

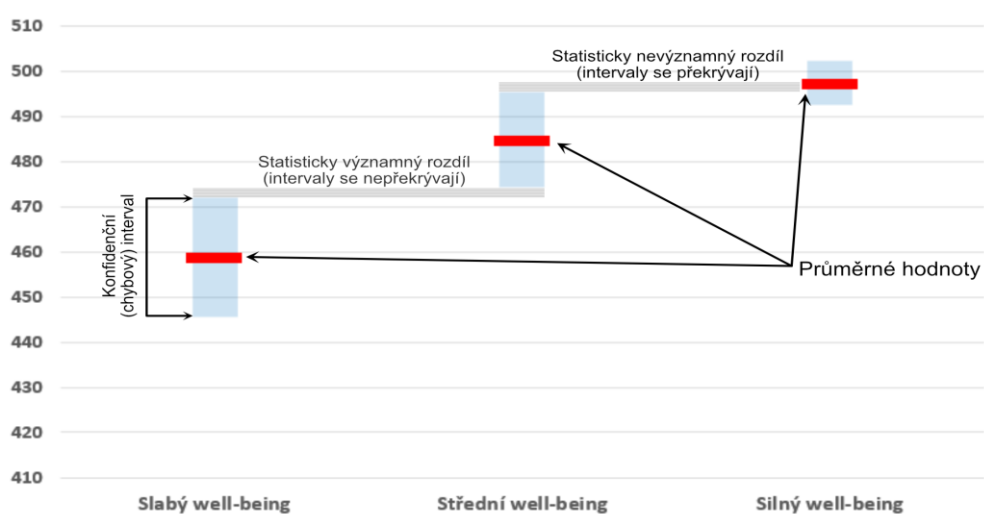
Pro výpočet statistické významnosti v případě dvoustupňového výběru používáme IDB Analyzer, balíček „PV Module“ pro program Stata a balíček „intsvy“ a „WeMix“ pro R. Statistická významnost je vizuálně znázorněna v grafech.

INTERVAL SPOLEHLIVOSTI

Interval spolehlivosti či konfidenční interval je velmi úzce spojen s testováním hypotéz a statistickou významností. Protože pracujeme s výběrovým vzorkem, a nikoliv s celou populací, neměli bychom výsledná zjištění prezentovat ve formě bodových odhadů, nýbrž jako intervalové odhady (rozsah hodnot).

Jako příklad můžeme uvést odhad v grafu níže, který zobrazuje průměrné skóre žáků ve třech skupinách dle míry jejich pocitu well-being. Každá průměrná hodnota (červený pruh) je doprovázena intervalem spolehlivosti, ve kterém se skutečná hodnota populačního průměru může pohybovat na zvolené hladině spolehlivosti. Volba 95% hladiny spolehlivosti znamená, že odhadovaný populační parametr (zde průměr) pokryje 95 ze 100 vytvořených intervalů spolehlivosti. Spodní hodnota intervalu se označuje jako dolní interval, horní naopak jako horní interval.

Příklad interpretace běžného deskriptivního grafu z této zprávy



Poznámka: Světle modrá část představuje interval spolehlivosti, červená linka určuje střední hodnotu mezi horním a dolním intervalem, skutečný výsledek se nachází v rozmezí těchto dvou hraničních hodnot.

HIERARCHICKÝ REGRESNÍ MODEL

Pokud mají data hierarchickou strukturu (např. žáci v rámci škol nebo učitelé v rámci škol), jednoduchá lineární regrese není pro jejich analýzu vhodná. Z tohoto důvodu se v edukačních vědách používají tzv. hierarchické regresní modely. Základní hierarchické modely počítají s náhodnou konstantou, která se mění v závislosti na tzv. klastru. Klastr v našem případě představuje škola. První úroveň je v hierarchickém modelu učitel, v druhé úrovni pak proměnné na úrovni školy – ředitelský dotazník či učitelské proměnné po agregaci.

Vhodným statistickým softwarem pro používání hierarchických regresních modelů jsou Stata, R a MPlus. Ve studii používáme všechny tyto programy a kombinujeme jejich hlavní přednosti. V případě programu Stata doporučujeme nainstalovat několik modulů. Hlavními balíčky jsou PV MODULE (pro výpočty s plausibilními hodnotami a vážení), MLT pro výpočet R2 pro první a druhou úroveň, ICCVAR pro výpočet vnitroskupinové korelace. V případě R používáme balíček lme4 Andrew Gelmana. Obecně potom platí, že výsledné hierarchické modely jsou konzervativní, protože klastrování a vážení na dvou úrovních zvyšují standardní chyby pro výpočet statistické významnosti regresních koeficientů.

ICC

Vnitrotřídní koeficient korelace (intra-class correlation coefficient, ICC) tvoří nedílnou část hierarchického modelování. Pomocí jeho výpočtu jsme schopni v první fázi výzkumu rozhodnout, zda je pro analýzu našich dat žádoucí použít hierarchické modely. V případě předkládané zprávy byl ICC vypočítán pro proměnnou testového skóre. První úroveň představoval žák, druhou úroveň jednotlivé školy. Hodnota ICC se pohybuje v rozmezí 0–1. Výsledná hodnota ICC značí rozptyl proměnné na druhé úrovni.