

FIRMA DELLO STUDENTE

**SECONDA PROVA INTERMEDIA DI STATISTICA**

CLEA, CLEFIN, CLAPI (COD. 5047/4038)

CLELI 4038

371/377

**14 Gennaio 2004**

*Cognome*

*Nome*

*Numero di matricola*

*Codice corso*

**COMPITO A1**

Ai fini della valutazione si terrà conto solo ed esclusivamente di quanto riportato negli appositi spazi. 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**).

**APPROSSIMARE TUTTI I CALCOLI ALLA QUARTA CIFRA DECIMALE**

- 1) La Befana ha preparato dei sacchetti con dei dolci, che possono essere di due dimensioni:
- sacchetti piccoli, il cui peso in grammi è distribuito come una normale, di *media* 100 e *scarto quadratico medio* pari a 5;
  - sacchetti grandi, il cui peso in grammi è distribuito come una normale, di *media* 200 e *varianza* 75.
- a) Calcolare la probabilità che scelto un sacchetto piccolo a caso tra quelli preparati dalla Befana, questo pesi più di 115 grammi. **[1 punto]**

- b) Scrivere la distribuzione del peso di un pacco regalo, costituito come insieme di un sacchetto piccolo e un sacchetto grande scelti indipendentemente l'uno dall'altro. **[1 punto]**

- c) Calcolare la probabilità che un pacco regalo pesi meno di 280 grammi. **[1 punto]**

2) Il vettore aleatorio  $(X,Y)$  ha una distribuzione di probabilità rappresentata da

	<b>X \ Y</b>	<b>0</b>	<b>2</b>	
<b>1</b>		0,1	0,3	0,4
<b>3</b>		0	0,6	0,6
		0,1	0,9	1

a) Dare la distribuzione di  $W=X-Y$  [1 punto]

b) Calcolare  $Cov(W,X)$  [1 punto]

a)	b)
----	----

3) Sia  $X$  una popolazione distribuita secondo la legge Bernoulliana di parametro  $p$ . Sia inoltre  $(X_1, X_2, X_3, X_4, X_5)$  un campione casuale semplice estratto da  $X$ .

a) Dimostrare (esplicitando il procedimento seguito) che

$$T_5 = \frac{1}{5}(4X_1 - 3X_2 + X_3 + X_4 + X_5)$$

è distorto per  $p$  e valutare la distorsione [2 punti]

b) Calcolare la varianza di tale stimatore [2 punti]

c) Partendo da  $T_5$  trovare uno stimatore non distorto per  $p$  [1 punto]

d) Supponendo che la popolazione sia di 1000 elementi, trovare uno stimatore non distorto per  $1000p$  (totale di  $X$  per la popolazione) a partire dallo stimatore calcolato in c) [1 punto]

4) In un concorso di Miss Italia viene rilevata, tra le altre informazioni, l'altezza (in cm) delle partecipanti ( $X$ ). Si ipotizza che  $X \sim N(\mu, \sigma^2)$ , con  $\mu$  e  $\sigma^2$  ignote. Estratto un campione Bernoulliano di 5 Miss, la realizzazione campionaria ottenuta è la seguente: (176, 184, 169, 190, 181).

Determinare l'intervallo di confidenza per  $\mu$  a livello  $1-\alpha = 0,95$  indicando:

a) l'espressione analitica in simboli dell'intervallo di confidenza [1 punto]

--

b) valore del percentile letto sulle tavole [1 punto]

--

c) intervallo di confidenza (basato sulla realizzazione campionaria sopra riportata) specificando il valore di tutte le quantità coinvolte [2 punti]

--

d) La lunghezza dell'intervallo cambierebbe se avessimo osservato un altro campione (sempre di ampiezza 5)? Perché?

[2 punti]

5) In un sacchetto di mandarini ci sono dei mandarini guasti e dei mandarini buoni. Basandoci su un campione casuale semplice di 4 estrazioni (in cui ad ogni estrazione si associa 1 se si è pescato un mandarino guasto e 0 altrimenti) si vuole valutare l'ipotesi che la probabilità di trovare un mandarino guasto sia  $\frac{1}{4}$  contro l'ipotesi che la probabilità di trovare un mandarino guasto sia  $\frac{3}{4}$ . ( $H_0: p = \frac{1}{4}$ ;  $H_1: p = \frac{3}{4}$ )

Allo scopo è stata proposta la seguente regione di rifiuto:  $R = \{(x_1, x_2, x_3, x_4) : (x_1 + x_2 + x_3 + x_4) \geq 3\}$ .

a) Se osserviamo  $x_1=1, x_2=0, x_3=1, x_4=1$  accettiamo o rifiutiamo l'ipotesi nulla? Perché? [2 punti]

--

b) Definire, basandosi sulla regione di rifiuto sopra scritta, l'errore di primo tipo [2 punti] e l'errore di secondo tipo. [1 punto]

--	--

c) Calcolare  $\alpha$ . [2 punti]

--

d) Se sulla base di una realizzazione campionaria accetto  $H_0$ , posso affermare che  $H_0$  è vera? Perché?

[2 punti]

--

6) Osservate il seguente output di excel, ottenuto dalla regressione lineare di Y su X1:

<i>Statistica della regressione</i>	
R multiplo	0,8448
R al quadrato	0,7137
R al quadrato corretto	0,7112
Errore standard	27,2427
Osservazioni	117

	<i>Coefficienti</i>	<i>Errore standard</i>	<i>Stat t</i>
Intercetta	348,45	12,49	27,89
Variabile X1	-5,87	0,35	-16,93

a) Scrivete l'equazione della retta di regressione lineare di Y su X1 [1 punto]

b) Sulla base della tabella accettereste l'ipotesi che  $\beta_1=0$ ? Scrivere il ragionamento seguito e il/i valore/i da cui si capisce. [2 punti]

c) Sulla base della tabella, ritenete che la relazione lineare tra Y e X1 sia forte o debole? Da quale valore si capisce? [1 punto]