UNIVERZITA KONŠTANTÍNA FILOZOFA V NITRE

PEDAGOGICKÁ FAKULTA

Katedra pedagogickej a školskej psychológie

ŠTATISTIKA PRE ZAČIATOČNÍKOV

Základy štatistických analýz pre študentov učiteľstva

Eva Ballová Mikušková

NITRA 2021

Názov: Štatistika pre začiatočníkov.
Základy štatistických analýz pre študentov učiteľstva

Autor: PhDr. Eva Ballová Mikušková, PhD.

Recenzenti: Mgr. Michal Kohút, PhD.

Prof. Peter Halama, PhD.

Schválené edičnou radou Pedagogickej fakulty Univerzity Konštantína Filozofa v Nitre.

ISBN 978-80-558-1823-8 (printová verzia)

ISBN 978-80-558-1824-5 (online verzia)

Obsah

Úvod

Štatistika nie je medzi laikmi ani medzi mnohými študentmi veľmi populárna. Prvá predstava pri pomyslení na štatistiku je často „čísla, čísla, tabuľky, grafy, čísla". Popularitu štatistiky nezvyšuje ani skutočnosť, že mnohé učebnice a knihy o štatistike sú naozaj plné náročných pojmov, definícií—a na zdesenie mnohých—aj vzorcov, ktoré sa neraz tiahnu cez niekoľko riadkov. Študenti si volia humanitné predmety z rôznych dôvodov, a keď sa ich pýtam na motiváciu, okrem záujmu o prácu s ľuďmi, pomoc iným, či iné poslanie, ktoré v sebe pociťujú, zohráva istú úlohu aj predstava (trúfam si povedať, že aj zbožné prianie), že pri štúdiu sa „vyhnú" matematike a príbuzným predmetom. Aké hlboké je potom ich sklamanie, keď zistia, že k ukončeniu štúdia potrebujú absolvovať predmet štatistika, a dokonca, že poväčšine potrebujú štatistiku aj reálne aplikovať pri vypracovávaní záverečnej práce. Každý rok vidím rovnaký výraz na tvárach tých niekoľkých študentov, ktorí sa neostýchajú navonok prejaviť svoje zdesenie. A to aj napriek tomu, že mnohí z nich majú vynikajúce predpoklady k vedeckej práci, k vedeckému uvažovaniu, a k tomu, aby realizovali kvalitné výskumy vrátane štatistických analýz.

Našťastie, v súčasnosti je možné učiť (sa) štatistiku aj inak ako pomocou memorovania vzorcov, úmorného počítania a hrabania sa v tabuľkách. Nielenže sa každým rokom objavujú rôzne štatistické programy, pomocou ktorých je možné relatívne jednoducho aplikovať štatistické analýzy, súbežne vznikajú aj publikácie, ktoré približujú štatistiku priateľským spôsobom a pomáhajú lepšie porozumieť tomu, ako možno výsledky analýz interpretovať. Takúto snahu má aj učebnica, ktorú držíte v rukách.

Ak zvažujete, či sa vrhnúť do vôd štatistiky (aj za pomoci tejto učebnice), musím vás upozorniť, že k uspokojivému porozumeniu základov štatistiky je nevyhnutné disponovať aspoň základmi metodológie výskumu. Metodológia a štatistika totiž k sebe neoddeliteľne patria, idú ruka v ruke na ceste k poznaniu. Ak ste sa s metodológiou výskumu doteraz nestretli, odporúčam vám najskôr naštudovať si aspoň jej základy, v knihách ako sú napríklad *Kvantitatívny výskum v pedagogických vedách* od Tomšika (2017), či *Úvod do metodológie psychologického výskumu* od Ferjenčíka (2000) a *Metodológia projektovania psychologického výskumu* od Ritomského (2016). Učebnica, ktorú práve držíte je totiž zameraná na štatistiku, respektíve na základy štatistiky v konkrétnom štatistickom programe. A hoci v úvode prináša stručný prehľad metodologických pojmov, ide skôr o pripomenutie a osvieženie vedomostí, než o podklad k štúdiu metodológie.

Postupne v jednotlivých kapitolách predstavím krok po kroku základné štatistické postupy: od ich teoretického vysvetlenia až po praktické využitie. Pre tých, ktorí chcú analyzovať dáta pomocou štatistických analýz, sú v súčasnosti k dispozícii viaceré programy a prostredia. Od Excelu, cez SPSS, JASP či JAMOVI až po programovanie v R-ku. V tejto učebnici sa zameriame na program JASP, ktorý je voľne dostupný, užívateľsky priateľský a relatívne intuitívny. Teda ako stvorený pre študentov a všetkých, ktorí so štatistikou začínajú. Ukážeme si, ako v JASP-e vypočítať základné štatistické ukazovatele (deskriptívnu štatistiku), ako overiť či sú rozdiely medzi skupina participantov alebo medzi rôznymi meraniami významné, ale aj ako preskúmať vzťahy medzi skúmanými premennými.

Cieľom knihy je priblížiť študentom, kolegom aj nadšencom z radov laikov štatistiku tak, aby sa ich obavy a strach postupne vytratili, aby v sebe našli nadšenie (alebo aspoň pochopenie) pre možnosti, ktoré im štatistika ponúka v ich vedeckej práci a aj v reálnom živote, a hlavne, aby boli schopní samostatne spracovať a analyzovať dáta vo svojom výskume a porozumieť výsledkom, ktoré získajú.

Táto publikácia by nevznikla bez podpory zo strany vedenia Katedry pedagogickej a školskej psychológie PF UKF. Nesporne veľkou pomocou mi boli študenti, ktorých otázky a postrehy za posledných pár rokov mi boli cenným usmernením v tom, ako čo najzrozumiteľnejšie o štatistike rozprávať a písať. A v neposlednej rade ďakujem rodine a blízkym za ich porozumenie, trpezlivosť a podporu.

V Nitre, 13.10.2021 Eva Ballová Mikušková

1 ŠTATISTIKA

Štatistika je veda zaoberajúca sa zberom, analýzou, interpretáciou a prezentáciou dát. Štatistika sa zaoberá hromadnými javmi, to znamená, že súbory údajov zhromažďuje, utrieďuje, analyzuje, charakterizuje a porovnáva, a na základe týchto analýz sa snaží vyvodzovať všeobecne platné závery (Aron et al., 2014; Foster et al., 2018; Rimarčík, 2007). Štatistika je v podstate spôsob, akým hľadáme pravdu, pomáha nám odhadnúť pravdepodobnosť, s akou sú naše tušenia či predpoklady pravdivé a rovnako nám pomáha odhadnúť pravdepodobnosť, s akou nejaký jav nastane v budúcnosti.

Vo všeobecnosti môžeme štatistiku charakterizovať ako odvetvie matematiky, ktoré sa zameriava na organizáciu, analýzu a interpretáciu skupiny dát, alebo čísiel (slovo štatistika má ešte iný význam, štatistika je aj číslo, resp. štatistický údaj, napríklad priemer skupiny je štatistika, pričom táto a iné štatistiky podnietili vznik odboru štatistika). Používať štatistiku a porozumieť jej je potrebné, pretože predstavuje spôsob, akým vo vede komunikujeme. Slúži na to, aby sme prepájali nápady a predpoklady s využiteľnými závermi; bez štatistiky by nebolo možné interpretovať a porozumieť veľkému množstvu informácií, ktoré sú obsiahnuté v dátach (Foster et al., 2018). Predstavte si tabuľku s 200 stĺpcami a 600 riadkami plnú čísiel – v takejto tabuľke je množstvo dát, no bez štatistického spracovania dáta iba ťažko prinesú informácie.

Aj v bežnom živote máme možnosť stretávať sa so štatistickými údajmi. Napríklad, keď čítame alebo počujeme, že:

- 9 z10 zubárov odporúča,
- takmer 60% ochorení súvisí s fajčením,
- vakcína je účinná v 95% prípadov,
- ženy zarobia 80 eur na každých 100 eur, ktoré zarobí muž na tej istej pozícii,
- je 80% šanca, že v miestnosti, v ktorej je 30 ľudí, sa nájdu najmenej dvaja, ktorí majú narodeniny v ten istý deň.

Štatistické údaje sú relatívne často prezentované, aby dodali dôveryhodnosť rôznym argumentom, predpokladom, tvrdeniam, či odporúčaniam. Štatistika nám pomáha porozumieť záverom, s ktorými sa stretneme nielen vo vedeckých výstupoch, ale aj v bežnom živote.

Predmetom skúmania štatistiky sú **štatistické jednotky** (štatistickými jednotkami v spoločenských a humanitných vedách sú najčastejšie ľudia, prípadne aj školy, triedy, organizácie a pod.), ktoré tvoria **štatistický súbor**.

V rámci výskumného procesu rozlišujeme tri typy štatistickej praxe: zber dát, exploráciu dát a štatistické usudzovanie. ***Zber dát*** je predmetom skúmania v rámci metodológie. ***Exploračná analýza dát*** predstavuje prvý stupeň štatistického spracovania dát, označujeme ju aj **deskriptívna** alebo opisná štatistika a predstavuje spôsob, ako zobraziť základné charakteristiky skúmaného súboru a sledovaných premenných (Kapitola 4 *Deskriptívna štatistika*), týka sa opisu jednotiek vo vzorke teda opisu konkrétnej skupiny. ***Štatistické usudzovanie*** už ide nad rámec deskriptívnej štatistiky, ide o testovanie hypotéz, zovšeobecňovanie zistení zo vzorky na celkovú populáciu, skúmanie spoľahlivosti záverov prostredníctvom pravdepodobnosti. Tento typ štatistiky nazývame aj **inferenčná** (induktívna) **štatistika** (Kapitola 6 *Inferenčná štatistika*). Inferenčná štatistika dáva odpovede o na otázky, ako napríklad premenné súvisia s inými premennými, za akých podmienok sa menia hodnoty premenných, či do akej miery sa líšia dve skupiny navzájom.

Navyše, v rámci štatistiky rozlišujeme tri úrovne štatistiky, podľa toho, koľko premenných je zahrnutých do analýzy:

1. ***univariačná*** štatistika: v rámci univariačnej štatistiky využívame štatistické metódy na analýzu jednej premennej,
2. ***bivariačná*** štatistika: v rámci bivariačnej štatistiky využívame štatistické metódy na analýzu súvislostí dvoch premenných,
3. ***multivariačná*** štatistika: v rámci multivariačnej štatistiky využívame štatistické metódy na analýzu súvislostí troch a viacerých premenných.

1.1 Základné pojmy

Ešte predtým, ako sa naplno ponoríme do štatistiky, v krátkosti si pripomenieme niekoľko základných pojmov z metodológie výskumu, ktoré sú kľúčové pre porozumenie štatistike a štatistických analýz, respektíve pojmy, s ktorými sa môžeme pri práci so štatistikou stretnúť. (Pojmy sú usporiadané *abecedne* a definície sa opierajú o Coolican, 2019; Dancey & Reidy, 2017; Foster et al., 2018; Hanna & Dempster, 2012; Kalina et al., 2010; Lichner, 2020; Miles & Banyard, 2007; Pathak, 2011; Ritomský, 2015; Tomšik, 2017.)

Hypotéza. Presné vyjadrenie predpokladaného vzťahu medzi premennými (pričom vzťah je tu chápaný v širšom zmysle slova, t.z. ak aj skúmame rozdiely medzi skupinami, stále skúmame vzťah nezávislej premennej, prostredníctvom ktorej delíme vzorku na skupiny, a závislej premennej).

Experimentálny dizajn. Experiment predstavuje dizajn, pomocou ktorého môžeme overiť kauzálne vzťahy medzi premennými, a to tak, že manipulujeme nezávislou premennou (pričom všetky ostatné premenné sú kontrolované) a sledujeme vplyv tejto manipulácie na závislú premennú (resp. závislé premenné). Jedným z hlavných znakov experimentálneho dizajnu je náhodné pridelenie participantov do podmienok nezávislej premennej. Rozlišujeme dva základné experimentálne dizajny: a) *vnútrosubjektový* dizajn, kedy porovnávame skóre jednotlivcov v jednej podmienke nezávislej premennej s ich skóre v inej podmienke nezávislej premennej, a b) *medzisubjektový* dizajn, kedy porovnávame skóre jednej skupiny ľudí v jednej podmienke nezávislej premennej so skóre inej skupiny ľudí v inej podmienke nezávislej premennej.

Kvázi-experimentálny dizajn. Kvázi-experiment je „takmer“ experiment, ktorý nespĺňa niektorú z hlavných podmienok experimentálneho dizajnu. Ide o dizajn podobný experimentálnemu, participanti však napríklad nie sú do podmienok nezávislej premennej zaraďovaní náhodne (napr. nemôžeme manipulovať premennou pohlavie), alebo nezávislá premenná nie je, alebo nemôže byť, kontrolovaná (napr. v prípade prirodzeného experimentu).

Meranie. Vlastnosť, ktorú chceme skúmať pomocou merania identifikujeme v podobe hodnoty (v prípade kvantitatívnych znakov hodnota predstavuje zväčša numerický údaj).

Populácia. Populácia je základný súbor štatistických jednotiek, pozostáva zo všetkých možných osôb (alebo predmetov), ktoré majú konkrétnu charakteristiku. Študovať celú populáciu je väčšinou nemožné, preto pozorujeme vybrané štatistické jednotky, ktoré tvoria *vzorku* (tento proces nazývame *výber*).

Premenná. Premenné sú pozorovateľné alebo hypotetické štatistické znaky, atribúty, ktoré sa môžu meniť a ktorých zmeny je možné merať. Ide o číselnú reprezentáciu štatistického znaku. Keď premennou manipulujeme, nazývame ju *nezávislá* premenná (alebo *prediktor*). V experimente sa pokúšame odhaliť vplyv zmien v nezávislej premennej na zmeny iných premenných. Tieto premenné nazývame aj *závislé*, inými slovami, závislá premenná je premenná, na ktorú nezávislá premenná pôsobí, jej hodnota závisí od hodnoty nezávislej premennej.

Premenné môžu dosahovať rôzne hodnoty z určitého rozsahu. Premenné môžeme merať alebo kategorizovať na rôznych úrovniach:

Nominálna úroveň: kategoriálna úroveň merania, pri ktorej hodnoty premennej predstavujú kategórie a štatistické jednotky sú roztrieďované do týchto kategórií podľa stanoveného kritéria. Všetky kategórie sú si rovnocenné, nie je možné usporiadať ich do poradia a nie je možné realizovať s nimi žiadne matematické operácie. V prípade, že má premenná iba dve možné kategórie, hovoríme o *binárnej* premennej.

Poradová (ordinálna) úroveň: úroveň merania, pri ktorej možno štatistické jednotky usporiadať do poradia vzhľadom na veľkosť ich meranej charakteristiky. Je možné určiť iba, ktoré jednotky sú väčšie či menšie, nie je možné určiť presné rozdiely.

Intervalová (kardinálna) úroveň: úroveň merania, pri ktorej má každá štatistická jednotka hodnotu v rámci stupnice, vďaka čomu vieme určiť presné rozdiely medzi štatistickými jednotkami a vykonávať rozmanité matematické operácie. Ak stupnica má začiatok v bode nula a hodnoty nemôžu klesnúť pod tento bod, hovoríme o *pomerovej* premennej.

Reliabilita. Reliabilita (konzistencia, presnosť) je miera, do akej môžu byť merania opakované pri zachovaní rovnakých (čo najpodobnejších) výsledkov. Rozlišujeme reliabilitu časovú (stabilita nástroja v čase) a internú (všetky položky merajú rovnaký konštrukt).

Validita. Validita (spoľahlivosť) meracieho nástroja je miera, do akej nástroj meria to, na meranie čoho bol skonštruovaný. Validita experimentu, výskumného dizajnu, je miera, do akej sa môžeme spoľahnúť, že výskumný efekt je skutočný.

Výskum. Rozlišujeme niekoľko základných typov výskumu. Podľa toho, či máme dostatočnú poznatkovú bázu, o ktorú by sme opierali hypotézy, rozlišujeme:

Exploračný výskum. Pomocou explorativneho výskumu skúmame nejasné javy, snažíme sa o porozumenie konceptom, resp. o objasnenie povahy riešeného problému. V rámci exploračného výskumu kladieme výskumné otázky. Výsledkom explorativneho výskumu býva formulácia hypotéz.

***Verifikačný výskum*.** Verifikačný, alebo aj potvrdzovací či konfirmačný výskum slúži na overovanie stanovených hypotéz (stanovenie hypotéz predpokladá v danej oblasti existenciu primeranej poznatkovej bázy).

Podľa dizajnu výskumu, výskumného plánu ďalej rozlišujeme:

Deskriptívny výskum. Cieľom výskumu je zorientovanie sa v problematike, popísanie stavu vecí, javov.

Korelačný výskum. Korelačný výskum slúži na overenie vzťahov medzi dvoma alebo viacerými premennými.

Komparačný výskum. Komparačný, alebo porovnávací výskum sa zameriava na skúmanie prípadných rozdielov medzi výberovými súbormi alebo medzi premennými.

Kauzálny výskum. Cieľom kauzálneho výskumu je hľadanie príčinných vzťahov, pri realizácii kauzálneho výskumu je potrebné použiť experimentálnu metódu.

Longitudinálny výskum. V rámci longitudinálneho výskumu sa sledujú zmeny v premenných v jednom súbore v priebehu času.

Vzorka. Výberový súbor, alebo vzorka predstavuje typickú časť populácie, ktorá sa vyberá na základe určitých stanovených charakteristík.

Nezávislé vzorky. Nezávislé vzorky sú také vzorky, v ktorých sú rôzne štatistické jednotky (najčastejšie sa jedná o rôznych ľudí). To znamená, že v nezávislých vzorkách sú také jednotky, ktoré môžu byť iba v jednej z nich a nie v inej. Napríklad máme nezávislé vzorky muži a ženy. Jednotka (osoba), ktorá sa nachádza v jednej skupine, sa nemôže zároveň nachádzať aj v druhej skupine.

Závislé vzorky. Závislé vzorky sú také vzorky, v ktorých sa nachádzajú rovnaké štatistické jednotky (rovnakí ľudia), no premenné, ktoré u nich sledujeme sú zachytené za rôznych podmienok (navodenie pozitívnej nálady – navodenie negatívnej nálady: u tej istej skupiny) alebo v rôznych časových obdobiach (tá istá skupina pred kurzom – po kurze).

1.2 Ako počítať štatistiku (štatistické programy)

Dáta z výskumu možno spracovávať viacerými spôsobmi. Časy, kedy sa dáta analyzovali a spracovávali „ručne“, teda formou pero – papier, sú dávno preč. V súčasnosti máme mnoho možností, ako s dátami pracovať, od Excelu, cez rôzne voľne dostupné štatistické programy, až po licencované štatistické programy.

1.2.1 Excel

Pre študentov bol donedávna najdostupnejší Excel. V rámci Excelu sa základné štatistické analýzy nachádzajú v záložke *Údaje* (*Data*) v nástroji *Data Analysis*, ktorý je však doplnkový a je potrebné si ho manuálne doinštalovať.

Výhodou štatistického spracovania dát v Exceli je okamžitá editácia a formátovanie výstupov – tabuliek, ako aj jednoduché vytváranie grafov. Napriek tomu Excel nie je vhodný pre všetkých. Štatistické analýzy v Exceli vyžadujú od užívateľa, aby mal minimálne základné znalosti a skúsenosti s prácou v tomto tabuľkovom prostredí. To predpokladá aj porozumenie logických krokov pri zadávaní funkcií na výpočet jednotlivých analýz. Bez týchto znalostí je Excel ako štatistický program skôr trápením ako pomocou.

1.2.2 SPSS

Veľmi často v odborných kruhoch používaný a obľúbený, hoci najmenej dostupný kvôli svojej cene, je licencovaný program SPSS. Ide o modulový program, ktorý môžu používať tak začiatočníci ako aj skúsení výskumníci. Program sa vyznačuje širokým spektrom štatistických metód, v rámci ktorých je možné nastavovať jednotlivé požiadavky a tým prispôsobiť analýzy potrebám výskumu. Je relatívne užívateľsky priateľský a intuitívny.

1.2.3 Voľne dostupné programy

Vzhľadom na relatívne vyššiu užívateľskú náročnosť Excelu a finančnú náročnosť SPSS, sa v posledných rokoch objavili iniciatívy viacerých nadšencov, a ich výsledkom je niekoľko voľne dostupných štatistických programov, ktoré stačí jednoducho stiahnuť a nainštalovať. Výhodou týchto programov, alebo aplikácií, je, že sa neustále aktualizujú a dopĺňajú, často na podnet samotných výskumníkov a študentov. Medzi tieto programy patria napríklad R, JASP, JAMOVI, či PSPP.

Jedným z najpopulárnejších voľne dostupných softwarových nástrojov je *R*[1]. *R* označuje programovací jazyk aj prostredie, v ktorom možno vykonávať štatistické výpočty a operácie. Svojou otvorenosťou a množstvom balíkov pre rôzne typy analýz je veľmi populárny, niekoho však môže odrádzať „programovanie“ potrebné pre prípravu analýz. Voľne dostupným je aj program *PSPP*[2], ktorý vznikol ako alternatíva k SPSS. PSPP je otvorený a voľne dostupný, nemá však takú obsiahlu databázu štatistických analýz ako SPSS. Ďalšími obľúbenými programami sú *Jamovi*[3] a *JASP*[4]. Oba programy sú užívateľsky veľmi priateľské a ich používanie je intuitívne. Vyznačujú sa jednoduchou a prehľadnou grafikou, výstupy sú kopírovateľné a okamžite použiteľné. Navyše, analýzy sú interaktívne – ako meníte parametre, vidíte, ako sa menia výsledky v tabuľkách. V tejto učebnici budú jednotlivé štatistické analýzy prezentované a ilustrované pomocou programu JASP.

[1] https://www.r-project.org/

[2] https://www.gnu.org/software/pspp/

[3] https://www.jamovi.org/

[4] https://jasp-stats.org/

2 Organizácia dát

Keď chceme skúmať rôzne fenomény a javy, využívame rôzne spôsoby získavania informácií, pričom tieto informácie nazývame údaje, dáta. Od participantov pri zbere dát získavame nespracované údaje, ktoré sú v podobe komplexných tabuliek (často s desiatkami riadkov a stĺpcov). V tejto kapitole si ukážeme, ako sa pracuje s dátami v rámci ich prípravy a organizácie. Niektoré kroky vám možno budú známe, najmä ak máte skúsenosti s prácou v Exceli. Dokonca aj niektoré pojmy a výpočty nebudú pre vás novinkou, pretože sa ich učíme už na základných a stredných školách (napríklad aritmetický priemer).

2.1 Príprava a kontrola dát

Možností, ako zozbierané údaje do tabuľky zapísať je niekoľko, všetko závisí od spôsobu, akým dáta zbierame. Ak zbierame dáta spôsobom pero – papier, v prvom kroku údaje v záznamových hárkoch vizuálne skontrolujeme, či sú vyplnené všetky nevyhnutné údaje, či odpovede nie sú systematické (napr. participant zaškrtával iba jednu hodnotu vo všetkých otázkach), či nie je vynechaná nejaká časť položiek (napr. preskočená strana). V prípade, že participanti vypĺňali naše metódy na viacerých hárkoch, alebo vo viacerých krokoch (počas viacerých dní), je potrebné zabezpečiť, aby boli údaje od jedného participanta vzájomne priraditeľné.

Pre štatistické spracovanie získaných údajov je potrebné ich usporiadať do „tabuľkovej" podoby, to znamená, že údaje musíme aj prepísať do tabuľky. Najčastejšie preto v prípravnej fáze pracujeme v Exceli – tu si môžeme vytvoriť prvú tabuľku s dátami. Prvé odporúčanie znie, že údaje každého subjektu, štatistickej jednotky (participanta, prípadne skupiny, organizácie) sú v jednom riadku a každá sledovaná položka/údaj je v jednom stĺpci (Obrázok 1).

Druhé odporúčanie je, aby mal každý participant priradené svoje jedinečné identifikačné číslo (môže to byť kód, ktorý si vytvorí, alebo poradové číslo, ktoré mu priradíte), pričom toto číslo by malo byť tak na papierovej verzii dotazníkov ako aj v tabuľke dát.

Ďalšia vec, ktorú pravdepodobne pri príprave dát budete zvažovať, je, či do stĺpcov uvádzať sumárne alebo priemerné skóre participantov z nejakého dotazníka, alebo či vkladať do tabuľky všetky údaje (jednotlivé odpovede). Tretie odporúčanie je: ak je to možné, vždy nahadzujte do tabuľky všetky údaje (surové dáta), teda každú jednu odpoveď (na každú položku, otázku) každého participanta. Pri takto pripravených dátach získate viac možností, ako s dátami pracovať, samozrejme vrátane výpočtu spomínaných sumárnych alebo priemerných skóre. Pri on-line zbere dát sú už údaje stiahnuteľné v uvedenej podobe, v každom prípade je potrebná kontrola, prípadne odstránenie stĺpcov s nepotrebnými údajmi.

Obrázok 1 Ukážka tabuľky pred vkladaním zozbieraných údajov

Občas sa stane, že participanti nevyplnia niektorú otázku alebo položku. Je niekoľko spôsobov, ako s chýbajúcimi údajmi naložiť. Najjednoduchšia rada je: nerobte nič. Pri mnohých analýzach chýbajúce údaje nie sú problémom. V prípade, že potrebujete využiť analýzu, ktorá pri chýbajúcich údajoch vyhlasuje chybu, môžete zvoliť niektorý z nasledovných postupov:

- pre danú premennú (stĺpec) vypočítate priemer a túto priemernú hodnotu vložíte do prázdnej bunky (niekedy sa odporúča vložiť nie priemer, ale najčastejšiu hodnotu),
- pre danú štatistickú jednotku (riadok) vypočítate priemerné skóre danej štatistickej jednotky (napr. participanta) pre danú premennú (dotazník, škálu, test) a tento priemer vpíšete do prázdnej bunky (znovu, niekedy sa odporúča vložiť nie priemer, ale najčastejšiu hodnotu),
- do prázdnych buniek vpíšete nulu (to ale môže skresľovať výsledky, napr. pri škálach skórovaných od 1).

V Exceli vypočítame priemer pomocou funkcie *Priemer* v ponuke *Domov* → *Úpravy* → *Automatický súčet*. Postavíme sa na bunku, do ktorej chceme vypočítať priemer, zvolíme funkciu *Priemer*. Zobrazí sa vzorec =AVERAGE(A1:A10), ktorý znamená, že sa bude počítať priemer z buniek A1 až A10 – samozrejme rozsah buniek, z ktorých chceme počítať priemer môžeme upraviť podľa potreby.

2.2 Zmena úrovne

Aby sme vedeli, ako môžeme s jednotlivými dátami pracovať, vždy je potrebné ujasniť si, aký typ dát máme k dispozícii, teda na akej úrovni sme premenné merali (pre pripomenutie, rozlišujeme nominálnu, poradovú a intervalovú úroveň). Podľa toho dáta pripravujeme a nakoniec aj analyzujeme. Niektoré premenné je možné merať iba na jednej úrovni, iné typy premenných možno merať na viacerých úrovniach a je na výskumníkovi, ktorú úroveň pri projektovaní výskumu zvolí. Zmena premennej z jednej úrovne merania na inú je možná, ale iba smerom nadol! Navyše, je potrebné pamätať si, že vyššia úroveň merania poskytuje viac informácií o skúmanom fenoméne. Preto v prípade, ak je možné premennú merať na viacerých úrovniach, je vhodné zvoliť najvyššiu možnú úroveň. Takéto dáta môžeme potom transformovať na nižšie úrovne merania. Opačným smerom to nie je možné. Vezmime si príklad s rokmi praxe.

Roky praxe môžeme merať ako intervalovú premennú, kedy participantov požiadame, aby uviedli, koľko presne majú rokov praxe vo svojom odbore. S touto premennou vieme robiť všetky analýzy. Navyše, vieme ju v prípade potreby transformovať na poradovú premennú, kedy každého participanta zaradíme do kategórie, napríklad:

$$\begin{gathered}\text{menej ako 1 rok}\\ 1-5\text{ rokov}\\ 6-10\text{ rokov}\\ \text{viac ako 10 rokov}\end{gathered}$$

Roky praxe ako poradová premenná síce umožňujú usporiadať participantov od neskúsených po veľmi skúsených, neumožňujú však presne určiť, aký je napríklad priemerný počet rokov praxe participantov v našej vzorke alebo aký je presný rozdiel v praxi medzi dvoma participantmi. A nakoniec, intervalovú premennú vieme transformovať aj na nominálnu premennú: v prípade rokov praxe každého participanta môžeme zaradiť do jednej z dvoch kategórií „má prax“ – „nemá prax“.

3 Zoznámte sa, JASP

Ako už bolo v úvode spomenuté, **JASP** je voľne dostupný štatistický program, ktorý si každý môže stiahnuť do svojho počítača. Ide o program, ktorý sa neustále vyvíja, od svojho vzniku prešiel rôznymi úpravami (v čase dokončovania tejto učebnice bola najaktuálnejšia verzia 0.15.0.0), a preto je možné, že vy už pracujete s niektorou z novších verzií. Z tohto dôvodu sa niektoré ilustračné obrázky v tejto učebnici nemusia úplne zhodovať s tým, čo vidíte vo svojom počítači. Našťastie JASP zachováva základnú logiku svojho prostredia, takže aj napriek malým odchýlkam vám aj s odstupom času môže byť táto učebnica pomôckou pri orientácii v základných štatistických analýzach. Program je užívateľsky priateľský a intuitívny – za predpokladu, že máte aspoň základy štatistiky. K dispozícii je aj príručka, ktorá podrobne vysvetľuje, ako s jednotlivými analýzami pracovať[5].

3.1 Otvorenie dátového súboru

Predtým, ako si dátový súbor pripravený v podobe excelovskej tabuľky otvoríte v programe JASP, uistite sa, že:

- □ v dátovom súbore sú všetky zozbierané údaje,
- □ dáta sú skontrolované a participanti, u ktorých je podozrenie na nesprávne alebo neúplné vyplnené dotazníky, sú vylúčení z analýz,
- □ máte uloženú kópiu súboru s hrubými dátami.

Takto pripravený dátový súbor v Exceli uložíme. Keďže však program JASP nedokáže načítať dáta z klasického excelovského súboru, je potrebné spraviť kópiu súboru vo formáte *.csv*. Až v tomto momente môžeme pristúpiť k tomu, aby sme začali pracovať s dátami v programe JASP.

Dátový súbor uložíme vo formáte .csv tak, že zvolíme možnosti *Súbor* →*Uložiť ako*, v okne na uloženie v riadku *Názov súboru* vpíšeme názov nášho dátového súboru a v riadku *Uložiť vo formáte* zvolíme možnosť *CSV(MS-DOS)*.

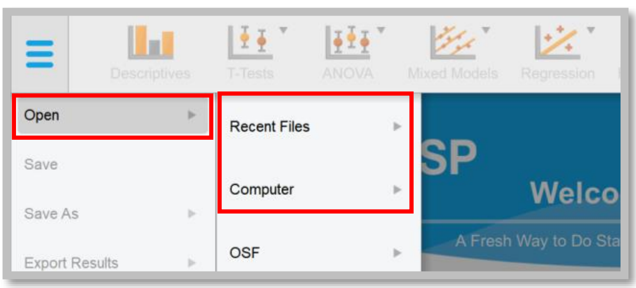
---

[5] Voľne stiahnuteľné príručky k aktuálnym verziám programu sú dostupné na https://jasp-stats.org/jasp-materials/

Po spustení aplikácie JASP je potrebné **otvoriť dátový súbor**, s ktorým chceme pracovať (pripomínam, že by mal byť v podobe excelovskej tabuľky uloženej vo formáte *.csv*, prípadne v inom podporovanom formáte[6]). Pre otvorenie dát vyberte v ľavom hornom rohu v ponuke menu (ikona ≡) možnosť *Open* a následne jednu z možností (Obrázok 2):

- *Recent Files*, v prípade, že sa chceme vrátiť k niektorému z posledných používaných dokumentov,
- *Computer*, v prípade, že chceme otvoriť dokument uložený v našom počítači
- *OSF*, v prípade, že chceme otvoriť dokument z účtu na OSF[7],
- *Data Library*, v prípade, že nemáme vlastné dáta a chceme sa učiť používať jednotlivé analýzy podľa príručky JASP (Goss-Sampson, 2020a); JASP ponúka vo svojom programe vlastnú knižnicu dátových súborov vhodných k učeniu sa jednotlivých analýz.

