

## Seminární úkol č. 1

Autoři: Klára Čapková (406803), Markéta Peschková (414906)

Zdroj dat: EU Kids Online Survey

### 1. T-test pro nezávislé výběry

$H_1$ : Dívky a chlapci se liší v míře problémů s vrstevníky.

$H_0$ : Dívky a chlapci se neliší v míře problémů s vrstevníky.

Předpoklady pro použití t-testu pro nezávislé výběry:

- Data vybraných skupin (tj. chlapců a dívek) jsou na sobě navzájem nezávislá, teda skóre účastníka z jedné skupiny není možné systematicky přiřadit ke skóru účastníka z druhé skupiny.
- Vzhledem k velikosti vzorku ( $N_1=9320$ ;  $N_2=9327$ ) není nutné zjišťovat, zda jsou proměnné normálně rozložené.
- Homoskedasticita sledovaných proměnných bude testována níže pomocí Levenova testu pro rovnost rozptylů.

Hypotézy byly testovány na 5% hladině významnosti, tj.  $\alpha=0,05$ .

Tabulka 1

#### Popisné statistiky

	Chlapci			Dívky		
	M	SD	N	M	SD	N
SDQ problémů s vrstevníky	1,36	0,33	9320	1,32	0,31	9327

Tabulka 2

*T-test pro nezávislé výběry*

	Levenův test pro rovnost rozptylů		T-test pro nezávislé výběry				
	F	Sig.	t	df	Sig. (2-tailed)	95 % CI pro rozdíl průměrů	
						Min	Max
SDQ problémů s vrstevníky	12,812	0,000	1,32	18606,03	0,000	0,02	0,05

Z výsledků Levenova testu pro rovnost rozptylů ( $F=12,812$ ,  $p<0,05$ ) vyplývá, že se skupiny ve svých rozptylech liší a proto tedy nenaplnují požadavek homoskedasticity. Z tohoto důvodů nadále pracujeme s hodnotami upravenými pro variantu, kde si rozptyly nejsou rovné.

Na základě výsledků zamítáme  $H_0$ . Byl prokázán statisticky významný rozdíl mezi mírou problémů s vrstevníky u chlapců ( $M=1,36$ ;  $SD=0,33$ ) a mírou problémů s vrstevníky zaznamenaných u dívek ( $M=1,32$ ;  $SD=0,31$ ), kde  $t(18645) = 7,81$ ,  $p = 0,000$ . Chlapci vykazují průměrně vyšší míru problémů s vrstevníky než dívky. Cohenovo  $d=0,12$ , jedná se tedy o malý efekt.

Pro rozdíl průměrů SDQ problémů s vrstevníky u chlapců a dívek byl zkonstruován 95% CI (0,02; 0,05).

## 2. Párový t-test

$H_1$ : Míra behaviorálních problémů jedince se liší od jeho zjištěné míry emocionálních problémů.

$H_0$ : Míra behaviorálních problémů jedince je rovna jeho zjištěné míře emocionálních problémů.

Předpoklady pro použití párového t-testu:

- Sledovaná data pochází z 2 měření jedné skupiny osob. Jednotlivé skóry z měření míry behaviorálních problémů tedy lze systematicky přiřadit ke skórum z měření míry emocionálních problémů, protože vždy odpovídají jedné osobě.
- Vzhledem k velikosti vzorku ( $N=18639$ ) není nutné zjišťovat, zda jsou proměnné normálně rozložené.

Hypotézy byly testovány na 5% hladině významnosti, tj.  $\alpha=0,05$ .

Tabulka 3

*Popisné statistiky*

SDQ behaviorálních problémů			SDQ emocionálních problémů		
M	SD	N	M	SD	N
1,4	0,30	18639	1,39	0,38	18639

Tabulka 4

*Korelace*

	N	r	Sig.
SDQ behaviorálních problémů & SDQ emocionálních problémů	18639	0,31	0,000

Tabulka 5

*Párový t-test*

	Párový t-test				
	t	df	Sig. (2-tailed)	95 % CI pro rozdíl průměrů	
				Min	Max
SDQ behaviorálních problémů & SDQ emocionálních problémů	3,04	18638	0,002	0,003	0,015

Výsledky testů ukazují na významný rozdíl mezi skórem dosaženým jedinci na škále behaviorálních problémů ( $M=1,4$ ;  $SD=0,3$ ) a skórem získaným na škále emocionálních problémů ( $M=1,39$ ;  $SD=0,38$ ), kde  $t(18638)=3,04$ ,  $p=0,002$ . Nulovou hypotézu tedy zamítáme.

Pro rozdíl průměrů SDQ behaviorálních problémů a SDQ emocionálních problémů byl sestaven 95% CI (0,003; 0,15).

### 3. Chí kvadrát test homogenity

$H_1$ : Existuje rozdíl mezi dívkami a chlapci v tom, zda si od rodičů přejí větší zájem o své online aktivity.

$H_0$ : Neexistuje rozdíl mezi dívkami a chlapci v tom, zda si od rodičů přejí větší zájem o své online aktivity.

Předpoklady pro použití chí kvadrát testu nezávislosti:

- Nanejvýš 20 % očekávaných četností je menších než 5
- Žádná očekávaná četnost není menší než 1
- $N > 40$

Hypotézy byly testovány na 5 % hladině významnosti, tj.  $\alpha=0,05$ .

Tabulka 6

#### Kontingenční tabulka

			Přání většího/menšího zájmu		Celkově
			rodičů		
			Větší	Menší	
Pohlaví dítěte	Chlapec	Počet	1156	1096	2252
		Očekávaný počet	1177,5	1074,5	2252,0
	Dívka	Počet	1152	1010	2162
		Očekávaný počet	1130,5	1031,5	2162,0
Celkově	Počet	2308	2106	4414	
	Očekávaný počet	2308,0	2106,0	4414,0	

Tabulka 7

#### Chí kvadrát test nezávislosti

$\chi^2$	df	Asymp. Sig. (2-sided)
1,684	1	0,194

Mezi proměnnými nebyl prokázán statisticky významný vztah.  $\chi^2 (1, 4414) = 1,684, p > .05$ .

Cramerovo  $V = 0,02$ . V platnosti ponecháváme nulovou hypotézu a předpokládáme, že se chlapci a dívky neliší v tom, zda si od svých rodičů přejí větší zájem o své online aktivity.

*Snad už je to po semináři s těmi tabulkami jasné. Po statistické stránce OK.*

*SJ*