# Chí kvadrát test dobrej zhody

Metódy riešenia úloh z pravdepodobnosti a štatistiky
Beáta Stehlíková, FMFI UK

# Opakovanie:
# Chí kvadrát test dobrej zhody

- Chceme overiť, či naše dáta pochádzajú z **konkrétneho pravdepodobnostného rozdelenia.**
- Napríklad:
  - – Je kocka pravidelná? Padá na nej 1, ..., 6 rovnako často?
  - – Je pravdepodobnosť narodenia chlapca a dievčaťa rovnaká?

# Opakovanie:
# Chí kvadrát test dobrej zhody

- Príklad: *Môžeme kocku považovať za pravidelnú?*
- Výsledky 600 hodov:

| padnuté číslo | početnosť |
|---|---|
| 1 | 85 |
| 2 | 99 |
| 3 | 91 |
| 4 | 108 |
| 5 | 119 |
| 6 | 98 |

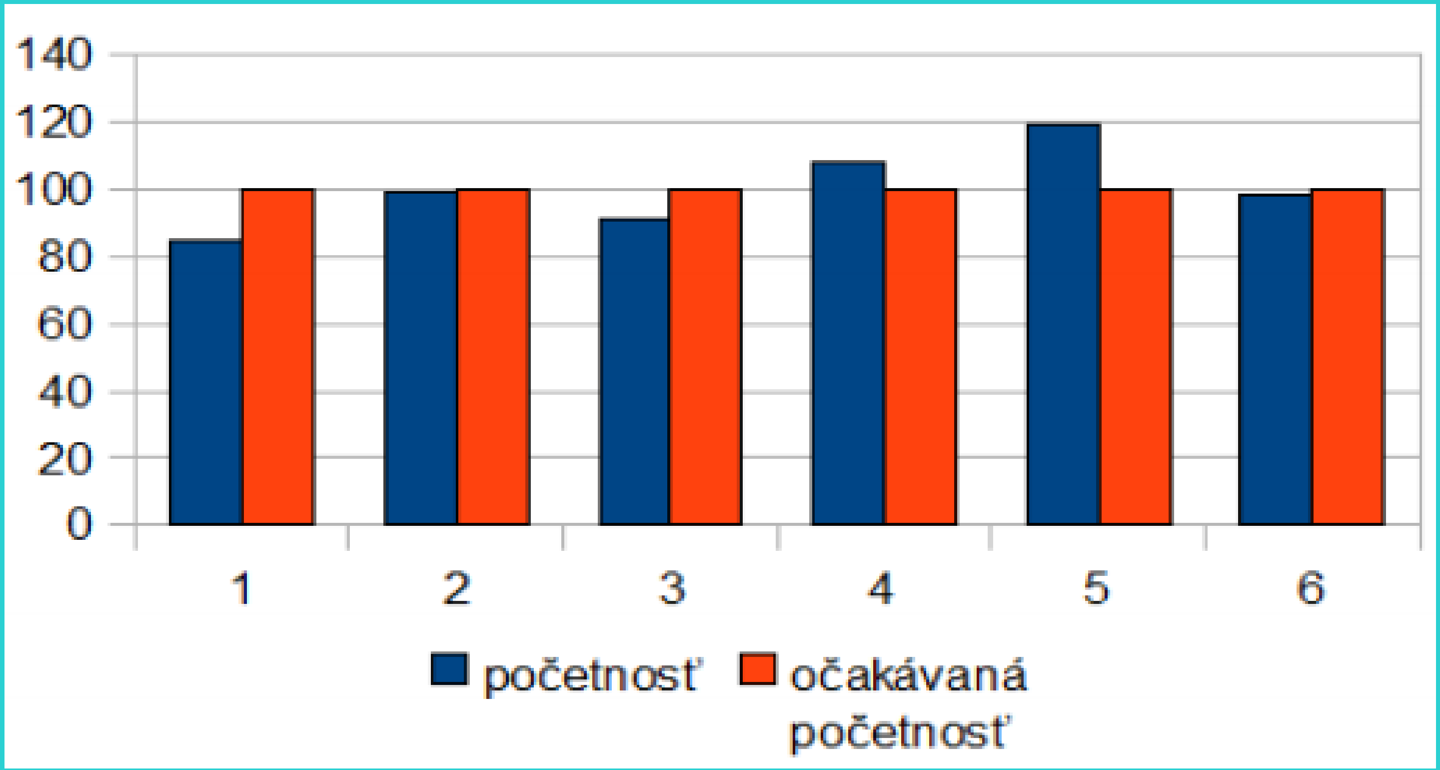
# Opakovanie:
# Chí kvadrát test dobrej zhody

