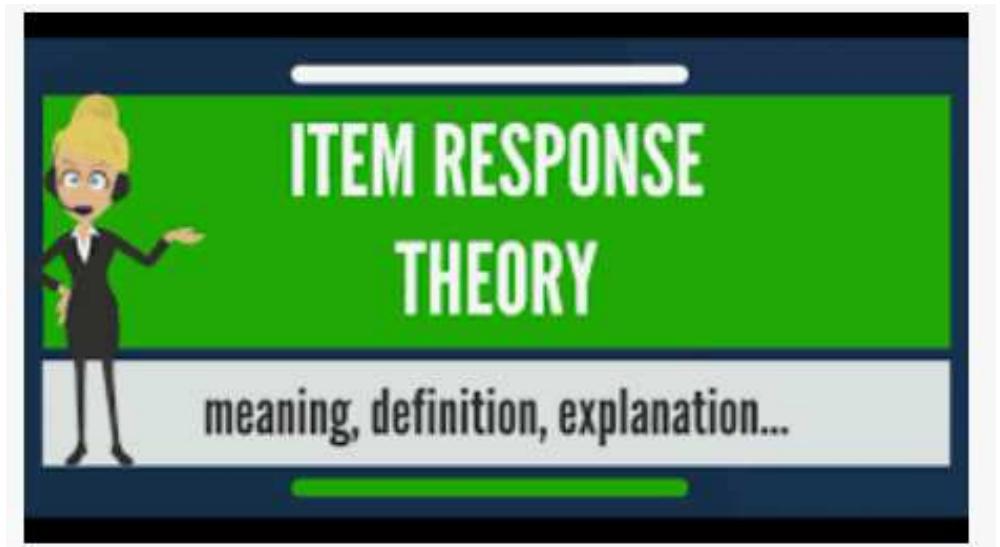


Statistica e Valutazione della Formazione Slides

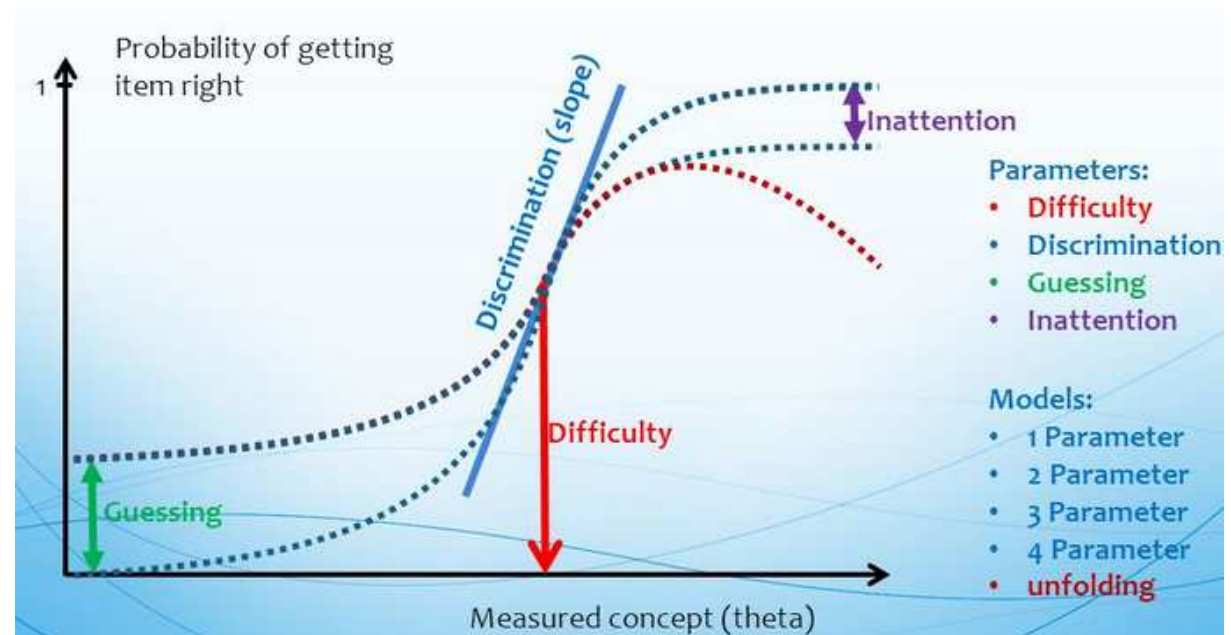
A.A. 2025-2026

Docente: ANNA LINA SARRA

I modelli di risposta all'item



I modelli dell'Item Response Theory (IRT)



Il problema della misurazione

- ✓ L'obiettivo finale è quello di misurare la conoscenza, competenza degli studenti, che è una variabile non osservabile (**variabile latente**).
- ✓ Lo strumento di misura è un test contenente un insieme di n item
- ✓ TCT vs. IRT
- ✓ Nella TCT, si utilizza il punteggio grezzo osservato (es. somma delle risposte corrette, se gli item sono binari). L'Alpha di Cronbach e la correlazione punto biseriale sono strumenti utili per studiare le caratteristiche degli item e del test

ITEM RESPONSE THEORY (IRT)

- ✓ Moderno approccio di studio delle proprietà degli item e di analisi dei risultati
- ✓ Il focus principale dell'IRT è l'**item** (da qui il nome IRT) invece del test e del punteggio al test
- ✓ Definisce un modello statistico per la probabilità di risposta ad ogni item del test

Si compone di due fasi:

1. **CALIBRAZIONE** (stima delle proprietà psicometriche degli item)
2. **SCORING** (stima delle abilità degli individui)

UN PO' DI STORIA.....

- **1943-44 Lawley** propone i primi modelli di IRT ma che si basano ancora su condizioni molto restrittive
- **1946 Tucker** è il primo psicometrico ad introdurre una funzione del livello di performance rispetto ad una variabile indipendente
- **1952-53 Lord** segna la nascita dell'IRT sviluppando per primo un metodo di stima e di applicazione del modello (basato sulla normale cumulata)
- **1952 Birnbaum** introduce il modello logistico e quello normale
- **1960-66 Rasch** sviluppa il modello applicato agli *slow readers* per cercare di confrontare i loro punteggi
- **1968 Lord e Novick** pubblicazione del testo fondamentale con 5 capitoli dedicati all'IRT (4 sono di Birnbaum)
- **1980 Lord** pubblicazione del testo considerato la pietra angolare della moderna IRT
- **Testi di riferimento:** van der Linden e Hambleton (1997), Embretson e Reise (2000), Baker e Kim (2004), De Ayala (2009)

AMBITI DI APPLICAZIONE

Educational and psychological measurement

Alcuni esempi:

- **Graduate Record Examination** (GRE): test standardizzato + CAT, requisito di ammissione a molte graduate schools negli Stati Uniti. Abilità rilevate: verbal reasoning, quantitative reasoning, analytical writing and critical thinking skills.
- **Graduate Management Admission Test** (GMAT): test standardizzato + CAT, utilizzato come requisito di ammissione alle business graduate schools negli Stati Uniti (es. per prendere un MBA) . Abilità rilevate: matematica e lingua Inglese
- **Test of English as a Foreign Language** (TOEFL)
- Prove **INVALSI** (solo calibrazione e non scoring...per ora!)

CARATTERISTICHE MODELLI IRT

Metodo basato su un modello

- ❑ Il modello esprime la probabilità di rispondere in un certo modo ad un item, in funzione delle proprietà dell'item stesso e dell'abilità latente dell'individuo
- ❑ Comparabilità dei risultati
- ❑ Indipendenza dal campione
- ❑ Su questi modelli si basano tecniche avanzate di
 - **LINKING** (comparabilità tra test diversi),
 - **EQUATING** (comparabilità tra punteggi ottenuti in test diversi)
 - **SCALING** (trasformazione del punteggio)
- ❑ Studio del **DIF** (Differential Item Functioning)

La teoria della Risposta all'Item (IRT)

Il concetto che sta alla base della IRT è che la risposta di un soggetto ad un test può essere spiegata da un set di fattori (o tratti latenti) e parametri.

Si definisce **fattore o tratto**



quella **variabile latente** che si assume influenzi e determini la riuscita in una dato compito.

Esempio: l'abilità spaziale può essere considerata un tratto necessario per risolvere test cognitivi.

Si definiscono parametri



le caratteristiche che sono possedute dalla prova sottoposta al soggetto.

La teoria della Risposta all'Item (IRT)

Convenzionalmente il tratto latente viene denominato nella IRT abilità ed è rappresentato dalla lettera greca theta (Θ). Ogni soggetto possiede una diversa quantità di abilità, per cui ogni individuo sottoposto ad un stesso compito avrà una performance diversa.

La risposta del soggetto ad un compito tuttavia non può essere spiegata solo dal livello di abilità (necessaria per quel compito) posseduto dal soggetto, ma dipende anche da **alcuni parametri:**

- livello di difficoltà dell'item
- livello di discriminatività dell'item
- incidenza del caso, o parametro di guessing
- trascuratezza (carelessness)

Modelli della Teoria della Risposta dell'Item

La Teoria della Risposta all'Item raggruppa **all'interno di sé una serie di modelli** che si differenziano in funzione:

- **NUMERO DI ABILITÀ MISURATE DAL TEST**
- **NUMERO DI PARAMETRI PRESI IN CONSIDERAZIONE**
- **FORMATO DI ITEM CHE SONO SOMMINISTRATI**

Modelli della Teoria della Risposta dell'Item in base al numero di abilità misurate

NUMERO DI ABILITA' MISURATE
DAL TEST

Quando un test misura solo un'abilità, o in altre parole solo un'abilità si ipotizza essere la variabile latente necessaria per risolvere un test, allora si parla di Modello TRI **UNIDIMENSIONALE**.

Viceversa nel caso in cui un test misuri un set di abilità, allora si parla di Modello TRI **MULTIDIMENSIONALE**.

Modelli della Teoria della Risposta dell'Item in base al numero di parametri

NUMERO DI PARAMETRI PRESI IN
CONSIDERAZIONE

- È possibile identificare quattro diversi tipi di modelli. Si definisce Modello Logistico ad un parametro (1PL) o Modello di Rasch il modello in cui si ipotizza che solo il parametro di **difficoltà** possa influenzare la risposta del soggetto all'item.
 - Quando oltre alla difficoltà si considera anche il parametro di **discriminatività** dell'item, allora si parla di Modello Logistico a due parametri (2PL).
 - Si definisce Modello Logistico a tre parametri (3PL) il modello in cui sono considerati contemporaneamente tutti i parametri degli item ovvero **difficoltà, discriminazione e guessing**.
 - Infine abbiamo il Modello Logistico a quattro parametri (4PL) che considera il caso in cui un individuo con abilità molto elevate potrebbe non presentare una probabilità di successo uguale a 1.
-

Modelli della Teoria della Risposta dell'Item in base al tipo di risposta

FORMATO DI ITEM CHE SONO
SOMMINISTRATI

È possibile infine fare una distinzione in funzione del formato degli item, che può essere di tipo **dicotomico o politomico**.

Per **item dicotomico** si intende un item che ha **due alternative di risposta**.

Tuttavia possono rientrare in questa categoria anche gli item che presentano un ventaglio di alternative tra cui comunque c'è una risposta giusta, e quindi le alternative presenti possono essere ricondotte alla forma dicotomica.

Gli **item politomici** propriamente detti sono invece quegli item che presentano più alternative di risposta. È questo il caso di test di personalità o scale di atteggiamento.

I modelli per dati dicotomici

Se si considerano gli **item dicotomici**, è possibile individuare **due famiglie di modelli**

I MODELLI LOGISTICI

**I MODELLI
DELL'OGIVA
NORMALE**

In entrambi i casi

lo scopo è quello di **calcolare la probabilità di rispondere correttamente ad un item in funzione del livello di abilità del soggetto e dei parametri degli item.**

In entrambe le famiglie

è possibile individuare quattro diversi modelli in base al numero di parametri presi in considerazione, in cui il calcolo della probabilità avviene con funzioni matematiche diverse, che producono comunque risultati molto simili.

Dati dicotomici: esempio

L'esempio canonico proviene da test educativi in cui gli item sono domande d'esame e i punteggi sono: $Y_{ik} = 1$ se l'esaminato i ha risposto correttamente all'item k e $Y_{ik} = 0$ altrimenti.

	Item 1	Item 2	Item 3	Item 4	Item 5
Person 1	1	1	1	1	1
Person 2	0	1	1	1	1
Person 3	0	0	1	1	1
Person 4	0	0	0	1	1
Person 5	0	0	0	0	1

Fonte: Chong Ho Yu (2016) A Simple Guide to the Item Response Theory (IRT) and Rasch Modeling - <http://www.creative-wisdom.com/computer/sas/IRT.pdf>

Dati dicotomici: esempio

	Item 1	Item 2	Item 3	Item 4	Item 5	Average
Person 1	1	1	1	1	1	1
Person 2	0	1	1	1	1	0.8
Person 3	0	0	1	1	1	0.6
Person 4	0	0	0	1	1	0.4
Person 5	0	0	0	0	1	0.2
Average	0.8	0.6	0.4	0.2	0	

La media della colonna è ottenuta come $1 - \frac{\sum_{i=1}^n Y_{ik}}{n}$; e denota la difficoltà dell'item.

