

---

Uno studio sui livelli di competenza: la Rasch  
Analysis dei dati OCSE-PISA 2012 per la misura delle  
differenze tra gli studenti delle regioni italiane

---

Tullio Menini

[menini@unior.it](mailto:menini@unior.it)

Dipartimento di Scienze Umane e Sociali,  
Università degli studi di Napoli L'Orientale

---

# Problema

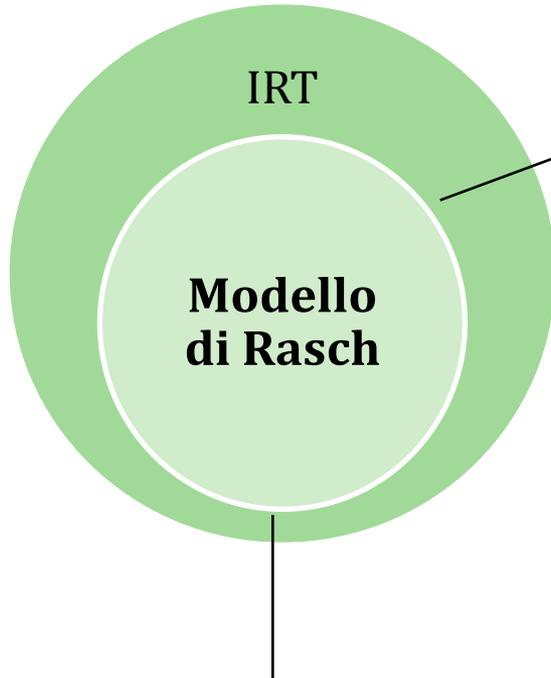
- *variabili latenti corrispondono a concetti astratti come abilità, comportamenti o stati mentali*
  - Studiare le competenze degli studenti mediante le prove cognitive (item) del questionario PISA, significa collocare il livello di tali competenze lungo un continuum numerico, in modo tale che sia possibile esprimere un giudizio quantitativo sulla posizione occupata sia in termini assoluti che relativi.
  - Presenza di item con un funzionamento significativamente differente tra le Regioni
-

---

# Metodologia

- ❑ *Item Response Theory* - IRT (Baker e Kim 2004)
  - ❑ Modello di Rasch – MR (1960)
  - ❑ *Partial Credit Model* (Masters 1982)
-

# Strumento di analisi utilizzato: Modello di Rasch



PCM: formalizzazione

$$\Pr\{X = x_k \mid \beta_j, \delta_i\} = \frac{\exp(\beta_j - \delta_{ik})}{1 + \exp(\beta_j - \delta_{ik})}$$

$\beta_j$  posizione del soggetto sul continuum

$\delta_{ik}$  regola il passaggio dalla categoria  $k - 1$  alla  $k$

Modello di Rasch

$$P(X_{ji} = x_{ji} \mid \beta_j, \delta_i) = \frac{\exp[x_{ji}(\beta_j - \delta_i)]}{1 + \exp(\beta_j - \delta_i)}$$

---

# *Diagnostica Rasch (1/2)*

- ❑ *cronbach alpha* misura di affidabilità globale del test
  - ❑ *person reliability index* capacità del test di distinguere i soggetti con alte prestazioni da quelli con basse prestazioni
  - ❑ statistica *outfit* (unweighted mean square statistic)  
valore medio dei residui standardizzati al quadrato
  - ❑ statistica *infit* (weighted mean square statistic),  
valore medio dei residui standardizzati al quadrato ponderati con le rispettive varianze.
-

# Diagnostica Rasch (2/2)

- *Differential Item Functioning* (DIF). strumento per comprendere se, relativamente all'item  $i$ -esimo, la probabilità di risposta di gruppi di soggetti (omogenei rispetto a qualche criterio qualitativo o quantitativo), varia soltanto in funzione del valore del tratto latente, da essi posseduto, o varia anche in funzione dei valori o dei livelli delle caratteristiche (*person factors*) dei rispondenti.

$$t_i = \frac{\hat{\delta}_{ig} - \hat{\delta}_{ig'}}{\sqrt{\text{var}(\hat{\delta}_{ig}) + \text{var}(\hat{\delta}_{ig'})}}$$

se il valore della statistica cade al di fuori dell'intervallo  $[-1,96 ; 1,96]$ , l'item ha un funzionamento statisticamente differente fra le classi ad un livello di significatività del 5%.

