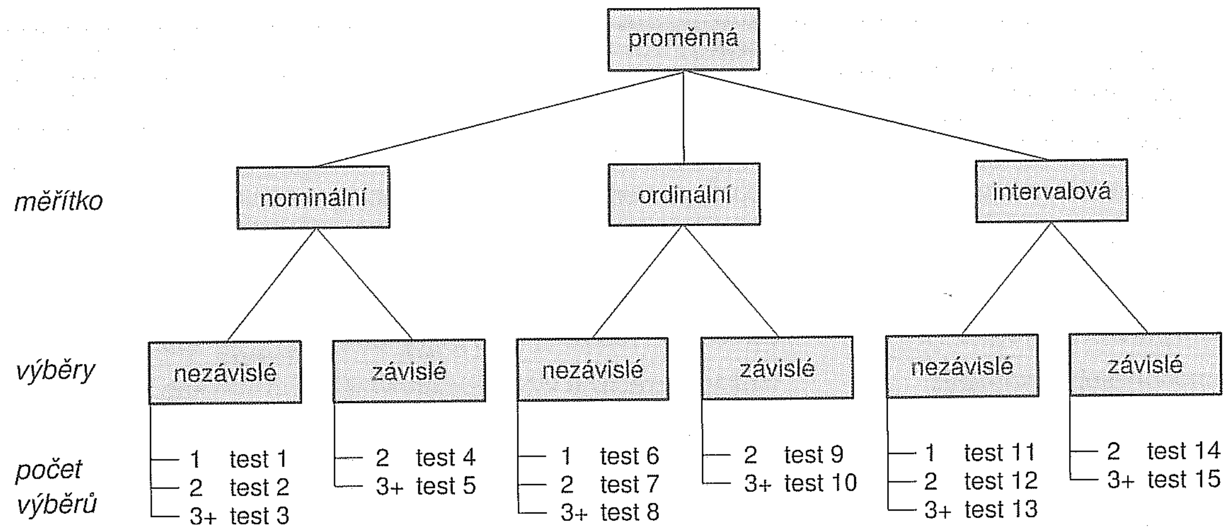
Seminář 9 – Volba statistických testů

- 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.

Obr. 12.1 Volba statistického testu pro porovnání skupin a hodnocení jedné skupiny



Předpokládá se splnění ostatních podmínek procedur.

Tab. 12.1 Označení jednotlivých testů

Test	Kapitola	Strana	Název testu
1	8.2	314	χ^2-test dobré shody
2	8.3.1	321	χ^2-test homogenity
3	8.3.1	321	χ^2-test homogenity
4	8.3.2 a 8.3.3	328 a 331	McNemarův test, Bowkerův test
5	8.3.3	331	Q test podle Cochrana
6	6.4.2	233	znaménkový pořadový test podle Wilcoxona
7	6.4.6	239	Wilcoxonův test, Mannův-Whitneyův test
8	9.1.5	358	Jonckheerův-Terpstrův test Kruskal-Wallis H
9	6.4.2	233	znaménkový pořadový test podle Wilcoxona
10	9.3.1	370	Friedmanův test, Pageův test
11	6.1	214	*t*-test pro jeden výběr, také 1 a 6
12	6.2.1–3	220–221	*t*-testy pro dva výběry, také 7
13	9.1	349	analýza rozptylu
14	6.2.4	224	*t*-test pro párové rozdíly
15	9.3	367	dvoufaktorová analýza rozptylu s opakováním měření

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

One-Sample Test

	Test Value = 10					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
P1	-6,222	83	,000	-2,607	-3,44	-1,77

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

Ranks

		N	Mean Rank	Sum of Ranks
p2 - p1	Negative Ranks	55[a]	43,58	2397,00
	Positive Ranks	26[b]	35,54	924,00
	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 školného zastoupeni 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		
			pro	proti	Total
typ_studia	pereznční	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 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: $\mu_s = \mu_č$ neboli $\delta = \mu_s - \mu_č = 0$ 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 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.

$$t = \frac{m_s - m_č}{s_d}$$

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

$d - {}_{0,975}t(\_\_\_)s_d < \delta < d + {}_{0,975}t(\_\_\_)s_d$

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

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

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: $\mu_s = \mu_č$ neboli $\delta = \mu_s - \mu_č = 0$ a hladinu významnosti zvolíme $\alpha = 0{,}05$

2. Rozdíl průměrů nezávislých skupin má *t*-rozložení s $n - 1$ stupni volnosti, středem v δ a směrodatnou chybou $s_d = \sqrt{\frac{1}{n}(s_1^2 + s_2^2 - 2rs_1s_2)}$

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*).

