

Seminář 6 – statistické testy

Část I. – Volba správného testu

- Chceme zjistit, zda se střeďeční a čtvrtední seminární skupiny 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 skóreů 1. průběžné písemky má průměr 10 (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 uvěřit hypotézu, že ve spokojených vztazích se míra romantičnosti obou partnerů neliší.

Úkol

- a) pro každou situaci najít ten správný test
- ... b) najít kód receptu Oseckých

Testy na rozdíly středních hodnot

Nominální závislá

- párový test: binomický znaménkový test
- nezávislé skupiny: χ^2 -kvadrát

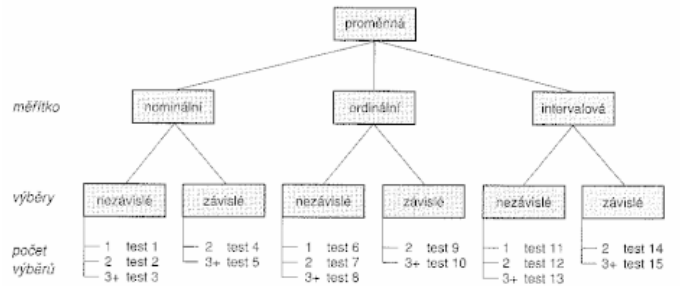
Ordinální závislá

- párový test: Wilcoxonovo T
- nezávislé skupiny: Mann-Whitney U

Intervalová závislá

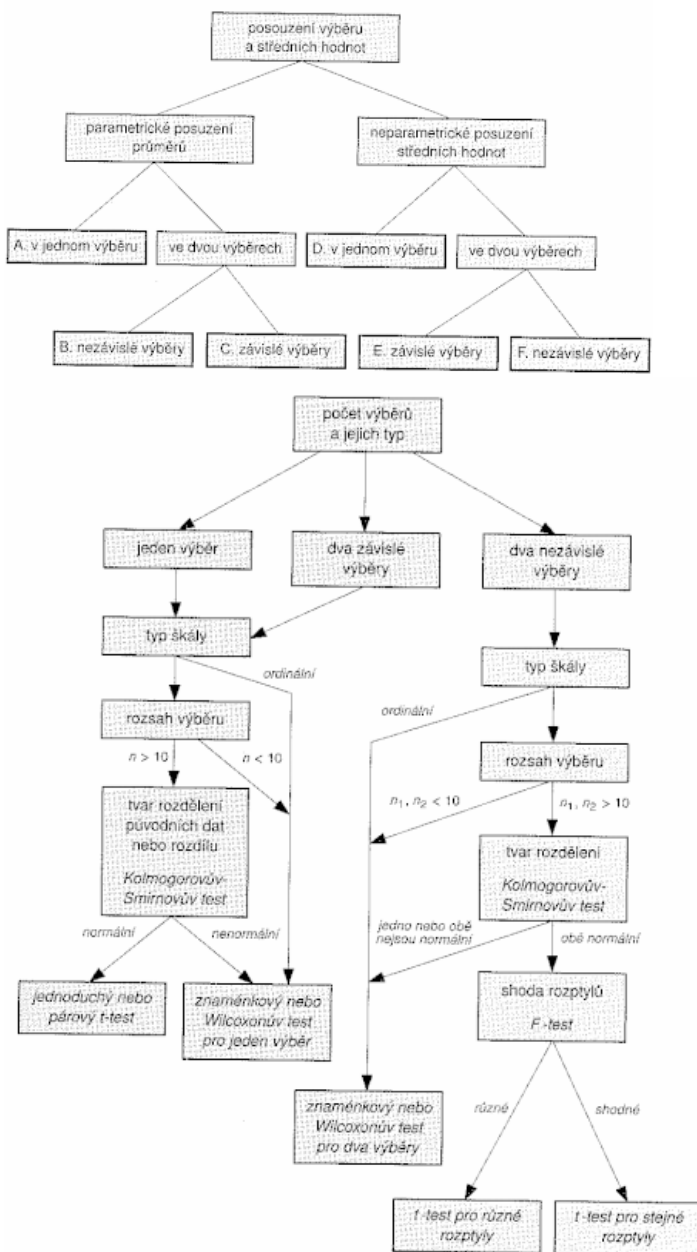
- párový test: párový t -test
- nezávislé skupiny:
 - známý rozptyl v populaci: z -test
 - neznámý rozptyl v populaci: t -test pro nezávislé skupiny
 - varianta pro stejné a nestejně rozptly mezi skupinami

AI sign test, chi-square, Wilcoxon T , Mann-Whitney U , paired(-samples) t -test (dependent, repeated measures), one-sample t -test, independent samples t -test



