

# Psicometria

1-Panoramica delle tecniche: Spiegazione intuitiva  
vers. 1.1

Germano Rossi<sup>1</sup>

`germano.rossi@unimib.it`

Giovanni Battista Flebus<sup>1</sup>

`giovannibattista.flebus@unimib.it`

<sup>1</sup>Dipartimento di Psicologia, Università di Milano-Bicocca

2008-2008

# Psicometria

- La parola “psicometria” significa sostanzialmente “misura della psiche” o “misura psichica” e viene usata in due ambiti disciplinari
- **In parapsicologia:** capacità di percepire la storia di un oggetto, toccandolo
- **In psicologia:** la misurazione della psiche o meglio **la misurazione in ambito psicologico**. E' in realtà un termine *contenitore* che va dalla *statistica descrittiva* alle *analisi multivariate* più complicate alla *teoria e tecnica dei test*
- Ovviamente, noi la useremo nel secondo significato, cioè di *analisi multivariate*

# Definizioni

**Variabile** le varie misurazioni di una certa caratteristica relativamente a diversi casi statistici

**Costante** una misurazione di una certa caratteristica che non cambia nei diversi casi statistici

**Variabile categoriale** una variabile misurata a livello nominale o ordinale

**Variabile quantitativa** una variabile misurata a livello intervallo o rapporto

**Variabile indipendente** una caratteristica che non dipende da altre (ad es. il sesso non dipende dall'intelligenza, dalla cultura. . .)

**Variabile dipendente** una caratteristica che può essere influenzata da altre caratteristiche (ad es. la ricchezza del vocabolario può dipendere dall'educazione e dalla famiglia di origine)

# Definizioni

**Variabile osservata o misurata** una caratteristica che si può misurare in modo diretto (es.: età, reddito, numero di errori)

**Variabile latente** una caratteristica che non può essere misurata direttamente ma che si ipotizza avere dei legami con altre variabili osservabili (es.: intelligenza, cultura, inconscio)

- Queste varie classificazioni vengono applicate in base al contesto
- Una variabile considerata indipendente in un certo contesto, può diventare dipendente in un'altro

# Modellazione grafica

E' possibile rappresentare graficamente le relazioni fra le variabili (osservate/latenti, dipendenti/indipendenti, categoriali/quantitative)

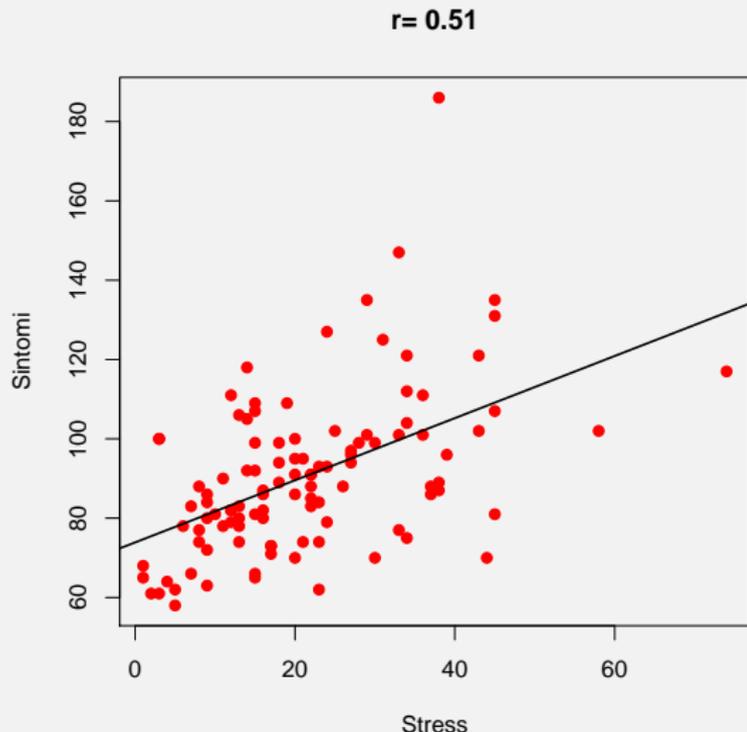
Variabile osservata o manifesta <i>quantitativa</i>	->	quadrato	
Variabile osservata <i>categoriale</i>	->	quadrati	
Variabile latente	->	cerchio	
Relazione generica	->	segmento	
Influenza, spiegazione	->	freccia	
Relazione reciproca (correlazione)	->	doppia freccia	
Interazione	->		

