

Kvantitatív tudásmérés

Dr. Balázs Béla

LEXINFO Informatikai Nyelvvizsga-központ

ÖSSZEFOGLALÁS

- Az utolsó 25-30 évben a tudásmérés területén végzett kutatások intenzitása exponenciálisan növekedett.
- Állandó gondot jelent azonban, hogy a vizsgáztatás/tesztelés mindennapos gyakorlata távolról sem fejlődik olyan gyorsan, mint a tesztkészítés elmélete.
- A gyakorlati szakemberek arra panaszkodnak, hogy az elméleti cikkeket nehéz megérteni, és azok gyakran számukra irrelevánsnak tűnnek, vagy legalábbis napi praxisuktól távol eső témákról szólnak.
- A kutatók és a gyakorlati szakemberek ritkán kooperálnak. Már érzékelhető azonban, hogy a légkör lassan változik, mindenekelőtt azért, mert a *Winsteps* programcsomag (és hasonló) segítségével a *Rasch*-modell elméleti és gyakorlati előnyei – akár nagy adathalmazok esetében – már standard PC-n is kihasználhatók.

Quantitative Methods in Testing Knowledge Proficiency

Dr. Béla A. Balázs

LEXINFO Language Examination Center for Informatics

Abstract

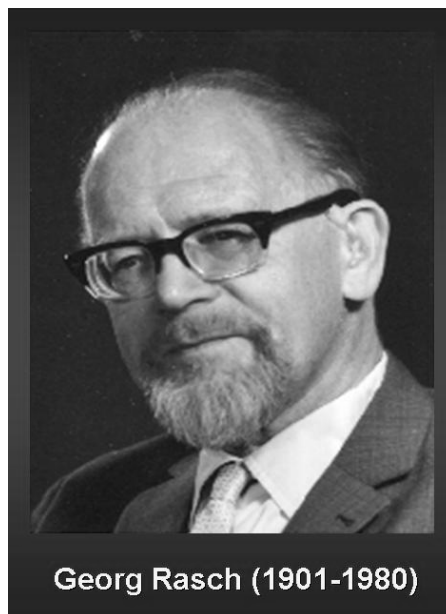
- In the last 25 - 30 years, research in the field of knowledge testing has grown exponentially.
- A perennial concern, however, is that the day-to-day practice of examining/testing has not changed as quickly as our understanding of the theory of knowledge testing.
- Practitioners complain that research reports are hard to understand and often address matters that are irrelevant or at least far removed from their daily concerns.
- Researchers and practitioners seldom co-operate. We sense, however, that this climate is changing, pre-eminently because now with the aid of *Winsteps* (or similar programs) the theoretical and practical advantages of the *Rasch*-model can be easily utilized even for large data sets, with the convenience of using a standard PC.

BEVEZETÉS

Magyarországon a klasszikus tesztelméleti módszerekkel történő elemzéseknek jelentős múltja van, de az utóbbi évek nemzetközi vizsgálatainak elemzései rávilágítanak egy alapjaiban más módszerekkel, más alapokon nyugvó tesztelmélet fontosságára.

Ez a más módszer a tesztelméletek újabb generációját képező, valószínűségi számítású alapon alapozó tesztelmélet (Item Response Theory [IRT], magyarul látens vonás elmélet), amely a vizsgalányok és a vizsgaanyagok fundamentális építőkövei -- az itemek -- tulajdonságait valószínűségelméleti eszközökkel jellemzi.

A látens vonás elmélet azzal foglalkozik, hogy standardizált pszichometriai tesztek eredményeiből hogyan következtethetünk különböző személyiségparaméterekre (pl. az informatikai staktudásra). Az idevágó modellek közül itt a *Georg Rasch* (1. ábra) dán matematikus által kidolgozott modell vázlatos ismertetésére kerül sor.



1. ábra

LÁTENS VONÁS ELMÉLET I.

- Egy vizsgánál válasszunk egy átlagos képességű standard személyt.
- Vegyünk egy átlagos nehézségű standard itemet*.
- A kiválasztás úgy történjen, hogy a standard személy a standard itemnél 50%-os valószínűséggel sikeres.
- Adott itemhalmaznál az n-edik személy sikerének valószínűsége az i-edik itemnél: P_{ni} .

A valószínűségi számításból tudjuk, hogy az esély valamely esemény bekövetkezési valószínűségének és be nem következési valószínűségének hányadosa.

