předmět 12: počítáte si z textu MPSO – viz slajd 12

cvičení 12: χ^2-test (čteme: test chí kvadrát); neparametrické testy pořadí hodnot

Nejprve se naučíme řecké písmeno „chí" (anglicky se čte [kai]):

$$\alpha\ \beta\ \gamma\ \delta\ \varepsilon\ \zeta\ \vartheta\ \eta\ \iota\ \kappa\ \lambda\ \mu\ \nu\ \xi\ o\ \pi\ \rho\ \sigma\ \tau\ \upsilon\ \varphi\ \chi\ \psi\ \omega$$

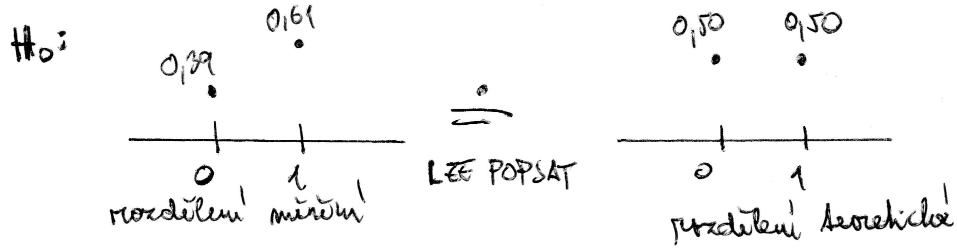
Tedy písmeno χ je sice velikostně velké, ale zasahuje pod linku, podobně jako naše české p (jako česká k mají velikost pouze $\beta, \zeta, \vartheta, \lambda, \xi, \delta$).

Př. 6 (MPSO, str. 127, př. 5.3) Honza Kovář pracuje v mincovně – jeho úkolem je zajistit, aby mince byly dobře vyvážený, aby např. rub a líc padaly stejně často. Proto hodí stovkou desetikorun a padne mu 61 líců. Tiskne stroj vyvážené mince?

Řešení: netiskne, to je vidět na první pohled. Odpovězme ovšem pravděpodobnostně statisticky, výpočtem. Provedeme statistický test dobré shody (GFT = good-fit test):

označme X = počet líců ze 100 hodů (hodili jsme 100 různých mincí)

krok 1: H_0:



H_1: Měření NELZE POPSAT naším teoretickým modelem

krok 2: Naším kritériem bude měřené četnosti rubů a líců $f_m = (39 \text{ rubů}, 61 \text{ líců})$, které budeme porovnávat s teoretickými četnostmi $f_t = (50 \text{ rubů}, 50 \text{ líců})$ teoretického modelu: mince jsou vyvážený vyvážené

Respektive rozumíme jejich „součet kvadratických rozptýlení těchto četností"

$$\text{součet odchylek} := \sum_0^1 \frac{(f_t - f_m)^2}{f_t} = \text{... dosazení nastane v kroku 5}$$

krok 3: $$\sum_0^1 \frac{(f_t - f_m)^2}{f_t} \overset{H_0}{\sim} \chi^2(1)$$

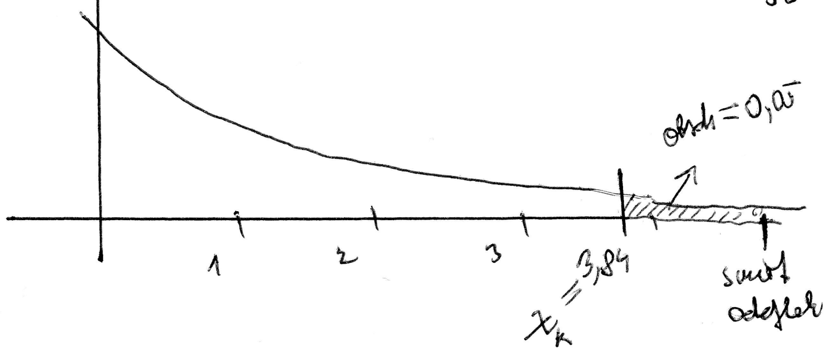
počet stupňů volnosti = počet párů MINUS jedna $= 2-1 = 1$

krok 4: $\alpha = 0{,}05$ je zvolená hladina významnosti testu; ze tvaru H_0, H_1 nyní není tolik zřejmé, že se jedná o jednostranný test – testy dobré shody jsou vždy jednostranné,