# Modellazione grafica

Variabile osservata <i>categoriale</i> (2 valori)	->	quadrati	
Variabile osservata <i>categoriale</i> (più valori)	->	quadrati	
Variabile osservata <i>categoriale</i> (più valori, dicotomizzati)	->	quadrati	

# Correlazione

Al corso di statistica abbiamo visto la correlazione fra due variabili.  
C'è una correlazione fra i sintomi fisici e lo stress?  
E il numero dei sintomi spiega l'indice dello stress?



# Regressione lineare

Tramite una funzione matematica (tipo quella di una retta) si cerca di vedere se:

- una variabile osservata quantitativa viene spiegata da un'altra variabile osservata quantitativa (*Regressione lineare semplice*) o da altre variabili osservate quantitative (*Regressione lineare multipla*)
- più variabili osservate quantitative vengono spiegate da altre variabili osservate quantitative (*Regressione lineare multivariata*)

# Concetto di regressione

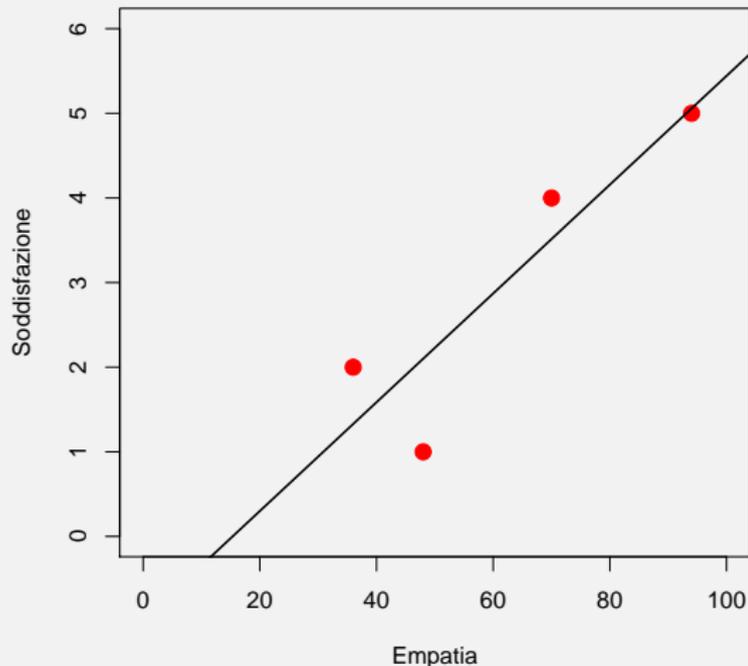
Empatia	Soddisfazione
70	4
94	5
36	2
48	1

Un ricercatore sta studiando la relazione fra il **grado di empatia degli psicoterapeuti** e la **soddisfazione dei loro pazienti**. In uno studio pilota, utilizza 4 coppie paziente-terapeuta. Con un test misura l'empatia del terapeuta (maggiore il valore, maggiore l'empatia) e chiede al paziente di giudicare il proprio grado di soddisfazione su una scala da 1 (poco) a 5 (molto)

# Rappresentazione grafica

All'aumentare del grado di empatia del terapeuta aumenta la soddisfazione del paziente.

I dati possono essere facilmente associati ad una retta, anche se non cadono esattamente su quella retta. Più vicini sono, meglio la retta è in grado di "predire" la soddisfazione sulla base dell'empatia.



