

Matrice dei dati

Lo strumento fondamentale di organizzazione delle informazioni per l'analisi statistica è rappresentato dalla matrice dei dati o casi per variabili ($C \times V$), ovvero il quadro degli cv elementi disposti su C righe (C = numero dei **casi**) e V colonne (V = numero delle **Variabili**).

Tali elementi sono i **dati** (che indichiamo qui con "x"), unità informative codificate relative a ciascuna variabile, per ciascun caso. In pratica, la matrice è una tabella in cui vengono riportate tutte le informazioni raccolte per tutti i casi studiati. Tali informazioni sono organizzate disponendo i casi in riga e le variabili in colonna.

ID	V1	V2	...	V _v
C1	X _{1,1}	X _{1,2}	...	X _{1,v}
C2	X _{2,1}	X _{2,2}	...	X _{2,v}
...	...	...	...	...
c	X _{c,1}	X _{c,2}	...	X _{c,v}

Per poter effettuare sulla matrice le necessarie operazioni algebriche richieste dall'analisi statistica, la matrice non deve presentare celle vuote. Di conseguenza:

- dobbiamo usare le stesse variabili – e tutte – per ogni caso (altrimenti, avremmo celle vuote in corrispondenza delle informazioni non rilevate);
- i casi della matrice devono essere tutti dello stesso tipo: non è possibile infatti rilevare le stesse informazioni per unità di analisi diverse (individui e province, ad esempio), e di conseguenza avremmo celle vuote.
- si dovrà prevedere un codice per i cosiddetti **Valori mancanti (Missing Values)** (o *missing values*)

Se in una indagine è necessario lavorare con tipi diversi di unità di analisi, si dovrà costruire (almeno) una matrice per ciascun tipo di unità di analisi.

Mentre un tempo l'inserimento dei dati nella matrice avveniva a mano, utilizzando grandi fogli quadrettati, oggi si utilizza un qualunque foglio di calcolo (come Excel o Access), o anche il foglio di lavoro incluso nei più comuni packages per l'analisi statistica dei dati.

From:

<https://www.agnesevardanega.eu/wiki/> - **Ricerca Sociale con R**

Permanent link:

https://www.agnesevardanega.eu/wiki/glossario/matrice_dei_dati?rev=1319355806

Last update: **23/10/2011 07:43**