- Aké sú očakávané početnosti:

| | A | B | C | D |
|---|---|---|---|---|
| 1 | padnuté číslo | početnosť | teoretická pravd. | očakávaná početnosť |
| 2 | 1 | 85 | 0,1667 | =+C2*$B$9 |
| 3 | 2 | 99 | 0,1667 | 100 |
| 4 | 3 | 91 | 0,1667 | 100 |
| 5 | 4 | 108 | 0,1667 | 100 |
| 6 | 5 | 119 | 0,1667 | 100 |
| 7 | 6 | 98 | 0,1667 | 100 |
| 8 | | | | |
| 9 | spolu | 600 | | |

# Opakovanie:
# Chí kvadrát test dobrej zhody

- Aké sú skutočné a očakávané početnosti:



# Opakovanie:
# Chí kvadrát test dobrej zhody

- Rozdiely medzi skutočnými a očakávanými početnosťami:

| početnosť | teoretická pravd. | očakávaná početnosť | početnosť – očakávaná |
|---|---|---|---|
| 85 | 0,1667 | 100 | -15 |
| 99 | 0,1667 | 100 | -1 |
| 91 | 0,1667 | 100 | -9 |
| 108 | 0,1667 | 100 | 8 |
| 119 | 0,1667 | 100 | 19 |
| 98 | 0,1667 | 100 | -2 |

# Opakovanie:
# Chí kvadrát test dobrej zhody

- Kladné a záporné odchýlky sa nemôžu navzájom zrušiť
- Rozdiel -15 nemá rovnakú váhu, keď je očakávaný počet 100 a keď je 1000

| | D | E | F |
|---|---|---|---|
| 1 | očakávaná početnosť | početnosť – očakávaná | (početnosť – očakávaná)^2/(očakávaná) |
| 2 | 100 | -15 | =+(E2^2)/D2 |
| 3 | 100 | -1 | 0,01 |
| 4 | 100 | -9 | 0,81 |
| 5 | 100 | 8 | 0,64 |
| 6 | 100 | 19 | 3,61 |
| 7 | 100 | -2 | 0,04 |

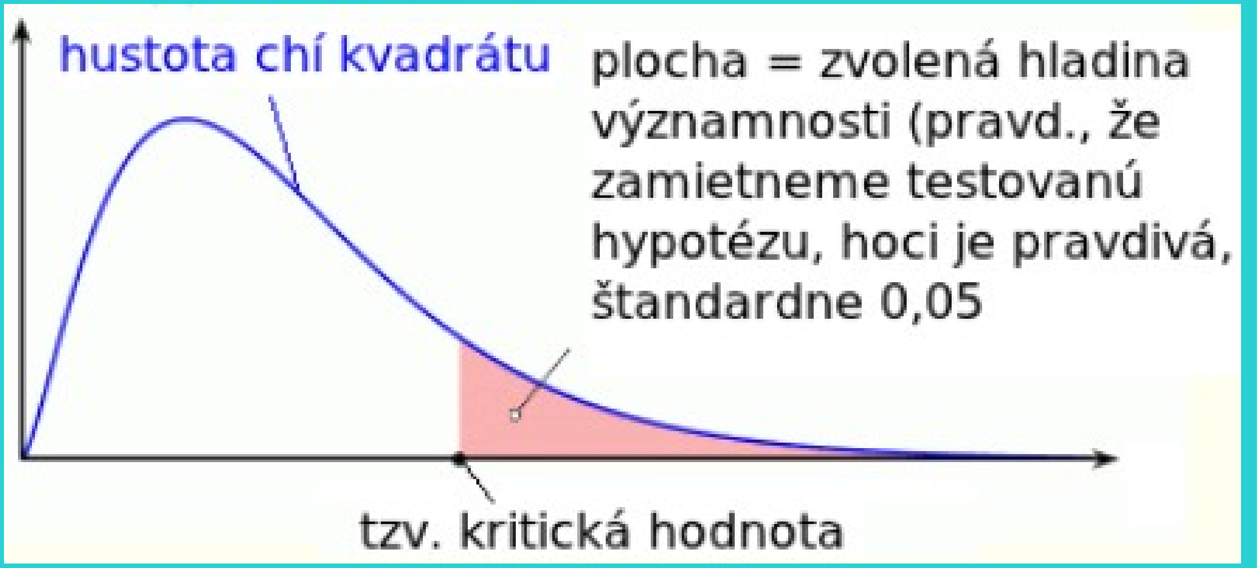
# Opakovanie:
# Chí kvadrát test dobrej zhody

