ESAME 12 Gennaio 2015 COMPITO C *Cognome* *Nome*
Numero di matricola
1) Approssimare tutti i calcoli alla quarta cifra decimale. ***2) Ai fini della valutazione si terrà conto solo ed esclusivamente di quanto riportato negli appositi spazi.*** 3) ***Al termine della prova, è OBBLIGATORIO consegnare il presente foglio ed il foglio di brutta (DI CUI NON SI TERRÀ CONTO AI FINI DELLA VALUTAZIONE).***

1. Il preside di una scuola sospetta che gli alunni che frequentano la sua scuola abbiano problemi di apprendimento. In particolare vuole verificare la preparazione in ortografia, propone un dettato a tutti i bambini. Detta un testo di 20 parole e conta gli errori in ciascun testo. Gli alunni di seconda elementare in Italia in un dettato di 20 parole commettono, in media, 4 errori.
Nella tabella seguente sono riportati i risultati ottenuti esaminando il dettato di 120 alunni:

Numero di errori	n_i
1	10
2	15
3	15
4	30
5	20
6	15
7	10
8	5

x	n	x*n	x^2	n*x^2	fi	Fi
1	10	10	1	10	0,083333	0,083333
2	15	30	4	60	0,125	0,208333
3	15	45	9	135	0,125	0,333333
4	30	120	16	480	0,25	0,583333
5	20	100	25	500	0,166667	0,75
6	15	90	36	540	0,125	0,875
7	10	70	49	490	0,083333	0,958333
8	5	40	64	320	0,041667	1
	120	505		2535		
					0	
		4,208333		21,125		
				3,414931		

a) Determinare media, moda, mediana e varianza del “Numero di errori in un dettato di 20 parole”

Media (1punto) **4.2083**	**Mediana (1punto)** **4**	**Moda(1punto)** **4**
Varianza (1punto) **$Var(x)=21,125-4.2083^2=3.4149$**		

b) La frequenza relativa con cui un alunno commette meno (o uguale) di 3 errori

Frequenza (1punto)

Freq(X≤3)=0.333

Sia ora X la variabile aleatoria “numero di errori”, e si assumono non note media e varianza

a) Determinare una stima puntuale di μ, numero medio di errori in un dettato di 20 parole **(1punto)**

Stima μ

$\bar{x} = 4.2083$

b) Definire le proprietà dello stimatore utilizzato. **(2punti)**

- Non distorto
- consistente
- asintoticamente normale

c)Utilizzando i dati del campione costruire l’intervallo di confidenza al 90% per μ.**(2punti)**

Intervallo di confidenza

$$90\%CI = \left[\bar{x} - z_{1-\alpha/2}\frac{s}{\sqrt{120}}; \bar{x} + z_{1-\alpha/2}\frac{s}{\sqrt{120}}\right] = \left[4.2083 - 1.645\frac{1.848}{\sqrt{120}}; 4.2083 + 1.645\frac{1.848}{\sqrt{120}}\right] = [3.9308, 4.4858]$$

$z_{1-\alpha/2} = 1.645$

$S^2 = 3.4149$

$s = 1.848$

d) Come cambia l'ampiezza dell'intervallo di confidenza se, a parità di altre condizioni, aumentiamo il livello di confidenza? **(1punto)**

Ampiezza dell’intervallo di confidenza aumenta

e) Verificare, ad un livello di significatività α = 0.01, l'ipotesi che il valore medio del numero di errori nella scuola di interesse sia uguale al valore medio nazionale, contro l’alternativa che sia superiore [H_0: μ = 4 vs H_1: μ >4] **(2punti)**

e)Verifica di ipotesi

$$R = \left\{ \frac{\bar{x}-4}{\frac{s}{\sqrt{120}}} \geq 2.326 \right\} \qquad \frac{\bar{x}-4}{\frac{s}{\sqrt{120}}} = \frac{4.2083-4}{\frac{1.848}{\sqrt{120}}} = 1.235 \quad ACCETTO$$

$$p-value = P\left(\bar{X} \geq 4.2083\right) = P\left(\frac{\bar{X}-4}{\frac{s}{\sqrt{120}}} \geq \frac{4.2083-4}{\frac{1.848}{\sqrt{120}}} \right) = P(Z > 1.235) = 0.108$$

f) Quali conclusioni può trarre il preside? Gli alunni della scuola in questione sono effettivamente meno bravi in ortografia, è quindi necessario incentivare la scrittura? **(1punto)**

Gli studenti della scuola NON sono meno preparati in ortografia.
NON è necessario un intervento per incentivare la scrittura