# *Descrizione dei dati e software*

- ❑ Dataset disponibile (<http://pisa2012.acer.edu.au>)

All'indagine PISA del 2012 hanno partecipato circa 510.000 studenti provenienti da 65 paesi diversi, tra i quali 34 paesi dell'OECD e 31 paesi partner. L'Italia vi ha partecipato con 31.073 studenti, provenienti da 1.194 scuole rappresentative sia per macroaree geografiche (Nord Ovest, Nord Est, Centro, Sud, Sud Isole) che per indirizzi di studio (licei, istituti tecnici, istituti professionali, scuole medie, formazione professionale).

- ❑ Dataset studenti italiani: analisi separatamente per gli item di lettura, di matematica e di scienze.
- ❑ L'analisi è condotta con il programma Winsteps di J. M. Linacre

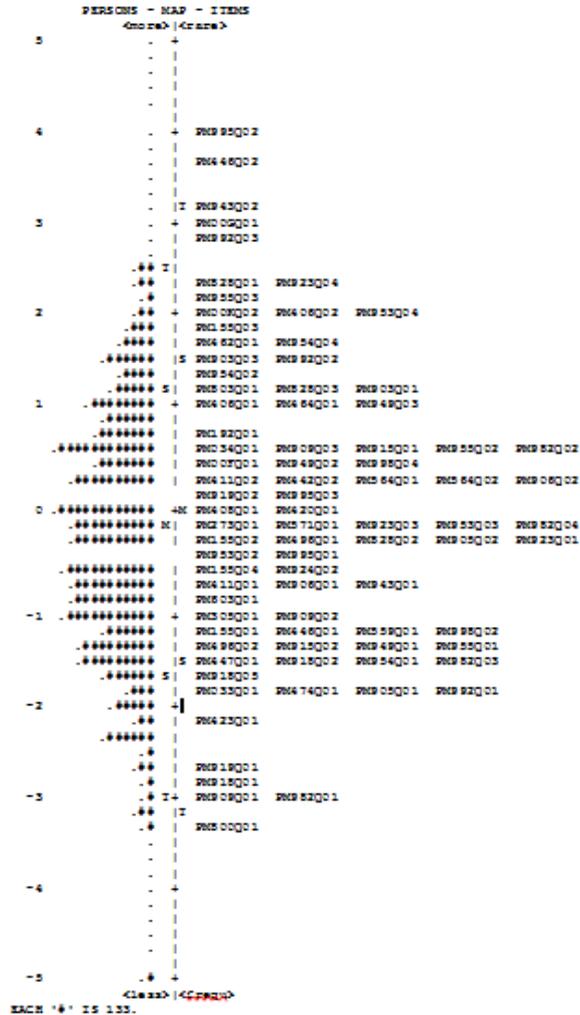


# Analisi dei dati - Mathematics (1/3)

- ❑ *cronbach alpha*       $\alpha = 0,71$
- ❑ *person reliability index* = 0,83
- ❑ *item reliability* = 1
- ❑ le statistiche *infit* non presentano per nessun item valori fuori dal range [0,8; 1,2].

# Analisi dei dati - Mathematics (2/3)

## Item-person map



# Analisi dei dati - Mathematics (3/3)

- I valori della statistica t al quadrato per ogni gruppo possono essere sintetizzati in una statistica chi-quadro per ogni item che rivela se la DIF osservata è dovuta solo al caso. L'ipotesi nulla è che la DIF sia statisticamente zero tra le classi (regioni).

Tabella 4 - Mathematics; DIF confronto fra le regioni italiane

```
INPUT: 31073 PERSONS 109 ITEMS MEASURED: 30948 PERSONS 84 ITEMS 176 CATS
```

```
-----  
DIF class specification is: DIF=@REGIONE
```

```
-----  
| PERSON          SUMMARY DIF          ITEM          |  
| CLASSES        CHI-SQUARE    D.F.  PROB.  Number Name |  
|-----|-----|-----|-----|-----|  
|      21          44.0225        20  .0015        1 PM00FQ01 |  
|      21          36.2504        20  .0144        2 PM00GQ01 |  
|      21          33.6313        20  .0287        3 PM00KQ02 |  
|      21          30.3363        20  .0646        4 PM033Q01 |  
|      21          27.1351        20  .1315        5 PM034Q01 |  
|      21          19.5765        20  .4847        6 PM155Q01 |  
|      21          40.1285        20  .0048        7 PM155Q02 |  
|      21          31.5214        20  .0487        8 PM155Q03 |  
|      21          28.5899        20  .0961        9 PM155Q04 |  
|      21          35.6243        20  .0170       10 PM192Q01 |  
|      21          80.4427        20  .0000       11 PM273Q01 |  
|-----|-----|-----|-----|-----|
```