- „Mierou rozdielu“ medzi dátami a testovaným rozdelením bude súčet týchto rozdielov
- Ak bude veľký, hypotézu zamietneme

| | E | F |
|---|---|---|
| tnosť | početnosť – očakávaná | (početnosť – očakávaná)^2/(očakávaná) |
| | -15 | 2,25 |
| | -1 | 0,01 |
| | -9 | 0,81 |
| | 8 | 0,64 |
| | 19 | 3,61 |
| | -2 | 0,04 |
| | | |
| | **spolu** | **=SUM(F2:F7)** |

# Opakovanie:
# Chí kvadrát test dobrej zhody

- Čo znamená „veľký“:
  - Väčší ako kvantil chí kvadrát rozdelenia s q-1 stupňami voľnosti, kde q je počet skupín, do ktorých sú dáta zaradené (teraz teda q=6)
  - Test sa dá použiť, ak je očakávaný počet dát v každej skupine aspoň 5



# Opakovanie:
# Chí kvadrát test dobrej zhody

- Kritická hodnota pre náš príklad:

```
> qchisq(0.95, df = 5)
[1] 11.0705
```

```
octave:2> chi2inv(0.95, 5)
ans = 11.070
```

```
=CHIINV(0,05;5)
```

- Záver:

| F |
|---|
| (početnosť – očakávaná)^2/(očakávaná) |
| 2,25 |
| 0,01 |
| 0,81 |
| 0,64 |
| 3,61 |
| 0,04 |
| |
| =SUM(F2:F7) |

7,36

- pravidelnosť kocky nezamietame

# Príklad 1: Weldonov pokus

- Walter F. L. Weldon (1860 – 1906)
  - Hodil 12 kockami, zopakoval to 26 306 krát
  - Úspech – padnutie 5 alebo 6
  - Teoretická pravdepodobnosť 1/3
  - Prakticky vyššia početnosť, pripisoval to konštrukcii bodiek, ktorá mala vplyv na váhu stien
- Budeme testovať hypotézu, že počet úspechov je Bin(12, 1/3)

# Príklad 1: Weldonov pokus

- Weldonove výsledky máme v tabuľke
- Niekoľko známych mien zo štatistiky:
  - Weldon poslal dáta v liste **Galtonovi**
  - Analyzoval ich aj **Pearson**, ktorý je autorom aj tohto chí kvadrát testu dobrej zhody

| 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 |

# Príklad 1: Weldonov pokus

- Dokončíme výpočet:

| úspechov | početnosť | očakávaná početnosť |
|---|---|---|
| 0 | 185 | 202,75 |
| 1 | 1149 | 1216,5 |
| 2 | 3265 | 3345,37 |
| 3 | 5475 | 5575,61 |
| 4 | 6114 | 6272,56 |
| 5 | 5194 | 5018,05 |
| 6 | 3067 | 2927,2 |
| 7 | 1331 | 1254,51 |
| 8 | 403 | 392,04 |
| 9 | 105 | 87,12 |
| 10 | 14 | 13,07 |
| 11 | 4 | 1,19 |
| 12 | 0 | 0,05 |
| spolu | 26306 | |

treba zlúčiť na "10 a viac" kvôli požiadavke na očakávaný počet

# Príklad 2: Farby M&M‘s

- Katarína Kocsisová, bakalárska práca, 2015

**Názov:** Experimenty a vlastné dáta pri vyučovaní pravdepodobnosti a štatistiky
*Experiments and own data in teaching probability and statistics*

**Cieľ:** Cieľom práce je spracovať "pokusy", ktoré sa dajú využiť pri vyučovaní pravdepodobnosti a štatistiky: experimenty motivujúce výpočty určitých pravdepodobností, získanie vlastných dát na testovanie hypotéz a pod. Konkrétne, práca sa bude zaoberať nasledovnými témami, ku ktorým sa pridá niekoľko podobných podľa vlastného výberu študenta:

(a) Vizualizácia distribučnej funkcie pomocou slovníka a testovanie zhody dvoch distribučných funkcií [1]

(b) Odhad veľkosti populácie (základná myšlienka: máme N guľôčok očíslovaných 1, 2, ..., N, vytiahneme n z nich a chceme na základe tohto výberu odhadnúť N) [2], [3]

(c) Hádzanie šípok a testovanie vplyvu dominantnej ruky [4]

(d) Hra s kockami Hog [5]

(e) Chí kvadrát test dobrej zhody a farby M&M's [6]

# Príklad 2: Farby M&M's

Naším cieľom bude štatisticky testovať, že pomery farieb v balíčkoch sú naozaj také, o akých hovoria vyjadrenia z rokov 1991 a 2007 spomínané vyššie.

