# SOLUZIONI

**della Prova Scritta di Statistica del 9 Febbraio 2005**
**Modalità B1 e B2**

## Esercizio 1

### punto a

#### Modalità B1/B2

Ricordiamo che la mediana può essere calcolata per caratteri misurati almeno su scala ordinale; il rapporto di concentrazione $R$ può essere invece calcolato solo per caratteri quantitativi trasferibili. Il carattere $B$ è qualitativo ordinale; dunque per esso può essere calcolata solo la mediana. Il carattere $X$ è quantitativo discreto non trasferibile; dunque per esso può essere calcolata solo la mediana.

### punto b

#### Modalità B1

Utilizzando i totali forniti e la formula abbreviata per il calcolo della varianza, si ottiene:

$$\begin{aligned}
\mu(Y) &= \frac{178}{200} = 0.89 \\
\sigma^2(Y) &= \frac{538}{200} - (0.89)^2 = 1.8979 \\
\sigma(Y) &= \sqrt{1.8979} = 1.3776
\end{aligned}$$

#### Modalità B2

Utilizzando i totali forniti e la formula abbreviata per il calcolo della varianza, si ottiene:

$$\begin{aligned}
\mu(Y) &= \frac{92}{200} = 0.46 \\
\sigma^2(Y) &= \frac{116}{200} - (0.46)^2 = 0.3684 \\
\sigma(Y) &= \sqrt{0.3684} = 0.6070
\end{aligned}$$

### punto c

#### Modalità B1

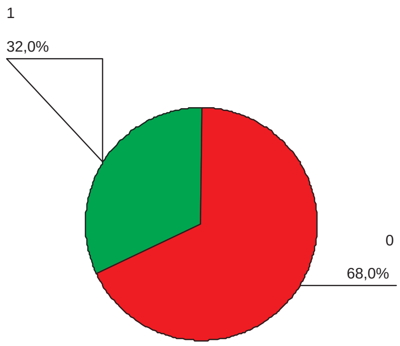
Il carattere $A$ è di tipo qualitativo nominale; la sua distribuzione può essere desunta dal totale dei valori:

$$A \equiv \begin{cases} 0 & 1 \\ 136 & 64 \end{cases}$$

o, usando le frequenze relative,

$$A \equiv \begin{cases} 0 & 1 \\ 0.68 & 0.32 \end{cases}$$

La rappresentazione grafica piu' adeguata è il diagramma a torta:




### Modalità B2

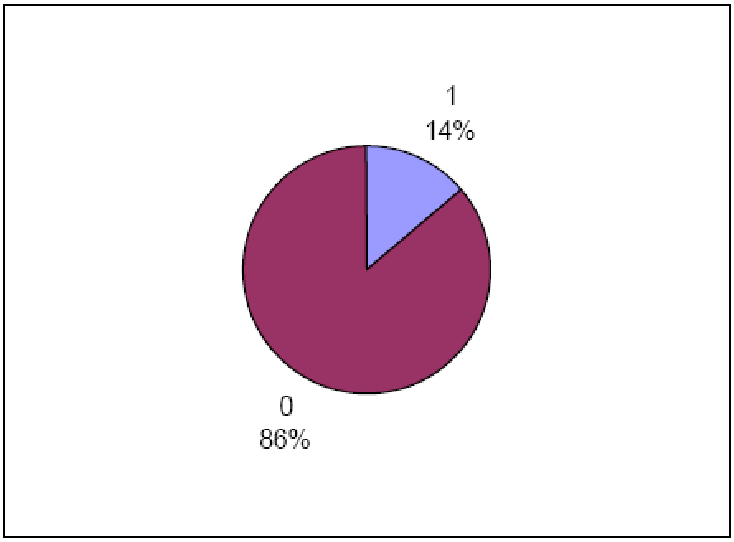
Il carattere $A$ è di tipo qualitativo nominale; la sua distribuzione può essere desunta dal totale dei valori:

$$A \equiv \begin{cases} 0 & 1 \\ 172 & 28 \end{cases}$$

o, usando le frequenze relative,

$$A \equiv \begin{cases} 0 & 1 \\ 0.86 & 0.14 \end{cases}$$

La rappresentazione grafica piu' adeguata è il diagramma a torta.




## punto d

### Modalità B1

Perchè si abbia indipendenza regressiva di $Y$ da $B$, tutte le medie condizionate di $Y$ fissato $B$ devono essere uguali e devono coincidere con $\mu(Y)$. Poichè $\mu(Y|B = 3) = 1.22$ è diverso da $\mu(Y) = 0.89$, $Y$ non è regressivamente indipendente da $B$.

