

Cvičenie 11 – Testovanie hypotéz 2

Riešené príklady

a) Medián : Wilcoxonov test

Príklad 1

Na kurze zdravej výživy sa účastníci učili intuitívne odhadovať svoj kalorický príjem. Cieľom bolo dosiahnuť (bez predbežného merania) optimálny denný príjem 7725 kJ. Dodatočne sa pre kontrolu vždy odmerala skutočná kalorická hodnota prijatej potravy.

Na skúškach sa meral skutočný kalorický príjem 11 náhodne vybraných účastníkov. Dosiahli sa nasledujúce hodnoty (usporiadané podľa veľkosti):

[5260, 5470, 5640, 6180, 6390, 6515, 6805, 7515, 7515, 8230, 8770]

Kurz sa považuje za úspešný, ak sa medián denných výsledkov účastníkov zhoduje s optimom, alebo presnejšie povedané, aby sa nedala na predpísanej hladine významnosti zamietnuť hypotéza o tom, že stanovená optimálna hodnota (7725) je mediánom.

Na hladine významnosti $\alpha = 0.05$ (5%), v konfrontácii s nameranými hodnotami, s použitím Wilcoxonovho testu, testujte hypotézu $\text{med.} = 7725$.

Riešenie:

Už na prvý pohľad vidíme, že hodnota 7725 sa v rade dosiahnutých hodnôt nachádza medzi 9. a 10. číslom, teda medián to určite nie je. Úlohou testovania je zistiť, či táto odchýlka sa dá vysvetliť náhodou alebo je znakom nie najlepších výsledkov kurzu. Testujeme na hladine $\alpha=0.05$ (5%), použijeme Wilcoxonov test, ku ktorému potrebujeme tabuľky Wilcoxonovho rozdelenia.

Krok 1: Odpočítame od všetkých hodnôt vo vektore hypotetický medián (7725):

[-2465 -2255 -2085 -1545 -1335 -1210 -920 -210 -210 505 1045]

Krok 2: Získané čísla v absolútnej hodnote usporiadame (iba v hlave) od najmenšieho po najväčšie a pridáme im (na papier) poradové čísla. Ak by dve alebo viaceré čísla mali rovnaké poradie, pridáme im aritmetický priemer poradových čísel, na ktoré „majú nárok“. V našom príklade sa to týka hodnoty 210, ktorá je tu dvojmo a „berie“ tak prvé dve miesta. Pridáme obom 1.5:

[	-2465	-2255	-2085	-1545	-1335	-1210	-920	-210	-210	505	1045	]
	11	10	9	8	7	6	4	1.5	1.5	3	5	

Krok 3: Sčítame poradové čísla osobitne pri kladných a záporných číslach.

$s+ = 8$, $s- = 58$

Asi tušíme, že hypotetický medián má tým väčšie šance na uznanie, čím menej sa $s+$ a $s-$ od seba líšia. Ideál by bol 33 a 33 (lebo súčet všetkých 11 poradí je 66), naopak najhoršie by bolo 0 a 66.

Krok 4: Vyberieme minimum z oboch súčtov

$\min\{s+, s-\} = 8$

Čím väčšie s (max. hranicou je 33), tým lepšie pre testovaný hypotetický medián.

Krok 5: Zhodnotenie

Máme dané $\alpha=0.05$, vo vektore je 11 údajov, preto si v tabuľkách zaostríme zrak na 11-ty riadok a stĺpec 0.05 (ak má dve časti, vyberáme časť *two-tailed test*):

$w_{11,0.05} = 10$

Získané číslo 10 je kritická hranica. Už vieme, že čím menšie s , tým horšie.

Naše s je $s = 8$, a ako vidíme, 8 je menej než 10, teda hodnota 8 je **pod** kritickou hranicou.

Zle je – hypotézu $\text{med.} = 7725$ zamietame na hladine 0.05.

Poznámka

Ak by sme testovali na hladine významnosti 0.01, kritická hodnota bude 5 a hypotézu nezamietneme.

Príklad 2

Riaditeľ štátnej pošty v krajine XYZ vyhlásil, že „pred okienkami“ sa čaká mediánovo 12 minút. Uvedenú hypotézu novinár ABC testoval tak, že na 12 náhodne vybraných poštách sa postavil do radu s pohľadnicou a čakal + meral čas (na celé minúty), kedy príde na rad. Výsledky sú zhrnuté vo vektore (ešte neusporiadanom):

[5 6 7 18 4 12 6 5 10 13 16 8]

Testujte riaditeľovu hypotézu na hladine významnosti 0.05.

