

# Item Response Theory for beginners

**Item time!**

---

Dr. Ottavia M. Epifania

University of Trento, 17 November 2023

Rovereto (TN)

- ① Item fit
- ② Differential Item Functioning

## Item fit

---

Evaluate the fit of each item to the model:

- $S - X^2$  (Orland & Thissen, 2000): Statistics based on the  $\chi^2$ . If significant, the item does not fit to the model
- *Root Mean Squared Deviation* (RMSD): Difference between what expected under the model and real data (the lower the better). Values under .15 define an acceptable fit of the item to the model, values under .10 are optimal

# Differential Item Functioning

---

# Differential Item Functioning

---

## An example

La massa dello Iodio 131 decresce di  $1/2$  ogni 8 giorni per decadimento radioattivo

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

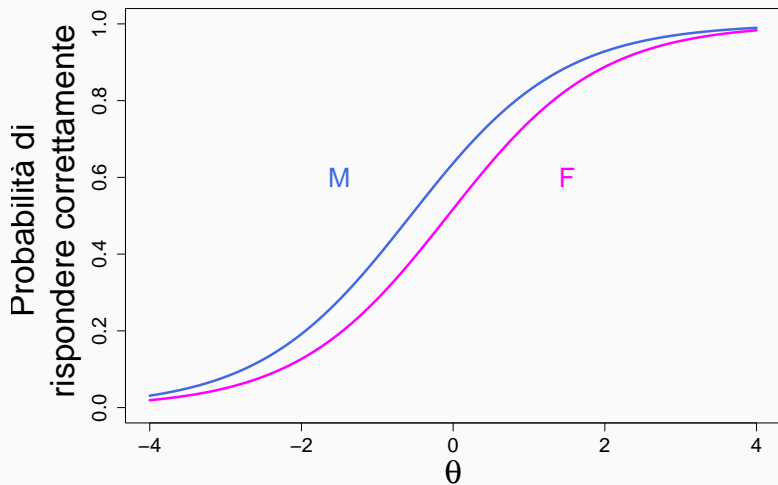
La massa dello Iodio 131 decresce di  $1/2$  ogni 8 giorni per decadimento radioattivo

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

Risposta corretta: 0.5gr



## Item Characteristic Curve (ICC)



The same item presented to two different groups paired for their level of latent trait. . . does not have the same probability of being endorsed!

The subjects are paired according to their level of the latent trait. **Are there any differences in the performance to the item?**

Theoretically: Different subjects but with the same level of the latent trait (i.e., **paired**) should have similar performances on the item

If this expectation is not met → **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

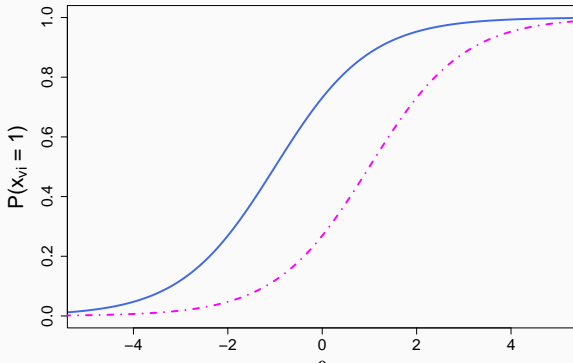
---

## 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



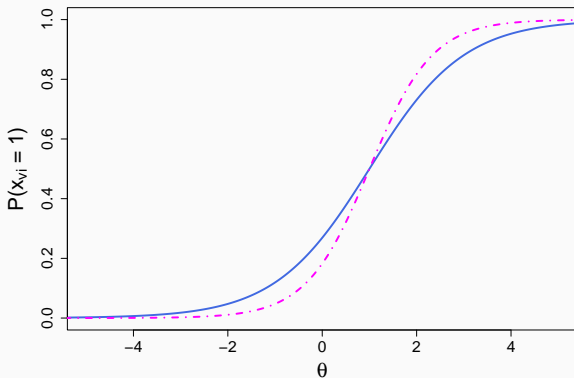
# Differential Item Functioning

---

## 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