---

# Analisi dei dati - Reading (1 / 3)

- ❑ *cronbach alpha*       $\alpha = 0,76$
  - ❑ *person reliability index* = 0,78
  - ❑ *item reliability* = 1
  - ❑ le statistiche *infit* non presentano per nessun item valori fuori dal range [0,8; 1,2].
-



# Analisi dei dati - Reading (3/3)

- diversi item utilizzati per valutare la comprensione del testo hanno un funzionamento statisticamente differente ad un livello di significatività del 5%.

Tabella 8 - Reading; DIF confronto fra le regioni italiane

INPUT: 31073 PERSONS 44 ITEMS MEASURED: 21261 PERSONS 44 ITEMS 89 CATS

DIF class specification is: DIF=@REGIONE

| PERSON  | SUMMARY DIF |      |       | ITEM   |          |
|---------|-------------|------|-------|--------|----------|
| CLASSES | CHI-SQUARE  | D.F. | PROB. | Number | Name     |
| 21      | 113.526     | 20   | .0000 | 1      | PR220Q01 |
| 21      | 35.1583     | 20   | .0193 | 2      | PR220Q02 |
| 21      | 35.1732     | 20   | .0192 | 3      | PR220Q04 |
| 21      | 24.7826     | 20   | .2099 | 4      | PR404Q03 |
| 21      | 50.0193     | 20   | .0002 | 5      | PR404Q06 |
| 21      | 160.694     | 20   | .0000 | 6      | PR404Q07 |
| 21      | 49.2649     | 20   | .0003 | 7      | PR404Q10 |
| 21      | 48.4404     | 20   | .0004 | 8      | PR404Q11 |
| 21      | 69.8937     | 20   | .0000 | 9      | PR406Q01 |
| 21      | 58.4720     | 20   | .0000 | 10     | PR406Q02 |
| 21      | 21.4198     | 20   | .3728 | 11     | PR406Q05 |
| 21      | 35.1862     | 20   | .0191 | 12     | PR412Q01 |
| 21      | 53.0641     | 20   | .0001 | 13     | PR412Q05 |
| 21      | 91.9540     | 20   | .0000 | 14     | PR412Q06 |
| 21      | 30.0532     | 20   | .0690 | 15     | PR412Q08 |
| 21      | 80.2470     | 20   | .0000 | 16     | PR420Q02 |
| 21      | 37.3668     | 20   | .0106 | 17     | PR420Q06 |
| 21      | 61.4067     | 20   | .0000 | 18     | PR420Q09 |
| 21      | 24.1873     | 20   | .2343 | 19     | PR420Q10 |
| 21      | 26.5743     | 20   | .1477 | 20     | PR424Q02 |
| 21      | 34.6767     | 20   | .0219 | 21     | PR424Q03 |