Riešenie:

Krok 0: $x = \text{sort}([5 \ 6 \ 7 \ 18 \ 4 \ 12 \ 6 \ 5 \ 10 \ 13 \ 16 \ 8])$

$x = [4 \ 5 \ 5 \ 6 \ 6 \ 7 \ 8 \ 10 \ 12 \ 13 \ 16 \ 18]$

Krok 1: $x - 12$

$[-8 \ -7 \ -7 \ -6 \ -6 \ -5 \ -4 \ -2 \ 0 \ 1 \ 4 \ 6]$

Krok 2:

[	-8	-7	-7	-6	-6	-5	-4	-2	0	1	4	6	]
	12	10.5	10.5	8	8	6	4.5	3	1	2	4.5	8	

Máme tu až tri šestky – nebojíme sa, spriemerujeme poradia 7,8,9 a všetky dostanú osmičku.

Krok 3:

Máme tu nulu – nie je záporná ani kladná. Z hľadiska detailnej stability metódy je najlepšie jej poradové číslo 1 rozdeliť po polovičke do oboch súčtov.

$s+ = 15$

$s- = 63$

Krok 4: $\min\{s+, s-\} = 15$

Krok 5: $w_{12,0.05} = 13$

$15 > 13$, tj. 15 (tesne) nekleslo pod kritickú hodnotu. Hypotézu nezamietame, čo však neznamená, že jej bezhranične dôverujeme.

b) Medián : Znamienkový test

Na rozdiel od Wilcoxonovho testu nás bude zaujímať len to, či sú hodnoty vpravo alebo vľavo od hypotetického mediánu, nebudeme rozlišovať ich poradie.

Príklad 3

Riešme príklad 1 zadaný vyššie pomocou znamienkového testu.

Riešenie:

Krok 1: Každému číslu vpravo (tj. väčším) od mediánu priradíme hodnotu 1 (reprezentuje znamienko +), číslam vľavo pridelíme 0 (znamienko-).

Hypotet. med. = 7725

5260	5470	5640	6180	6390	6515	6805	7515	7515	8230	8770
0	0	0	0	0	0	0	0	0	1	1

Krok 2: Počet jednotiek

$$Y = 2$$

Krok 3: Kľúčový vzorec (n je počet prvkov v dátovom vektore)

$$U = (Y - n/2) / \sqrt{(n/4)}$$

$$U = (2 - 5.5) / \sqrt{2.75} = -2.1106$$

Krok 4: Zhodnotenie – $|U|$ porovnáme s $F^{-1}(1-\alpha/2)$ – normálne rozdelenie

$$F^{-1}(1-0.05/2) = F^{-1}(0.975) = 1.96$$

1.96 je kritická hranica. Čo je medzi -1.96 a 1.96, nezamietame, to čo je mimo, zamietame.

$2.1106 > 1.96$ – sme vonku z normálu (na „chvoste“), hypotézu zamietame

Poznámka 1

Ak by sme testovali na hladine významnosti 0.01, kritická hodnota bude

$$F^{-1}(0.995) = 2.575$$

Vtedy bude 2.1106 „v norme“, teda menšie ako kritická hranica.

Poznámka 2

Výsledky vyšli rovnako ako pri Wilcoxonovom teste.

Príklad 4

Riešme príklad 2 zadaný vyššie pomocou znamienkového testu.

Riešenie:

Krok 1: med. = 12
x = [4 5 5 6 6 7 8 10 12 13 16 18]
 0 0 0 0 0 0 0 0 - 1 1 1

Číslo 12 presne trafilo do hypotézy. Bežne sa v takýchto prípadoch zvykne vynechávať (tj. berieme do úvahy len 11 výsledkov). *Ak by však takýchto hodnôt bolo viac, môže takýto postup skresliť výsledky a vtedy je lepšie ich rátať s ohodnotením 0.5.*

Krok 2: Počet jednotiek

$$Y = 3$$

Krok 3: Kľúčový vzorec (beriem n=11)

$$U = (3 - 5.5) / \sqrt{2.75} = -1.5076$$

Krok 4: Zhodnotenie

$$F^{-1}(1-0.05/2) = F^{-1}(0.975) = 1.96 \quad - \text{kritická hranica.}$$

$$1.5076 < 1.96 \quad - \text{sme v normále, hypotézu nezamietame}$$

Výsledok vyšiel opäť rovnako ako pri Wilcoxonovej metóde. Samozrejme, sú príklady, kde sa obe metódy môžu v hodnotení rozchádzať (jedná tesne zamietne a druhá tesne nezamietne).

c) χ^2 -test dobrej zhody

Týmto testom porovnávame hypotetické rozdelenie diskkrétnej náhodnej veličiny s experimentálne zistenými relatívnymi početnosťami jej hodnôt.

