Odpovědi (téma 11)

1.1 Testuje, zda mediány dvou vzorků jsou totožné. Jinými slovy, zda oba vzorky pochází z jedné populace. H_0: $\sim\mu_1 = \sim\mu_2$.

1.2 Testuje, zda se populační medián (parametr) rovná zvolené hodnotě *k* (často 0). H_0: $\sim\mu = k$ ($\sim\mu$ je populační medián, *k* je zvolená hodnota, konstanta).

2.1 všechny čtyři, a-d

2.2 b

2.3 b

2.4 c

2.5 c

2.6 ano

2.7 b

2.8 ${}_{0,95}X_7^2 = 14,07$

2.9 Ne, p > 0,05; ${}_{0,95}X_4^2 = 9,49$

3.1 0,5

3.2 0,05

3.3 cca u 2% průzkumů

3.4 0,08 ; 0,04; 0,02

3.5 Vzroste-li velikost vzorku 4x, směrodatná chyba relativní četnosti klesne na polovinu.

4.1 (0,4; 0,6)

4.2 c

5.1 Pokud počet sourozenců považujeme za poměrovou škálu, můžeme ke srovnání vlastníků s nevlastníky titulu použít t-test pro nezávislé skupiny. Pokud si uvědomíme vysoké pozitivní zešikmení rozložení počtu sourozenců, můžeme se na malém vzorku přiklonit k neparametrickému testu pro srovnání dvou nezávislých skupin – Mann-Whitney U. Pokud bychom brali proměnnou počet sourozenců přesně tak, jak je v zadání, tj. kategoriální proměnnou s 5 kategoriemi: „0“, „1“, „2“, „3“ a „4 a více“, pak bychom použili chí-kvadrát test nezávislosti.

5.2 Úroveň dosaženého vzdělání je obvykle ordinální proměnná. Pohlaví je dichotomická – dává 2 porovnávané nezávislé skupiny. Takže M-W U. Pokud vaše zdůvodněné soudy přisuzovaly proměnné úroveň dosaženého vzdělání jinou úroveň měření, akceptovali jsme to.

6.1 0,25

6.2 ${}_{0,95}X_3^2 = 7,82$

6.3 ano, ano, ne

7.

2. $\chi^2 = \dfrac{110(255 - 1520)^2}{(55)(55)(53)(57)} = 19.26$

$\chi^2_{crit} = 3.84 \qquad \chi^2(1, N = 110) = 19.26, p < .05$

8.1 – 8.4

3. a. $df = 1$
   b. $df = 6$
   c. $df = 12$
   d. $df = 2$

9.1 Ho: Není vztah mezi typem nálepky a zastavením policistou.
H1: Proměnné jsou ve vztahu (nejsou nezávislé)

9.2 – 9.4

b. and c.

	Stop Brutality Sticker	*Smile Sticker*	
Stopped	18	5	23
Not Stopped	7	20	27
	25	25	$N = 50$

f_o	f_e	$(f_o - f_e)^2$	$(f_o - f_e)^2/f_e$
18	11.50	42.25	3.67
5	11.50	42.25	3.67
7	13.50	42.25	3.13
20	13.50	42.25	3.13
			$\chi^2 = 13.60$

$\chi^2_{crit} = 3.84$

Řidiči s nálepkami o brutalitě jsou zastavování signifikantně více častěji než ti, co mají nálepky s usmíváním $\chi^2 (1, N = 50) = 13{,}60$, $p < 0{,}05$.

10.

5. a. (f_o)

30	50	20	20	120
10	30	40	20	100
40	80	60	40	$N = 220$

(f_e)

21.82	43.64	32.73	21.82
18.18	36.36	27.27	18.18

b. (f_0)

7	7	14
5	11	16
12	18	$N = 30$

(f_e)

5.60	8.40
6.40	9.60

11.

8.

