

Studentův t-test

t-test pro dva nezávislé výběry

Pro použití **t-testu** je zapotřebí splnit určité podmínky:

- Normální rozložení souboru (nad 100 respondentů však začíná platit „centrální limitní věta“)
- Podobný rozptyl v obou porovnávaných podsouborech (nad 100 respondentů už to však nemusíme řešit)

KDY MŮŽEME POUŽÍT PARAMETRICKÉ TESTY?

1. Zjišťování normálního rozdělení proměnných:

- (A) V nabídce „**Statistics**“ zvolíme: **Šikmost** (*skewness*) a **špičatost** (*kurtosis*) – když obojí vydělíme jejich standardními chybami (*Std. Error*), tak by výsledek v absolutní hodnotě neměl překročit hodnotu 1,96. Je-li vyšší, nejde o normálně rozložená data. (Při početnějším souboru $N > 200$ se stačí vejít do hodnoty 2,58.)
- (B) **Kontrola normálního rozložení proměnné** („okometricky“) - V JASP nahoře na liště: „**Descriptives**“:
- Do pravého horního okénka („**variables**“) přesuneme kontrolovanou proměnnou (získáme její průměr a směrodatnou odchylku)
 - Níže nalevo v nabídce otevřeme „**plots**“ a pod tím napravo označíme „**Distribution plots**“. Dostaneme histogram. Odpovědi by měly rozhazovat do většiny variant, neměly by být jen v jedné části grafu.
- (C) **Pro zjištění normality** můžeme použít také **Shapiro-Wilkův test**, který je v nabídce „**Statistics**“.
- Zpravidla však pracujeme s položkami z celého dotazníku, které sčítáme (spíš zprůměrujeme) jako celek a výsledný graf se pak podobá Gaussově křivce mnohem víc.

V případě nenormálního rozdělení závisle proměnné budeme používat místo parametrického Studentova t-testu neparametrický **Mann Whitney U-test** (tj. pro pořadové/ordinální proměnné a pro kardinální/spojité proměnné s velmi nenormálním rozdělením hodnot). Výpočet U se objeví v tabulce s t-testem ve druhém řádku. U Mann Whitneyeho testu se místo

hodnoty „t“ do diplomky uvádí hodnota „U“, kterou najdeme ve druhém řádku v sloupečku „Statistic“.¹

2. **Má-li závisle proměnná normální rozdělení hodnot**, provádíme dále **kontrolu podobnosti rozptylů v obou porovnávaných podsouborech**. V JASP to provádíme už **rovnou při výpočtu t-testu**. Zvolíme nahoře „t-tests“ => „Independent samples t-test“:

- Do pravého horního okénka („Variables“) přesuneme kontrolovanou proměnnou
- Do okénka pod tím („Grouping Variable“) přesuneme proměnnou, která definuje porovnávané podsoubory (při zkoumání rozdílů mezi muži a ženami to bude „pohlaví“)
- Níže vpravo zaškrtneme „equality of variances“ – **pokud vyjde výpočet Leveneho F jako nevýznamný (hodnota „p“ je vyšší než 0,05), můžeme použít výpočet t-testu v horní tabulce**. Pokud by byl výsledek statisticky významný, bylo by v případě malých souborů vhodnější místo t-testu použít **Welchův test**, který je robustní i při porušení předpokladu shody rozptylů (tj. SD).
- Chceme-li vědět o rozdílech mezi sledovanými skupinami a jejich podskupinami víc, zadáme ještě „location parameter“ a díváme se na „Mean difference“. Jde o kódování vzestupné, takže např. 0 versus 1 (tj. kladná hodnota znamená, že první skupina má vyšší průměr než ta druhá).
- Pro interpretaci a popis dat je nutné říci si i o popisnou statistiku s průměry a směrodatnými odchylkami „Descriptive statistics“.²
- V nabídce nalevo je i možnost „Descriptives“, díky níž získáme popisná data pro oba z porovnávaných podsouborů (průměry/means a směrodatné odchylky/SD)

Pokud nevíme, jestli použít parametrický t-test či neparametrický Mann Whitneyův U-test, je na místě provést oba. Když vyjde významnost jen v t-testu, ale nevyjde v Mann Whitneyově testu, pak se řídíme tímto neparametrickým testem (a přijímáme závěr, že se oba podsoubory statisticky významně neliší).

Párový t-test

Používá se, pokud **porovnáváme jednu skupinu respondentů ve stejném dotazníku naměřeném ve dvou časových obdobích** (např. *výsledky ve stejném znalostním testu u školní*

¹ Ve statistickém software SPSS, který mají studenti k dispozici na katedře učitelství, se při výpočtu t-testu vytvoří dva výpočty. Druhý z vypočítaných t-testů je právě pro případ, že Leveneho F test vyjde statisticky významně, tedy že se rozptyly obou podsouborů významně liší (výpočet je pod názvem „Equal variances not assumed“).

