**Seminář 9 – Volba statistických testů**

* Chceme zjistit, zda se seminární skupiny A a B od sebe liší ve výsledcích v 1. průběžné písemce ze statistiky.
* Chceme zjistit, zda 1. průběžná písemka ze statistiky byla stejně těžká jako 2. průběžná.
* Chceme zjistit, zda populační rozložení skórů 1. průběžné písemky má průměr 8 (pro který byl konstruován).
* Chceme zjistit podle známek v ISu, jestli je statistika stejně těžká jako vývojová psychologie 2.
* Chceme zjistit podle známek v ISu, zda je statistika stejně těžká pro muže a ženy.
* Chceme zjistit, zda jsou v populaci všechny základní barvy (b,čr,čv,z,m,ž,o,h) stejně oblíbené.
* Chceme zjistit, zda se kombinovaní a prezenční studenti psychologie liší v preferenci placeného vysokoškolského studia.
* Chceme na vzorku 30 rodin se dvěma školou povinnými dětmi zjistit, zda mladší i starší sourozenci jsou stejně populární ve své třídě.
* Chceme zjistit, zda výkonnost ve statistice (1. průběžná) roste s dobou přípravy (v hodinách).
* Chceme zjistit, zda platí, že čím více chodí lidé do kina, tím méně jsou pro školné na VŠ.
* Chceme zjistit, zda se milovníci různých základních barev liší ve výkonnosti ve statistice (1. průběžná).
* Chceme na vzorku 30 spokojených partnerů ověřit hypotézu, že ve spokojených vztazích se míra romantičnosti obou partnerů neliší.



**Část II. Příklady výstupů k jednotlivým testům.**

1. t-test pro nezávislé skupiny

Chceme zjistit, zda se seminární skupiny A a B liší ve výsledcích v 1. a 2. průběžné písemce.

|  |
| --- |
| **Group Statistics** |
|  | seminar | N | Mean | Std. Deviation | Std. Error Mean |
| p1 | A | 18 | 7,17 | 2,684 | ,633 |
| B | 12 | 5,83 | 2,406 | ,694 |
| p2 | A | 17 | 4,41 | 2,917 | ,707 |
| B | 9 | 5,67 | 3,571 | 1,190 |

|  |
| --- |
| **Independent Samples Test** |
|  | Levene's Test for Equality of Variances | t-test for Equality of Means |
| F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference |
| Lower | Upper |
| p1 | Equal variances assumed | ,606 | ,443 | 1,388 | 28 | ,176 | 1,333 | ,961 | -,635 | 3,302 |
| Equal variances not assumed |  |  | 1,419 | 25,480 | ,168 | 1,333 | ,939 | -,600 | 3,266 |
| p2 | Equal variances assumed | ,984 | ,331 | -,966 | 24 | ,343 | -1,255 | 1,298 | -3,935 | 1,425 |
| Equal variances not assumed |  |  | -,906 | 13,790 | ,380 | -1,255 | 1,385 | -4,229 | 1,719 |

2. párový t-test

Chceme zjistit, zda 1. průběžná písemka ze statistiky byla stejně těžká jako 2. průběžná.

|  |
| --- |
| **Paired Samples Statistics** |
|  | Mean | N | Std. Deviation | Std. Error Mean |
| Pair 1 | p1 | 6,00 | 69 | 2,776 | ,334 |
| p2 | 4,28 | 69 | 3,115 | ,375 |
| **Paired Samples Correlations** |
|  | N | Correlation | Sig. |
| Pair 1 | p1 & p2 | 69 | ,218 | ,072 |

|  |
| --- |
| **Paired Samples Test** |
|  | Paired Differences | t | df | Sig. (2-tailed) |
| Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference |
| Lower | Upper |
| Pair 1 | p1 - p2 | 1,725 | 3,694 | ,445 | ,837 | 2,612 | 3,878 | 68 | ,000 |

3. jednovýběrový t-test

Chceme zjistit, zda populační rozložení skórů 1. průběžné písemky má průměr 7.