2. Un'azienda vuole analizzare l'efficacia della campagna pubblicitaria adottata negli ultimi 10 mesi sui propri profitti. A riguardo, sono stati raccolti a fine di ogni mese i dati relativi alla spesa pubblicitaria (X) ed alle vendite (Y) (entrambi espressi in migliaia di euro). I dati a disposizione sono i seguenti:

$$\sum_{i=1}^{10} x_i = 50 \quad \sum_{i=1}^{10} y_i = 120 \quad \sum_{i=1}^{10} x_i^{\,2} = 410 \quad \sum_{i=1}^{10} (y_i - \bar{y})^2 = 1000 \quad \sum_{i=1}^{12} x_i y_i = 920$$

		media		
somma x	50	5	covarianza	32
sommay	120	12	var x	16
somma x2	410	41	var y	100
somma (y-ymedio)2	1000	100		
somma xy	920	92	corr	0,8

b1	2
b0	2

a) Specificare un modello di regressione lineare semplice che descriva la relazione tra le vendite (Y) e la spesa pubblicitaria (X) e stimarne i coefficienti con il metodo dei minimi quadrati. **(2punti)**

y=2+2x

b) Calcolare il coefficiente di correlazione tra la spesa pubblicitaria e le vendite. **(2punti)**

$$\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{32}{\sqrt{100 \times 16}} = 0.8$$

c)Sulla base dei risultati, quali conclusioni può trarre l'azienda? **(1punto)**

Le vendite aumentano all'aumentare delle spese

c) secondo il modello, quali sono le vendite attese nel prossimo mese se si pianifica di spendere 6 mila euro in pubblicità? **(1punto)**

$$\hat{y} = 2 + 2 \times 6 = 14$$

3. Data una popolazione con media μ e varianza σ^2 e un campione casuale semplice di quattro variabili, considerare il seguente stimatore per μ

$$T_1 = -\frac{1}{5}X_1 + \frac{2}{5}X_2 - X_3 + \frac{3}{5}X_4 .$$

a) Indicare se T è uno stimatore corretto per μ**(1punto)**

b) Calcolare l'errore quadratico medio di T**(1punto)**

a)

$$E(T_1) = -\frac{1}{5}E(X_1) + \frac{2}{5}E(X_2) - E(X_3) + \frac{3}{5}E(X_4) = \mu\left(-\frac{1}{5} + \frac{2}{5} - 1 + \frac{3}{5}\right) = -\frac{1}{5}\mu$$

EQM

$$Var(T_1) = \frac{1}{25}Var(X_1) + \frac{4}{25}Var(X_2) + Var(X_3) + \frac{9}{25}Var(X_4) = \frac{39}{25}\sigma^2$$

$$B(T_1) = E(T_1) - \mu = -\frac{1}{5}\mu - \mu = -\frac{6}{5}\mu$$

$$EQM(T_1) = Var(T_1) + B(T_1)^2 = \frac{39}{25}\sigma^2 + \frac{36}{25}\mu^2$$

4. Enunciare il teorema di Bayes

5. Si indichi se le seguenti affermazioni sono Vere o False

Affermazione	**Vero / Falsa?**
Due eventi indipendenti con probabilità positiva non possono essere incompatibili.	
La covarianza tra due variabili aleatorie è sempre minore di 1	

6. E' noto che il punteggio – in unità convenzionali - riportato dagli studenti nella prova d'ingresso in una certa Università si distribuisce secondo una legge normale con media pari a 180 e varianza 1600.

a) Si determini la percentuale di individui che ottengono un punteggio superiore a 200; **(1punto)**

b)Se l'Università è interessata ad ammettere il 70% dei candidati, quale dovrà essere il punteggio di soglia perché uno studente sia ammesso? **(1punto)**

a)

```
1-pnorm(200,180,sqrt(1600))
[1] 0.3085375
```

b)

30-esimo percentile distribuzione X

```
159.2
```

30-esimo percentile distribuzione normale standardizzata X

-0.52

Soglia=180-0.52*40=159.2

c) Scegliendo a caso tre studenti che hanno sostenuto la prova d'ingresso, si calcoli la probabilità che almeno due di essi siano stati ammessi. **(1punto)**

d) Scegliendo a caso 150 studenti che sosterranno la prova d'ingresso, ci calcoli la probabilità che non più di 100 studenti verranno ammessi? **(1punto)**

c)

```
> dbinom(2,3,.7)
[1] 0.441
> dbinom(3,3,.7)
[1] 0.343
> sum(dbinom(2:3,3,.7))
[1] 0.784
```

d)

```
> pnorm(100,150*.7,sqrt(.7*.3)*sqrt(150))
[1] 0.1864992
```