

## Statistická analýza dat II. – PSY 252

Ondřej Sedlák (414830), Lenka Tarabíková (414635)

### Seminární úkol č. 2

Pro analýzu prostřednictvím mnohonásobné lineární regrese jsme použili data použitá při hodině Statistické analýzy dat II., dostupná ve studijních materiálech v datovém souboru Long1.sav. Výzkumníci se zabývali faktory ovlivňujícími míru deprese u mládeže. Výzkumný vzorek zahrnuje celkem 768 respondentů. Pro analýzu mnohonásobné lineární regrese jsme zvolili závislou proměnnou míru deprese ( $M = 2.01$ ,  $SD = 0.47$ ). Pro určení statisticky významných ( $p < .05$ ) nezávislých proměnných jsme použili metodu Enter a stanovili hypotézu, že na míru deprese působí negativní pocity (osamělost, strach, vina; únava, stres), životní spokojenost, self-esteem, zdravotní potíže psychosomatického rázu a nehody s rodiči. V analýze jsme vynechali celkem 74 respondentů, u nichž jsme neměli kompletní údaje týkající se závislé a nezávislých proměnných.

Tabulka 1

Základní popisné statistiky proměnných negativní pocity (osamělost, strach, vina), negativní pocity (únava, stres), životní spokojenost, self-esteem, zdravotní potíže psychosomatického rázu, nehody s rodiči

	M	SD
Negativní pocity (osamělost, strach, vina)	.01	1.01
Negativní pocity (únava, stres)	- .01	1.01
Životní spokojenost	2.91	.47
Self-esteem	3.07	.47
Zdravotní potíže psychosomatického rázu	25.47	4.81
Neshody s rodiči	3.13	.57

Zkontrolovali jsme, zda jsou splněny předpoklady použití mnohonásobné lineární regrese. Závislá i všechny nezávislé proměnné byly měřeny na intervalové úrovni. Vzájemné vztahy mezi závislou proměnnou a jednotlivými nezávislými proměnnými jsme vyhodnotili jako lineární. Na základě platnosti centrálního limitního teorému předpokládáme normální rozložení proměnných. V datech se vyskytuje několik extrémních hodnot, ale žádní outlieři. Rovněž jsme prokázali absenci vysoké kolinearity, jelikož korelace mezi nezávislými proměnnými nepřesahuje  $r > .62$ ,  $VIF < 2$  a tolerance  $> .56$ . Nezávislost reziduí jsme splnili díky hodnotě 1.98 v testu Durbin-Watson.

Následně jsme provedli postupnou mnohonásobnou regresi s cílem zjistit nejlepší model proměnných pro predikování míry deprese. Negativní pocity osamělosti, strachu a viny byly vybrány do analýzy jako první, jelikož tato proměnná má nejvyšší korelaci s mírou deprese ( $r = -.61$ ) a vysvětluje 38 % rozptylu míry deprese. Ostatní uvedené nezávislé proměnné jsme do analýzy přidali na základě vysokých hodnot parciální korelace a standardizovaných  $b$  hodnot.

Tabulka 2

Regresní koeficienty proměnných negativní pocity (osamělost, strach, vina), negativní pocity (únava, stres), životní spokojenost, self-esteem, zdravotní potíže psychosomatického rázu, neshody s rodiči

	B	$\beta$	t
Constant	3.06		21.29
Negativní pocity (osamělost, strach, vina)	-.15	-.32	-9.77
Negativní pocity (únava, stres)	.05	.09	3.57
Životní spokojenost	-.15	-.15	-4.31
Self-esteem	-.19	-.19	-5.37
Zdravotní potíže psychosomatického rázu	.1	.14	4.79
Neshody s rodiči	-.13	-.16	-5.35

Sada prediktorů, kterou jsme vytvořili, vysvětluje celkem 53 % rozptylu míry deprese ( $R = .73$ ,  $R^2 = .53$ , adjusted  $R^2 = .53$ ). Náš model regrese predikuje míru deprese významně

dobře ( $F = 128.49$ ,  $p < .001$ ). Naši hypotézu, že míru deprese ovlivňují negativní pocity (osamělost, strach, vina; únava, stres), životní spokojenost, self-esteem, zdravotní potíže psychosomatického rázu a neshody s rodiči, jsme přijali.