V rozhraní JASP je na hornej lište hlavná ponuka modulov analýz (*Descriptives*, *T-Tests*, *ANOVA*, *Mixed Models*, *Regression*, *Frequencies*, a *Factor*), z ktorých vyberáme podľa toho, ktoré analýzy chceme aplikovať. Vpravo na lište je navyše symbol + pomocou ktorého môžeme rozšíriť ponuku modulov, napr. o modul *Reliability*, ktorý budete určite využívať.



Obrázok 2 Otvorenie dátového súboru v JASP

[6] *.csv*, *.tsv*, *.txt* (akékoľvek textové súbory so stĺpcami oddelenými čiarkou / bodkočiarkou / dvojbodkou / tabulátorom), alebo *.sav* a *.por* (SPSS súbory), prípadne *sas7bdat* (SAS súbory), *.ods* (Open Document Spreadsheet formáty), a *.dta* (Stata súbory)

[7] OSF – platforma Open science framework (https://osf.io/)

3.2 Transformácia dát

Po otvorení dátového súboru sa zobrazí tabuľka s našimi údajmi. V tomto kroku je vhodné skontrolovať, či majú jednotlivé premenné prisúdenú správnu ikonku podľa toho, či sú merané na

 nominálnej úrovni (*nominal*) ,

 poradovej úrovni (*ordinal*),

 intervalovej úrovni (*scale*).

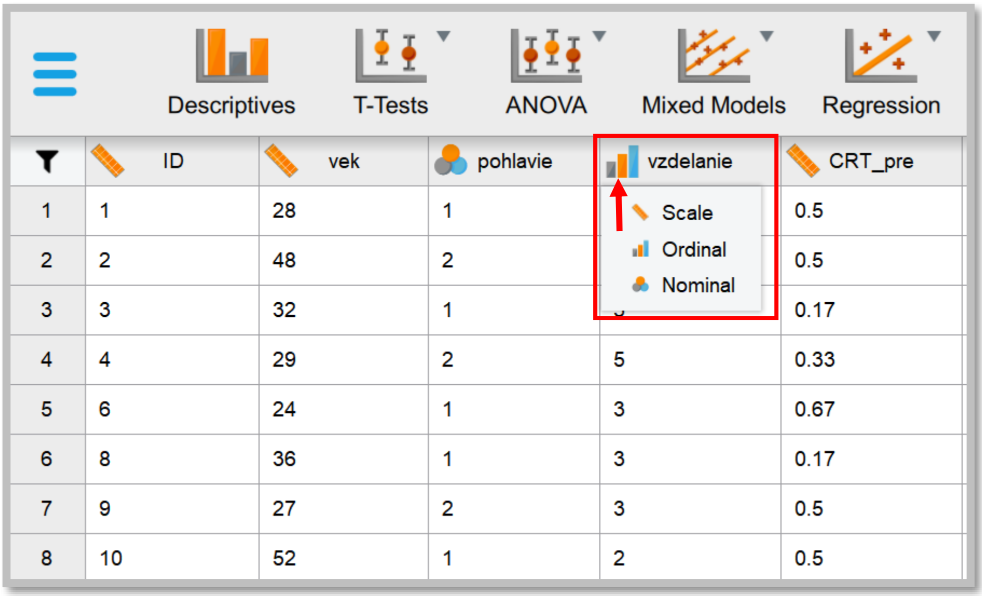
Ak niektorá premenná nemá správne označenie, môžeme manuálne nastaviť úroveň premennej tak, že klikneme na ikonku pri názve premennej a zvolíme správnu úroveň (Obrázok 3).

Aj napriek všetkej snahe sa môže stať, že v súbore objavíte chybu, alebo sa vám hodí doplniť stĺpec s údajmi. Našťastie, údaje je možné upraviť. Autori JASP uvádzajú, že dvojklikom na ktorékoľvek miesto tabuľky JAPS sa dostanete do pôvodného excelovského súboru, kde môžete chybu opraviť, prípadne doplniť údaje. Niekedy však toto prepojenie zlyháva. Vtedy je lepšie a bezpečnejšie JAPS zatvoriť, otvoriť zdrojový *.csv* súbor s dátami, chybu napraviť, súbor uložiť a nanovo ho otvoriť v JASP.

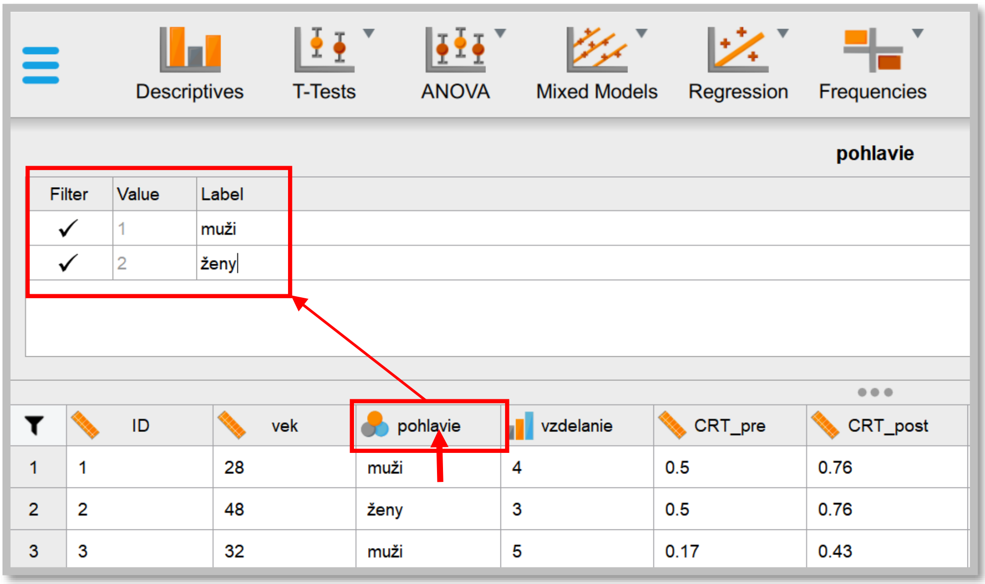
V prípade nominálnych a poradových premenných je vhodné priradiť jednotlivým číselným kódom aj **popis**, aby sa v tabuľkách s výsledkami zobrazovali názvy kategórií a nie iba kódy (napr. muži – ženy a nie 1 a 2). Postupujeme tak, že klikneme na názov premennej a zobrazí sa okno s troma stĺpcami. Pre opis kategórie pracujeme v stĺpci *Label*, kde jednoducho klikneme, zmažeme pôvodný text (číslo), vpíšeme náš opis (Obrázok 4) a potvrdíme tlačidlom enter[8].

Chvíľu sa ešte pristavíme pri okne, ktoré sme otvorili kliknutím na názov premennej. V prvom stĺpci *Filter* sa vyskytuje možnosť odfiltrovať skupinu, ktorú nechceme zaradiť do analýz a následne ju kedykoľvek môžeme zakliknutím opäť do analýz zaradiť. Takto môžeme v prípade potreby vykonávať separátne analýzy.

[8] Už na začiatku práce s dátovým súborom odporúčam vytvoriť si formálny kódovací kľúč – dokument, do ktorého si značíte, čo jednotlivé kódy (čísla) pre tú-ktorú premennú znamenajú (napr. že pre premennú pohlavie 1 = muž, 2 = žena alebo pre škálu 1 = vôbec nesúhlasím, 5 = úplne súhlasím). Je to pomôcka, ku ktorej sa môžete kedykoľvek aj s odstupom času vrátiť, prípadne pomôcka, ktorá slúži aj iným ľuďom pri orientácii vo vašom dátovom súbore.



Obrázok 3 Zmena úrovne merania premennej



Obrázok 4 Opis kategórií nominálnych a poradových premenných

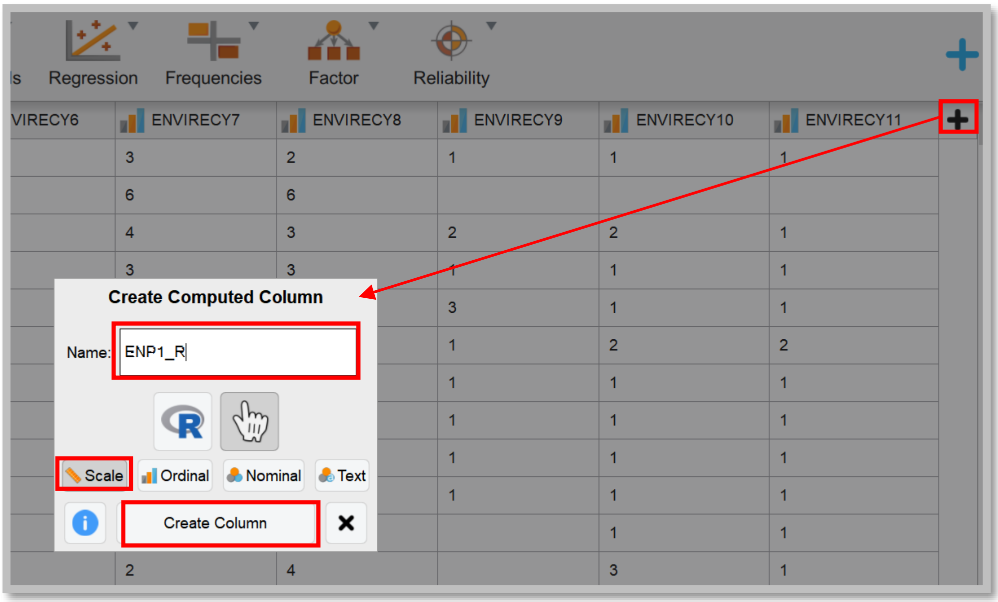
V tomto kroku (keď už máte súbor v programe JASP skontrolovaný, premenné správne označené a kategórie pomenované), môžete si dátový súbor uložiť aj vo formáte JASP – tak zostanú zachované všetky vaše nastavenia a úpravy (ktoré sa inak do *.csv* súboru nepreklopia). Pre uloženie dát vyberte v ľavom hornom rohu v ponuke menu (ikona ≡) možnosť *Save as*, zvoľte umiestnenie *Computer* (ak chcete súbor uložiť vo svojom počítači) a pomocou funkcie *Browse* vyberte konkrétne umiestnenie v pamäti vášho zariadenia.

3.2.1 Re-kódovanie dát

V niektorých nástrojoch sa nachádzajú položky, ktoré sú formulované reverzne, teda v opačnom smere, ako sú konštruované ostatné položky. Napríklad v dotazníku merajúcom úroveň šťastia sa okrem položiek formulovaných na zachytenie šťastia môžu vyskytovať aj položky ako „Často sa cítim smutný/á.“, teda opačne formulované, reverzné. Takéto položky je potrebné reverzne skórovať, re-kódovať. Ak participanti hodnotia spomínané výroky na škále od 1 (úplne súhlasím) po 5 (vôbec nesúhlasím), reverzné položky re-kódujeme podľa kľúča:

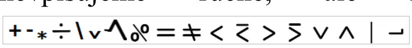
$$
\begin{aligned}
&\textit{skóre } 1 \textit{ na skóre } 5\\
&\textit{skóre } 2 \textit{ na skóre } 4\\
&\textit{skóre } 4 \textit{ na skóre } 2\\
&\textit{skóre } 5 \textit{ na skóre } 1\\
&\textit{skóre } 3 \textit{ nie je potrebné upravovať}
\end{aligned}
$$

V programe JASP re-kódujeme položky nasledovne: premennú (položku), ktorú chceme re-kódovať musíme označiť ako meranú na intervalovej úrovni. Následne si pomocou ikony v pravom rohu vytvoríme novú premennú (nový stĺpec, v ktorom bude nami re-kódovaná premenná). Otvorí sa malé okno, kde do rámčeka *Name* vpíšete názov novej premennej (pri re-kódovaných premenných odporúčam použiť ten istý názov ako je originálna premenná s dodatkom *_R*), napr. pre premennú, ktorú re-kódujeme z premennej *ENP1* použijeme názov *ENP1_R* (Obrázok 5). Označíme, že premenná je meraná na intervalovej úrovni (*Scale*) a potvrdíme vytvorenie premennej tlačidlom *Create Column*.

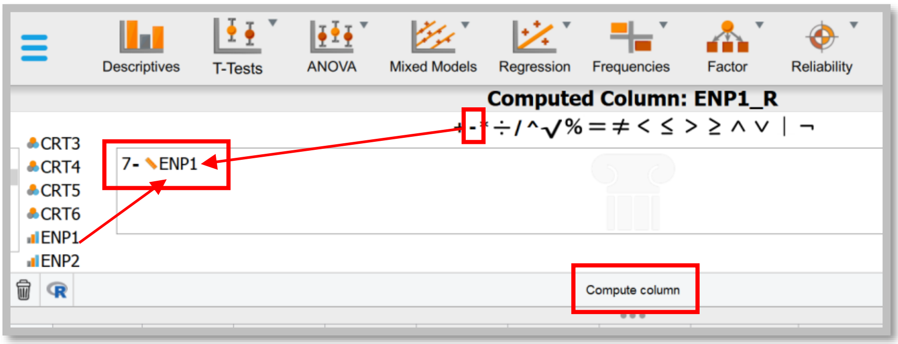


Obrázok 5 Vytvorenie stĺpca pre novú premennú

Vznikne nový stĺpec, ktorý je umiestnený na konci dátového súboru. Zároveň s vytvorením stĺpca sa otvorí nad dátovým súborom okno na výpočet hodnôt do tohto stĺpca (pre novú premennú). Do rámčeka je potrebné zadať vzorec, príkaz. Keď chceme re-kódovať položku, napr. ENP1, potrebujeme vedieť, na akej škále bola meraná, resp. aká je najvyššia hodnota tejto škály. Vieme, že ENP bolo merané na škále od 1 po 6, takže maximálna hodnota pre každú položku ENP je 6 (teda aj pre ENP1). Re-kódovať položku môžeme jednoducho tak, že namerané skóre odčítame od čísla, ktoré je o +1 bod vyššie ako maximálna možná hodnota. V našom príklade teda namerané skóre odčítame od čísla 7 (7 je o +1 bod vyššie číslo ako maximálna možná dosiahnuteľná hodnota 6). Týmto jednoduchým výpočtom dosiahneme, že nameraná hodnota 1 sa re-kóduje na hodnotu 6 (lebo $7 - 1 = \mathbf{6}$), hodnota 2 na hodnotu 5 (lebo $7 - 2 = \mathbf{5}$), atď. až po zmenu hodnoty 7 na hodnotu 1 (lebo $7 - 6 = \mathbf{1}$).

Podľa tejto logiky zadáme do rámčeka vzorec $7 - ENP1$, pričom tento nevpisujeme ručne, ale najskôr si z ponuky nad rámčekom  vyberieme možnosť - . Tým sa do rámčeka presunie vzorec ... - ..., ktorý musíme „naplniť“ hodnotami. Kliknutím na prvé trojbodky vo vzorci ... sa zobrazí kurzor a môžeme vpísať hodnotu, od ktorej budeme odčítať pôvodnú premennú, v našom príklade napíšeme číslo 7. Kliknutím na druhé trojbodky vo vzorci ... sa znova zobrazí kurzor, tentokrát ale nič nevpisujeme, ale z ponuky naľavo vyberieme premennú, ktorú chceme re-kódovať[9]. Tá sa prenesie do vzorca. Teraz klikneme na tlačidlo *Compute column* (Obrázok 6). Hneď uvidíme, ako sa prázdne bunky novej premennej ENP1_R vyplnia vypočítanými re-kódovanými hodnotami.

Tento postup zopakujeme pri všetkých premenných, ktoré je potrebné re-kódovať.



Obrázok 6 Výpočet re-kódovanej premennej

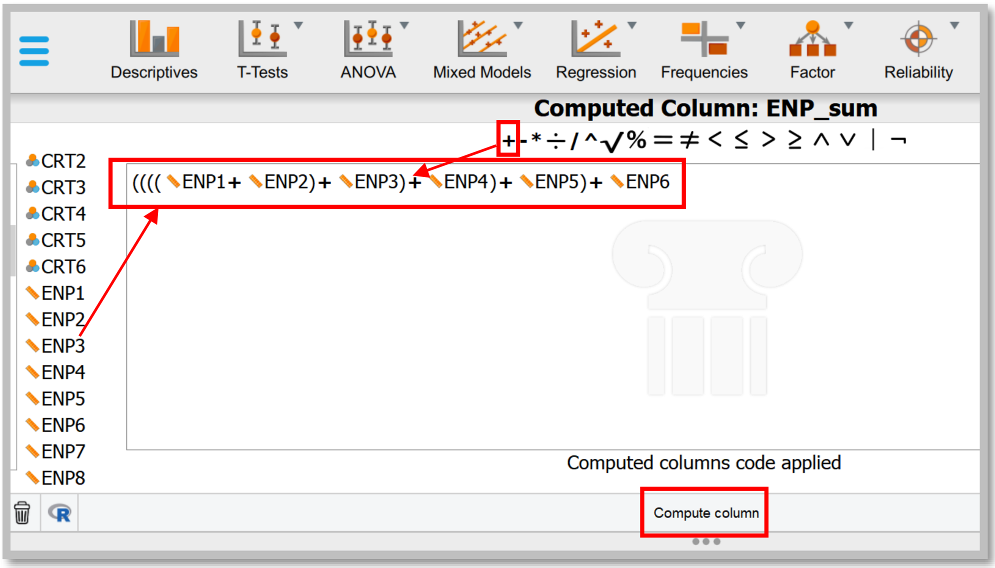
3.2.2 Výpočet nových premenných

Keď máme všetky premenné (položky v stĺpcoch) pripravené, je pravdepodobné, že budeme potrebovať z niektorých z nich počítať skóre, či už priemerné alebo sumárne (to závisí od použitého meracieho nástroja). Znova musíme vytvoriť nové stĺpce pre premenné, ktoré budeme počítať (postup viď 3.2.1 Re-kódovanie dát – Obrázok 5).

Po vytvorení nového stĺpca sa otvorí okno, kam zadáme vzorec výpočtu (v prípade, že si najskôr vytvoríte všetky stĺpce pre plánované nové premenné, stačí, ak v novom stĺpci kliknete na symbol f_x a otvorí okno na zadanie vzorca). Vzorec na výpočet **sumárneho skóre** zadáme jednoducho: položky, z ktorých má byť sumárne skóre počítané postupne presúvame do okna (jednoducho na ne klikneme v zozname naľavo) s tým, že za každou položkou je potrebné kliknúť

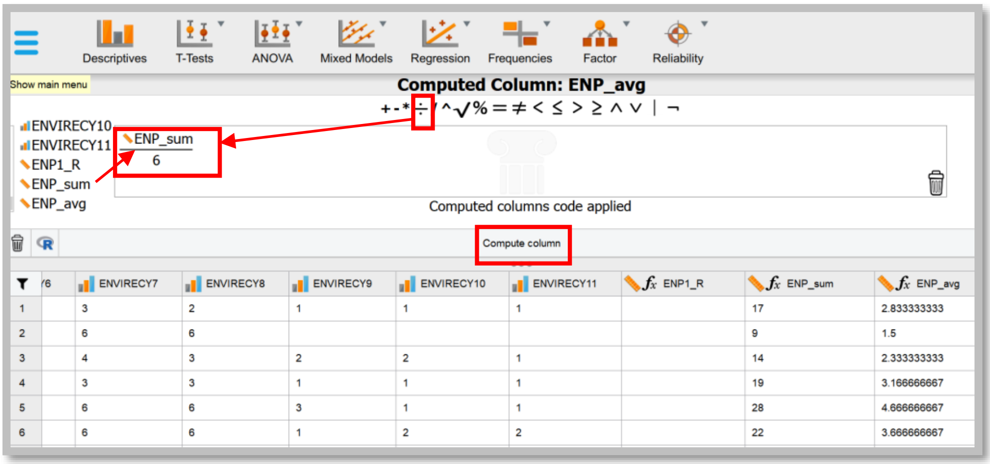
[9] Re-kódované položky musia byť označené, že sú merané na intervalovej úrovni (*Scale*).

na znamienko + z ponuky nad rámčekom +-*÷/^√%=≠<≤>≥∧∨|¬ (Obrázok 7). Ako keď zadávate sčítanie na kalkulačke, iba namiesto čísiel zadávame jednotlivé položky. Výpočet potvrdíte tlačidlom *Compute column*. V príklade bolo počítané sumárne skóre z prvých 6 položiek ENP, premennú sme označili *ENP_sum*.



Obrázok 7 Výpočet novej premennej – sumárne skóre

Ak potrebujete vypočítať **priemerné skóre**, je potrebné spraviť jeden medzikrok – výpočet už spomínaného sumárneho skóre. Takže na výpočet priemerného skóre musíme vytvoriť dva nové stĺpce – jeden pre sumárne skóre (medzikrok) a jeden pre finálne priemerné skóre. V príklade sme premennú označili *ENP_avg*. Po výpočte sumárneho skóre z príslušných položiek, zadáme do nového stĺpca pre priemerné skóre nasledovný vzorec: sumárne skóre / počet položiek, z ktorých sa sumárne skóre počítalo. V našom príklade do okna presunieme sumárne skóre, teda *ENP_sum* (jednoduchým kliknutím na premennú v zozname naľavo), následne klikneme na znamienko ÷ z ponuky nad rámčekom +-*÷/^√%=≠<≤>≥∧∨|¬ (Obrázok 8) a na záver napíšeme číslo zodpovedajúce počtu položiek, z ktorých sa počítalo sumárne skóre, teda v príklade je to číslo 6. Výpočet potvrdíte tlačidlom *Compute column*. Priemerné skóre bolo počítané z prvých 6 položiek ENP (teda za pomoci ENP_sum), a vzorec vyzeral nasledovne $\frac{ENP\_sum}{6}$.



Obrázok 8 Výpočet novej premennej – priemerné skóre

4 Deskriptívna štatistika

Dáta, ktoré sme pri zbere získali potrebujeme v prvom kroku organizovať, opísať, charakterizovať a sumarizovať. V tejto kapitole preto predstavím postupy, pomocou ktorých môžete zhrnúť a popísať dáta, inými slovami, deskriptívnu štatistiku. Pôjde o miery polohy alebo stredných hodnôt (*central tendency*) a miery variability (*variability*), ale aj postupy, ako sa zhodnotiť distribúciu dát a ako pracovať s frekvenciami dát.

4.1 Miery polohy (centrálnej tendencie)

Jeden z najčastejšie používaných štatistických príkladov hovorí, že keď by sme mali dostatočne veľký súbor osôb, tak prevažná väčšina ľudí v tomto súbore by mala v skutočnosti viac nôh ako by bol priemerný počet nôh pre túto skupinu ľudí (priemer by bol napríklad 1,9). Je to spôsobené tým, že nie každá osoba v súbore by mala obe nohy, niekto by mal iba jednu nohu, a niekto by dokonca nemusel mať žiadnu. Priemerný počet nôh by preto v tomto súbore vyšiel nižší ako 2 (napr. spomínaných 1,9). Teda ak má väčšina ľudí v skupine 2 nohy, potom väčšina má reálne viac nôh (2) ako je priemer skupiny (1,9).

V kvantitatívnom výskume je väčšinou k dispozícií príliš veľa informácií na to, aby sme ich mohli prezentovať všetky. Predstavte si, že máte v triede 25 študentov a študentiek. Keď sa o nich s niekým rozprávate, nepoviete o každom jednom z nich, koľko má rokov, ale poviete niečo v duchu „väčšina z nich má 20 rokov“, alebo „majú v priemere 19,5 roka“. Toto je veľmi zjednodušený spôsob základného štatistického opisu dát za použitia mier **polohy** alebo **centrálnej tendencie**. K základným mieram polohy patria *priemer*, *medián* a *modus*. Miery polohy údajov sú väčšinou prvou a najbežnejšou formou deskriptívnej štatistiky. Prostredníctvom jednej hodnoty, jedného čísla, zobrazujú stred skupiny dát (Coolican, 2019; Dancey & Reidy, 2017; Hanna & Dempster, 2012). Inými slovami, miery polohy pomáhajú vytvoriť predstavu o distribúcii dát, predstavu o tom, ako sú dáta usporiadané, kde sú ich stredné hodnoty. Najskôr si ich predstavíme a následne si spolu s mierami variability, ukážeme, ako ich v JASP vypočítať.

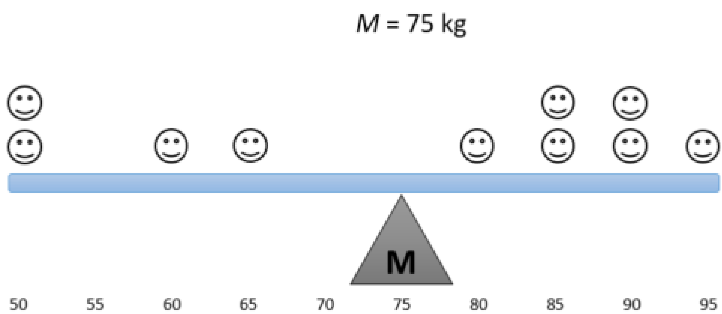
4.1.1 Priemer (M, $\bar{x}$)

Aritmetický priemer je najznámejšou mierou polohy a používame ho pri premenných meraných na intervalovej úrovni. Priemer sa učíme počítať už na základnej škole a neraz ho využívame aj v bežnom živote. Priemer predstavuje sumár všetkých skóre (dát) vydelený ich počtom. Ak napríklad máme známky 10 žiakov v 3.A triede z písomky z matematiky (3, 2, 2, 2, 3, 1, 1, 2, 1, 3), potom priemer vypočítame nasledovne:

$$M = \frac{3 + 2 + 2 + 2 + 3 + 1 + 1 + 2 + 1 + 3}{10} = \frac{20}{10} = 2$$

Zistíme, že priemerná známka triedy z písomky je 2 – informácie o hodnotení výkonu každého žiaka v triede sú zhrnuté do jedného čísla, do jednej známky, ktorá dáva predstavu o priemernom výkone žiakov v danej písomke. Priemer je stredom všetkých hodnôt, je typickou hodnotou sledovanej premennej v danej skupine. Priemer si môžeme predstaviť ako bod na udržanie rovnováhy dosky, na ktorej sú umiestnené sledované subjekty (Obrázok 9), napríklad ľudia s rôznou váhou. Matematicky predstavuje priemer bod, od ktorého sumárna vzdialenosť všetkých štatistických jednotiek pod ním je rovnaká ako sumárna vzdialenosť všetkých štatistických jednotiek nad ním (Aron et al., 2014).

V Obrázku 9 to znamená, že súčet vzdialeností od priemeru u všetkých ľudí pod 75 kg, teda pod priemerom (4 osoby, ktorých váha sa od priemeru líši nasledovne: 25 + 25 + 15 + 10 = *75*) je rovnaký ako súčet vzdialeností od priemeru u všetkých ľudí nad 75 kg, teda nad priemerom (6 osôb, ktorých váha sa od priemeru líši nasledovne: 5+ 10 + 10 + 15 + 15 + 20 = *75*).



Obrázok 9 Ilustrácia priemeru distribúcie váhy 10 osôb

Priemer má nesporne svoje výhody, no spája sa s ním aj niekoľko nevýhod. Hlavnou jeho nevýhodou je skutočnosť, že je veľmi ovplyvniteľný extrémnymi hodnotami premennej. Vráťme sa k 3.A triede a k známkam jej žiakov (10 žiakov so známkami: 3, 2, 2, 2, 3, 1, 1, 2, 1, 3). Zistili sme, že priemerná známka v tejto triede je 2. Predstavme si ďalšiu triedu, 3.B, o ktorej vieme, že priemerná známka v nej je tiež 2. Pri bližšom pohľade na konkrétne známky v 3.B však zistíme, že napriek rovnakej priemernej známke v oboch triedach sa výkony žiakov odlišujú:

3. A: 3, 2, 2, 2, 3, 1, 1, 2, 1, 3

3. B: 1, 1, 1, 2, 1, 1, 5, 2, 1, 5

V 3.A je viacero žiakov s priemerným výkonom, a v 3.B je viacero výborných žiakov a dvaja veľmi slabí žiaci. Žiaci týchto tried sa od seba výkonom odlišujú, priemerné známky však majú rovnaké. Aby sme sa vyhli podobnému skreslenému obrazu o rozložení dát (o ich polohe), odporúča sa využívať aj ďalšie miery polohy – medián a modus.

4.1.2 Medián (Md)

Druhou zo základných mier polohy je medián – stredná hodnota, ktorú získame, ak usporiadame dáta od najnižšej hodnoty po najvyššiu. Ak je stredná hodnota dát medzi dvoma hodnotami, potom mediánom bude priemer týchto hodnôt. Inými slovami, ak poznáme medián, vieme povedať, že 50% dát je menších ako hodnota mediánu a 50% dát je väčších ako hodnota mediánu. Vráťme sa k našim triedam. Známky z písomky v 3.A (3, 2, 2, 2, 3, 1, 1, 2, 1, 3) usporiadame od najnižšej po najvyššiu a nájdeme strednú hodnotu

$$1, 1, 1, 2, 2 \text{ — } 2, 2, 3, 3, 3$$

Stredná hodnota, teda medián je 2. V 3.B spravíme to isté. Známky (1, 1, 1, 2, 1, 1, 5, 2, 1, 5) tak isto zoradíme a nájdeme strednú hodnotu

$$1, 1, 1, 1, 1 \text{ — } 1, 2, 2, 5, 5$$

a určíme medián = 1. Napriek tomu, že obe triedy majú rovnakú priemernú známku, vidíme, že majú odlišný medián. V 3.A triede má polovica žiakov známku lepšiu ako 2 a druhá polovica má známku horšiu ako 2, a v 3.B triede má polovica žiakov známku lepšiu ako 1 (teda má 1) a druhá polovica má známku horšiu ako 1. Keď si dáta takto prerozprávame, získavame presnejšiu predstavu o distribúcii známok v triedach.

4.1.3 Modus (Mo)

Tretím ukazovateľom miery polohy je modus. Modus predstavuje hodnotu, ktorá je v sledovaných dátach (v distribúcii) najčastejšia, má najvyššiu frekvenciu výskytu. V 3. A triede je spomedzi známok 3, 2, 2, 2, 3, 1, 1, 2, 1, 3 najčastejšie zastúpená známka 2, vyskytuje sa v súbore dát 4-krát. Takže modus sledovaných dát je 2. V 3. B triede je spomedzi známok 1, 1, 1, 2, 1, 1, 5, 2, 1, 5 najčastejšia známka 1, teda modus tejto triedy je 1.

Keď si zhrnieme všetky miery polohy pre 3. A a 3. B, tak vieme povedať, že hoci priemerná známka v oboch triedach je rovnaká (2), v 3. A má polovica žiakov známku 2 alebo horšiu známku (medián) a známka 2 je aj najčastejšia v tejto triede (modus) a v 3. B má polovica žiakov známku 1 alebo horšiu (medián) a známka 1 je aj najčastejšia (modus). Obraz o dvoch triedach sa vďaka trom mieram polohy dát vyostril.