Obr. 9: Testované cukríky

# Príklad 2: Farby M&M's

- Podiely jednotlivých farieb:

Málokto by sa však zrejme zamýšľal, či existuje nejaké pravidlo na rozdelenie cukríkov do jednotlivých balíčkov. Ešte donedávna zverejňovala spoločnosť Mars na svojej internetovej stránke percentuálne zastúpenie každej farby. Podľa článku [1], ktorý sa venuje testovaniu pomerov farieb $M\&M's$, balenie v roku 1991 obsahovalo: 30% hnedej, 20% žltej, 20% červenej a oranžovú, zelenú a žltohnedú po 10% z každej farby. Ako uvádza zdroj [4], prieskum spoločnosti Mars ukázal, že posledná menovaná farba je najmenej obľúbená, preto oslovili verejnosť, aby hlasovala za novú. Z vyše 10 miliónov ľudí sa takmer 55% vyjadrilo v prospech modrej. Žlhohnedú nahradila s totožným percentuálnym vyjadrením.

Článok [2] popisuje novšie tvrdenie spoločnosti Mars z roku 2007, ktoré uvádza nasledujúce pravdepodobnostné rozdelenie cukríkov: 13% hnedej, 14% žltej, 13% červenej, 24% modrej, 20% oranžovej a 16% zelenej.[1]

# Príklad 2: Farby M&M's

- Testovanie:

V nasledujúcej časti budeme skúmať tvrdenie o farbách a ich počtoch v balíčkoch na zozbieraných dátach. Tvrdenie spoločnosti Mars uvedené na začiatku kapitoly hovorí, že balíček obsahuje 30% hnedých, 20% žltých a červených, po 10% z modrých, oranžových a zelených cukríkov. Hypotézu $H_0$ teda zapíšeme: $p_1 = 0{,}3$, $p_2 = p_3 = 0{,}2$ a $p_4 = p_5 = p_6 = 0{,}1$.

Tabuľka 8: Dáta v balíčkoch podľa farieb

| | hnedá | žltá | červená | modrá | oranžová | zelená | spolu |
|---|---|---|---|---|---|---|---|
| balíček 1 | 12 | 52 | 35 | 46 | 18 | 59 | 222 |
| balíček 2 | 15 | 48 | 41 | 44 | 18 | 55 | 221 |
| spolu | 27 | 100 | 76 | 90 | 36 | 114 | 443 |

# Príklad 2: Farby M&M's

- Záver:

$$\chi^2 = \frac{(27-132,9)^2}{132,9} + \frac{(100-88,6)^2}{88,6} + \frac{(76-88,6)^2}{88,6} + \frac{(90-44,3)^2}{44,3} + \frac{(36-44,3)^2}{44,3} + \frac{(114-44,3)^2}{44,3} = 246,01 \quad (2)$$

Tabuľka 9: $\chi^2$ test (1991)

| farby | hnedá | žltá | červená | modrá | oranžová | zelená | spolu |
|---|---|---|---|---|---|---|---|
| očakávané | 132,9 | 88,6 | 88,6 | 44,3 | 44,3 | 44,3 | 443 |
| skutočné | 27 | 100 | 76 | 90 | 36 | 114 | 443 |
| $\frac{(X_i-Np_i)^2}{Np_i}$ | 84,39 | 1,47 | 1,79 | 47,14 | 1,56 | 109,66 | 246,01 |

Ako sme ukázali, $\chi^2 = 246,01 > \chi^2_{6-1}(0,05) = 11,07$, čiže hypotézu $H_0$ zamietame.

# Príklad 2: Farby M&M's

- Druhá sada pravdepodobností:

Následne overíme aj pravdivosť novšieho tvrdenia spomínaného v úvode kapitoly, kde hypotéza $H_0$ hovorí: $p_1 = 0{,}13$, $p_2 = 0{,}14$, $p_3 = 0{,}13$, $p_4 = 0{,}24$, $p_5 = 0{,}2$ a $p_6 = 0{,}16$. Postupujeme analogicky ako pri prvom testovaní. Potrebné údaje sú uvedené v tabuľke 10.