Fonte: Chong Ho Yu (2016) A Simple Guide to the Item Response Theory (IRT) and Rasch Modeling - <http://www.creative-wisdom.com/computer/sas/IRT.pdf>

Dati dicotomici: esempio

Due persone hanno lo stesso punteggio grezzo

Person 4	0	0	0	1	1	0.4
Person 5	0	0	0	0	1	0.2
Person 6	1	1	0	0	0	0.4

Fonte: Chong Ho Yu (2016) A Simple Guide to the Item Response Theory (IRT) and Rasch Modeling - <http://www.creative-wisdom.com/computer/sas/IRT.pdf>

Dati dicotomici: esempio

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Average
Person 1	1	1	1	1	1	0	0.83
Person 2	0	1	1	1	1	0	0.67
Person 3	0	0	1	1	1	0	0.50
Person 4	0	0	0	1	1	0	0.33
Person 5	0	0	0	0	1	1	0.33
Average	0.8	0.6	0.4	0.2	0	0.8	

La media della colonna è ottenuta come $1 - \frac{\sum_{i=1}^n Y_{ik}}{n}$, e denota la difficoltà dell'item.

Ci sono due item (Item1 e Item6) che hanno la stessa media

Fonte: Chong Ho Yu (2016) A Simple Guide to the Item Response Theory (IRT) and Rasch Modeling - <http://www.creative-wisdom.com/computer/sas/IRT.pdf>

Il Modello di Rasch

Il Modello di Rasch è il modello più semplice e il primo ad essere formulato negli anni '60.

La risposta del soggetto viene ipotizzata essere

- influenzata **dal livello di abilità posseduto dal soggetto**
- e da **un solo parametro dell'item, il livello di difficoltà**

La funzione matematica che descrive questo modello è la seguente:

$$P(X_{ik} = 1 | \theta_i, \beta_k) = \frac{e^{(\theta_i - \beta_k)}}{1 + e^{(\theta_i - \beta_k)}}$$

Il Modello di Rasch

L'idea di base è molto semplice.

La risposta del soggetto è governata da due componenti (le incognite del modello, o “parametri”).

La prima componente è la capacità o abilità del soggetto nella scelta della risposta.

Esempio:

preparazione in matematica di uno studente. Maggiore è la sua preparazione, maggiore è la probabilità di dare la risposta esatta.

La seconda componente è rappresentata dalla difficoltà delle domande (item).

Esempio:

una sottrazione ha una maggiore probabilità di ricevere una risposta corretta rispetto a un'equazione.

Obiettivo:

assegnare un valore a ogni soggetto sulla base della sua abilità e un valore a ogni item sulla base della sua difficoltà.

Il Modello di Rasch

La risposta del soggetto scaturisce dall'interazione tra capacità e difficoltà:

maggiore è la preparazione rispetto a una determinata domanda, maggiore è la probabilità di fornire la risposta corretta

simmetricamente, maggiore è la difficoltà rispetto alla preparazione di un dato studente, minore è la probabilità di una risposta corretta.

Ma come si quantifica questa interazione?

Attraverso un confronto tra capacità e difficoltà, che si esegue con un semplice sottrazione tra le due entità (che ovviamente utilizzano la stessa unità di misura).

Il Modello di Rasch

Da questa prima operazione possono scaturire tre casi:

a. capacità = difficoltà

P (probabilità di una risposta corretta) = 50%

b. capacità maggiore di difficoltà

P > 50%

c. capacità minore di difficoltà

P < 50%

Unità di misura: il Logit

Il Modello di Rasch in realtà risulta essere una evoluzione di un'altra formula che Rasch teorizzò negli anni '60.

Agli inizi degli anni '60 Rasch interpretò la performance di un soggetto ad un compito come una variabile dipendente dalla differenza tra il parametro dell'abilità e quello della difficoltà.

Questo concetto fu formalizzato nel seguente modo:

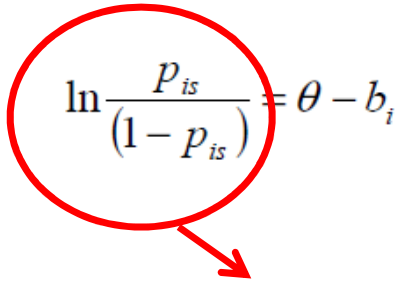
$$\ln \frac{P_{is}}{(1 - P_{is})} = \theta - b_i$$

dove:

P_{is} = numero di successi (superamento dell'item).

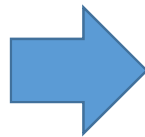
$(1 - P_{is})$ = numero di insuccessi (non superamento dell'item).

Unità di misura: il Logit

$$\ln \frac{p_{is}}{(1 - p_{is})} = \theta - b_i$$


logaritmo naturale di un odds

espresso dal rapporto tra numero di successi e numero di fallimenti), che può essere spiegato sempre dalla differenza tra il livello di theta posseduto dal soggetto e il livello di difficoltà dell'item.



Il logaritmo naturale di un odds rappresenta un'unità di misura lineare detta LOGIT, che permette di misurare sullo stesso continuum sia i soggetti (ovvero il loro livello di theta), sia la difficoltà.



In un secondo momento Rasch sostituisce il logaritmo naturale di un odds con il concetto di probabilità, che essendo un concetto più intuitivo, ha reso il modello matematico più facile da capire.

Unità di misura: il Logit

OSSERVAZIONE:

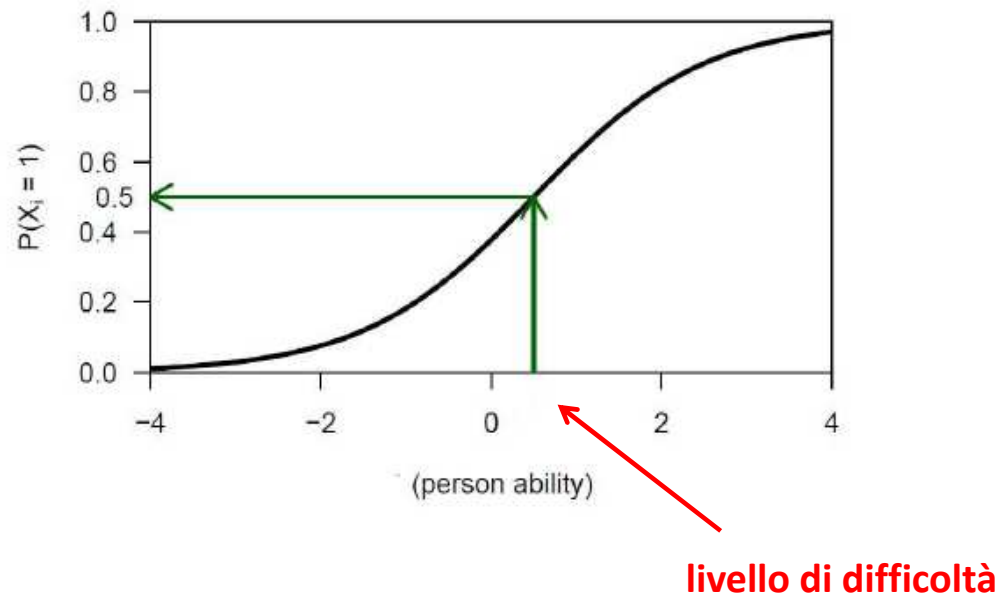
La probabilità di un evento può assumere valori compresi tra **0 (l'evento non si verifica)** e **1 (evento certo)**, mentre la differenza tra abilità e difficoltà può assumere valori compresi tra **$-\infty$** e **$+\infty$** .



Quindi pur mantenendo come unità di misura il logit, non nella forma di $\ln$ di odds ma come $(\theta - b_i)$, si rese necessaria una correzione della formula al fine di portare la differenza in un range compreso tra 0 e 1. Rasch quindi considerò $(\theta - b_i)$ come esponente di “e” così da avere $e^{(\theta - b_i)}$. In questo modo la differenza oscillava tra 0 e $+\infty$. Ma ciò ancora non bastava; perciò fu divisa l’espressione esponenziale per un “fattore di normalizzazione”, cioè $1 + e^{(\theta - b_i)}$.

Curva caratteristica dell' Item (Item Characteristic Curve)

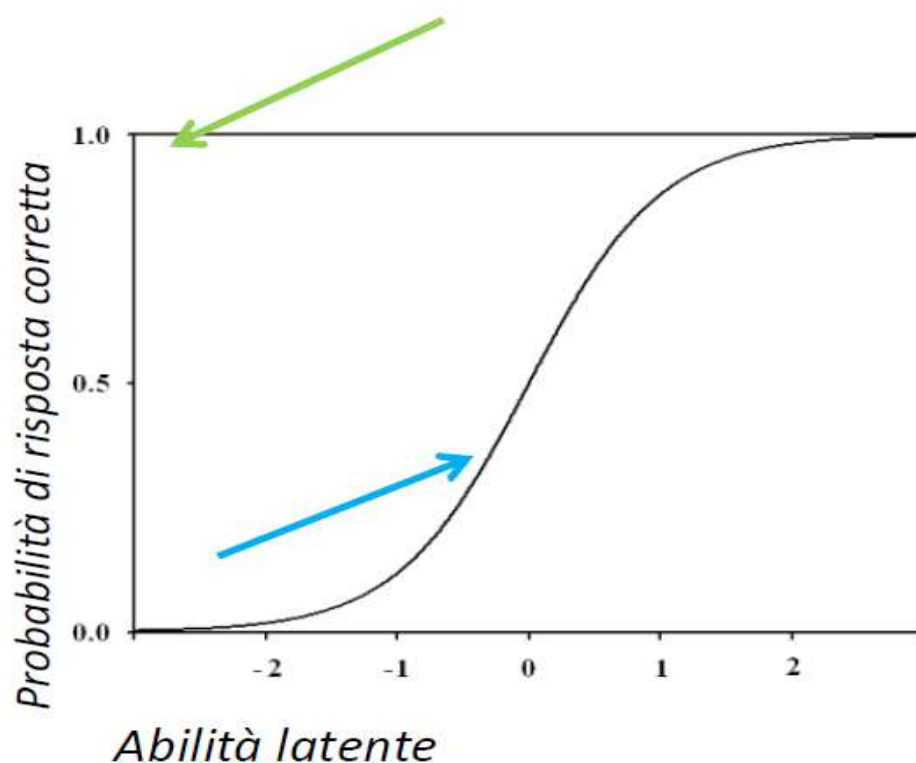
È possibile rappresentare graficamente su piano cartesiano la relazione espressa dalla funzione matematica del modello, mediante la Curva Caratteristica dell'Item (ICC), ponendo in ascissa l'abilità ed in ordinata la probabilità di risposta corretta.



Come si interpreta la curva?

Più è alta, più l'item è facile

*Più **sale rapidamente** più è discriminativa*



La formula del Modello di Rasch, utilizza come unità di misura il logit per cui la relazione ipotizzata tra le variabili non può essere spiegata con una relazione lineare ma logistica.

Curva Caratteristica dell'Item (Item Characteristic Curve)

La **Curva Caratteristica dell'Item** appare come un funzione monotona crescente, e rappresenta la probabilità di rispondere correttamente ad un item per i diversi livelli di abilità in funzione della difficoltà dell'item stesso.