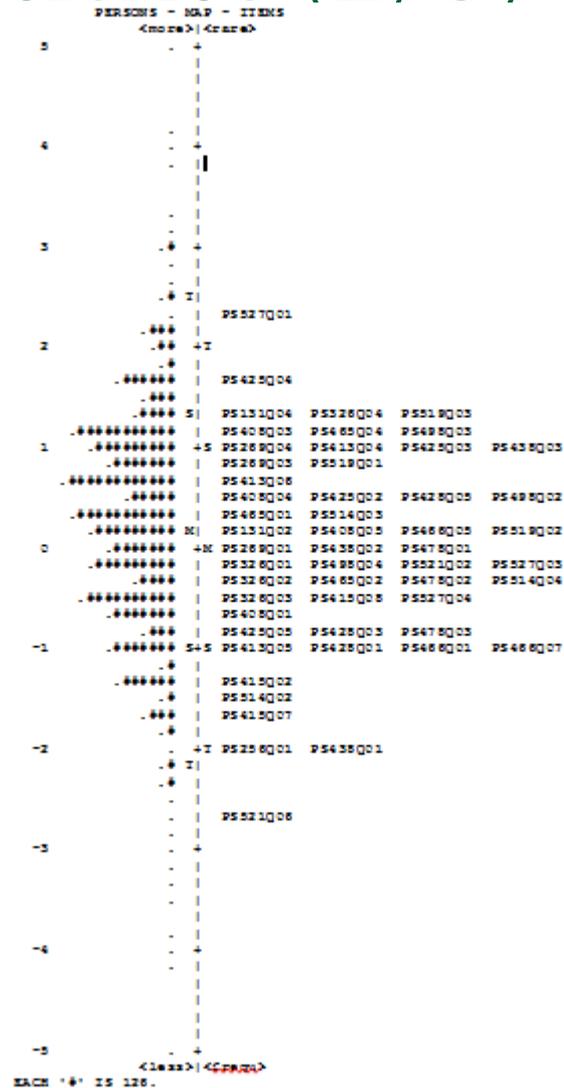
---

# Analisi dei dati - Science (1/3)

- ❑ *cronbach alpha*       $\alpha = 0,74$
  - ❑ *person reliability index* = 0,79
  - ❑ *item reliability* = 1
  - ❑ le statistiche *infit* non presentano per nessun item valori fuori dal range [0,8; 1,2].
-

# Analisi dei dati - Science (2/3)

## Item-person map



# Analisi dei dati - Science (3/3)

- diversi item utilizzati per valutare la comprensione del testo hanno un funzionamento statisticamente differente ad un livello di significatività del 5%.

Tabella 12 - Science; DIF confronto fra le regioni italiane

INPUT: 31073 PERSONS 53 ITEMS MEASURED: 21209 PERSONS 53 ITEMS 109 CATS

DIF class specification is: DIF=@REGIONE

| PERSON CLASSES | SUMMARY DIF CHI-SQUARE | D.F. | PROB. | ITEM Number | ITEM Name |
|----------------|------------------------|------|-------|-------------|-----------|
| 21             | 28.2179                | 20   | .1043 | 1           | PS131Q02  |
| 21             | 59.2642                | 20   | .0000 | 2           | PS131Q04  |
| 21             | 54.5288                | 20   | .0000 | 3           | PS256Q01  |
| 21             | 67.4230                | 20   | .0000 | 4           | PS269Q01  |
| 21             | 65.7281                | 20   | .0000 | 5           | PS269Q03  |
| 21             | 34.0376                | 20   | .0259 | 6           | PS269Q04  |
| 21             | 40.3539                | 20   | .0045 | 7           | PS326Q01  |
| 21             | 64.3161                | 20   | .0000 | 8           | PS326Q02  |
| 21             | 27.6412                | 20   | .1182 | 9           | PS326Q03  |
| 21             | 47.6798                | 20   | .0005 | 10          | PS326Q04  |
| 21             | 38.8255                | 20   | .0070 | 11          | PS408Q01  |
| 21             | 40.1224                | 20   | .0048 | 12          | PS408Q03  |
| 21             | 68.8558                | 20   | .0000 | 13          | PS408Q04  |
| 21             | 29.3261                | 20   | .0815 | 14          | PS408Q05  |
| 21             | 45.3482                | 20   | .0010 | 15          | PS413Q04  |
| 21             | 43.7717                | 20   | .0016 | 16          | PS413Q05  |
| 21             | 32.7003                | 20   | .0364 | 17          | PS413Q06  |
| 21             | 21.8827                | 20   | .3469 | 18          | PS415Q02  |
| 21             | 46.5444                | 20   | .0007 | 19          | PS415Q07  |
| 21             | 18.7094                | 20   | .5408 | 20          | PS415Q08  |
| 21             | 20.0615                | 20   | .4541 | 21          | PS425Q02  |
| 21             | 56.9684                | 20   | .0000 | 22          | PS425Q03  |

---

# Conclusioni

- La capacità dello strumento di produrre gli stessi risultati in prove ripetute é pressoché identica nei tre casi.
- Tutti gli item utilizzati mostrano un buon adattamento (le statistiche inflat non presentano per nessun item valori fuori dal range)
- Le item map mostrano che il test di matematica risulta complessivamente più difficile rispetto alle altre discipline.
- Sono presenti item che hanno un funzionamento significativamente differente tra le Regioni

dunque

- sono auspicabili politiche di intervento indirizzate rendere omogenei i livelli di competenza dei quindicenni italiani

**(PROGETTO OCSE PISA 2015-OBIETTIVO 500)**

---

---

# Prospettive

- Analisi dei dataset regionali.
- ?

