

Chí kvadrát test dobrej zhody

Diskrétné náhodné premenné

Príklad 1: Budeme sa zaoberať výsledkami, ktoré pri hádzaní kockou získal Walter F. L. Weldon (1860 – 1906). Ten tvrdil, že väčším počtom bodiek pri týchto hodnotách je spôsobená jej nepravidelnosť. Táto nepravidelnosť by mohla byť taká výrazná, že hodnoty 5 a 6 nebudú padať rovnako často ako ostatné.

Pri svojich pokusoch hádzal dvanástimi kockami. Ako “úspech” označil padnutie päťky alebo šestky a zaznamenával počet úspechov v jednotlivých pokusoch. Tých spravil spolu 26 306 s nasledujúcimi výsledkami:

úspechov	početnosť
0	185
1	1149
2	3265
3	5475
4	6114
5	5194
6	3067
7	1331
8	403
9	105
10	14
11	4
12	0
spolu	26306

Ak je minca pravidelná, tak počet úspechov v pokuse má binomické rozdelenie. Otestujeme chí kvadrát testom dobrej zhody, že hodnoty v tabuľke pochádzajú z tohto rozdelenia.

Poznámky:

- Kvôli požiadavke na minimálny počet očakávaných úspechov bude potrebné zlúčiť niekoľko posledných riadkov.
- Pravdepodobnosti binomického rozdelenia v R sa dajú vypočítať funkciou **dbinom**.
- Pre zaujímavosť: Tieto dáta analyzoval aj Peason, podľa ktorého je tento test dobrej zhody pomenovaný.

Chí kvadrát test dobrej zhody v R

Takéto ručné počítanie štatistiky je užitočné kvôli pochopeniu jej princípu, prakticky ale užitočné poznať funkciu, ktorá tento test spraví.

Pre príklad o kocke zo slajdov (či je pravdepodobnosť hodnôt 1, 2, 3, 4, 5, 6) rovnaká:

```
> chisq.test(c(85, 99, 91, 108, 119, 98), p = rep(1/6, 6))
```

Chi-squared test for given probabilities

data: c(85, 99, 91, 108, 119, 98)
X-squared = 7.36, df = 5, p-value = 0.1952

pozorované početnosti jednotlivých možností ich pravdepodobnosti pri platnosti testovanej hypotézy

hodnota štatistiky počet stupňov voľnosti p-hodnota

Všimnime si rovnakú hodnotu štatistiky ako vyšla “ručne”.

P-hodnota je pravdepodobnosť, že za platnosti testovanej hypotézy (tzv. nulovej hypotézy) “takáto zlá” hodnota štatistiky. Ak je táto pravdepodobnosť malá, tak to nasvedčuje tomu, že hypotéza neplatí. Porovnať p-hodnotu napríklad s 5 percentami je to isté, ako porovnať štatistiku s kritickou hodnotou získanou pre pravdepodobnosť 0,05.

Poznámka:

V tomto prípade je p-hodnota väčšia ako 0,05, takže na tejto hladine významnosti hypotézu nezamietame.

Defaultné pravdepodobnosti, ktoré **chisq.test** testuje, sú rovnaké pre všetky hodnoty. Takže v tomto prípade stačí spraviť:

```
> chisq.test(c(85, 99, 91, 108, 119, 98))
```

Chi-squared test for given probabilities

data: c(85, 99, 91, 108, 119, 98)
X-squared = 7.36, df = 5, p-value = 0.1952

Spojité náhodné premenné

Potrebuje dve veci:

- Spraviť tabuľku s početnosťami hodnôt zo zvolených intervalov.
- Pravdepodobnosti pri zadanom rozdelení, ktoré testujeme.
 - Ak ide o známe rozdelenie, ktoré je implementované v R, môžeme ich vypočítať pomocou distribučnej funkcie (funkcie typu **pnorm**, **pexp** a pod.).
 - Ak máme len zadanú hustotu, pravdepodobnosť získame pomocou integrálu, R vie numericky integrovať.

Tabuľka s početnosťami - napríklad:

```
> set.seed(123)
> x <- rexp(1000, rate = 1)
> table(cut(x, breaks = c(0, 0.5, 1, 1.5, 2, 2.5, 3, Inf)))
```

(0,0.5]	(0.5,1]	(1,1.5]	(1.5,2]	(2,2.5]	(2.5,3]	(3,Inf]
376	231	162	94	51	29	57

Pravdepodobnosť

Vypočítame pravdepodobnosť, že exponenciálne rozdelenie s parametrom 1 (teda aj so strednou hodnotou 1) nadobudne hodnotu z intervalu (1.5, 2), resp. z intervalu, do ktorého patrí niektorý z krajných bodov, alebo aj obidva:

```
> pexp(2, rate = 1) - pexp(1.5, rate = 1) # P(Exp(1) \in (1.5, 2))
[1] 0.08779488
```

Integrovanie:

```
> integrate(function(x) exp(-x), lower = 1.5, upper = 2)
0.08779488 with absolute error < 9.7e-16
> integrate(function(x) exp(-x), lower = 1.5, upper = 2)[[1]]
[1] 0.08779488
```

Príklad 2: Bertrandov paradox.