Príklad 5

Vezmeme 5 kusov 1-eurových mincí a hodíme na stôl. Zrátame počet jednotiek a zapíšeme. Tento elementárny experiment opakujeme 30-krát. Výsledky (tj. koľkokrát padol aký počet jednotiek) zhrnieme do tabuľky početností:

x_i (počet jednotiek)	0	1	2	3	4	5	spolu $n=$
n_i (početnosť)	1	5	5	10	8	1	30

Na hladine významnosti $\alpha=0.05$ testujte hypotézu, že mince sú „pocitivé“ (tj. číslo a znak padajú 50:50).

Riešenie:

Hypotézu, ktorú ideme skúmať, odborne preformulujeme – otázkou bude, či zistené početnosti sú v súlade s hodnotami presne generovanými binomickým rozdelením, ktoré je teoretickým základom experimentu. Inými slovami, či odchýlky od ideálu binomického rozdelenia vieme zvaliť na náhodu alebo nie – ak nie, zrejme nejaké mince nebudú „pocitivé“ alebo hádzanie nebolo dosť náhodné.

Porovnáme získané početnosti s teoreticky očakávanými hodnotami. Pozor, vôbec sa nesmieme dať vyrušovať tým, že očakávané hodnoty nebudú celočíselné. Do tabuľky teraz zapíšeme pod seba:

Experimentálne získané početnosti	n_i
Pravdepodobnostná funkcia binomického rozdelenia pre 5 mincí	p_i
Očakávané početnosti podľa teoretického modelu	$p_i \cdot n$
Rozdiely medzi realitou a teóriou	$\Delta_i = n \cdot p_i - n_i $
Relatívne rozdiely medzi realitou a teóriou	$\kappa_i = (\Delta_i)^2 / (p_i \cdot n)$

$x_i = i$	0	1	2	3	4	5	spolu
n_i	1	5	5	10	8	1	30
p_i	$1/32$	$5/32$	$10/32$	$10/32$	$5/32$	$1/32$	1
$p_i \cdot n$	0.9375	4.6875	9.375	9.375	4.6875	0.9375	30
Δ_i	0.0625	0.3125	4.375	0.625	3.3125	0.0625	
κ_i	0.0041667	0.020833	2.0417	0.041667	2.3408	0.0041667	4.4533

Kľúčová hodnota C, ktorá o všetkom rozhodne, je súčtom čísel v poslednom riadku:

$$C = 4.4533$$

V tabuľkách chí-kvadrátu zaostríme zrak na piaty riadok (počet stĺpcov v našej tabuľke mínus 1):

$\chi^2(v)$										
v	0,005	0,01	0,025	0,05	0,1	0,9	0,95	0,975	0,99	0,995
5	0,41	0,55	0,83	1,15	1,61	9,24	11,07	12,83	15,09	16,75

Hodnota C pri $\alpha=0.05$ musí ostať v bezpečnej zóne, tj. nesmie prekročiť kritickú hodnotu 11.07, zodpovedajúcu stĺpcu $1-\alpha=0.95$. Ako vidíme, naše C je „s prehľadom“ v bezpečnej zóne, takže hypotézu nezamietame.

Príklad 6

Vezmeme 4 kusy 1-eurových mincí a hodíme na stôl. Zrátame počet jednotiek a zapíšeme. Tento elementárny experiment opakujeme 300-krát. Výsledky (tj. koľkokrát padol aký počet jednotiek) zhrnieme do tabuľky početností:

x_i (počet jednotiek)	0	1	2	3	4	spolu $n=$
n_i (početnosť)	29	95	95	64	17	300

Na hladine významnosti $\alpha=0.05$ testujte hypotézu, že mince sú „pocťivé“ (tj. číslo a znak padajú 50:50).

Riešenie:

Postup je rovnaký ako v predošlom príklade. Vzhľadom na vysoký počet pokusov (300) budeme za odchýlky od očakávaní „platiť“ viac. Miesto tabuľky napíšeme len pod seba príslušné vektory:

n_i [29 95 95 64 17]

p_i [1 4 6 4 1] / 16 (binomické rozdelenie – Pascalov trojuholník, 4 mince)

$n \cdot p_i$ [18.750 75.000 112.500 75.000 18.750]

Δ_i [10.25 20 17.5 11 1.75]

κ_i [5.60333 5.33333 2.72222 1.61333 0.16333]

$C = 15.436$

V tabuľkách skenujeme riadok 4:

	$\chi^2(v)$									
v	0,005	0,01	0,025	0,05	0,1	0,9	0,95	0,975	0,99	0,995
4	0,21	0,30	0,48	0,71	1,06	7,78	9,49	11,14	13,28	14,86

Vidíme, že C je nad kritickou hodnotou 9.49 (pre $\alpha=0.05$), a dokonca aj pre $\alpha=0.005$ by ešte bolo mimo bezpečnej zóny. Hypotézu zamietame na predpísanej hladine 0.05, ale keby sme mali vyššie ambície, dalo by sa zamietnuť aj na 0.005.