Jelöljük az m személy sikerének esélyét a standard itemen b_m -el. Így

$$\frac{P_{m0}}{1 - P_{m0}} = b_m \quad (1)$$

Jelöljük továbbá a standard személy kudarcának esélyét az i itemen d_i -vel. Azaz

*Az item olyan fundamentális teszt-építőkö, amely egy vagy több kérdést ill. feleletet tartalmaz.

$$\frac{1 - P_{0i}}{P_{0i}} = d_i \quad (2)$$

Definíciók:

Legyen az m személy „képesége”,

$$\log(b_m) = \log\left(\frac{p_{m0}}{1 - p_{m0}}\right) = \theta_m \quad (3)$$

az i item „nehézsége”,

$$\log(d_i) = \log\left(\frac{1 - p_{0i}}{p_{0i}}\right) = \delta_i$$

és így a Rasch-modellre:

$$\log\left(\frac{P_{mi}}{1 - P_{mi}}\right) = \theta_m - \delta_i \quad (4)$$

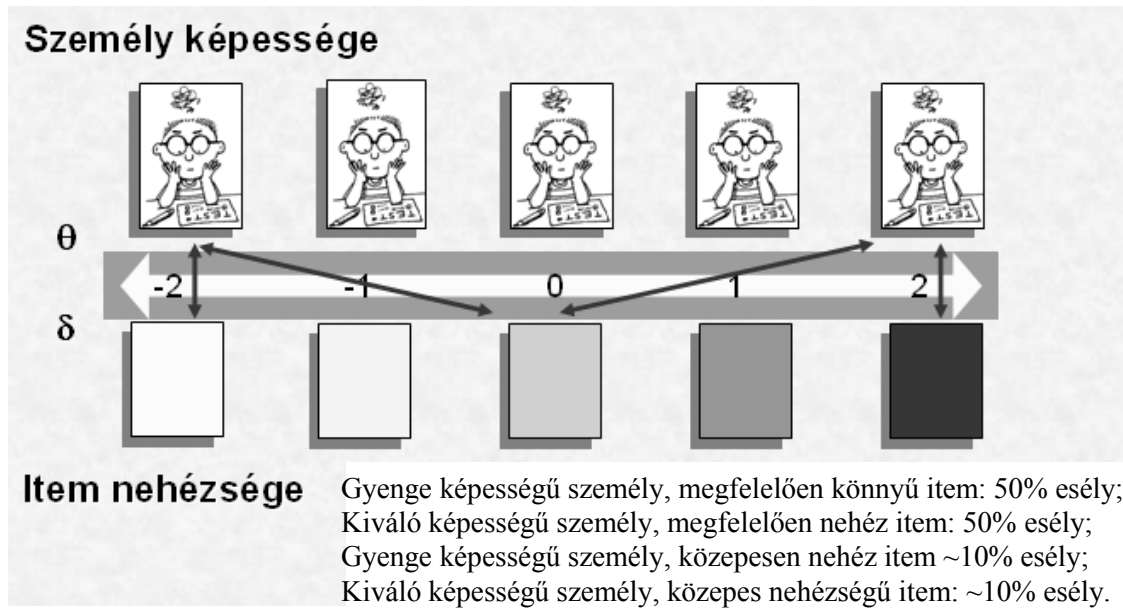
Azaz valamely személy sikerének logaritmusos esélye az i itemen egyenlő a személy képességének és az item nehézségének különbségével.

Egyébként az IRT modellek közül csak a Rasch-modellnél független két tesztszemély képességének eltérése attól, hogy melyik itemeket használjuk, és egyedül itt teljesül, hogy az itemek nehézség-különbsége nem függ a tesztelt személyek képességétől.