**Modalità B2**

Perchè si abbia indipendenza regressiva di $Y$ da $B$, tutte le medie condizionate di $Y$ fissato $B$ deno essere uguali e devono coincidere con $\mu(Y)$. Poichè $\mu(Y|B = 3) = 0.33$ è diverso da $\mu(Y) = 0.46$, $Y$ non è regressivamente indipendente da $B$.

## punto e

**Modalità B1**

In primo luogo, sostituiamo alle classi utilizzate per il carattere $X$ il loro valore centrale:

| $X \backslash Y$ | 0 | 1 | 4 | |
|---|---|---|---|---|
| 65 | 0.24 | 0.02 | 0 | 0.26 |
| 80 | 0.32 | 0.09 | 0 | 0.41 |
| 95 | 0 | 0.18 | 0.15 | 0.33 |
| | 0.56 | 0.29 | 0.15 | 1 |

Al fine di applicare la formula $\text{Cov}(X, Y) = \mu(XY) - \mu(X)\,\mu(Y)$ per il calcolo della covarianza, nella prossima tabella riportiamo, per ciascuna cella, il prodotto delle corrispondenti modalità dei due caratteri:

| $X \backslash Y$ | 0 | 1 | 4 |
|---|---|---|---|
| 65 | 0 | 65 | 260 |
| 80 | 0 | 80 | 320 |
| 95 | 0 | 95 | 380 |

Calcoliamo ora:

$$\begin{aligned}
\mu(XY) &= (65)(0.02) + (80)(0.09) + (95)(0.18) + (380)(0.15) = 82.6 \\
\mu(X) &= (65)(0.26) + (80)(0.41) + (95)(0.33) = 81.05 \\
\text{Cov}(X, Y) &= 82.6 - (81.05)(0.89) = 10.4655 \\
\sigma^2(X) &= (65)^2(0.26) + (80)^2(0.41) + (95)^2(0.35) - (81.05)^2 = 131.6475 \\
\sigma(X) &= \sqrt{131.6475} = 11.4738
\end{aligned}$$

Infine, si ottiene

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\,\sigma(Y)} = \frac{10.4655}{(11.4738)(1.3776)} = 0.6621.$$

**Modalità B2**

In primo luogo, sostituiamo alle classi utilizzate per il carattere $X$ il loro valore centrale:

| $X \backslash Y$ | 0 | 1 | 3 | |
|---|---|---|---|---|
| 21 | 0.2 | 0.33 | 0 | 0.53 |
| 26 | 0.23 | 0.07 | 0.02 | 0.32 |
| 29 | 0.15 | 0 | 0 | 0.15 |
| | 0.58 | 0.4 | 0.02 | 1 |

Al fine di applicare la formula $\mathrm{Cov}(X,Y) = \mu(XY) - \mu(X)\,\mu(Y)$ per il calcolo della covarianza, nella prossima tabella riportiamo, per ciascuna cella, il prodotto delle corrispondenti modalità dei due caratteri:

| $X \backslash Y$ | 0 | 1 | 3 |
|---|---|---|---|
| 21 | 0 | 21 | 63 |
| 26 | 0 | 26 | 78 |
| 29 | 0 | 29 | 87 |

Calcoliamo ora:

$$\begin{aligned}
\mu(XY) &= (21)(0.33) + (26)(0.07) + (78)(0.02) = 10.31 \\
\mu(X) &= (21)(0.53) + (26)(0.32) + (29)(0.15) = 23.8 \\
\mathrm{Cov}(X,Y) &= 10.31 - (23.8)(0.46) = -0.638 \\
\sigma^2(X) &= (21)^2(0.53) + (26)^2(0.32) + (29)^2(0.15) - (23.8)^2 = 9.76 \\
\sigma(X) &= \sqrt{9.76} = 3.1241
\end{aligned}$$

Infine, si ottiene

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma(X)\,\sigma(Y)} = \frac{-0.638}{(3.1241)(0.6070)} = -0.3364.$$

# Esercizio 2

## punto a

### Modalità B1

Dal testo si evince che il numero degli impiegati che lavorano part-time, a prescindere dal sesso, è pari a 80+7=87. Pertanto, in base alla definizione classica di probabilità, otteniamo

$$\begin{aligned}
P(A \cap B) &= \frac{80}{250} \\
P(B) &= \frac{87}{250}
\end{aligned}$$

da cui si ottiene

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{80/250}{87/250} = \frac{80}{87} = 0.9195.$$

### Modalità B2

Dal testo si evince che il numero degli operai che fanno il turno notturno, a prescindere dalla provenienza geografica, è pari a 40+10=50. Pertanto, in base alla definizione classica di probabilità, otteniamo

$$\begin{aligned}
P(A \cap B) &= \frac{10}{250} \\
P(B) &= \frac{50}{250}
\end{aligned}$$

da cui si ottiene

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{10/250}{50/250} = \frac{1}{5} = 0.2.$$

