**Esercitazione del 25 marzo 2014**

**Esercizio** (prosecuzione dall'esercitazione del 20 marzo sull'indipendenza in media)
Un gruppo di 500 dipendenti della Pubblica Amministrazione è stato classificato secondo la qualifica (*X*) e ed il numero di assenze nell'ultimo mese (*Y*). Si è ottenuta la seguente tabella di contingenza:

| | *Y* | | | |
|---|---|---|---|---|
| *X* | 0 | 1 | 2 | 3 |
| Impiegato | 34 | 64 | 115 | 35 |
| Funzionario | 35 | 88 | 26 | 15 |
| Dirigente | 58 | 20 | 10 | 0 |

Calcolare il coefficiente $\eta^2_{Y|X}$, ove *Y* è il numero di giorni di ferie ed *X* è la funzione ricoperta per valutare il grado di dipendenza in media tra le due variabili.

**Esercizio 1 dalla prima parte del tema d'esame del 10.02.2014**
(conclusione dell'esercizio svolto il 6 marzo 2013)

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *X* | 2 | 5 | 1 | 3 | 4 | 1 | 5 | 3 | 4 | 2 |
| *Y* | 50 | 57 | 41 | 54 | 54 | 38 | 63 | 48 | 59 | 46 |
| *Z* | 46 | 53 | 48 | 40 | 42 | 42 | 60 | 51 | 52 | 43 |

Si calcoli la correlazione tra le variabili *X* ed *Y* così come tra *X* e *Z.* In quale stato la campagna pubblicitaria ha dato maggiori risultati?

***<u>Soluzione completa con i valori esatti del coeff. di correlazione</u>***
Il coefficiente di correlazione lineare, *ρ*, tra due variabili *X* ed *Y* è definito come, usando le solite notazioni per medie e varianze:

$$\rho_{X,Y} = \frac{n^{-1}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sigma_x^2\sigma_y^2}}$$

Con semplici passaggi algebrici si vede che:

$$\rho_{X,Y} = \frac{n^{-1}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sigma_x^2\sigma_y^2}} = \frac{\sum_{i=1}^{n}x_iy_i - n\bar{x}\bar{y}}{\sqrt{\left(\sum_i x_i^2 - n\bar{x}^2\right)\left(\sum_i y_i^2 - n\bar{y}^2\right)}}$$

Con i nostri dati:

$$\bar{x}=3,\ \bar{y}=51,\ \bar{z}=47.7$$
$$\sum_i x_i^2 = 110,\ \sum_i y_i^2 = 26576,\ \sum_i z_i^2 = 23111$$
$$\sum_i x_iy_i = 1629$$

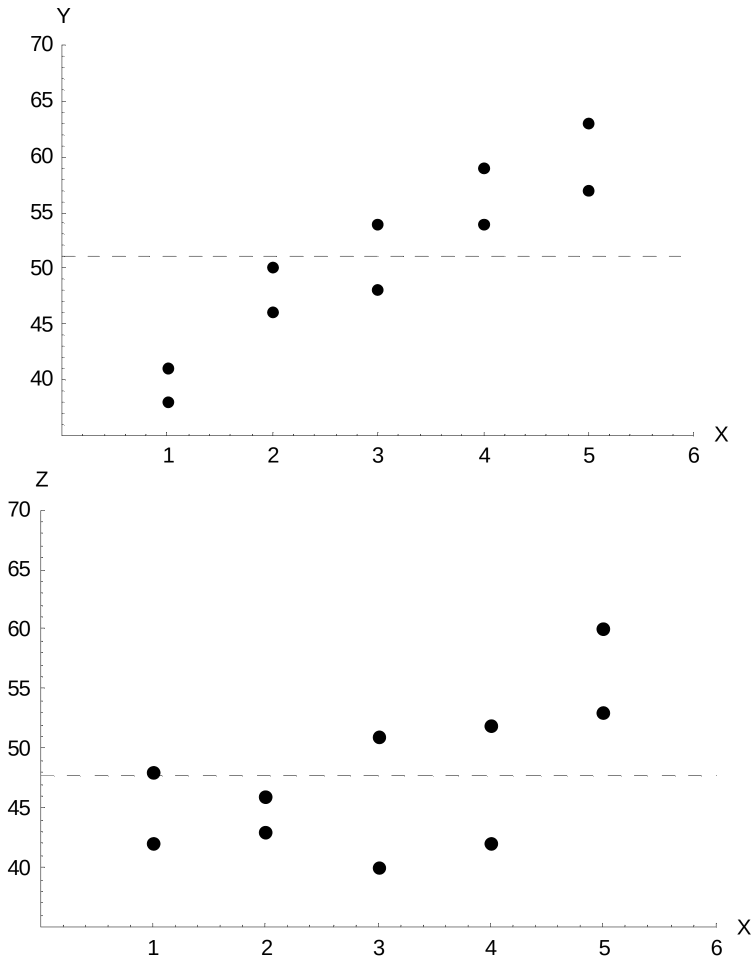
quindi:

$$\rho_{X,Y} = \frac{1629 - 10\times 3\times 51}{\sqrt{(110-10\times 3^2)(26576-10\times 51^2)}} = 0.93$$

e, allo stesso modo:

$$\rho_{X,Z} = \frac{1482 - 10 \times 3 \times 47.7}{\sqrt{(110 - 10 \times 3^2)(23111 - 10 \times 47.7^2)}} = 0.60$$