² **Je rozdíl, zjištěný mezi podskupinami (tj. odlišnosti mezi průměry a SD), výrazný, nebo malý?** Na to odpovídá některá z hodnot „velikosti účinku“ (Effect Size). U t-testů se používá „Cohenovo d“ (kolem 0,5 by už šlo o střední významnost, ale my máme jen 0,15). Dá se vypočítat i intervaly spolehlivosti pro dolní a horní hranici – kde se nachází 95 % souboru - „Confidence Interval“ (95% CI). Interpretujeme to jednak k ve vztahu ke Cohenovu „d“ a také ve vztahu k dolní hranici „CI“.

třídy před výukovým experimentem a po něm, nebo klima třídy (či postoje k šikaně) před stmelovacím kurzem a po něm, nebo dotazník životní spokojenosti při příchodu do domova pro seniory a po nějaké době pobytu apod. Měří se takto také účinky terapie před zahájením a po nějaké době trvání apod.)

V tomto případě musíme do Excelu přepsat číselné údaje od daného respondenta z obou dotazníků/testů na jeden řádek – Excelový list s daty pak bude mít dvojnásobný počet sloupců a bude tím pádem širší.

Výpočet párového t-testu v JASP: Zvolíme nahoře „**t-tests**“ => „**Paired Samples t-test**“:

- Do pravého horního okénka („**Variables**“) přesuneme kontrolovanou proměnnou v prvním čase měření i ve druhém čase měření.
- V nabídce nalevo využijeme možnost „**Descriptives**“ (v nabídce vlevo uprostřed), díky níž získáme popisná data pro oba z porovnávaných podsouborů (průměry/**means** a směrodatné odchylky/**SD**)

Nevýznamná hodnota „**p**“ (více než 0,05) znamená, že se hodnoty neliší a že v čase nedošlo k posunu ve sledované proměnné. Je-li podle „**p**“ rozdíl mezi prvním a druhým měřením významný, znamená to, že se sledovaná proměnná změnila. Získáním hodnot průměrů v možnosti „**Descriptives**“ zjistíme, zda sledovaná proměnná narostla, nebo se zmenšila.

ANOVA (a WELCHŮV či BROWN-FORSYTŮV TEST)

Pokud porovnáváme více než dva podsoubory, nemůžeme provést jednotlivé t-testy, protože při mnoha výpočtech bychom riskovali, že neoprávněně přijmeme alternativní hypotézu.

Při tomto výpočtu potřebuju jako **závisle proměnnou**: kardinální/spojitou, nebo ordinální/pořadovou (= „**dependent variable**“), dále proměnnou třídící, resp. **nezávisle proměnnou**: buď nominální, nebo ordinální, složenou z několika málo kategorií (= „**fixed factors**“).

Opět platí tytéž podmínky jako u t-testu:

- (1) Nejprve **kontrolujeme normalitu rozložení hodnot závisle proměnné**. Pokud by závisle proměnná byla příliš vychýlená od normálního rozdělení, volíme výpočet **Kruskall-Walis**, který najdeme v ANOVA pod nabídkou „**Nonparametrics**“.
- (2) Nyní musíme **určit, zda jsou rozptyly v podskupinách shodné, nebo ne**, abychom věděli, zda můžeme použít výsledek získaný parametrickou statistikou. Pod „**Assumption checks**“ jsou testy pro toto ověření => „**Homogeneity test**“. Je-li výsledek významný, liší se rozptyly mezi zkoumanými podskupinami. Měli bychom tedy místo

původní parametrické hodnoty Fischerova F-testu použít buď **Brown-Forsythe test**, nebo **Welchův test**. Výsledek vychází významně i zde, takže se některé ze sledovaných podskupin statisticky významně odlišují. Hodnota p „na minus něco“ znamená že je ta hodnota úplně minimální, tj. že mezi alespoň některými skupinami existuje statisticky významný rozdíl.

INTERPRETACE VÝSLEDKŮ:

Abychom mohli výsledky interpretovat, **musíme si říct o popisnou statistiku** („**Descriptive statistics**“) a podívat se na průměry.