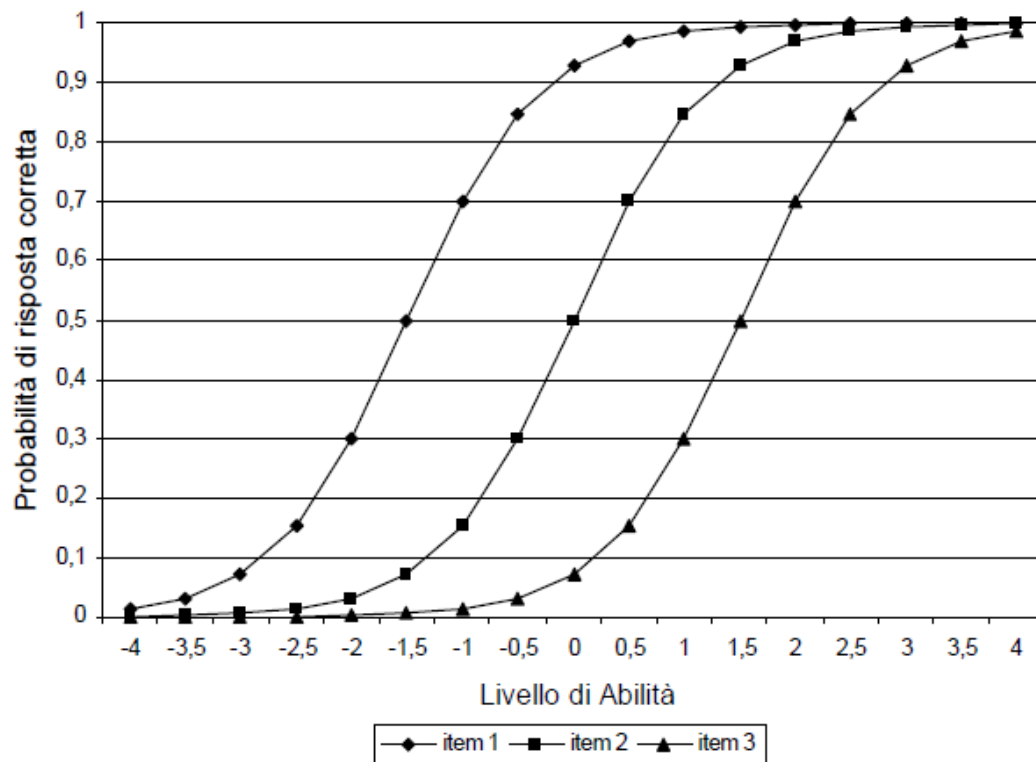
La curva tende ad approssimarsi asintoticamente

- ✓ a $-\infty$ per valori molto bassi di theta (cui è associata una probabilità di rispondere correttamente che tende a 0),
- ✓ e a $+\infty$ valori molto alti di theta (cui associata una probabilità di rispondere correttamente che tende a 1)

Anche il **parametro difficoltà** può assumere concettualmente valori compresi tra - ∞ e + ∞ ma convenzionalmente si utilizza anche in questo caso un range più circoscritto compreso tra -4 e $+4$.

Curva Caratteristica dell'Item (Item Characteristic Curve)

Item diversi con diversi livelli di difficoltà avranno quindi diverse ICC come rappresentato nella seguente Figura



Il parametro difficoltà è direttamente responsabile del minore o maggiore spostamento della curva verso destra sulla scala di abilità, quindi viene definito location della curva.

Il Modello Logistico ad un parametro (1PL)

Il Modello Logistico ad un parametro (1PL), fu introdotto nel 1968 da Birnbaum.

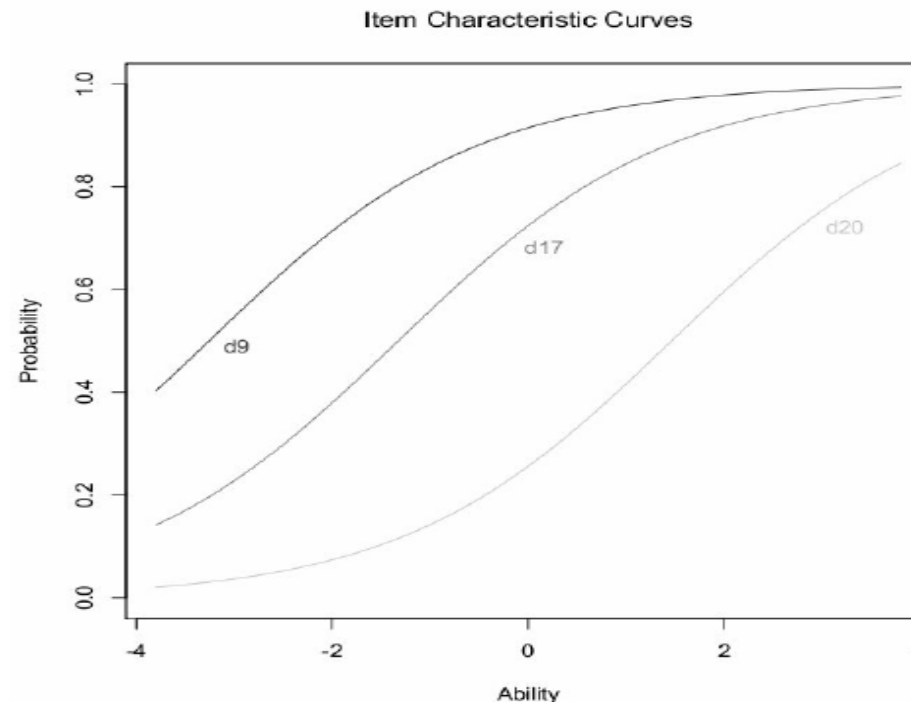


Il modello equivale sostanzialmente al modello di Rasch con la differenza che introduce il parametro di discriminatività ponendolo uguale a 1.

Il Modello Logistico ad un parametro (1PL): Item Characteristic Curves

Nel Modello Logistico ad un parametro (1PL) la difficoltà è l'unico parametro che varia, in quanto il parametro di discriminazione è vincolato ad 1.

Si avrà che la ICC assume una forma influenzata solo da questo parametro



Il Modello Logistico a due parametri (2PL)

Il Modello Logistico a due parametri (2PL), fu teorizzato nel 1968 da Birnbaum.

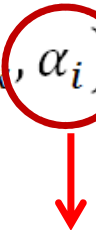
Il modello è una trasformazione della versione sviluppata nel 1952 da Lord, che si basa su una distribuzione normale cumulata e prende il nome di Modello della Ogiva Normale a due parametri.



Questa seconda classe di modelli del TRI considera la possibilità che gli item differiscano tra di loro non solo per il livello di difficoltà ma anche per come discriminano tra i soggetti ed introduce un secondo parametro: il parametro di discriminatività

Il Modello Logistico a due parametri (2PL)

La versione logistica del modello è quella più diffusa perché gode di proprietà statistiche più desiderabili. L'equazione matematica del 2PL è la seguente:

$$P(X_{ik} = 1 | \theta_i, \beta_k, \alpha_i) = \frac{e^{\alpha_i(\theta_i - \beta_k)}}{1 + e^{\alpha_i(\theta_i - \beta_k)}}$$


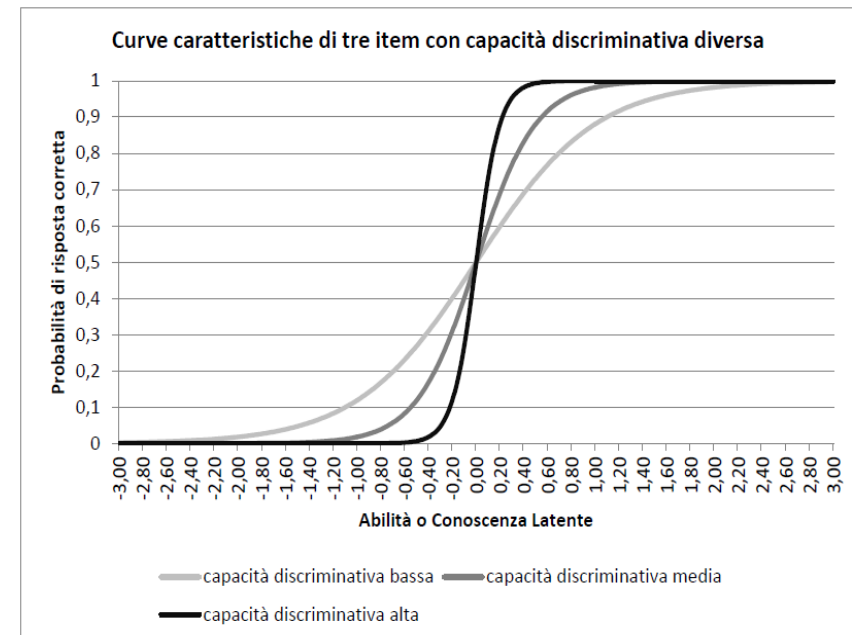
parametro di discriminazione dell'item.

Teoricamente può assumere valori compresi tra $\pm \infty$ anche se convenzionalmente varia tra 0.5 e 2.5

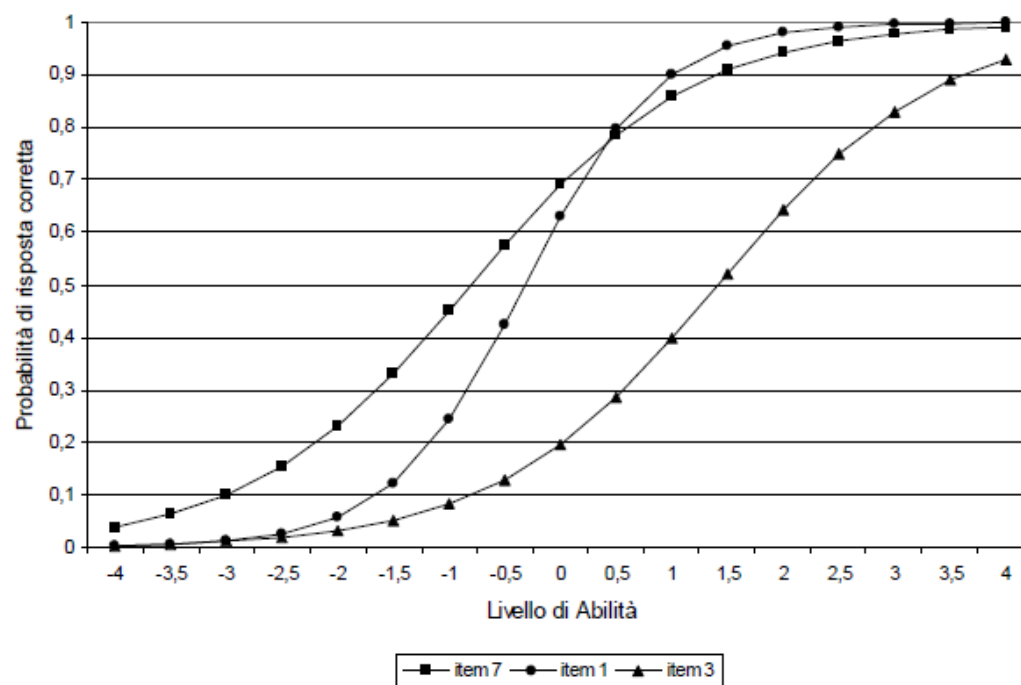
Il Modello Logistico a due parametri (2PL)

La **discriminazione** indica il potere dell'item di discriminare i soggetti sopra e sotto un livello di theta scelto, ovvero la **capacità dell'item di distinguere tra soggetti che presentano livelli differenti di abilità.**

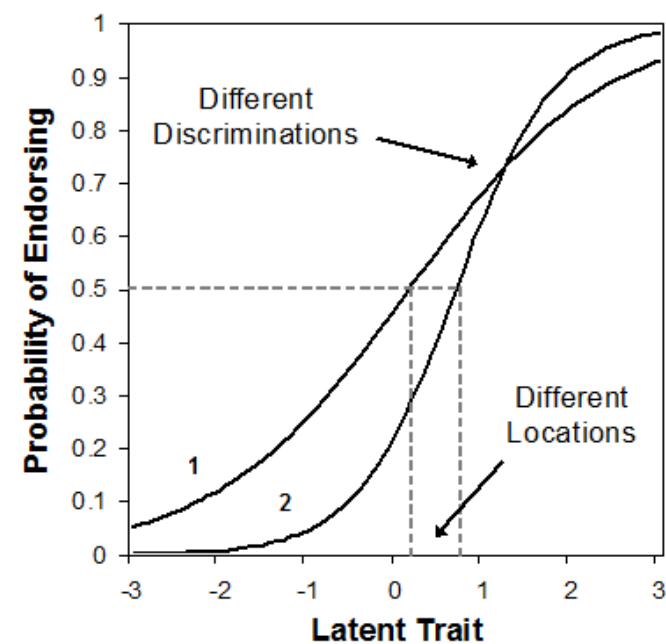
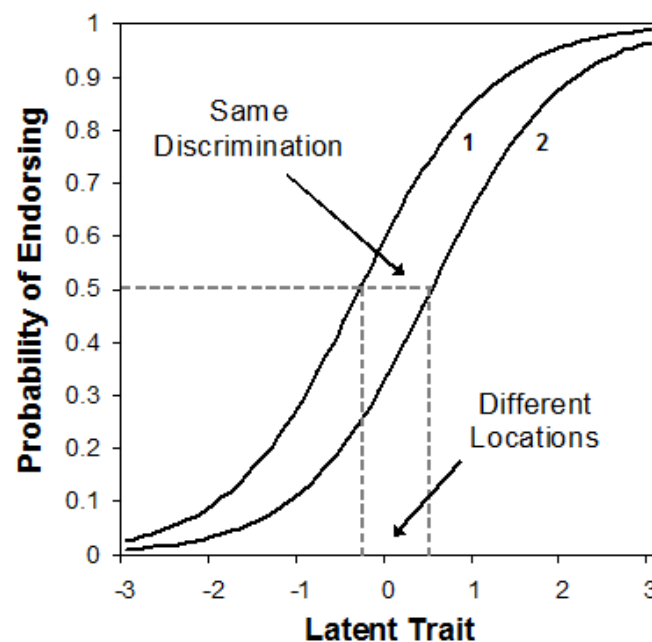
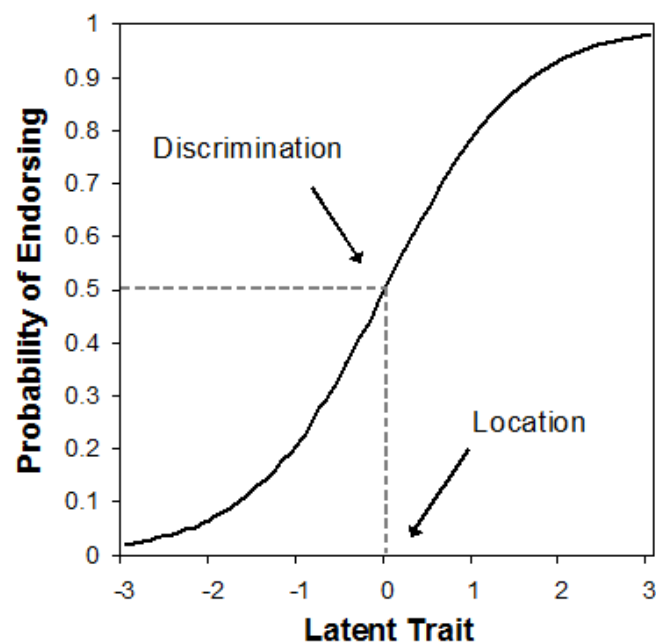
Quindi le ICC di tre item con diverso potere discriminante (ma con uguale indice di difficoltà) si presentano nel seguente modo:



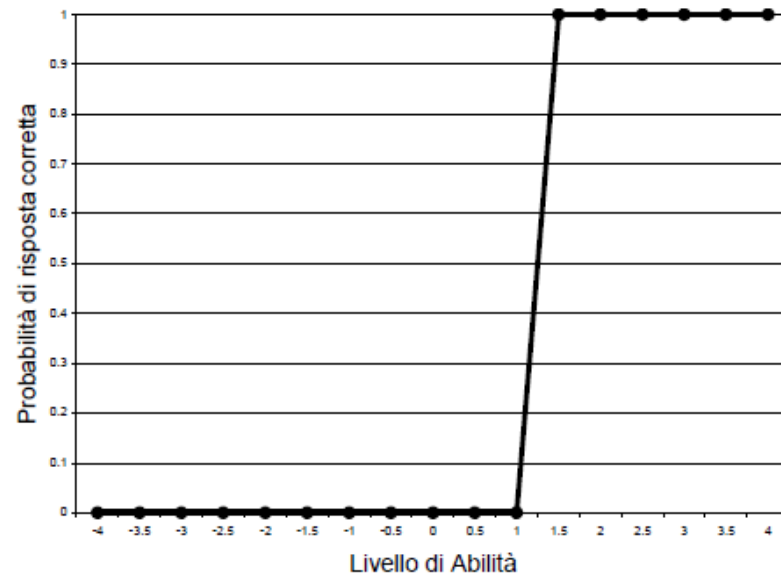
ICC di tre item con diverso livello di difficoltà e discriminazione



Parametro di discriminazione



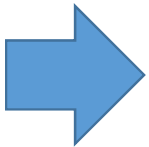
ICC di un item con perfetta discriminazione.



In questo caso, considerando come livello theta di interesse il valore di 1, l'item discrimina perfettamente i soggetti con un livello di abilità inferiore a tale valore da quelli con livello superiore.

Discriminazione negativa

Le ICC che abbiamo visto sono tutte curve monotone crescenti, ovvero all'aumentare del livello di abilità aumenta anche la probabilità di rispondere correttamente all'item. È possibile che la curva possa presentare **un andamento anomalo:**



all'aumentare del livello di abilità si assiste ad una diminuzione della probabilità di rispondere correttamente.

In questo caso si parla di **discriminazione negativa**



Per risposte corrette



Per risposte non corrette

Il modello Logistico a tre parametri (3PL)

Guessing effect:

Finora abbiamo considerato il caso in cui al decrescere dell' abilità di un soggetto, la sua probabilità di superare l'item tenda ad annullarsi completamente. Si possono, però, dare casi in cui tale limite asintotico si attesta su valori della probabilità superiori a zero.

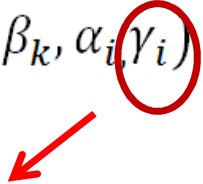
Il Modello Logistico a tre parametri (3PL) risale a Birnbaum (1968), ed è il caso in cui si considerano simultaneamente tre parametri degli item:

- ✓ Difficoltà
- ✓ Discriminazione
- ✓ Guessing

Il modello Logistico a tre parametri (3PL)

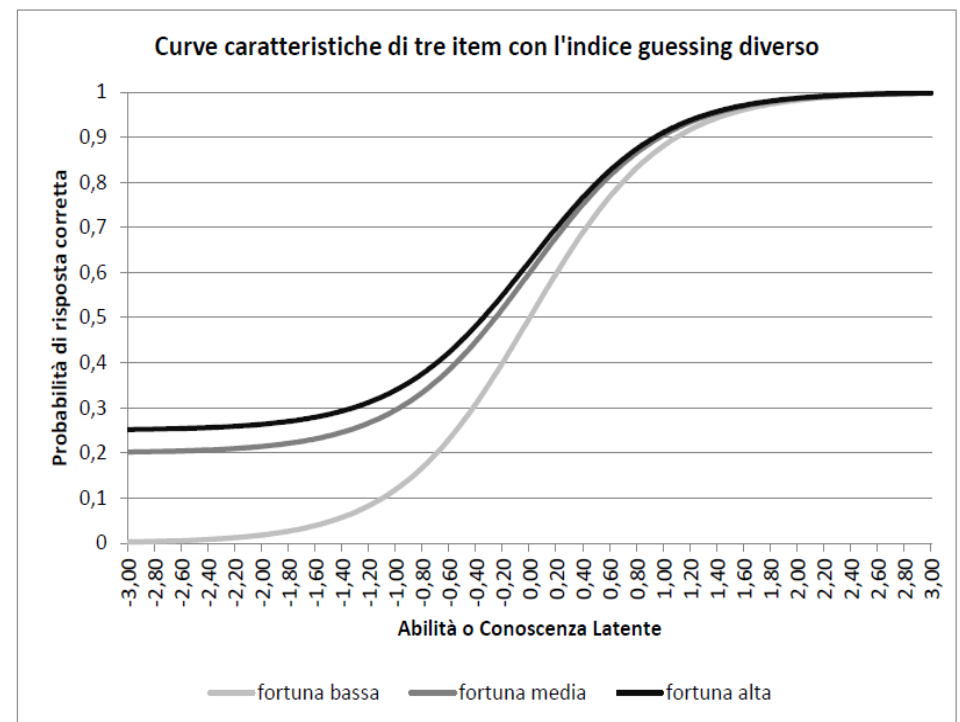
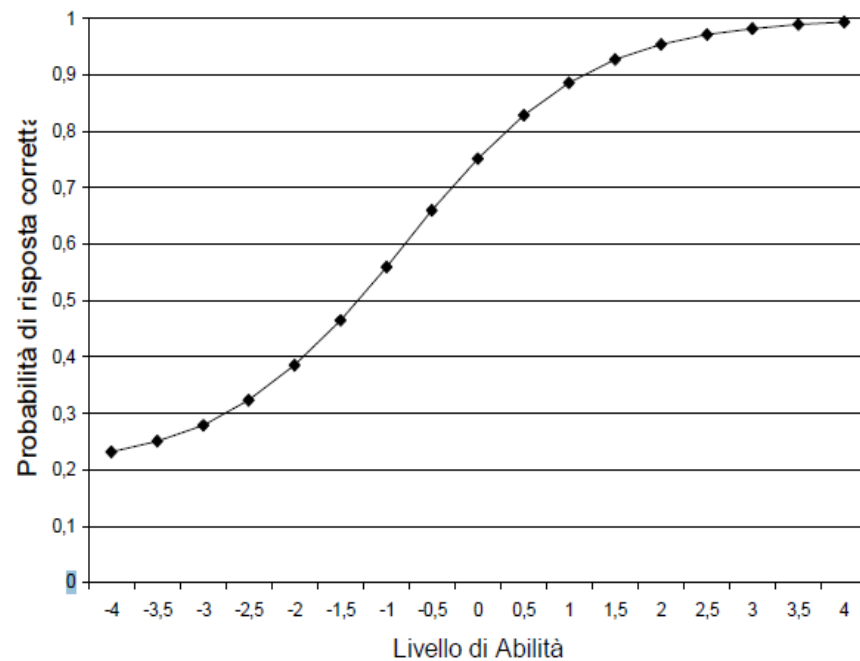
Il parametro guessing viene introdotto nell'equazione per rendere ragione del fatto che un soggetto con basso livello di abilità può rispondere correttamente ad un item semplicemente tirando ad indovinare; quindi questo parametro quantifica l'incidenza del caso nel fornire una risposta soprattutto nel caso di item a scelta multipla.

L'equazione matematica del 3PL
è la seguente

$$P(X_{ik} = 1 | \theta_i, \beta_k, \alpha_i, \gamma_i) = \gamma_i + (1 - \gamma_i) \frac{e^{\alpha_i(\theta_i - \beta_k)}}{1 + e^{\alpha_i(\theta_i - \beta_k)}}$$


asintoto quando quando theta tende a - ∞

La ICC di un item secondo il Modello 3PL



Il modello Logistico a quattro parametri (4PL)

Carelessness (inaccuratezza) effect:

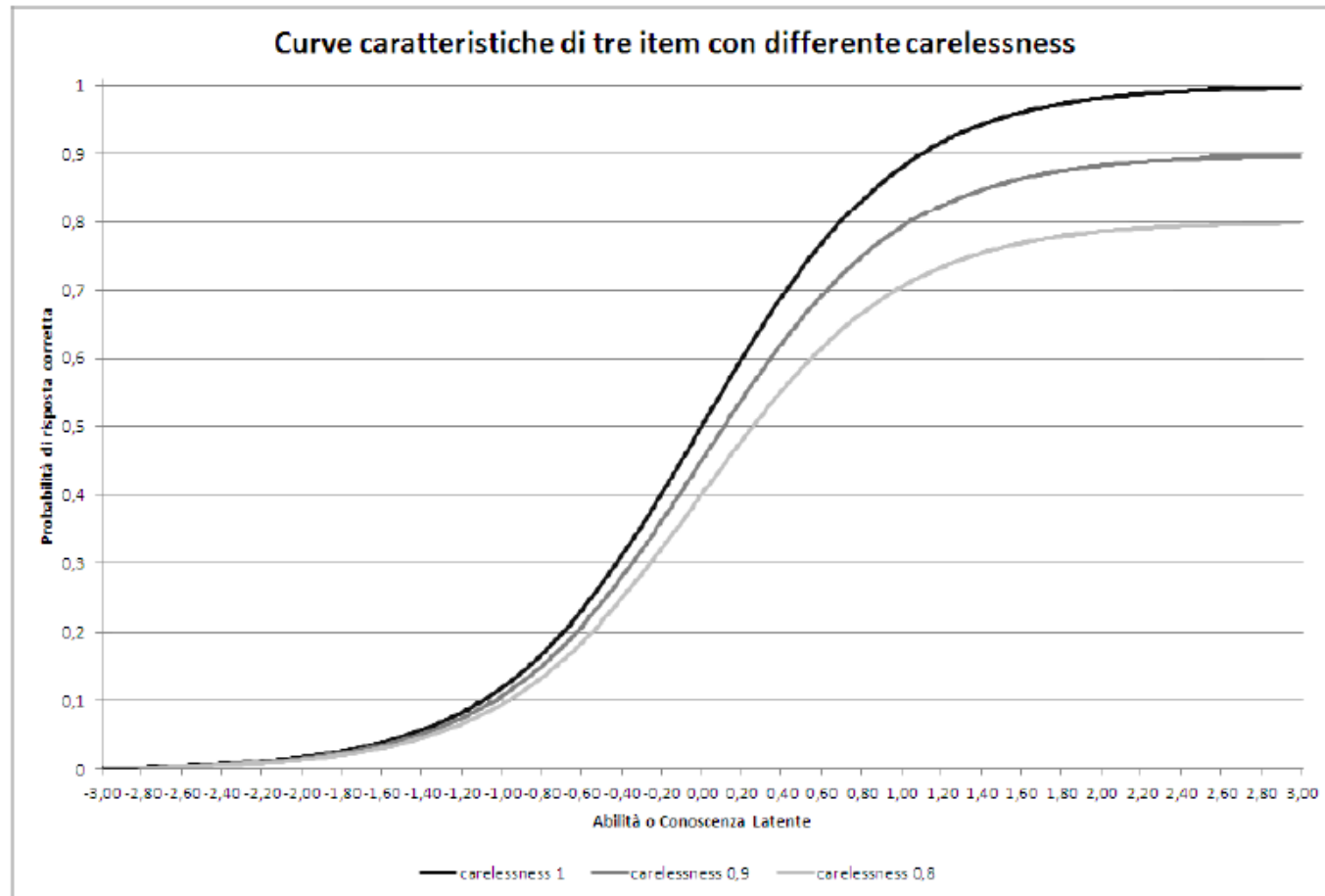
Quando i modelli logistici vengono applicati in campo valutativo si può prendere in considerazione anche il caso in cui un soggetto con alti livelli di abilità non necessariamente rispondono in modo corretto con un' elevata probabilità.
In questo caso viene introdotto un quarto parametro che determina un asintoto superiore.