**¨**

 **One-Sample Statistics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|   | N | Mean | Std. Deviation | Std. Error Mean |
| P1 | 84 | 7,39 | 3,840 | ,419 |



4. neparametrický test pro dva nezávislé výběry – Mann-Whitney U

Chceme zjistit, zda se středeční a čtvrteční seminární skupiny liší ve výsledcích v 1. průběžné písemce.

 … a nevěříme tak úplně dobře intervalovosti svého měření

 **Ranks**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|   | sk | N | Mean Rank | Sum of Ranks |
| P1 | A | 45 | 43,78 | 1970,00 |
| B | 39 | 41,03 | 1600,00 |
| Total | 84 |   |   |

 **Test Statistics(a)**

|  |  |
| --- | --- |
|   | P1 |
| Mann-Whitney U | 820,000 |
| Wilcoxon W | 1600,000 |
| Z | -,518 |
| Asymp. Sig. (2-tailed) | ,604 |

a Grouping Variable: sk

5. neparametrický párový test – Wilcoxon T

Chceme zjistit, zda 1. průběžná písemka ze statistiky byla stejně těžká jako 2. průběžná (loňská data).

 … a nevěříme tak úplně dobře intervalovosti svého měření





6. Chí-kvadrát test dobré shody

Chceme zjistit, zda jsou v populaci studentů odpůrci a příznivci školného zastoupeni rovnoměrně.





7. Chí kvadrát test rozdílu rozložení mezi dvěma populacemi / nezávislosti mezi dvěma kategoriálními proměnnými.

Chceme zjistit, zda je poměr příznivců/odpůrců stejný mezi prezenčními a kombinovanými studenty.







**Část III. Ruční počítání statistických testů**

A) t-test pro nezávislé skupiny

Chceme zjistit, zda se středeční a čtvrteční seminární skupiny liší ve výsledcích v 1. průběžné písemce.

|  |
| --- |
|  |
|  | seminar | N | Mean | Std. Deviation | Std. Error Mean |
| p1 | A | 18 | 7,17 | 2,684 | ,633 |
| B | 12 | 5,83 | 2,406 | ,694 |

1. *H*0: **s = **č neboli ** = **s – **č = 0 a hladinu významnosti zvolíme ** = 0,05

2. Rozdíl průměrů nezávislých skupin má *t*-rozložení s *n*1 + *n*2 – 2 stupni volnosti, středem v  ** a směrodatnou chybou 

3. Nyní spočítáme testovou statistiku, což je *t*, které vyjadřuje jak je zjištěný rozdíl veliký v jednotkách své směrodatné chyby.

 

4. Jaká je pravděpodobnost, že nám při náhodném výběru z *t*-rozložení s \_\_\_ stupni volnosti a průměrem 0 vyjde standardizovaná hodnota *t* =\_\_\_\_\_ nebo větší?

 2\*(1–T.DIST(\_\_\_\_\_;\_\_\_\_\_;1)) =

5. Vyšla nám pravděpodobnost vyšší než je zvolená hladina statistické významnosti. To znamená, že kdyby byla nulová hypotéza platná, tak by tak velký rozdíl, jaký nám vyšel, mohl vyjít se \_\_\_\_% pravděpodobností. Nulovou hypotézu tedy na 5% hladině významnosti \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

6. Interval spolehlivosti

 *d* – 0,975*t*(\_\_\_)*s*d < ** < *d* + 0,975*t*(\_\_\_)*s*d

7. Co nám SPSS nespočítalo - velikost účinku – Cohenovo *d*

 B) Párový t-test (*t*-test pro korelované vzorky)

Chceme zjistit, zda 1. průběžná písemka ze statistiky byla stejně těžká jako 2. průběžná.

|  |
| --- |
|  |
|  | Mean | N | Std. Deviation | Std. Error Mean |
| Pair 1 | p1 | 6,00 | 69 | 2,776 | ,334 |
| p2 | 4,28 | 69 | 3,115 | ,375 |
| **Paired Samples Correlations** |
|  | N | Correlation | Sig. |
| Pair 1 | p1 & p2 | 69 | ,218 | ,072 |

