



# multivariata, analisi

Enciclopedie on line

**multivariata, analisi** In statistica, l'insieme dei metodi statistici e delle tecniche usati nello studio della variazione simultanea di due o più variabili casuali (nel caso di una variabile → varianza). Date le distribuzioni congiunte di due o più variabili, il metodo più utile per analizzare i dati è quello di rappresentarli sotto forma di tabella o matrice dei dati in cui ciascuna riga rappresenta le varie caratteristiche osservate sul generico elemento, mentre ciascuna colonna rappresenta la variabilità della stessa caratteristica sui vari elementi. In particolare, se la matrice dei dati è esprimibile linearmente in funzione di una matrice di variabili indipendenti, si parla di modello lineare multivariato.

Una distinzione essenziale nell'analisi m. è se si tratti di analizzare un sistema di covariazione tra gruppi (gli oggetti da esaminare appartengono a più di una popolazione) oppure entro gruppi (gli oggetti appartengono a un'unica popolazione). Le tecniche usate nel primo caso mirano a porre in evidenza un'eventuale disomogeneità dei campioni, e lo scopo dell'analisi m. può essere quello di determinare se i vettori medi delle popolazioni siano o meno uguali (ed è il caso dell'analisi m. della varianza, o MANOVA), oppure quello di trovare combinazioni lineari delle variabili che massimizzino le differenze tra gruppi preesistenti (analisi discriminante).

Esistono inoltre metodi che consentono di raggruppare gli oggetti a seconda del grado di somiglianza che essi presentano; questi metodi vanno generalmente sotto il nome di *analisi dei clusters*, e lo scopo principale dell'analisi è quello di risolvere una raccolta eterogenea e unica di oggetti in una serie di suddivisioni omogenee al loro interno. Quando gli oggetti da esaminare sono tratti da un'unica popolazione, l'attenzione è rivolta innanzitutto ai modi di semplificare la struttura dei dati. Così, mentre l'analisi

dei *clusters* tende a raggruppare quelle variabili che sono altamente correlate, l'analisi dei fattori ha lo scopo di risolvere le correlazioni tra variabili in quelle che si ritengono le loro cause determinanti, cioè di esprimere la covariazione in termini di un certo numero di fattori (inferiori al numero delle variabili) che spieghino una grande parte della varianza e della covarianza delle variabili originarie.

Diversamente dall'analisi dei fattori, quella delle componenti principali è rivolta alla rappresentazione della variabilità trovata nel campione, più che all'interpretazione delle correlazioni osservate tra variabili misurate; ma, analogamente a essa, opera attraverso una riduzione del numero di variabili coinvolte a quelle maggiormente informative. Se, infine, le variabili possono essere suddivise in maniera logica in due o più insiemi, per stabilire se esistono relazioni lineari tra questi insiemi, con l'analisi della correlazione canonica è possibile valutare il grado di associazione esistente tra di essi considerati a due a due, essendo ciascun insieme ridotto al proprio componente principale.