Il Modello Logistico a quattro parametri (4PL) introdotto da Barton e Lord (1981)

considerano simultaneamente quattro parametri degli item:

- ✓ Difficoltà
- ✓ Discriminazione
- ✓ Guessing
- ✓ Carelessness

ICC per un modello Logistico a quattro parametri (4PL)



Assunzioni e proprietà dei modelli di TRI

Partendo dall'idea di base che la probabilità di rispondere correttamente ad un item sia condizionata dal livello di abilità posseduto dal soggetto e da una serie di parametri che caratterizzano l'item stesso, è possibile individuare una serie di assunzioni e proprietà contenuti nella IRT, indipendentemente dal tipo di modello considerato.

Assunzioni dell'IRT

Monodimensionalità

L'assunzione di monodimensionalità o unidimensionalità, assume che una sola abilità sia misurata dal test. Di fatto possono esserci diversi fattori che concorrono a determinare la risposta all'item ma c'è una sola dimensione dominante.

Nel caso di modelli multidimensionali, naturalmente questa assunzione diventa l'assunzione di multidimensionalità, ovvero si assume che una serie di fattori (e sono specificati quali) sono effettivamente misurati dal test.

Assunzioni:

Indipendenza locale o specifica oggettività

L'altra assunzione molto importante e strettamente connessa con la precedente è quella di indipendenza locale.

Assumendo che un test misura una sola abilità latente (caso più semplice, ovvero unidimensionale), si può ragionevolmente ipotizzare che gli item sono correlati fra loro visto che misurano aspetti diversi dello stesso costrutto latente; e quindi anche le risposte di un soggetto ai diversi item di uno stesso test saranno correlate.

Se però si mantiene costante il livello di abilità di un soggetto, gli item non hanno più ragione di essere correlati, poiché si parzializzano rispetto al parametro che li lega. In questo caso, le risposte del soggetto ad un set di item non devono essere più correlate **ma indipendenti**.

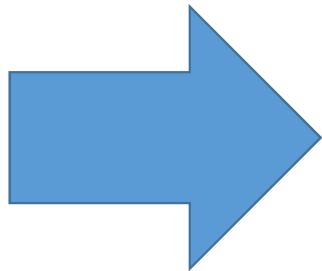
Assunzioni:

Monotonicità

La probabilità di **rispondere correttamente ad un item aumenta** in modo **monotono all'aumentare del livello di abilità posseduto** dal soggetto.

Ne consegue che le Curve Caratteristiche degli Item avranno sempre un andamento logistico crescente.

Violazioni di questa assunzione inficiano i risultati dei test di bontà dell'adattamento del modello ai dati.



Per applicare i modelli IRT è necessario **che i test somministrati non siano test di velocità,** ovvero che valutano l'abilità del soggetto a realizzare il test velocemente.

È necessario infatti che **la risposta errata sia dovuta alla carenza di abilità nel soggetto e non all'assenza di tempo a disposizione.**

Matrice dei risultati di un test

Prima di passare all'esame delle principali proprietà dei modelli della IRT è opportuno riportare nei suoi tratti essenziali la matrice delle risposte (soggetti x items)

studenti	item							
	1	2	...	j	...	k-1	k	
1	X_{11}	X_{12}	...	X_{1j}	...	X_{1k-1}	X_{1k}	Γ_1
2	X_{21}	X_{22}	...	X_{2j}	...	X_{2k-1}	X_{2k}	Γ_2
...								...
i	X_{i1}	X_{i2}	...	X_{ij}	...	X_{ik-1}	X_{ik}	Γ_i
...								...
N-1	X_{N-11}	X_{N-12}	...	X_{N-1j}	...	X_{N-1k-1}	X_{N-1k}	Γ_{N-1}
N	X_{N1}	X_{N2}	...	X_{Nj}	...	X_{Nk-1}	X_{Nk}	Γ_N
	S_1	S_2	...	S_j	...	S_{k-1}	S_k	

Risposta osservata ←

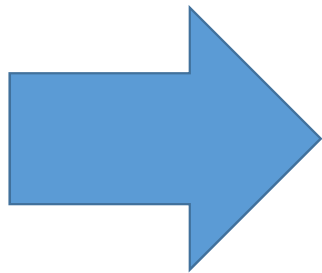
→ Punteggio complessivo di ciascun soggetto

→ Punteggio totale registrato da ciascun item

Proprietà:

Scala omogenea per item e soggetti

Il livello di abilità del soggetto e il livello di difficoltà dell'item vengono espressi con la stessa metrica, utilizzando la stessa scala di misura.



Questo permette una riduzione dell'errore di misurazione e una migliore stima del livello di abilità posseduto dal soggetto.

Proprietà: *Invarianza*

Se le assunzioni di unidimensionalità e di indipendenza locale vengono rispettate e se il modello si adatta ai dati empirici allora i parametri stimati godono della proprietà di invarianza. Questa proprietà implica

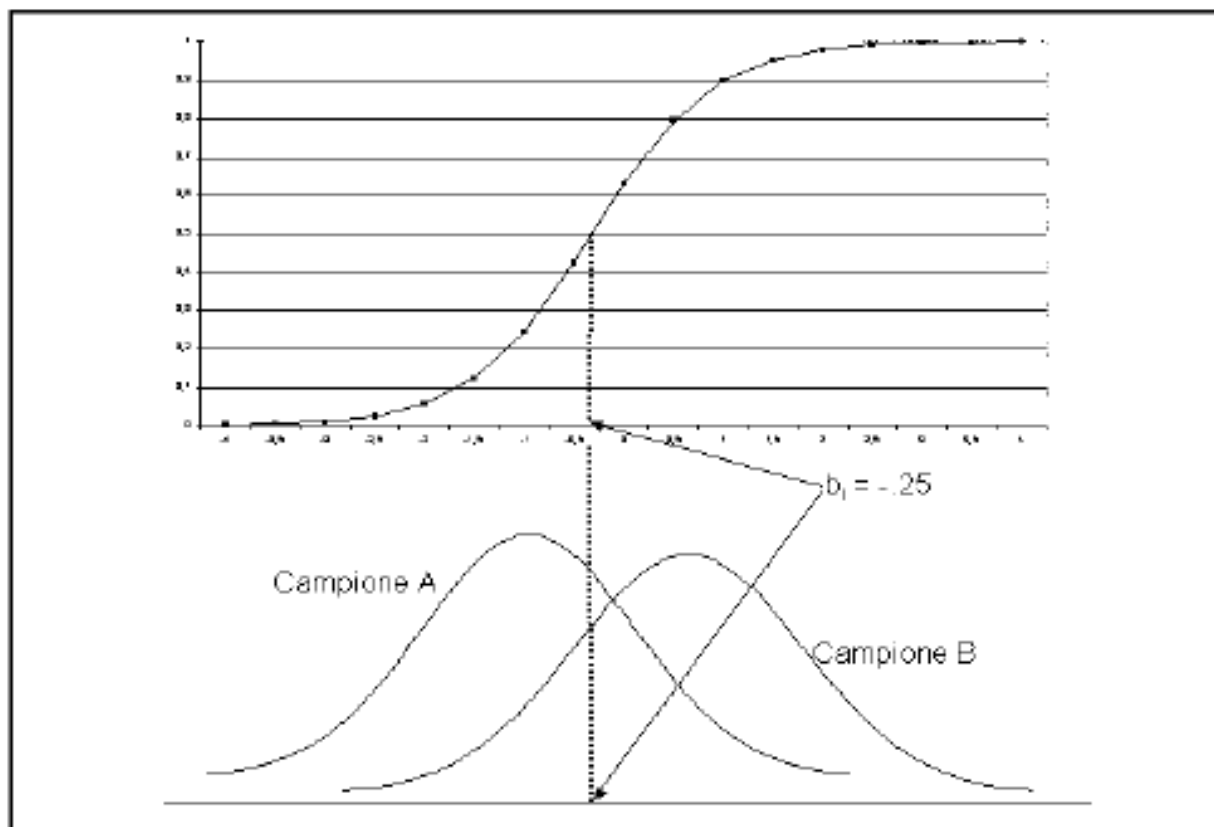


la stima del livello di abilità posseduta dal soggetto è indipendente dal test somministrato e utilizzato per tale stima (***test-free***).



In modo analogo anche il processo di stima dei parametri degli item è indipendente rispetto alla distribuzione di abilità del campione a cui sono somministrati (***sample-free***).

Invarianza



Invarianza ed oggettività specifica nel modello di Rasch

Il modello di Rasch gode

1. proprietà DELL'INVARIANZA DELLA MISURAZIONE (si estende anche agli altri modelli TRI)
2. sia della OGGETTIVITÀ SPECIFICA (è specifica del modello di Rasch)



✓ per stimare l'abilità di un soggetto non è necessario conoscere la difficoltà dell'item

✓ e viceversa per stimare la difficoltà dell'item non è necessario conoscere la distribuzione dell'abilità nel campione.

livello di difficoltà dell'item e numero complessivo di risposte corrette fornite allo stesso item sono

STATISTICHE SUFFICIENTI



Rasch tuttavia sviluppa ulteriormente il concetto di invarianza:

OGGETTIVITA' SPECIFICA

è possibile confrontare soggetti diversi indipendentemente dagli item somministrati, e confrontare item che misurano lo stesso costrutto ma che sono stati somministrati a soggetti diversi.

Oggettività specifica: esempio

Si consideri una situazione di comune riscontro in ambito scolastico: un test di calcolo a mente.

Si considerino gli item “2+2” e radice (169) cui si dà un punteggio dicotomico, vale a dire successo (1) o insuccesso (0).

Un allievo risponde ai due item. Egli riesce nell’item “2+2” e ottiene il punteggio di 1; lo stesso allievo fallisce nell’item radice di 169.

Appare chiaro che i punteggi ottenuti da una persona di capacità data dipendono dalla difficoltà dell’item.

Analogamente si considerino **due soggetti che rispondono all’item radice (169)**. Il primo risponde correttamente e ottiene il punteggio 1, il secondo dà una risposta sbagliata e ottiene il punteggio 0. Appare chiaro che i punteggi ottenuti da un item di difficoltà data dipendono dalla capacità della persona.

I punteggi ottenuti non soddisfano il criterio di oggettività specifica.

Il modello di Rasch soddisfa il criterio di oggettività specifica perché fornisce una misura della capacità della persona che non varia in funzione della difficoltà dell’item ed una misura della difficoltà dell’item che non varia in funzione della capacità della persona.

Modelli per dati politomici

Verso la fine degli anni '60 cominciano ad essere teorizzati dei modelli che hanno l'obiettivo di analizzare i dati politomici.

Per dati politomici

si intendono quei dati che presentano delle alternative di risposta che non vengono semplicemente dicotomizzate nelle categorie giusto o sbagliato.

Le alternative di risposta possono venire ordinate in modo **graduale** (come avviene nella scale tipo Likert) e sono molto utilizzate per misurare tratti di personalità, atteggiamenti e così via, dove non può esserci una risposta corretta.

Modelli per dati politomici

I modelli per dati politomici possono essere raggruppati in modelli indiretti e modelli diretti (Thissen e Steinberg, 1986).

Modelli indiretti



I modelli indiretti sono quelli in cui la probabilità che un soggetto con un livello j di θ ha di scegliere una precisa alternativa va calcolata attraverso **due step**. Nel primo step si calcolano tutte le probabilità associate ad ogni singola alternativa di risposta, nel secondo step si fa la differenza tra due probabilità di due alternative di risposta. Sono modelli indiretti il **Graded Response Model di Samejima (1969)** e il **Modified Graded Response Model di Muraki (1990)**.

