcal{P}\Big(\bigcap_{i=1}^{n} A_i\Big).$$

Príklad 1.9

Hádžeme dvoma hracími kockami. Výberový priestor $\Omega = \{(i,j) : i,j = 1,2,\dots,6\}$ má 36 prvkov.

1. Aká je pp., že na aspoň jednej kocke padne 6 ?
   Označme A resp. B udalosť, že na 1. resp. 2. kocke padne 6.
   $A = \{(6,k) : k = 1,2,\dots,6\}, B = \{(k,6) : k = 1,2,\dots,6\},$
   $\mathcal{P}(A) = \frac{1}{6}, \mathcal{P}(B) = \frac{1}{6}, \mathcal{P}(A \cap B) = \frac{1}{36},$
   $\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A \cap B) = \frac{11}{36}.$
2. Aká je pp., že na kockách padne súčet 10 alebo rozdiel 1 ?
   Označme C udalosť, že na kockách padne súčet 10 a D udalosť, že na kockách padne rozdiel čísel 1.
   $C = \{(i,j) \in \Omega : i + j = 10\}, D = \{(i,j) \in \Omega : |i - j| = 1\},$
   $\mathcal{P}(C) = \frac{3}{36} = \frac{1}{12}, \mathcal{P}(D) = \frac{10}{36} = \frac{5}{18}, \mathcal{P}(C \cap D) = \frac{0}{36} = 0,$
   $\mathcal{P}(C \cup D) = \mathcal{P}(C) + \mathcal{P}(D) = \frac{13}{36}.$

Príklad 1.10

Hádžeme troma hracími kockami. Aká je pravdepodobnosť, že aspoň na jednej kocke padne 6 ?

Označme udalosť A_i = „na i-tej kocke padne 6“. Očividne ;-)

$$\mathcal{P}(A_1) = \mathcal{P}(A_2) = \mathcal{P}(A_3) = \frac{1}{6},$$

$$\mathcal{P}(A_1 \cap A_2) = \mathcal{P}(A_1 \cap A_3) = \mathcal{P}(A_2 \cap A_3) = \frac{1}{6 \cdot 6},$$

$$\mathcal{P}(A_1 \cap A_2 \cap A_3) = \frac{1}{6 \cdot 6 \cdot 6}.$$

Podľa tvrdenia 1f) dostaneme

$$\mathcal{P}(A_1 \cup A_2 \cup A_3) = \mathcal{P}(A_1)+\mathcal{P}(A_2)+\mathcal{P}(A_3)-\mathcal{P}(A_1 \cap A_2)-\mathcal{P}(A_1 \cap A_3)$$

$$-\mathcal{P}(A_2 \cap A_3) + \mathcal{P}(A_1 \cap A_2 \cap A_3) = 3\frac{1}{6} - 3\frac{1}{36} + \frac{1}{216} = \frac{91}{216}.$$

Cvičenie

1.1 Pri hode tromi mincami vypočitajte pp. náhodnej udalosti $A_i, (i = 0, 1, 2, 3)$, že padne i-krát znak.

1.2 Aká je pp., že štvordetná rodina má zhodné zastúpenie dievčat a chlapcov? [Pre jednoduchosť predpokladajte, že dievčatá i chlapci sa rodia s rovnakou pp. (čo nie je pravda :-)]

1.3 V triede s x chlapcami a y dievčatami sú losom náhodne vybraní dvaja hovorcovia. Aká je pp., že budú rovnakého pohlavia?

*1.4 V sérii finálových zápasov dvoch družstiev vyhráva to družstvo, ktoré vyhrá 4 vzájomné zápasy. Aká je pp., že séria bude mať $4, 5, 6, 7$ zápasov? [Predokladajte, že o výhre súperov rozhoduje náhodný hod spravodlivou mincou]

*1.5 Na večierku je n ľudí. Niekto sa chce staviť o 100 eur, že medzi nimi existujú dvaja ľudia s narodeninami v ten istý deň. Pri akom minimálnom počte ľudí sa mu oplatí prijať takúto stávku pri 5% riziku neúspechu?