Tabuľka 10: $\chi^2$ test (2007)

| farby | hnedá | žltá | červená | modrá | oranžová | zelená | spolu |
|---|---|---|---|---|---|---|---|
| očakávané | 57,59 | 62,02 | 57,59 | 106,32 | 88,6 | 70,88 | 443 |
| skutočné | 27 | 100 | 76 | 90 | 36 | 114 | 443 |
| $\frac{(X_i - Np_i)^2}{Np_i}$ | 15,79 | 23,29 | 5,59 | 2,42 | 31,56 | 26,04 | 104,69 |

# Príklad 3: Bertrandov paradox

- Andrea Ječmenová, bakalárska práca, 2017

**Názov:** Don't just solve it; fight it! Príklady z pravdepodobnosti
*Don't just solve it; fight it! Probability problems*

**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 myšienka 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

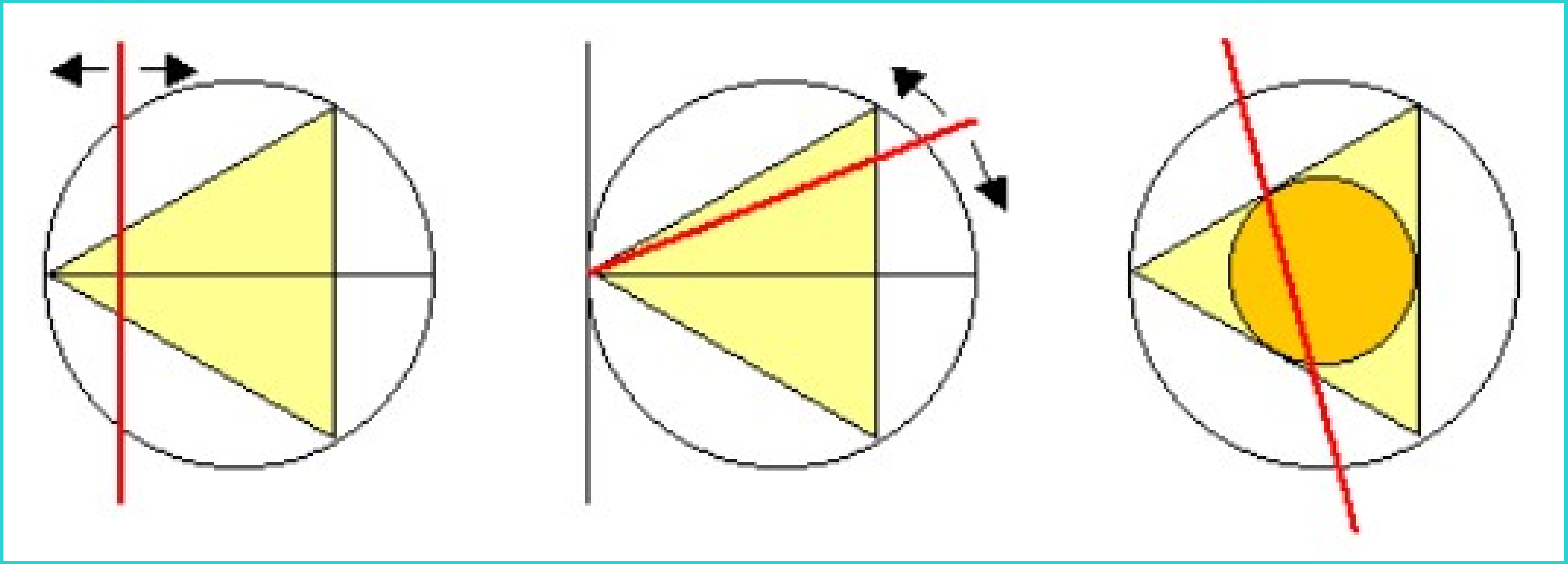
# Príklad 3: Bertrandov paradox

- Bertrand (1889):

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

# Príklad 3: Bertrandov paradox

- V závislosti od toho, čo znamená „náhodná tetiva“, dostaneme tri rôzne výsledky (1/2, 1/3, 1/4)