	Sun.	Mon.	Tues.	Wed.	Thurs.	Fri.	Sat.
f_o	56	29	17	22	25	15	33
f_e	28.14	28.14	28.14	28.14	28.14	28.14	28.14

f_o	f_e	$(f_o - f_e)^2$	$(f_o - f_e)^2/f_e$
56	28.14	776.18	27.58
29	28.14	.74	.03
17	28.14	124.10	4.41
22	28.14	37.70	1.34
25	28.14	9.86	.35
15	28.14	172.66	6.14
33	28.14	23.62	.84
			$\chi^2 = 40.69$

$\chi^2_{crit} = 12.59$

$\chi^2(6, N = 197) = 40.69, p < .05$

12. Ne. Porušeny jsou předpoklady nezávislosti pozorování, neboť některé subjekty jsou reprezentovány ve více než jednom políčku.

13. Existuje asociace mezi diabetem a prodlouženým hojením ran, neboť u diabetiků se častěji vyskytuje protrahované hojení, χ^2 (df=1) = 137,08, CHIDIST(137;1)= 1,2 . 10^{-31} , p < 0,05.

14. Není zde žádný odlišující efekt jednotlivých druhů léčby χ^2 (df=3) = 0,75, =CHIDIST(0,75;1)=0,86, p > 0,05.

15.1 H_0: $Md_{KBT} = Md_{PA}$ (=6,5) H_1: $Md_{KBT} \neq Md_{PA}$ (*Md* je zde parametr)

15.2 Jde o 2 nezávislé skupiny, pořadová data, tj. Mann-Whitney U (nebo mediánový test, Wilcoxonův test pro nezávislé výběry)
Výsledky testu Manna-Whitneyho v podání SPSS:

Ranks	skupina	N	Mean Rank	Sum of Ranks
poradi	KBT	6	8,50	51,00
	PA	6	4,50	27,00
	Total	12		

Test Statistics(b)	poradi
Mann-Whitney U	6,000
Wilcoxon W	27,000
Z	-1,922
Asymp. Sig. (2-tailed)	,055
Exact Sig. [2*(1-tailed Sig.)]	,065(a)
Exact Sig. (2-tailed)	,065
Exact Sig. (1-tailed)	,032
Point Probability	,012

a Not corrected for ties.
b Grouping Variable: skupina

15.3 Na 5% hladině významnosti nemůžeme nulovou hypotézu zamítnout, tj. musíme se držet toho, že rozdíl mezi skupinami se nám nepodařilo prokázat. Kdyby mezi typy výcvikových skupin nebyly rozdíly v dovednosti studentů dělat rozhovory, pak ty rozdíly, které nám vyšly (studenti KBT v našem vzorku byli lepší než PA) mohly být způsobeny náhodou (=výběrová chyba) s pravděpodobností 0,055 (nebo 0,065 podle přesnosti určení). Nicméně pravděpodobnost takto extrémního nebo extrémnějšího výsledku je poměrně malá a od zvolené hladiny významnosti se příliš neliší. Proto by bylo dobré pokus zopakovat, ideálně na větším vzorku.

16.

1. f_o	f_e	$(f_o - f_e)^2$	$(f_o - f_e)^2/f_e$	R
70	79.60	92.16	1.16	−1.08
160	159.20	.64	.004	.06
168	159.20	77.44	.49	.70
30	20.40	92.16	4.52	2.13
40	40.80	.64	.016	−.13
32	40.80	77.44	1.90	−1.38
			$\chi^2 = \mathbf{8.09}$	
			$\chi^2_{crit} = 5.99$	

Vidíme, že je zde více neúspěšných odpovědí na hypnózu, než by bylo očekáváno na základě náhody, χ^2 (2, N = 500) = 8,09. Pole neúspěšná odpověď na hypnózu dává signifikantní přínos k signifikantnímu χ^2.