Modelli diretti



I modelli diretti calcolano la probabilità di scegliere una alternativa attraverso un'unica formula, ovvero il rapporto tra l'esponenziale e la somma degli esponenziali (per tale ragione i modelli diretti sono anche definiti *divide-by-total*). Appartengono a questo gruppo il **Nominal Response Model di Bock (1972)**, il **Partial Credit Model di Masters (1982)** e il **Generalized Partial Credit Model di Muraki (1992. 1993)**.

Curva di Risposta di Categoria

È possibile calcolare le ICC per ogni alternativa di risposta dello stesso item. Tali curve vengono definite **Curve di Risposta di Categorie (CRC)**, e rappresentano la probabilità di scegliere le diverse alternative condizionate dal livello di theta posseduto dal soggetto.

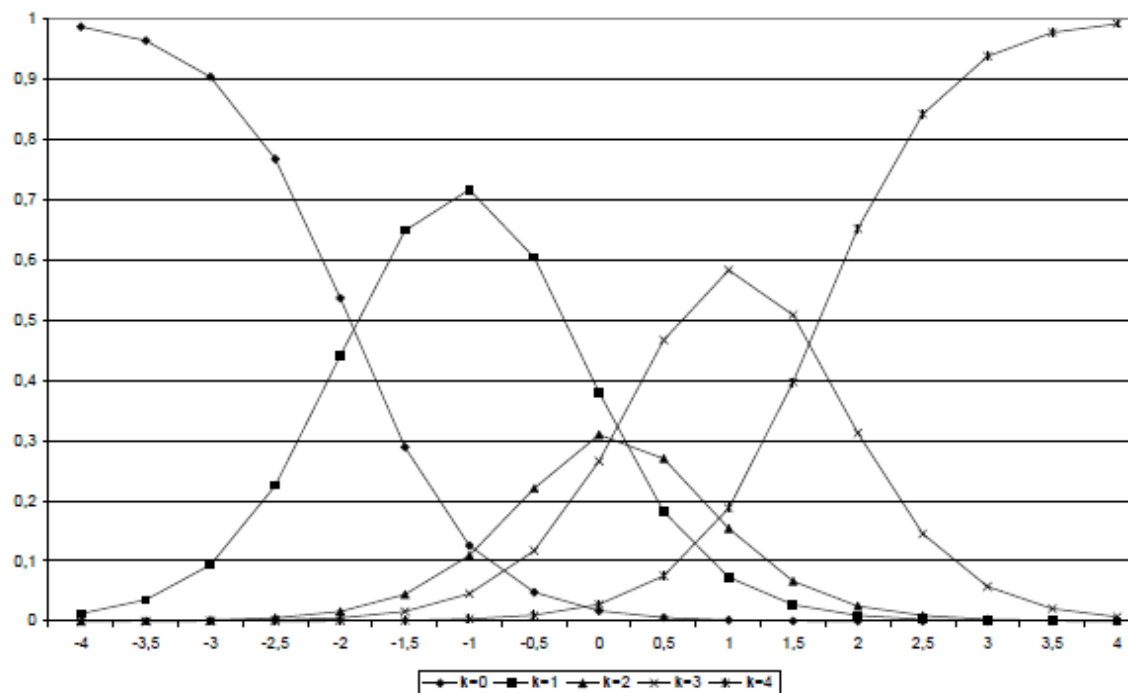
✓ **ogni alternativa di risposta** dell'item ha una sua *funzione di categoria di punteggio*, ovvero una funzione che esprime la probabilità per un soggetto con un dato livello di abilità (j), di fare una scelta o di ottenere un punteggio.

✓ **ogni curva** rappresenta la probabilità che un soggetto ha di scegliere un'alternativa piuttosto che quella successiva, intesa come probabilità di dare una risposta sopra o sotto una soglia di difficoltà.

✓ il **parametro di difficoltà** viene **interpretato come la soglia** che permette di dividere le diverse alternative di risposte

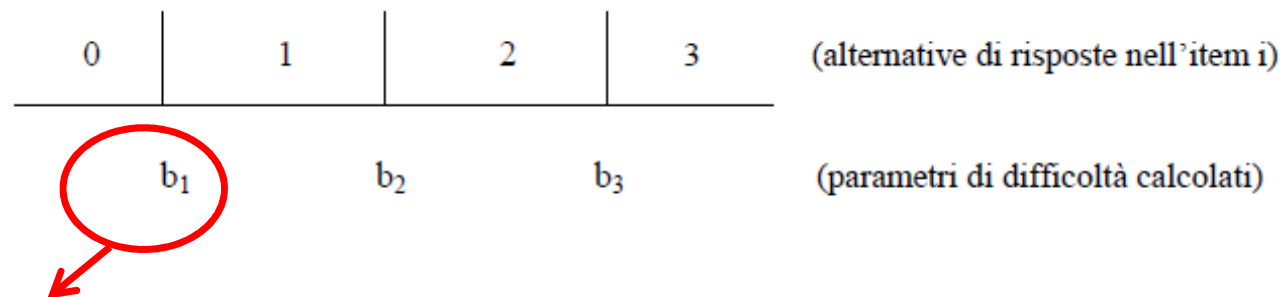
Curva di Risposta di Categoria

Curve di risposta della categoria relative alle cinque alternative dello stesso item



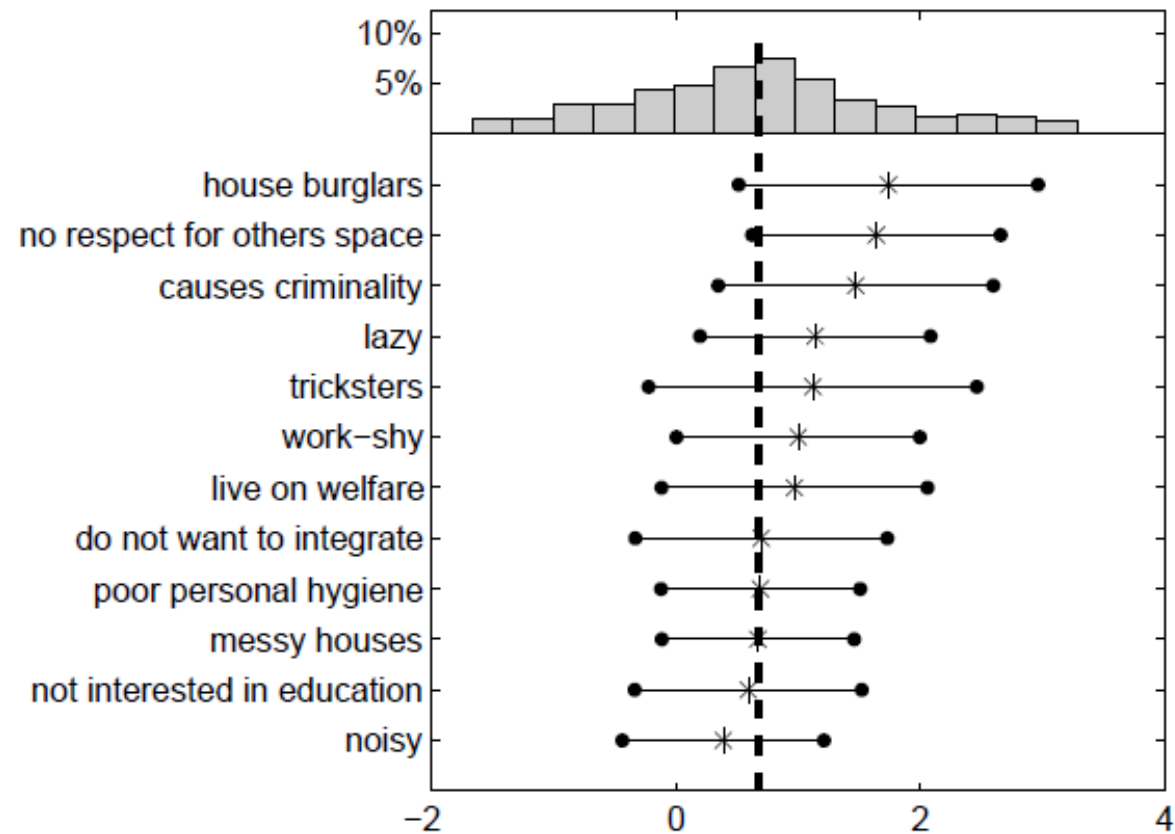
Esempio: scala tipo Likert con quattro alternative di risposta

Nel caso di una scala tipo Likert con quattro alternative di risposte, si calcolano tre livelli di difficoltà che permettono di dividere queste alternative.



- sotto tale livello c'è più probabilità di scegliere l'alternativa 0
- sopra questo livello c'è più probabilità di scegliere l'alternativa 1
- mentre a b_1 corrisponde una probabilità di 0,5 di scegliere l'alternativa 0 o 1.

Person-item map



Verifica delle assunzioni e delle proprietà

La Teoria della Risposta all'Item può essere sottoposta a **tre tipi di verifiche**:

- I. verifica delle assunzioni
- II. verifica delle proprietà
- III. verifica dell'adattamento del modello ai dati (o valutazione del fit)

Verifica assunzione UNIDIMENSIONALITA'

Se si considera il modello unidimensionale, per verificare l'assunzione di unidimensionalità si possono utilizzare diverse tecniche.

1) ANALISI FATTORIALE non lineare sulla matrice di correlazione interitem (che contiene correlazioni tetracoriche se gli item sono dicotomici, o policoriche se gli item sono politomici) e studiare i RESIDUI.

2) Si possono esaminare i RESIDUI calcolati come differenza tra punteggi osservati e punteggi previsti.

3) Si può anche applicare UN'ANALISI IN COMPONENTI PRINCIPALI sulla matrice di correlazione interitem per calcolare il plot degli autovalori

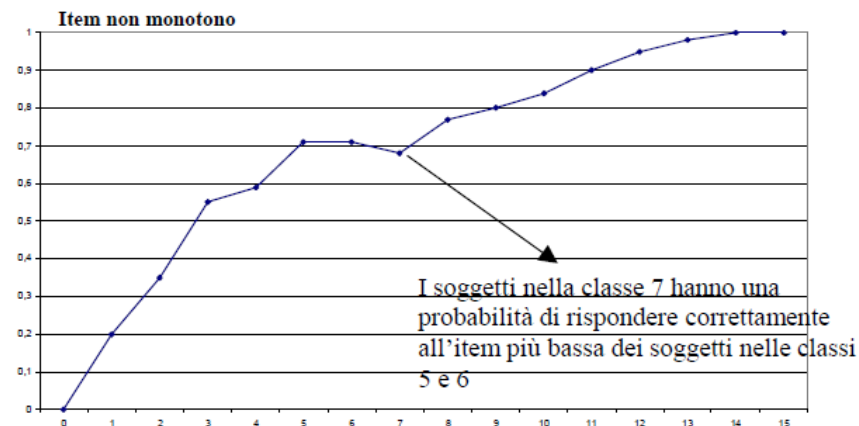
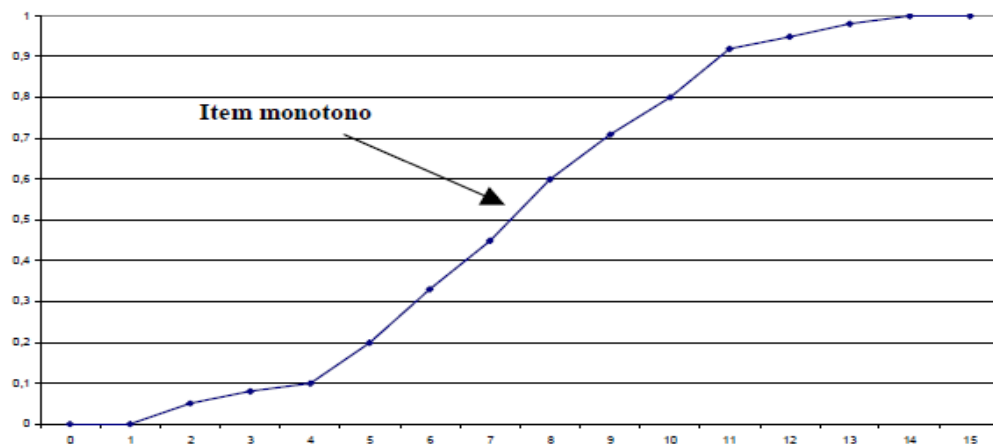
Verifica assunzione INDIPENDENZA LOCALE

Per verificare l'indipendenza locale

- 1) si può ispezionare la matrice di varianza-covarianza o la matrice di correlazione tra i soggetti classificati entro uno stesso intervallo di theta. Se l'assunzione di indipendenza locale è verificata, le correlazioni fuori la diagonale principale saranno prossime allo zero.
- 2) nel caso del modello logistico ad uno e a due parametri, si può verificare se il parametro guessing tende a zero. Per verificare questa assunzione si può osservare la performance di un soggetto con basso livello di abilità. Se il livello della performance è basso (quindi il soggetto risponde in modo errato all'item), è lecito supporre che il parametro guessing tende a zero.

