

Università di Cassino
Corso di Statistica 1
Esercitazione del 26/11/2007
Dott. Alfonso Piscitelli

Esercizio 1

Nella tabella seguente sono riportati sia i costi della pubblicità che le vendite effettuate da un'azienda in otto mesi di attività.

Pubblicità	Vendite
1	30
3	40
5	40
4	50
2	35
5	50
3	35
2	25

Verificare l'esistenza di una relazione statistica tra le due variabili.

Considerando che entrambe le variabili sono di natura quantitativa, l'indice più opportuno per verificare l'esistenza di una relazione tra le variabili è il coefficiente di correlazione lineare.

Il coefficiente di correlazione è una misura dell'INTERDIPENDENZA lineare tra due fenomeni e varia tra -1 ed +1 (perfetta correlazione inversa, perfetta correlazione diretta).

Il coefficiente di correlazione di Bravais-Pearson è dato da:

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu_X)(y_i - \mu_Y)}{\sqrt{\mu_{X^2} - \mu_X^2} \sqrt{\mu_{Y^2} - \mu_Y^2}} = \frac{\frac{1}{N} \sum_{i=1}^N x_i y_i - \mu_X \mu_Y}{\sqrt{\mu_{X^2} - \mu_X^2} \sqrt{\mu_{Y^2} - \mu_Y^2}}$$

Il numeratore si chiama **Covarianza** perché è una misura della contemporanea variazione di X e Y in rapporto alle rispettive medie.

Mesi (N)	X	Y	XY	X ²	Y ²
1	1	30	30	1	900
2	3	40	120	9	1600
3	5	40	200	25	1600
4	4	50	200	16	2500
5	2	35	70	4	1225
6	5	50	250	25	2500
7	3	35	105	9	1225
8	2	25	50	4	625
Totale	25	305	1025	93	12175

Si inizia con il calcolare la media per le due variabili

$$\mu_x = 25/8 = 3,125$$

$$\mu_y = 305/8 = 38,125$$

$$\text{Cov}(X, Y) = \frac{1}{N} \sum_{i=1}^N x_i y_i - \mu_x \mu_y = \frac{1025}{8} - 3,125 * 38,125 = 128,125 - 119,1406 = 8,9844$$

$$\sigma_x = \sqrt{\mu_{x^2} - \mu_x^2} = \sqrt{11,625 - 9,765} = 1,3635$$

$$\sigma_y = \sqrt{\mu_{y^2} - \mu_y^2} = \sqrt{1521,875 - 1453,516} = 8,2679$$

$$\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \frac{8,9844}{1,3635 * 8,2679} = 0,7969$$

Siamo in presenza di un'alta correlazione positiva tra i costi della pubblicità e le vendite effettuate.

Esercizio 2

Nella tabella seguente sono riportati sia il CONSUMO (watt) che la DURATA (ore) di trenta lampade.

		DURATA				
		70 -100	100 -150	150 -300	300 -400	
CONSUMO	5 -10	0	0	0	8	8
	10 -15	0	1	8	0	9
	15 -20	8	5	0	0	13
		8	6	8	8	30

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu_X)(y_i - \mu_Y)}{\sqrt{\mu_{X^2} - \mu_X^2} \sqrt{\mu_{Y^2} - \mu_Y^2}} = \frac{\frac{1}{N} \sum_{i=1}^N x_i y_i - \mu_X \mu_Y}{\sqrt{\mu_{X^2} - \mu_X^2} \sqrt{\mu_{Y^2} - \mu_Y^2}}$$