1. *H*0: **s = **č neboli ** = **s – **č = 0 a hladinu významnosti zvolíme ** = 0,05

2. Rozdíl průměrů nezávislých skupin má *t*-rozložení s *n* – 1 stupni volnosti, středem v  ** a směrodatnou chybou 

3. Nyní spočítáme testovou statistiku, což je *t*, které vyjadřuje jak je zjištěný rozdíl veliký v jednotkách své směrodatné chyby (jinými slovy, rozdíl průměrů převedeme na standardizovaný skór *t*, což je něco jako *z*).

 

4. Jaká je pravděpodobnost, že nám při náhodném výběru z *t*-rozložení s 87 stupni volnosti a průměrem 0 vyjde standardizovaná hodnota 3,73 nebo větší?

 2\*(1–T.DIST(3,73;82;1) = 0,0003 (=TDIST(3,73;82;2)

5. Vyšla nám pravděpodobnost nižší než je zvolená hladina statistické významnosti. To znamená, že kdyby byla nulová hypotéza skutečně platná, tak by tak by pravděpodobnost toho, že nám vyjde tak velký nebo větší rozdíl, než jaký nám vyšel, byla velmi nízká cca 0,03%. Nulovou hypotézu tedy na 5% hladině významnosti zamítáme.

6. Interval spolehlivosti

 *d* – 0,975t(87)*s*d < ** < *d* + 0,975t(87)*s*d

7. Co nám SPSS nespočítalo - velikost účinku – Cohenovo *d*

 

C) Chí-kvadrátový test nezávislosti proměnných

Chceme zjistit, zda je poměr příznivců/odpůrců stejný mezi prezenčními a kombinovanými studenty.

 **typ\_studia \* skolne Crosstabulation**

|  |  |  |  |
| --- | --- | --- | --- |
|   |   | skolne | Total |
| pro | proti | pro |
| typ\_studia | prezenční | Count | 17 | 45 | 62 |
| Expected Count | 21,7 | 40,3 |  |
| kombinované | Count | 12 | 9 | 21 |
| Expected Count | 7,3 | 13,7 |  |
| Total | Count | 29 | 54 | 83 |
|  |  |  |  |

1. *H*0:  Kdyby bylo procento příznivců stejné mezi prezenčními i kombinovanými studenty (35% ku 65%), očekávali bychom *abcd* přibližně 22, 40, 7, 14. Nulová hypotéza je tedy, že mezi očekávanými četnostmi a skutečně získanými četnostmi není žádný rozdíl. Konkrétním vyjádřením těchto rozdílů je jejich speciální součet zvaný chí-kvadrát, jehož výběrové rozložení známe

 

Očekávaná hodnota (průměr) chí-kvadrát rozložení je rovna jeho stupňům volnosti ** = (*i*–1)(*j*–1)

*H*0: **2 ≤ **(ano, jednostranný test)**a hladinu významnosti zvolíme ** = 0,01

2. Spočítáme testovou statistiku

3. Jaká je pravděpodobnost **2 s jedním stupněm volnosti?

 1–CHISQ.DIST(6,1;1;1) = 0,0135 (=CHIDIST(6,1;1))

4. *H*0 na 1% hladině významnosti podržíme; rozdíly nejsou dost velké na to, aby se přihodily náhodou.

5. Interval spolehlivosti bychom mohli počítat pro jednotlivé relativní četnosti (viz níže) nebo pro poměr šancí (OR, odds ratio, viz PSY252).

6. Velikost účinku je zde např. *r*, nebo obecněji Cramerovo V, Pearsonův koeficient kontingence

 *r* = C) Interval spolehlivosti a test hypotézy o relativních četnostech

*p* má přibližně normální rozložení s průměrem ** a 

1. činitel v čitateli zohledňuje, jak velkou část populace máme ve vzorku. Je-li populace vzhledem k vzorku obrovská(nekonečná), nemusíme ho používat.