řecké písmeno ný

obsah α hustoty f „usekáváme" vždy jen vpravo

V tabulce χ^2 řádek $\nu = 1$, sloupec $q = 0{,}05$ } $x_K = 3{,}84146$ pro $\nu = 1$, $\alpha = 0{,}05$ pravostranný test krok 5: $\text{součet odchylek} = \frac{(50-39)^2}{50} + \frac{(50-61)^2}{50} = \frac{242}{50} = \underline{\underline{4{,}84}}$



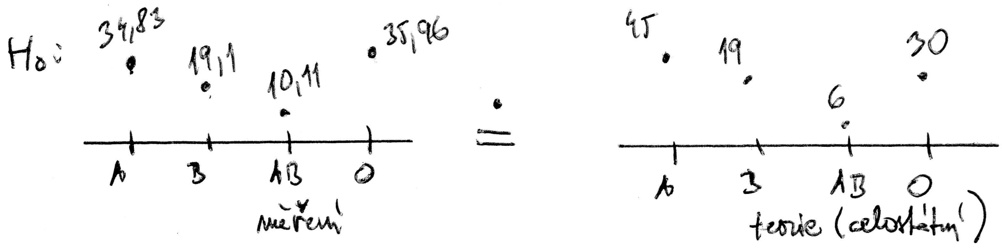
⇓ H_0 zamítáme, vyváženost výsledku hodu mincí nelze popsat modelem $\left(\frac{1}{2}, \frac{1}{2}\right)$

Př. 7. V celé populaci se vyskytuje v ČR 45 % lidí s krevní skupinou A
19 % — " — B
6 % — " — AB
30 % — " — 0. Výzkumník má k dispozici
89 pacientů, z nichž 31 má skupinu krve A, 17 skupinu B, 9 skupinu AB, 32 skupinu 0.
Ověřte statistickým testem, zda výběr daných pacientů je reprezentativní vzhledem k celé populaci v ČR.

Řešení – totéž jako př. 6, pouze v bledě modrém: naměřené $f_m = (31, 17, 9, 32)$ porovnáváme s „teoretickým rozdělením" $(45, 19, 6, 30) = f_t$. Abychom mohli porovnat, musíme f_m převést na procenta: $f_m = \left(\frac{31}{89}, \frac{17}{89}, \frac{9}{89}, \frac{32}{89}\right) = (34{,}83;\ 19{,}1;\ 10{,}11;\ 35{,}96)$.

krok 1:

H_0:



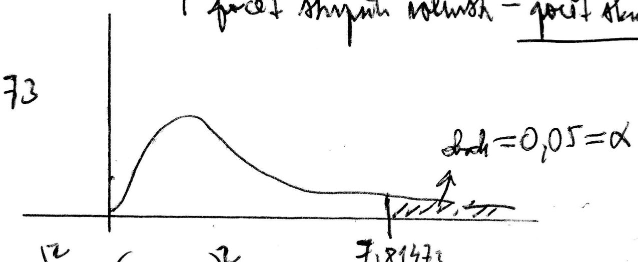
H_1: měření NELZE popsat naším teoretickým modelem

krok 2–3

kritérium $\displaystyle\sum_{1}^{4} \frac{(f_t - f_m)^2}{f_t} \overset{H_0}{\sim} \chi^2(3)$

↑ počet stupňů volnosti = <u>počet skupin MINUS 1</u>

krok 4: pro $\alpha = 0{,}05$: $\chi_k(3) = 7{,}81473$



krok 5: $\text{součet odchylek} = \frac{(45-34{,}83)^2}{45} + \frac{(19-19{,}1)^2}{19} + \frac{(6-10{,}11)^2}{6} + \frac{(30-35{,}96)^2}{30} = \underline{\underline{6{,}29835}}$

$6{,}29835 \in (0;\ 7{,}81473) \Rightarrow H_0$ nezamítáme, daný vzorek pacientů je celorepublikově reprezentativní!