„VIZSGAPÁLYA”, SZEMÉLY-ITEM TÉRKÉP

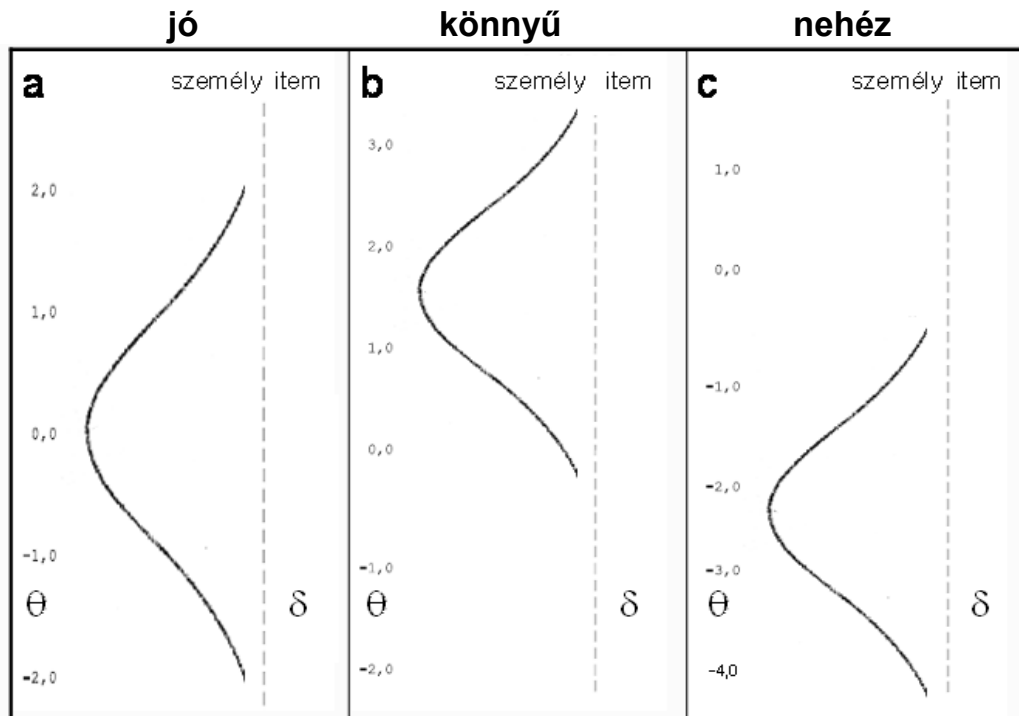
Mindenegyik vizsgázó számos képességgel rendelkezik, de ezek közül egyszerre egyet tesztelünk. Ezért az eredmény egyenes mentén – ún. logit skálán – modellezhető. {Az L logit érték a siker esélyének logaritmusosa: $L = \text{logit}(p) = \log(p/[1-p]) = \log(p) - \log(1-p)$.}

Képzeljünk el egy fokozatosan nehezedő itemekkel ellátott vizsgapályát, amelyen különböző képességű vizsgázók versenyeznek. Az előbbiek szerint az egyes itemeken való sikeres áthaladás logaritmusos esélye egyenlő a vizsgázó képességének és az item nehézségének különbségével (2. ábra).



2. ábra

A várható vizsgázói képesség-eloszlásnak megfelelő itemhalmaz esetén a teljesítmények a 0 nehézségi szint körül szórnak, míg túl könnyű feladatok esetén az értékek szignifikánsan pozitív, túl nehéz tételeknél pedig negatív középértéket mutatnak. Miután a vizsgaszintek előre rögzítettek, és a felkészülésnek, valamint a vizsgatételeknek ezekhez kell alkalmazkodniuk, a b. esetben az átlagosnál jobb, a c. esetben viszont gyengébb felkészültségű vizsgázókkal van dolgunk (lásd a 3. ábrát).



3. ábra.
Személy-item térkép.

ITEMNEHÉZSÉGI GÖRBÉK

A tudásszintmérő tesztek itemei leggyakrabban logisztikusak. A logisztikus jellegű görbének három szakasza van: a gyenge összpontszámok tartományában a görbe lassan emelkedik, majd valahol hirtelen meredekké válik, végül a magasabb összpontszámoknál ellaposodik. Általános alakja:

$$P(t) = a \frac{1 + me^{-t/\tau}}{1 + ne^{-t/\tau}} \quad (5)$$

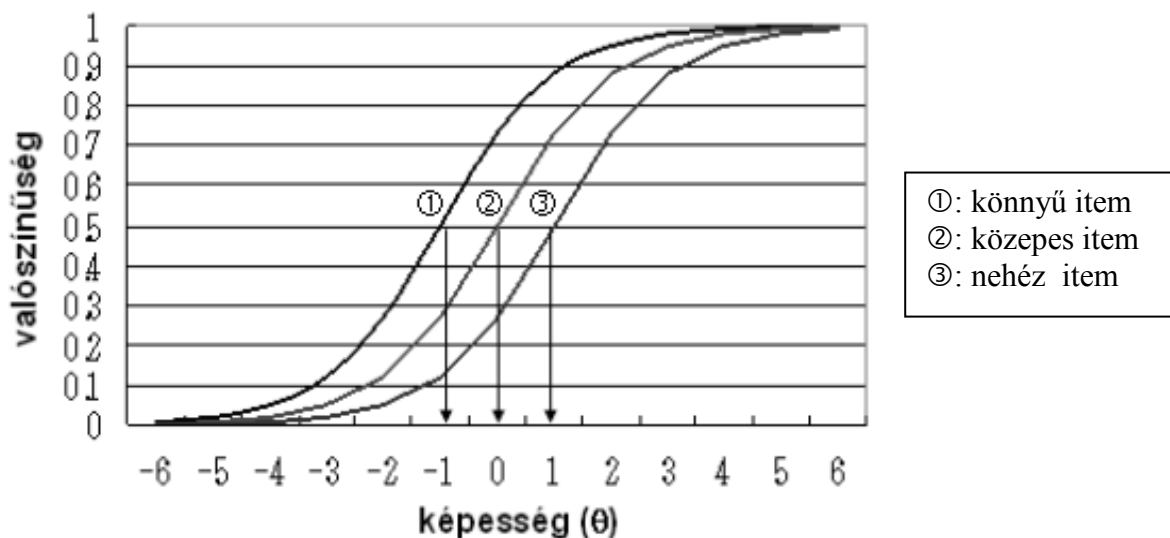
Itt a , m , n és τ valós paraméterek.

A sokféle logisztikus görbe meredekségében, illetve abban különbözik egymástól, hogy melyik képességtartományba esik a meredek szakasz. Az itemjellegfüggvény logisztikus, monoton növekvő, de csak 0 és 1 közötti értékeket vehet fel (mivel a függő változó valószínűség), értelmezési tartománya viszont az egész számsíakra.

A legegyszerűbb olyan függvény, amely 0-tól 1-ig nő, ha a független változó 0-tól végtelenig növekszik, az $f(x) = x/(1 + x)$ függvény.

Az itemjellegfüggvény is logisztikus, monoton növekvő, de csak 0 és 1 közötti értékeket vehet fel (mivel a függő változó valószínűség), értelmezési tartománya viszont az egész számsíakra. A helyes válasz valószínűsége a Rasch-modellen belül:

$$P = f(\theta, \delta) = [1 + \exp(-(\theta - \delta))]^{-1} \quad (6)$$



4. ábra

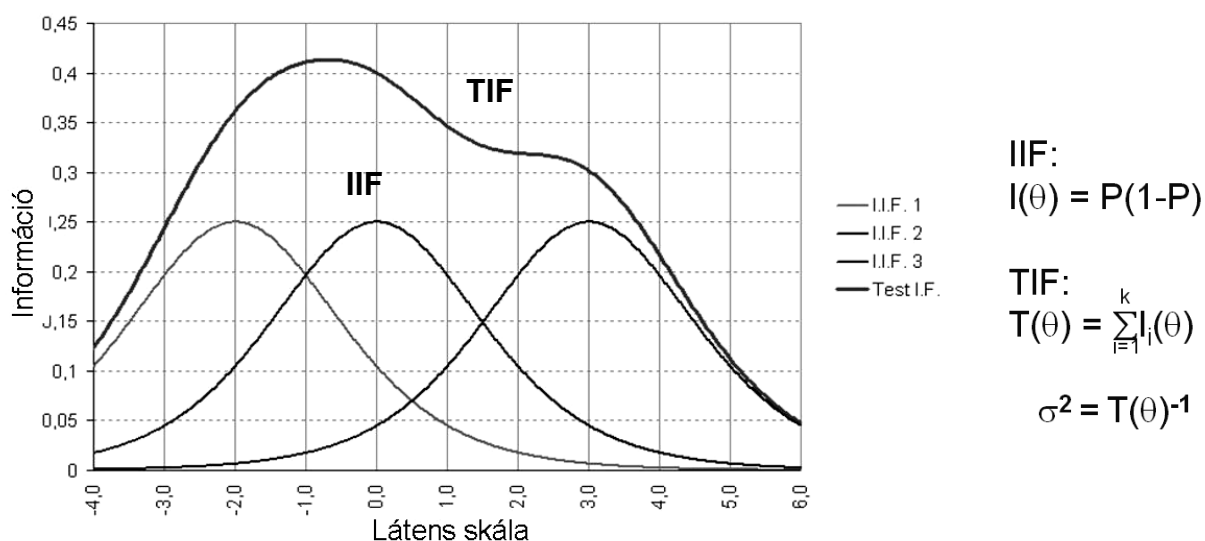
A Rasch-modell három tipikus logisztikus itemnehézségi görbéje

ITEM ÉS TESZT INFORMÁCIÓS FÜGGVÉNY

A klasszikus eljárásokkal szemben a valószínűségszámítási alapozású tesztelmélet – és ezen belül a Rasch-modell – módot talált arra, hogy a mérési hiba nagyságát a jelöltek képességeinek függvényében határozza meg.

