

# Item Response Theory for beginners

**Item time!**

---

Dr. Ottavia M. Epifania

Corso IRT @ Università Libera di Bolzano, 17-18 Gennaio 2023

Bressanone

- ① Fit degli item
- ② Differential Item Functioning

## Fit degli item

---

Si può valutare la fit di ogni item nel modello

Statistiche più usate:

- $S - \chi^2$  (Orland & Thissen, 2000): Statistica basata sul  $\chi^2$ . Se è significativo, l'item non è adeguato
- *Root Mean Squared Deviation* (RMSD): Distanza tra quanto atteso dal modello e i dati. Si vuole che questa distanza sia minima. Valori al di sotto di .15 sono accettabili, al di sotto di .10 sono ottimali e indicano ottima fit dell'item

# Differential Item Functioning

---

# Differential Item Functioning

---

## Un esempio

La massa dello Iodio 131 decresce di  $1/2$  ogni 8 giorni.

In un laboratorio ci sono 2gr di Iodio 131. Quanti grammi ci sono dopo 16 giorni?

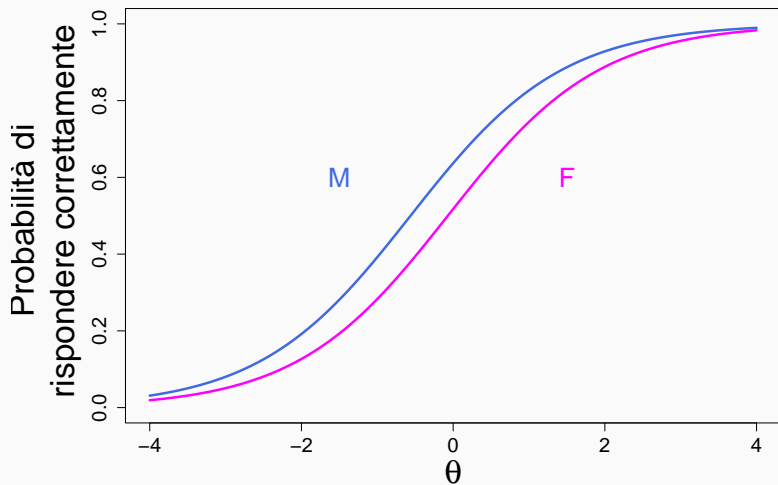
La massa dello Iodio 131 decresce di  $1/2$  ogni 8 giorni.

In un laboratorio ci sono 2gr di Iodio 131. Quanti grammi ci sono dopo 16 giorni?

Risposta corretta: 0.5gr



# Item Characteristic Curve (ICC)



Lo stesso item presentato a due gruppi appaiati per livello di tratto latente. . . non funziona nello stesso modo

I soggetti vengono appaiati in funzione del loro livello di tratto latente. **Ci sono delle differenze nella performance su un item?**

Teoricamente: Soggetti di gruppi diversi ma con lo stesso livello di tratto latente (i.e., **appaiati**) dovrebbero mostrare performance simili sull'item

Se così non è → **DIFFERENTIAL ITEM FUNCTIONING (DIF)**

Solitamente ci si basa su due gruppi:

- **reference group** (Gruppo di riferimento): è la base del confronto. Spesso è il gruppo maggioritario (e.g., italiani) o in caso di traduzioni del test il gruppo a cui viene somministrata la versione originale
- **focal group** (Gruppo focale): Gruppo focus dell'analisi. Spesso è il gruppo minoritario (e.g., immigrati di prima generazione) o il gruppo a cui viene somministrata la versione tradotta del test

# Differential Item Functioning

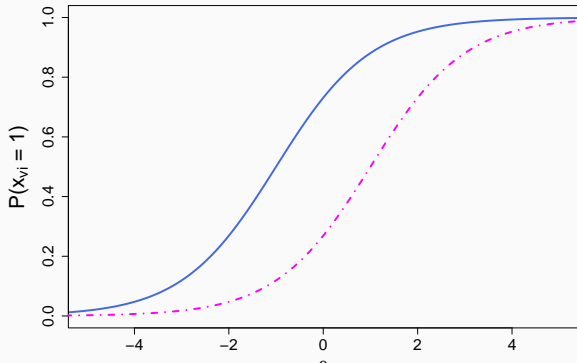
---

## DIF uniforme

L'item funziona in modo diverso nel focal group rispetto al reference group (o viceversa)

Nello specifico, avvantaggia sistematicamente uno dei due gruppi

Il vantaggio nei confronti del gruppo è costante (**uniforme**) per tutti i livelli del tratto latente



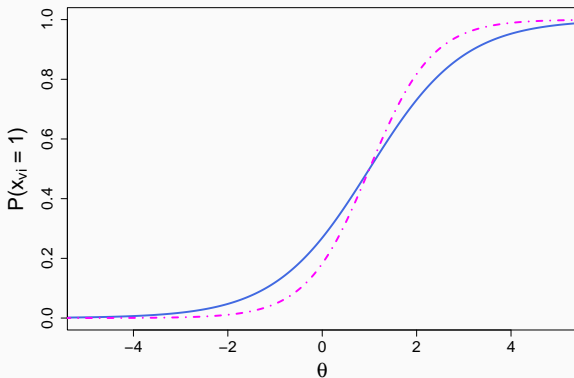
# Differential Item Functioning

---

## DIF Non uniforme

L'item funziona in modo diverso nel gruppo focale rispetto al gruppo di riferimento (o viceversa), ma il vantaggio che dà a uno dei due non è costante per tutti i livelli del tratto

L'item avvantaggia un gruppo per i livelli più bassi del tratto e avvantaggia il gruppo opposto per i livelli più alti del tratto.



# Differential Item Functioning

---

## Valutare il DIF



Esistono metodi basati sui modelli IRT e metodi basati sul punteggio osservato per valutare il DIF

I metodi basati sui modelli IRT utilizzano la stima del tratto latente ottenuta tramite il modello per appaiare i soggetti

La DIF uniforme viene valutata sui parametri di difficoltà, mentre la DIF non uniforme viene valutata sui parametri di discriminatività

#### DIF uniforme

- 1PL
- 2PL
- 3PL

#### DIF non uniforme

- 2PL
- 3PL

# Differential Item Functioning

---

## Likelihood Ratio Test

Thissen, Steinberg, & Wainer (1988)

Vengono stimati due modelli IRT sui dati:

- ➊ Nel modello “no-DIF”, i parametri vengono vincolati ad essere uguali nel reference e focal group
- ➋ Nel modello “DIF”, si lasciano liberi i parametri di cambiare nei due gruppi

I due modelli vengono confrontati. Se la differenza tra l'adattamento dei due modelli è statisticamente significativa, c'è DIF

Di fatto, funziona come una regressione lineare, in cui il gruppo di appartenenza viene inserito come predittore

# Differential Item Functioning

---

Lord

## Lord (1980)

Si stimano i parametri nel gruppo di riferimento e nel gruppo focale e si calcola la differenza.

Se la differenza è statisticamente significativa, c'è DIF tra gli item

Oltre alla significatività, si può ottenere anche un'informazione circa la dimensione dell'effetto della differenza (Delta di Lord):

- $< 1.00$ : DIF trascurabile
- $1.00 < d < 1.5$ : DIF moderata
- $> 1.5$ : DIF grande

# Differential Item Functioning

---

## L'area di Raju

## Raju (1988)

Si calcola l'area tra le curve caratteristiche dell'item nel gruppo di riferimento e nel gruppo focale.

Se l'area è uguale a 0, non c'è DIF

Più ci si allontana da 0, maggiore è il DIF