4.2 Miery variability

Miery variability opisujú, ako sa jednotlivé hodnoty navzájom podobajú, respektíve nakoľko sa odlišujú. Inými slovami, aká veľká je variabilita distribúcie (rozloženia) dát. Keď sa vrátime k triedam 3. A a 3. B, tak k vytvoreniu predstavy o tom, ako sú známky jednotlivých žiakov v triedach distribuované nám slúžia základné miery variability – *rozpätie* a *štandardná odchýlka*.

4.2.1 Rozpätie

Jedným z najrýchlejších a najjednoduchších spôsobov, ako získať informáciu o tom, ako sú skóre premennej rozptýlené, je porovnanie minimálnej a maximálnej hodnoty. Hodnota, ktorú získame, keď vypočítame rozdiel medzi minimálnym a maximálnym skóre, sa nazýva **variačné rozpätie** alebo variačná šírka (*range*, označujeme ho aj *R*). Pre 3. A triedu by bolo rozpätie známok z matematiky 2 (najvyššia známka je 3, najnižšia 1; $3 - 1 = 2$) a pre 3. B by bolo rozpätie 4 (najvyššia známka je 5, najnižšia 1; $5 - 1 = 4$). Hoci je rozpätie užitočná miera (ohraničuje dáta), nenapovie veľa toho, čo sa s dátami v rámci tohto rozpätia deje, ako sú usporiadané. Rozpätie môže byť veľmi skresľujúce, podobne ako pri priemere, dva súbory s rôznym rozložením dát môžu mať rovnakú hodnotu rozpätia, no distribúcia dát v rámci týchto rozpätí môže byť odlišná. Viac informácií o distribúcií dát poskytuje napríklad rozptyl a štandardná odchýlka.

4.2.2 Rozptyl (δ^2) a štandardná odchýlka (SD; δ)

Na rozdiel od rozpätia, štandardná odchýlka poskytuje informácie o tom, čo sa deje s dátami medzi minimálnou a maximálnou hodnotou. **Štandardná odchýlka** je miera, do akej sa hodnota odlišuje od priemernej hodnoty. Aby sme sa dopracovali k štandardnej odchýlke, potrebujeme poznať niekoľko ďalších štatistík – priemernú odchýlku a rozptyl.

Každá hodnota (napr. skóre každého participanta) sa od priemernej hodnoty odlišuje o určitú hodnotu, a ak od každej hodnoty odčítame priemer, získame údaj o tom, ako ďaleko je každá hodnota od priemeru (Tabuľka 1). Priemer týchto rozdielov nazývame *priemerná odchýlka*.

Tabuľka 1 Vzdialenosť každej známky od priemernej známky (príklad pre 3. A)

	známky										priemer
hodnoty (známky)	3	2	2	2	3	1	1	2	1	3	2
rozdiel hodnoty od priemeru	1	0	0	0	1	-1	-1	0	-1	1	0

Problém s priemernou odchýlkou je v tom, že polovica rozdielových hodnôt je negatívna a polovica pozitívna, a spolu tvoria nulový súčet (spomeňte si na Obrázok 9 s vyvažovaním dosky). Táto informácia nie je teda veľmi výpovedná. Pre výpočet štandardnej odchýlky je preto potrebné odchýlky od priemeru umocniť (tým aj negatívne skóre získajú pozitívne hodnoty) a z týchto mocnín vypočítať priemer. Táto hodnota sa nazýva **rozptyl**. Konkrétne, v 3. A všetky rozdiely hodnôt od priemeru umocníme na druhú

$$1^2, 0, 0, 0, 1^2, -1^2, -1^2, 0, -1^2, 1^2 \ \rightarrow \ 1, 0, 0, 0, 1, 1, 1, 0, 1, 1$$

a vypočítame rozptyl (priemer: 6 / 10 = 0,6; Tabuľka 2).

Ale ani rozptyl, tým, že sa jedná o priemer mocnín, nie je konečnou stanicou na ceste za štandardnou odchýlkou, pretože je priemerom umocnených hodnôt. V poslednom kroku preto odmocníme rozptyl a tým získame **štandardnú odchýlku** – mieru veľkosti variácie alebo rozptylu súboru hodnôt, teda ako široko sú hodnoty rozptýlené. V 3. A odmocnením rozptylu 0,6 prídeme k štandardnej odchýlke 0,77. S interpretáciou štandardnej odchýlky na chvíľu počkáme, pretože pre jej správne porozumenie je potrebné porozumieť distribúcii dát, ktorej je venovaná nasledovná kapitola.

Tabuľka 2 Vzdialenosť každej známky od priemernej známky (príklad pre 3. A)

	známky										
hodnoty (známky)	3	2	2	2	3	1	1	2	1	3	priemer = 2
rozdiel hodnoty od priemeru	1	0	0	0	1	-1	-1	0	-1	1	priemerná odchýlka = 0
rozdiel hodnoty od priemeru	1	0	0	0	1	1	1	0	1	1	rozptyl = 0,6

4.2.3 Distribúcia dát

Doteraz sme sa na dáta pozerali cez optiku mier polohy a variability. Dáta, s ktorými pracujeme však môžeme nielen organizovať, ale aj zobrazovať. Najčastejšie sa tak deje pomocou tabuliek a grafov. Na to je potrebné, aby sme pracovali s frekvenciami dát, pravdepodobnosťami a rozložením, teda s **distribúciou**.

To, že na väčších súboroch dát môžeme pozorovať distribúciu dát znamená, že môžeme pozorovať opakovanie rovnakých alebo podobných hodnôt – môžeme sledovať **frekvenciu** výskytu jednotlivých hodnôt. V príklade s 3.A a 3.B frekvencia znamená, že sledujeme, koľko žiakov v triede dostalo akú známku. Sledujeme frekvenciu výskytu každej známky od 1 po 5. Ukážka frekvenčnej tabuľky pre obe triedy je v Tabuľke 3.

Tabuľka 3 Príklad frekvenčnej tabuľky

	3.A			3.B		
známka	frekvencia výskytu	% podiel výskytu	kumulatívne %	frekvencia výskytu	% podiel výskytu	kumulatívne %
1	3	30%	30%	6	60%	60%
2	4	40%	70%	2	20%	80%
3	3	30%	100%	0	0%	80%
4	0	0%	100%	0	0%	80%
5	0	0%	100%	2	20%	100%

V stĺpci *známka* uvádzame jednotlivé druhy hodnôt (známok), ktoré môžu žiaci dosiahnuť; v stĺpci *frekvencia* uvádzame počet žiakov, ktorí danú hodnotu dosiahli; v stĺpci s *%* percentuálny podiel počtu participantov; a v poslednom stĺpci *kumulatívne %* sú uvedené percentá participantov, ktoré sa vyskytli až po danú hodnotu (známku, teda v 3. A 70% žiakov dostalo známku 2 alebo lepšiu).

Pomocou frekvenčnej tabuľky môžeme identifikovať ukazovatele ako sú *percentily* a *kvartily*. Pozrite si teraz Tabuľku 4. V prvom stĺpci je počet bodov, ktoré ste mohli dosiahnuť v teste. Vy ste napríklad dosiahli 60 bodov. Aby ste si vytvorili predstavu o tom, čo vašich 60 bodov znamená, môžete sa pozrieť, koľko bodov bolo maximum (70) a porovnať svoj výkon s maximom. Alebo sa môžete porovnať s ostatnými spolužiakmi a to tak, že zistíte, koľko ľudí dosiahlo tiež 60 bodov tak, ako vy (60 bodov ste dosiahli vy a ďalší 2 z 80 spolužiakov). Ešte lepším ukazovateľom vášho výkonu je **percentil**, na ktorom ste sa umiestnili. Ide o bod, ktorý rozdeľuje populácia na dve skupiny (percentil hľadáme v stĺpci s kumulatívnymi percentami). Keď ste dosiahli 60 bodov, nachádzate sa na 98,8. percentile – to znamená, že iba 1,2% spolužiakov dosiahlo lepší výkon, ako vy. Niekto, kto skončil na 20. percentile, podal výkon, ktorý prekonalo až 80% jeho/jej spolužiakov.

Tabuľka 4 Príklad percentilov a kvartilov

počet bodov	n	%	kumulatívne %	
10	4	5,0	5,0	
15	5	6,3	11,3	1. kvartil
20	6	7,5	18,8	
25	8	10,0	28,8	2. kvartil
30	10	12,5	41,3	
35	11	13,8	55,0	3. kvartil
40	11	13,8	68,8	
45	9	11,3	80,0	
50	7	8,8	88,8	
55	5	6,3	95,0	
60	3	3,8	98,8	4. kvartil
70	1	1,3	100,0	
spolu	80	100%	100%	

Z frekvenčnej tabuľky môžeme ďalej vyčítať **kvartily**, teda body, ktoré rozdeľujú vzorku na štyri skupiny, a to podľa 25., 50., a 75. percentilu. Takéto delenie možno využiť napríklad pri rovnomernom zadeľovaní participantov do podskupín podľa ich skóre v jednej premennej. V Tabuľke 4 je 25. percentilom medzi 20 a 25 bodmi, 50. percentil je medzi 30 a 35 bodmi a 75. percentil medzi 40 a 45 bodmi. Do prvého kvartilu by sme zaradili študentov, ktorí dosiahli skóre 20 a menej, do druhého kvartilu tých, ktorí skórovali medzi 25 a 30 bodmi, do tretieho tých, ktorí získali 35 – 40 bodov, a do štvrtého kvartilu by boli zaradení tí, ktorí skórovali 45 a viac bodmi.

To, ako sú vaše dáta rozložené, distribuované, je dôležité, pretože mnohé zo štatistických testov, ktoré budete využívať, stavajú na predpokladoch o tom, ako sú dáta distribuované (Dancey & Reidy, 2017). Ideálna alebo optimálna podoba distribúcie dát (s ktorou sa v realite nie vždy stretnete) je tzv. normálne rozdelenie, **normálna distribúcia**. Aby sme distribúciu mohli označiť za normálnu, mali by byť

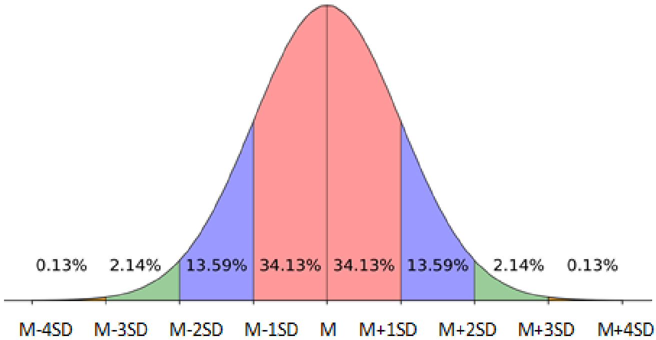
- □ dáta usporiadané symetricky okolo priemeru,
- □ priemer, medián a modus by mali byť identické,
- □ konce distribúcie by sa mali limitne blížiť k osy x, a
- □ dáta by mali byť rozložené do zvonovitého tvaru (Dancey & Reidy, 2017; McLaughlin & McGill, 2017).

Príklad normálnej distribúcie je na Obrázku 10. Hoci nie vždy budú distribúcie identické s tou na obrázku, pokým distribúcia spĺňa vyššie uvedené podmienky, môžeme ju označiť za normálnu. Mnoho prirodzene sa vyskytujúcich premenných je v populácii normálne distribuovaných (napr. výška a váha ľudí).

Normálna distribúcia je charakterizovaná tým, že je funkciou priemeru a štandardnej odchýlky. Platí, že pri normálnej distribúcii sa

- hodnoty **68,27%** participantov nachádzajú v intervale $< M - 1\,SD;\ M + 1\,SD >$
- hodnoty **95,45%** participantov nachádzajú v intervale $< M - 2\,SD;\ M + 2\,SD >$
- hodnoty **99,70%** participantov nachádzajú v intervale$< M - 3\,SD;\ M + 3\,SD >$

Na chvíľu sa vrátime k príkladu so známkami v 3.A triede – k priemeru a štandardnej odchýlke týchto známok. Zistili sme, že v 3.A je priemerná známka 2 a štandardná odchýlka 0,77. To znamená, že známky 68,27% žiakov sú rozložené v okruhu 0,77 bodu od priemernej známky 2, teda v intervale **1,23** (2 – 0,77) až **2,77** (2 + 0,77).

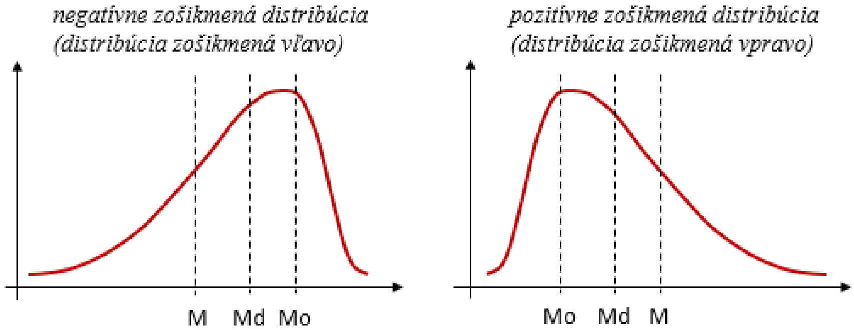


Obrázok 10 Normálna distribúcia dát

Či sú dáta distribuované normálne zistíme v niekoľkých krokoch. Ako prvá by mala byť vizuálna kontrola histogramu[10] dát – tvar krivky napovie ako sú dáta rozložené. Následne je vhodné skontrolovať ukazovatele mier polohy (priemer, medián, modus) a variability (štandardná odchýlka). Ak je priemerná hodnota súboru zároveň aj strednou a najčastejšou hodnotou ($M = Md = Mo$), znamená to, že dáta sú pravdepodobne rozložené normálne. Prax však ukazuje, že ideálne symetricky rozložené dáta nie sú pravidlom, preto pre čo najpresnejší obraz o distribúcii dát je potrebné sledovať aj ďalšie ukazovatele normality rozloženia dát. Tými sú napríklad hodnoty šikmosti (*skewness*) a špicatosti (*kurtosis*).

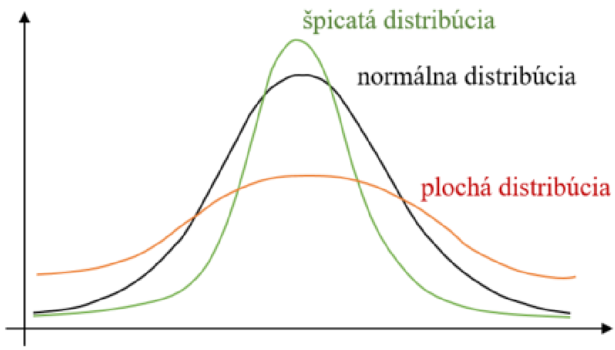
Šikmosť rozloženia dát hovorí o tom, či je distribúcia dát deformovaná doprava alebo doľava, teda či je viac dát na jednej alebo druhej strane distribúcie (Obrázok 11). Ak je hodnota šikmosti negatívne číslo, „chvost" distribúcie viac predĺžený v ľavej časti a hovoríme, že distribúcia je zošikmená vľavo (negatívne zošikmená distribúcia). Ak je hodnota šikmosti pozitívne číslo, „chvost" distribúcie je viac predĺžený v pravej časti, hovoríme, že distribúcia je zošikmená vpravo pozitívne zošikmená).

[10] Histogram je stĺpcový graf, kde výška stĺpca reprezentuje veľkosť alebo početnosť nameraných hodnôt.



Obrázok 11 Šikmosť distribúcie

Strmosť distribúcie je ukazovateľom toho, nakoľko sú dáta nahustené okolo priemeru (Obrázok 12). Ak je distribúcia je *špicatá*, hodnota strmosti je pozitívna, krivka je úzka a vysoká a znamená to, že dáta sú nahustené okolo priemeru. Ak je krivka *plochá*, hodnota strmosti je negatívna, dáta sú rozložené ďalej od priemeru – relatívne veľké množstvo dát je rozložené aj na „chvostoch" krivky, kde sú extrémne hodnoty.



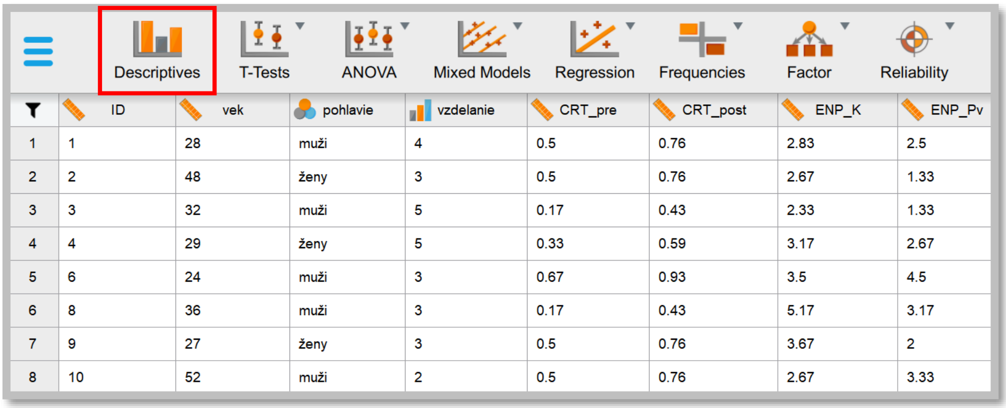
Obrázok 12 Strmosť distribúcie

Pre označenie distribúcie za normálnu je potrebné, aby sa hodnoty šikmosti a strmosti približovali nule: čím viac sa hodnota šikmosti blíži nule, tým je krivka symetrickejšia. Pre vyhodnotenie dát ako normálne distribuovaných pokladajú niektorí autori (napr. Dancey & Reidy, 2017) za prijateľné rozmedzie od -1 do +1, iní autori (napr. George & Mallery, 2010) sú v posudzovaní hodnôt šikmosti a strmosti benevolentnejší a za prijateľné podkladajú aj hodnoty medzi -2 a +2.

Poslednou kontrolou normality rozloženia dát je overenie prostredníctvom Shapiro-Wilkovho testu, ktorý overuje, nakoľko sa rozloženie našich dát odlišuje od normálnej distrubúcie (výpočet Shapiro-Wilkovho testu nájdete v kapitole 7 *Testovanie rozdielov*). Treba mať na pamäti, že hodnota Shapiro-Wilkovho testu závisí od veľkosti vzorky, pri väčších vzorkách je vyššia pravdepodobnosť, že vyjde signifikantný výsledok. Preto sa na normalitu rozloženia dát pozerám komplexne, pomocou všetkých dostupných ukazovateľov: mier polohy a variability, hodnôt šikmosti, strmosti, veľkosť vzorky a výsledku Shapiro-Wilkovho testu.

4.3 Výpočet deskriptívnych štatistík

Vo všeobecnosti v JASP je prostredie usporiadané tak, že po výbere analýzy sa v ľavej časti okna objaví ponuka premenných a špecifík analýzy a v pravom okne je priestor na výsledky (*Results*). V štatistickom programe JAPS vypočítame deskriptívne štatistiky—miery polohy a miery variability—v sekcii *Descriptives* (Obrázok 13).

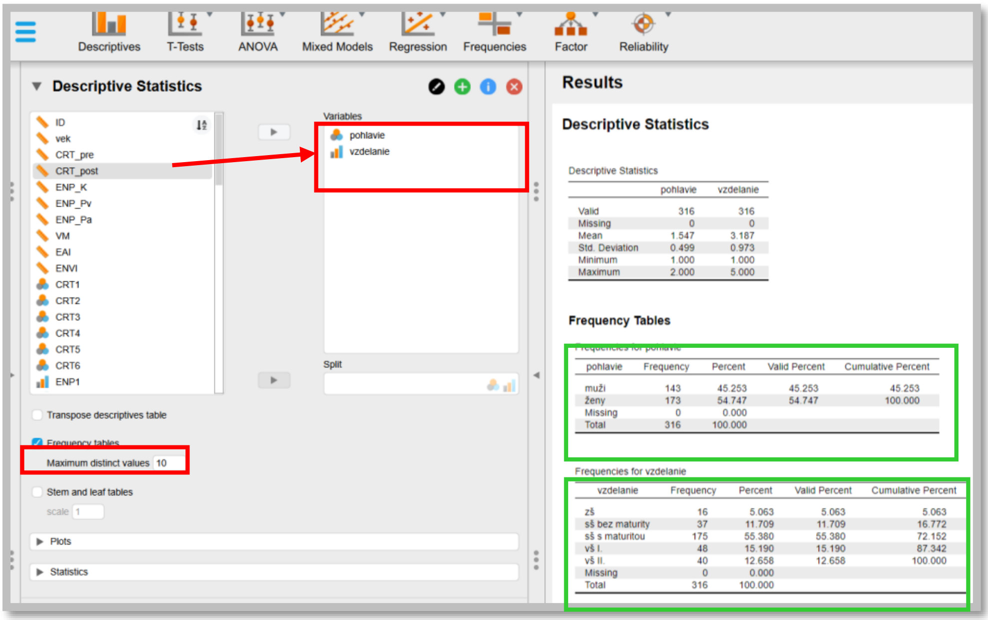


	ID	vek	pohlavie	vzdelanie	CRT_pre	CRT_post	ENP_K	ENP_Pv
1	1	28	muži	4	0.5	0.76	2.83	2.5
2	2	48	ženy	3	0.5	0.76	2.67	1.33
3	3	32	muži	5	0.17	0.43	2.33	1.33
4	4	29	ženy	5	0.33	0.59	3.17	2.67
5	6	24	muži	3	0.67	0.93	3.5	4.5
6	8	36	muži	3	0.17	0.43	5.17	3.17
7	9	27	ženy	3	0.5	0.76	3.67	2
8	10	52	muži	2	0.5	0.76	2.67	3.33

Obrázok 13 Deskriptívna štatistika – 1. krok

Zobrazí sa okno *Descriptive Statistic*, kde zo zoznamu premenných vyberáme a presúvame do okna *Variables* tie premenné, ktoré chceme podrobiť výpočtom (Obrázok 14).

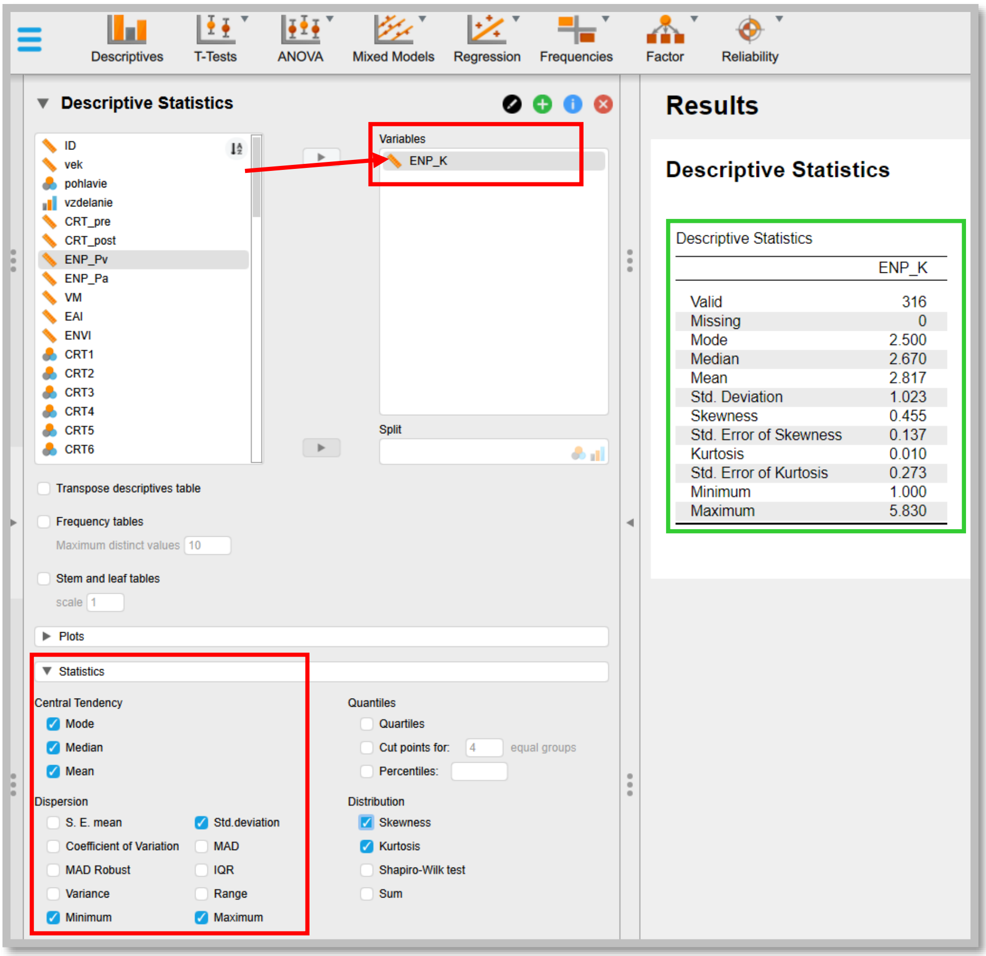
V prvom kroku sa pozrieme na premenné merané na nominálnej a ordinálnej úrovni (v našom príklade *pohlavie* a *vzdelanie*). Pri nominálnych a ordinálnych premenných si vo výsledkovej časti (*Results*) nevšímame tabuľku *Descriptive Statistics*, ale si zvolíme (zaškrtneme) možnosť *Frequency tables* pod zoznamom premenných (Obrázok 14). Vo výsledkovej časti sa zobrazia frekvenčné tabuľky (*Frequency Tables*), z ktorých vieme vyčítať početnosti jednotlivých kategórií každej sledovanej premennej. V našom príklade to znamená, že výskumu sa zúčastnilo 143 mužov (45,25%) a 173 žien (54,75%); 16 participantov (5,06%) malo základné vzdelanie, 37 (11,71%) mala stredoškolské bez maturity, 175 (55,38%) mala stredoškolské s maturitou, 48 (15,19%) vysokoškolské 1. stupňa, a 40 (12,66%) vysokoškolské 2. stupňa.



Obrázok 14 Deskriptívna štatistika – nominálne a ordinálne premenné

V druhom kroku do okna *Variables* presunieme premenné na kardinálnej úrovni (v našom príklade napríklad priemerné skóre konšpiračných presvedčení *ENP_K*). V základnej tabuľke sa objavia výsledky uvádzajúce počet participantov (*valid*), počet chýbajúcich údajov (*missing*), a priemerné skóre (*mean*). Aby sme získali údaje ku všetkým potrebných deskriptívnym štatistikám, je potrebné rozbaliť možnosť *Statistics* a navoliť požadované údaje (Obrázok 15). Môžeme zvoliť výpočet mediánu (*median*), modusu (*mode*), štandardnej odchýlky (*std. deviation*), ale aj minimálne skóre (*minimum*), maximálne skóre (*maximum*), či iné ukazovatele. Rovnako môžeme zvoliť aj miery šikmosti (*skewness*) a špicatosti (*kurtosis*).

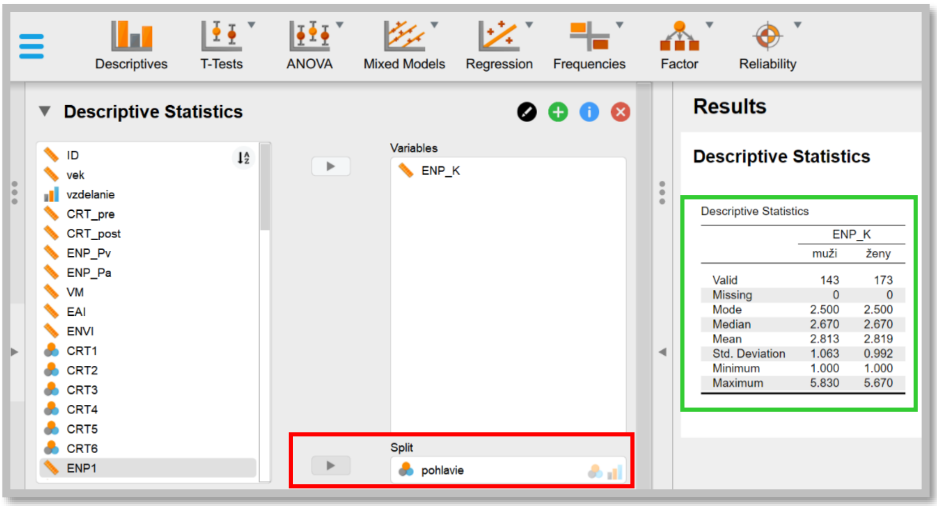
V našom príklade vo výsledkoch vidíme, že výskumu sa zúčastnilo 316 participantov, priemerné skóre konšpiračných presvedčení bolo 2,82 a štandardná odchýlka 1,02; medián bol 2,67 (polovica participantov dosiahla skóre nižšie ako 2,67 bodov), modus 2,5 (najčastejšie dosiahnuté skóre bolo 2,5). Čo sa týka normality rozloženia dát, šikmosť (0,46) aj strmosť (0,01) boli relatívne nízke, takže môžeme usudzovať na normálne rozloženie dát.



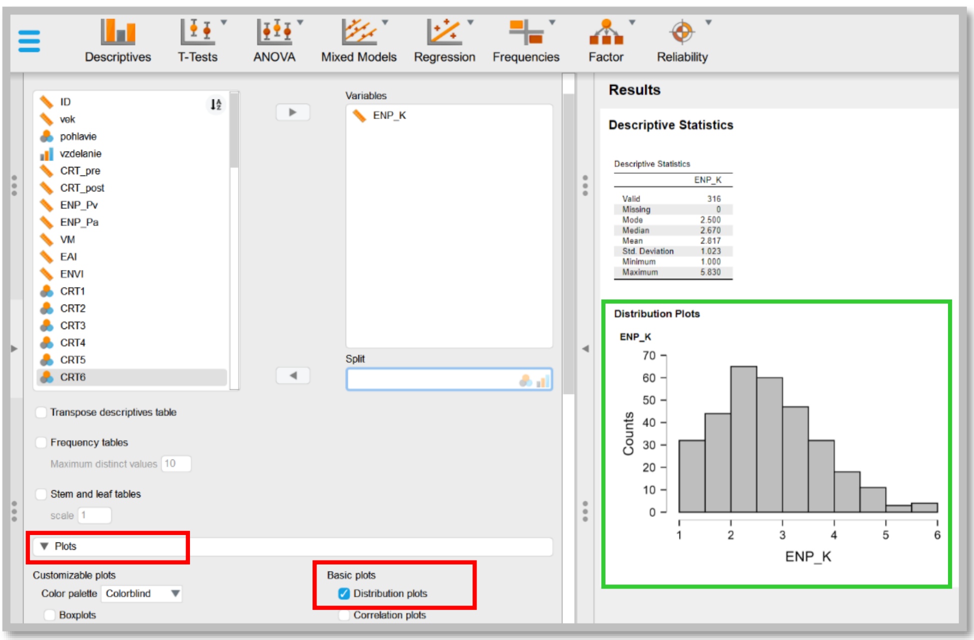
Obrázok 15 Deskriptívna štatistika – kardinálne premenné

V niektorých prípadoch je potrebné alebo užitočné poznať deskriptívne štatistiky nie len celého súboru, ale aj jednotlivých podskupín, napríklad mužov a žien, alebo kontrolnej a experimentálnych skupín. Na to slúži funkcia *Split* (Obrázok 16) – do tejto kolónky zadáme premennú, podľa ktorej bude celý súbor rozdelený do jednotlivých podskupín. Vo výsledkov časti potom uvidíme jednotlivé deskriptívne štatistiky pre každú podskupinu (v našom príklade pre mužov a pre ženy).

Spomínali sme, že **normalitu rozloženia dát** kontrolujeme cez miery polohy a variability, ale aj vizuálne prostredníctvom grafu. Histogram (stĺpcový graf) rozloženia dát vytvoríme tak, že pri výpočte deskriptívnych štatistík (*Descriptives*) pre vybrané premenné pod oknom s premenným rozbalíme možnosť *Plots* a zvolíme graf *Distribution plots*. V okne výsledkov sa objaví graf zobrazujúci distribúciu dát (Obrázok 17).