Príklad 7

Experimentálny biológ v jednu decembrovú nedeľu poobede sledoval a zaznamenával čas, ktorý strávili nadšenci vo vode na neďalekom jazere. Výsledky zaznamenal do tabuľky početností:

Čas (min.)	(0,3]	(3,6]	(6,9]	(9,12]	(12,15]	(15,18]	(18,21]	(21,24]
Počet	9	4	15	17	10	7	3	1

Biológ na základe údajov predpokladá, že náhodná veličina, ktorá určuje čas pobytu vo vode, sa riadi normálnym rozdelením $N(10,25)$. Na hladine 5% testujte túto hypotézu.

Riešenie:

Najprv potrebujeme zistiť, aké počty nadšencov by sa v jednotlivých intervaloch vyskytli, ak by všetko išlo podľa hypotetického $N(10,25)$.

V prvom kroku vypočítame pre každý interval $(a,b]$ hodnotu $F(b)-F(a)$ – to bude ten povestný riadok p_i . Dá sa to z tabuliek, z on-line kalkulačky alebo nasledovne (Octave):

$x=[0\ 3\ 6\ 9\ 12\ 15\ 18\ 21\ 24]$

$xL=x(1:8)$

$xP=x(2:9)$

$\text{normcdf}(xP, 10, 5) - \text{normcdf}(xL, 10, 5)$

ľavé hranice intervalov

pravé hranice intervalov

p_i [0.058007 0.131099 0.208885 0.234681 0.185923 0.103856 0.040896 0.011348]

Očakávané počty podľa hypotézy, tj. $n \cdot p_i$ (kde $n=66$):

$n \cdot p_i$ [3.82843 8.65252 13.78640 15.48898 12.27092 6.85449 2.69913 0.74899]

Rozdiely

Δ_i [5.17157 4.65252 1.21360 1.51102 2.27092 0.14551 0.30087 0.25101]

κ_i [6.9859244 2.5016897 0.1068312 0.1474077 0.4202676 0.0030888 0.0335388 0.0841221]

$S = 10.283$

Sledujeme 7. riadok v tabuľke pre chí-kvadrát:

$\chi^2(v)$										
v	0,005	0,01	0,025	0,05	0,1	0,9	0,95	0,975	0,99	0,995
7	0,99	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,48	20,28

Hodnota C je menšia (aj s rezervou) voči predpísanej 14.07. Na hladine 5% preto hypotézu nezamietame.

Neriešené príklady

1a. Na odbernom mieste vo Svetlých Ležiakovciach zaznamenali v priebehu 14 dní nasledujúce hodnoty pozitivity testov (v percentách).¹

[4.7 5.1 14.7 7.2 6.1 2.7 3.1 19.1 11.2 9.7 4.7 3.1 8.6 4.7]

Starosta obce tvrdí, že dlhodobo dosahujú pozitivitu 4%, zatiaľčo krčmár je presvedčený, že pozitivita je 5%. Považujme tieto čísla za hypotetický medián všetkých hodnôt.

Na hladine významnosti 0.05 *Wilcoxonovým testom* overte mediánovú hypotézu

a) starostu.

b) krčmára.

1b. Na hladine významnosti 0.05 *znamienkovým testom* overte mediánovú hypotézu

a) starostu.

b) krčmára.

2. Zmrzlinársky reťazec XYZ ponúka 4 druhy zmrzliny. Doterajšia štruktúra predaja bola nasledovná: vanilková 62%, čokoládová 18%, jahodová 12% a pistáciová 8%.

Včera otvorili prevádzku v Prievidzi. V prvý deň sa predalo 120 vanilkových, 40 čokoládových, 18 jahodových a 22 pistáciových.

Na základe vhodného štatistického testu (na hladine 5%) sa vyjadrite k možnej zhode alebo odlišnosti štruktúry predaja v Prievidzi a celkovou štruktúrou.

3. Experimentálny klimatológ 40 rokov meral denne teplotu vody v jazere za svojím observatóriom. Na Mikuláša zostavil tabuľku početností teplôt nameraných práve v tento deň:

Teplota	(0,1]	(1,2]	(2,3]	(3,4]	(4,5]	(5,6]	(6,7]	(7,8]	(8,9]
Počet záznamov	0	1	2	4	7	9	11	3	3

Na hladine významnosti 5% (prípadne skúste aj iné hladiny) testujte hypotézu, že skúmanú náhodnú veličinu určuje normálne rozdelenie $N(6,2)$.

¹ V rozpore s reálnymi skúsenosťami predpokladáme, že tu miera pozitivity nezávisí od vonkajších faktorov, takže vybraných 14 dní môže adekvátne reprezentovať aj dlhšie obdobie meraní.