### punto b

#### Modalità B1

Dal testo si evince che il numero degli impiegati che *non* lavorano part-time, a prescindere dal sesso, è pari a 250-87=163. Inoltre, il numero delle donne impiegate che *non* lavorano part-time è pari a 180-80=100. Pertanto, in base alla definizione classica di probabilità, otteniamo

$$\begin{aligned} P(A) &= \frac{180}{250} \\ P(\bar{B}) &= \frac{163}{250} \\ P(A \cap \bar{B}) &= \frac{100}{250} \end{aligned}$$

da cui si ottiene

$$P(A \cup \bar{B}) = P(A) + P(\bar{B}) - P(A \cap \bar{B}) = \frac{180}{250} + \frac{163}{250} - \frac{100}{250} = \frac{243}{250} = 0.972.$$

Nota Bene: Utilizzando le leggi di De Morgan, si può ugualmente calcolare:

$$P(A \cup \bar{B}) = 1 - P(\bar{A} \cap B) = 1 - \frac{7}{250} = \frac{243}{250}.$$

#### Modalità B2

Dal testo si evince che il numero degli operai che *non* fanno il turno notturno, a prescindere dalla provenienza geografica, è pari a 250-50=200. Inoltre, il numero degli operai extracomunitari che *non* fanno il turno notturno è pari a 25-10=15. Pertanto, in base alla definizione classica di probabilità, otteniamo

$$\begin{aligned} P(A) &= \frac{25}{250} \\ P(\bar{B}) &= \frac{200}{250} \\ P(A \cap \bar{B}) &= \frac{15}{250} \end{aligned}$$

da cui si ottiene

$$P(A \cup \bar{B}) = P(A) + P(\bar{B}) - P(A \cap \bar{B}) = \frac{25}{250} + \frac{200}{250} - \frac{15}{250} = \frac{210}{250} = 0.84.$$

Nota Bene: Utilizzando le leggi di De Morgan, si può ugualmente calcolare:

$$P(A \cup \bar{B}) = 1 - P(\bar{A} \cap B) = 1 - \frac{40}{250} = \frac{210}{250}.$$

### punto c

Ricordiamo che la funzione di densità della legge Uniforme Continua sull'intervallo $(a, b)$ è

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{altrove} \end{cases}$$

### Modalità B1

Nel nostro caso la funzione di densità di $X$ è dunque

$$f(x) = \begin{cases} \frac{1}{4} & 0 < x < 4 \\ 0 & \text{altrove} \end{cases}$$

Pertanto, possiamo calcolare

$$\begin{aligned} \Pr\{(X < 1) \cup (X > 3.5)\} &= \Pr\{X < 1\} + \Pr\{X > 3.5\} = \\ &= \int_0^1 \frac{1}{4}\,\mathrm{d}x + \int_{3.5}^4 \frac{1}{4}\,\mathrm{d}x = \\ &= \frac{1}{4}(1-0) + \frac{1}{4}(4-3.5) = 0.375 \end{aligned}$$

### Modalità B2

Nel nostro caso la funzione di densità di $X$ è dunque

$$f(x) = \begin{cases} \frac{1}{10} & 0 < x < 10 \\ 0 & \text{altrove} \end{cases}$$

Pertanto, possiamo calcolare

$$\begin{aligned} \Pr\{(X < 5) \cup (X > 8.5)\} &= \Pr\{X < 5\} + \Pr\{X > 8.5\} = \\ &= \int_0^5 \frac{1}{10}\,\mathrm{d}x + \int_{8.5}^{10} \frac{1}{10}\,\mathrm{d}x = \\ &= \frac{1}{10}(5-0) + \frac{1}{10}(10-8.5) = 0.65. \end{aligned}$$

# Esercizio 3

## punto a

Ricordiamo che l'espressione dell'intervallo di confidenza per $\mu$ (essendo $\sigma^2$ non noto) è

$$\left(\bar{x} - t_{1-\alpha/2}^{(n-1)} \frac{s_c}{\sqrt{n}} \quad ; \quad \bar{x} + t_{1-\alpha/2}^{(n-1)} \frac{s_c}{\sqrt{n}}\right)$$

### Modalità B1

Nel nostro caso si ha:

$$\begin{aligned} n &= 6 \\ \bar{x} &= \frac{41+40+38+45+40+39}{6} = 40.5 \\ s_c^2 &= \frac{(41-40.5)^2 + (40-40.5)^2 + (38-40.5)^2 + (45-40.5)^2 + (40-40.5)^2 + (39-40.5)^2}{5} = \\ &= 5.9 \\ s_c &= \sqrt{5.9} = 2.4290 \\ t_{0.95}^{(5)} &= 2.015 \end{aligned}$$

e dunque

$$\left(40.5 - 2.015\frac{2.4290}{\sqrt{6}} \quad ; \quad 40.5 + 2.015\frac{2.4290}{\sqrt{6}}\right)$$

ossia (38.5019; 42.4981).