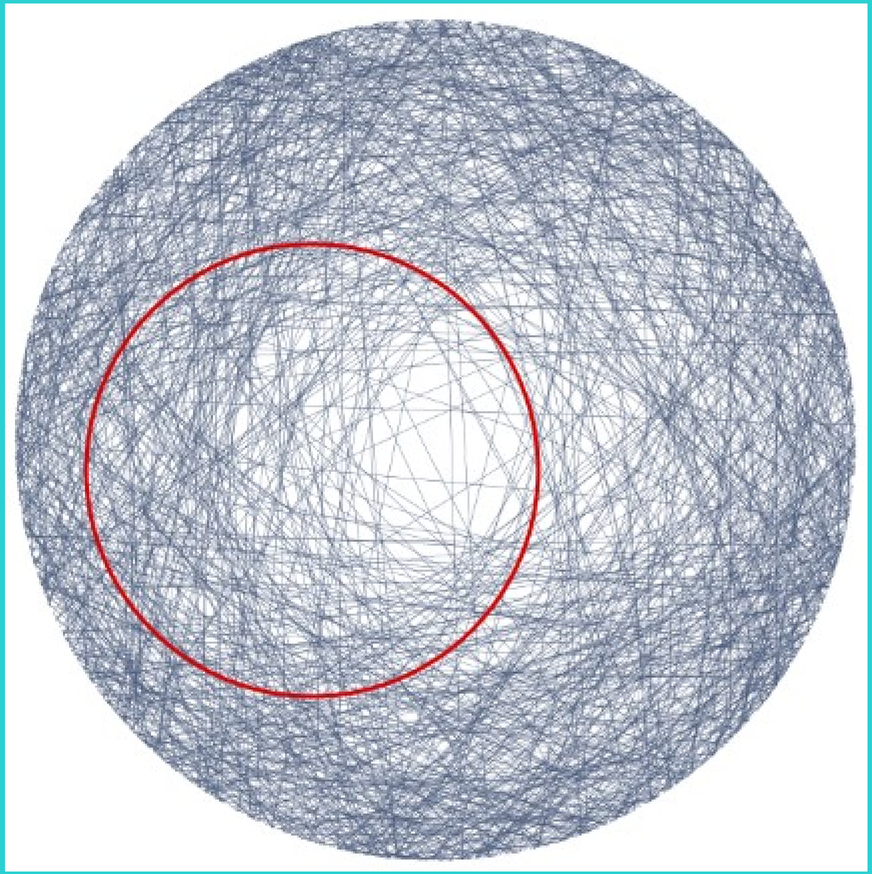


Obrázok: http://utenti.quipo.it/base5/probabil/bertrand.htm

# Príklad 3: Bertrandov paradox

- Jaynes (1973)
  - Rozdelenie dĺžky tetivy musí byť invariantné na polohu a veľkosť kruhu

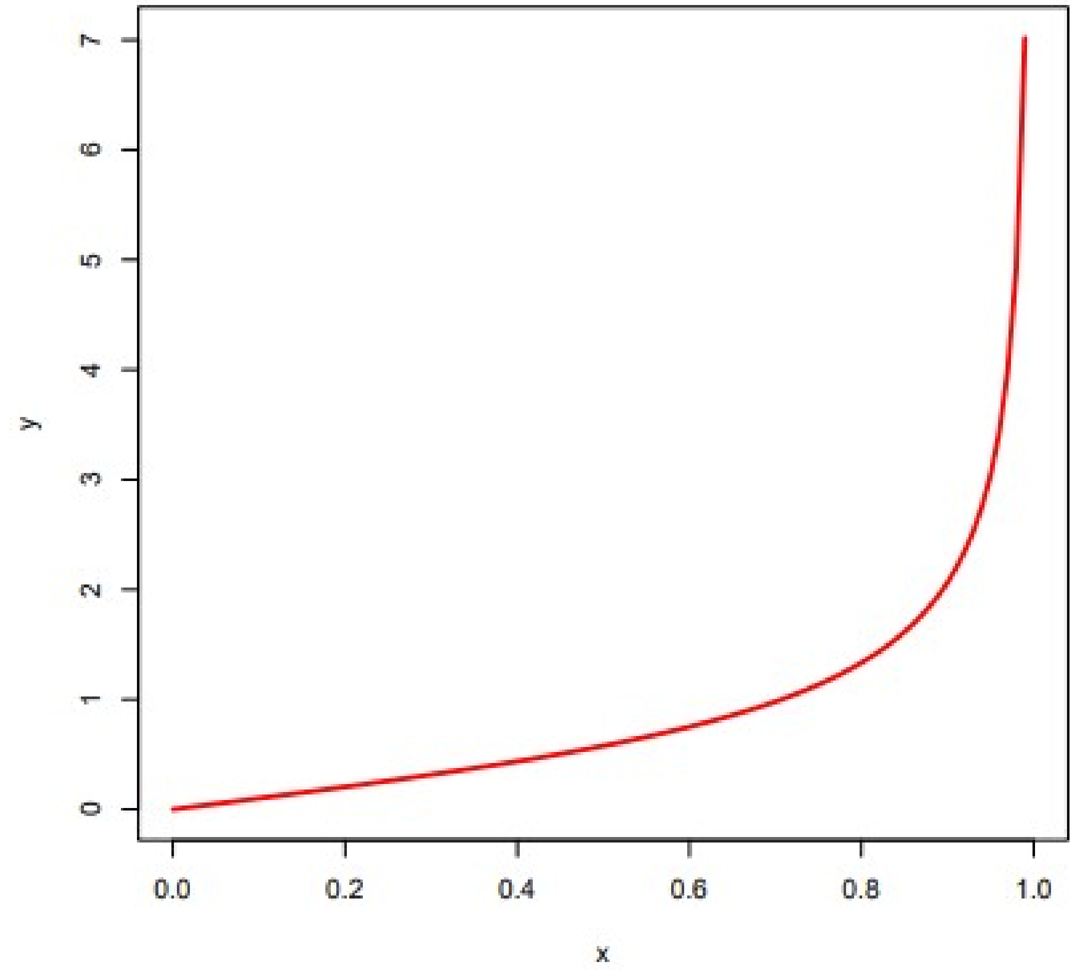
    Obrázok: https://en.wikipedia.org/wiki/Bertrand_paradox_(probability)#/media/File:Bertrand3-translate_ru.svg
  - Odvodil hustotu rozdelenia:



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

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

# Príklad 3: Bertrandov paradox




**Obr. 4.1.4:** Priebeh odvodenej funkcie hustoty tvaru $p(x) = \dfrac{x}{\sqrt{1-x^2}}$

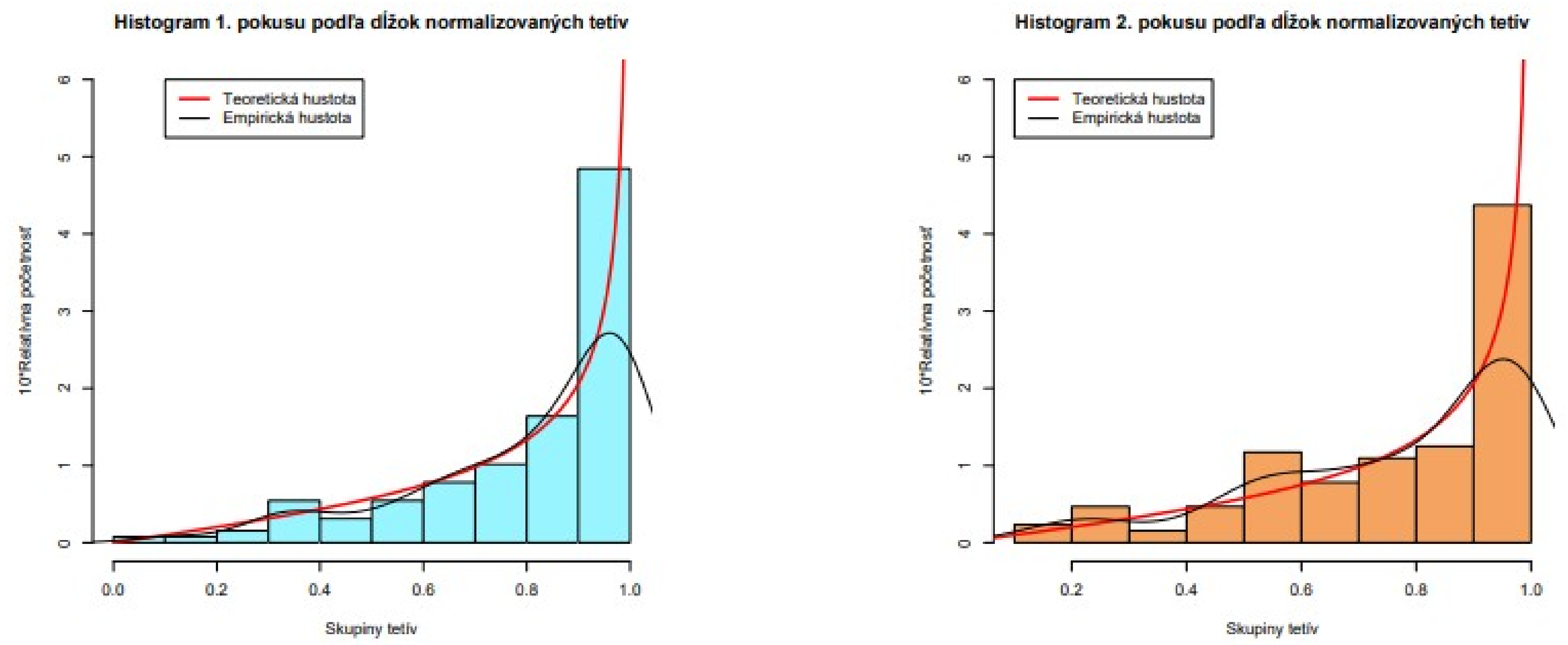
# Príklad 3: Bertrandov paradox

- Experiment v bakalárke:

model. My sme sa rozhodli realizovať experiment popísaný v [15], teda hádzať slamku (v našom prípade špajdľu) 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.

# Príklad 3: Bertrandov paradox

- Výsledky:




**Obr. 4.2.1:** Histogramy vykreslené z dát spolu s funkciou očakávanej (teoretickej) a vypočítanej (empirickej) hustoty. Na osi x sú rovnomerne rozdelené skupiny tetív a na osi y početnosť.

# Príklad 3: Bertrandov paradox

- Testovanie hustoty chí kvadrát testom:

| Číslo skupiny | Skupina tetív | Očak. počet |
|---|---|---|
| **1** | $(L/2R \leq 0.5)$ | 17.1487 |
| **2** | $(0.5 < L/2R \leq 0.6)$ | 8.45125 |
| **3** | $(0.6 < L/2R \leq 0.7)$ | 10.9897 |
| **4** | $(0.7 < L/2R \leq 0.8)$ | 14.6103 |
| **5** | $(0.8 < L/2R \leq 0.85)$ | 9.37182 |
| **6** | $(0.85 < L/2R \leq 0.9)$ | 11.6343 |
| **7** | $(0.9 < L/2R \leq 0.925)$ | 7.15812 |
| **8** | $(0.925 < L/2R \leq 0.95)$ | 8.6678 |
| **9** | $(0.95 < L/2R \leq 0.975)$ | 11.5258 |