Verifica MONOTONICITÀ di un item

È necessario ed importante verificare anche la monotonicità di un item, rappresentando graficamente la proporzione di risposte ad ogni item rispetto ai diversi livelli del punteggio grezzo (considerato come una stima del punteggio nel tratto latente).

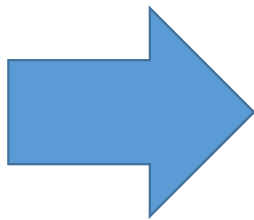


Verifica proprietà INVARIANZA

Per verificare l'invarianza della stima del livello di abilità rispetto al test



possono essere considerati due test di cui uno facile e uno difficile ed in funzione delle risposte fornite viene stimato il loro livello di abilità



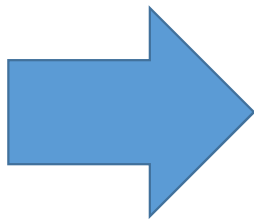
se l'invarianza è verificata, le stima di abilità saranno uguali nei due test, mentre ciò che cambierà sarà l'errore standard associato a tale stima

Verifica proprietà INVARIANZA

Per verificare l'invarianza della stima dei parametri rispetto al campione su cui viene effettuata



possono essere considerati gruppi diversi ma con uguale numerosità



se l'invarianza è verificata, allora le stime dei parametri dovranno essere uguali nei due gruppi. La verifica dell'invarianza può avvenire solo i due gruppi presentano la stessa numerosità campionaria

I PARAMETRI e la loro STIMA

I parametri da STIMARE sono



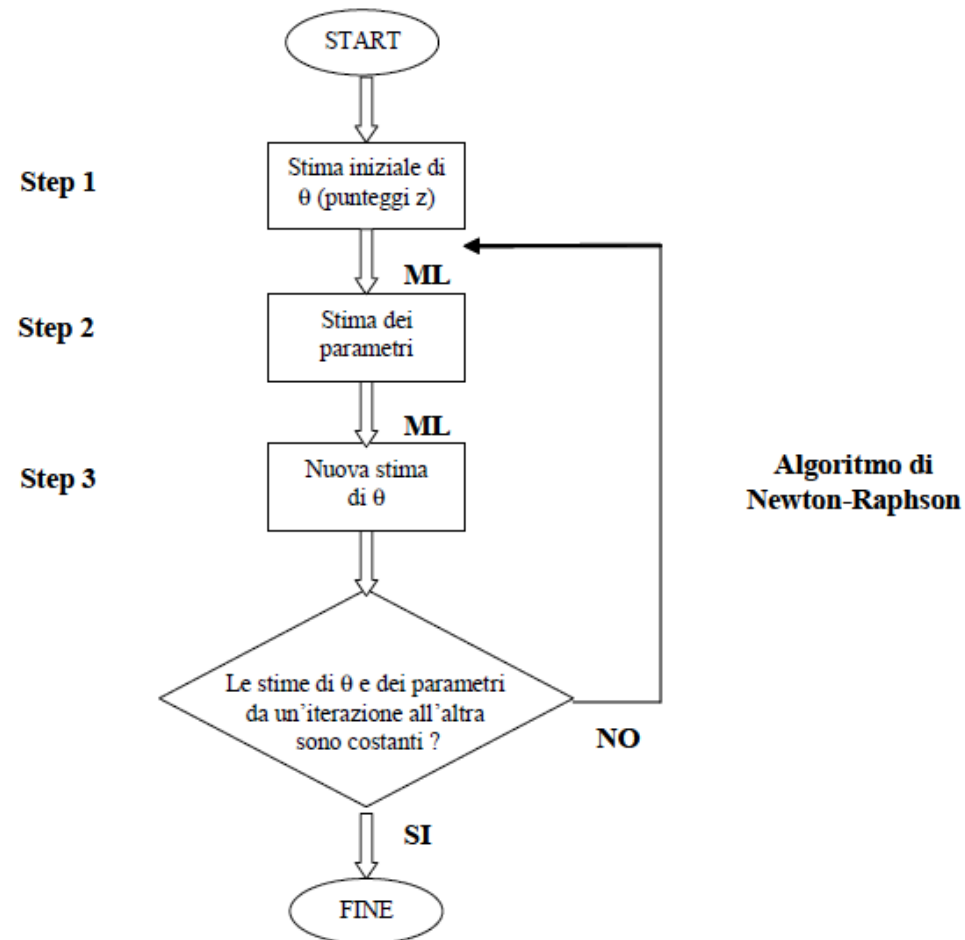
OSSERVAZIONE:

il ricercatore conosce solo il punteggio ottenuto dal soggetto in un test

Metodo della Massima Verosimiglianza

Il metodo che quindi viene utilizzato per stimare i parametri nella TRI è il metodo di *massima verosimiglianza (ML)*, che può essere implementato da algoritmi iterativi aggiuntivi che hanno lo scopo di raggiungere stime che massimizzano la funzione di verosimiglianza.

Diagramma di flusso che rappresenta il processo di stima congiunta



Stima congiunta: casi limite

Nel processo di stima congiunta non vengono considerati due casi.

1. Il primo caso è quando un soggetto **risponde correttamente o in modo sbagliato a tutti gli item**.
2. Il secondo caso è quando tutti i soggetti hanno risposto in maniera corretta o in maniera errata ad un item.

Questi casi sono eliminati perché non danno informazioni sufficienti.

Per **superare i limiti della stima congiunta** di massima verosimiglianza, successivamente sono stati proposti dei metodi alternativi come

metodo di massima verosimiglianza condizionale

la stima bayesiana

la stima marginale massima di verosimiglianza.

Proprietà delle stime

I processi di stima più utilizzati nella TRI applicano il metodo di massima verosimiglianza, è utile sottolineare che indipendentemente dalla tecnica applicata, e nel caso di campioni molto numerosi, le stime che sono prodotte godono delle seguenti proprietà:

Consistenza = c'è una convergenza al valore vero all'aumentare della numerosità campionaria.

Efficienza = l'errore standard ha un rapporto inversamente proporzionale con la numerosità campionaria.

Normalità = per infinite misurazioni l'errore standard tende ad assumere una distribuzione normale con media 0 e varianza 1

Valutazione del fit

L'esame del fit (adattamento) può essere fatto a tre livelli e in particolare:

- valutazione globale del fit del modello (model fit)
- valutazione del fit di ogni singolo item (item fit)
- valutazione del fit del soggetto (person fit)

Model fit: statistica del chi-quadrato

L'esame del fit si pone come obiettivo quello di valutare quale modello si adatta meglio ai dati (ad uno, a due a tre, o quattro parametri), e quindi che rappresenta un'immagine più verosimile della realtà.

Per la valutazione del fit si può utilizzare la statistica del chi-quadrato

$$\chi^2 = \frac{\sum_j [p(\theta_j) - P(\theta_j)]^2}{P(\theta_j)Q(\theta_j)}$$

dove

$p(\theta_j)$ = proporzione osservata di risposte corrette

$P(\theta_j)$ = proporzione attesa di risposte corrette

$Q(\theta_j) = 1 - P(\theta_j)$

Model fit: residui

Un altro modo per valutare il model fit è quello di considerare i residui.

Il residuo è definito come la differenza tra risposta osservata e il punteggio atteso previsto dal modello.

Maggiore è il valore dei residui, e maggiore sarà la discrepanza tra i dati osservati e il modello ipotizzato.

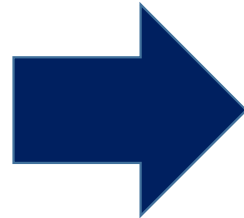
Il residuo può essere espresso da diversi tipi di differenza ma quella più utilizzata è la differenza standardizzata, per i quali si assume che abbiano media 0 e varianza 1.

Se l'adattamento è buono i residui standardizzati avranno una dispersione minima intorno allo zero per i diversi livelli di theta.

La funzione informativa dell'Item (Item Information Function)

Fondamentale per qualsiasi analisi inferenziale è che le stime conservino una quota rilevante dell'informazione originale.

L'informazione in generale rappresenta l'inverso della varianza; quindi più il valore di informazione è elevato, migliore sarà la precisione nella stima dei parametri del modello.



$$I(\theta_i) = \frac{1}{\sigma^2(\theta_i)}$$

Nei modelli della IRT è possibile calcolare l'informazione associata a ogni valore dello stimatore dell'abilità giungendo a una curva continua lungo tutto il suo dominio; graficamente rappresenta la Funzione di Informazione (Item Information Curve, IIC) .

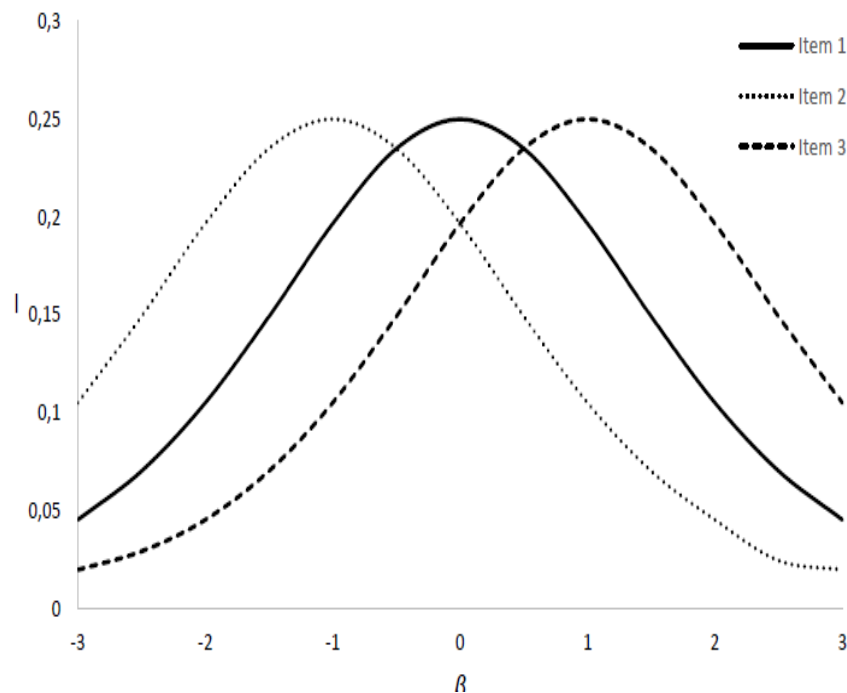
La funzione informativa dell'Item (Item Information Function)

La funzione informativa dell'item, fornisce un'idea sulla posizione rispetto alla scala di abilità dove l'item è maggiormente informativo:



un item, quindi, è molto informativo solo per uno specifico livello di tratto nel quale la stima sarà più precisa e meno inficiata da errore; di conseguenza, lo stesso item avrà un errore standard maggiore nei livelli di tratto in cui il suo potere informativo è più debole.

Funzioni informative dell'item



Nel caso più semplice di modello, quello a un parametro, la funzione informativa dell'item è definita in ogni punto dal prodotto tra la rispettiva probabilità di superare l'item e quella, reciproca, di non superarlo.

$$I_j(\theta_i) = P(y_{ij} = 1)P(y_{ij} = 0)$$

La funzione informativa del test

Sommando in ogni punto le funzioni informative degli item, si ottiene la funzione informativa del test.



In questo caso si osserva un punto di massimo per un'abilità di 0 e quindi si può concludere che il test, nel suo insieme, è più preciso nello stimare il livello di abilità di individui collocati proprio in quel punto.

Funzioni informative dell'item e del test

Le funzioni d'informazione degli item e del test si qualificano come strumenti molto importanti per analizzare la natura dei dati e sfruttare le potenzialità dei modelli.

In linea teorica, l'obiettivo preferibile sarebbe costruire test la cui capacità informativa si mantenga stabile lungo tutto il continuum della proprietà e per qualsiasi individuo, a prescindere dal suo livello di abilità. In altre parole, tale condizione garantirebbe un'eguale precisione di "misurazione" per qualsiasi soggetto, indipendentemente dalle sue caratteristiche. Ciò non è facile da ottenere nella pratica, poiché sarebbe necessario trovare una perfetta combinazione di item, la somma delle cui funzioni informative fosse costante in ogni punto.

La teoria classica dei test vs IRT

Le principali differenze tra le due teorie riguardano soprattutto

1. i concetti relativi all'errore standard di misurazione
2. l'attendibilità della misurazione
3. la difficoltà dell'item
4. la stima del livello di abilità e interpretazione dei punteggi
5. la relazione tra variabile dipendente e indipendente
6. la bontà dell'item

La teoria classica dei test vs IRT: L'errore standard di misurazione

TEORIA CLASSICA DEI TEST



l'errore standard di misurazione è costante rispetto al campione su cui viene stimato, quindi tutti i soggetti appartenenti ad un unico campione hanno lo stesso livello di errore

IRT



l'errore standard di misurazione varia in funzione del livello di abilità posseduto dal soggetto: tende ad essere minimo intorno al valore di θ in cui l'informazione del test (e quindi la precisione della misurazione) è massima, e al contrario tende ad aumentare per quei valori di abilità in cui l'informazione del test è minima.

