**Statistická analýza dat II (PSY252) – seminární skupina A**

Úkol 3

**Zpracovaly:** Denisa Lakomá, 414925

 Kristýna Cejthamrová, 362244

K analýze jsme využily datový soubor Long1.sav. Celkový vzorek byl 768 respondentů. Jako závislou proměnnou jsme si vybraly pohlaví jedince. Ve vzorku bylo 317 mužů (41,3%) a 449 žen (58,6%), 2 lidé pohlaví neuvedli.

 Z proměnných zahrnutých do výzkumu jsme vybraly 5 prediktorů na základě výsledků logistické regrese. Použitými prediktory tedy jsou *deprese, self – efficacy, důvěrnost s přáteli, vřelost matky a vřelost otce*.

U deprese jsme předpokládaly, že by mohla být dobrým prediktorem pohlaví, protože v období adolescence se dá předpokládat vyšší míra depresivity u dívek (více řeší svůj vzhled, problémy spojené s dospíváním). Self – efficacy jsme naopak předpokládaly vyšší u chlapců. Důvěrnost s přáteli by podle nás měla být vyšší u dívek, protože dívky častěji řeší svoje osobní problémy (dospívání, kluky, rodiče) se svými kamarádkami než chlapci. Vřelost matky a vřelost otce pak předpokládáme vyšší u dívek, protože u chlapců v tomto věku není přílišná vřelost ze strany rodičů žádaná a rodiče si to většinou uvědomují.

Regrese byla provedena za použití SPSS verze 21.

Tabulka 1. *Popisné statistiky prediktorů – ženské pohlaví*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|   | Minimum | Maximum | Průměr | Směrodatná odchylka |
| vřelost matky | 1,18 | 4,00 | 3,30 | 0,46 |
| vřelost otce | 1,00 | 4,00 | 3,05 | 0,64 |
| důvěrnost s přáteli | 1,33 | 4,00 | 2,82 | 0,48 |
| self-efficacy | 1,33 | 4,00 | 2,68 | 0,54 |
| deprese | 1,05 | 3,40 | 2,06 | 0,49 |

Tabulka 2. *Popisné statistiky prediktorů – mužské pohlaví*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|   | Minimum | Maximum | Průměr | Směrodatná odchylka |
| vřelost matky | 1,64 | 4,00 | 3,13 | 0,45 |
| vřelost otce | 1,00 | 4,00 | 3,12 | 0,48 |
| důvěrnost s přáteli | 1,17 | 4,00 | 2,51 | 0,49 |
| self-efficacy | 1,33 | 4,00 | 2,83 | 0,48 |
| deprese | 1,06 | 3,05 | 1,94 | 0,41 |

**Předpoklady:**

Ze scatterplotů lze usuzovat na linearitu vztahů. Předpoklad kolinearity je také splněn, korelace mezi jednotlivými prediktory se pohybovala mezi 0,03 a 0,49, tolerance mezi 0,1 a 1 a hodnoty VIF se pohybovaly kolem 1. Předpokládáme nezávislost reziduí, výsledek Durbin – Watsonova testu je 1,8. U spojitých proměnných není pro každou hodnotu alespoň jeden respondent, nikde však nebyla nalezena větší mezera v datech, ve výsledcích se nedodržení tohoto předpokladu také nějak výrazně neprojevilo. Nezávislou proměnnou nelze perfektně predikovat jedním nebo kombinací více prediktorů. Rozptyl není větší, než bylo očekáváno z modelu, protože rezidua jsou nezávislá a směrodatné chyby nejsou příliš malé. Je možné říci, že data splňují podmínky pro mnohonásobnou regresi.

**Výsledky:**

Test modelu se všemi prediktory proti modelu bez prediktorů vyšel signifikantně, χ2(5, N= 768) = 125,98; p < 0,001. Model byl schopen správně klasifikovat 50,9% mužů a 82,7% žen, celkově tedy správně zařadil 70,1% respondentů.

Tabulka 3 poskytuje informace o regresních koeficientech B, jejich směrodatných chybách, Exp (B) a jejich intervalech spolehlivosti pro všechny prediktory. Z výsledků vyplývá, že všechny prediktory signifikantně přispívají k lepšímu modelu predikce.

Tabulka 3. *Regresní koeficienty*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|   | koeficient B | Směrodatná chyba | Exp (B) | C.I. pro Exp (B) na 95% hladině | Signifikance |
| Deprese | 0,526 | 0,208 | 1,692 | 1,125 | 2,545 | 0,012 |
| Self-efficacy | -0,702 | 0,177 | 0,496 | 0,350 | 0,701 | 0,000 |
| Důvěrnost s přáteli | 1,361 | 0,190 | 3,899 | 2,687 | 5,657 | 0,000 |
| Vřelost matky | 1,574 | 0,245 | 4,825 | 2,987 | 7,794 | 0,000 |
| Vřelost otce | -0,701 | 0,202 | 0,496 | 0,334 | 0,737 | 0,001 |
| Konstanta | -5,254 | 1,122 | 0,005 |   |   | 0,000 |

Hodnota -2LL byla po zahrnutí prediktorů 790,53; Nagelkerke R2 = 0,23. Nejsilnějším prediktorem je na základě námi zvoleného modelu vřelost matky s tím, že s každým jeho nárůstem o jednu jednotku se šance na to, že respondent je žena, zvýší 4,83 krát. Důvěrnost s přáteli tuto šanci zvyšuje 3,90 krát a deprese 1,69 krát. Oproti tomu nárůst self-efficacy mírně zvyšuje šanci, že respondent je chlapec, což je v souladu s naším předpokladem. Naopak oproti původnímu očekávání se ukazuje, že stejný vliv má i vyšší vřelost otce. Šance, na to, že respondent je mužského pohlaví, se s nárůstem o jednu jednotku v tomto případě zvyšuje 0,50 krát.

*Dobrý pokus.*