Tab. 12.1 Označení jednotlivých testů

Test	Kapitola	Strana	Název testu
1	6.2	304	χ^2 -test dobré shody
2	8.3.1	311	χ^2 -test homogenity
3	6.3.1	311	χ^2 -test homogenity
4	8.3.2 a 6.3.3	318 a 321	McNemarův test, Bowkerův test
5	8.3.3	321	Q test podle Cochran
6	6.4.2	223	znaménkový pořadový test podle Wilcoxon
7	6.4.6	229	Wilcoxonův test, Mann-Whitneův test
8	9.1.5	348	Jockheere-Terpstra test
9	6.4.2	223	znaménkový pořadový test podle Wilcoxon
10	9.3.1	360	Friedmanův test, Pegoův test
11	6.1	204	t -test pro jeden výběr, také 1 a 6
12	6.2.1-3	210-211	t -testy pro dva výběry, také 7
13	9.1	339	analýza rozptly
14	6.2.4	214	t -test pro párové rozdíly
15	9.3	357	dvoufaktorová analýza rozptly s opakovaním měření



Druhá z proměnných	Jedna z proměnných				
	spojitá (normální)	spojitá (ne normální)	ordinální pořadová	ordinální kategoriální	nominální dichotomická
spojitá (normální)	regrese, korelace (kap. 7)				
spojitá (ne normální)	regrese, pořadová korelace podle Spearmana (kap. 7, 7.2.6)	pořadová korelace podle Spearmana (kap. 7.2.6)			
ordinální pořadová	pořadová korelace podle Spearmana (kap. 7.2.6)	pořadová korelace podle Spearmana	pořadová korelace podle Spearmana		
ordinální kategoriální	Kendallův koeficient korelace (kap. 7.2.7)	Kendallův koeficient korelace	Kendallův koeficient korelace	Kendallův koeficient korelace	
nominální	analýza rozptly (kap. 9.1)	test Kruskala a Wallise (kap. 9.1.4)	test Kruskala a Wallise	χ^2 -test (kap. 6.3.1), Fisherův přesný test (kap. 6.3.1)	přesný Fisherův test, nebo χ^2 -test pro kontingenční tabulku
dichotomická	t -test (kap. 6.2), z -test (kap. 6.2)	z -test pro velké výběry (kap. 6.2.1), Wilcoxonův test (kap. 6.4.6)	Wilcoxonův test (kap. 6.4.6)	přesný Wilcoxonův test, χ^2 -test trendu v kontingenční tabulce	χ^2 -test, Fisherův přesný test, χ^2 -test, Fisherův přesný test

Část II. Příklady výstupů k jednotlivým testům.

1. t-test pro nezávislé skupiny

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

Group Statistics

sk_num	N	Mean	Std. Deviation	Std. Error Mean
p1 >= 3	48	8,0833	3,26055	,47062
< 3	53	10,2264	2,47014	,33930

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	2,387	,126	-3,744	99	,000	-2,14308	,57238	-3,27880	-1,00736
	Equal variances not assumed			-3,694	87,252	,000	-2,14308	,58018	-3,29620	-,98996

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	9,7045	88	2,65295	,28281
p2	8,3523	88	3,04389	,32448

Paired Samples Correlations

	N	Correlation	Sig.
Pair 1 p1 & p2	88	,295	,005

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	1,35227	3,39723	,36215	,63247	2,07208	3,734	87	,000

3. jednovýběrový t-test

Chceme zjistit, zda populační rozložení skóre 1. průběžné písemky má průměr 10.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
p1	101	9,2079	3,05391	,30387

One-Sample Test

		Test Value = 10					
		t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
						Lower	Upper
p1		-2,607	100	,011	-,79208	-1,3950	-,1892

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

Chceme zjistit, zda se středční a čtvrtční seminární skupiny liší ve výsledcích v 1. průběžné písemce.
... a nevěříme tak úplně dobře intervalovosti svého měření

Ranks

	den seminární skupiny	N	Mean Rank	Sum of Ranks
p1	středa	53	60,60	3212,00
	čtvrtek	48	40,40	1939,00
	Total	101		

Test Statistics^a

	p1
Mann-Whitney U	763,000
Wilcoxon W	1939,000
Z	-3,485
Asymp. Sig. (2-tailed)	,000

a. Grouping Variable: den seminární skupiny

5. neparametrický párový test – Wilcoxon T

Chceme zjistit, zda 1. průběžná písemka ze statistiky byla stejně těžká jako 2. průběžná.
... a nevěříme tak úplně dobře intervalovosti svého měření