**Modalità B2**

Nel nostro caso si ha:

$$
\begin{aligned}
n &= 5 \\
\bar{x} &= \frac{122+127+125+130+122}{5} = 125.2 \\
s_c^2 &= \frac{(122-125.2)^2+(127-125.2)^2+(125-125.2)^2+(130-125.2)^2+(122-125.2)^2}{4} = \\
&= 11.7 \\
s_c &= \sqrt{11.7} = 3.4205 \\
t_{0.995}^{(4)} &= 4.604
\end{aligned}
$$

e dunque

$$\left(125.2 - 4.604\frac{3.4205}{\sqrt{5}} \quad ; \quad 125.2 + 4.604\frac{3.4205}{\sqrt{5}}\right)$$

ossia $(118.1573; 132.2427)$.

## punto b

**Modalità B1**

Per il collegamento tra gli intervalli di confidenza e i test bilaterali, poichè il valore $\mu_0 = 42$ cade nell'intervallo determinato al punto precedente, l'ipotesi nulla viene accettata.

**Modalità B2**

Per il collegamento tra gli intervalli di confidenza e i test bilaterali, poichè il valore $\mu_0 = 125$ cade nell'intervallo determinato al punto precedente, l'ipotesi nulla viene accettata.

## punto c

Ricordiamo che, per le ipotesi $\mathrm{H}_0 : \mu \leq \mu_0$ contro $\mathrm{H}_1 : \mu > \mu_0$ la regione critica ottimale è data da

$$R \equiv \left\{(x_1, \ldots, x_n) : \frac{\bar{x} - \mu_0}{s_c/\sqrt{n}} \geq t_{1-\alpha}^{(n-1)}\right\}.$$

Pertanto, indicata con

$$T^{(n-1)} = \frac{\bar{X} - \mu_0}{S_c/\sqrt{n}}$$

la statistica-test utilizzata e con $t_{oss}$ il suo valore osservato, il $p-$value si calcola come

$$\Pr\{T^{(n-1)} \geq t_{oss} \,|\, \mu = \mu_0\};$$

la probabilità sopra si può calcolare notando che, sotto l'ipotesi $\mu = \mu_0$, la statistica-test $T^{(n-1)}$ si distribuisce secondo la legge T di Student con $n-1$ gradi di libertà.

**Modalità B1**

Nel nostro caso si ha:

$$\begin{aligned} n &= 6 \\ t_{oss} &= \frac{40.5 - 36.5}{2.4290/\sqrt{6}} = 4.0337 \end{aligned}$$

e dunque il $p-$value è dato da

$$\Pr\{T^{(5)} \geq 4.0337 | \mu = 36.5\} = 1 - 0.995 = 0.005$$

**Modalità B2**

Nel nostro caso si ha:

$$\begin{aligned} n &= 5 \\ t_{oss} &= \frac{125.2 - 120.95}{3.4205/\sqrt{5}} = 2.7783 \end{aligned}$$

e dunque il $p-$value è dato da

$$\Pr\{T^{(4)} \geq 2.7783 | \mu = 120.95\} = 1 - 0.975 = 0.025$$

## punto d

**Modalità B1**

Poichè $\alpha = 0.01$ è maggiore del $p-$value l'ipotesi nulla viene rifiutata.
Nota Bene: Se si è utilizzato il valore 0.042, $\alpha = 0.01$ è minore del $p-$value e dunque l'ipotesi nulla viene accettata.

**Modalità B2**

Poichè $\alpha = 0.05$ è maggiore del $p-$value l'ipotesi nulla viene rifiutata.
Nota Bene: Se si è utilizzato il valore 0.073, $\alpha = 0.05$ è minore del $p-$value e dunque l'ipotesi nulla viene accettata.

## punto f

**Modalità B1/B2**

In questo caso la popolazione (indicata nuovamente con $X$) si distribuisce secondo la legge bernoulliana di parametro $p$. Uno stimatore ottimale per questo parametro è la *proporzione campionaria*

$$\hat{P} = \frac{1}{n} \sum_{i=1}^{n} X_i;$$

la sua varianza è data dalla seguente espressione:

$$\mathrm{Var}(\hat{P}) = \frac{p(1-p)}{n}.$$