La teoria classica dei test vs IRT: attendibilità della misurazione

Collegato al concetto di errore standard vi è il concetto di attendibilità della misurazione, ovvero la precisione della misurazione stessa. Errore standard e attendibilità hanno un rapporto inversamente proporzionale, ovvero all'aumentare della precisione della misura diminuisce l'incidenza dell'errore commesso e viceversa.

TEORIA CLASSICA DEI TEST



Nella CTT la precisione di una misurazione viene chiamata attendibilità, che tende ad aumentare all'aumentare della lunghezza del test

La teoria classica dei test vs IRT: attendibilità della misurazione

IRT

Nella IRT non si parla invece di attendibilità ma

Funzione Informativa dell'Item
(IIF= Item Information Function),

consente al ricercatore di scegliere gli item che hanno un elevato potere informativo per una regione di theta specifica.

Funzione Informativa del Test
(TIF= Test InformationFunction)

Sommatoria delle singole IIF

La teoria classica dei test vs IRT: difficoltà dell'item

TEORIA CLASSICA DEI TEST



Il parametro di difficoltà (e in generale tutti i parametri degli item) nella CTT è dipendente dal campione su cui viene stimato.

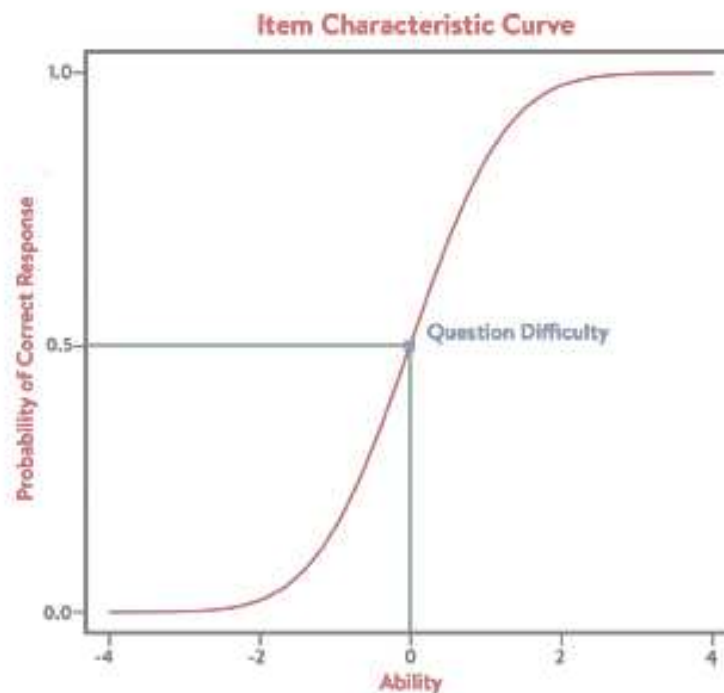
IRT



Nella IRT invece la difficoltà dell'item non dipende in alcun modo dalla distribuzione di abilità del campione a cui viene somministrato l'item, e quindi i parametri vengono definiti come indipendenti dal campione su cui vengono stimati.

La teoria classica dei test vs IRT: difficoltà dell'item

In particolare nella IRT la difficoltà viene interpretata come quel punto nella scala di abilità in cui la probabilità di risposta corretta è uguale a 0,5.



La teoria classica dei test vs IRT: stima del livello di abilità

TEORIA CLASSICA DEI TEST



Per quanto riguarda la stima del livello di abilità posseduto dal soggetto, nella CTT tale stima dipende dal test che viene somministrato al soggetto. Ne consegue che se si somministrano due test di diversa difficoltà allo stesso soggetto, la stima del livello di abilità posseduto dal soggetto sarà diversa.

IRT



al contrario nella IRT la stima dell'abilità del soggetto è indipendente dal test che viene somministrato

La teoria classica dei test vs IRT: Interpretazione dei punteggi

TEORIA CLASSICA DEI TEST



Nella CTT i punteggi ottenuti da un soggetto non vengono interpretati in modo assoluto ma vengono confrontati con il gruppo normativo (composto da soggetti che hanno caratteristiche simili a quelle del soggetto in questione) attraverso i punteggi standard (z, T, centili).

IRT



Nella IRT invece i punteggi del soggetto vengono interpretati considerando la posizione del soggetto rispetto al tratto latente. Poiché i modelli IRT godono della proprietà di invarianza della misurazione, la stima del livello di abilità del soggetto non è interpretata facendo riferimento ad alcun gruppo normativo.

La teoria classica dei test vs IRT:

relazione tra variabile dipendente e indipendente

TEORIA CLASSICA DEI TEST



Nella CTT il punteggio ottenuto ad un test X o ad un item i viene considerato come una variabile dipendente spiegata dalla **sommatoria di due variabili indipendenti**: il punteggio vero (V) e la componente d'errore (E), secondo la nota formula:

$$X=V+E$$

La relazione ipotizzata è quindi una regressione lineare e additiva, rappresentata su un piano cartesiano da una retta.

IRT



Nella IRT invece la **variabile dipendente** è **rappresentata dalla probabilità di rispondere correttamente ad un item**, che può essere spiegata dal livello di abilità del soggetto e dai parametri dell'item (che rappresentano quindi le variabili indipendenti). La relazione ipotizzata tra i predittori è una regressione di tipo logistico, che viene rappresentata da una curva logistica.

La teoria classica dei test vs IRT: La bontà dell'item

TEORIA CLASSICA DEI TEST



nella CTT la bontà di un item viene espressa dal **coefficiente di correlazione item-totale** (interpretato in termini di quanto bene l'item misura il costrutto in esame e quindi quanto incide sull'attendibilità del test in generale)

IRT



nella IRT l'item viene valutato in funzione del **parametro di discriminazione** (quanto un item riesce a differenziare soggetti che possiedono un livello di abilità superiore o inferiore ad un cut-off di riferimento).

In sintesi..

	CTT	IRT
Assunzione fondamentale	Un individuo con il tratto latente più alto risponderà correttamente a un maggior numero di domande	Risposto correttamente a una certa domanda, con più probabilità si risponderà correttamente a domande più facili
Normalizzazione	Risente molto del campione di normalizzazione	È indipendente dal campione di normalizzazione
	Prima si fanno le norme poi si possono valutare altri soggetti	Normalizzazione e valutazione dei soggetti avvengono nello stesso momento
	Si valuta il punteggio test	Si valuta ogni item
Calcolo del punteggio	Solitamente si sommano i punteggi di ogni singolo item	Si ipotizza un complesso modello probabilistico
Interpretazione del punteggio	Confronto tra il punteggio del soggetto con delle norme	Confronto tra gli item a cui si è risposto
	Non tiene conto della difficoltà degli item	Tiene conto della difficoltà degli item
Errore di misura	Ipotizzato costante per tutti i livelli di tratto	Relazione matematica che varia in funzione del livello di tratto

Applicazioni dei modelli TRI:

Test computerizzati di tipo adattivo(CAT)

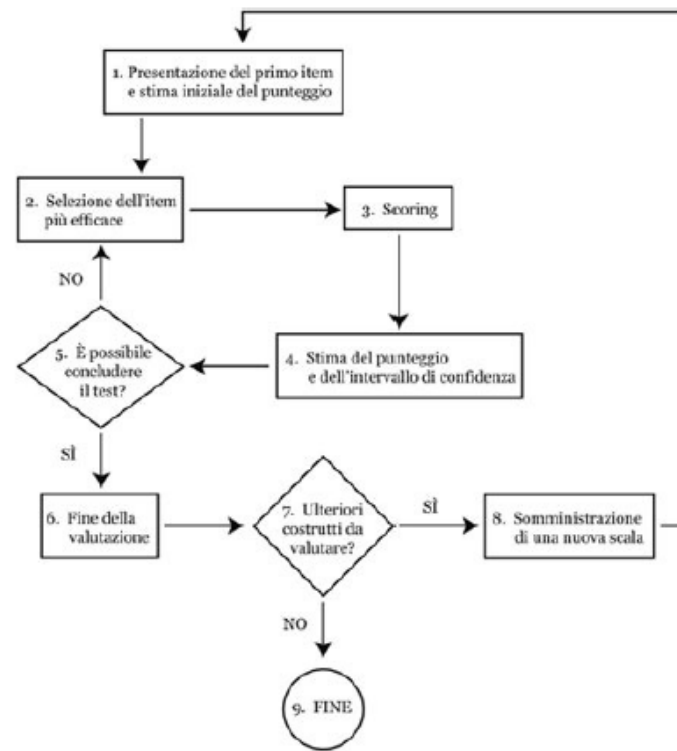
L'obiettivo dei test informatizzati è quello di riuscire a presentare ad ogni soggetto un gruppo di item che sia in grado di rilevare il massimo livello di conoscenze su quel soggetto.

La possibilità di sviluppare strumenti siffatti si poggia sull'apposita creazione di una banca dati contenente un ampio numero di item altamente discriminanti con parametri di difficoltà in grado di coprire l'intero continuum dei possibili atteggiamenti rispetto al concetto oggetto di studio, dalle posizioni più estreme in un senso alle più estreme nel senso opposto.

Logica dei test computerizzati di tipo adattivo (CAT)

1. somministrazione di un item di media difficoltà, capace di ricoprire una vasta gamma di livelli e di fornire le basi per una prima stima del punteggio.
2. selezione dell'item più informativo per ogni soggetto da un pool di item calibrati attraverso l'applicazione della modellizzazione IRT,
3. dopo lo scoring relativo al secondo compito, il processo prosegue con una nuova stima del punteggio e con il calcolo dell'intervallo di confidenza.
4. l'applicazione di tale principio viene seguita finché l'errore standard di misurazione, non sarà abbastanza esiguo da permettere la conclusione dell'indagine

Logica dei test computerizzati di tipo adattivo (CAT)



CASO DI STUDIO

Studenti iscritti ai corsi di laurea ad indirizzo scientifico

Dati raccolti nelle sessioni d'esame dal 2018 al 2019 di circa 1300 studenti iscritti a diversi corsi di laurea ad indirizzo scientifico e che hanno frequentato insegnamenti attinenti ai fondamenti di informatica.



somministrazione di un questionario
con item di tipo *multiple-response*:
ad ogni domanda si presentano n
alternative di risposta

5 MACRO-ARGOMENTI (ITEM):

1. Fondamenti (elementi di base)
2. Glossario
3. Prolog (capacità di programmazione)
4. Prolog1 (sintassi)
5. Prolog2 (semantica)

CASO DI STUDIO

Risultati per ciascuna domanda di 256 studenti esaminati in una sessione.

Percentuali dei punteggi ottenuti per ciascun item

Punteggi	<i>Fondamenti</i> (%)	<i>Glossario</i> (%)	<i>Prolog</i> (%)	<i>Prolog1</i> (%)	<i>Prolog2</i> (%)
1	4	7	8	4	13
2	16	30	47	50	57
3	33	32	36	42	20
4	48	31	7	4	9

CASO DI STUDIO

Stima parametri item modello TRI

Item		Punteggi	Difficoltà	Discriminazione
<i>Fondamenti</i>	1		-2.51	0.49
	2		0.23	
	3		2.58	
<i>Glossario</i>	1		-3.31	0.51
	2		-1.49	
	3		0.09	
<i>Prolog</i>	1		-2.81	0.63
	2		-0.60	
	3		0.87	
<i>Prolog1</i>	1		-3.68	1.07
	2		0.21	
	3		3.83	
<i>Prolog2</i>	1		-2.78	1.74
	2		1.25	
	3		3.27	

ITEM PIU' DISCRIMINANTI

ITEM PIU' DIFFICILI

CASO DI STUDIO

Stima parametri ABILITA' modello IRT

Classificazione degli studenti in base all'abilità

<i>Abilità</i>	<i>Voto</i>	<i>Numero studenti</i>	<i>Percentuale studenti</i>
<-0.85	<18	24	9
-0.85 —0.61	18	18	7
-0.61 —0.40	19	37	15
-0.40 — 0.19	20	26	10
-0.19 — 0.02	21	37	15
0.02 — 0.22	22	26	10
0.22 — 0.43	23	21	8
0.43 — 0.64	24	15	6
0.64 — 0.85	25	21	8
0.85 — 1.05	26	5	2
1.05 — 1.26	27	10	4
1.26 — 1.47	28	6	2
1.47 — 1.65	29	5	2
≥ 1.65	30+	5	2
		256	100