Ranks

	N	Mean Rank	Sum of Ranks
p2 - p1	Negative Ranks	55 ^a	43,58
	Positive Ranks	26 ^b	35,54
	Ties	7 ^c	
	Total	88	

a. $p2 < p1$

b. $p2 > p1$

c. $p2 = p1$

Test Statistics^b

	p2 - p1
Z	-3,482 ^a
Asymp. Sig. (2-tailed)	,000

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

6. Chí-kvadrát test dobré shody

Chceme zjistit, zda jsou v populaci studentů odpůrci a příznivci školního zastoupení rovnoměrně.

skolne

	Observed N	Expected N	Residual
pro	29	41,5	-12,5
proti	54	41,5	12,5
Total	83		

Test Statistics

	skolne
Chi-Square ^a	7,530
df	1
Asymp. Sig.	,006

- a. 0 cells (,0%) have expected frequencies less than 5. The minimum expected cell frequency is 41,5.

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.

typ_studia * skolne Crosstabulation

		skolne		Total	
		pro	proti		
typ_ studia	perezdní	Count	17	45	62
		Expected Count	21,7	40,3	62,0
		% within typ_studia	27,4%	72,6%	100,0%
		% within skolne	58,6%	83,3%	74,7%
		% of Total	20,5%	54,2%	74,7%
		Residual	-4,7	4,7	
		Adjusted Residual	-2,5	2,5	
kombinované		Count	12	9	21
		Expected Count	7,3	13,7	21,0
		% within typ_studia	57,1%	42,9%	100,0%
		% within skolne	41,4%	16,7%	25,3%
		% of Total	14,5%	10,8%	25,3%
		Residual	4,7	-4,7	
		Adjusted Residual	2,5	-2,5	
Total		Count	29	54	83
		Expected Count	29,0	54,0	83,0
		% within typ_studia	34,9%	65,1%	100,0%
		% within skolne	100,0%	100,0%	100,0%
		% of Total	34,9%	65,1%	100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	6,097 ^b	1	,014		
Continuity Correction ^a	4,859	1	,027		
Likelihood Ratio	5,896	1	,015		
Fisher's Exact Test				,018	,015
Linear-by-Linear Association	6,023	1	,014		
N of Valid Cases	83				

a. Computed only for a 2x2 table

b. 0 cells (,0%) have expected count less than 5. The minimum expected count is 7,34.

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	-,271	,014
	Cramer's V	,271	,014
N of Valid Cases		83	

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Část III. Ruční počítání statistických testů

A) t-test pro nezávislé skupiny

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

Group Statistics

	sk_num	N	Mean	Std. Deviation	Std. Error Mean
p1	>= 3	48	8,0833	3,26055	,47062
	< 3	53	10,2264	2,47014	,33930

1. $H_0: \mu_s = \mu_{\check{c}}$ neboli $\delta = \mu_s - \mu_{\check{c}} = 0$ a hladinu významnosti zvolíme $\alpha = 0,05$

2. Rozdíl průměrů nezávislých skupin má t -rozložení s $n_1 + n_2 - 2$ stupni volnosti, středem v δ a

$$\text{směrodatnou chybou } s_d = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

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

$$t = \frac{m_s - m_{\check{c}}}{s_d}$$

4. Jaká je pravděpodobnost, že nám při náhodném výběru z t -rozložení s 99 stupni volnosti, průměrem 0 a směrodatnou chybou (odchylkou) 0,57 vyjde hodnota 3,74 nebo větší?

$$\text{TDIST}(3,74;99;2) = 0,000308$$

5. Vyšla nám pravděpodobnost menší než je zvolená hladina statistické významnosti. To znamená, že kdyby byla nulová hypotéza skutečně platná, bylo by ve velmi nepravděpodobné, aby nám vyšel tak velký rozdíl, jaký nám vyšel. Nulovou hypotézu tedy na 5% hladině významnosti zamítáme.

6. Interval spolehlivosti

$$d - 0,975t(99)s_d < \delta < d + 0,975t(99)s_d$$

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

$$\text{Cohenovo } d = \frac{d}{s_{pooled}}$$

B) 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

$$\chi^2 = \sum \frac{(f - f_o)^2}{f_o}$$

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

$H_0: \chi^2 > \nu$ (ano, jednostranný test) a hladinu významnosti zvolíme $\alpha = 0,05$

