

Glossario

❑ **DISTRIBUZIONE DOPPIA DI FREQUENZE**

e' una tabella nella quale vengono elencate tutte le modalità con le quali si presentano due caratteri (qualitativi e/o quantitativi); ad ogni coppia di modalità (*i-esima* del carattere X, e *j-esima* del carattere Y) viene associata una frequenza assoluta, che rappresenta il numero di volte in cui si sono presentate contemporaneamente la *i-esima* modalità del carattere X e la *j-esima* modalità del carattere Y.

❑ **DISTRIBUZIONI MARGINALI**

sono rappresentate dai totali di riga e di colonna in una tabella a doppia entrata. In pratica si tratta delle distribuzioni di frequenza dei singoli caratteri.

❑ **DISTRIBUZIONI CONDIZIONATE**

sono le colonne e le righe interne alle tabelle. Si dicono condizionate in quanto rappresentano le frequenze di un determinato carattere "condizionatamente" al fatto che il secondo carattere assuma determinati valori.

❑ **INDIPENDENZA STATISTICA**

quando la conoscenza della modalità con cui si manifesta uno dei caratteri non migliora la "previsione" delle modalità dell'altro carattere

❑ **INDIPENDENZA PERFETTA**

due caratteri X ed Y si diranno indipendenti se le distribuzioni relative condizionate di un carattere rispetto alle modalità dell'altro sono uguali

❑ **DIPENDENZA PERFETTA**

si dice che un carattere Y dipende perfettamente da X se ad ogni modalità di X è associata una ed una sola modalità del carattere Y; viceversa, un carattere X dipende perfettamente da Y se ad ogni modalità di Y è associata una ed una sola modalità del carattere X

❑ **INTERDIPENDENZA PERFETTA**

quando ad ogni modalità del carattere X corrisponde una ed una sola modalità del carattere Y e, simultaneamente, ad ogni modalità del carattere Y corrisponde una ed una sola modalità del carattere X

❑ **TABELLA QUADRATA**

è una tabella che presenta lo stesso numero di righe e di colonne

❑ **FREQUENZE OSSERVATE**

sono le frequenze che abbiamo effettivamente osservato nell'analisi delle unità statistiche considerate

❑ **FREQUENZE TEORICHE (DI INDIPENDENZA)**

sono le frequenze che ci aspetteremmo di trovare nel caso in cui i due caratteri considerati fossero tra loro indipendenti. Sono date dalla seguente relazione: $n'_{ij} = \frac{n_i \cdot n_j}{n}$, dove n_i rappresenta il totale della *i-esima* riga, n_j il totale della *j-esima* colonna e n il totale generale.

❑ **IL CHI-QUADRATO (χ^2)**

è una misura dell'associazione tra due caratteri (qualitativi o quantitativi). E' dato dalla

seguente formula: $\chi^2 = \sum_{i=1}^H \sum_{j=1}^K \frac{(n_{ij} - n'_{ij})^2}{n'_{ij}}$, nella quale vengono confrontate le frequenze

osservate con quelle teoriche. Può variare tra un minimo di zero ed un massimo determinato dal χ^2 massimo (che più sotto viene definito)

❑ **IL CHI-QUADRATO MASSIMO (χ^2_{\max})**

è il valore massimo valore raggiungibile dal Chi-quadrato ed è dato dal prodotto tra la numerosità totale e il minimo tra le righe e le colonne diminuito di uno:

$$\chi^2_{\max} = n \cdot \min[(H - 1); (K - 1)]$$

❑ **LA V DI CRAMER**

è una misura relativa dell'associazione, calcolata a partire dal χ^2 . E' data

dalla formula: $V = \sqrt{\frac{\chi^2}{\chi^2_{\max}}}$, e può variare tra 0 (indipendenza tra i due

caratteri) ad 1 (massima dipendenza tra i due caratteri).