| **10** | $(0.975 < L/2R \leq 1)$ | 28.4422 |

**Tabuľka 4.2.1:** Druhé rozdelenie normalizovaných tetív do skupín na testovanie s ich očakávanými početnosťami (červenou) podľa hustoty 4.1.4.

# Príklad 3: Bertrandov paradox

- Testovanie hustoty chí kvadrát testom:

| | Pokus č.1 | | Pokus č.2 | | Pokus č.3 | | Pokus č.4 | |
|---|---|---|---|---|---|---|---|---|
| **Skupiny** | **Početnosť** | **TS** | **Početnosť** | **TS** | **Početnosť** | **TS** | **Početnosť** | **TS** |
| **1** | 15 | 0.2692 | 17 | 0.0012 | 5 | 8.6065 | 11 | 2.2046 |
| **2** | 7 | 0.2492 | 15 | 5.0745 | | | | |
| **3** | 10 | 0.0891 | 10 | 0.0891 | | | | |
| **4** | 13 | 0.1774 | 14 | 0.0254 | | | | |
| **5** | 10 | 0.0421 | 5 | 2.0393 | | | | |
| **6** | 11 | 0.0345 | 11 | 0.0345 | 17 | 2.4746 | 15 | 0.9736 |
| **7** | 6 | 0.1873 | 7 | 0.0034 | 8 | 0.0990 | 10 | 1.1282 |
| **8** | 6 | 0.8211 | 9 | 0.0127 | 11 | 0.6275 | 5 | 1.5520 |
| **9** | 10 | 0.2019 | 10 | 0.2019 | 15 | 1.0472 | 10 | 0.2019 |
| **10** | 40 | 4.6966 | 30 | 0.0853 | 39 | 3.9190 | 38 | 3.2118 |
| **Súčet** | 128 | **6.7688** | 128 | **7.5679** | 128 | **20.3906** | 128 | **11.8703** |

```
octave:1> pkg load statistics
octave:2> chi2inv(0.95, 9)
ans = 16.919
octave:3> chi2inv(0.99, 9)
ans = 21.666
```

**Tabuľka 4.2.2:** Početnosť a testová štatistika pre skupiny tetív a pokusy.