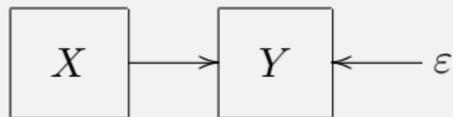
# Causa-effetto

- Sostanzialmente, quando diciamo che una variabile “predice” o “spiega” un’altra variabile, intendiamo dire che stiamo cercando di stabilire una relazione di causa-effetto. La variabile “spiegata” è l’effetto, quella che spiega (o che influenza) è la causa.
- La regressione lineare utilizza il principio matematico della retta per “cercare” questa relazione causale.

$$Y = bX + a$$

# Regressione lineare semplice

Possiamo rappresentare graficamente la regressione lineare semplice in questo modo:



La variabile osservata  $X$  spiega (o influenza) la variabile osservata  $Y$ . E' però possibile che non tutta la variabile  $X$  serva a spiegare  $Y$  e che la variabile  $Y$  non venga interamente spiegata da  $X$  (la parte non spiegata è chiamata errore,  $\varepsilon$ ).

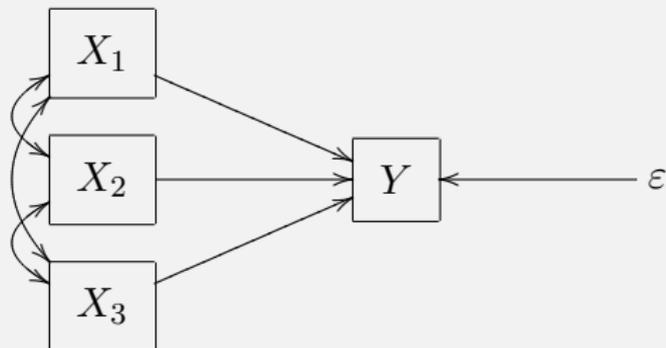
Questo modello dà origine ad un'equazione, che in termini linguistici, diventa:  $Y$  è formata da una parte di  $X$  e da un errore.

In termini algebrici:  $Y = bX + \varepsilon$ .

Tuttavia questo tipo di modello è troppo semplicistico.

# Regressione lineare multipla

Se complichiamo il modello ipotizzando che più variabili concorrano a spiegare  $Y$ , abbiamo la regressione multipla:

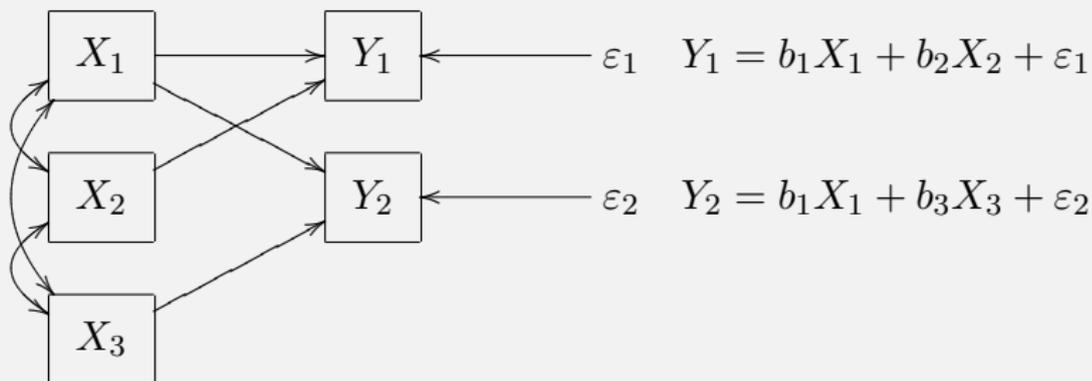


Le variabili osservate  $X$  influenzano la variabile osservata  $Y$  (anche qui sono possibili gli errori). E la formulazione algebrica diventa:

$Y = b_1X_1 + b_2X_2 + b_3X_3 + \varepsilon$  e cioè, ogni  $X$  contribuisce a spiegare una parte di  $Y$ . In pratica, ogni freccia è un “peso”, un “parametro”.