$$t = \frac{m_1 - m_2}{s_d}$$

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

$$d - {}_{0,975}t(87)s_d < \delta < d + {}_{0,975}t(87)s_d$$

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

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

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

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

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

H_0: $\chi^2 \leq \nu$ (ano, jednostranný test) a hladinu významnosti zvolíme $\alpha = 0{,}01$

2. Spočítáme testovou statistiku

3. Jaká je pravděpodobnost χ^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_ϕ, nebo obecněji Cramerovo V, Pearsonův koeficient kontingence

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

C) Interval spolehlivosti a test hypotézy o relativních četnostech

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

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

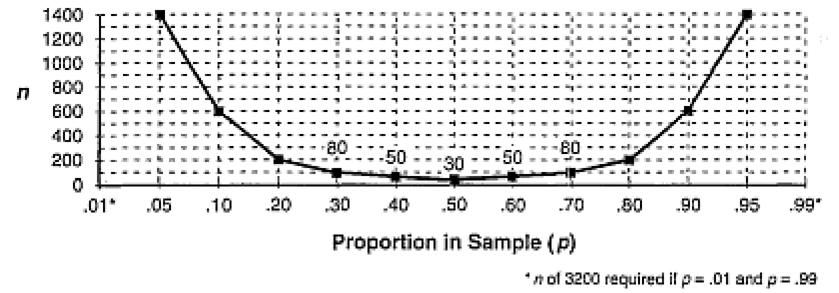


FIGURE 13.4 Minimum Sample Size (n) Needed for Use of the Normal Approximation When Setting Confidence Intervals for the Proportion (π) in the Population.

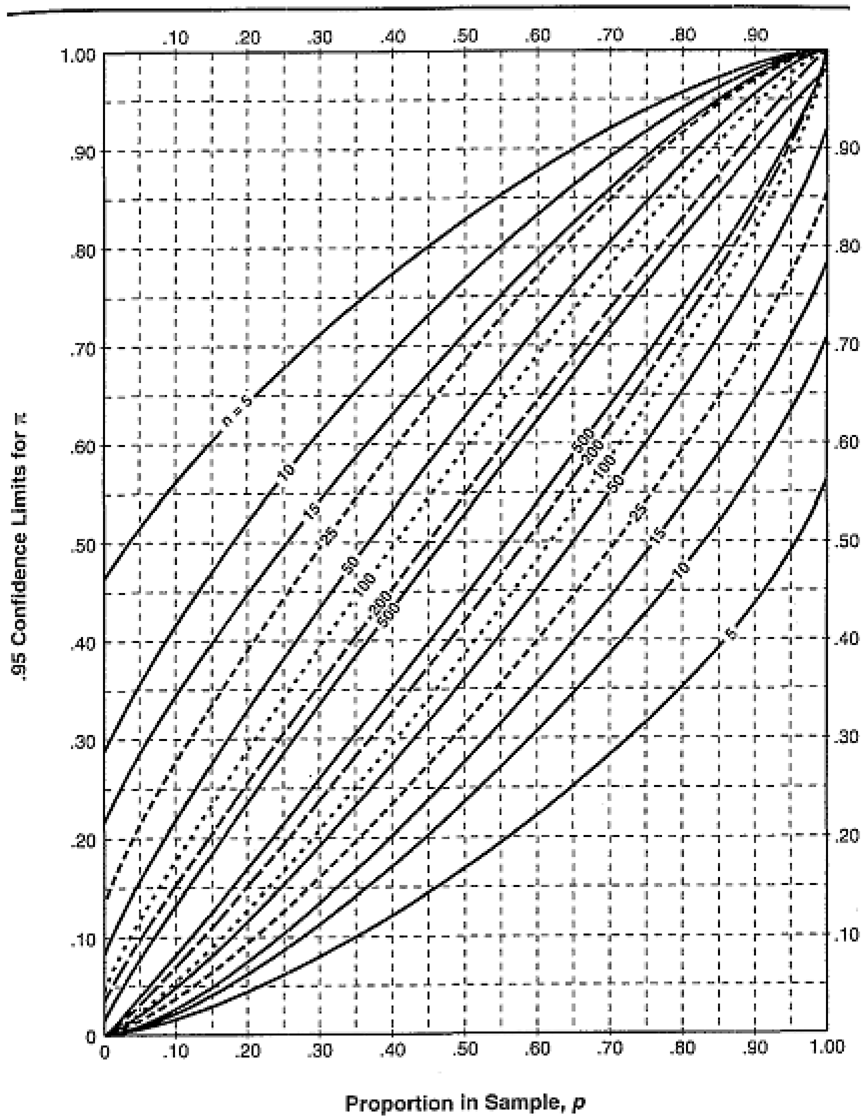


FIGURE 13.5 Graph for .95 Confidence Limits for the Parameter, π, from p and n.