

POST HOC TESTY JEDNODUCHÉHO TŘÍDENÍ

Jestliže provádíme analýzu rozptylu jednoduchého třídění a zamítneme hypotézu o rovnosti středních hodnot ve skupinách, obvykle se dále zajímáme o to, které ze skupin se od sebe navzájem odlišují. Pokud bychom však každou z těchto dílčích hypotéz testovali samostatně na 95% hladině spolehlivosti, celková spolehlivost složeného závěru nebude 95 %, ale nižší. Problém nárůstu pravděpodobnosti celkové chyby prvního druhu při mnohonásobných srovnáních řeší níže uvedené post hoc testy.

ZAVEDENÍ NĚKTERÝCH POJMŮ

VYVÁŽENOST TŘÍDĚNÍ

Jednoduché třídění nazýváme *vyvážené*, jestliže je v každé skupině stejný počet případů.

HOMOGENITA ROZPTYLŮ

Shoda rozptylů závislé proměnné ve všech skupinách.

KONZERVATIVNÍ, LIBERÁLNÍ, RESP. PŘESNÝ TEST

Testy založené na porovnávání dvojic dělíme na konzervativní, liberální a přesné podle jejich dosažené pravděpodobnosti celkové chyby I. druhu, kterou označujeme FWE (familywise error). Testy, které dosahují vyšší pravděpodobnosti chyby I. druhu než je stanovená hladina, nazýváme *liberální*. Testům, které dosahují nižší hladiny, než je požadováno, říkáme *konzervativní*. Dosahuje-li test právě předem stanovenou hladinu, nazýváme ho *přesný*. Pro pravděpodobnosti chyb I. druhu tedy platí:

$$P_{\text{konzervativní}}(\text{chyba I. druhu}) < P_{\text{přesný}}(\text{chyba I. druhu}) < P_{\text{liberální}}(\text{chyba I. druhu}).$$

Liberální testy mají silnější tendenci nesprávně detekovat významně odlišné dvojice než konzervativní testy. Na druhou stranu, konzervativní testy jsou někdy tak slabé, že žádné odlišné dvojice neidentifikují.

HOMOGENNÍ PODMNOŽINA

podmnožina skupin, pro niž usuzujeme, že se střední hodnoty ve skupinách neliší.

NABÍDKA POST HOC TESTŮ V PROCEDUŘE ONE-WAY ANOVA

One-Way ANOVA: Post Hoc Multiple Comparisons

Equal Variances Assumed

☐ LSD ☐ S-N-K ☐ Waller-Duncan

☐ Bonferroni ☐ Tukey Type I/Type II Error Ratio: 100

☐ Sidak ☐ Tukey's-b ☐ Dunnett

☐ Scheffe ☐ Duncan Control Category: Last

☐ R-E-G-W F ☐ Hochberg's GT2

☐ R-E-G-W Q ☐ Gabriel

Test: ☒ 2-sided ☐ < Control ☐ > Control

Equal Variances Not Assumed

☐ Tamhane's T2 ☐ Dunnett's T3 ☒ Games-Howell ☐ Dunnett's C

Null Hypothesis test

☒ Use the same significance level [alpha] as the setting in Options

☐ Specify the significance level [alpha] for the post hoc test

Level: 0,05

Continue Cancel Help

Software IBM SPSS Statistics poskytuje širokou škálu post hoc testů pro analýzu rozptylu jednoduchého třídění. Nabídka je rozdělena na dvě části podle toho, zda metoda předpokládá stejné rozptyly ve skupinách. Mezi testy nevyžadující shodu rozptylů (*Equal Variances Not Assumed*) patří Tamhane's T2, Dunnett's T3, Games-Howell a Dunnett's C. Třída testů předpokládajících stejné rozptyly ve skupinách (*Equal Variances Assumed*) je rozsáhlejší a lze ji rozdělit na tři části:

PÁROVÉ TESTY

Slouží k testování hypotéz o rovnosti středních hodnot jednotlivých dvojic skupin. Tyto testy poskytují rovněž intervaly spolehlivosti pro rozdíly středních hodnot těchto dvojic, a umožňují tak učinit závěry typu „průměr ve skupině A je větší než průměr ve skupině B“. Mezi párové testy řadíme: LSD, Bonferroni, Sidak, Scheffe, Tukey, Hochberg's GT2, Gabriel a Dunnett. Test Dunnett však na rozdíl od ostatních neprovádí všechna párová srovnání skupin, ale porovnává ostatní skupiny vzhledem ke kontrolní skupině. Některé párové testy umožňují rovněž nalézt homogenní podmnožiny skupin, tj. podmnožiny, v nichž se průměry neliší (viz tabulka níže).

