

Quando lo studio di un fenomeno richiede l'analisi contemporanea di due o più caratteri, è importante esaminare la relazione che li lega.

Variabili Qualitative

Nel caso di una **Mutabile Doppia**, si prendono in esame due caratteri:

U con modalità $u_1, u_2 \dots u_i \dots u_r$

V con modalità $v_1, v_2 \dots v_k \dots v_c$
rilevati congiuntamente su **N** soggetti

♣ Se le coppie di modalità $[u_i, v_k]$ si presentano con frequenza non-sempre unitaria, allora la sintesi tabellare è una **distribuzione doppia di frequenze** che, nel caso di variabili qualitative, è detta **Tabella di Contingenza** e si presenta nella forma seguente.

In essa si indica con il simbolo n_{ik} la **frequenza congiunta**, ossia il numero di casi in cui si è presentata congiuntamente la *i*-esima modalità, u_i della mutabile **U** e la *k*-esima modalità, v_k della mutabile **V**.

U	V					Totali
	v_1	v_k	...	v_c	
u_1	n_{11}	n_{1k}	n_{1c}	n_{v1}
u_2	n_{21}	n_{2k}	n_{2c}	n_{v2}
:	:	:	:	:	:	
u_i	n_{i1}	n_{ik}	n_{ic}	n_{vi}
:	:	:		:		
u_r	n_{r1}	n_{rk}	n_{rc}	n_{vr}
Totali	n_{u1}	\\	n_{u2}	\\	n_{uc}	N

Lo studio delle relazioni che possono esistere tra le modalità di due Mutabili si basa unicamente sull'analisi del comportamento delle frequenze.

♣ La prima verifica da effettuare è la **verifica di indipendenza** tra le due mutabili. In generale *due mutabili sono indipendenti* se :

$$n_{ik} = \frac{n_{vi} \times n_{uk}}{N}$$

Questa formula esprime la condizione della Perfetta Indipendenza Stocastica o matematica. Si indica con il nome di *frequenza teorica* e con il simbolo v_{ik} il prodotto del totale riga per il totale colonna”, diviso il totale generale.

♣ Analisi delle Contingenze

Nel caso in cui la tavola a doppia entrata, è una tavola formata da r righe e c colonne, la verifica di indipendenza deve essere fatta su tutte le frequenze della tavola: utilizzando la formula [a] si calcolano le *frequenze teoriche* v_{ik} e si confrontano con le “frequenze empiriche” della tabella, calcolando le **contingenze**:

$$C_{ik} = n_{ik} - v_{ik}$$

Nel caso di associazione nulla tra i due caratteri, le contingenze sono tutte nulle: *perfetta indipendenza*. Se le contingenze sono diverse da zero $C_{jk} \neq 0$, vi è un legame reciproco tra le due mutabili che può essere misurato tramite alcuni indici.

Table (r x c)		
Indici	formula	caratteristiche
Indice di Associazione di Pearson	$X^2 = \sum_{i=1}^r \sum_{k=1}^c \left[\frac{(n_{ik} - v_{ik})^2}{v_{ik}} \right]$	$X^2 \geq 0$ $X^2 = 0$ assenza di associazione $X^2 > 0$ associazione Il massimo valore che può assumere l'indice X^2 è pari a: N per [numero minimo tra $(r - 1)$ e $(c - 1)$].
Coefficiente di Contingenza Quadratica	$C = \sqrt{\frac{X^2}{N + X^2}}$	$0 \leq C < 1$ se $C = 0$ i due caratteri sono <u>indipendenti</u> ; e più C si avvicina a 1 più l'associazione tra i due caratteri aumenta.

♣ **Tabelle Dicotomiche**

Se le due mutabili considerate possiedono entrambe *due sole modalità*, che esauriscono tutte le modalità possibili, ossia sono *esaustive*, la tabella è chiamata *tavola dicotomica*: per misurarne l'associazione, esistono degli indici specifici.

U	V		Totali
	v₁	v₂	
u₁	a	b	a + b
u₂	c	d	c + d
Totali	a + c	b + d	N

Nel caso di tavole dicotomiche la verifica sull'esistenza di associazione si basa sul confronto tra *a* e la frequenza teorica *a**

$$a^* = [(a + b) \times (a + c)] / N$$

Se *a = a** non c'è associazione;

se *a > a** vi è associazione positiva;

se, invece *a < a** l'associazione è di tipo inverso, o negativa.

Tavole Dicotomiche		
Indici	Formule	caratteristiche
Indice di associazione di Yule	$Q = \frac{ad - bc}{ad + bc}$	- 1 ≤ Q ≤ + 1
Indice di Collegamento	$G = \frac{\sqrt{ad} - \sqrt{bc}}{\sqrt{ad} + \sqrt{bc}}$	- 1 ≤ G ≤ + 1
Indice di Differenza (Eduards)	$E = \frac{d}{(c+d)} - \frac{b}{(a+b)}$	- 1 ≤ E ≤ + 1
Indice di Pearson	$V = \frac{ad - bc}{\sqrt{(a+b)(c+d)(b+d)(a+c)}}$	- 1 ≤ V ≤ + 1 L'indice di Pearson è il più preciso

Gli indici considerati sono compresi tra **- 1** e **+ 1**:

V = 0 assenza di associazione: le due mutabili sono indipendenti;

V = - 1 tra le due mutabili vi è massima associazione negativa, o inversa;

V = + 1 tra le due mutabili vi è massima associazione positiva, o diretta;