Obrázok 16 Deskriptívna štatistika – funkcia Split



Obrázok 17 Deskriptívna štatistika – graf rozloženia dát

5 Reliabilita

V rámci deskriptívnej štatistiky zvykneme reportovať aj údaje o reliabilite meracích nástrojov v zmysle ich **vnútornej konzistencie**, teda toho, nakoľko jednotlivé položky nástroja spolu súvisia, v akom tesnom sú vzťahu. Ak by chcela napríklad vyučujúca zhodnotiť vaše vedomosti z predmetu štatistika, nezadala by vám iba jednu úlohu, alebo by vám nepoložila iba jednu otázku. Takéto hodnotenie by totiž nemuselo byť presné – otázka by vám nemusela sadnúť, alebo by sa týkala témy, ktorá vám robí problémy, no neodzrkadľovala by vaše celkové vedomosti v danom predmete. Takže aj vo výskume si pripravíme viac položiek, aby sme sa vyhli chybe pri hodnotení nejakej vlastnosti, preferencie, schopnosti, zručnosti a pod. Keďže máme viac položiek, musíme vedieť, nakoľko sa skóre z týchto položiek približuje ku skutočnému skóre, teda nakoľko je použitý nástroj dobrý, presný. Pri škálovaných položkách (meraných na škále s rozpätím, napr. od 1 do 5) sa najčastejšie overuje vnútorná konzistencia pomocou koeficientu **Cronbachova alfa (α)**, ktorý je chápaný ako odhad korelácie medzi skutočným skóre a nameraným skóre. Cronbachova alfa hodnotí nakoľko položky nástroja zdieľajú spoločnú varianciu, je ukazovateľom toho, nakoľko má skutočné skóre (latentná premenná) kauzálny efekt na merané položky.

Koeficient alfa nie je dokonalým koeficientom, nie vždy je jej použitie zmysluplné a nie vždy je vhodne interpretovaný. Kritika alfy sa týka viacerých oblastí (Dunn et al., 2014), tie najhlavnejšie si zhrnieme. Hoci vo všeobecnosti platí úzus, že na to, aby sme mohli označiť nástroj za reliabilný, hodnoty α by mali byť nad hranicou 0,7 (alfa môže dosahovať hodnoty v rozmedzí 0 – 1), ide iba o „zvyk“ bez racionálneho základu. Treba mať na pamäti, že vysoká alfa nemusí byť ukazovateľom výbornej vnútornej konzistencie nástroja, ale skôr prílišnej obsahovej príbuznosti položiek. Napríklad alfa viac ako 0,9 môže byť ukazovateľom toho, že položky navzájom príliš tesne korelujú, čo naznačuje, že sú takmer identické (Miles & Banyard, 2007). Rovnako vysoká alfa nie je ukazovateľom jednodimenzionality nástroja – nezabezpečuje, že nástroj pozostáva len z jedného faktora. Ďalším slabým miestom Cronbachovej alfy je skutočnosť, že v spoločenských a humanitných vedách je väčšina nástrojov a ich položiek meraných na ordinálnej úrovni a alfa je založená na Pearsonových koreláciách (viac o koreláciách v kapitole 9 *Korelácie*), preto môže dochádzať k skresleniu. Alfa je navyše závislá od veľkosti súboru a je citlivá aj na heterogénnosť položiek.

Vzhľadom na uvedenú kritiku sa v súčasnosti odporúča na meranie vnútornej konzistencie nástroja skôr používanie koeficientu **McDonaldova omega (ω)**, ktorý na výpočet zdieľanej variancie využíva faktorovo-analytický prístup (Dunn et al., 2014).

Špecifickým prípadom, kedy nie je vhodné počítať ani alfu ani omegu sú dichotomické položky, teda také, kde skórujeme iba dve hodnoty. Napríklad test, kde môže participant získať 0 bodov za nesprávnu odpoveď a 1 bod za správnu odpoveď. V takomto prípade sa odporúča na meranie vnútornej konzistencie nástroja využiť vzorec Kuderu a Richardsona. V každom prípade platí, čím viac sa koeficient reliability približuje 1, tým lepšia je vnútorná konzistencia nástroja.

Ak sú v nástroji položky s rôznou obťažnosťou, používame *Kuderov-Richardsonov vzorec č. 20*

$$KR_{20} = \frac{\frac{K}{K-1} * (1 - \sum p * q)}{\delta^2}$$

K – počet položiek

p – podiel správnych odpovedí na položku $(\frac{počet\ správnych\ odpovedí}{počet\ všetkých\ odpovedí})$

q – počet nesprávnych odpovedí na položku $(1 - p)$

($\Sigma p*q$ – sumár $p*q$ pre všetky položky)

δ^2 – rozptyl (štandardná odchýlka na druhú)

Ak sú v nástroji položky s rovnakou obťažnosťou používame *Kuderov-Richardsonov vzorec č. 21*

$$KR_{21} = \frac{K}{K-1} * (1 - \frac{M * (K - M)}{K * \delta^2})$$

K – počet položiek

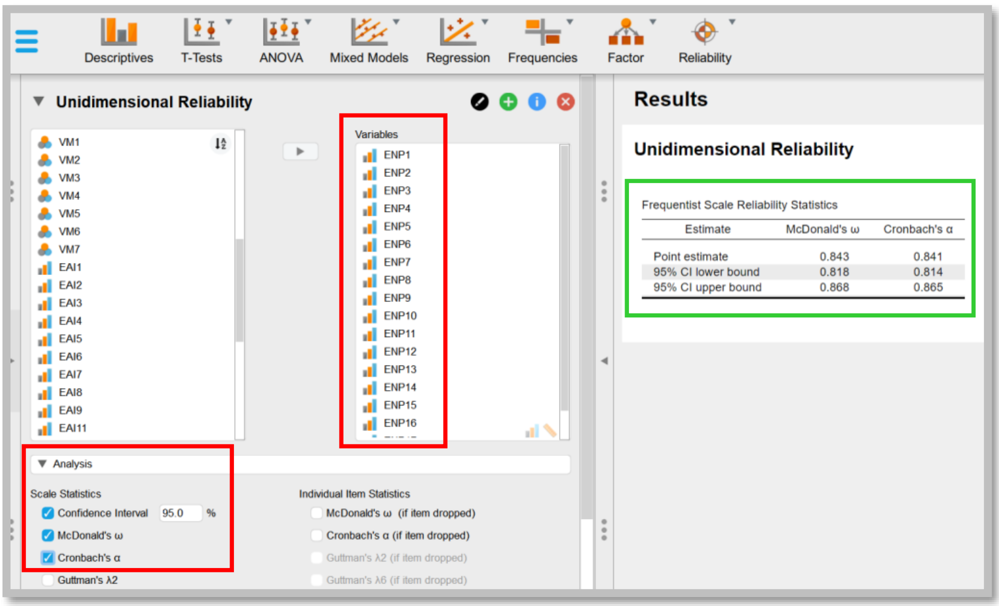
M – priemer sumárnych výsledkov

δ^2 – rozptyl (štandardná odchýlka na druhú)

Odhad reliability meracieho nástroja možno robiť viacerými spôsobmi, ako už spomínanú vnútornú konzistenciu nástroja, ale aj ako mieru stability v čase a mieru ekvivalencie. Prehľad metód na overovanie jednotlivých typov reliability nájdete v kapitole 12 *Prehľad analýz*.

5.1 Výpočet vnútornej konzistencie

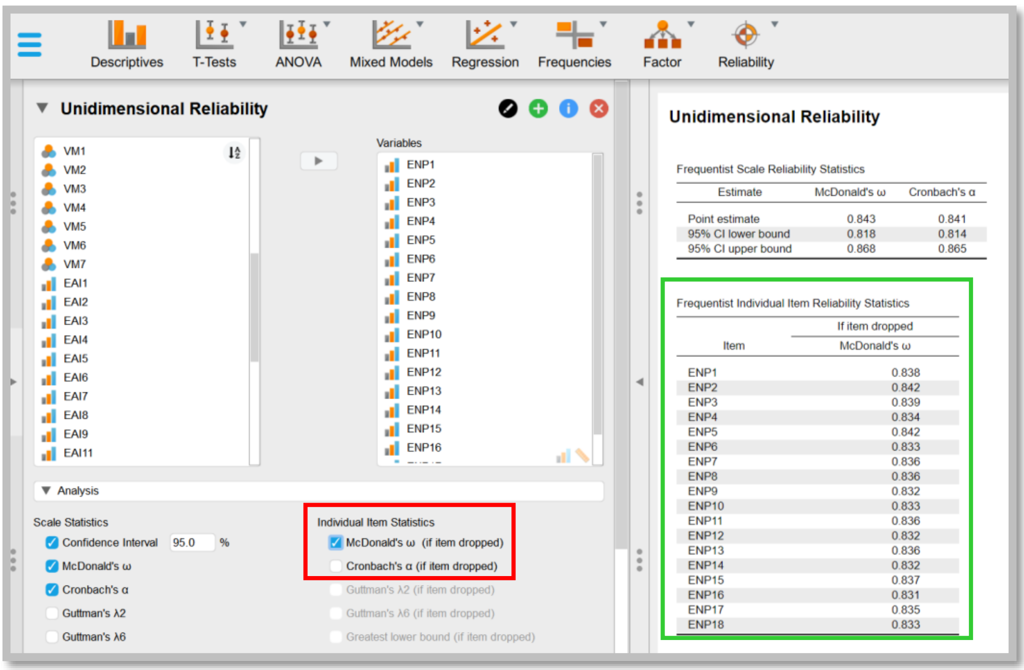
Na výpočet **vnútornej konzistencie - reliability** meracieho nástroja potrebujete využiť funkciu *Reliability*. V starších verziách JASP sa výpočet reliability nachádzal v ponuke *Descriptives*. A aktuálnej verzii je už samostatným modulom, nie je však v základnej ponuke modulov. Preto je potrebné najskôr kliknúť na ikonu + vpravo hore a zvoliť možnosť *Reliability*. Následne v rámci modulu *Reliability* zvolíme možnosť *Unidimensional Reliability* (vo verzii JASP 0.14.1 ju nájdete pod názvom *Classical – Single-Test Reliability Analysis*; Obrázok 18). Zo zoznamu premenných vyberieme a presunieme do okna *Variables* nie premenné, ale konkrétne položky nástroja, ktorý chceme overiť. V našom príklade budeme overovať reliabilitu dotazníka ENP, preto do sekcie *Variables* presunieme všetky položky dotazníka ENP (Obrázok 18). Rozbalíme možnosti *Analysis*, kde môžeme pre porovnanie zvoliť okrem McDonaldovej omegy aj výpočet Cronbachovej alfy.



Obrázok 18 Reliabilita – výpočet Cronbachovej alfy a McDonaldovej omegy

Vo výsledkoch z tabuľky odčítame hodnotu Cronbachovej alfy a McDonaldovej omegy – vidíme, že v tomto prípade je Cronbachova alfa 0,841 a McDonaldova omega 0,843. Hodnoty sú dostatočne vysoké na to, aby sme mohli konštatovať, že dotazník ENP má dobrú vnútornú konzistenciu.

V prípade, že by mal nástroj nízku hodnotu vnútornej konzistencie, je možné spraviť dodatočný krok a podrobiť položky nástroja detailnejšej analýze, aby sme zistili, či nie je v nástroji položka, ktorá je „problémová“, ktorej prítomnosť znižuje vnútornú konzistenciu celého nástroja. V takomto prípade môžeme využiť funkciu *If item dropped*. Táto funkcia odhalí položky, ktoré znižujú vnútornú konzistenciu a po vylúčení alebo preformulovaní (v budúcom výskume) ktorých koeficient omega, resp. alfa, stúpne (na Obrázku 19 je ukážka pre dotazník ENP). V možnostiach *Analysis* v časti *Individual Item Statistics* zvolíme *McDonald's ω (if item dropped)*. Vo výsledkoch vidíme, že McDonaldova omega pre ENP je ω = 0,843. V druhej tabuľke sú zobrazené jednotlivé položky a hodnoty, na ktoré by omega stúpla, ak by sme tú – ktorú položku vylúčili. V tomto prípade by reliabilita nestúpla nad 0,843 ani v jednom prípade, takže nie je potrebné žiadnu položku vylúčiť ani preformulovať.



Obrázok 19 Reliabilita – funkcia *If item dropped*

6 Inferenčná štatistika – úvod do testovania hypotéz

Deskriptívna štatistika slúži na opísanie dát, s ktorými pracujeme. No vo výskume potrebujeme väčšinou spraviť krok ďalej a položiť si aj zložitejšie otázky. Výskumy sa realizujú, aby sme overili, otestovali teoretické princípy, aby sme hľadali odpovede na výskumné otázky, aby sme overovali vyslovené hypotézy. V takýchto prípadoch sa opierame o inferenčnú štatistiku, ktorá ide nad rámec deskriptívnej štatistiky a slúži na vytváranie inteligentných odhadov (Miles & Banyard, 2007). Inferenčná štatistika slúži na to, aby sme robili závery týkajúce veľkých súborov (populácií) na základe zistení získaných z malých súborov (výskumných vzoriek; Hanna & Dempster, 2012).

Ak zistíme, že za rôznych podmienok sa dáta správajú odlišne, môže to mať dve príčiny. Buď ide o to, že rôzne podmienky majú vplyv (efekt) na sledované premenné (napríklad, žiaci, ktorí sa prežívajú zvýšenú úzkosť, podávajú slabšie výkony), alebo ide iba o náhodu, náhodnú chybu, ktorá nereprezentuje reálne vzťahy medzi premennými. Aby sme vedeli odlíšiť, kedy sa jedná o efekt a kedy o náhodnú chybu, využívame práve inferenčnú štatistiku. A keďže inferenčná štatistika stavia na pravdepodobnosti, v krátkosti si pravdepodobnosť predstavíme.

6.1 Pravdepodobnosť

Na prvý pohľad sa pravdepodobnosť javí ako niečo jasné a samozrejmé. V bežnom živote hovoríme o pravdepodobnosti toho, či bude pršať, že sa pár žijúci v susedstve rozíde, alebo že na kocke hodíme číslo 5. Napriek tomu mnoho ľudí s pravdepodobnosťou nepracuje správne. Príčinou je skutočnosť, že pravdepodobnosť je kontra-intuitívna, ide proti spôsobu, akým bežne premýšľame. Keď hráme *Človeče, nehnevaj sa*, a niekoľko kôl za sebou hodíme nízke čísla, hovoríme si, že už predsa musí padnúť šestka, lebo už dlho nebola. Toto je naše intuitívne rozmýšľanie. V skutočnosti pravdepodobnosť, že hodíme šestku je pri každom jednom hode rovnaká. Kocka totiž „neberie do úvahy"—na rozdiel od nás—to, aké čísla sme hodili v predchádzajúcich kolách.

V rámci štatistiky môže pravdepodobnosť dosahovať hodnoty od 0 po 1 (v bežnom živote často hovoríme, že je 10% pravdepodobnosť prehánok, alebo že je šanca 50 : 50, že padne znak keď hodíme mincou). Udalosť, ktorá má pravdepodobnosť 0 je udalosť, ktorá sa určite nestane, za žiadnych okolností.

Na druhej strane, udalosť, ktorá má pravdepodobnosť 1 je udalosť, ktorá určite nastane, bez ohľadu na okolnosti. Vypočítať pravdepodobnosť udalosti si vyžaduje, aby sme poznali počet želaných výsledkov (napríklad k výhre potrebujem, aby som na kocke hodila číslo 6, teda počet želaných udalostí je 1) a počet všetkých možných výsledkov (v tej istej hre je počet možných výsledkov 6, pretože kocka má 6 strán a môže padnúť ktorákoľvek z nich). Pravdepodobnosť, že vyhrám hru tým, že na kocke hodím číslo 6, je $1 : 6$, teda $0{,}17$. Keby v tej istej hre potrebujem k výhre hodiť číslo 4 alebo 6 (počet želaných udalostí je 2, lebo k výhre mi pomôže tak 4 ako aj 6), tak pravdepodobnosť výhry je $2 : 6$, teda $0{,}33$. Hodnoty pravdepodobnosti si vieme premeniť aj na % jednoduchým vynásobením pravdepodobnosti hodnotou 100 – takže pravdepodobnosť, že hodím 6 je $0{,}17 * 100 = 17\%$, resp. že hodím buď 4 alebo 6 je $0{,}33 * 100 = 33\%$.

Okrem takejto jednoduchej pravdepodobnosti sa v živote, štatistiku nevynímajúc, stretneme aj s iným typom pravdepodobnosti, s *podmienenou pravdepodobnosťou*. Ide o pravdepodobnosť udalosti, ktorej výskyt závisí od inej udalosti. Napríklad, u fajčiarov je 5-násobne vyššia pravdepodobnosť, že dostanú infarkt. To znamená, výskyt udalosti „infarkt" závisí od udalosti „fajčenie" (ale samozrejme aj ďalších, ako genetika či životný štýl).

A ako teda reálne využívame pravdepodobnosti vo výskume a v štatistike? Z metodológie vieme, že bežne výskumy nerealizujeme na celých populáciách, ale iba na ich reprezentatívnych častiach – na výskumných súboroch. Napriek všetkým výhodám, ktoré pozorovanie výskumných súborov prináša, musíme si byť vedomí, že to, čo odpozorujeme na výskumnom súbore, nemusí vždy reflektovať realitu, teda to, čo by sme pozorovali, keď by sme sledovali celú populáciu. Ak je rozdiel medzi tým, ako sa správa výskumný súbor a ako populácia, hovoríme o *výberovej chybe* (Dancey & Reidy, 2017). Ideálna situácia pre výskumníka je, ak k takejto výberovej chybe nedochádza a vzorka reprezentuje celú populáciu. To je však vzácny jav. Preto ako výskumníci potrebujeme poznať (vedieť vypočítať), aká je pravdepodobnosť, že to, čo nameriame, je v zhode s realitou, respektíve aká je pravdepodobnosť, že keď sa budeme riadiť našimi zisteniami, že sa mýlime (a ako veľmi sa mýlime)? A tu prichádza k slovu spomínaná pravdepodobnosť, a doteraz nespomínané testovanie hypotéz.

6.2 Testovanie hypotéz

V rámci inferenčnej štatistiky je našou snahou robiť závery z dát, z toho, čo pozorujeme. Vo výskume sa snažíme porozumieť vzťahom a príčinám. Na to využívame pozorovanie výberových súborov z populácie, a ako už vieme, keď pracujeme so vzorkou, je možné, že sa dopustíme výberovej chyby a následne nevieme, či zistenia z našej vzorky reflektujú to, čo sa deje aj v populácii alebo sú iba výsledkom chyby. Pravdepodobnosť nám na tomto mieste môže veľmi dobre poslúžiť. Vďaka tomu, že dokážeme vypočítať pravdepodobnosť chyby, vieme sa rozhodnúť, či sú naše zistenia skôr v súlade s realitou v populácii alebo sú skôr výsledkom chyby. Ak je šanca, že zistenia sú výsledkom náhody malá, potom si môžeme dovoliť vysloviť záver, že naša vzorka s vysokou pravdepodobnosťou reflektuje populáciu. Pri testovaní hypotéz hľadáme dôkazy o tom, či sú sledované vzťahy štatisticky a vecne významné.

6.2.1 Štatistická významnosť

Do výskumu vstupujeme s nejakým výskumným problémom, máme stanovený výskumný cieľ, na základe ktorého vyslovujeme hypotézy o populácii. Tieto hypotézy následne testujeme, a na základe výsledkov hypotézy podporujeme alebo zamietame[11]. Vo všeobecnosti je **hypotéza** predpoklad, tvrdenie, pravdivosť ktorého overujeme výskumom, pričom testovanie hypotéz (a z toho vyplývajúce vyvodzovanie záverov) má vždy iba pravdepodobnostný charakter. Testovanie hypotéz je systematický proces, ktorý vedie k rozhodnutiu, či výsledky výskumu podporujú hypotézu platnú pre populáciu. V každom výskume si musíme v prvom kroku stanoviť nulovú a alternatívnu hypotézu, hoci sa za bežných okolností v závere výskumu vyjadrujeme iba k alternatívnej hypotéze.

Nulová hypotéza (H_0) je tvrdenie o tom, že skúmaný efekt v populácii neexistuje, že sledované populácie (skupiny) sú identické alebo že priemery populácií (skupín) sú rovnaké alebo že korelácia je nulová. Nulová hypotéza nie je predpoklad o tom, čo *budeme* pozorovať, je to tvrdenie o tom, aká *je* populácia (vzhľadom na skúmané javy), z ktorej vyberáme výskumnú vzorku (Coolican, 2019). Ak je nízka

[11] Keď sa chceme dopracovať k poznaniu, jeden výskum nie je postačujúci, jedným výskumom nie je možné získať jednoznačné odpovede na výskumný problém, otázku. V praxi sa preto realizuje séria výskumov, ktoré nie len že prinášajú odpovede, ale často generujú nové otázky a hypotézy. Tie sa následne testujú v nových výskumoch. Takouto systematickou prácou sa možno dopracovať k množstvu informácií – dôkazov, na ktorých možno stavať závery (Hanna & Dempster, 2012).

pravdepodobnosť toho, že je nulová hypotéza platná, potom túto nulovú hypotézu zamietame v prospech *alternatívnej hypotézy*. **Alternatívna hypotéza** je predpoklad o tom, že skúmaný efekt *budeme* pozorovať v rámci výskumného súboru (že budú rozdiely medzi skupinami, priemery populácií nebudú rovnaké, korelácia nebude nulová; Coolican, 2019). Nulová hypotéza nevyjadruje to, čo očakávate, že zistíte; vaše očakávania budú vyjadrené alternatívnou hypotézou. Napriek tomu, je to práve nulová hypotéza, ktorú v rámci inferenčnej štatistiky testujeme (Miles & Banyard, 2007).

Pri testovaní rozdielov môže mať alternatívna hypotéza dve podoby: jednosmerná (*one-tailed*) alebo dvojsmerná (*two-tailed*). *Jednosmerná hypotéza* vyjadruje, špecifikuje smer efektu, napríklad že jedna skupina dosiahne vyššie skóre ako druhá skupina, alebo že vzťah medzi premennými bude pozitívny. Nulová hypotéza k jednosmernej alternatívnej hypotéze by tvrdila, že rozdiel medzi skupinami neexistuje alebo je v opačnom smere, ako predpokladáme v alternatívnej hypotéze, resp. vzťah medzi premennými neexistuje alebo má opačný smer ako predpokladáme v alternatívnej hypotéze. Napríklad, ak by sme predpokladali, že žiaci 3.A dosiahnu *lepšie* známky z matematického testu ako žiaci 3.B, nulová hypotéza by znela, že medzi žiakmi 3.A a 3.B nie sú rozdiely v známkach z matematického testu alebo žiaci 3.A majú horšie známky ako žiaci 3.B. Iný príklad. Ak by sme predpokladali, že medzi známkami z matematiky a z fyziky existuje pozitívny vzťah, nulová hypotéza by znela, že neexistuje vzťah medzi známkami z matematiky a z fyziky alebo je tento vzťah negatívny.

Niekedy ale vo výskume predpokladáme efekt jednej premennej na druhú, ale nevieme určiť aký bude mať smer. V tom prípade formulujeme *dvojsmernú hypotézu*, v ktorej predpokladáme rozdielne skóre medzi dvoma skupinami/meraniami alebo predpokladáme vzťah medzi premennými, a to bez toho aby sme špecifikovali smer vzťahu. V takomto prípade nulová hypotéza znie: rozdiel medzi skupinami/meraniami neexistuje, resp. vzťah medzi premennými neexistuje.

V štatistike sa musíme zmieriť so skutočnosťou, že výsledky inferenčnej štatistiky nám nepovedia, ako sa veci reálne majú, povedia nám „iba“, **aká je pravdepodobnosť, že by sme pozorovali (namerali) to, čo sme pozorovali v našom výbere za podmienky, že je nulová hypotéza pravdivá**. Inými slovami, netestujeme pravdepodobnosť, že je nulová hypotéza pravdivá, ale pravdepodobnosť toho, že by pozorovaný jav opakovane nastal v prípade, že by bola nulová hypotéza pravdivá. V psychológii a v príbuzných disciplínach je ustálená hraničná hodnota pravdepodobnosti na **0,05** (na 5%), označujeme ju aj *hladina významnosti alfa* (α).

Hladinu významnosti si stanovujeme vopred, aj keď väčšinou sa držíme konvencie a ponechávame ju na zaužívanej úrovni 0,05[12].

Pomocou štatistických testov počítame hodnotu *štatistickej signifikancie **p***: ak je pravdepodobnosť nižšia ako uvedených 0,05, hovoríme, že výsledok je *štatisticky významný* (signifikantný),. Znamená to, že existuje pravdepodobnosť menšia ako 5%, že udalosť nastane za podmienky, že je nulová hypotéza pravdivá (Coolican, 2019; Dancey & Reidy, 2017; Foster et al., 2018; Miles & Banyard, 2007).

Predstavme si bazén naplnený guľôčkami modrej a zelenej farby, pričom guľôčky majú rôznu váhu. V našej nulovej hypotéze *tvrdíme*, že medzi modrými a zelenými guľôčkami *nie sú* významné rozdiely vo váhe. Predpokladáme však (alternatívna hypotéza), že sa modré a zelené guľôčky v našom výbere *budú* odlišovať čo sa týka ich váhy. Spravíme výskum. Z bazéna náhodne vytiahneme 100 guľôčok (snažíme sa, aby sme mali rovnaký počet zelených a modrých) a poznačíme si farbu a váhu každej z nich. Zistíme, že priemerná váha modrých guľôčok je vyššia ako priemerná váha zelených guľôčok, a že tento rozdiel je významný na hladine významnosti $p = 0{,}030$. Späť k hypotézam: ak platí nulová hypotéza, a teda že guľôčky v bazéne majú (približne) rovnakú váhu, tak pravdepodobnosť toho, že by sme namerali rozdielne váhy modrých a zelených guľôčok je iba 3%, teda veľmi nízka. Uvedený výsledok preto nie je o tom, že sme mali „šťastie“ a namerali sme rozdiely, ale smeruje skôr k tomu, že v bazéne sú naozaj guľôčky rôznej váhy a že váha súvisí s farbou guľôčky (a že nulová hypotéza neplatí). Zamietame nulovú hypotézu v prospech alternatívnej hypotézy.

Ukážme si to ešte na inom príklade. Predstavte si, že učíte matematiku a chcete zistiť, či má pohlavie vplyv na úspešnosť v matematických úlohách. Otestujete si teda žiakov niekoľkými matematickými úlohami. V tomto prípade by sme v nulovej hypotéze tvrdili, že výkon chlapcov a dievčat v matematických úlohách je vyrovnaný (nie sú žiadne rozdiely medzi chlapcami a dievčatami). Alternatívna hypotéza by znela nasledovne: predpokladáme, že budú rozdiely medzi chlapcami a dievčatami vo výkone v matematických úlohách (hypotézu by sme samozrejme opierali o relevantné zdroje). Zistili by sme, že namerané rozdiely sú štatisticky významné, pretože $p = 0{,}020$. To znamená, že pravdepodobnosť toho, že by sme namerali rozdiely medzi chlapcami a dievčatami—za predpokladu, že v populácii tieto rozdiely neexistujú (ak platí nulová hypotéza)—je iba 2%. Ide o nízku pravdepodobnosť, a záver znie, že je pravdepodobné, že aj v populácii tieto rozdiely reálne sú. Zamietame nulovú hypotézu

[12] V niektorých prípadoch je možné stanoviť si hranicu štatistickej významnosti aj na úroveň 0,01.

(nie sú rozdiely medzi chlapcami a dievčatami) a našimi zisteniami podporujeme alternatívnu hypotézu (predpoklad o rozdieloch medzi chlapcami a dievčatami).

6.2.2 Chyby typu I a II

Ako ste už zrejme postrehli, v psychológii a v príbuzných vedách zbierame dôkazy, ktoré sa týkajú predpokladov, hypotéz, no napriek tomu nemáme 100% istotu, že naše zistenia sú naozaj pravdivé. Preto si stanovuje hladinu významnosti, aby sme si nastavili mieru, do akej si dovolíme mýliť sa – pri hladine významnosti 0,05 je to maximálne 1 z 20 prípadov, kedy zamietneme nulovú hypotézu keď tá je v skutočnosti pravdivá. Stále však treba mať na pamäti, že aj keď dosiahneme významný výsledok, môže sa jednať iba o výsledok náhody. V takomto prípade hovoríme o **chybe typu I**, chybe, že odmietneme nulovú hypotézu, keď je pravdivá (Coolican, 2019). Stanovením hladiny významnosti alfa robíme rozhodnutie o tom, aká je najvyššia akceptovaná pravdepodobnosť, že sa dopustíme chyby typu I. Aby sme ako výskumníci eliminovali chybu typu I, jednotlivé výskumy replikujeme, opakujeme. Ak získame významný výsledok v jednom výskume, môže ísť o náhodu, ak podobný výsledok získame v dvoch, troch, desiatich výskumoch, šanca, že ide o výsledok náhody sa významne znižuje (eliminujeme tak chybu typu I).

Iným spôsobom, ako môžeme redukovať pravdepodobnosť chyby typu I je zníženie hladiny významnosti (napr. na úroveň 0,01). Tým sa však vystavujeme riziku, že sa dopustíme inej chyby, **chyby typu II**. Chyba typu II nastáva vtedy, keď sa držíme nulovej hypotézy, keď tá v skutočnosti nie je pravdivá, resp. keď je alternatívna hypotéza pravdivá, ale my sme získali výsledok, ktorý naznačuje, že je skôr pravdivá nulová hypotéza. Chyby typu II sa môže dopustiť iba v prípade, že je nulová hypotéza v skutočnosti nepravdivá (chybu typu II označujeme aj ako β; Foster et al., 2018). Tradične sa akceptuje pravdepodobnosť chyby typu II nie viac ako 0,20 (20%). Aby sme redukovali pravdepodobnosť chyby typu II, môžeme zväčšiť veľkosť súboru (ak napríklad v skutočnosti existuje rozdiel medzi dvoma skupinami, potom čím väčšie vzorky porovnáme, tým je pravdepodobnejšie, že dosiahneme významný výsledok). Sumár možných výsledkov pri testovaní hypotéz je zobrazený v Tabuľke 5. To, že sa dopustíme chyby typu I alebo II však v konečnom dôsledku nevypovedá o tom, že by sme spravili niečo nesprávne, tieto chyby sú proste nevyhnutnou súčasťou štatistických analýz (Hanna & Dempster, 2012), no je potrebné mať ich na zreteli najmä pri vyslovovaní záverov a interpretácii zistení.

Tabuľka 5 Možné správne a nesprávne rozhodnutia v testovaní hypotéz

		nulová hypotéza je v skutočnosti:	
		pravdivá	*nepravdivá*
záver z testovania hypotéz:	*alternatívna hypotéza je podporená*	**chyba typu I (α)**	správne rozhodnutie
	nesmieme zamietnuť nulovú hypotézu	správne rozhodnutie	**chyba typu I (β)**

6.2.3 Veľkosť efektu a sila

Niekedy sa môže stať, že výsledky signifikancie tesne atakujú hranicu 0,05, inokedy sú omnoho nižšie ako 0,01, to však neznamená, že môžeme o jednom výsledku povedať, že je významnejší ako druhý. Hodnota významnosti totiž závisí od viacerých faktorov: od typu testu, od veľkosti vzorky, aj od veľkosti efektu, ktorý skúmame. Predstavte si, že máme dve skupiny študentov, A a B, v každej je po 100 študentov. Obom skupinám predstavíme nového učiteľa, pričom skupine B navyše povieme, že sa nový učiteľ popri práci venuje charitatívnej činnosti v zariadení pre ľudí bez domova. Následne študentov požiadame, aby zhodnotili, do akej miery je nový učiteľ sympatický na škále od 1 (vôbec) do 10 (veľmi). Zistíme, že v skupine A bola priemerná hodnota sympatie učiteľa 6,8 a v skupine B 7,6. Po testovaní hypotézy sme zistili, že môžeme zamietnuť nulovú hypotézu, teda že sú naše výsledky „signifikantné“. Teraz si predstavme, že priemerné hodnotenie v skupine B by bolo 7,2, a že testovanie hypotéz by nám opäť ukázalo, že môžeme zamietnuť nulovú hypotézu, že rozdiely sú „významné“. Čiže vidíme, že štatistická signifikancia nám ukazuje, že existuje efekt, no nehovorí nám nič o jeho veľkosti.