Př. 8. Ve městě byl po dobu 60 dnů evidován počet dopravních nehod v průběhu každého dne (X = počet nehod v daný den) a získány četnosti

x_i	0	1	2	3	4	5	6
f_i	4	28	10	7	6	4	1

Odhadněte, o jaké rozdělení asi se jedná a proveďte test dobré shody, že váš odhad je správný.

Řešení: to už je příklad analogický Vašemu nápověrnému příkladu číslo 2: z povahy věci, že různé nehody jsou na sobě nezávislé, a nastávají naprosto náhodně:

počet výskytů těchto nehod (za jednotku času) lze popsat rozdělením Poissonovým (model D5 z předn. 8)

se střední hodnotou, kterou odhadneme z průměru: $\lambda \doteq \frac{1}{60}(0\cdot4+1\cdot28+2\cdot10+3\cdot7+4\cdot6+5\cdot4+6) = \underline{1{,}9833}$

Odtud lze spočítat teoretické psti $p(k) = \frac{\lambda^k}{k!}\cdot e^{-\lambda}$. Udělejme to rychle v programu R (instrukce viz předn. 8 pro jednotlivá rozdělení)

x <- c(0:6) ENTER ... vytvoří vektor 0 až 6

px <- dpois(x, 1.9833) ENTER ... vypočte psti $p(k)$ v bodech, které nás zajímají

px ENTER ... opíšeme tyto hodnoty z obrazovky

pro $k = 0,1,2,3,4,5,6$ vyde $p(k) = (0{,}1376;\ 0{,}2729;\ 0{,}2707;\ 0{,}1789;\ 0{,}0887;\ 0{,}0352;\ 0{,}0116)$.

To jsou vlastně „teoretické relativní četnosti", když je vynásobíme počtem měření 60, který chceme porovnávat, dostaneme

$60\cdot p(k) \doteq (8.257;\ \underset{17}{16.376};\ 16.239;\ \underset{11}{10.736};\ 5{,}323;\ 2.111;\ 0.7)$

Tyto hodnoty mají modelovat celočíselné výskyty – mohli bychom je zaokrouhlit, ale součet zaokrouhlených čísel je 59, nikoli 60 (celému teoretickému modelu neceločíselné hodnoty nevadí, tak je můžeme nechat); provedeme statistický test:

krok 1: H_0: měření (4, 28, 10, 7, 6, 4, 1) se statisticky významně nelišší od četností (8,257; 16,376; 16,239; 10,736; 5,323; 2,111; 0,7)

H_1: ... liší ...

POZOR!

Poznámka: test χ^2 dobré shody vyžaduje, aby jednotlivé četnosti byly aspoň 5: pokud nejsou, sousední třídy četností sloučíme: správně bychom měli testovat shodu

$(\underbrace{4,28}_{32},10,7,6,\underbrace{4,1}_{5})$, tedy $\underline{(32,10,7,6,5)}$, s modelem

$(\underbrace{8,257;16,376}_{24,633};16,239;10,736;5,323;\underbrace{2,111;0,7}_{2,811})$, tedy $\underline{(24,633;16,239;10,736;5,323;2,811)}$:

reformulace H_0: měření $(32,10,7,6,5)$ se neliší od četností — " —

H_1: ... liší ...

krok 2+3: $$\sum_1^5 \frac{(f_t-f_m)^2}{f_t} \overset{H_0}{\sim} \chi^2(4)$$

počet tříd MINUS 1 = 5 − 1 = 4

krok 4:

$$\chi_k(4,\ 0,05) = \underline{9,48773}$$

krok 5: součet odchylek $= \frac{(24,633-32)^2}{24,633} + \frac{(16,239-10)^2}{16,239} + \frac{(10,736-7)^2}{10,736} + \frac{(5,323-6)^2}{5,323} + \frac{(2,811-5)^2}{2,811} =$

$= 7,691 \in (0;\ 9,48773) \Rightarrow$ H_0 nezamítáme,

náš teoretický model $Po(\lambda = 1,983)$

dobře popisuje získaná data!!!