17.1 s_p=√[*10*(100-*10*)/*100*]= √9=3
(10-normsinv(0,025)*s_p; 10+normsinv(0,025)*s_p) = (10-1,96*3; 10+1,96*3)=(4; 16)

17.2 s_p=√[*90*(100-*90*)/*100*]= √9=3
(90-normsinv(0,025)*s_p; 90+normsinv(0,025)*s_p) = (90-1,96*3; 90+1,96*3)=(84; 96)

18.1 chí-kvadrát

18.2
A. f = 40 f_0 = 33,3 $(f-f_0)^2/f_0$ = 1,33
B. f = 32 f_0 = 33,3 $(f-f_0)^2/f_0$ = 0,05
C. f = 28 f_0 = 33,3 $(f-f_0)^2/f_0$ = 0,85 χ^2 = 2,24
χ^2_{crit} =CHIINV(0,05;2) = 5,99
χ^2(2) = 2,24; p > 0,05 – nulová hypotéza nebyla zamítnuta, rozdíl mezi kandidáty není na 5% hladině statisticky významný

19.1 H_0: $\sim\mu_{starší} = \sim\mu_{mladší}$; H: $\sim\mu_{starší} \neq \sim\mu_{mladší}$

19.2 důvody, proč byl v tomto případě zvolen neparametrický Wilcoxonův test, mohou být v zásadě dva: (1) důvěra ve vztahu s rodiči je výzkumníkem chápána jako ordinální proměnná, (2) pokud byla data sebrána na adolescentech docházejících do terapie, lze očekávat problematické rozložení hodnot (výrazná nenormalita + výskyt outlierů); variantu pro dva závislé výběry bylo třeba použít, protože porovnávané skupiny na sobě nejsou nezávislé (jde o sourozenecké dvojice); oproti jiným testům v této kategorii (např. znaménkovému testu) má Wilcoxonův test větší statistickou sílu.

19.3 mladší sourozenci mají signifikantně větší důvěru ve vztahu s rodiči než starší

19.4 na základě prezentovaných statistik nelze o velikosti případného rozdílu nic určit

19.5 lze použít znaménkový test pro dva závislé výběry; pokud bychom chápali důvěru jako intervalovou či poměrovou proměnnou, potom by bylo možné zvážit i použití párového t-testu

20.

	Občasní hráči	Pravidelní hráči	Závislí hráči	Σ
Skupina 1	10 (f_e=15)	20 (f_e=25)	30 (f_e=20)	60
Skupina 2	20 (f_e=15)	30 (f_e=25)	10 (f_e=30)	60
Σ	30	50	40	120

a. χ^2=15,33 df=2, p=0,000; H_0 zamietame
b. Cramerovo V = 0,357
c. rozdiel medzi skupinami sa dal spočítať aj neparametrickým testom pre porovnanie dvoch nezávislých súborov v jednej premennej, a to konkrétne Mann-Whitneyho U testom, nakoľko kategorizovaná premenná čas trávený hraním hier predstavuje ordinálnu premennú.

21. Testujeme, či farba vybranej farbičky je volená rovnako často. Použijeme chí kvadrát test dobrej zhody:

Červená	245
Čierna	225

Modrá	225
Žltá	305

Očakávané frekvencie budú v danom prípade mať hodnotu 250, chí kvadrát bude rovný 17,2 pri 3 stupňoch voľnosti a dosiahnutá signifikancia bude rovná 0,000643. Nulovú hypotézu zamietneme, farby nie sú volené rovnako často.

22. Čím sú stupne voľnosti vyššie v rozložení χ^2, tým sa toto rozloženie bude viac a viac podobať normálnemu rozloženiu. Takže z daných možností bude správna odpoveď d.

23. Každá farba predstavuje ¼ z balíčka a potom každá očakávaná (expected) početnosť každej farby cukríkov je (1/4)*40=10.

24. Podľa vzorca: $(10-8)^2/10 + (10-5)^2/10 + (10-12)^2/10 + (10-15)^2/10 = 5,8$

25.

a) očakávaná početnosť žien v sociálnych vedách je počítaná ako súčin celkového počtu žien a celkového počtu odpovedí v spoločenských vedách delený celkovým počtom respondentov (22*34)/57 = 13,12

b) $\chi^2 = 2,2$