Abychom zjistili, **které z podskupin jsou vzájemně statisticky významně odlišné**, použijeme výpočet „**post hoc tests**“. Do otevřeného okénka přeneseme zkoumanou proměnnou. Při shodě rozptylů používáme defaultně nastavený „**test**“. Máme-li neshodné rozptyly (což je náš případ), používáme test „**Games-Howell**“. ³

ANOVA umožňuje do výpočtu zahrnout více než jednu nezávisle proměnnou. Mohli bychom např. zkoumat, zda „spravedlivost“ požadují ve stejné míře žáci a studenti různých tříd a zároveň i všechny dívky a chlapci. V tom případě do horní kolonky „**Fixed factors**“ zadáme obě dvě proměnné (pohlaví i příslušnost ke školní třídě). Ve výsledcích dostaneme výpočty pro obě proměnné i pro jejich **interakci** (z ní se můžeme dozvědět za určitých okolností např. to, že spravedlivost učitele je extrémně důležitá pro chlapce z některých konkrétních tříd).

- Tento výpočet interakce můžeme zrušit. Interakce je vyjádřena proměnnou: „**pohlaví * třída**“. V nabídce „**Model**“ ji odstraníme tím, že tuto proměnnou vymažeme z pravého okénka.

ANCOVA

Jde o analýzu variance/rozptylu (ANOVA), kdy kombinujeme nominální/kategoriální nezávisle proměnné s nezávisle proměnnou spojitou/kardinální (té se říká „kovariát“). ANCOVA je tedy kombinací ANOVY s lineární regresí.

- mezi nezávisle proměnné se zařazuje proměnná nominální/kategoriální (kolonka „**fixed factors**“), ale i proměnná spojitá/kardinální (kolonka „**covariates**“). Výsledek pak zachycuje, jakou hrají tyto nezávisle proměnné roli u proměnné závislé, když je sledujeme obě najednou. Je jasné, že zahrnování více proměnných do výpočtu přináší

³ Zajímá-li nás i reálná **velikost účinku** (**Effect size**), tedy to, kolik procent variability závisle proměnné vysvětluje zvolená nezávisle proměnná (tj. fixed factor), můžeme zvolit některý z těchto testů: buď „**éta na druhou**“ (η^2), nebo „**omega na druhou**“ (ω^2).

přesnější výsledky, protože lépe odpovídají naší realitě. V běžném životě na nás také působí mnohem víc vlivů než jen jeden.

- ANCOVA umožňuje i výpočet s více nezávisle proměnnými, jak nominálními, tak i spojitými.

KORELAČNÍ VÝPOČTY (SOUVISLOSTI MEZI PROMĚNNÝMI)

Výpočty korelací používáme pro zjišťování, zda mezi dvěma (a více) proměnnými existuje významný lineární vztah. Tento vztah se může podobat přímé úměrnosti (a pak očekáváme pozitivní korelační koeficient), nebo nepřímé úměrnosti (pak by šlo o zápornou hodnotu).

Klikneme na: „Regression“ => „Correlation“. Do horního okénka přesuneme obě proměnné, jejichž vzájemný vztah chceme zkoumat.⁴ Získáváme výpočet parametrického **Pearsonova korelačního koeficientu** (r) a pod ním hodnotu významnosti p .⁵

Korelační vztahy mezi proměnnými je možné si rozkreslit v grafu: „Plots“ => „Scatter plot“. Uvidíme jednotlivé „respondenty“ v bodech, kde se v prostoru nacházejí na osách x a y (=nezávisle a závisle proměnné).

Pokud korelujeme víc proměnných, můžeme si říct o schéma „Heatmap“, které ukáže Pearsonovy hodnoty mezi každými z nich. Podle intenzity barvy daného políčka znázorní i těsnost korelačního vztahu mezi jednotlivými dvojicemi proměnných, přičemž modrá značí kladný korelační vztah a červená vztah záporný.

V případě pořadových/ordinálních proměnných a také u kardinálních/spojitých proměnných s výrazně nesymetrickým (nenormálním) rozložením hodnot volíme výpočty pořadových koeficientů: **Spearmanův** (ρ), nebo **Kendallův** (τ). Za těmito účely i zde testujeme normalitu rozložení těchto dat (pomocí „Assumption checks“ => „Shapiro“ pro dvě proměnné – „pairwise“).

⁴ Dolní okénko „condition on“ (=„v závislosti na“) slouží pro případ výpočtu parciálních korelací. To v běžných pracích příliš nevyužíváme.

⁵ **Velikost účinku (Effect size) v případě korelací** vyplývá z hodnoty korelačního koeficientu. Z něj se počítá tzv. koeficient determinace: R^2 . Jde o druhou mocninu hodnoty r vynásobenou hodnotou **100**. Tato hodnota uvádí, kolik procent variability jedné proměnné vysvětluje druhá z nich (a naopak). Možná si vzpomenete, že už jsme o tom mluvili v metodologickém semináři - I.

TERMINOLOGIE a práce s daty v Excelu pro JASP:

missing = chybějící hodnota (pokud respondent nějakou položku v dotazníku nevyplní).