

Anno 2005/2006

Laurea specialistica in Comunicazione istituzionale e d'impresa
Insegnamento: valutazione di campagne di comunicazione

**Analisi dei dati relativi alla
diffusione del personal computer
nelle imprese**

Introduzione

L'**obiettivo** di questa relazione è capire se c'è una relazione, un'influenza tra la localizzazione geografica delle imprese considerate nel territorio italiano (carattere X) e l'uso del personal computer nelle imprese (carattere Y).

Il **metodo** utilizzato per verificare l'effettiva esistenza di una influenza tra i due caratteri è quello dell'*indice di connessione* χ^2 e dell'*indice di Cramér*.

La **fonte** da cui sono stati recuperati i dati è il sito internet dell'Istituto Nazionale di Ricerca – ISTAT (<http://www.istat.it>) nella sezione Industria e servizi: sviluppo e innovazione da una rilevazione sulle tecnologie dell'informazione e della comunicazione nelle imprese nel 2005.

I **dati** di partenza si riferiscono alla diffusione del personal computer nelle imprese con almeno 10 addetti per attività economica, classe di addetti e ripartizione geografica e sono relativi all'anno 2005.

I **caratteri** che sono stati presi in considerazione per studiare il fenomeno sono:

- la ripartizione geografica nelle seguenti modalità: Nord-Est, Nord-Ovest, Centro, Sud e isole;
- tipologia di utilizzo del personal computer da parte degli addetti nelle modalità: addetti che utilizzano computer e addetti che utilizzano computer connessi a Internet.

La seguente tabella è stata tratta dal sito internet dell'ISTAT:

Diffusione del personal computer nelle imprese con almeno 10 addetti per attività economica, classe di addetti e ripartizione geografica - Anno 2005
(valori percentuali sul totale delle imprese con almeno 10 addetti e degli addetti)

Ripartizioni geografiche	Addetti che utilizzano computer	Addetti che utilizzano computer connessi ad Internet
Nord-ovest	42,5	28,4
Nord-est	35,5	22,7
Centro	42,7	27,6
Sud e Isole	25,3	14,3
	146,0	93,0

Fonte: Istat, Rilevazione sulle tecnologie dell'informazione e della comunicazione nelle imprese, anni 2004 e 2005

Svolgimento

TABELLA DI PARTENZA

Ripartizioni geografiche	Addetti che utilizzano computer	Addetti che utilizzano computer connessi ad Internet	
Nord-ovest	42,5	28,4	70,9
Nord-est	35,5	22,7	58,2
Centro	42,7	27,6	70,3
Sud e Isole	25,3	14,3	39,6
	146,0	93,0	239,0

Per valutare la relazione tra i due caratteri e individuare un'eventuale influenza tra i due, si sostituisce ad ogni frequenza osservata (n_{ij}) la frequenza teorica sotto indipendenza (\hat{n}_{ij}):

$$\hat{n}_{ij} = \frac{n_{i\cdot} \cdot n_{\cdot j}}{N} \quad \begin{array}{l} i = 1, \dots, r \\ j = 1, \dots, c \end{array}$$

TABELLA DI INDIPENDENZA

Ripartizioni geografiche	Addetti che utilizzano computer	Addetti che utilizzano computer connessi ad Internet	
Nord-ovest	43,3	27,6	70,9
Nord-est	35,6	22,6	58,2
Centro	42,9	27,4	70,3
Sud e Isole	24,2	15,4	39,6
	146,0	93,0	239,0

In questo caso, confrontando la tabella di partenza e quella dei valori in condizione di indipendenza osserviamo che i valori non coincidono esattamente, ciò significa che i caratteri X e Y sono connessi. Occorre però misurare anche il livello di connessione tra i due caratteri osservati: la connessione sarà maggiore quando la tabella di partenza si discosterà notevolmente dai valori della tabella di indipendenza.

Per confrontare le due tabelle si crea la tabella di contingenza:

$$c_{ij} = n_{ij} - \hat{n}_{ij} = n_{ij} - \frac{n_{i\cdot} \cdot n_{\cdot j}}{N} \quad \begin{array}{l} i = 1, \dots, r \\ j = 1, \dots, c \end{array}$$

TABELLA DI CONTINGENZA

Ripartizioni geografiche	Addetti che utilizzano computer		Addetti che utilizzano computer connessi ad Internet	
		C_{ij}^2		C_{ij}^2
Nord-ovest	-0,8	0,7	0,8	0,7
Nord-est	-0,1	0,0	0,1	0,0
Centro	-0,2	0,1	0,2	0,1
Sud e Isole	1,1	1,2	-1,1	1,2
	0,0		0,0	

Per sintetizzare in un unico indice i risultati della tabella delle contingenze si può utilizzare l'*indice chi quadro*:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \left(\frac{c_{ij}}{\hat{n}_{ij}} \right)^2 \hat{n}_{ij} = \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$$

In questo caso $\chi^2 = 0,173548$.

Questo indice assume valori compresi tra 0 e il suo valore massimo:

$$N \cdot \min[(r-1), (c-1)]$$

Nel caso osservato $\max \chi^2 = 239$.

Se $\chi^2 = 0$ → indipendenza tra i due caratteri;

Se $\chi^2 = \max \chi^2$ → perfetta dipendenza tra i caratteri X e Y.

Per poter descrivere l'andamento della connessione tramite un indice relativo, si può fare riferimento all'*indice di Cramér*:

$$C = \sqrt{\frac{\chi^2}{N \cdot \min[(r-1), (c-1)]}}$$

Nel nostro caso $C = 0,026946$.

Conclusioni

Dal confronto tra la tabella dei dati osservati e quella dei dati calcolati sotto un'ipotetica condizione di indipendenza, si osserva che i valori delle due tabelle non sono significativamente distanti e diversi. Si può quindi ritenere che la *localizzazione geografica* delle imprese considerate influenzi minimamente la *modalità di uso del computer* da parte degli addetti.

Nella tabella di contingenza si notano alcuni valori negativi e precisamente nei casi:

- addetti che utilizzano il computer nel Nord-Ovest (-0,8), nel Nord-Est (-0,1), nel Centro (-0,2);
- addetti che utilizzano il computer connesso a Internet nel Sud e isole (-1,1).

Possiamo affermare che il numero di addetti che utilizzano semplicemente il computer è nella realtà inferiore rispetto a quello osservato in condizione di indipendenza (ad eccezione del Sud e isole) e che il numero di addetti che utilizzano il computer connesso a Internet è superiore rispetto a quello ipotizzato sotto indipendenza, con l'eccezione del Sud e isole.

L'*indice di connessione* χ^2 è pari a 0,173548 e il valore massimo che questo può raggiungere è uguale a $\max \chi^2 = 239$. Si può notare quindi, dato il valore di χ^2 e il suo valore massimo che, nel caso osservato la connessione tra i due caratteri è molto modesta e si avvicina quasi alla situazione di indipendenza. L'*indice* χ^2 inoltre è un indice bilaterale cioè ci dice che il carattere X è poco dipendente dal carattere Y così come il carattere Y è poco dipendente da X .

Per poter valutare in modo migliore il livello di questa influenza tra i caratteri X e Y occorre fare anche riferimento all'*indice di Cramér*. Nella situazione osservata C equivale a 0,026946. È un indice relativo ovvero è compreso tra 0 e 1: in questo caso quindi l'*indice di Cramér* è molto più vicino a 0. Questo significa che tra la *ripartizione geografica dell'impresa* e il *tipo di utilizzo di Interne* c'è poca connessione.