

MA012 Statistika II

cvičení 5

Ondřej Pokora (pokora@math.muni.cz)

Ústav matematiky a statistiky, Přírodovědecká fakulta, Masarykova univerzita, Brno

(podzim 2015)



U všech příkladů vykreslujte i sloupcový graf empirických a teoretických četností.

Příklad 1 (z přednášky)

Ze souboru rodin s pěti dětmi bylo náhodně vybráno 84 rodin a zjištěn počet chlapců, data jsou v souboru [rodiny.csv](#). Na asymptotické hladině významnosti 0,05 testujeme hypotézu, že počet chlapců v rodinách s 5 dětmi má binomické rozdělení $Bi(5; 0,5)$.

[[reseni-05-1.R](#) Hypotézu nezamítneme.]

Příklad 2 (z přednášky)

Byla sledována doba (v minutách), jakou 70 klientů jisté firmy strávilo čekáním na obsluhu (od vyzvednutí pořadového lístku). Data jsou v souboru [fronta.csv](#). Na asymptotické hladině významnosti 0,05 testujeme hypotézu, že doba čekání má exponenciální rozdělení pravděpodobnosti, a to testem dobré shody a jednoduchým testem.

[[reseni-05-2.R](#) Test dobré shody hypotézu nezamítne, jednoduchý test zamítne.]

Příklad 3

Na nádraží byl sledován počet přijíždějících vlaků za dobu 1 h. Pozorování bylo prováděno celkem 5 dnů, tj. 360 h. Výsledky jsou zaznamenány v souboru `vlaky.csv`. Na asymptotické hladině významnosti 0,05 testujeme hypotézu, že počet přijíždějících vlaků za 1 h se řídí Poissonovým rozdělením pravděpodobnosti, a to testem dobré shody a jednoduchým testem.

$[K = 9,7, \chi_{0,95}^2(6) = 12,6; Q = 331,1.$
Ani jeden z obou testů hypotézu nezamítne.]

Příklad 4

Soubor `Brno.csv` obsahuje počty obyvatel města Brna (údaje z roku 2001) podle měsíců narození. Na asymptotické hladině významnosti 0,05 testujeme hypotézu, že počty narozených jsou pro všechny měsíce stejné (uvažujte počty dní v měsících nepřestupného roku).

$[K = 1506,2, \chi_{0,95}^2(11) = 19,7.$ Hypotézu zamítneme, a prokážeme tak, že rozdělení měsíců narození není rovnoměrné.]

Příklad 5

V R si vygenerujte náhodný výběr rozsahu alespoň 100 z nějakého rozdělení pravděpodobnosti se konkrétními parametry. Následně testem dobré shody a Kolmogorovovým-Smirnovovým testem s $\alpha = 0,05$ testujte, že vygenerovaný náhodný výběr odpovídá zvolenému rozdělení pravděpodobnosti s danými parametry.

U testu dobré shody můžete intervaly volit podle dělicích bodů z histogramu. Prozkoumejte: `h <- hist (X,plot=FALSE), h$breaks, h$counts`.

Příklad 6

J. G. Mendel ve svých pokusech pozoroval 10 rostlin hrachu a na každé z nich počet žlutých a zelených semen. Výsledky pokusu jsou v souboru `JGMendel.csv`. Podle teoretického genetického modelu je pravděpodobnost výskytu žlutého semene 0,75 a zeleného 0,25. Na asymptotické hladině významnosti 0,05 testujeme hypotézu, že výsledky Mendelových pokusů se shodují s teoretickým modelem.

$$[K = 1,8, \chi_{0,95}^2(9) = 16,9. \text{ Hypotézu nezamítneme.}]$$

Příklad 7

Vyzkoušejte si použití různých testů normality na datech ze cvičení 1–3 (ANOVA).

Příklad 8

Při 60 hodech hrací kostkou byly sledovány četnosti jednotlivých padnutých čísel; ty jsou uvedeny v souboru [kostka.csv](#). Na asymptotické hladině významnosti 0,05 testujeme hypotézu, že kostka je homogenní (spravedlivá).

$$[K = 2,8, \chi_{0,95}^2(5) = 11,1. \text{ Hypotézu nezamítneme.}]$$

Příklad 9

Soubor [pohotovost.csv](#) uvádí počty pacientů na lékařské pohotovosti během 8hodinových směnách v celkem 75 dnech. Jednoduchým testem s $\alpha = 0,05$ testujeme hypotézu, počty mají Poissonovo rozdělení.

$$[Q = 1158,6. \text{ Hypotézu nezamítneme.}]$$

Příklad 10

Při 4096 hodech 12 hracími kostkami byly sledováno, kolikrát padne šestka. Výsledky jsou uvedeny v souboru [kostky.csv](#). Na asymptotické hladině významnosti 0,05 testujeme hypotézu, že počet šestek v jednom hodu má binomické rozdělení $Bi(12; \frac{1}{6})$.

$$[K = 5,5, \chi_{0,95}^2(7) = 14,1. \text{ Hypotézu nezamítneme.}]$$

Příklad 11

Do rybníka bylo umístěno pět pastí, svítících bílým, žlutým, modrým, zeleným a červeným světlem. Počty ryb chycených do jednotlivých pastí obsahuje soubor [rybnik.csv](#). Na asymptotické hladině významnosti 0,05 testujeme hypotézu, že barva světla nemá vliv na počet chycených ryb.

$$[K = 14,1, \chi_{0,95}^2(4) = 9,5. \text{ Hypotézu zamítneme, barva má vliv na počet chycených ryb.}]$$

Příklad 12

300 zákazníků jistého řetězce supermarketů bylo dotázáno, který den v týdnu chodí nejčastěji nakupovat. Soubor `supermarkety.csv` obsahuje tyto počty. Na asymptotické hladině významnosti 0,05 testujeme hypotézu, že všechny dny v týdnu jsou zákazníci k nákupu využívány rovnoměrně.

$$[K = 78, \chi_{0,95}^2(6) = 12,6.]$$

Hypotézu zamítneme, preference jednotlivých dnů jsou nerovnoměrné.]

Příklad 13

Byl zjišťován celkový počet vstřelených branek v základní hrací době 84 fotbalových zápasů. Počty jsou zaznamenány v souboru `fotbal.csv`. Testem dobré shody na asymptotické hladině významnosti 0,05 testujeme hypotézu, že počet vstřelených branek má Poissonovo rozdělení pravděpodobnosti.

$$[K = 2,1, \chi_{0,95}^2(3) = 7,8. \text{ Hypotézu nezamítneme.}]$$