

Statistické testy při srovnání skupin u kvantitativních dat

Pro skupiny s **Poissonovskými daty** (diskrétní data popisující počet výskytů sledované události na jednotku času/prostoru), kde $\bar{x} > 30$ (nebo v nejhorším 15), můžeme spočítat 95% **interval spolehlivosti průměru**: $\bar{x} \pm 2 \times (\sqrt{\sum x} / n)$. Skupiny, jejichž intervaly spolehlivosti se nepřekrývají, jsou odlišné na hladině významnosti 0,05.

Pro **spojitá data** vybereme test jednak podle designu experimentu a jednak podle toho, jestli je rozložení dat normální nebo ne.

Typ srovnání	Parametrický test	Neparametrický test
2 skupiny dat nepárově (nezávislé skupiny vzorků)	t-test	Mann-Whitney
2 skupiny dat párově (stejná sada vzorků ve dvou stavech)	párový t-test	pro symetrická rozložení Wilcoxon , jinak znaménkový test
Více skupin dat	ANOVA	Kruskal-Wallis

Testy, které nevyžadují znalost rozdělení, nazýváme **neparametrické**. Místo původních hodnot pracují nejčastěji jen s jejich pořadím. Mají menší sílu než parametrické testy (které obvykle vyžadují normální rozdělení) a při jejich použití je proto dobré mít větší výběr (20 hodnot ve skupině). Proto je nejprve potřeba ověřit normalitu rozložení dat minimálně vizuálně, např. pohledem na histogram, případně i s použitím testu na normalitu (např. **Shapirův-Wilkův**: pokud dá $P > 0,05$, nezamítáme hypotézu o normalitě). **Parametrické** testy také obvykle předpokládají rovnost rozptylů dat v jednotlivých skupinách. Např. u t-testu musíme v případě nerovnosti rozptylů použít tzv. **Welchovu korekci**.

Hodnocení rozptylů (jejich rovnosti nebo nerovnosti) využijeme také v případě, že srovnáváme dvě metody co do přesnosti. V tom případě použijeme **F-test** (dostupný v Excelu).

Testy dále mohou být **jednostranné** (one-tailed) nebo **oboustranné** (two-tailed). První případ nastává, pokud z logiky věci (ne na základě dat) musí být jedna skupina stejná nebo větší než druhá, ale nikdy menší.

Pokud srovnáváme víc než dvě skupiny, nemůžeme použít sadu testů pro srovnání dvou skupin, protože by rostla pravděpodobnost falešně pozitivního výsledku. Proto použijeme **analýzu rozptylu** (ANOVA, Analysis of Variance). ANOVA je parametrický test, kromě normálního rozložení předpokládá nezávislost vzorků a rovnost rozptylů v jednotlivých skupinách. Ale pokud je vyrovnaná velikost skupin, tak případná různost rozptylů nevádí. A pokud je v každé skupině víc než 20 vzorků, nevádí ani porušení normality. ANOVA zkoumá, jestli rozptyl průměrů jednotlivých skupin od celkového průměru je významně větší než rozptyl jednotlivých opakování od průměru skupiny. Pokud ne, platí nulová hypotéza, skupiny se navzájem neliší a všechny mají vlastně jednu společnou střední hodnotu. Pokud ano, alespoň některé skupiny se vzájemně liší. Chceme-li zjistit které, použijeme tzv. **post-hoc test**. Post-hoc test vybíráme nejsilnější možný. Při srovnání všech skupin mezi sebou je nejčastější **Tukey test**, při srovnání několika skupin vůči jedné kontrolní použijeme **Dunnett test**.

Nejjednodušším případem je sledování vlivu jedné nezávisle proměnné (neboli parametru neboli faktoru), která rozdělí data do několika skupin (kategorií) – nezávisle proměnná je tzv. kategoriální, tedy nominální nebo ordinální (pro binární by stačil t-test). Tento design se nazývá **one-way ANOVA**. Druhý nejčastější případ nastává, když sledujeme dva parametry, u kterých předpokládáme možnou interakci – tzn. že při různých úrovních prvního parametru reagují data na druhý parametr různě. V tomto případě použijeme **faktorovou ANOVU**, která umí vyhodnotit i interakci faktorů (parametrů).

Neparametrická varianta ANOVY je **Kruskalův-Wallisův test**. Nevyžaduje normální rozdělení, ale striktně vyžaduje nezávislost pozorování. Post-hoc test se v tomto případě jmenuje **test mnohonásobného porovnání**.