Dunque, tra *X* ed *Y* c'è una relazione lineare positiva molto forte (0.93), mentre tra *X* e *Z* c'è una relazione lineare positiva media (0.60). La figura seguente mostra il diagramma di dispersione di *Y* verso *X* e di *Z* verso *X*, avendo considerato *X* come variabile *indipendente* in entrambi i casi e *Y* e *Z*, rispettivamente, come variabile *dipendente.*




Dalla figura si vede come la relazione lineare tra *X* ed *Y* sia piuttosto chiara (sopra), mentre nel caso di *X* e *Z* la minore linearità è visualizzata in un andamento delle entrate tutto sommato "costante" rispetto al numero di spot trasmessi: i punti sono un po' sopra un po' sotto la media delle entrate settimanali, indicata dalla linea tratteggiata orizzontale, anche per valori più alti del numero degli spot. Possiamo quindi concludere che la campagna sembra avere dato maggiori effetti nello stato *Y* dove è più evidente l'aumentare delle entrate al crescere di spot trasmessi: per i valori più alti del numero di spot i punti sono sopra la media. E tale aumentare è fortemente lineare.

**Esercizio**
Relativamente alla prima sessione d'esame dei corsi del secondo anno del corso di laurea in Economia Aziendale si ha che la probabilità

- di superare Statistica è 0.4;
- di superare Micro Economia è 0.5
- di superare Macro Economia è 0.3
- di superare sia Statistica che Micro E. è 0.35
- di superare sia Micro E. che Macro E. è 0.25
- di superare sia Statistica che Macro E. è 0.2
- di superare tutti e tre gli esami è 0.15.

Determinare la probabilità che alla prima sessione d'esame di quest'anno uno studente scelto a caso

a. non superi l'esame di Statistica
b. superi Statistica ma non Micro Economia
c. superi almeno un esame
d. non superi alcun esame.

Il superamento dei tre esami forma una terna di eventi indipendenti? Se uno studente ha già passato Statistica, qual è la probabilità che superi Micro Economia?

**Esercizio**
Un'urna chiusa di cui non sia possibile vedere l'interno contiene 6 palline nere, 5 palline rosse e 4 palline bianche. Le palline sono tutte della stessa dimensione, indistinguibili al tatto. Estraendo a caso una pallina dall'urna, qual è la probabilità che

a. la pallina sia nera
b. la pallina sia rossa o nera
c. la pallina non sia rossa.

Estraiamo ora due palline consecutivamente rimettendo ogni volta la pallina nell'urna: calcolare la probabilità che nessuna delle due palline sia rossa. Determinare la stessa probabilità nel caso in cui le palline **non** vengano rimesse nell'urna.

**Esercizio**
Si lancia un dado equilibrato e si vince se esce il numero 1 od il numero 6. Qual è la probabilità di vincere in un singolo lancio del dado?
Si lancia il dado 5 volte e si vince *ogni volta che*[1] esce un 1 od un 6.

a. Qual è la probabilità che la prima vittoria sia al terzo lancio?
b. Qual è la probabilità di vincere al primo ed al terzo lancio ma non negli altri?
c. Senza fare nessun calcolo, dire qual è la probabilità di vincere al secondo ed al quarto lancio ma non negli altri.
d. Qual è la probabilità di non vincere mai?

**Esercizio**
Andate in libreria alla ricerca di due libri, *It* e *Shining.* La probabilità che la libreria abbia *It* è 0.5, che abbia *Shining* è 0.6, che sia sguarnita di almeno uno dei due è 0.65.
Usando la notazione per eventi: I = *la libreria possiede una copia di It;* S = *la libreria possiede una copia di Shining,* formalizzate l'evento *la libreria è sguarnita di almeno uno dei due libri*. Calcolate

[1] Per chi era a lezione: specifico che considero il caso delle telefonate compulsive, da perditempo, al centralino bagagli smarriti della Malpensa. Il testo è scritto qui correttamente per questo caso.

la probabilità che la libreria si sfornita di entrambi. La vendita dei due libri forma una coppia di eventi indipendenti?