Z bakalárskej práce Kataríny Koscisovej, ktorú som viedla v roku 2017:

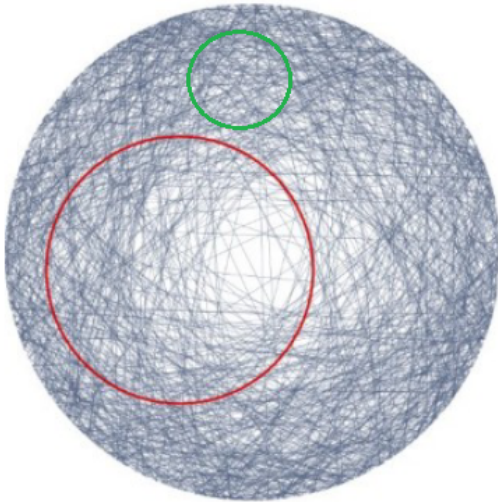
Názov:	Don't just solve it; fight it! Príklady z pravdepodobnosti <i>Don't just solve it; fight it! Probability problems</i>
Cieľ:	Téma práce je malou obmenou citátu Paula Halmosa o dôkazoch: "Don't just read it; fight it! Ask your own questions, look for your own examples, ..." Jeho myšlienka sa dá vzťahovať aj na riešenie príkladov. To sa nemusí skončiť vyriešením určitého príkladu, mnohé príklady dávajú možnosť zamýšľať sa nad nimi ďalej, klásť si otázky a hľadať na ne odpovede. Cieľom práce je venovať sa takýmto spôsobom niekoľkým úlohám z oblasti pravdepodobnosti: (a) Príklad 293 zo zbierky [1] a súvislosť získaných pravdepodobností s pravdepodobnosťami získanými z ratingov typu [3] (b) Známy Bertrandov paradox - porovnanie s experimentom popísaným v článku [4] (c) Úloha o spojeniach kariet z [3] - výpočet pravdepodobnostného rozdelenia

Paradox spočíva v riešení úlohy, ktorú Bertrand sformuloval v roku 1889:

Majme kruh, v ktorom náhodne zvolíme tetivu. Aká je pravdepodobnosť, že zvolená tetiva bude dlhšia ako strana rovnostranného trojuholníka vpísaného do daného kruhu?

Keďže "náhodne zvolíme tetivu" nie je jednoznačne popísané, môžeme ju voliť rôznym spôsobom, pričom môžeme dostať rôzne pravdepodobnosti. Môžeme si však predstaviť aj

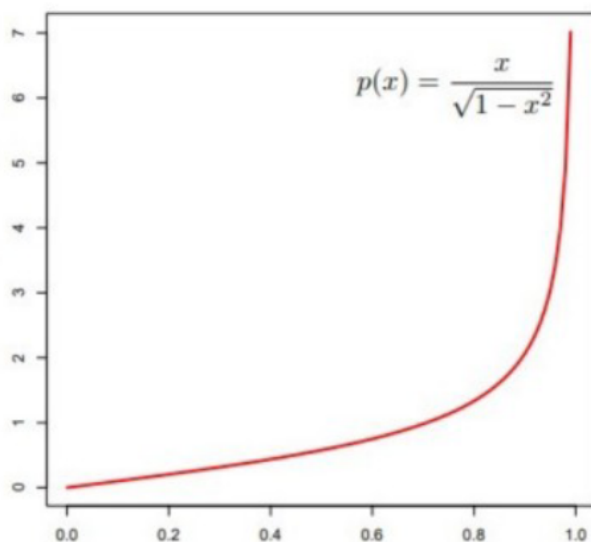
to, že fyzicky hádzame špajdlu na kružnicu nakreslenú na zemi a zaznamenávame dĺžku tetivy, ktorá takto vznikne. Otázka je, aká voľba “náhodného výberu tetivy” tomu zodpovedá. Napríklad takáto simulácia z wikipédie zrejme nevyhovuje, čakali by sme určitú pravidelnosť:



V článku Jaynesa (1973) sa ako podmienka dáva, že rozdelenie dĺžky tetivy (relatívne k polomeru kruhu) musí byť invariantné na polohu a veľkosť kruhu. V bakalárskej práci sa uvádza hustota, ktorú odvodil:

$$p(x) = \frac{x}{\sqrt{1-x^2}}, \quad (4.1.2)$$

pričom x predstavuje dĺžku normalizovanej tetivy, teda $x = \frac{L}{2 \cdot R}$ (L je dĺžka tetivy, R je polomer kruhu a $0 \leq x \leq 1$).



a tiež výsledky experimentov, v ktorých ľudia reálne hádzali na kružnicu špajdlu:

model. My sme sa rozhodli realizovať experiment popísaný v [15], teda hádzať slamku (v našom prípade špajdl'u) na kruh s priemerom 5 palcov = 12.7 cm. Výsledky podľa informácií z článku potvrdili nízkymi testovými štatistikami (v tomto prípade s veľkou jednoznačnosťou) rozdelenie tetív podľa T. Jaynesa, čo si teraz overíme realizáciou rovnakého experimentu. Tohto experimentu sa zúčastnili 3 ľudia, ktorí hádzali slamku zostoja na kruh nakreslený na papieri a položený na zemi. Výsledkom boli 4 pokusy so 128 meraniami. Získané merania sme následne normalizovali, teda predelili ich priemerom kruhu a rozdelili sme ich do 10 skupín.

Namerané hodnoty:

Číslo skupiny	Skupina tetív	Pokus č.1	Pokus č.2
		Početnosť	Početnosť
1	$(L/2R \leq 0.5)$	15	17
2	$(0.5 < L/2R \leq 0.6)$	7	15
3	$(0.6 < L/2R \leq 0.7)$	10	10
4	$(0.7 < L/2R \leq 0.8)$	13	14
5	$(0.8 < L/2R \leq 0.85)$	10	5
6	$(0.85 < L/2R \leq 0.9)$	11	11
7	$(0.9 < L/2R \leq 0.925)$	6	7
8	$(0.925 < L/2R \leq 0.95)$	6	9
9	$(0.95 < L/2R \leq 0.975)$	10	10
10	$(0.975 < L/2R \leq 1)$	40	30

Testom dobrej zhody otestujeme hypotézu, že výsledky pokusu pochádzajú z rozdelenia s horeuvedenou hustotou.

Cvičenia

Cvičenie 1.

Naprogramujte jedno kolo nasledovnej hry (to znamená, že ignorujeme inštrukciu “*he is also the starter of the new game*” – vždy sa odohrá len jedna hra). Spravte simulácie hry a zaznamenávajte, ktorý hráč vyhral (prvý, druhý alebo tretí). Testom dobrej zhody testujte hypotézu, že pravdepodobnosti víťazstva všetkých troch hráčov sú rovnaké.

Poznámka: Ak je počet dát veľmi veľký, aj malá odchýlka od teoretických pravdepodobností spôsobí, že sa hypotéza zamietne. Preto sa pri obrovských počtoch dát používajú iné

metódy, ako je tento test dobrej zhody. Pri simuláciách je preto tiež dobré neísť do “príliš” veľkého počtu simulácií (typu 100 000 alebo 1 000 000).

◆ DESCRIPTION OF THE GAME ◆

We first describe the game with three players. Each player contributes 7 units of stake into a “pool”. Thus, the total initial amount in the pool is 21 units. A player is randomly selected to start the first game. Each time six *Mang Kung Dice* are thrown simultaneously. The total face value of these dice is counted. This total face value (F) is then compared with the total amount in the pool (P). The possible outcomes and the corresponding player's actions are listed in Table 1.



Table 1. The three possible situations in the Mang Kung Dice Game.

Outcomes	Player's Actions
$F < P$	The Player can take the amount F from the pool. He then passes the dice to the next player on his left to continue the game.
$F > P$	The Player has to top up the pool by the difference, $(F - P)$. He then passes the dice to the next player to continue the game.
$F = P$	The Player wins the game. He receives the amount P from the pool. Moreover, each of the other players has to pay him an amount F as “bonus”. He is also the starter of the next new game.

Cvičenie 2. Spravte simulácie nasledovnej úlohy a do histogramu doplňte ponúkané hustoty. Vyberte si tú, ktorá sa vám zdá najbližšia k získaným dátam zo simulácie a otestujte zhodu dát s týmto rozdelením pomocou chí kvadrát testu dobrej zhody:

A company offers earthquake insurance. Annual premiums are modeled by an exponential random variable with mean 2. Annual claims are modeled by an exponential random variable with mean 1. Premiums and claims are independent. Let X denote the ratio of claims to premiums. What is the density function of X ?

- (A) $\frac{1}{2x+1}$ (B) $\frac{2}{(2x+1)^2}$ (C) e^{-x} (D) $2e^{-2x}$ (E) xe^{-x}