Veľkosť efektu (*effect size*) je miera, ktorá poukazuje na skutočný efekt v populácii, na skutočný rozdiel medzi priemermi dvoch skupín (ako veľmi sa niečo zmení za špecifických podmienok). Poznať veľkosť efektu je potrebné najmä vo výskumoch, z ktorých chceme usudzovať na praktické dôsledky a odporúčania (Aron et al., 2014). Napríklad, ak by sme sledovali efekt intervenčného programu zameraného na zvyšovanie empatie, pozorovali by sme dve skupiny, kde prvá skupina by absolvovala program (experimentálna skupina) a druhá nie. V oboch skupinách by sme v prvom kroku zmerali úroveň empatie a zistili by sme, že priemerná úroveň všetkých zúčastnených je pod priemerom populácie (o to viac by sme vnímali zrealizovanie intervenčného programu ako dôležité). Zrealizovali by sme experiment (teda prvú skupinu by sme previedli intervenčným programom, druhú by sme ponechali bez zásahu) a znovu by sme u všetkých zmerali úroveň empatie.

Testovaním by sme zistili, že došlo k posunu v prvej skupine v zmysle zvýšenia úrovne empatie, a že tento posun je štatisticky významný. No pri bližšom pohľade na výsledky by sme zistili, že úroveň empatie je stále pod priemerom v populácii, a to aj v experimentálnej skupine. Čiže hoci by boli výsledky štatisticky významné, reálna veľkosť efektu (hovoríme o tzv. *vecnej významnosti*) by bola malá – napriek posunu v experimentálnej skupine by si boli obe skupiny účastníkov reálne stále veľmi podobné (v nízkej úrovni empatie).

Veľkosť efektu si vieme na hrubo vypočítať ako rozdiel medzi priemerom jednej a druhej skupiny. Takto zistený rozdiel ale neumožňuje porovnať naše zistenia s inými výskumami (či už našimi alebo iných výskumníkov). Preto je potrebné mať výsledky v podobe, ktorá je porovnateľná. Pre posúdenie veľkosti efektu sa využívajú rôzne koeficienty, podľa toho, aké štatistické analýzy využívame:

- pri porovnávaní dvoch skupín sa najčastejšie využíva koeficient *Cohenovo d*, prípadne Hedgevo *g* alebo Glassova *delta*,
- pri overovaní vzťahov prostredníctvom korelácií sa využíva *koeficient determinácie* r^2; pri regresných analýzach využívame zas *mnohonásobný koeficient determinácie* R^2,
- pri porovnávaní viacerých skupín (pri analýze rozptylu) sa využíva najčastejšie *eta* η^2 alebo *omega* ω^2,
- pri skúmaní nominálnych premenných sa využíva koeficient *phi* – Φ, *Cramerovo V* alebo *kontingenčný koeficient*.

Viac o jednotlivých koeficientoch veľkosti efektu sa dočítate pri jednotlivých analýzach v nasledujúcich kapitolách. Porovnanie jednotlivých koeficientov vzhľadom na veľkosť efektu nájdete v kapitole 12 *Prehľad analýz.*

S veľkosťou efektu veľmi úzko súvisí **štatistická sila**. Je to pravdepodobnosť, že sa dopracujeme k signifikantným výsledkom, ak je alternatívna hypotéza pravdivá (Aron et al., 2014). Pozor, dôležitá je podmienka – ak je alternatívna hypotéza pravdivá (v opačnom prípade by sa jednalo o chybu typu I; Coolican, 2019). V praxi štatistická sila vyjadruje pravdepodobnosť, že so vzorkou, ktorú mám, dosiahnem signifikantný výsledok. Vo všeobecnosti je aktuálne vnímaná ako prijateľná sila aspoň na úrovni 80%.

Nesmieme zabúdať na fakt, že aj keď je alternatívna hypotéza pravdivá, nie vždy musíme získať významné výsledky, napríklad keď náš výber participantov z nejakého dôvodu nereprezentuje populáciu. **Veľkosť výskumného súboru** je preto ďalší faktor, ktorý ovplyvňuje štatistickú silu výskumu. Takže tu máme veľkosť efektu a veľkosť vzorky, ktoré významne ovplyvňujú štatistickú silu. V spomínanom príklade s aplikovaním intervencie na zvýšenie empatie sme mali dve skupiny po 100 ľudí. Pri očakávanej malej veľkosti efektu (napr. 0,25) by bola sila 94%, to znamená, že by bola pravdepodobnosť 94%, že ľudia zo skupiny B by mali po intervencii štatisticky vyššie skóre empatie ako ľudia zo skupiny A. Odhadovanú silu si vieme vypočítať pomocou rôznych online kalkulačiek a simulátorov[13], rovnako ako aj silu efektu, či veľkosť vzorky. Tieto kalkulačky sú užitočné. Keď vopred viete, akú veľkú veľkosť efektu a sily by ste chceli dosiahnuť, viete vypočítať, akú veľkú vzorku budete vo výskume potrebovať.

[13] Napr. https://rpsychologist.com/d3/nhst/

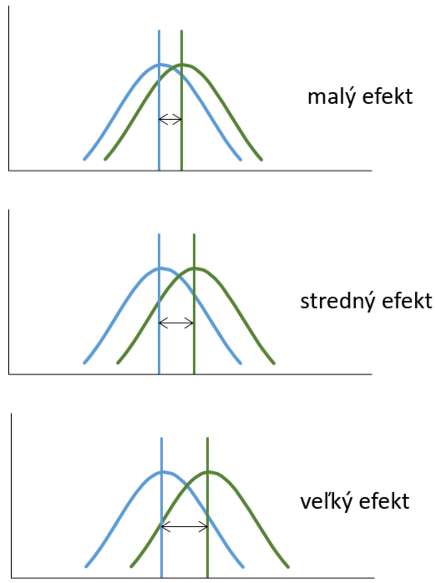
7 Testovanie rozdielov (testovanie pomocou t-testov)

Pri testovaní hypotéz sa veľmi často stretnete so situáciou, kedy budete potrebovať testovať rozdiely medzi dvoma súbormi dát. Buď sú to dva súbory dát od jednej výskumnej vzorky za dvoch rôznych podmienok – hovoríme o *závislých výberoch* (napr. meranie pred a po intervencii v jednej skupine osôb), alebo sú to dva súbory dát od dvoch rôznych výskumných vzoriek – hovoríme o dvoch *nezávislých výberoch* (napr. meranie v dvoch skupinách, ktoré chceme porovnať). V špecifických prípadoch máme k dispozícii iba jeden súbor dát, ktorý chceme porovnať s nejakou známou hodnotou – hovoríme o *jednom výbere*.

Pri výbere vhodného štatistického testu máme k dispozícii dva základné balíky testov. Ak máme primerane veľkú vzorku a naše dáta sú distribuované približne normálne, využívame takzvané **parametrické testy**. V prípade, že máme malú vzorku a/alebo dáta nezodpovedajú normálnemu rozloženiu, využívame tzv. **neparametrické testy**. V nasledovných kapitolách si prestavíme základné štatistické parametrické testy a aj ich neparametrické alternatívy.

Na zhodnotenie rozdielov medzi skupinami za použitia parametrických testov sa využívajú tzv. **t-testy**, ktorých výsledkom je hodnota ***t***. *Štatistickú významnosť* výsledku t-testu (hodnoty *t*) posudzujeme podľa štatistickej významnosti ***p***, teda podľa pravdepodobnosti, že získame hodnotu testovacej štatistiky, ktorá je väčšia alebo rovná ako skutočne získaná hodnota za predpokladu, že je nulová hypotéza pravdivá. Na zhodnotenie *sily efektu* sa pri porovnávaní skupín za použitia parametrických testov najčastejšie využíva koeficient *Cohenovo* ***d*** (Aron et al., 2014; Cohen, 1988), ktorý môže nadobúdať akúkoľvek hodnotu. Cohenovo *d* vyjadruje, do akej miery sa sledované skupiny ľudí prekrývajú (zhodujú) v ich skóre. Pre zhodnotenie toho, čo je malý, stredný alebo veľký efekt sa opierame o konvenciu (Obrázok 20):

- malá sila efektu: *d* do 0,2 – 0,5, prekryv skupín približne 85% (väčšina osôb jednej skupiny sa zhoduje s väčšinou ľudí druhej skupiny),
- stredná sila efektu: *d* medzi 0,5 – 0,8, prekryv skupín približne 67%,
- veľká sila efektu: *d* nad 0,8. prekryv skupín približne 53%.



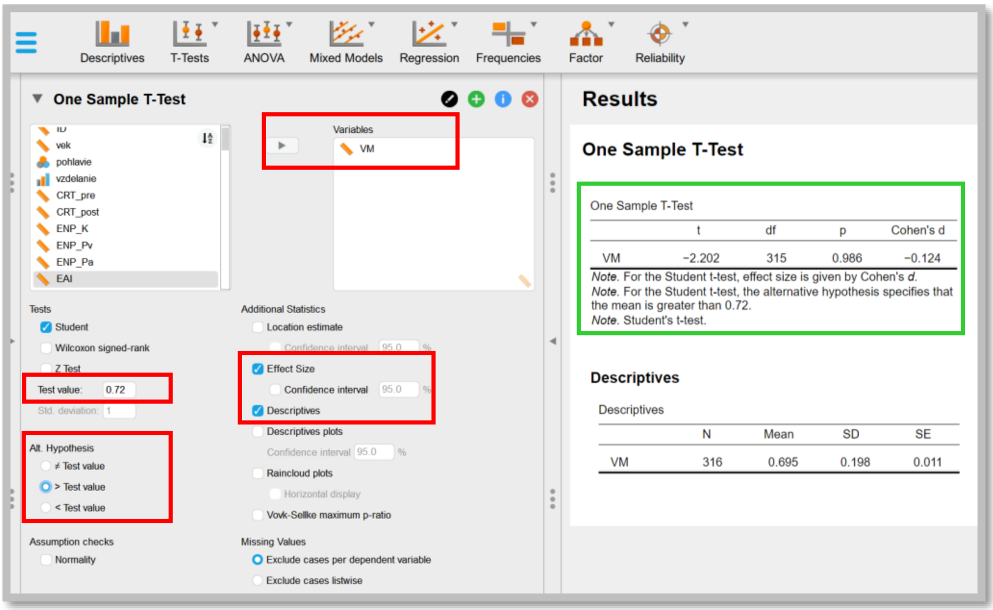
Obrázok 20 Veľkosť efektu – Cohenovo *d*

7.1 T-test pre jeden výber(one sample t-test)

Ak máme dáta od jednej vzorky a chceme túto vzorku porovnať v sledovanej premennej s nejakou známou konštantou, použijeme **t-test pre jeden výber** (*one sample t-test*). Napríklad, máme skupinu študentov – budúcich učiteľov, u ktorých sledujeme úroveň vedeckého myslenia a chceme porovnať úroveň ich vedeckého myslenia s úrovňou vedeckého myslenia v populácii. Informácie o vzorke máme z merania a údaje o populácii sme získali z článku o vedeckom myslení Slovákov. Takže v tomto prípade, keď máme jednu vzorku a poznáme konštantu (priemerné skóre populácie). Predpokladáme, že úroveň vedeckého myslenia študentov bude na vyššej úrovni ako je priemer populácie (alternatívna jednosmerná hypotéza). Aplikujeme pre porovnanie *t-test pre jeden výber*, ktorým testujeme nulovú hypotézu: priemer výberového súboru je zhodný s priemerom populácie (známeho výberového súboru).

7.1.1 Výpočet t-testu pre jeden výber

Výpočet *t-testu pre jeden výber* si ukážeme na príklade s vedeckým myslením. Okrem samotných dát musíme disponovať aj informáciou o výške priemerného skóre z testu vedeckého myslenia populácie. Dajme tomu, že úroveň vedeckého myslenia v populácii je 0,72 (z maximálnej hodnoty 1, pričom 1 znamená 100% úspešnosť v testoch vedeckého myslenia). V programe JASP v sekcii *T-tests* vyberieme možnosť *One Sample T-Test* (Obrázok 21). Zobrazí sa okno na výpočet *t-testu pre jeden výber*. Do okna *Variables* navolíme premennú, ktorú sledujeme (v našom prípade vedecké myslenie *VM*) a v ktorej chceme porovnávať vzorku so známou konštantou. V sekcii *Tests* do okna *Test value* vpíšeme známu konštantu, s ktorou chceme porovnávať súbor, teda hodnotu 0.72 (pozor, v programe sa namiesto desatinnej čiarky píše desatinná bodka). Keďže sme si formulovali jednosmernú hypotézu, v sekcii *Alt. Hypothesis* zvolíme možnosť > *Test value*, ktorá vyjadruje náš predpoklad, že úroveň vedeckého myslenia študentov bude na *vyššej* úrovni ako je priemer populácie (v prípade dvojsmernej hypotézy by sme ponechali prednastavenú možnosť ≠ *Test value*). Ďalej, v sekcii *Additional Statistics* zaškrtneme možnosti *Effect size* pre výpočet Cohenovho *d* a *Descriptives* pre výpočet deskriptívnych štatistík (Obrázok 21).



One Sample T-Test

	t	df	p	Cohen's d
VM	−2.202	315	0.986	−0.124

Note. For the Student t-test, effect size is given by Cohen's *d*.
Note. For the Student t-test, the alternative hypothesis specifies that the mean is greater than 0.72.
Note. Student's t-test.

Descriptives

	N	Mean	SD	SE
VM	316	0.695	0.198	0.011

Obrázok 21 Jednovzorkový t-test (*One Sample T-Test*)

Vo výsledkoch vidíme v prvej tabuľke výsledky t-testu: hodnotu *t*, stupne voľnosti *df*[14], hodnotu pravdepodobnosti *p* (štatistická významnosť) a veľkosť efektu (*Cohen's d*). V druhej tabuľke vidíme deskriptívne štatistiky súboru. Priemerné skóre vedeckého myslenia *VM* je 0,70 bodov[15] a priemerné skóre porovnávaného súboru je 0,72, priemerný rozdiel medzi priemerom vzorky a zadanou známou konštantou je -0,02. Výsledok t-testu ukazuje, že tento rozdiel nie je štatisticky významný (p = 0,986): pravdepodobnosť toho, že by sme nenamerali rozdiely medzi populáciou a našou vzorkou za predpokladu, že v populácii reálne tieto rozdiely ani neexistujú (teda že platí nulová hypotéza), je až 98,6% (p = 0,986). Môžeme konštatovať, že nami sledovaný súbor nemá štatisticky významne vyššie skóre vedeckého myslenia (M = 0,70; SD = 0,20) ako je priemer populácie (M = 0,72).

Pri reportovaní výsledkov uvádzame hodnoty t-testu (*t*), stupne voľnosti (*df*), hodnotu signifikancie (*p*) a hodnotu Cohenovho *d* v podobe (*t*(hodnota *df*) = hodnota t-testu; *p* = hodnota signifikancie; *d* = hodnota Cohenovho *d*).

Príklad:

Jednovzorkový t-test nepreukázal signifikantný rozdiel vo vedeckom myslení v porovnaní s priemerom populácie ($t(315)$ = 2,202; p = 0,986; d = 0,124), participanti dosahovali približne rovnako vysoké skóre (M = 0,70; SD = 0,20) ako je priemer v populácii (M = 0,72).

7.2 Studentov t-test pre dva nezávislé výbery

Predstavte si, že vo svojom výskume sledujete výkon participantov a chcete overiť rozdiely napríklad medzi mužmi a ženami. Ide o prípad, keby potrebujete porovnať priemery dvoch súborov (skupín, vzoriek; medziskupinový dizajn). Podmienkou je, že tieto súbory musia byť navzájom nezávislé, to znamená, že každý člen môže byť súčasťou iba jednej skupiny. Hovoríme o nezávislých súboroch.

[14] Stupne voľnosti označované aj *df* (*degree of freedom*) počítame ako N-1 (teda pri vzorke *N* = 316, je *df* = 315). Stupne voľnosti vyjadrujú počet skóre, ktoré sú „voľné“ k variovaniu. Tento koncept sa opiera o myšlienku, že ak poznáme priemer populácie a poznáme aspoň jednu hodnotu (skóre), potom všetky ostatné môžeme „variovať“, aby sme dosiahli uvedený priemer populácie. Napríklad, ak máme zmerané 3 hodnoty, ktorých priemer je 4 a jedna z týchto hodnôt je 8, potom zvyšné dve môžu byť 2 a 2, alebo 1 a 3 – zvyšné dve hodnoty môžeme variovať; ak poznáme priemer a jednu z troch hodnôt, potom máme 2 stupne voľnosti (3 – 1 = 2).

[15] priemer a štandardnú odchýlku zaokrúhľujeme štandardne na dve desatinné miesta

V prípade, že máte dostatočne veľký výskumný súbor a hodnoty šikmosti a strmosti nepresahujú kritické hranice, môžete pri testovaní hypotéz použiť parametrický test – **Studentov t-test pre dva nezávislé výbery**. V prípade, že si nie ste istí, či je váš súbor dostatočne veľký, alebo máte pochybnosti o hodnotách šikmosti a strmosti, môžete ešte overiť normalitu distribúcie dát a rovnosť variabilít (odchýlok) závislej premennej pre obe skupiny. Ak sa stane, že niektorá z podmienok normálnej distribúcie a rovnosti variabilít nie je splnená, je vhodnejšie použiť neparametrický test – **Mann-Whitneyho U test**. V JASP sa výpočet t-testu a overenie normality distribúcie dát a rovnosti variabilít vykonáva v spoločnom kroku (kapitoly 7.2.1 *Výpočet t-testu pre dva nezávislé výbery* a 7.2.2 *Mann-Whitneyho U-test*).

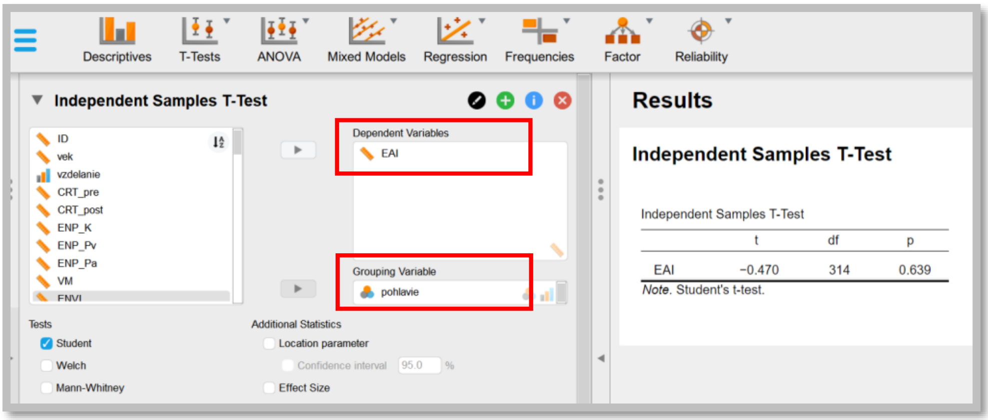
7.2.1 Výpočet t-testu pre dva nezávislé výbery

Ak by sme porovnávali výkon mužov a žien (dvoch nezávislých skupín) v ich postojoch k ochrane životného prostredia, mohli by sme formulovať alternatívnu (dvojsmernú) hypotézu nasledovne: predpokladáme rozdiely v postojoch k ochrane životného prostredia medzi mužmi a ženami. Našim zámerom by teda bolo testovať nulovú hypotézu, ktorá tvrdí, že medzi mužmi a ženami nie je významný rozdiel v ich postojoch k ochrane životného prostredia.

V prvom kroku je vhodné pozrieť sa na deskriptívne štatistiky (pozri kapitolu 4.3 *Výpočet deskriptívnych štatistík*) závislej premennej (postojoch k ochrane životného prostredia) u oboch skupín. Keďže vo výberovom súbore máme dostatočne veľkú vzorku (143 mužov a 173 žien) a ani jedna z hodnôt šikmosť (-0,040 pre mužov a -0,116 pre ženy) a strmosť (0,465 pre mužov a 0,074 pre ženy) nie je podozrivo vysoká, môžeme sa spoľahnúť na Studentov t-test pre dva nezávislé výbery.

Pre výpočet t-testu v sekcii *T-tests* vyberieme možnosť *Independent Sample T-Test*. Zobrazí sa nám okno na výpočet t-testu (Obrázok 22). Do okna *Variables* zvolíme premennú (alebo aj viacero premenných), ktorú sledujeme a v ktorej chceme porovnávať skupiny – v našom prípade postoje k ochrane životného prostredia EAI. Do okna *Grouping Variable* presunieme premennú, ktorá rozdeľuje vzorku na dve nezávislé skupiny—väčšinou s jedná o nominálnu premennú—v našom príklade pohlavie (môže to byť aj premenná, ktorá rozdeľuje vzorku na kontrolnú a experimentálnu skupinu).

V prípade, že by sme mali malý výberový súbor, alebo by hodnoty šikmosti a strmosti boli podozrivo vysoké, spravíme ešte medzi-krok – testovanie normality distribúcie dát a rovnosti variabilít, aby sme zhodnotili, či je vhodné použiť parametrický alebo neparametrický test. Normálna distribúcia závislej premennej sa overuje pomocou *Shapiro-Wilkovho testu*. T-test je relatívne robustný typ analýzy a malé odchýlky od normy sú prípustné, to však neplatí, ak sú porovnávané skupiny nerovnako veľké (odporúča sa držať pravidla, že rozdiel vo veľkosti skupín by nemal presahovať 1,5 násobok, teda ak je v jednej skupine 40 osôb, tá druhá by nemala byť menšia ako 27). Ak sa nepotvrdí normálna distribúcia dát, potom je potrebné využiť neparametrický *Mann-Whitneyho U test*.



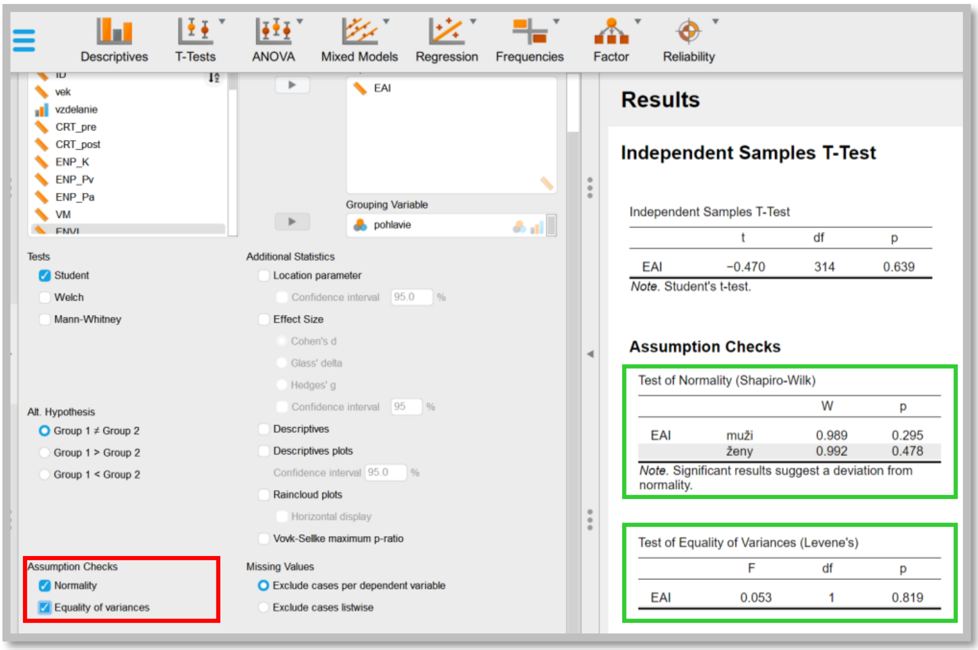
Obrázok 22 Studentov t-test pre dva nezávislé výbery – 1. krok

Rovnosť variabilít závislej premennej v oboch skupinách sa overuje pomocou *Levenovho testu*. Ak tento test preukáže signifikantné výsledky, je to indikátor toho, že skupiny nemajú rovnakú variabilitu závislej premennej a my musíme pri ich porovnávaní využiť upravenú t-štatistiku podľa *Welchovej metódy*.

Pri testovaní normality distribúcie dát a rovnosti variabilít v JASP zostávame stále v rovnakom okne; v časti *Assumption Check* zvolíme možnosti *Normality* a *Equality of variance* a pozrieme sa na výsledky (Obrázok 23).

V našom príklade Shapiro-Wilkov test normality vyšiel pre obe skupiny nesignifikantný (obe hodnoty *p* sú vyššie ako 0,05), to znamená, že aj u mužov aj u žien sú dáta pre postoje k životnému prostrediu distribuované normálne. Ak by jedna, alebo obe z hodnôt signifikancie *p* boli významné (< 0,05), potom by bolo potrebné zvážiť použitie *Mann-Whitneyho U testu* pre porovnanie priemerov oboch skupín.

V druhej tabuľke sú výsledky Levenovho testu rovnosti variabilít. Aj tento vyšiel nesignifikantný (*p* = 0,819), takže môžeme povedať, že nie sú prítomné rozdiely vo variabilite sledovaných skupín. Ak by Levenov test vyšiel významný (*p* < 0,05), potom by sme sa pri výpočte t-testu orientovali nie na Studentov t-test, ale na *Welchov t-test*.

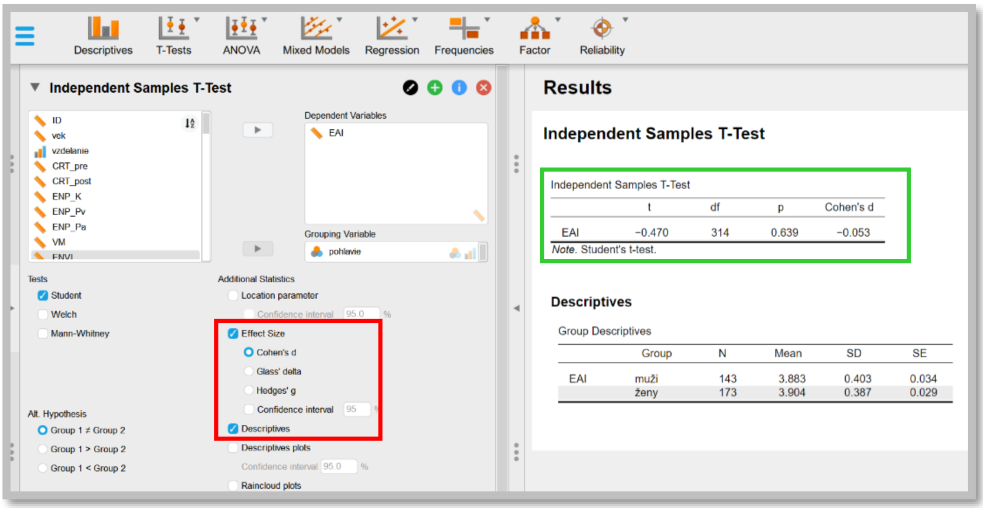


Obrázok 23 Testovanie normálnej distribúcie a rovnosti variabilít

Teraz sa môžeme pozrieť na samotné overenie rozdielov priemerov mužov a žien v ich postojoch k ochrane životného prostredia (Obrázok 24). V časti *Tests* je prednastavená voľba na Studentov t-test (zaškrtnutá možnosť *Student*). Toto je aj časť, kde si v prípade potreby (nerovnosti variabilít) môžeme zvoliť Welchovu korekciu hodnoty t-testu (možnosť *Welch*) namiesto Studentovho testu. Ďalej, v sekcii *Additional Statistics* zaškrtneme možnosti *Effect Size* (*Cohen's d*) na overenie vecnej významnosti a *Descriptives* na výpočet deskriptívnych štatistík.

Vo výsledkoch vidíme v prvej tabuľke výsledky Studentovho t-testu (v prípade, že sme zvolili Welchovu korekciu *t* hodnoty, všímame si riadok Welch): hodnotu *t* (*Statistic*), stupne voľnosti *df*, hodnotu signifikancie *p*, Cohenovo *d*. V tabuľke *Descriptives* vidíme deskriptívne štatistiky oboch skupín pre premennú postoje k ochrane životného prostredia (*EAI*). Priemerné skóre EAI mužov je 3,88 a žien 3,90. Už na pohľad ide o veľmi podobné skóre, čo potvrdzuje aj výsledok Studentovho t-testu: rozdiely medzi mužmi a ženami nie sú štatisticky významné (*p* = 0,639).

Pravdepodobnosť toho, že by sme namerali vyrovnané výsledky u mužov a u žien za predpokladu, že reálne medzipohlavné rozdiely neexistujú (teda že platí nulová hypotéza), je až 63,9% (p = 0,639). Hodnota Cohenovho d, vecná významnosť, je takmer nulová (d = 0,053[16]), teda rozdiel nemá takmer žiadnu vecnú významnosť, skóre skupín sa vo vysokej miere prekrýva.



Obrázok 24 Studentov t-test pre dva nezávislé výbery – 2. krok

Pri reportovaní výsledkov uvádzame hodnoty t-testu (t), stupne voľnosti (df), hodnotu signifikancie (p) a hodnotu Cohenovho d v podobe (t(hodnota df) = hodnota t-testu, p = hodnota signifikancie, d = hodnota Cohenovko d).

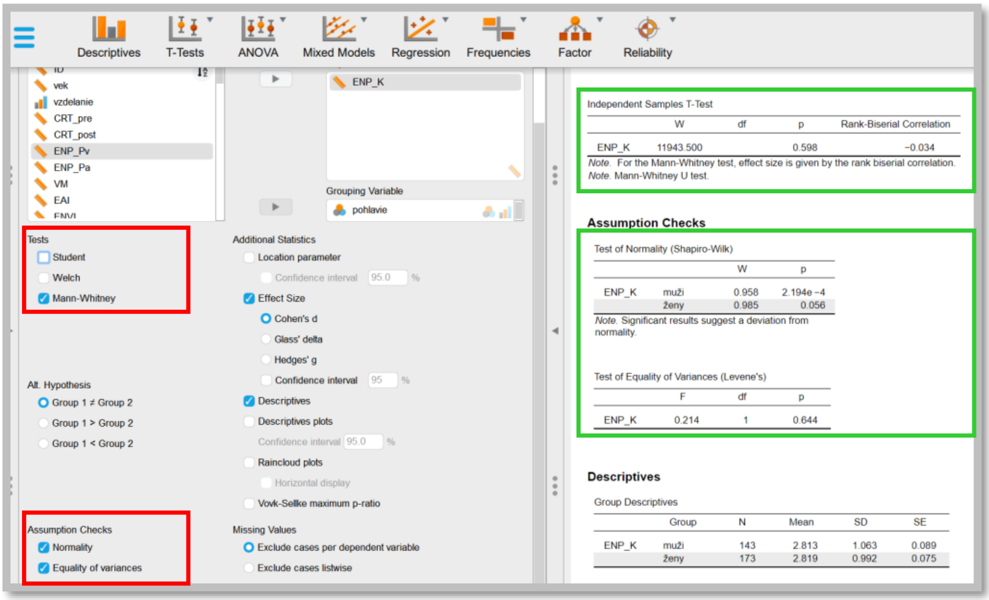
Príklad:

Studentov t-test pre dva nezávislé výbery nepreukázal signifikantný rozdiel medzi mužmi a ženami (t(314) = 0,470; p = 0,639; d = 0,053) v ich postojoch k ochrane životného prostredia.

