

PROVA SCRITTA DI STATISTICA (COD 4038-5047-371-377)

26 Marzo 2004

MODALITÀ A

APPROSSIMARE TUTTI I CALCOLI ALLA QUARTA CIFRA DECIMALE

ESERCIZIO 1 (9 punti)

a)

Di seguito sono riportati alcuni dati relativi ai clienti di un'azienda che opera nel settore delle macchine utensili.

Per ciascuna azienda cliente sono riportati:

AZIENDA Codice azienda
 FATTURATO Fatturato in milioni di euro
 AREA Area Geografica (NO=Nord Ovest, NE=Nord Est, C=Centro, S=Sud e Isole)

NUMERO DIPENDENTI IN CLASSI

ANZIANITA' COME CLIENTI (in anni)

SETTORE DI APPARTENENZA (quattro modalità distinte S1, S2, S3, S4)

Azienda	FATTURATO	AREA	NUMERO DIPENDENTI IN CLASSI	ANZIANITA' COME CLIENTE (anni)	SETTORE	FATTURATO ²	ANZIANITA' COME CLIENTE ²
A	15	NO	[1-5)	5	S1	225	25
B	12	NE	[5-10)	6	S2	144	36
C	11	NO	[10-50]	12	S4	121	144
D	13	C	[10-50]	10	S1	169	100
E	9	NO	[5-10)	11	S2	81	121
F	11	C	[5-10)	9	S3	121	81
G	10	C	[5-10)	7	S4	100	49
H	7	S	[1-5)	11	S3	49	121
I	8	S	[1-5)	6	S1	64	36
L	20	S	[1-5)	12	S1	400	144
SOMMA	116			89		1474	857

a1) Fornire un'opportuna rappresentazione grafica del carattere NUMERO DIPENDENTI IN CLASSI.

a2) Calcolare media e mediana del FATTURATO.

a3) La distribuzione del FATTURATO è simmetrica? Giustificare la risposta.

a4) Il FATTURATO è regressivamente indipendente dal SETTORE? Giustificare la risposta.

b)

Si consideri la tabella seguente, che riporta la distribuzione bidimensionale di due variabili quantitative X e Y . Si dispongano le frequenze congiunte relative in modo che vi sia perfetta connessione ma non perfetta dipendenza lineare:

$X \backslash Y$	6	8	10	
0				
1				
2				
				1

ESERCIZIO 2 (6 punti)

a) Considerando 30 titoli azionari, solo 5 di questi hanno attualmente una quotazione superiore a 10 anni fa. Si estraggono a caso con reimmissione 6 titoli.

Qual è la probabilità che

- a1) Nessun titolo estratto abbia una quotazione superiore a 10 anni fa.
- a2) Almeno 2 titoli estratti abbiano una quotazione inferiore a 10 anni fa.

b) Si consideri una variabile aleatoria X con varianza $V(X)=3$. Si calcoli la varianza della variabile aleatoria $Y=-3X + 4$.

c) Si enunci il teorema di Bayes.

ESERCIZIO 3 (12 punti)

a) Si consideri una popolazione con distribuzione uniforme continua tra 2 e 6 e un campione bernoulliano di ampiezza 100 estratto da tale popolazione. Si calcoli la varianza dello stimatore media campionaria.

b)

b1) Si consideri una popolazione normale X con media m e con varianza nota e pari a 4. Si fornisca l'espressione dell'intervallo di confidenza per m con livello di confidenza fissato ad $\alpha=5\%$

b2) Si calcoli l'intervallo di confidenza corrispondente alla realizzazione campionaria (3,4 ; 4,5 ; 4,8 ; 3,8 ; 3,1).

b3) Quale deve essere la numerosità campionaria minima per garantire che la lunghezza dell'intervallo proposto al punto (b1) non superi in ampiezza 1,5?

c)

c1) Si consideri una popolazione bernoulliana X di parametro p . Ipotizzando n sufficientemente grande per utilizzare i risultati asintotici, si fornisca l'espressione della regione di rifiuto per provare:

$$H_0 : p \geq 0,75 \text{ contro } H_1 : p < 0,75$$

con una dimensione del test fissata al livello $\alpha=5\%$.

c2) Avendo estratto un campione bernoulliano di ampiezza $n=200$ e avendo riscontrato un numero di successi pari a 120, accettereste o rifiutereste l'ipotesi H_0 ?

c3) Si calcoli la probabilità di errore di seconda specie corrispondente a $p=0,65$.

d) Si definisca la proprietà di consistenza di uno stimatore in senso forte.