2. Spočítáme testovou statistiku

3. Jaká je pravděpodobnost, χ^2 s jedním stupněm volnosti?

$$\text{CHIDIST}(6,1;1)=0,0135$$

4. H_0 na 5% hladině významnosti zamítáme; rozdíly jsou příliš velké na to, aby se přihodily náhodou.

5. Interval spolehlivosti zde nepočítáme.

6. Velikost účinku je zde např. r_ϕ , nebo Cramerovo V (interpretujeme jako r^2)

$$r_\phi = \sqrt{\frac{\chi^2}{n}}$$

C) Interval spolehlivosti a test hypotézy o relativních četnostech

p má přibližně normální rozložení s průměrem π a $\sigma_p = \sqrt{\frac{(1 - \frac{n}{N}) \cdot p \cdot (1 - p)}{n}}$

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

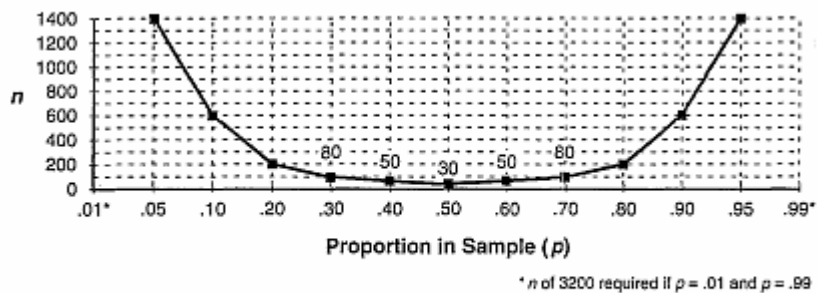


FIGURE 13.4 Minimum Sample Size (n) Needed for Use of the Normal Approximation When Setting Confidence Intervals for the Proportion (π) in the Population.

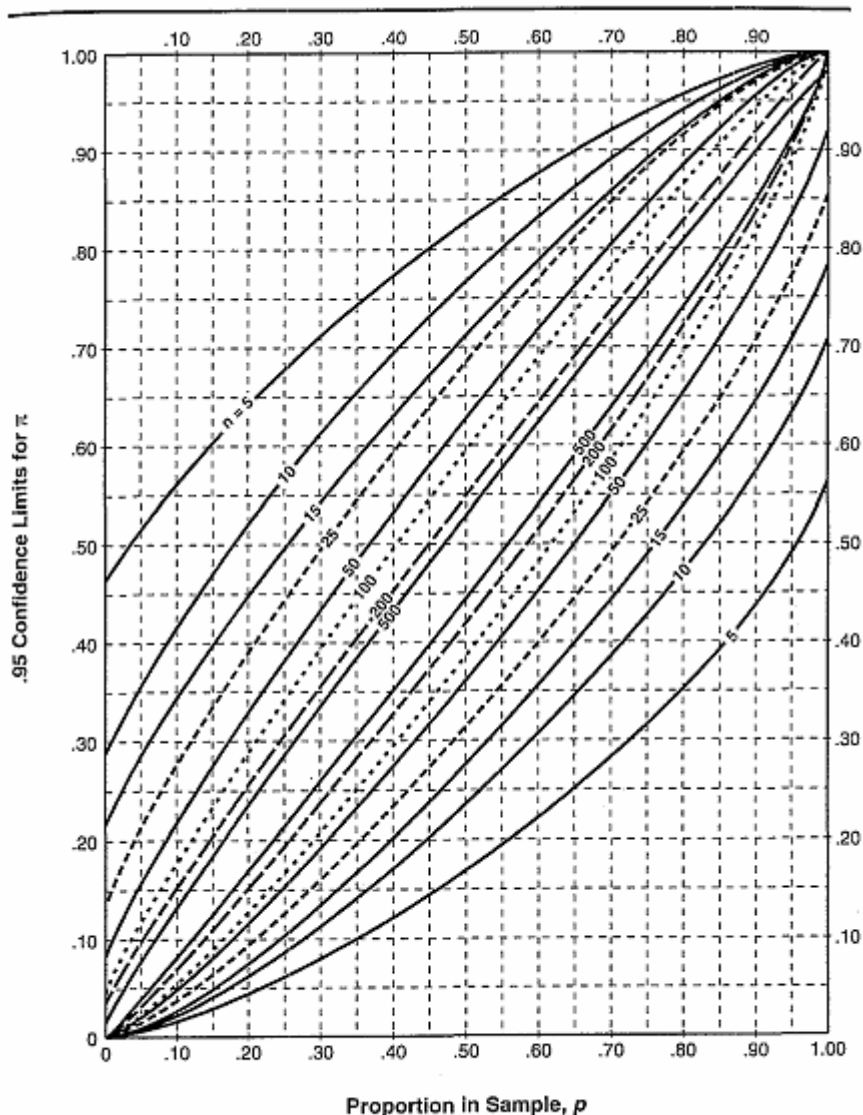


FIGURE 13.5 Graph for .95 Confidence Limits for the Parameter, π , from p and n .