[16] V tabuľke je negatívna hodnota, to ale znamená iba to, že prvá porovnávaná skupina má nižšie priemerné skóre ako druhá; pri reportovaní stačí uvádzať absolútnu hodnotu. Toto platí aj pre hodnotu t-testu (negatívna hodnota t-testu vyjadruje iba to, že prvá skupina má nižšie skóre ako druhá).

7.2.2 Mann-Whitneyho U test

V prípade malého výberového súboru, vysokých hodnôt šikmosti a strmosti a/alebo signifikantného výsledku Shapiro-Wilkovho testu normality je potrebné použiť neparametrický test – *Mann-Whitneyho U test*. Napríklad, ak by sme šli porovnávať mužov a ženy v ich nepodložených konšpiračných presvedčeniach (*ENP_K*), postup je rovnaký ako pri Studentovom t-teste, iba v časti *Tests* si navolíme možnosť *Mann-Whitney* (Obrázok 25), keďže test normality vyšiel signifikantný ($p < 0{,}001$ pre mužov).



Obrázok 25 Mann-Whitneyho U test

Vo výsledkoch sú v prvej tabuľke výsledky U-testu: hodnota Mann-Whitneyho U test (*W*), hodnota signifikancie *p* a hodnota *Rank-Biserial Correlation* (alternatíva Cohenovo *d* na zhodnotenie veľkosti efektu). Výsledok U-testu ukazuje, že rozdiely medzi mužmi a ženami nie sú štatisticky významné ($p = 0{,}598$). Rovnako aj hodnota *Rank-Biserial Correlation*[17] je veľmi nízka (0,034).

[17] Rank-Biserial Correlation interpretujeme ako korelačný koeficient (kapitola 9 *Korelácie*).

Pri reportovaní výsledkov uvádzame hodnoty U-testu (*W*) a hodnotu signifikancie (*p*) v podobe (*W* = hodnota U-testu, *p* = hodnota signifikancie). Pri neparametrických testoch vo výsledkoch reportujeme mediány, nie priemery.

Príklad:

Mann-Whitney U-test nepreukázal signifikantný rozdiel medzi mužmi (*Md* = 2,67) a ženami (*Md* = 2,67) v úrovni ich nepodložených konšpiračných presvedčení (*W* = 11943,50; *p* = 0,598).

7.3 Studentov t-test pre dva závislé výbery (párový t-test)

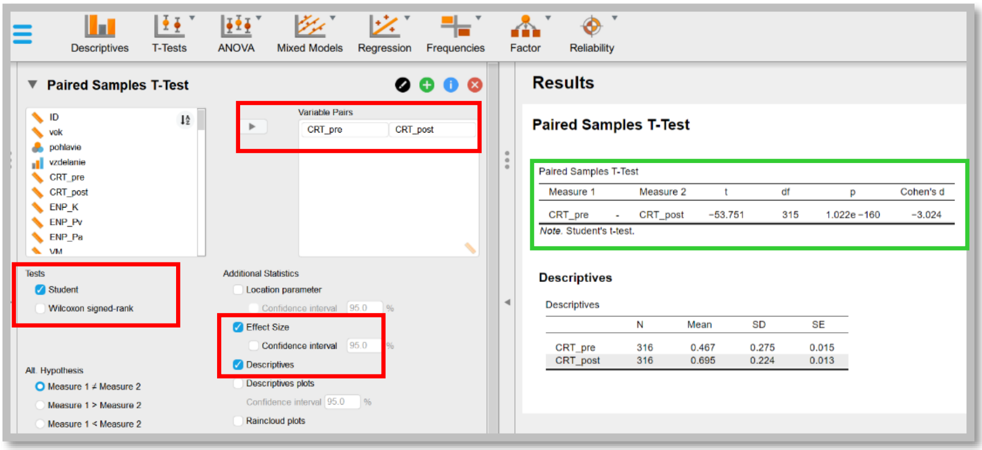
Ak máte vo výskume jednu skupinu participantov, pri ktorej sledujete zmenu vo výkone v teste kognitívnej reflexie pred a po absolvovaní kurzu zameraného na rozvoj analytického myslenia, tak každý participant je testovaný tým istým testom dvakrát, pred kurzom v kontrolnej podmienke a po kurze v experimentálnej podmienke. Hovoríme o opakovanom meraní, o vnútro-subjektovom dizajne (*within-subjects design*) a získavame párované dáta. Tieto párované dáta z dvoch meraní porovnávame pomocou párového t-testu.

Pri testovaní zhody v opakovaných meraniach je našim zámerom overiť predpoklad o rozdiele medzi prvým a druhým meraním (dvojsmerná alternatívna hypotéza). Nulová hypotéza by znela: medzi meraniami sa nič nezmení, výkon (priemer) participantov je v prvom a druhom meraní rovnaký. Podmienkou pre testovanie je, aby nezávislá premenná pozostáva z 2 kategórií – skupín a každý participant musí byť priradený do oboch skupín. Ďalej, aby rozdiel medzi prvým a druhým meraním bol približne normálne distribuovaný a aby neboli prítomné žiadne extrémne odchýlky v rozdiele medzi prvým a druhým meraním.

Ak dáta spĺňajú uvedené kritériá, použijeme parametrický párový test **Studentov t-test pre dva závislé výbery** alebo aj **párový t-test**. Opäť platí, že ak súbor nie je dostatočne veľký alebo hodnoty šikmosti a strmosti sú vysoké, overujeme normalitu distribúcie dát a rovnosť variabilít (odchýlok) závislej premennej pre obe merania. Ak sa stane, že niektorá z podmienok normálnej distribúcie a rovnosti variabilít nie je splnená, je vhodnejšie použiť neparametrický test – **Wilcoxon test**.

7.3.1 Výpočet t-testu pre dva závislé výbery

V sekcii *T-tests* vyberieme možnosť *Paired Sample T-Test*. Zobrazí sa okno na výpočet t-testu (Obrázok 26). Do okna *Variable pairs* navolíme dvojicu premenných (alebo aj viacero dvojíc premenných), v ktorých chceme skupinu porovnávať. V našom príklade, kedy chceme porovnať kognitívnu reflexiu pred (*CRT_pred*) a po absolvovaní kurzu (*CRT_po*) si ako dvojicu zvolíme premenné *CRT_pred* a *CRT_po* a presunieme ich okna *Variable pairs*. V sekcii *Additional Statistics* zaškrtneme možnosti *Effect Size* na overenie veľkosti efektu a *Descriptives* na výpočet deskriptívnych štatistík. Keďže naša hypotéza je dvojsmerná, ponechávame nastavenie v časti *Alt.Hypothesis* na možnosti *Measure 1 ≠ Measure 2*.



Obrázok 26 Párový t-test (Studentov t-test pre dva závislé výbery)

V prvej tabuľke sledujeme hodnoty Studentovho t-testu (riadok *Student*). Priemerné skóre kognitívnej reflexie pred kurzom je 0,47 a po kurze 0,70. V našom príklade výsledok Studentovho párového testu ukazuje, že rozdiely sú štatisticky významné: pravdepodobnosť toho, že by sme namerali rozdiely medzi meraniami pred a po kurze za predpokladu, že reálne tieto rozdiely neexistujú (teda že platí nulová hypotéza), je menšia ako 0,1% (p < 0,001). Musíme si všímať aj hodnotu vecnej významnosti. Tá je v tomto prípade vysoká (3,*024*) a môžeme konštatovať, že absolvovanie kurzu rozvíjajúceho analytické myslenie malo za následok nárast skóre v teste kognitívnej reflexie a že prekryv skóre pred a po kurze je minimálny.

Pri reportovaní výsledkov Studentovho t-testu uvádzame hodnoty t-testu (*t*), stupne voľnosti (*df*), hodnotu signifikancie (*p*) a hodnotu Cohenovho *d* v podobe (*t*(hodnota *df*) = hodnota t-testu, *p* = hodnota signifikancie, *d* = hodnota Cohenovho *d*).

Príklad:

Studentov t-test pre dva závislé výbery preukázal signifikantný rozdiel v kognitívnej reflexii pred a po absolvovaní kurzu (t(315) = 53,751; p < 0,001; d = 3,024): participanti dosahovali signifikantne nižšie skóre (M = 0,467; SD = 0,275) pred absolvovaním kurzu ako po ňom (M = 0,695; SD = 0,224), pričom výsledok je vysoko vecne významný.

7.3.2 Wilcoxonov test (neparametrická alternatíva)

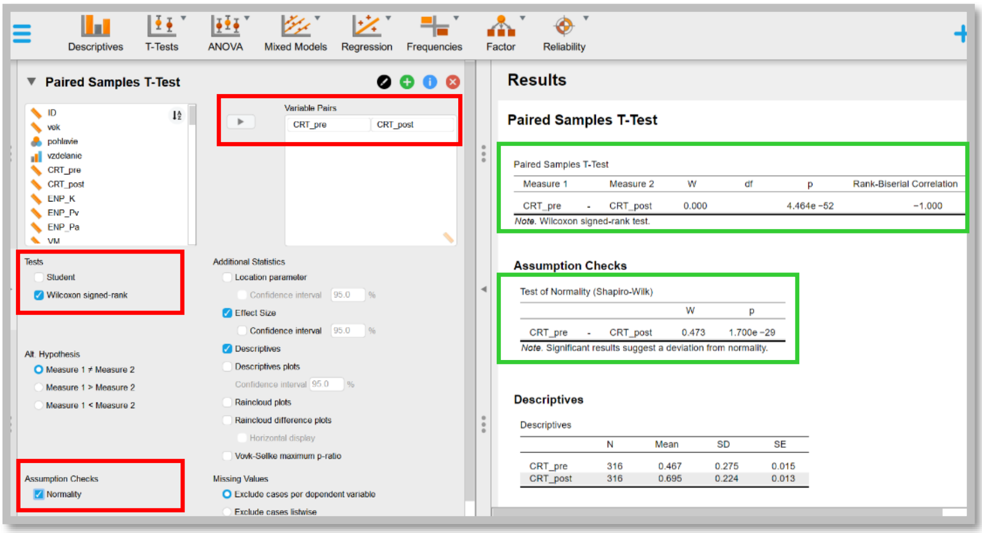
V prípade, že je potrebné overiť normálne rozloženie rozdielov v meraniach (malá vzorka, vysoké hodnoty šikmosti, strmosti) zaškrtneme v časti *Assumption Checks* možnosť *Normality* (Obrázok 27). V takomto prípade si najskôr pozrieme tabuľku *Test of Normality* (Shapiro-Wilk), aby sme vedeli, či budeme sledovať Studentov t-test alebo Wilcoxonov test. V našom príklade test normality naznačuje, že rozdiely v meraniach nie sú normálne distribuované, preto je potrebné použiť neparametrický *Wilcoxonov test* (ak by test normality preukázal normálnu distribúciu, tak by sme si zvolili iba Studentov t-test).

V našom príklade (Obrázok 27) výsledky testu normality odporúčajú použitie Wilcoxonovho neparametrického testu. V prvej tabuľke sledujeme teda riadok *Wilcoxon* a všímame si hodnotu testu (*Statistic*), stupne voľnosti *df*, hodnotu signifikancie *p*, a hodnotu veľkosti efektu (*Effect Size*). Hodnota signifikancie naznačuje, že rozdiely medzi meraniami sú v našom príklade štatisticky významné (p < 0,001). Hodnota veľkosti efektu (meraná prostredníctvom *Rank-Biserial Correlation*) je tiež ukazovateľom toho, že efekt bol významný.

Pri reportovaní výsledkov uvádzame hodnoty Wilcoxonovho testu (*W*) a hodnotu signifikancie (*p*) v podobe (*W* = hodnota U-testu, *p* = hodnota signifikancie). Pri Wilcoxonovom teste vo výsledkoch reportujeme namiesto priemerov mediány.

Príklad:

Wilcoxonov test preukázal signifikantný rozdiel v kognitívnej reflexii pred (Md = 0,50) a po absolvovaní kurzu (Md = 0,76), pričom výsledok je vysoko vecne významný (W = 0,000; p < 0,001; *Rank-Biserial Correlation* = 1,000).



Obrázok 27 Wilcoxonov test

8 Analýza variancie

V predchádzajúcej kapitole sme si ukázali, ako otestovať rozdiely v priemeroch dvoch výberov, dvoch skupín. Vo výskumnej praxi sa môžeme stretnúť aj s prípadom, kedy potrebujeme porovnať priemery viac ako dvoch skupín. Niekomu môže napadnúť, že ak chceme porovnať napríklad priemery troch tried, tak by sme mohli vypočítať t-testy pre nasledovné kombinácie tried:

$$\text{trieda A} - \text{trieda B}$$
$$\text{trieda A} - \text{trieda C}$$
$$\text{trieda B} - \text{trieda C}$$

Takýto postup by však vystavil naše výsledky vyššej pravdepodobnosti chyby typu I. Uvažujme, že by vo všetkých troch prípadoch vyšla signifikancia $p < 0{,}05$, teda pravdepodobnosť, že sa nedopustíme chyby typu I by bola v každom prípade 95%. Zároveň však treba brať do úvahy, že pre samostatných t-testoch by sa jednalo o nezávislé testy, preto by sme celkovú pravdepodobnosť chyby typu I počítali nasledovne:

$$0{,}95 * 0{,}95 * 0{,}95 = 0{,}86$$

To znamená, že kumulatívne by pravdepodobnosť, že sa nedopustíme chyby typu I bola iba 86%. Aby sme sa podobnej chybe vyhli, pri porovnávaní viac ako dvoch skupín využívame *analýzu rozptylu*, skrátene *ANOVA* (*analysis of variance*). ANOVA pracuje ako súhrnná analýza, ktorej výsledkom je hodnota ***F***. Pri ANOVA testujeme nulovú hypotézu, že neexistuje žiadny významný rozdiel medzi priemermi sledovaných skupín. ANOVA sa počíta v dvoch krokoch. V prvom kroku výsledok ANOVA ukáže, či je medzi skupinami významný rozdiel, v druhom kroku ukáže medzi ktorými skupinami a aké veľké tieto rozdiely sú. Na to využívame v rámci ANOVA dodatočné testy—*post-hoc testy*—viacnásobné párové porovnania skupín, pričom na zhodnotenie *štatistickej významnosti* rozdielov medzi jednotlivými skupinami (post-hoc testov) sa využíva výsledok t-testu ***t*** a jeho štatistická významnosť ***p***.

ANOVA analyzuje rôzne zdroje, z ktorých pramenia variácie v skóre, jej cieľom je zistiť, či existujú rozdiely v priemeroch skupín (Dancey & Reidy, 2017). V prvom kroku sa v rámci analýzy vypočíta „veľký“ priemer a následne sa zisťuje, do akej miery sa jednotlivci odlišujú od tohto priemeru. Tieto odlišnosti, resp. variabilita, má dve zložky: variabilita skóre medzi skupinami a variabilita v rámci skupín. Úlohou ANOVA je odhadnúť, akou mierou prispievajú tieto dve zložky k celkovej variabilite. Variabilita v rámci skupín sa počíta ako odchýlka každého jednotlivca od priemeru danej skupiny. Variabilita medzi skupinami sa počíta ako odchýlka priemerov jednotlivých skupín od „veľkého" priemeru. Inými slovami, variabilita poukazuje na to, ako veľmi sa skupiny odlišujú od seba navzájom, a ako sa zároveň odlišujú od celkového spoločného priemeru. Záverečný výpočet zahŕňa nájdenie priemeru medzi-skupinovej variancie, vnútro-skupinovej variancie a celkovej variancie. Na zhodnotenie *štatistickej významnosti* rozdielov medzi skupinami sa využíva hodnota ***F*** ako výsledok analýzy variancie (významnosť ktorej posudzujeme podľa jej štatistickej významnosti ***p***).

ANOVA predstavuje veľmi flexibilnú a všeobecnú techniku (nemá obmedzenia na počet skupín), má veľmi široké možnosti využitia, jednotlivé aplikácie analýzy rozptylu si však vyžadujú dodržanie rôznych požiadaviek na dáta (Goss-Sampson, 2020b; Miles & Banyard, 2007). My si v tejto kapitole predstavíme najzákladnejší typ analýzy rozptylu – ANOVA pre jeden faktor, alebo jednosmerná, jednocestná ANOVA (*one way ANOVA*), kedy porovnávame viac ako dve skupiny, či merania v jednom faktore. Tento jeden faktor je závislá premenná, nazýva sa aj výsledná premenná (*outcome variable*) – závisí od prediktora, teda premennej, ktorú manipulujeme (napríklad zaradenie do experimentálnych skupín) alebo ktorú meriame (napríklad rôzne ročníky). Tieto prediktory predstavujú nezávislú premennú.

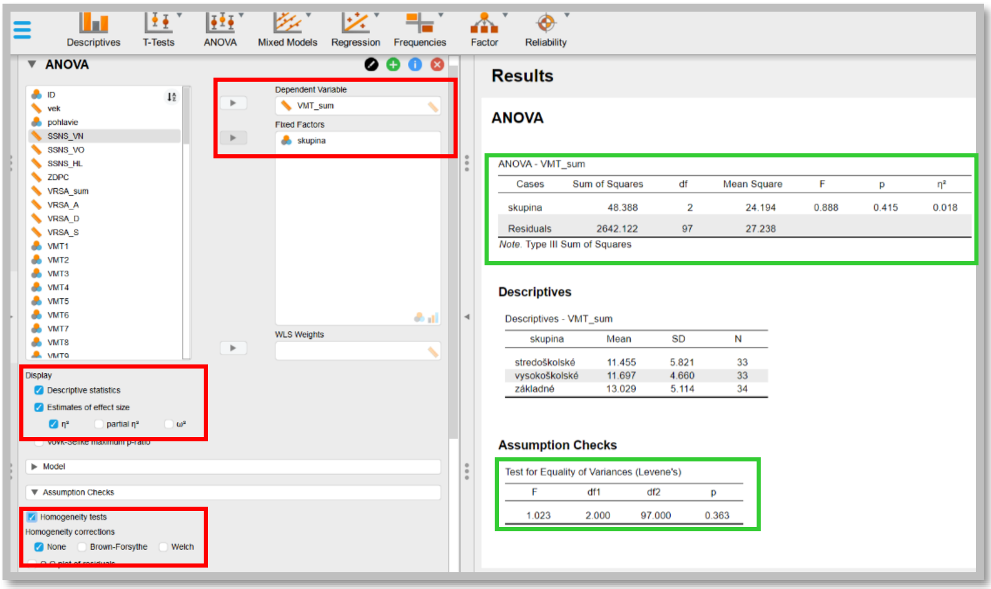
Predpokladom pre použitie jednocestnej analýzy rozptylu je, aby nezávislá premenná bola meraná na nominálnej úrovni, aby skupiny boli na sebe nezávislé, aby závislá premenná mala približne normálnu distribúciu, v dátach neboli žiadne extrémne hodnoty a aby medzi skupinami existovala rovnosť rozptylov (inak hodnota signifikancie p pre štatistiku F nemusí byť spoľahlivá; Goss-Sampson, 2020b).

8.1 Výpočet analýzy rozptylu pre jeden faktor

Predstavte si, že sledujete populáciu s rôznymi stupňami vzdelania (*základné, stredné, vysokoškolské*) a chcete zistiť, či sú medzi ľuďmi s rôznym stupňov vzdelania rozdiely v ich kognitívnych schopnostiach (*VMT*). Alternatívna hypotéza by znela: predpokladáme, že úroveň kognitívnych schopností sa bude odlišovať u ľudí so základným, stredoškolským a vysokoškolským vzdelaním. Nulová hypotéza by znela: neexistujú rozdiely v kognitívnych schopnostiach medzi ľuďmi so základným, stredoškolským a vysokoškolským vzdelaním.

V sekcii ANOVA zvolíme možnosť *ANOVA*. Zobrazí sa okno na výpočet analýzy rozptylu (Obrázok 28). Do okna *Dependent Variable* presunieme závislú premennú, teda faktor, v ktorom porovnávame sledované skupiny (v našom príklade *kognitívne schopnosti VMT*), a do okna *Fixed Factors* presunieme nezávislú premennú (v našom prípade premennú *skupina*). V tomto základnom kroku si ešte v časti *Display* zvolíme aj výpočet deskriptívnych štatistík pre sledované skupiny (*Descriptive statistics*) a aj výpočet veľkosti efektu (*Estimates of effect size*), ktorý počítame prostredníctvom druhej mocniny koeficientu eta (η^2)[18]. Tento koeficient možno interpretovať nasledovne:

- $\eta^2 = 0{,}01$ – malý efekt
- $\eta^2 = 0{,}06$ – stredný efekt
- $\eta^2 = 0{,}14$ – veľký efekt



Obrázok 28 Analýza rozptylu (ANOVA)

Tak ako pri t-testoch, aj pri výpočte ANOVA, ak máme malé vzorky, alebo hodnoty šikmosti a strmosti sú vysoké, testujeme rovnosť variabilít závislej

---

[18] JASP ponúka tri možnosti výpočtu veľkosti efektu. Základným ukazovateľom je druhá mocnina koeficientu *eta* η^2, ktorá sa interpretuje ako podiel (%) v celkovej odchýlke závislej premennej, ktorý je možný prisúdiť nezávislej premennej(Goss-Sampson, 2020b). Druhou možnosťou je *partial eta* $\eta_p{}^{2}$', využíva sa skôr v prípadoch, kedy do analýzy vstupuje viacero premenných, pre analýzu variancie s jedným faktorom je však zhodná s druhou mocninou koeficientu *eta* η^2. Poslednou možnosťou je omega ω^2, ktorá sa využíva v prípade malých skupín (n < 30).

premennej vo všetkých skupinách pomocou *Levenovho testu* (pripomeňme si, že ak Levenov test preukáže signifikantné výsledky, znamená to, že skupiny nemajú rovnakú variabilitu sledovaného faktora a je potrebná korekcia). V tomto prípade rozrolujeme časť *Assumption Checks* a najskôr zvolíme možnosť *Homogenity Tests* (Obrázok 28). Ak vo výsledkovej tabuľke pre Levenov test vyjde nesignifikantný výsledok, tak v rámci *Homogenity corrections* zvolíme možnosť *None* (teda žiadna korekcia). Ak vyjde signifikantné výsledok, zvolíme jednu z dvoch ponúkaných korekcií – *Brown-Forsythe* alebo *Welch;* vo väčšine situácií sa odporúča Welchova korekcia, v prípade, že dáta sú zošikmené, potom je vhodnejšia Brown-Forsytheova korekcia (Glantz et al., 2016). Tu je na mieste upozornenie – pokým nezvolíte aspoň jeden z typov korekcie (none, Brown-Forsythe, alebo Welch), vo výsledkovej tabuľke sa nezobrazia žiadne výsledky pre ANOVA. V našom príklade (Obrázok 28) je výsledok nesignifikantný (p = 0,363), preto sme nezvolili žiadnu korekciu (možnosť *None* v *Homogenity correction*).

Teraz sa môžeme pozrieť na výsledky analýzy variancie ANOVA. V tabuľke s deskriptívnymi štatistikami sa na prvý pohľad zdá, že skupiny nie sú v kognitívnych schopnostiach rovnaké. Ľudia so základným vzdelaním dosiahli priemerné skóre 11,46 (SD = 5,82), ľudia so stredoškolským vzdelaním 11,70 (SD = 4,67) a ľudia s vysokoškolským vzdelaním dosiahli priemerné skóre 13,03 (SD = 5,11). To, do akej miery sú tieto rozdiely významné nám ale povie výsledok ANOVA v prvej tabuľke. V nej vidíme, že hodnota signifikancie p nie je nižšia ako 0,05, čo znamená, že rozdiely medzi skupinami s rôznou úrovňou vzdelania nie sú štatisticky významné. Inými slovami, za predpokladu, že reálne v populácii tieto rozdiely neexistujú (platí nulová hypotéza), je pravdepodobnosť toho, že by sme namerali u ľudí s rôznym vzdelaním významne rozdielne skóre až 41,5% (p = 0,415). Hodnota druhej mocniny koeficientu eta (η^2 = 0,018) poukazuje na malý efekt.

Pri reportovaní výsledkov uvádzame hodnoty F-testu (F), stupne voľnosti (df), hodnotu signifikancie (p) a hodnotu eta η^2 v podobe (F(obe hodnoty df) = hodnota F-testu, p = hodnota signifikancie, η^2 = hodnota eta).

Príklad:

Analýza variancie ANOVA pre jeden faktor nepreukázala signifikantný rozdiel medzi ľuďmi s rôznym stupňom vzdelania v ich kognitívnych schopnostiach (F(2, 97) = 0,888; p = 0,415; η^2 = 0,018).

V prípade, že ANOVA nepreukáže signifikantné rozdiely, v tomto kroku analýza končí. Ak sa však preukážu štatisticky významné rozdiely ($p < 0,05$), potom v analýze pokračujeme ďalej. Keď totižto odhalíme rozdiely medzi sledovanými skupinami, zaujímajú nás detaily – aké sú tieto rozdiely, medzi ktorými skupinami sú, resp. nie sú. Na to slúžia dodatočné t-testy, označované aj ako *post-hoc testy*.

8.2 Post-hoc testy

Post-hoc testy sú testy založené na princípe t-testov, a slúžia na to, aby sme zistili, či rozdiely, na ktoré poukazuje ANOVA sú spôsobené rozdielmi medzi všetkými sledovanými skupinami navzájom, alebo sa odlišuje iba skupina 1 od skupiny 2, nie však od skupiny 3, alebo sa odlišuje skupina 2 od skupiny 1 a 3, pričom skupiny 1 a 3 sú bez rozdielov, a podobne. Mohli by sme spraviť separátne t-testy pre každú dvojicu sledovaných skupín, no ako už bolo spomenuté, viedlo by to k zvýšenej pravdepodobnosti, že sa dopustíme chyby typu I. V rámci post-hoc testov sa vykonajú viacnásobné t-testy, ktoré pracujú s rôznymi typmi korekcie, aby sme zamedzili práve zvýšenej pravdepodobnosti chyby typu I, pričom JASP ponúka 5 typov korekcie(Goss-Sampson, 2020b):

- *Tukey* – alebo aj *Honest Significant Difference* (HSD) je jeden z najčastejšie používaných testov, poskytuje dobrú kontrolu chyby typu I pri skupinách s rovnakou veľkosťou a rovnosťou variabilít (niektoré programy pre nerovnako veľké skupiny ponúkajú korekciu Tukey-Kramer test), porovnáva v jednom čase iba jeden pár
- *Bonferroni* – poskytuje zaručenú kontrolu chyby typu I avšak s rizikom, že sa zníži štatistická sila, nepredpokladá nezávislosť porovnaní, môže sa využívať aj pri rôzne veľkých skupinách, využíva sa pri plánovanom porovnávaní
- *Scheffe* – kontroluje celkovú úroveň spoľahlivosti aj pri odlišne veľkých skupinách, v jednom čase porovnáva viacero párov, poskytuje ale nižšiu štatistickú silu a má menšiu šancu detekovať všetky skutočné rozdiely medzi skupinami
- *Holm* – alebo aj Holm-Bonferroni, menej konzervatívny ako pôvodný Bonferroniho test
- *Šidák* – podobný ako Bonferroni, vykonáva viacnásobné porovnania, predpokladá, že porovnania sú nezávislé, trochu silnejší ako Bonferroni

Okrem toho, existujú aj viaceré typy post-hoc testov, v JASP sú v ponuke štyri:

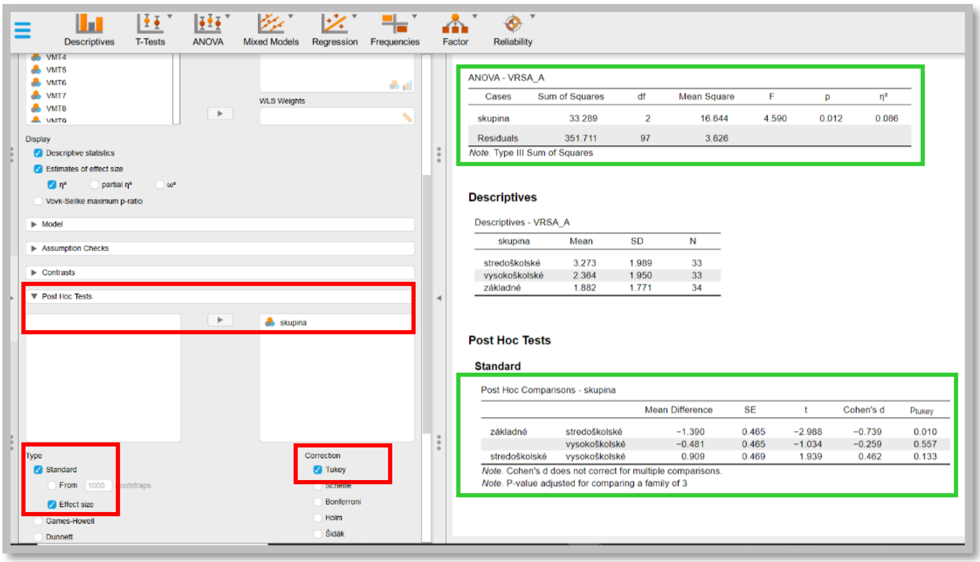
- *Standard* – prednastavený základný typ
- *Games-Howell* – používa sa, ak si nie ste istí rovnosťou variabilít

- *Dunnett's* – porovnáva všetky skupiny s jednou skupinou (napr. s kontrolnou)
- *Dunn* – neparametrická alternatíva post hoc testu pri testovaní malých skupín

Vezmime si nový príklad. Predstavte si, že chcete zistiť, či sú medzi ľuďmi s rôznym stupňov vzdelania rozdiely v ich rizikovom správaní (*VRSA_A*). Alternatívna hypotéza by znela: predpokladáme, že úroveň rizikového správania sa bude odlišovať u ľudí so základným, stredoškolským a vysokoškolským vzdelaním. Nulová hypotéza by znela: neexistujú rozdiely v rizikovom správaní medzi ľuďmi so základným, stredoškolským a vysokoškolským vzdelaním.

Vypočítame analýzu variancie ANOVA a zistíme signifikantný rozdiel medzi ľuďmi s rôznym stupňom vzdelania v ich rizikovom správaní ($F(2, 97) = 4{,}590$; $p = 0{,}012$; $\eta^2 = 0{,}086$), pričom efekt je stredne veľký. Vzhľadom na signifikantné rozdiely medzi skupinami musíme v nasledujúcom kroku overiť rozdiely medzi jednotlivými sledovanými skupinami v úrovni ich rizikového správania. Rozbalíme sekciu *Post-hoc tests*. Nezávislú premennú (*skupina*) presunieme z ľavého okna do pravého, vyberieme základný typ post-hoc testov *Standard* (môžeme si zvoliť aj zobrazenie veľkosti efektu *Effect size*) a vzhľadom na približne rovnako veľké skupiny zvolíme *Tukey*ho korekciu (Obrázok 29).

Vo výsledkovej tabuľke je porovnanie všetkých skupín a hodnoty poukazujúce na ich štatistickú a vecnú významnosť (podobnosť s t-testami vôbec nie je náhodná, vidíme hodnoty t-testu t, signifikancie p_{tukey}, aj Cohenovo d pre každý porovnávaný pár skupín). V našom príklade bol významné rozdiely iba medzi osobami so základným a stredoškolským vzdelaním ($t = 2{,}988$; $p = 0{,}010$; $d = 0{,}739$), pričom platí, že najvyššie skóre rizikového správania mali stredoškolsky vzdelaní ($M = 3{,}27$, $SD = 1{,}99$), za nimi nasledovali vysokoškolsky vzdelaní ($M = 2{,}36$, $SD = 1{,}95$) a najnižšie skóre mali osoby so základným vzdelaním ($M = 1{,}88$, $SD = 1{,}77$). Na tomto odseku už môžete vidieť aj príklad toho, ako výsledky post-hoc testov reportovať v práci alebo v článku.