# Regr. lineare multivariata

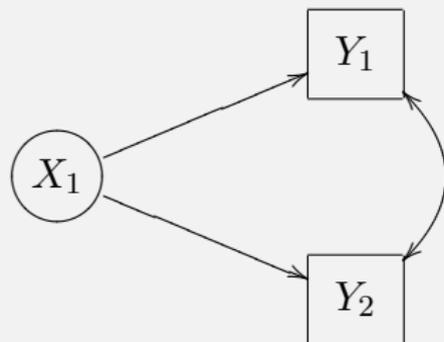
Ovviamente, anche questo modello è semplicistico. Nella realtà è più probabile che molte variabili siano legate in modo più o meno complesse a molte altre, a seconda del “modello” analizzato. Quindi che diverse  $X$  contribuiscano ad influenzare, in proporzioni diverse o contributi differenti, diverse  $Y$ .



Le variabili osservate  $X$  influenzano, in vario modo, le variabili osservate  $Y$ . Per ogni  $Y$ , si può costruire un'equazione.

# Analisi fattoriale

Quando abbiamo studiato la correlazione, abbiamo visto come possa essere interpretata come un indice di qualcosa che le due variabili hanno in comune. E abbiamo anche visto come sia possibile che due variabili correlino fra loro, anche senza una connessione logica e che tale correlazione possa invece dipendere da una terza variabile. Se questa terza variabile non l'abbiamo mai effettivamente misurata, è una variabile latente! (o fattore) e dovrà essere stimata sulla base dei suoi effetti.



# Analisi fattoriale

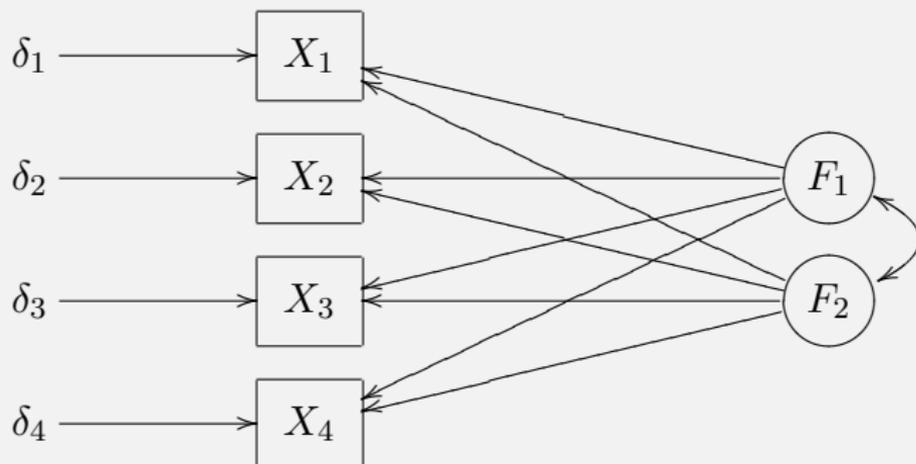
L'analisi fattoriale è una tecnica che cerca di

- associare una o più variabili latenti (che non si conoscono) ad un gruppo di variabili osservate. Si presuppone che le osservate abbiano qualche cosa in comune fra di loro e che questo qualcosa dipenda dalla variabile latente.
- Questo “qualcosa in comune” è stato storicamente chiamato **Fattore**

# AF Esplorativa: esempio 1

- In un prova (test) sono presenti item relativi a: 1) *Contare e ordinare sequenze numeriche*, 2) *Soluzione di problemi aritmetici*, 3) *Trovare errori di sintassi*, 4) *Trovare errori grammaticali*
- Tutti (o alcuni) di questi item hanno qualcosa in comune? Se sì, cosa? e con chi?
- Posso ipotizzare la presenza di due variabili latenti, una che ha qualcosa a che fare con i numeri e l'altra con la scrittura?
- Se effettuo una analisi fattoriale esplorativa posso calcolare la “relazione” che ciascuna prova ha con 1 o più fattori (var. latenti)

