Lecture 4: Ability Scale
(Embretson & Reise (2000), Chapter 6)

Metric of the scale

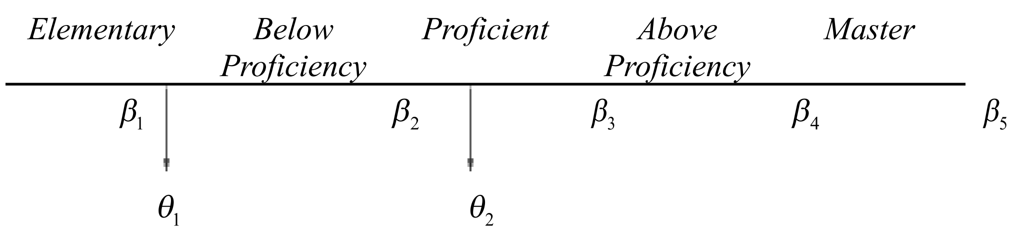
Scale in IRT is not uniquely defined. The model specification involves some indeterminacy. That is, two sets of parameters can give rise to the same vector of probabilities assuming appropriate transformation.

$$\{\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{c}\}' \rightarrow \{\boldsymbol{\theta}^*, \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \boldsymbol{c}^*\}' ; \; P(\theta) = P(\theta^*)$$

Additional constraints are needed to address this indeterminacy.

1) Imposed on the mean and variance of θ
2) Imposed on the mean and variance of β .

Example Regardless of where constraints are imposed,



Based on the locations of their ability, the examinee 1 has an elementary grasp of the domain, whereas the examinee 2 is relatively proficient in this area.

In IRT, θ and β are in a common scale. Thus, they are interpreted vis-à-vis each other.

1) Anchored on the item: $\theta > 0$ means that the person can answer items of average difficulty (i.e., $\beta = 0$) with probability greater than $.5(1+c)$.
2) Anchored on the ability: $\beta > 0$ means that person of average ability (i.e., $\theta = 0$) has less than $.5(1+c)$ chance of getting the item right.

Unit of Scale

1) Logit scale (θ')

Odds is the probability of success divided by the probability of failure □ $\dfrac{P_i(\theta)}{Q_i(\theta)}$, where $Q_i(\theta) = 1 - P_i(\theta)$

Strictly speaking, the "log-odds" transformation is meaningful only for the 1PL model.

$$P_i(\theta) = \frac{\exp(\theta - \beta_i)}{1 + \exp(\theta - \beta_i)}\ ; \quad Q_i(\theta) = \frac{1}{1 + \exp(\theta - \beta_i)}$$

Odds: $\dfrac{P_i(\theta)}{Q_i(\theta)} = \exp(\theta - \beta_i)$

logit: $(\theta - \beta_i)$

<u>Example</u> Consider two examinees with the following abilities: θ_1 and $\theta_2 = \theta_1 + 1$ (i.e., examinee 2's ability is one unit higher than examinee 1)

The logit for examinee 1 is:

The logit for examinee 2 is:

This means that if abilities differ by one point, there is a difference of one point on the logit ability scale corresponding to a factor of 2.72 in odds for success on the θ scale.

Similarly, every unit increases in difficulty, the same changes in the logit regardless of the ability level.

The logit is on an interval scale.

2PL: The logit is non-linear in θ. That is, it is equal to $\alpha_i(\theta - \beta_i)$. This means that for a unit increase in ability the increment in the logit is constant across the ability continuum but not across the items.

2) Odds scale

$$P_i(\theta) = \frac{\exp(\theta - \beta_i)}{1 + \exp(\theta - \beta_i)} = \frac{\exp(\theta)/\exp(\beta_i)}{1 + \exp(\theta)/\exp(\beta_i)} = \frac{\exp(\theta)/\exp(\beta_i)}{\frac{\exp(\beta_i) + \exp(\theta)}{\exp(\beta_i)}} = \frac{\exp(\theta)}{\exp(\beta_i) + \exp(\theta)}$$

Define $\theta^* = \exp(\theta)$ and $\beta_i^* = \exp(\beta_i)$.

$$P_i(\theta) = \frac{\theta^*}{\theta^* + \beta_i^*}; \quad Q_i(\theta) = \frac{\beta_i^*}{\theta^* + \beta_i^*}$$

Taking the odds for success, we get $\frac{\theta^*}{\beta_i^*}$.