Obrázok 29 Analýza rozptylu (ANOVA) – post-hoc testy

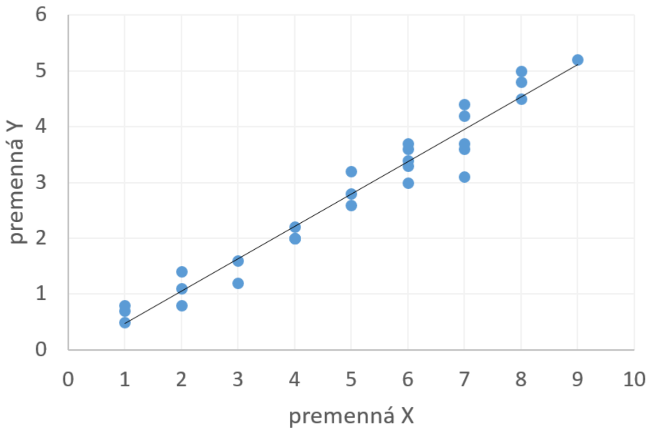
9 Korelácie (testovanie tesnosti vzťahov)

V tejto chvíli by ste mali vedieť, ako pomocou štatistiky overiť rozdiely medzi rôznymi skupinami alebo meraniami. Vo výskume však nezriedka potrebujeme skúmať aj vzťahy medzi premennými, napríklad či ľudia, ktorí skórujú vysoko v jednom dotazníku majú tendenciu skórovať vysoko (alebo nízko) v inom dotazníku. Štatistická procedúra, ktorá nám umožňuje skúmanie vzťahov medzi premennými, resp. medzi dvoma skupinami skóre u jednej skupiny, sa nazýva *korelačná analýza*. Korelačná analýza dáva odpoveď na otázku, aký tesný je vzťah medzi sledovanými premennými, do akej miery majú premenné tendenciu „meniť sa spolu" – kovariovať (Coolican, 2019; Miles & Banyard, 2007). Ak premenné spolu korelujú, sú na sebe závislé (zmena v jednej súvisí so zmenou druhej premennej; Dancey & Reidy, 2017). Takýto vzťah v štatistike nazývame **korelácia**.

Vzťahy medzi premennými, alebo korelácie, overujeme vždy, keď máme na mysli predpoklad typu „vyššia inteligencia súvisí s lepším výkonom v škole", alebo „rozvinuté kritické myslenie súvisí s nižšou mierou viery v hoaxy". Takéto vzájomné vzťahy vieme vizualizovať—na os X nanesieme jednu premennú a na os Y druhú a vytvoríme bodový diagram (*scatterplot*, Obrázok 30). V bodovom diagrame každý bod predstavuje jedného participanta s príslušnými hodnotami oboch sledovaných premenných. Z Obrázka 30 sa zdá, že premenné X a Y spolu súvisia – vyššie hodnoty X súvisia s vyššími hodnotami Y. Samotnú koreláciu by sme mohli graficky znázorniť ako priamku prechádzajúci cez tieto body tak, aby boli body od nej čo najmenej vzdialené.

Keď podrobujeme dáta korelačnej analýze, hovoríme o troch rôznych charakteristikách vzťahu, ktoré musíme zohľadňovať – tvar, smer a sila.

Tvar vieme vizuálne identifikovať už z bodového grafu. Pri výpočte korelácií takmer vždy predpokladáme, že sa jedná o *lineárny* vzťah. Lineárny vzťah medzi premennými je vtedy, ak po (pomyselnom) nakreslení čiary cez stredové body získame rovnú čiaru (ako na Obrázku 30). Vzťah medzi premennými môže mať aj podobu *krivky*, keď je čiara prechádzajúca stredom bodov zakrivená, napríklad v podobe exponenciálnej krivky, alebo v podobe krivky v tvare U, prípadne obráteného U. A niekedy sa nám môže stať, že pri pohľade na bodový graf nie je možné identifikovať žiaden tvar vzťahu, nie je možné viesť čiaru stredom bodov.



Obrázok 30 Bodový diagram – korelácia

Zaujímať sa o tvar korelácie (tvar čiary) je potrebné. Môže sa stať, že korelačný koeficient naznačí, že vzťah nie je prítomný, no pri pohľade na graf vidíme, že premenné vytvárajú nejaký „vzor“, čiara má tvar, ktorý by sme mohli interpretovať. Krivka v tvare U naznačuje, že vysoké hodnoty premennej Y sú iba v prípade extrémne nízkych a extrémne vysokých hodnôt premennej X. Napríklad, ak by sme sledovali náchylnosť na infekčné ochorenia vo vzťahu k veku, mohli by sme zistiť, že vzťah je veľmi slabý, no pri pohľade na bodový graf by sme odhalili, že vysoká náchylnosť na infekcie je prítomná vo veľmi nízkom a vo veľmi vysokom veku. To už je zistenie hodné povšimnutia, interpretácie a ďalšieho skúmania. A to ja napriek tomu, že samotný koeficient korelácie je nízky.

Smer vzťahu medzi dvoma premennými poukazuje na to, či sa premenné menia rovnakým alebo opačným spôsobom. Ak sa premenné menia rovnakým spôsobom, rovnakým smerom (pri stúpajúcom X stúpa aj Y, pri klesajúcom X klesá aj Y), hovoríme o *pozitívnom* vzťahu, pozitívnej korelácii. Ak sa premenné menia opačným smerom (pri stúpajúcom X premenná Y klesá, pri klesajúcom X premenná Y stúpa), hovoríme o *negatívnom* vzťahu, negatívnej korelácii. Smer vzťahu vieme odčítať z grafu, no vieme ho posúdiť aj z výsledkov korelačnej analýzy. Výsledkom korelačnej analýzy je buď Pearsonov korelačný koeficient r, alebo jeho neparametrická alternatíva – Spearmanovo rho ρ (označované aj r_s). Ak je hodnota korelačného koeficientu pozitívna, hovoríme pozitívnom vzťahu, ak je negatívna, hovoríme o negatívnom vzťahu.

Sila vzťahu vypovedá o tom, aký silný, tesný je vzťah medzi premennými, ktoré skúmame. Korelačný koeficient môže dosahovať hodnoty od -1 po 1. Pritom platí, že čím je hodnota koeficientu bližšie k 1 (resp. -1), tým je vzťah tesnejší, a čím je bližšie k 0, tým je vzťah slabší. Vo všeobecnosti sa uznáva nasledovné označovanie sily vzťahov (Miles & Banyard, 2007):

- *r* = 0 – 0,1 (resp. -0,1 – 0) – žiadny vzťah
- *r* = 0,1 (resp. -0,1) – slabý pozitívny (resp. negatívny) vzťah
- *r* = 0,3 (resp. -0,3) – stredne silný pozitívny (resp. negatívny) vzťah
- *r* = nad 0,5 (resp. -0,5) – silný pozitívny (resp. negatívny) vzťah

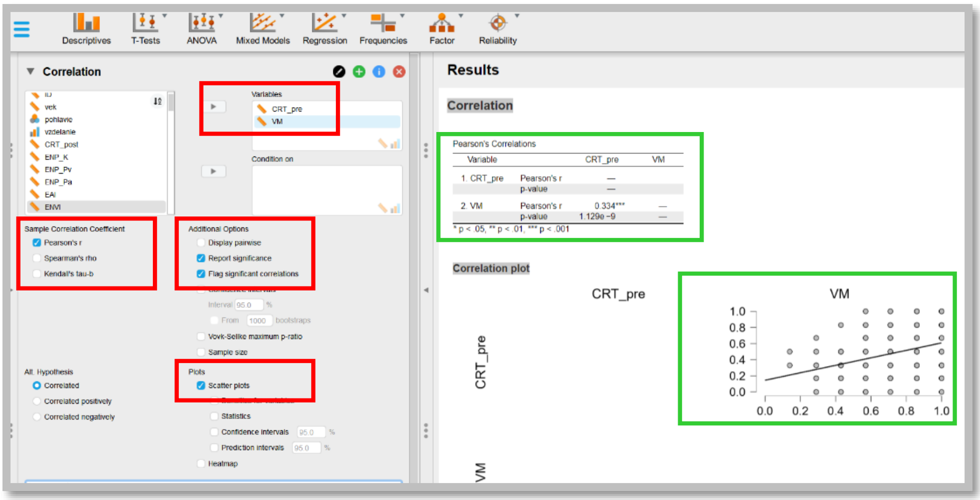
Pozor ale na koreláciu, ktorá sa blíži nule. Môže sa totiž stať, že hoci vám *r* vyjde veľmi nízke, graf vám ukáže také rozloženie dát, ktoré dáva zmysel. Už sme tento prípad spomínali vyššie: ide o prípad, kedy sú dáta usporiadané do krivky (linearita medzi premennými je narušená; Coolican, 2019). Takže pri korelačnej analýze by ste vždy mali pracovať aj s bodovým grafom.

Skôr ako sa pustíme do počítania korelácií, je potrebné spomenúť ešte dve zásadné skutočnosti, ktoré je potrebné si zapamätať o koreláciách. Za prvé, existencia akokoľvek silnej korelácie neznamená, že medzi premennými je aj príčinný (kauzálny) vzťah. Inými slovami **korelácia ≠ kauzalita**. To, že zmenou jednej premennej nastáva zmena druhej premennej, nám nehovorí nič o kauzalite – nevieme, ktorá premenná je príčinou a ktorá dôsledkom. Za druhé, pomocou korelačného koeficientu síce zistíme silu, tesnosť, vzťahu no nevypovedá to nič o štatistickej významnosti tohto výsledku (na to slúži hodnota signifikancie *p*).

9.1 Výpočet korelačnej analýzy

Miera tesnosti lineárneho vzťahu medzi dvoma premennými je najčastejšie vyjadrená Pearsonovým korelačným koeficientom *r*. Predstavme si, že nás zaujíma, či súvisí úroveň vedeckého myslenia s úrovňou kognitívnej reflexie. Na overenie tohto predpokladu je potrebné overiť tesnosť vzťahu medzi vedeckým myslením a kognitívnou reflexiou. Nulová hypotéza v tomto prípade tvrdí, že vzťah medzi sledovanými premennými neexistuje.

V sekcii *Regression* vyberieme možnosť *Correlation* (Obrázok 31). Zobrazí sa okno na výpočet korelačného koeficientu. Do okna *Variables* navolíme premenné, medzi ktorými chceme overiť vzťah – v tomto prípade premennú vedecké myslenie *VM* a premennú kognitívna reflexia *CRT_pre* (do okna *Variables* môžeme presunúť aj viacero premenných, tým získame vo výsledkoch korelačnú maticu, kde si vieme pozrieť všetky vzájomné vzťahy medzi zvolenými premennými).



Obrázok 31 Korelačná analýza

Všimnite si, že možnosť Pearsonovho korelačného koeficientu je prednastavená, takže za bežných okolností v tejto časti nemusíte nič meniť. Ďalej je vhodné sa pozrieť na graf – v sekcii *Plots* zaškrtneme možnosť *Scatter plots*. Vo výsledkovej časti sa zobrazí graf, podľa ktorého vieme posúdiť tvar korelácie. V ďalšom kroku môžeme v sekcii *Additional Statistics* zvoliť možnosti, ktoré nám uľahčia „čítanie" výsledkov. Odporúčam zaškrtnúť možnosť *Flag significant correlations*, vďaka čomu budú signifikantné výsledky v tabuľke označené hviezdičkami

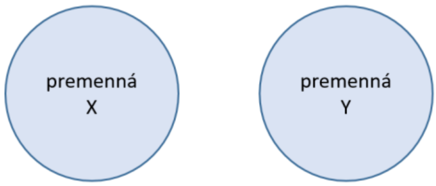
$$* \; p < 0{,}05$$

$$** \; p < 0{,}01$$

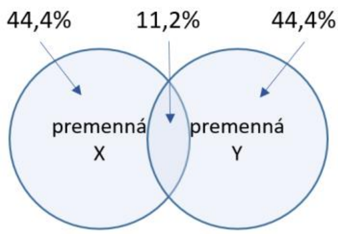
$$*** \; p < 0{,}001$$

Vo výsledkoch sa pozrieme na bodový graf, ktorý naznačuje lineárny vzťah (rovná čiara naprieč stredmi bodov). V tabuľke vidíme hodnoty *Pearson's r* (korelačný koeficient *r*) a *p-value* (hodnotu signifikancie *p*). Najskôr sa pozrieme na silu a smer vzťahu vyjadrené koeficientom *r*. Hodnota koeficientu *r* je 0,334: hodnota je pozitívna (smer vzťahu) na úrovni slabého až stredne silného vzťahu (sila vzťahu). Čo sa týka štatistickej významnosti, hodnota je označená troma hviezdičkami (pretože *p* je veľmi nízke, blíži sa 0), čo znamená, že pravdepodobnosť toho, že by sme namerali takto silný vzťah za predpokladu, že reálne neexistuje (teda že platí nulová hypotéza), je takmer nulová (p < 0,001). Môžeme zhodnotiť, že vzťah je pozitívny, slabý až stredný, a je štatisticky významný.

S korelačným koeficientom vieme spraviť ešte jednej krok, ktorý nám ukáže, do akej miery môžeme rozptyl v jednej premennej vysvetliť druhou premennou. Stačí, ak umocníme koeficient r na druhú a získame tak *koeficient determinácie* R^2 (0% znamená, že rozptyl v jednej premennej nemožno vôbec vysvetliť druhou premennou; 100% znamená, že celý rozptyl v jednej premennej možno vysvetliť druhou premennou). V prípade r = 0,334 je R^2 = 0,112, teda 11,2%. Takýmto spôsobom slúži r (presnejšie R^2) ako svoja vlastná miera veľkosti efektu (v zásade platí, že ak r > 0,10 ide o malý efekt, ak r je v rozmedzí 0,30 – 0,50 ide o stredný efekt, a ak r > 0,50 hovoríme o veľkom efekte). Koeficient determinácie 11,2% znamená, že variabilita vo vedeckom myslení prispieva 11,2% k variabilite v kognitívnej reflexii (slabý efekt) a naopak, variabilita v kognitívnej reflexii prispieva 11,2% k variabilite vo vedeckom myslení (slabý efekt). Pre lepšiu predstavivosť si premenné predstavme ako kruhy – v rámci nulovej hypotézy tvrdíme, že medzi nimi nie je žiaden vzťah, že sú nezávislé:



Ak majú premenné koreláciu na úrovni 0,334, majú zdieľanú variabilitu 11,2%. Netreba pritom opomínať, že zvyšných 88,8% variancie nevieme vysvetliť, hovoríme, že každá premenná má 44,4% unikátnej variability. Znova si to môžeme ilustrovať pomocou kruhov:



Pri reportovaní výsledkov uvádzame hodnoty korelačného koeficientu (t) a hodnotu signifikancie (p) v podobe (r = hodnota korelačného koeficientu, p = hodnota signifikancie). Môžeme uvádzať aj koeficient determinácie R^2.

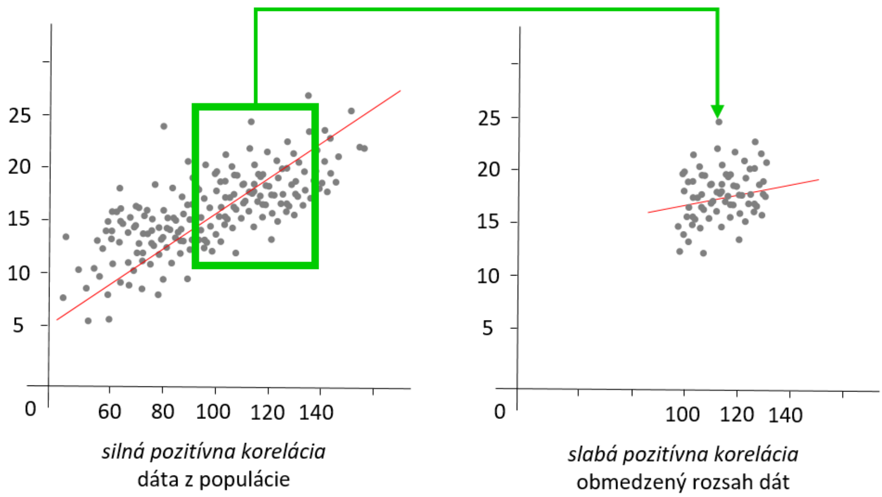
Príklad:

Pomocou Pearsonovej korelačnej analýzy sme odhalili pozitívny, slabý až stredne silný, štatisticky významný vzťah medzi vedeckým myslením a kognitívnou reflexiou (r = 0,334; p < 0,001), pričom variabilita vo vedeckom myslení prispieva 11,2% k variabilite v kognitívnej reflexii (stredný efekt).

9.2 Možné problémy

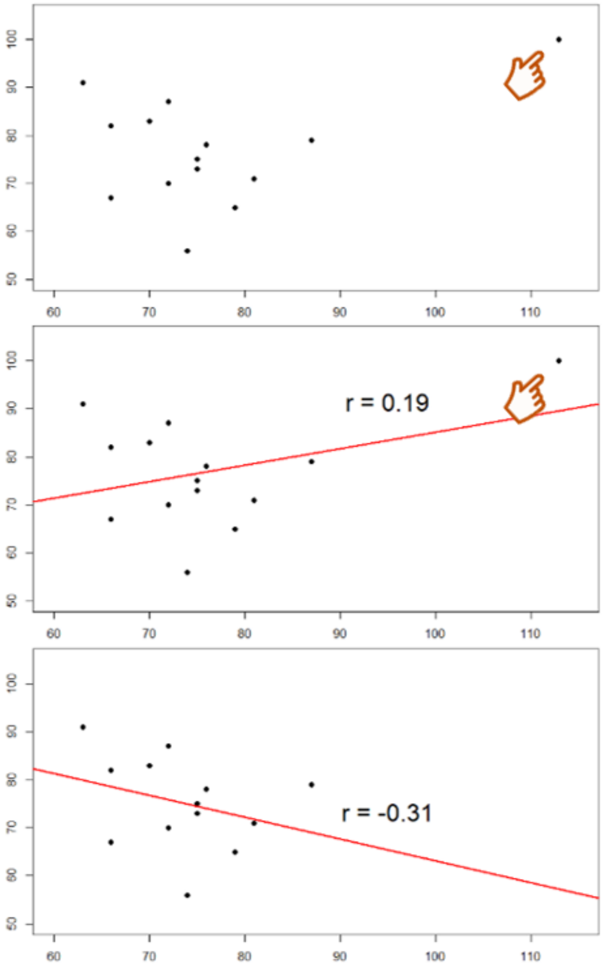
Hoci korelačná analýza sama o sebe nie je náročná, pri výsledkoch by sme mali zohľadňovať minimálne dve veci – rozsah súboru a extrémne hodnoty (*outliers*).

Rozsah súboru je dôležitý, pretože sila korelácie závisí od toho, aká je variabilita sledovaných premenných (Foster et al., 2018). Ak majú naše dáta obmedzenú variabilitu, znižuje to silu korelácie. Čo si ale predstaviť pod variabilitou premenných vzhľadom na rozsah súboru? Predstavte si, že by ste skúmali inteligenciu ľudí a jej vzťah k výkonu v matematickom teste. Výskum by ste realizovali na študentoch gymnázií. Tým, že limituje vzorku na študentov gymnázií obmedzujete variabilitu oboch premenných – dá sa predpokladať, že študenti gymnázií disponujú istou úrovňou inteligencie a matematických zručností a vzhľadom na populáciu bude takáto vzorka predstavovať iba malý výsek informácií.



Obrázok 32 Korelácie pri vzorke bez a s obmedzeným rozsahom

Keď sa pozriete na Obrázok 32, uvidíte, ako môže obmedzený rozsah súboru ovplyvniť silu korelácie. V praxi máte dve možnosti. Môžete dôkladne vyberať výskumný súbor, ktorý bude reprezentovať populáciu, čo však nie je vždy realizovateľné, alebo budete pri interpretácii výsledkov zohľadňovať rozsah súboru a jeho možné implikácie. Ďalšou otázkou, ktorú je potrebné brať do úvahy, sú extrémne hodnoty (*outliers*). Ak je extrémnych hodnôt v dátovom súbore veľa, alebo sú výrazne extrémne, môžu skresľovať výsledky. Extrémna hodnota je taká, ktorá je vzdialená od ostatných hodnôt. Môže sa objaviť ako omyl, náhodná alebo úmyselná chyba, niekedy je však odrazom reality (niekto v súbore reálne vykazuje extrémne hodnoty v porovnaní s ostatnými). Identifikovať pôvod extrémnej hodnoty je náročné, niekedy nemožné. Ak si nie sme istí pôvodom extrémnej hodnoty, je na stole otázka, či nie je lepšie údaj z dátového súboru odstrániť; ak pôvod poznáme, môžeme sa rozhodnúť, že si údaje ponecháme, pretože je odrazom reality. V každom prípade je potrebné uvedomiť si, čo taký extrémny údaj spraví s koreláciou: môže zmeniť silu korelácie (zoslabiť alebo posilniť ju), signifikanciu (zoslabiť alebo posilniť ju), alebo môže zmeniť smer korelácie (z pozitívnej na negatívnu a naopak). Príklady týchto zmien vplyvom extrémnej hodnoty možno vidieť na Obrázku 33.



Obrázok 33 Vplyv extrémnej hodnoty na koreláciu

10 Regresná analýza

Teraz sa pozrieme trochu ďalej za hranice korelačnej analýzy. V predchádzajúcej kapitole v úvode sme spomínali jedno dôležité pravidlo, ktoré si v súvislosti s koreláciami musíme pamätať, a to **korelácia ≠ kauzalita**. Toto pravidlo je veľmi užitočné, nielen na poli štatistiky, ale aj v každodennom živote. Mnoho omylov, chybných názorov, povier, skreslených presvedčení má svoje korene práve v zamieňaní si korelácie a kauzality. Mnoho ľudí si nesprávne vytvára závery o príčinnom vzťahu (kauzalite) o dvoch nesúvisiacich udalostiach iba na základe toho, že sa vyskytujú súčasne, alebo jedna nasleduje druhú. Napríklad, ak si niekto na väčšinu skúšok obliekol ružové bodkované ponožky a tieto skúšky aj úspešne zvládol, na konci skúškového obdobia si povie, že ponožky mu prinášajú šťastie na skúškach. Vyslovil záver o kauzálnom vzťahu – ponožky sú príčinou úspechu na skúškach. Nesporne, medzi nosením ponožiek a zvládaním skúšok je vzťah (korelácia), tento vzťah však nie je kauzálny (daný človek prisúdil zásluhy neprávom ponožkám namiesto svojim vedomostiam a príprave).

Veľmi známy je aj príklad so zmrzlinou a smrťou utopením. Štatistika ukazuje, že čím viac zmrzliny sa v populácii zje, tým viac ľudí sa utopí. To je pozitívna korelácia medzi množstvom skonzumovanej zmrzliny a počtom úmrtí utopením. Znamená to ale, že zvýšená konzumácia zmrzliny je príčinou zvýšeného počtu úmrtí utopením? Existuje medzi týmito premennými kauzálny vzťah? To by sme mohli overiť iba pomocou experimentu. Sledovali by sme u ľudí množstvo zmrzliny, ktoré zjedia a zároveň by sme sledovali počet úmrtí utopením v danej skupine ľudí. Údaje by sme potom podrobili štatistickej analýze, ktorú nazývame *regresná analýza*, a s vysokou pravdepodobnosťou by sme zistili, že vyššie alebo nižšie množstvo skonzumovanej zmrzliny nemá vplyv na počet utopení v sledovanej skupine. To znamená, že medzi týmito premennými nie je kauzálny vzťah, jedna nie je príčinou druhej, a zrejme do hry vstupe iná príčinná premenná. Nebudem vás napínať (zrejme ste na to prišli aj sami) – tou premennou je ročné obdobie, konkrétne leto. V lete sa skonzumuje viac zmrzliny a dochádza aj v vyššiemu počtu úmrtí utopením, keďže ľudia trávia viac voľnočasových aktivít pri a vo vode. Na tomto príklade krásne vidno, že hoci sa dve udalosti, dva javy vyskytujú v rovnakom čase (je medzi nimi korelácia), neznamená to, že je medzi nimi aj kauzálny vzťah. Ak sa pri používaní štatistiky naučíte byť obozretní pri vyslovovaní záverov ohľadom vzťahov (kedy môžeme hovoriť o kauzalite a kedy o korelácii), môže vám to v konečnom dôsledku priniesť mnoho výhod aj v osobnom živote.

Pri odlišovaní korelácie a kauzality treba byť naozaj obozretní. Je jedno, nakoľko logické, zmysluplné alebo očividné sa vám to zdá, korelačná analýza vám NIKDY nepovie NIČ o kauzalite. Jediný spôsob, ako overiť kauzalitu vzťahu je prostredníctvom dobre nadizajnovaného experimentu. Iba pomocou experimentu dokážete kontrolovať (manipulovať) jednu premennú (nezávislú) a sledovať ako sa následne mení druhá sledovaná premenná (závislá).

Vzájomnú závislosť dvoch[19] premenných opisuje regresná analýza, konkrétne *jednoduchá lineárna regresia* (alebo párová regresia). Je podobná ako korelačná analýza, no namiesto toho, ako veľmi premenné spolu súvisia, nám hovorí, ako veľmi sa zmení jedna premenná, ak sa zmení druhá premenná (Dancey & Reidy, 2017). Inými slovami, umožňuje odhadnúť—predikovať—hodnotu závislej premennej podľa toho, ako sa mení nezávislá premenná. Hoci tu hovoríme o kauzálnom vzťahu medzi premennými, nejde o to, že by jedna premenná bola reálnou priamou príčinou druhej premennej, ale o to, že premenné spolu súvisia v prediktívnom zmysle.

V rámci regresnej analýzy, ako ste už zrejme vytušili, veľmi záleží na tom, ktorá premenná je závislá a ktorá nezávislá. Premenná, ktorá je odhadovaná, predikovaná na základe inej premennej, sa nazýva *závislá* premenná (*DV*, *dependent variable*) alebo aj kritérium, označujeme ju *Y*. Premenná, ktorá určuje, predikuje závislú premennú sa volá *vysvetľujúca* premenná (*IV*, *explanatory variable*)[20] alebo aj prediktor, označujeme ju *X* (Dancey & Reidy, 2017; Hanna & Dempster, 2012). Pomocou korelačnej analýzy by sme mohli zistiť aký tesný je vzťah medzi počtom hodín strávených štúdiom a úspešnosťou na skúške, pomocou regresnej analýzy by sme ďalej mohli overiť, či a do akej miery počet hodín štúdia predikuje úspešnosť na skúške. Pri regresnej analýze je dôležité vopred vedieť a určiť, ktorá premenná je prediktor a ktorá závislá (tu sa opierame o teórie, či predchádzajúce výskumy; Hanna & Dempster, 2012).

Lineárna regresná analýza a korelačná analýza sú si veľmi podobné – tak ako pri korelačnej analýze aj pri regresnej analýze pracujeme s bodovým grafom. To znamená, že dáta nanášame na os X (vysvetľujúcu premennú) a na os Y (závislú

---

[19] Pre prípady, kedy chceme skúmať spôsoby, akými sú *viaceré* premenné (vysvetľujúce) vo vzťahu so závislou premennou, využívame *viacnásobnú regresnú analýzu* (tá nie je ale predmetom tejto učebnice).

[20] Označovanie v rámci regresnej analýzy môže trochu mätúce. Žiada sa povedať, že vysvetľujúca premenná – prediktor, je nezávislá premenná (aj sa označuje ako IV z anglického *independent variable*, Keďže však nie vždy ide o „pravú“ nezávislú premenné v zmysle definície experimentálneho dizajnu, uprednostňuje sa označenie prediktor, alebo vysvetľujúca premenná.

premennú). No namiesto pomyselnej čiary (ako pri korelačnej analýze), spravíme—samozrejme pomocou počítačového programu—reálnu čiaru, okolo ktorej sa dáta tesne zhlukujú (teda tak tesne, ako sa len dá). Ide o priamku, ktorá je najlepšia možná (*line of best fit;* Foster et al., 2018) a nazývame ju *regresná priamka* (Miles & Banyard, 2007). Preto hovoríme o lineárnej regresnej analýze. Vzdialenosť každého bodu (údaju) od tejto priamky nazývame chyba (v zmysle odchýlky) alebo rezíduum (*residuals*).

Keď už túto čiaru máme, vieme podľa jej sklonu (označujeme ho b; čím je b vyššie, tým je priamka strmšia) povedať o koľko jednotiek (najčastejšie bodov) sa zmení závislá premenná, ak sa vysvetľujúca premenná zmení napríklad o +1 bod.

Rovnica regresnej priamky je vyjadrená nasledovne:

$$Y = c + b * X$$

kde:

Y – hodnota závislej premennej
c – regresná konštanta, hodnota, pri ktorej sa regresná priamka dotýka osy y, resp. je to predpokladaná hodnota premennej Y v momente, kedy je hodnota premennej $X = 0$
b – regresný koeficient, určuje sklon regresnej priamky
X – hodnota vysvetľujúcej premennej – jej hodnotu využívame na odhad hodnoty závislej premennej

Pri reportovaní výsledkov vo vzorci namiesto X a Y použite názvy premenných, napríklad

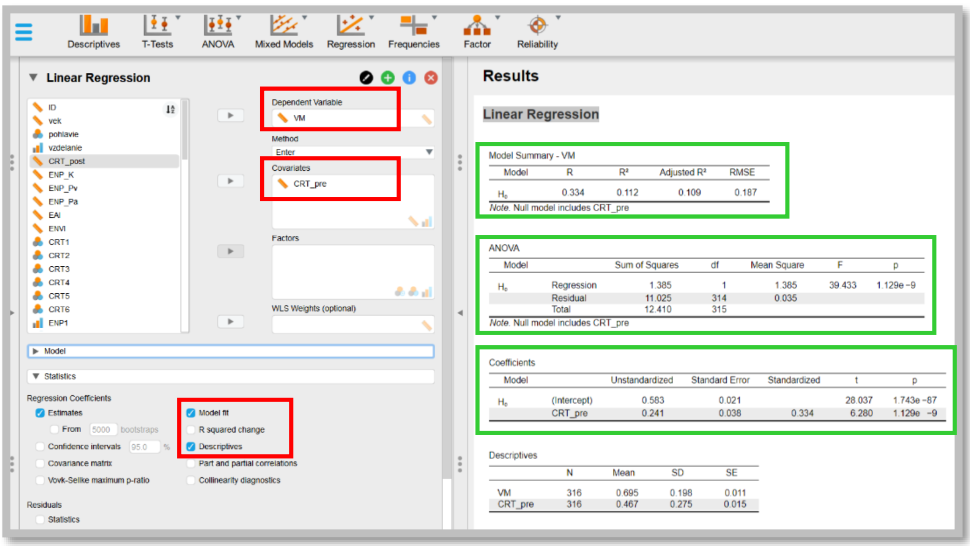
$$skóre\ na\ skúške = 33{,}4 + (1{,}8 * počet\ hodín\ štúdia)$$

To znamená, že pri 10 hodinách štúdia je predpokladané skóre na skúške 51,4 bodov, pri 5 hodinách štúdia je to 42,4 bodov, pri 15 hodinách by to bolo 60,4 bodov. Ak sa $b = 0$, tak priamka je rovnobežná s osou x a teda vysvetľujúca premenná (X) nemôže predpovedať hodnotu závislej premennej Y.

