

Fórum:Testy/Validita testu

Měří test to, co chceme, aby měřil? Snad nejdůležitější vlastností testu je, aby měřil skutečně tu znalost nebo dovednost, kterou měřit chceme. Pokud test měří něco jiného, mohou být jeho výsledky zcela nesprávně interpretovány. Míru, do jaké test měří skutečně to, co má, nazýváme **validita**. Přístupů, jak měřit validitu, je celá řada. Tyto přístupy se vzájemně doplňují a tvoří tak různé zdroje důkazů o validitě (podrobněji viz ^[1]). Jednodušeji pak osvětluje různé druhy validity kniha ^[2]. Zde pojednáme jen některé z nich.

Validita obsahová

Obsahová validita hodnotí vztah mezi obsahem testu a oborem, jehož znalost má měřit. Zkušený pedagog by měl posoudit, nakolik otázky obsažené v testu pokrývají zkoušenou látku, nakolik reprezentativní je zastoupení otázek z jednotlivých okruhů, zda všechny otázky spadají do zkoušené látky a nezkoušejí spíše nějakou jinou schopnost, vlastnost nebo znalost. Posouzení obsahové validity je zčásti zpětným pohledem na plán testu a zhodnocením jeho kvality, ale také posouzením, zda byl plán testu skutečně dodržen.

Validita kritériální

Na prokázání skutečnosti, že daný test skutečně měří znalost, kterou měřit chceme, samotná obsahová validita nepostačí. Ta může prokázat pouze skutečnost, že byla vyvinuta snaha pokrýt obsah daného kurikula. Pro doložení validity testu je obecně nutné ještě nějaké další kritérium měřeného atributu, nezávislé na našem didaktickém testu. Validita, které se v tomto kontextu říká validita kritériální, se pak odhaduje zkoumáním **vztahu výsledku testu k danému nezávislému kritériu**. Jak lze tušit, často největším problémem prokazování kritériální validity je právě nalezení vhodného kritéria.

- Máme-li k dispozici jiný, již ověřený test, můžeme zjišťovat **závislost mezi naším testem a zmíněným ověřeným testem**. Tomuto druhu kritériální validity se pak říká **validita souběžná** (*concurrent validity*). V praxi se odhaduje tak, že skupině studentů zadáme jak nový test, tak test ověřený (dosud používaný) a měříme závislost mezi dvěma celkovými skóre např. pomocí korelačního koeficientu.
- Představu o tom, do jaké míry náš test **předpovídá budoucí hodnoty nějakého kritéria**, nám dává další typ kritériální validity, tzv. **validita predikční**. Predikční validita je klíčovým parametrem všech přijímacích testů. Účelem přijímacích testů je vybrat studenty s nejlepšími dispozicemi pro budoucí studium. Je proto na místě zkoumat, zda používané testy skutečně predikují úspěšnost ve studiu. V praxi to znamená, že se zjišťuje korelace výsledků přijímacích zkoušek s úspěšností studia, nebo že se z dat odhaduje regresní model, kterým lze úspěšnost ve studiu předpovídat.
- Dále nás může zajímat, zdali daný test **přináší novou informaci nad tu, kterou získáme jiným prováděným testem**, tedy jaká je jeho **validita inkrementální** neboli přírůstková. V případě zmíněných přijímacích testů nás může například zajímat, zda přijímací testy přidávají novou informaci o budoucím studiu uchazeče nad tu, kterou nám poskytne jeho středoškolský prospěch. Např. studie ^[3] na základě dat studentů přijatých na 1. LF UK ukázala, že středoškolský prospěch vysvětlí zhruba 15 % variability úspěšnosti ve studiu. Výsledek z přijímací zkoušky zvýší procento vysvětlené variability úspěšnosti na 22 %, přidání informace o úspěšně absolvovaných profilových předmětech na střední škole na 25 % a informace o roku maturity (v roce přijetí nebo dříve) dokonce na 30 %. Všechny zmíněné efekty byly v modelu signifikantní (tedy statisticky průkazné), prokázala se tak jejich přírůstková validita.

Statistické metody použité k **odhadu kritériální validity** testu závisí především na typu proměnné, která vystupuje jako kritérium. Lze-li tuto proměnnou chápat jako spojitou (např. počet bodů v jiném testu, nebo úspěšnost měřená průměrným prospěchem v prvním roce vysokoškolského studia), můžeme k odhadu použít Pearsonův nebo jiný korelační koeficient. Obecněji lze závislost na kritériu uchopit pomocí **regresní analýzy** ^[4]. Odhadnutý regresní model se na základě našich dat snaží pro každý výsledek testu předpovědět hodnotu kritéria (počtu bodů v referenčním testu, průměrný prospěch v následném studiu, úspěšnost studia apod.). Testováním významnosti jednotlivých složek modelu lze prokázat validitu testu nebo jeho přírůstkovou validitu nad již používané testy. Podrobněji se odhadu validity (v kontextu přijímacích zkoušek na vysoké školy) věnuje Zvára v publikaci ^[5].

Odkazy

Reference

1. **Cite error: Invalid <ref> tag; no text was provided for refs named URBÁNEK**
2. SCHINDLER, Radek. Rukověť autora testových úloh. Vyd. 1. Praha: centrum pro zjišťování výsledků vzdělávání, 2006, 86 s. ISBN 80-239-7111-5., on-line: <http://www.cermat.cz/rukovet-autora-testovych-uloh-1404034186.html>
3. ŠTUKA, Čestmír, Patrícia MARTINKOVÁ a Karel ZVÁRA, et al. The prediction and probability for successful completion in medical study based on tests and pre-admission grades. *The New Educational Review* [online]. 2012, roč. -, vol. 28, no. 2, s. 138-152, dostupné také z http://www.educationalrev.us.edu.pl/vol/tner_2_2012.pdf. ISSN 1732-6729.
4. ZVÁRA, Karel. *Regrese*. 1. vydání. Praha : MATFYZPRESS, vydavatelství Matematicko-fyzikální fakulty

- Univerzity Karlovy v Praze, 2008. 254 s. ISBN 978-80-7378-041-8.
5. BYČKOVSKÝ, Petr a Karel ZVÁRA. *Konstrukce a analýza testů pro přijímací řízení*. 1. vydání. Praha : Univerzita Karlova v Praze, Pedagogická fakulta, 2007. 79 s. ISBN 978-80-7290-331-3.