KROKOVĚ SESTUPNÉ TESTY

Hledají homogenní podmnožiny skupin následujícím způsobem: nejprve seřadí skupiny podle průměru a provedou test shody průměrů pro skupinu s nejvyšším a nejnižším průměrem. Je-li hypotéza zamítnuta, rozdělí se další postup na dvě větve: v jedné je vyloučena skupina s nejnižším průměrem, v druhé skupina s nejvyšším průměrem. V obou větvích se následně na

zbývající množině opět pokračuje porovnáním průměrů skupiny s nejnižším a nejvyšším průměrem a v případě zamítnutí nulové hypotézy dochází k dalšímu větvení. Proces v každé větvi končí, pokud není zamítnuta nulová hypotéza (vznikne homogenní podmnožina), nebo již nezůstaly žádné skupiny, které by bylo možné porovnat. Postup však neumožňuje činit závěry o jednotlivých dvojicích skupin a rovněž nejsou k dispozici intervaly spolehlivosti. Mezi tyto testy patří R-E-G-W F, R-E-G-W Q, S-N-K, Tukey's – b a Duncan.

TEST ZALOŽENÝ NA BAYESOVĚ PŘÍSTUPU

Například Waller-Duncanův test. Bližší popis tohoto testu lze nalézt níže.

POPISY JEDNOTLIVÝCH TESTŮ

TESTY PŘEDPOKLÁDAJÍCÍ SHODU ROZPTYLŮ VE SKUPINÁCH:

- **LSD (least significant difference)** – provádí Fischerův LSD test založený na t-testu pro jednotlivé dvojice skupin, kde se používá stejná hladina (zadaná v poli *Significance level*). Jediný rozdíl proti opakovanému použití t-testu je ve způsobu odhadu směrodatné odchylky, která je založena na všech skupinách, ne pouze na dvou porovnávaných. Je velmi liberální a pro velké množství testovaných skupin dosahuje celkově vysoké chyby I. druhu.
- **Bonferroni** – vychází z Fischerova LSD testu. Pro jednotlivé t-testy je jejich hladina významnosti dělena celkovým počtem testovaných dvojic, aby bylo zohledněno mnohonásobné porovnávání. Tento test je poměrně konzervativní a často není schopen identifikovat významné rozdíly průměrů ve dvojicích.
- **Sidak** – Šidákova varianta korekce pro t-testy v metodě LSD. Hladina významnosti testů pro jednotlivé dvojice skupin se získá jako $1 - (1 - \alpha)^{1/\text{počet dvojic}}$, kde α je hodnota zadaná v poli *Significance level*. Je trochu méně konzervativní než Bonferroniho metoda, avšak předpokládá nezávislost testovaných hypotéz.
- **Scheffe** – Scheffého test zaručuje nepřekročení požadované hladiny významnosti nejen při porovnávání dvojic skupin, ale i při testování jakýchkoliv kontrastů (kontrast je lineární kombinace průměrů ve skupinách splňující, že součet koeficientů je nulový). Je založen na F-rozdělení poměru součtu čtverců. Vzhledem ke své konstrukci je velmi konzervativní a nevhodný, chceme-li testovat pouze dvojice.
- **R-E-G-W F** – Ryan-Einot-Gabriel-Welschův krokově sestupný test mnohonásobného porovnávání založený na F-testu.
- **R-E-G-W Q** – je analogií R-E-G-W F, ale využívá studentizované rozpětí průměrů ve skupinách (studentizované rozpětí se získá jako podíl rozpětí a směrodatné odchylky). Je výpočetně méně náročný než R-E-G-W F, který je ale o něco silnější.
- **S-N-K** – Student-Newman-Keulsův krokově sestupný test porovnává rozdíly středních hodnot dvojic skupin na základě studentizovaného rozpětí průměrů pro danou podmnožinu skupin.

