**Odpovědi (téma 11)**

1.1 Testuje, zda mediány dvou vzorků jsou totožné. Jinými slovy, zda oba vzorky pochází z jedné populace. *H0*: ~μ1 = ~μ2.

1.2 Testuje, zda se populační medián (parametr) rovná zvolené hodnotě *k* (často 0). *H0*: ~μ = *k* (~μ 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*χ*72 = 14,07

2.9 Ne, p > 0,05; 0,95*χ*42 = 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*χ*32 = 7,82

6.3 ano, ano, ne

7.



8.1 – 8.4



9.1 Ho: Není vztah mezi typem nálepky a zastavením policistou.

H1: Proměnné jsou ve vztahu (nejsou nezávislé)

9.2 – 9.4

Řidiči s nálepkami o brutalitě jsou zastavování signifikantně více častěji než ti, co mají nálepky s usmíváním χ2 (1, N = 50) = 13,60, p < 0,05.

10.



11.







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, **2 (df=1) = 137,08, CHIDIST(137;1)= 1,2 . 10-31 , p < 0,05.

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

15.1 H0: *Md*KBT = *Md*PA (=6,5) H1: *Md*KBT ≠ *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.



Vidíme, že je zde více neúspěšných odpovědí na hypnózu, než by bylo očekáváno na základě náhody, χ2 (2, N = 500) = 8,09. Pole neúspěšná odpověď na hypnózu dává signifikantní přínos k signifikantnímu χ2.

17.1 sp=√[*10*(100-*10*)/*100*]= √9=3

(10-normsinv(0,025)\*sp; 10+normsinv(0,025)\*sp) = (10-1,96\*3; 10+1,96\*3)=(4; 16)

17.2 sp=√[*90*(100-*90*)/*100*]= √9=3

(90-normsinv(0,025)\*sp; 90+normsinv(0,025)\*sp) = (90-1,96\*3; 90+1,96\*3)=(84; 96)

18.1 chí-kvadrát

18.2

A. f = 40 f0 = 33,3 (f-f0)2/f0 = 1,33

B. f = 32 f0 = 33,3 (f-f0)2/f0 = 0,05

C. f = 28 f0 = 33,3 (f-f0)2/f0 = 0,85 2 = 2,24

2crit =CHIINV(0,05;2) = 5,99

2(2) = 2,24; p > 0,05 – nulová hypotéza nebyla zamítnuta, rozdíl mezi kandidáty není na 5% hladině statisticky významný

19.1 H0: ~μstarší = ~μmladší; H: ~μstarší ≠ ~μmladší

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 (fe=15) | 20 (fe=25) | 30 (fe=20) | 60 |
| Skupina 2 | 20 (fe=15) | 30 (fe=25) | 10 (fe=30) | 60 |
| Σ | 30 | 50 | 40 | 120 |

a. *χ*2=15,33 df=2, p=0,000; H0 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 χ2, tým sa toto rozloženie bude viac a viac podobať normálnemu rozloženiu. Takže z daných možností bude správna odpoveď d.

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

b) Podľa vzorca: (10-8)2/10 + (10-5)2/10 + (10-12)2/10 + (10-15)2/10 = 5,8

c) Očekávané četnosti by podle této H byly č4 z4 o16 m16. Data nejsou v souladu s touto hypotézou (**2(3, N=40)=8,5, *p*=0,037)

d) Očekávané četnosti by podle této H byly č4 ostatní36. Data nejsou v souladu s touto hypotézou (**2(1, N=40)=4,22, *p*=0,040) a zdá se, že je na místě obsah červených lépe prověřit..

25.

1. 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
2. *χ*2 = 2,2