Př. 9 (MPSO – př. 5.4/str. 127) Rozdělení IQ v celé ČR lze popsat jako normální definované, v matematickém smyslu slova, tj. X = IQ náhodně vybraného člověka v ČR $\sim No(\mu = 100, \sigma^2 = 225)$. Ověřte testem dobré shody, že náhodně vybraný vzorek Brňanů je celorepublikově reprezentativní:

$\sigma = 15$

IQ	< 55	55–70	70–85	85–100	100–115	115–130	130–145	> 145	
četnost f_m	20	17	29	52	63	42	13	14	$\sum f_m = 250$ lidí

Řešení: Porovnáme naměřené četnosti s teoretickými – výpočet teoretických četností je celkem pracný, protože každého intervalu pst musíme pomocí tabulky distribuční funkce Φ.

Ale jinak to nejde:

Pro $X \sim No(\mu = 100, \sigma = 15)$ máme

a) $P(X < 55) = \Phi\left(\frac{55-100}{15}\right) = \Phi(-3) = 1 - \Phi(3) = 1 - 0,9986501 = \underline{0,0013499} \doteq 0,00135$

b) $P(X \in \langle 55;70\rangle) = \Phi\left(\frac{70-100}{15}\right) - \Phi\left(\frac{55-100}{15}\right) = \Phi(-2) - \Phi(-3) = \Phi(3) - \Phi(2) \doteq \underline{0,0214}$

c) $P(X \in \langle 70;85\rangle) = \Phi\left(\frac{85-100}{15}\right) - \Phi\left(\frac{70-100}{15}\right) = \Phi(-1) - \Phi(-2) = \Phi(2) - \Phi(1) \doteq \underline{0,1359}$

d) $P(X\in\langle 85,100\rangle)=\Phi\left(\frac{100-100}{15}\right)-\Phi\left(\frac{85-100}{15}\right)=\Phi(0)-\Phi(-1)=\frac{1}{2}-1+0{,}8413447\doteq \underline{0{,}341345}$

Protože rozdělení normální je symetrické vzhledem ke hodnotě $\mu=100$, následující čtyři psti jsou už „symetricky stejné":

e) $P(X\in\langle 100;115\rangle)=\Phi\left(\frac{115-100}{15}\right)-\Phi\left(\frac{100-100}{15}\right)=\Phi(1)-\Phi(0)\doteq\underline{0{,}341345}$

f) $P(X\in\langle 115;130\rangle)=\Phi(2)-\Phi(1)\doteq\underline{0{,}1359}$

g) $P(X\in\langle 130;145\rangle)=\Phi(3)-\Phi(2)\doteq\underline{0{,}0214}$

h) $P(X>145)=1-\Phi(3)\doteq\underline{0{,}00135}$

dostáváme tedy

teoretický model

IQ	< 55	55–70	70–85	85–100	100–115	115–130	130–145	>145
p_t	0,00135	0,0214	0,1359	0,341345	0,341345	0,1359	0,0214	0,00135
$250\cdot p_t=\underline{f_t}$	0,3375	5,35	33,975	85,33625	85,33625	33,975	5,35	0,3375

↑ celek = 250 lidí

z toho plyne poučení: pokud si délku a polohu intervalů volíme sami, zkonstruujeme je tak (například, aby μ je krajní mez při sudém počtu intervalů, nebo aby μ je střed jednoho z intervalů, pokud je jich lichý počet)

Provedeme nyní statistický test dobré shody:

1. H_0: měřené f_m a teoretické četnosti f_t se významně <u>ne</u>liší
   H_1 — || — liší
2. kritérium je $\sum_{1}^{8}\frac{(f_t-f_m)^2}{f_t}$
3. za předpokladu platnosti H_0 má naše kritérium rozdělení $\chi^2(7)$
4. pro $\alpha=0{,}05$: $\chi_K(7;0{,}05)=14{,}0671$
5. součet odchylek $=\frac{(0{,}3375-20)^2}{0{,}3375}+\frac{(5{,}35-17)^2}{5{,}35}+\frac{(33{,}975-29)^2}{33{,}975}+\frac{(85{,}33625-52)^2}{85{,}33625}+\frac{(85{,}33625-63)^2}{85{,}33625}+\frac{(33{,}975-42)^2}{33{,}975}+\frac{(5{,}35-13)^2}{5{,}35}+\frac{(0{,}3375-14)^2}{0{,}3375}=$ hodnota,

což tedy první zlomek je větší než 300, tj. hodnota $\notin(0;14{,}0671)$, výběr Brňáků není celorepublikově reprezentativní (H0 zamítáme)

<u>Př. 10</u> (Chráska: Metody pedagogického výzkumu, str. 78). Provádíme výzkum u 400 studentů, zda průměrný prospěch studentů závisí na tom, zda bydlí na VŠ kolejích nebo ne. Měření:

$\underline{f_m}$:

	prosp. 1–1,6	prosp. 1,6–2,1	prosp. 2,1–5
bydlí na koleji	39	107	93
nebydlí na koleji	41	73	47

Ověřte testem pro kontingenční tabulky, zda typ bydlení závisí na prospěchu průměrném.

Řešení: Test pro kontingenční tabulky je prakticky stejný jako test dobré shody, pouze porovnáváme naměřenou tabulku s „teoreticky zkonstruovanou tabulkou", v níž veličina v řádcích nezávisí na veličině ve sloupcích. Tu zkonstruujeme následujícím způsobem:

a) nejprve sečteme hodnoty v každé z podmínek v řádku či sloupci:

	B_1	B_2	B_3	
A_1				239 (0,5975)
A_2				161 (0,4025)
	80 (0,20)	180 (0,45)	140 (0,35)	400 studentů celkem

b) do závorky za každou ze sečtených hodnot dodáme její relativní četnost vzhledem k počtu studentů 400:

např. 239 studentů je 0,5975 z celku 400 studentů

180 studentů je 0,45 — " —

c) nyní sestavíme do závorek psti na průniku řádku a sloupce, které plynou z nezávislosti proměnných — zde si připomeneme vzorec z první poloviny semestru pro pst průniku nezávislých veličin:

$$P(A_i \cap B_j) = P(A_i)\cdot P(B_j)$$

(0,1195)	(0,268875)	(0,209125)	239 (0,5975)
(0,0805)	(0,181125)	(0,140875)	161 (0,4025)
80 (0,2)	180 (0,45)	140 (0,35)	

d) Z dílčích pstí v tabulce vyrobíme „absolutní teoretické četnosti" vynásobením 400:

$400 \cdot 0,1195 = 47,8$

$400 \cdot 0,181125 = 72,45$ atd. Dostaneme hledané NEZÁVISLÉ teoretické absolutní četnosti f_t:

f_t:

47,8	107,55	83,65
32,2	72,45	56,35

krok 1: H_0: četnosti f_m se podstatně neliší od četností f_t

H_1: — " — liší — " —

krok 2-3: kritérium $\sum_1^6 \frac{(f_t-f_m)^2}{f_t} \overset{H_0}{\sim} \chi^2$(volnost = (počet řádků MINUS 1)·(počet sloupců MINUS 1))

$= (2-1)\cdot(3-1) = 1\cdot 2 = 2$

krok 4: pro $\alpha = 0,05$ a volnost 2: $\chi_K(2; 0,05) = 5,99147$

krok 5:

$$\text{součet odchylek} = \frac{(47{,}8-39)^2}{47{,}8} + \frac{(107{,}55-107)^2}{107{,}55} + \frac{(83{,}65-93)^2}{83{,}65} +$$
$$+ \frac{(32{,}2-41)^2}{32{,}2} + \frac{(72{,}45-73)^2}{72{,}45} + \frac{(56{,}35-47)^2}{56{,}35} = 6{,}62856 \notin (0;5{,}99)$$

$\Downarrow$

H_0 zamítáme, hypotéza nezávislosti se nepotvrdila, tj. platí H_1, veličiny jsou závislé

(pohledem do tabulky měření je vidět, že lidé bydlící na kolejích mají horší průměrný prospěch).