# Analisi fattoriale esplorativa

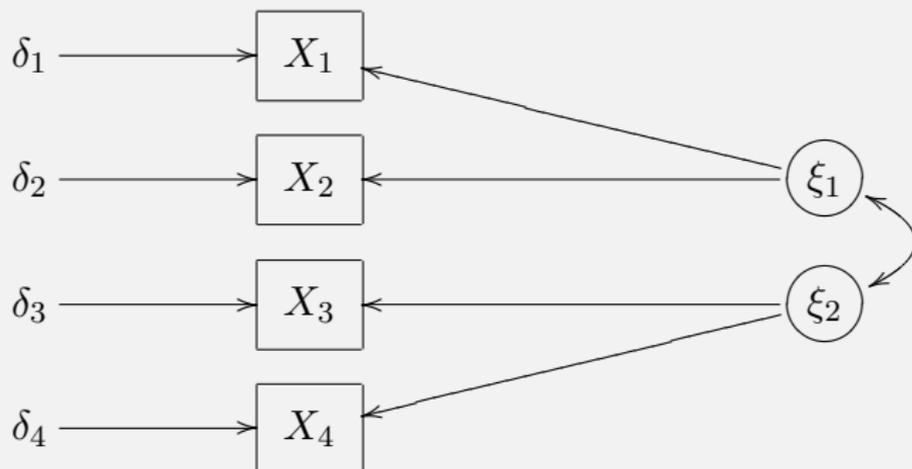


Tutte le  $X$  (ma in grado diverso) vengono spiegate da tutti i fattori (che possono anche essere correlati fra loro).

# Analisi fattoriale confermativa

- Serve per verificare che una o più variabili latenti che abbiamo ipotizzato siano effettivamente legate alle relative variabili osservate che gli sono associate
- Nell'esempio di prima, potremmo essere noi stessi ad ipotizzare che 1 e 2 sottostiano ad una variabile latente di *abilità numerica* e che 3 e 4 indichino una variabile latente di *abilità del linguaggio scritto*.
- L'analisi ci dirà quanto "buono" è il modello che abbiamo ipotizzato.

# Analisi fattoriale confermativa



Solo alcune  $X$  vengono spiegate dai Fattori (variabili latenti)

Anche in questo caso, ogni osservata (in quanto “riceve” una freccia) sarà oggetto di una equazione:

$$X_1 = b_1 \xi_1 + \delta_1$$

e così via

# Analisi della varianza

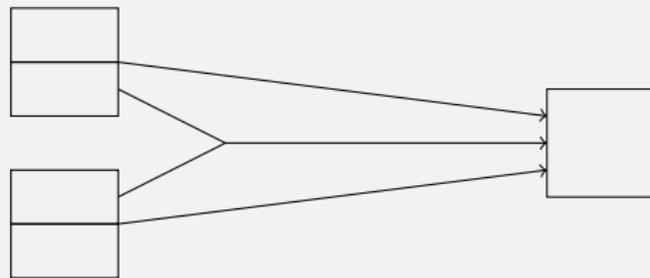
- Sempre in statistica abbiamo visto la possibilità di studiare se due gruppi sono stati o meno estratti dalla stessa popolazione. Per fare questo tipo di verifica abbiamo usato un test basato sulla differenza delle medie (genericamente chiamato T-Test).



- Abbiamo anche visto che, se ci sono più di due gruppi, bisogna usare una tecnica più avanzata: l'analisi della varianza.
- Il principio generale è che quello di comparare la varianza totale e la varianza generata dai singoli gruppi.

# Analisi della varianza

Secondo diverse teorie, il fondamentalismo religioso è più elevato tra le donne e nelle persone anziane. Perciò misuriamo il fondamentalismo religioso in un campione suddiviso per genere e 2 fasce d'età (giovani, anziani). Per verificare la nostra ipotesi, usiamo l'analisi della varianza.



L'ANOVA tiene in considerazione sia l'influenza diretta di ogni singola variabile (genere, fasce d'età) sia la loro interazione.

# Riepilogo

Dipendente	Indipendente	Tecnica
1 oss. quant.	1 oss. quant.	regr. lineare semplice
1 oss. quant.	2+ oss. quant.	regr. lineare multipla
2+ oss. quant.	2+ oss. quant.	regr. lineare multivariata
2+ oss. quant.	1+ lat. quant.	analisi fattoriale
1 oss. quant.	2+ oss. categ.	analisi varianza