θ két konzisztens becslésének összevetésekor azt tekinthetjük jobbnak, amelyeknek szórása kisebb. Minél kisebb a variancia (szórásnégyzet: σ^2), annál kevesebb mintavételre van szükség egy bizonyos pontosságú becslés realizálásához. Így kisebb becslés-variancia esetén a minta pontosabb „információt” ad, mint nagyobb variancia esetén.

Ebben az értelemben a minta „információtartalma” (melyet az ún. információfüggvénnyel fejezünk ki) fordítva arányos a becslés szórásnégyzetével. A Rasch modell esetén az egyes itemek információfüggvénye az $I(\theta) = P(1 - P)$ alakot ölti.* Tekintve, hogy az item-információk additívak, az egyes itemek információfüggvényeinek összege adja a teszt információfüggvényét: $T(\theta) = \sum I_i(\theta)$. Az információs függvények leggyakoribb alkalmazását a vizsgák és általában tesztek szerkesztésénél találjuk.



5. ábra

Három item információs függvény (IIF) és a teszt információs függvény (TIF)

ITEMSZERKESZTÉS, ITEMILLESZTÉS

Az információfüggvényen alapuló teszt szerkesztés menetét a következőkben összegezhethetjük:

* A látens vonás elméletben (és általában a pszichometriában) nem a Shannon-féle információfogalmat használják, hanem a Fisher-félét.

- Határozzuk meg a *teszt-információfüggvény* kívánt alakját, tekintetbe véve, hogy milyen pontosságú képességbecslésre van szükségünk az egyes képességszinteken. Eredményül kapjuk az un. *cél-információs görbét*.
- Szelektáljunk olyan itemeket, amelyek információs görbéi kielégítően kitöltik a célfüggvény alatt lefedendő területet.
- Az egymás után kiválasztott itemek információs görbéit rendre adjuk hozzá a korábbiak összegéhez, menet közben értékelve az egyre tökéletesedő teszt információfüggvényét.
- Mindaddig folytassuk az eljárást, amíg a cél-információs görbe alatti terület nincs elfogadhatóan kitöltve (azaz a teszt-információfüggvény a képességkontinuum minden számbajövő pontján elfogadható becslés-varianciát eredményez). A *Winsteps* programcsomag használata nélkül olyan mennyiségű élő munkára lenne szükség, ami már csak pénzügyi és időtényező okokból sem engedné meg a látens vonás modell alkalmazását. Segítségével viszont a Rasch-modell alkalmazása problémamentesen megoldható.

IRODALOM

1. Baker, F. B.: *Item banking in computer-based instructional systems*. Applied Psychological Measurement, 10, 405, 1986.
2. Balázs, B.: *A Rasch-modell szerepe a kvantitatív nyelvtudásmérésben*, Alkalmazott Nyelvtudomány, Vol. VII., No. 1-2., 177, 2007.
3. Horváth, Gy.: *A modern tesztmodellek alkalmazása*, Akadémiai Kiadó, Budapest, 1997.
4. Linacre, J. M.: *A User's Guide to Winsteps*, Program Manual, Chicago, 2007.
5. Molnár, Gy.: *Az ismeretek alkalmazásának vizsgálata modern tesztelméleti (IRT) eszközökkel*, Magyar Pedagógia, Vol.103, No. 4, 423, 2003.
6. Müller, H.: *Illustrationen zum Rasch-Modell*, 2002.
<http://user.uni-frankfurt.de/~hmvff/rabix/rabix101.pdf>
7. Pauen, P., Six, H-W.: *Informatikunterstützung für den weltweiten Sprachtest Deutsch als Fremdsprache (TestDaf)*, 2001.
<http://www.fernuni-hagen.de/se/PDFs/jahrbuchGFFNov2001.pdf>
7. Vale, C. D.: *Computerized Item Banking*. In: Downing, S. M., Haaladyna, T. M.: *Handbook of Test Development*, Routledge, 2006.
8. Verhelst, N. D.: *Az item-válasz-elmélet, KER szintillesztési módszertani segédlet, G. fejezet*, 2006.
http://www.nyak.hu/nyat/doc/modszertani_segedlet.pdf