- **Tukey** – zahrnuje Tukeyho test a Tukeyho HSD test. Tukeyho HSD test je modifikací Tukeyho testu pro rozdílné velikosti skupin. Oba jsou založené na porovnání všech dvojic skupin na základě rozdělení studentizovaného rozpětí průměrů pro všechny skupiny. Pro vyvážené třídění je pravděpodobnost chyby prvního druhu rovna požadované hladině. Pro nevyvážené třídění je test mírně konzervativní a slabší než R-E-G-W F test, ale na rozdíl od něj umožňuje konstrukci intervalů spolehlivosti.
- **Tukey's b** – kritická hodnota testu pro rozdíl průměrů ve skupinách je stanovena jako aritmetický průměr studentizovaných rozpětí používaných v testech S-N-K a Tukey. Je tedy kombinací liberálního a v případě nevyváženého třídění mírně konzervativního testu.
- **Duncan** – Duncanův krokově sestupný test postupuje analogicky jako S-N-K pomocí srovnání s rozdělením studentizovaného rozpětí průměrů pro danou podmnožinu skupin, ale kalibruje hladinu testů pro jednotlivé páry s ohledem na mnohonásobná porovnávání. Je liberálnější než S-N-K.
- **Hochberg's GT2** – je podobný jako Tukeyho HSD test, je však založen na rozdělení studentizované maximální hodnoty průměrů. Obvykle je více konzervativní než Tukeyho test.
- **Gabriel** – Gabrielův párový test vychází, stejně jako Hochsberg's GT2, z rozdělení studentizované maximální hodnoty průměrů ve skupinách. Pro vyvážené třídění tomuto testu odpovídá, pro mírně nevyvážené je méně konzervativní, ovšem pro výrazně nevyvážené se stává liberálním.
- **Waller-Duncan** – Waller-Duncanův test je založen na bayesovském přístupu. Místo požadované hladiny se určuje požadovaný poměr chyby prvního a druhého druhu, defaultní hodnota je 100. Metoda předpokládá vyváženost, jinak je použit harmonický průměr. Parametr 100 přibližně odpovídá hladině významnosti 0,05, parametry 1000 a 10 hladinám 0,01 a 0,1.
- **Dunnett** – Dunettův test je test mnohonásobného porovnávání založený na t-testu, který porovnává kontrolní skupinu s ostatními skupinami. IBM SPSS Statistics umožňuje jako kontrolní skupinu zvolit první nebo poslední z nich. Dále lze nastavit, jestli se má provést test oboustranný nebo jednostranný. Dunettův test je považován za jeden z nejméně konzervativních post hoc testů.

TESTY **NEVYŽADUJÍCÍ** SHODU ROZPTYLŮ VE SKUPINÁCH:

- **Tamhane's T2** – Tamhanův test je odvozen z t-testu pro rozdílné rozptyly ve skupinách. Je konzervativní.
- **Dunnett's T3** – Je analogický Tamhanově T2 testu, ale je založen na studentizované maximální hodnotě průměrů ve skupinách. Je rovněž konzervativní, ale méně než Tamhane's T2.
- **Games-Howell** – test odvozený z Tukeyoho testu. V mnoha případech je liberální, na druhou stranu je silnější než ostatní testy v této kategorii.

- **Dunnett's C** – test založený na studentizovaném rozpětí průměrů ve skupinách. Obecně je konzervativní – v případě malého rozsahu výběru je více konzervativní než Dunnettův T3 test, pro větší rozsahy výběru je méně konzervativní.

| Test | Vyžaduje shodu rozptylů | Vyžaduje vyváženost | Přístup k chybě I. druhu | K dispozici interval spolehlivosti | Konstrukce homogenních podmnožin |
|-----------------|-------------------------|---------------------|--|------------------------------------|----------------------------------|
| LSD | ANO | NE | nejvíce liberální | ANO | NE |
| Bonferroni | ANO | NE | konzervativní | ANO | NE |
| Sidak | ANO | NE | konzervativní | ANO | NE |
| Scheffe | ANO | NE | nejvíce konzervativní | ANO | ANO ² |
| R-E-G-W F | ANO | ANO ¹ | mírně konzervativní | NE | ANO |
| R-E-G-W Q | ANO | ANO ¹ | mírně konzervativní | NE | ANO |
| S-N-K | ANO | ANO ¹ | liberální | NE | ANO ² |
| Tukey | ANO | NE | přesný/mírně konzervativní ³ | ANO | ANO ² |
| Tukey's – b | ANO | ANO ¹ | χ^5 | NE | ANO ² |
| Duncan | ANO | ANO ¹ | liberální | NE | ANO ² |
| Hochsberg's GT2 | ANO | NE | konzervativní | ANO | ANO ² |
| Gabriel | ANO | NE | liberální/mírně konzervativní ⁴ | ANO | ANO ² |
| Waller-Duncan | ANO | ANO ¹ | χ^5 | NE | ANO ² |
| Dunnett | ANO | NE | mírně konzervativní | ANO | NE |

| | | | | | |
|---------------------|----|----|------------------|-----|----|
| Tamhane's T2 | NE | NE | konzervativní | ANO | NE |
| Dunnett's T3 | NE | NE | konzervativní | ANO | NE |
| Games-Howell | NE | NE | vesměs liberální | ANO | NE |
| Dunnett's C | NE | NE | konzervativní | ANO | NE |