<u>Example</u> Consider two examinees with the following abilities responding to an item: θ_1^* and $\theta_2^* = 2\theta_1^*$.

The odds for success for examinee 1 is:

The odds for success for examinee 2 is:

Taking the ratio of the odds for success is:

The ratio scale property for θ^* - and β_i^* - scales holds only for the 1PL.
What are the possible values of θ^* ?

3) True score scale (limited in this class to dichotomous response)

For a fixed person with ability θ , the total score can be written as $T = \sum X_i$.

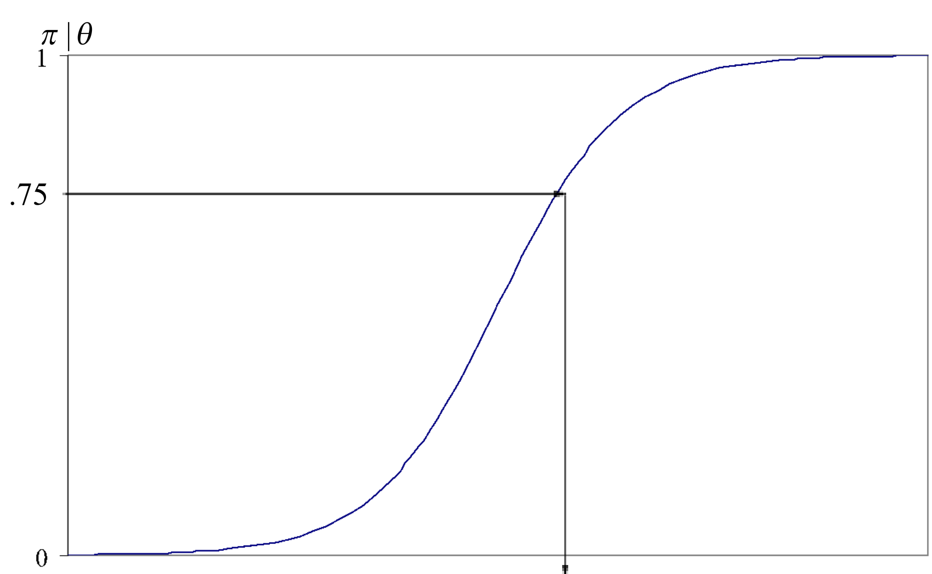
The expected value of the total score is $E(T) = E(\sum X_i) = \sum E(X_i) = \sum P_i(\theta)$

The expected total score for a fixed ability θ is given by $\tau(\theta) = \tau|\theta = \sum P_i(\theta)$, and is referred to as the true score. $\tau|\theta$ emphasizes the fact that the true score assumes different values depending on the value of θ .

The maximum value of $\tau|\theta$ is n , the number of items; the minimum value is 0 for the 1PL and 2PL and $\sum c_i$ for the 3PL.

Instead of $\tau|\theta$, we can compute $\pi|\theta = \frac{\tau|\theta}{n}$ which is the true proportion correct (or domain score).

Converting $\theta \rightarrow \pi$ and $\pi \rightarrow \theta$ should be straightforward.



θ

$\theta^{(*)}$

For example, if a domain score of .75 is the required passing score, one can easily find the $\theta^{(*)}$ such that $\pi \mid \theta^{(*)} = .75$. Conversely, given a specific value of $\theta^{(*)}$, one can find the person's domain score.

As shown above, scores in IRT can be reported in various scales. For our purposes, we will use the following classification:

I. Scores independent of the test (can be reported regardless of the test used)

A. Logit scale - θ'
1) Log metric
2) Normal metric - values can be interpreted like z-scores typically in conjunction with a mean of 0 and standard deviation of 1; used as the default scale

B. Odds scale - $\theta* = \exp(\theta)$

Interpretations of log metric and odds scale reflect the meanings of their scales when 1PL model is used. That is, a unit increase in θ will correspond to a unit increase in the log-odds of a correct response (i.e., logistic); and doubling $\theta*$ will also double the odds of a correct response.

II. Scores specific to a test

A. True score - expected number of correct responses: $\tau(\theta) = \sum P_i(\theta)$
B. True Proportion correct score - expected proportion of correct responses: $\pi(\theta) = \sum P_i(\theta) / n$