**Měření reliability**

Zkoumání reliability lze provádět pomocí celé řady matematicko-statistických metod založených na faktu, že každé měření se skládá z pravé a chybové komponenty. Teoreticky lze reliabilitu vyjádřit jako podíl pravého a celkového (skládajícího se z pravého a chybového) rozptylu výsledků testu. V praxi bývá výše uvedený teoretický model různě modifikován a pro výpočet reliability existuje řada matematických metod.

**Metoda opakovaného měření** může být použita tam, kde lze měření u stejného souboru opakovat za stejných podmínek. Koeficient reliability se pak určuje jako korelační koeficient mezi prvním a druhým měřením. V pedagogicko-psychologické praxi není tato metoda běžná, protože tu je velmi obtížné vytvořit dvakrát po sobě stejné podmínky měření, protože výsledky druhého měření bývají ovlivněny zkušeností získanou při prvním měření.

Máme-li k dispozici dvě ekvivalentní formy téhož testu, je nejjednodušší použít **metodu paralelního měření**. Pro výpočet reliability se tu použije korelační koeficient mezi výsledky těchto dvou paralelních testů. Úskalím této metody je požadavek na dvě skutečně ekvivalentní formy testu.

U didaktických testů je velmi rozšířený výpočet reliability **metodou půlení.** U této metody se výsledky testu rozdělují na dvě části, z nichž každá se vyhodnotí samostatně a poté se výsledky obou částí korelují. Získaný korelační koeficient je pak třeba ještě upravit podle **Spearman** – **Brownova** vzorce

, obecně , kde nový test je L krát větší

Největší nevýhodou metody půlení je závislost koeficientu reliability na tom, jakým způsobem test rozpůlíme.

Zjišťování koeficientu reliability na základě výpočtu korelačního koeficientu metodou opakovaného nebo paralelního měření se setkává s kritikou z důvodu opominutí otázky soudržnosti (vnitřní konzistence) testu. Proto jsou poslední dobou nejrozšířenějšími metodami pro výpočet koeficientu reliability tzv. **metody vnitřní konzistence.** Reliabilitu tu stanovujeme jednak výpočtem podle **Kuderova-Richardsonova vzorce** (vhodný pro didaktické testy úrovně, které jsou složeny z obsahově homogenních úloh)

,

kde ***k*** je počet úloh v testu, ***s*** směrodatná odchylka pro celkové výsledky žáků v testu, Pi úspěšnost i-té úlohy a Qi = 1 - Pi, nebo pomocí **Cronbachova** **alfa**

***a* *=*** , kde si je rozptyl i-té položky

Cronbachovo alfa má několik interpretací. Nejčastější je představa, že jde o korelaci mezi naším testem a všemi ostatními možnými testy z hypotetické množiny všech testů pro zjišťování dané proměnné se stejným počtem položek. Často používaným koeficientem reliability je také **standardizované Cronbachovo alfa**, které získáváme, standardizujeme-li všechny položky testu na stejný rozptyl rovný jedné. Standardizované Cronbachovo alfa se počítá podle vzorce

***a* *=*** ,

kde ***k*** je počet položek testu a ***r*** průměrný korelační koeficient mezi položkami.

Jestliže má v některých úlohách bodová stupnice více než dvě hodnoty (např. boduje se i částečné řešení zlomkem bodu nebo je za úlohu možné získat např. 0 až 5 bodů), je nutné použít koeficient Cronbachovo alfa. Kuder-Richardsonův vzorec lze použít pouze pro testy, ve kterých jsou odpovědi hodnoceny "správně" - "spatně".

 Obecně platí, že čím je test delší, tím je větší jeho reliabilita. Reliabilitu testu lze tedy za jinak stejných podmínek (nezmění se průměrná korelace mezi jednotlivými položkami testu) zvětšit pouhým zvýšením položek testu.

Mnohdy nás může zajímat, jaký má každá položka testu vliv na celkovou reliabilitu testu. Za tímto účelem můžeme výpočet koeficientu reliability (používá se tu především Cronbachovo alfa) opakovat s tím, že postupně vynecháváme jednotlivé položky a sledujeme, zda po jejich vynechání Cronbachovo alfa klesá či vzrůstá. Pokud po vyjmutí některé položky reliabilita testu vzrůstá, je účelné ji vynechat.

**Validita** viz <https://is.muni.cz/el/fsps/jaro2007/d010/Vlastnosti.pdf>