Př. 11. Trochu tímto porušuji slib 6 stran na jedno cvičení či přednášku: ukázalo se, že plán na daný obsah předmětu je o jeden příklad větší než dostupný prostor na přednášku a cvičení.

Ovšem vzhledem k tomu, že předn. číslo 4 odpadla, snad jeden větší příklad navíc nebude tolik vadit (tento typ je v zápočtových příkladech, budete ho potřebovat):

zadání (Chráska: Metody pedagogického výzkumu, str. 93). Ve skupinách A, B žáků byly získány následující výsledky (počty bodů):

skupina A: 5, 7, 12, 13, 15 ($n_1 = 5$)

skupina B: 6, 9, 11, 14, 16, 18 ($n_2 = 6$)

Rozhodněte, zda mezi výsledky v obou skupinách je statisticky významný rozdíl

---

Řešení: U některých typů měření, zejména pokud se jedná o pořadí lidí, stupnice oblíbenosti výrobků, apod., už nelze předpokládat, že měřené veličiny mají normální rozdělení ani že je můžeme normálním rozdělením aproximovat. Testy předpokládající normalitu měřené veličiny se nazývají parametrické testy (protože testujeme hodnotu parametru μ nebo σ). Testy u veličin, kde normalitu předpokládat nemůžeme, se nazývají testy neparametrické;

například kdybychom místo počtu bodů měli jen pořadí –
seřaďme hodnoty podle velikosti:

5, 6, 7, 9, 11, 12, 13, 14, 15, 16, 18

A, B, A, B, B, A, A, B, A, B, B → kdybychom měli jen pořadí těchto skupin, aniž bychom znali skutečné počty bodů, i tak jsme schopni provést test – tzv. Mannův-Whitneyho statistický test:

krok 1: H_0: obě skupiny se (co se týká průměrného pořadí) v zásadě neliší

H_1: obě skupiny se průměrným pořadím liší

krok 2: kritérium zde je $\min(U_A, U_B)$,

kde $U_A = \sum_{\text{prvky } A}$ inverze vzhledem k množině B

(inverze vzhledem k B = počet prvků množiny B, které jsou PŘED daným prvkem)

$U_B = \sum_{\text{prvky } B}$ inverzí vzhledem k množině A

(inverze prvku z B vzhledem k množině A = počet prvků z A, které mají lepší pořadí než daný prvek)

krok 3: při platnosti H_0
lze kritické hodnoty pro $\min(U_A, U_B)$ najít v tabulkách, které spočítali pánové Mann a Whitney
(viz přednáška)

krok 4: pro $\alpha = 0{,}05$, $n_1 = 5$, $n_2 = 6$ máme: $\underline{\underline{U_k = 3}}$

krok 5:

$$\left.\begin{aligned} U_A &= 0+1+3+3+4 = 11 \\ U_B &= 1+2+2+4+5+5 = 19 \end{aligned}\right\} \min(U_A, U_B) = \underline{\underline{11}}$$

NYNÍ POZOR, ROZHODOVACÍ PROCES ZDE FUNGUJE JINAK:

$11 > U_k = 3 \Rightarrow$ $\underline{\underline{H_0 \text{ nezamítáme}}}$

(H_0 nezamítáme tehdy, když je $\min(U_A, U_B)$ dostatečně velké, větší než kritická hodnota)

Odpověď: výsledky obou skupin jsou zhruba stejné, neprokázal se významný rozdíl.

TOŤ VŠE! NYNÍ JSTE SCHOPNI SPOČÍTAT VŠECHNY ZÁPOČTOVÉ PŘÍKLADY
A TÝDEN PŘED SVOU ústní částí mi je pošlete!