

Amikor a (mérési) módszer definiálja a mérendőt

“Mi az IQ? Az, amit az IQ-teszt mér!”

Először röviden az esélyről (ami a “fogadási tét” arányszáma).

Az **esély** (o : **odds**) a *kedvező*/megvalósuló (wins/success) esemény-szám **aránya** a nem-megvalósuló/*kedvezőtlen* (losses/failure) események számához. Valószínűséggel (relatív gyakorisággal) kifejezve, pl. kockával a „6-os dobás” *esélye*

$$o = \frac{1}{5} = \frac{1/6}{5/6} = \frac{p}{q} = \frac{p}{1-p} \quad (p+q=1)$$

$$\text{Probability} = \frac{p}{p+q} \quad \blacklozenge p / \circ p q$$

és megfordítva, ennek *valószínűsége*

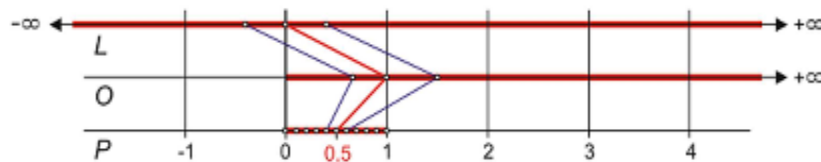
$$p = \frac{o}{1+o} = \frac{1}{6}$$

$$\text{Odds} = p : q \quad \blacklozenge p : \circ q$$

Csak kicsi esély közelíthető valószínűséggel (pl. ha $o = 1\% = 1/100$, akkor $p = 1/101 \approx o$).

Logit (log of odds - "logistic unit", L): az *esély* (természetes alapú) logaritmusa (analógiaként: az *arány* dB “egység”ben történő megadására gondoljunk)

Logit – Esély – Valószínűség



Az *esélyek aránya* logit “egység”ben különbség (“távolság”)

$$\log\left(\frac{p_1/(1-p_1)}{p_2/(1-p_2)}\right) = \log\left(\frac{p_1}{1-p_1}\right) - \log\left(\frac{p_2}{1-p_2}\right)$$

(tehát ún. intervallum skála, mint pl. amit hőmérséklet mérésnél használunk).

De hogyan kerül a csizma az asztalra?

Pszichológiai teszteknel az egyén *képessége* (f) és egy kérdés *nehézsége* (k) **arányát** gyakran, mint a [Rasch modell](#) esetén, az eséllyel teszik egyenlővé

$$\frac{f}{k} = \frac{p}{1-p}, \quad \text{vagyis} \quad p = \frac{f/k}{1+f/k}$$

a pozitív (helyes) válasz valószínűsége (p) az f/k értékkel arányos (és $f = k$ esetén $p = 0.5$).

A kérdés nehézségét a modell alapján definiálja, milyen képesség-szintű személy válaszol rá helyesen $p = 0.5$ valószínűséggel (*átlagos* képességű személy $p = 0.5$ valószínűséggel ad helyes választ).

A mérendő (f : képesség $\rightarrow \Theta$) és a „műszer paraméter” (k : nehézség $\rightarrow \delta$) aránya *logit* „egység”ben

$$L = \ln(f/k) = \ln(f) - \ln(k) = \Theta - \delta$$

Kérdéses f és k direkt összehasonlítása, hiszen ezek a tulajdonságok *nem* egymeműek.

Viszont, a kérdés *nehézsége* interpretálható úgy, hogy az nem más, mint az a *minimális képesség*, amely a pozitív (helyes) válaszhoz szükséges.

Mivel a modell a személy-képességet és a kérdés-nehézséget egymáshoz viszonyítva definiálja, a skálának nincs abszolút nulla pontja. 1 *logit* megfelel $\Theta - \delta = 1$ értéknek.

A képességet és a kérdés nehézségét a teszt adatokból kell becsülni: **a (mérési) művelet definiálja tehát a mérendőt** („a teszt mér valamit, nevezzük azt képességnek”).

A *helyes* válasz valószínűsége $p = \exp(\Theta - \delta) / (1 + \exp(\Theta - \delta)) = \exp(\Theta - \delta) / K$.

A mérendő (Θ) ilyen formában történő megadása (a konkrét mérési egyenlet helyett) a méréstechnikában szokatlan, és az is, hogy azt empirikus táblázatból (teszt adatokból) kell megadni.

Ha **a** személy (Θ_a képességgel) *helyesen* válaszol δ nehézségű kérdésre, viszont **b** személy (Θ_b képességgel) *nem*, ennek együttes valószínűsége: $p_{ab} = p_a (1 - p_b) = \exp(\Theta_a - \delta) / K_a K_b$, mivel feltételezhető a függetlenség.

Fordított esetben (**b** válaszol helyesen és **a** nem): $p_{ba} = \exp(\Theta_b - \delta) / K_a K_b$, így arányuk $p_{ab} / p_{ba} = \exp(\Theta_a - \Theta_b)$, vagyis $\Theta_a - \Theta_b = \ln(p_{ab}) - \ln(p_{ba})$. Két személy képességének „log távolsága” becsülhető relatív gyakorisággal, ami azt jelenti, hogy a képesség és a nehézség paraméter(ek) szeparálható(k).

A műszer paraméterek (= a k_i kérdés nehézségek) függetlennek tekinthetők egymástól, ezért egy **a** személynél a p_{ai} valószínűségek összegezzhetők (succes probability = total raw score).

A módszerrel-definiált mérendő (vagyis az, hogy művelet-specifikus az eredmény) és hogy a „műszer kalibráció” (k_i becslése) a teszt adatokból történik, különösen az anyagtudományok (mint anyag-keményység), a biológia (pl. fájdalom-fokozatok), a kémia (oktánszám) területén gyakori. Mivel nincs igazi „referencia módszer”, a bizonytalanság becslés „szürke terület”.