Aby sme mohli robiť validné porovnania zistení z rôznych štúdií, prevádza sa regresný koeficient b, označovaný aj ako neštandardizovaný koeficient (ukazujúci o koľko bodov sa zmení hodnota závislej premennej Y ak sa hodnota vysvetľujúcej premennej X zmení o jeden bod), na ***štandardizované skóre* beta *β***. Štandardizované skóre β vypovedá o koľko štandardných odchýlok sa zmení hodnota závislej premennej Y ak sa hodnota vysvetľujúcej premennej X zmení o jednu štandardnú odchýlku.

10.1 Výpočet lineárnej regresnej analýzy

Pri lineárnej regresii overujeme nulovú hypotézu, že neexistuje signifikantná predikcia závislej premennej na základe vysvetľujúcej premennej, prediktora. Korelačnou analýzou sme zistili stredne silný pozitívny vzťah (r = 0,334; p < 0,001) medzi vedeckým myslením a kognitívnou reflexiou. Keďže kognitívna reflexia ako súčasť analytického myslenia má potenciál predikovať úroveň vedeckého myslenia, tak v našom hypotetickom výskume predpokladáme, že kognitívna reflexia bude predikovať úroveň vedeckého myslenia. V programe JASP si v rámci balíka *Regression* zvolíme možnosť *Linear Regression*. Do okna *Dependent variable* presunieme závislú premennú, teda tú, ktorú chceme vysvetliť (*VM*, vedecké myslenie), a do okna *Covariates* presunieme vysvetľujúcu premennú (prediktor; *CRT_pre*, kognitívna reflexia). Následne v časti *Statistics* zaškrtneme okrem už prednastavených možností aj možnosť *Descriptives*, aby sme mali prehľad o základných parametroch sledovaných premenných (Obrázok 34).



Obrázok 34 Lineárna regresná analýza

Vo výsledkoch vidíme niekoľko tabuliek. Najskôr sa pozrieme na tabuľku *Model Summary*, kde si všímame hodnotu R – koreláciu medzi premennými (0,334 v našom príklade). Ďalej si všímame R^2, teda veľkosť efektu (0,112). Toto sú hodnoty, ktoré získame aj klasickou korelačnou analýzou.

V tabuľke *ANOVA* je dôležitý údaj hodnota F a jej signifikancia p. V našom prípade $F = 39{,}433$ a $p < 0{,}001$. To znamená, že model je štatisticky významný, že kognitívna reflexia predikuje vedecké myslenie so štatistickou významnosťou.

A nakoniec, v tabuľke *Coefficients* identifikujeme údaje do regresnej rovnice. Pre pripomenutie:

$$Y = c + b * X$$

Y – hodnota závislej premennej

c – regresná konštanta (v tabuľke výsledkov nájdete v stĺpci *unstandardized* a v riadku *intercept*)

b – regresný koeficient (v tabuľke výsledkov nájdete v stĺpci *unstandardized* a v riadku s názvom vysvetľujúcej premennej),

X – hodnota vysvetľujúce premennej

Takže, ak v našom príklade chceme odhadnúť, ako sa bude vyvíjať hodnota vedeckého myslenia (*Y*) pri zmenách v premennej kognitívna reflexia (*X*), aplikujeme hodnoty z poslednej tabuľky do vzorca, pričom za *X* dosadíme hodnotu kognitívnej reflexie, pre ktorú chceme odhadnúť hodnotu vedeckého myslenia. Ak vieme, že test kognitívnej reflexie môže dosahovať hodnoty od 0 po 6, tak môžeme zadať do vzorca napr. hodnotu 3 a potom maximálnu hodnotu 6:

Pri skóre 3 v kognitívnej reflexii by skóre vedeckého myslenia malo byť na úrovni 1,306.

$$Y = 0{,}583 + 0{,}241 * 3 = 1{,}306$$

a pri maximálnom skóre 6 v kognitívnej reflexii by skóre vedeckého myslenia malo byť na úrovni 2,029.

$$Y = 0{,}583 + 0{,}241 * 6 = 2{,}029$$

Pri reportovaní výsledkov uvádzame hodnoty F-testu (*F*), hodnotu signifikancie (*p*) a hodnotu koeficientu determinácie R^2 v podobe (*F(*obe hodnoty *df*) = hodnota F-testu, *p* = hodnota signifikancie, R^2 = hodnota koeficientu determinácie). Ideálne uvádzame aj regresnú rovnicu.

Príklad:

Lineárna regregresia odhalila, že úroveň kognitívnej reflexie signifikantne predikuje úroveň vedeckého myslenia ($F(1{,}314) = 39{,}433$, $p < 0{,}001$, $R^2 = 0{,}112$) podľa nasledovnej regresnej rovnice

$$vedecké\ myslenie = 0{,}583 + 0{,}241\ .\ kognitívna\ reflexia$$

Výsledky regresnej analýzy sa uvádzajú v tabuľkovej podobe nasledovne (Tabuľka 6).

Tabuľka 6 Lineárna regresia pre vedecké myslenie ako prediktor kognitívnej reflexie

	B	SE	β	t	p
(Constant)	0,583	0,021		27,037	< 0,001
vedecké myslenie	0,241	0,038	**0,334**	6,280	< 0,001
	$F(1,314) = 39{,}433$, $p < 0{,}001$, $R^2 = 0{,}112$				

11 Testovanie pomocou chí-kvadrátu

V záverečnej kapitole si predstavíme test, ktorý slúži na to, aby sme porozumeli frekvenčnému rozdeleniu jednej nominálnej premennej alebo aby sme preskúmali vzťah medzi dvoma nominálnymi premennými. Ide chí-kvadrát test (χ^2). Chí-kvadrát je test, ktorý sa má použiť, keď hľadáme súvislosť, asociáciu, resp. rozdiely vo frekvenciách nominálnych premenných.

Pripomeňme si, že nominálna úroveň merania premenných je špecifická tým, že údaje na tejto úrovni môžeme opísať iba ich názvom, zaradením do kategórie, napríklad muži – ženy, alebo trieda 1.A, 1.B, 1.C. Pri týchto premenných nevieme vypočítať priemer, štandardnú odchýlku, nevieme určiť poradie, vieme maximálne určiť frekvenciu, teda koľko jednotiek (ľudí) sa nachádza v tej-ktorej kategórii. Tým pádom pri nich nevieme využiť väčšinu štatistických analýz. Niekedy však náš výskumný problém vyžaduje, aby sme analyzovali aj nominálne premenné.

Predstavte si, že máte vzorku 100 žiakov a ich rozdelenie v triedach je nasledovné (Tabuľka 7):

Tabuľka 7 Rozdelenie žiakov v triedach

	1.A	1.B	1.C	spolu
počet žiakov	24	37	39	100

Toto je reálne rozloženie žiakov – v jazyku štatistiky hovoríme o *pozorovanej* frekvencii. Keď by sme chceli overiť, ako veľmi sa toto rozdelenie žiakov líši od toho, aké by malo byť (predpokladané rozloženie), potom využijeme chí-kvadrát test. Otázka v prvom kroku je, ako zistíme, aké by malo byť optimálne rozloženie? Intuícia nám nahovára, že v triedach by mal byť rovnaký počet žiakov. Ak máme iba jednu premennú, je jednoduché vypočítať toto optimálne rozloženie: celkový počet žiakov vydelíme počtom tried; vo všeobecnosti sa využíva vzorec

$$E = \frac{N}{C}$$

kde E je očakávaný počet (*expected value*), N je celkový počet jednotiek (ľudí v našom prípade), C je počet kategórií.

$$E = \frac{100}{3} = 33{,}33; \textit{teda po } 33 \textit{ žiakov v } 2 \textit{ kategóriách a } 34 \textit{ žiakov v } 1 \textit{ kategórii}$$

V kontexte hypotéz sú *očakávané* frekvencie také frekvencie, ktoré by sa vyskytli v realite, ak by platila nulová hypotéza. Teda testujeme nulovú hypotézu, že neexistuje žiadna súvislosť medzi kategorickými premennými.

V prípade tried sme mali v príklade 3 kategórie (3 triedy), ale môže sa nám stať, že chceme brať do úvahy aj ďalšiu premennú, napríklad pohlavie (chlapec – dievča). V tomto momente sa počet tried zvýši na 3 x 2 = 6 kategórií (Tabuľka 8).

Tabuľka 8 Rozdelenie žiakov v triedach – dievčatá

	1.A	1.B	1.C	spolu
dievčatá	11	15	28	54
chlapci	13	22	11	46
spolu	24	37	39	100

Výpočet očakávaných frekvencií pri dvoch a viac premenných je trošku iný, ako pri jednej premennej. Hoci intuitívne by niekto mohol využiť predchádzajúci vzorec

$$E = \frac{100}{6} = 16{,}67\ (teda\ po\ 17\ žiakov\ v\ 4\ kategóriách\ a\ po\ 16\ žiakov\ v\ 2\ kategóriách),$$

nie je to korektný postup. Pri výpočte očakávaných frekvencií pri viacerých premenných musíme brať do úvahy sumárne početnosti v riadkoch a v stĺpcoch. Výpočet vyzerá takto:

$$očakávaný\ model = \frac{spolu\ v\ riadku * spolu\ v\ stĺpci}{spolu}$$

chlapci v 1.A	=	(46 * 24) / 100 = 11
chlapci v 1.B	=	(46 * 37) / 100 = 17
chlapci v 1.C	=	(46 * 39) / 100 = 18
dievčatá v 1.A	=	(54 * 24) / 100 = 13
dievčatá v 1.B	=	(54 * 37) / 100 = 20
dievčatá v 1.C	=	(54 * 39) / 100 = 21

Frekvenčná tabuľka by vyzerala nasledovne (Tabuľka 8):

Tabuľka 9 Rozdelenie žiakov v triedach

		1.A	1.B	1.C	spolu
dievčatá	pozorované	11	15	28	54
	očakávané	*13*	*20*	*21*	*54*
chlapci	pozorované	13	22	11	46
	očakávané	*11*	*17*	*18*	*46*
spolu	pozorované	24	37	39	100
	očakávané	*24*	*37*	*39*	*100*

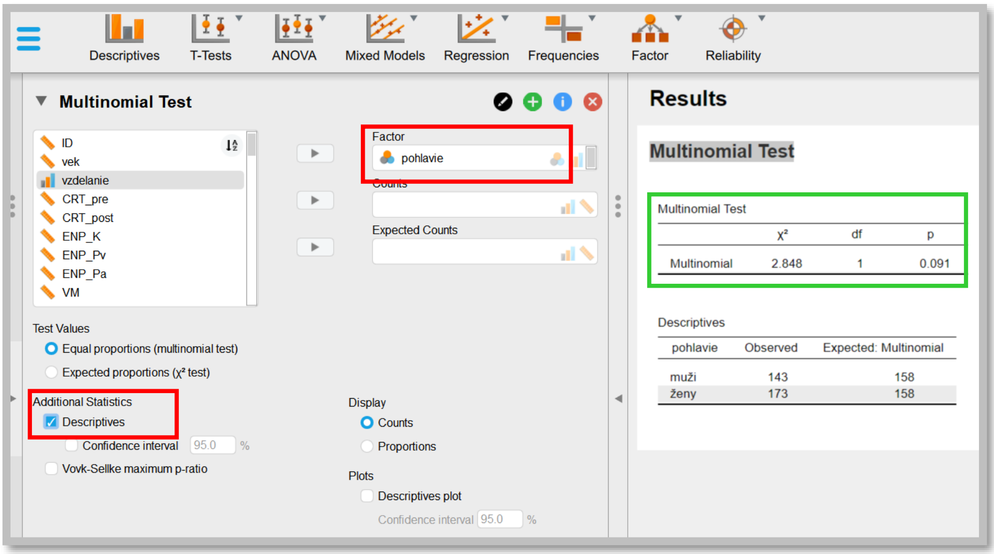
A teraz prichádza moment, kedy by sme chceli posúdiť, nakoľko je očakávané a pozorované rozloženie odlišné. Keď máme iba jednu premennú (napríklad pohlavie, alebo triedu), využijeme chí-kvadrát test. V prípade, že hľadáme zhodu, asociáciu medzi dvoma a viacerými premennými, z ktorých má každá dve alebo viac kategórií, využijeme *chí-kvadrát test nezávislosti*.

Podmienkou výpočtu chí-kvadrát testu je, že premenné musia byť merané na nominálnej úrovni a že každá premenná by mala obsahovať dve alebo viac nezávislých kategoriálnych skupín. Navyše platí, že chí-kvadrát test je validný iba vtedy, ak máte primerane veľkú vzorku, to znamená, že menej ako 20% kategórií (buniek) nemá očakávanú početnosť menšiu ako 5 a žiadna kategória nemá očakávanú početnosť menšiu ako 1.

11.1 Výpočet chí-kvadrát testu

Chí-kvadrát test testuje, či je alebo nie je frekvencia vzorky v rámci nominálnej premennej štatisticky odlišná od hypotetickej očakávanej frekvencie. Testujeme teda nulovú hypotézu, že frekvencie vzorky sa rovnajú očakávaným frekvenciám.

Ak máme iba jednu premennú, chí-kvadrát test vypočítame tak, že v rámci sekcie *Frequencies* zvolíme test *Multinomial Test*. Do okna *Factor* presunieme nominálnu premennú, ktorú chceme testovať (napríklad *pohlavie*). V rámci dodatočných štatistík (*Additional Statistics*) môžeme zvoliť aj možnosť deskriptívnych štatistík (*Descriptives*), aby sme videli, aké sú očakávané hodnoty v rámci sledovaných kategórií (Obrázok 35).



Obrázok 35 Chí-kvadrát test

V deskriptívnej tabuľke vidíme pozorované početnosti (*Observed*) a očakávané početnosti (*Expected: Multinomial*). V našom príklade je vo vzorke 143 mužov a 173 žien, očakávaný počet mužov je 158, rovnako aj očakávaný počet žien 158. Či je tento rozdiel medzi pozorovanými a očakávanými početnosťami významný alebo nie si pozrieme v prvej tabuľke (*Multinomial Test*). Hodnota chí-kvadrát testu je χ^2 = 2.848, a tento výsledok je nevýznamný, pretože p = 0,091.

Pri reportovaní výsledkov chí-kvadrát testu uvádzame hodnoty chí (χ^2), stupne voľnosti (*df*) a hodnotu signifikancie (p).

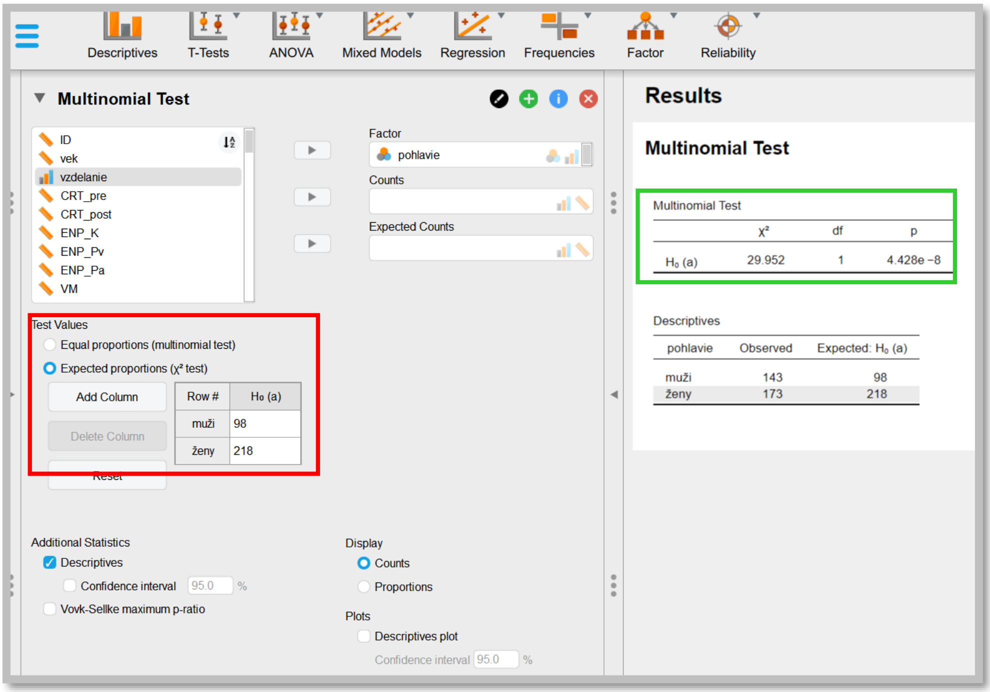
Príklad:

Chí-kvadrát test nepreukázal signifikantný rozdiel medzi pozorovanou frekvenciou mužov a žien vo výberovom súbore v porovnaní s očakávanou frekvenciou (χ^2 = 2.848; p = 0,091).

11.1.1 Chí-kvadrát test dobrej zhody

V niektorých prípadoch sa môžeme ocitnúť v situácii, kedy chceme porovnať nejaké vopred známe rozloženie s nami nameraným rozložením. Napríklad by sme vedeli, že v oblasti, kde sme zbierali dáta, je väčší podiel dievčat ako chlapcov v pomere 69 ku 31, a chceme porovnať naše rozloženie dát s uvedeným pomerom.

Využijeme *chí-kvadrát dobrej zhody* (*goodness-of-fit test*) a to tak, že zvolíme ten istý postup ako v pri výpočte chí-kvadrát testu pre jednu premennú, iba sekcii *Test Values* zvolíme možnosť *Expected proportions* (χ^2 *test)*. Objaví sa tabuľka v rámci ktorej zadáme hodnoty, ktoré poznáme. Ak platí, že v nami sledovanej oblasti je 69% dievčat a vieme, že naša vzorka má 316 participantov, tak vieme, že máme očakávať 218 dievčat a 98 chlapcov. Tieto hodnoty uvedieme do tabuľky a pozrieme sa na výsledky (Obrázok 36).



Obrázok 36 Chí-kvadrát test dobrej zhody

Vidíme, že hodnota chí-kvadrát testu je χ^2 = 29.952, a tento výsledok je signifikantný pretože p < 0,001. Pri reportovaní výsledkov chí-kvadrát testu uvádzame hodnoty chí (χ^2), stupne voľnosti (*df*) a hodnotu signifikancie (*p*).

Príklad:

Chí-kvadrát test preukázal signifikantný rozdiel medzi pozorovanou frekvenciou mužov a žien vo výberovom súbore v porovnaní s očakávanou frekvenciou opierajúcou sa o pomer mužov a žien v populácii (χ^2 = 29.952; p < 0,001).

11.1.2 Chí-kvadrát test nezávislosti

Chí-kvadrát test vieme využiť aj v situáciách, keď máme viacero nominálnych premenných, ktoré zaraďujú našich participantov do kategórií – napríklad ako sme si ukázali v Tabuľkách 7 – 9, kedy sme participantov kategorizovali podľa triedy a pohlavia. Keď chceme porovnávať očakávané početnosti s nameranými početnosťami v prípade, že sú participanti kategorizovaní podľa viacerých premenných využívame *chí-kvadrát test nezávislosti*.

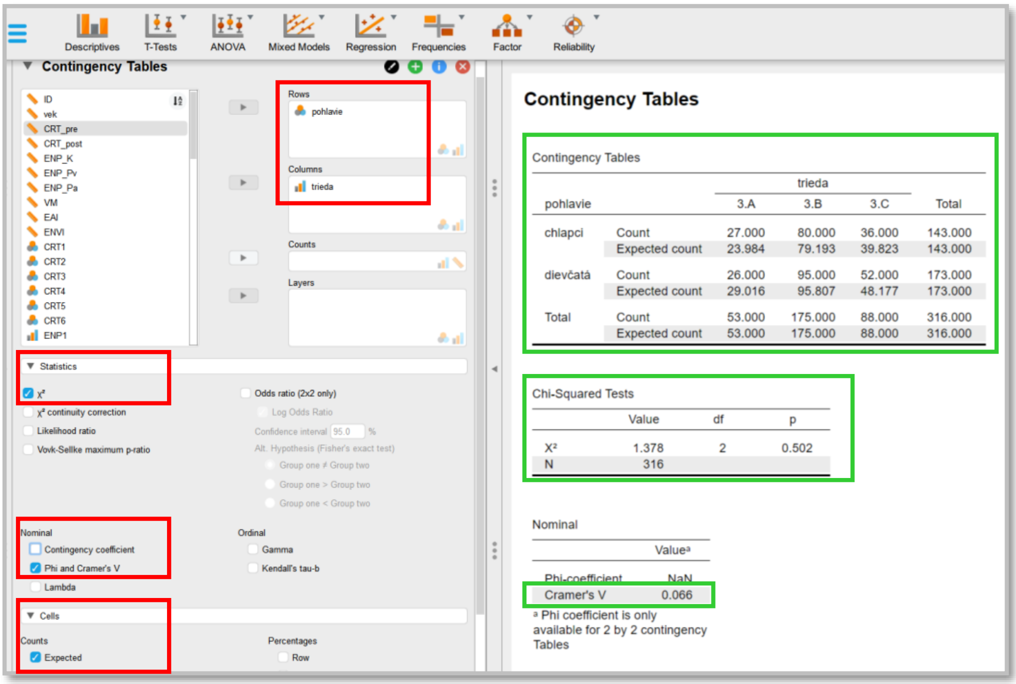
Pre výpočet tohto typu chí-kvadrát testu, keď máme viac premenných, si v sekcii *Frequencies* zvolíme *Contingency Tables* (Obrázok 37). Do okna *Rows* (riadky) presunieme premennú, ktorú chceme mať v riadkoch a do okna *Columns* (stĺpce) zas tú, ktorú chceme mať v stĺpcoch. V našom príklade do riadkov vložíme premennú pohlavie a do stĺpcov premennú trieda. V sekcii *Statistics* je predvolená možnosť χ^2 na výpočet chí-kvadrát testu[21]. V Sekcii *Cells* zvolíme možnosť *Expected*, čím sa v tabuľke zobrazia očakávané hodnoty pre každú bunku (kombináciu premenných) a zvolíme aj percentá (*Percentage)* pre riadky (*Row*) a stĺpce (*Column*).

Pre overenie veľkosti efektu sa pri chí-kvadrát teste využíva *Cramerovo V* – miera relatívnej sily asociácie medzi dvoma premennými. Cramerovo *V* vypočítame zvolením možnosti *Phi and Cramer's V* v sekcii *Nominal*. Koeficient sa pohybuje od 0 (žiadna asociácia) do 1 (dokonalá asociácia), pričom:

- Cramerovo *V* približne do 0,2 – slabá asociácia
- Cramerovo *V* približne od 0,2 do 0,6 – stredne silná asociácia
- Cramerovo *V* približne nad 0,6 – silná asociácia

Vo výsledkovej časti najskôr pozrieme na tabuľku *Contingency Tables*, v ktorej vidíme početnosti a percentá (pre riadky a stĺpce) v rámci jednotlivých buniek, kombinácií premenných. Napríklad, vidíme, že v 3.A je 27 chlapcov (riadok *Count*) a očakávaná početnosť je po zaokrúhlení 24 chlapcov. Vidíme aj, že v 3.C je viac dievčat ako chlapcov. Otázka je, či je spojitosť medzi pohlavím a triedou významná. Na to si pozrieme tabuľku *Chi-Squared Tests*, v ktorej vidíme, že hodnota χ^2= 1,378 a hodnota p = 0,502 (Cramerovo *V* je na úrovni 0,066, teda takmer žiadna asociácia). Spojitosť medzi sledovanými premennými nie je štatisticky významná.

[21] Ak sa stane, že by minimálne v jednej bunke bola početnosť nižšia ako 5, potom na výpočet chí-kvadrát testu zvolíme v sekcii *Statistics* možnosti *χ^2 continuity correction*. A ak máme celkovo malú vzorku (menej ako 30), potom zvolíme *Vovk-Sellke maximu p-ratio*.



Obrázok 37 Chí-kvadrát test nezávislosti

Pri reportovaní výsledkov chí-kvadrát testu uvádzame hodnoty chí (χ^2), stupne voľnosti (*df*), hodnotu signifikancie (*p*), hodnota Cramerovho *V*.

Príklad:

Chí-kvadrát test nepreukázal signifikantný rozdiel medzi pozorovanou frekvenciou mužov a žien v jednotlivých triedach v porovnaní s očakávanou frekvenciou (χ^2 = 1.378; p = 0,502; V = 0,066).

12 Prehľad analýz

Dostali ste sa na záver učebnice. V tejto chvíli by ste mali ovládať základné štatistické analýzy, u ktorých je vysoká pravdepodobnosť, že ich využijete na začiatku svojej vedeckej dráhy, alebo prinajmenšom pri vypracovávaní svojej záverečnej práce.

Pre uľahčenie rozhodovania, ktorý test je vhodné pre aké dáta použiť, prikladám na záver Tabuľku 10. V Tabuľke 11 nájdete prehľad jednotlivých koeficientov a ich interpretácií, ku ktorým sa dopracujete prostredníctvom štatistických analýz uvedených v tejto učebnici a v Tabuľke 12 je prehľad mier efektov a ich vzájomné porovnanie.

Dúfam, že vám predložená učebnica poskytla náhľad do základov štatistických analýz, že ste sa zbavili strachu zo štatistiky a že možno niektorí z vás objavili jej čaro a užitočnosť (nielen pre vedu a výskum).

Tabuľka 10 Prehľad analýz

skúmanie vzťahov	
nominálna úroveň merania	*chí-kvadrát test*
ordinálna úroveň merania	*Spearmanova korelačná korelácia*
kardinálna úroveň merania	*Pearsonova korelačná analýza*

skúmanie rozdielov			
2 skupiny / merania	nezávislé skupiny	parametrické	*t-test pre 2 nezávislé výbery*
		neparametrické	*Mann-Whitneyho U-test*
	závislé skupiny	parametrické	*t-test pre 2 závislé výbery*
		neparametrické	*Wilcoxonov test*
3 a viac skupín / meraní	1 závislá premenná	parametrické	*ANOVA pre jeden faktor*
		neparametrické	*Kruskal-Wallisov test*

Tabuľka 11 Prehľad koeficientov a ich interpretácia

názov koeficientu	skratka	možný rozsah[22]	interpretácia
signifikancia	p	< 0,05	signifikantný výsledok
		>0,05	nesignifikantný výsledok
korelačný koeficient	r	< 0,1 (-0,1)	žiadny vzťah
		0,1 (-0,1)	slabý vzťah
		0,3 (-0,3)	stredne silný vzťah
		0,5 (-0,5)	silný vzťah
vnútorná konzistencia	α, ω	< 0,7	slabá vnútorná konzistencia
	KR_{20}, KR_{21}	> 0,7	dobrá vnútorná konzistencia
veľkosť efektu	*Cohenovo d*	0,2 – 0,5	malý efekt
		0,5 – 0,8	stredný efekt
		nad 0,8	veľký efekt
	R^2	> 0,01	malý efekt
		0,09 – 0,25	stredný efekt
		> 0,25	veľký efekt
	Rank-Biserial Correlation	< 0,1	žiadny efekt
		0,1	slabý efekt
		0,3	stredne silný efekt
		0,5	silný efekt
	η^2	0,01	malý efekt
		0,06	stredný efekt
		0,14	veľký efekt
	Cramerovo V	< 0,2	slabá asociácia
		0,2 – 0,6	stredne silná asociácia
		> 0,06	silná asociácia

[22] Hranice jednotlivých koeficientov sú skôr orientačné, v literatúre sa môžete stretnúť aj s inými hodnotami.

Tabuľka 12 Najčastejšie miery efektu – porovnanie

	Cohenovo *d*	korelačný koeficient *r*	koeficient determinácie R^2
Veľký efekt	2,0	0,707	0,500
	1,9	0,689	0,474
	1,8	0,669	0,448
	1,7	0,648	0,419
	1,6	0,625	0,390
	1,5	0,600	0,360
	1,4	0,573	0,329
	1,3	0,545	0,297
	1,2	0,514	0,265
	1,1	0,482	0,232
	1,0	0,447	0,200
	0,9	0,410	0,168
	0,8	0,371	0,138
	0,7	0,330	0,109
	0,6	0,287	0,083
Stredný efekt	0,5	0,243	0,059
	0,4	0,196	0,038
	0,3	0,148	0,022
Malý efekt	0,2	0,100	0,010
	0,1	0,050	0,002
	0,0	0,000	0,000

13 Referencie

Aron, A., Aron, E., & Coups, E. (2014). *Statistics for psychology*. Pearson.

Cohen, J. (1988). *Statistical Power for Behavioral Sciences*. Academic Press.

Coolican, H. (2019). Research methods and statistics in psychology. In *Research Methods and Statistics in Psychology*. Routledge. https://doi.org/10.4324/9780203769836

Dancey, C., & Reidy, J. (2017). *Statistics without maths for psychology*. Pearson.

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*(3), 399–412. https://doi.org/10.1111/bjop.12046

Ferjenčík, J. (2000). *Úvod do metodológie psychologického výskumu*. Portál.

Foster, G. C., Lane, D., Scott, D., Hebl, M., Guerra, R., Osherson, D., & Zimmer, H. (2018). *An Introduction to Psychological Statistics*. Open Educational Resources Collection. https://irl.umsl.edu/oer/4

George, D., & Mallery, M. (2010). *SPSS for Windows Step by Step: A Simple Guide and Reference*. Pearson.

Glantz, S. A., Slinker, B. K., & Neilands, T. B. (2016). *Primer of Applied Regression & Analysis of Variance* (3rd ed.). McGraw-Hill Education.

Goss-Sampson, M. A. (2020a). *Statistical analysis in JASP. A guide for students* (4th ed.). https://doi.org/10.6084/m9.figshare.9980744

Goss-Sampson, M. A. (2020b). *Statistical analysis in JASP. A guide for students (JASP v 0.14)*. https://doi.org/10.6084/m9.figshare.9980744

Hanna, D., & Dempster, M. (2012). *Psychology statistics for dummies*. Wiley.

Kalina, M., Bacigál, T., & Schiesslová, A. (2010). *Základy pravdepodobnosti a matematickej štatistiky*. Slovenská technická univerzita.

Lichner, V. (2020). *Základy štatistiky v sociálnych vedách I. Teoretické východiská a deskriptívna analýza*. Univerzita Pavla Jozefa Šafárika v Košiciach, Filozofická fakulta, Katedra sociálnej práce.

McLaughlin, A. C., & McGill, A. E. (2017). Explicitly Teaching Critical Thinking Skills in a History Course. *Science Education, March,* 1–13. https://doi.org/10.1007/s11191-017-9878-2

Miles, J., & Banyard, P. (2007). *Understanding and using statistics in psychology. A* *practical introduction or, how I came to know and love the standard error*. Sage Publications.

Pathak, R. P. (2011). *Statistics in education and psychology*. Pearson.

Rimarčík, M. (2007). *Štatistika pre prax*. Rimarčík. www.rimarcik.com

Ritomský, A. (2015). *Metodologické a metodické otázky kvantitatívneho výskumu*. IRIS.

Ritomský, A. (2016). *Metodológia projektovania psychologického výskumu*. Aleš Čeněk.

Tomšik, R. (2017). *Kvantitatívny výskum v pedagogických vedách. Úvod do metodológie a štatistického spracovania*. Pedagogická fakulta, Univerzita Konštantína Filozofa v Nitre.

Názov: Štatistika pre začiatočníkov.
Základy štatistických analýz pre študentov učiteľstva

Autor: Eva Ballová Mikušková

Recenzenti: Mgr. Michal Kohút, PhD.

Prof. Peter Halama, PhD.

Cover design: Krause Mental Care Group

Vydanie: prvé

Rok: 2021

Rozsah: 100 s.

Vydavateľ: PF UKF v Nitre

Tlač: DMC, s.r.o.

Náklad: 120 ks

ISBN 978-80-558-1823-8 (printová verzia)

ISBN 978-80-558-1824-5 (online verzia)