PROVA SCRITTA DI STATISTICA (COD 4038-5047-371-377)

26 Marzo 2004

MODALITÀ B

APPROSSIMARE TUTTI I CALCOLI ALLA QUARTA CIFRA DECIMALE

ESERCIZIO 1 (9 punti)

a)

Di seguito sono riportati alcuni dati relativi ai clienti di un'azienda che opera nel settore delle macchine utensili.

Per ciascuna azienda cliente sono riportati:

AZIENDA Codice azienda
 FATTURATO Fatturato in milioni di euro
 AREA Area Geografica (NO=Nord Ovest, NE=Nord Est, C=Centro, S=Sud e Isole)
 ANZIANITA' COME CLIENTI (in anni)
 SETTORE DI APPARTENENZA (quattro modalità distinte S1, S2, S3, S4)

Azienda	FATTURATO	AREA	ANZIANITA' COME CLIENTE (anni)	SETTORE	FATTURATO ²	ANZIANITA' COME CLIENTE ²	FATTURATO x ANZIANITA' COME CLIENTE
A	15	NO	5	S1	225	25	75
B	12	NE	6	S2	144	36	72
C	11	NO	12	S4	121	144	132
D	13	C	10	S1	169	100	130
E	11	NO	11	S2	121	121	121
F	9	C	9	S3	81	81	81
G	10	C	7	S4	100	49	70
H	7	S	11	S3	49	121	77
I	8	S	6	S1	64	36	48
L	20	S	12	S1	400	144	240
SOMMA	116		89		1474	857	1046

a1) Costruire la distribuzione di frequenza, calcolando le frequenze relative, della variabile SETTORE.

a2) Calcolare mediana e scarto quadratico medio di ANZIANITA' COME CLIENTE.

a3) La distribuzione di ANZIANITA' COME CLIENTE è obliqua destra? Giustificare la risposta.

a4) Le variabili FATTURATO e ANZIANITA' COME CLIENTE sono linearmente indipendenti? Giustificare la risposta.

b)

Si consideri la tabella seguente, che riporta la distribuzione bidimensionale di due variabili quantitative X e Y . Si dispongano le frequenze congiunte relative in modo che Y risulti regressivamente indipendente da X ma che le due variabili non siano statisticamente indipendenti.

$X \backslash Y$	6	8	10	
0				
1				
2				
				1

ESERCIZIO 2 (6 punti)

a) Considerando 20 titoli azionari, solo 4 di questi hanno attualmente una quotazione superiore a 10 anni fa. Si estraggono a caso con reimmissione 5 titoli.

Qual è la probabilità che:

- a1) Tutti i titoli estratti abbiano una quotazione superiore a 10 anni fa
- a2) Almeno 2 titoli estratti abbiano una quotazione inferiore a 10 anni fa

b) Si consideri una variabile aleatoria X con valore atteso $E(X)=4$ e una variabile aleatoria Y con valore atteso $E(Y)=6$. Si calcoli il valore atteso della variabile aleatoria $Z=4X - 6Y$.

c) Si enunci il teorema delle probabilità totali.

ESERCIZIO 3 (12 punti)

a) Si consideri una popolazione con distribuzione bernoulliana di parametro $p=0,6$. e un campione bernoulliano di ampiezza 100 estratto da tale popolazione. Si calcoli la probabilità che lo stimatore media campionaria assuma valori superiori a 0,4.

b)

b1) Si consideri una popolazione normale X con media m e con varianza non nota. Si fornisca l'espressione dell'intervallo di confidenza per m con livello di confidenza fissato ad $\alpha=1\%$

b2) Si calcoli l'intervallo corrispondente alla realizzazione campionaria (5 ; 4,5 ; 5,8 ; 6,4 ; 3,3 ; 6,6).

b3) Che cosa si intende per *livello di confidenza*?

c)

c1) Si considerino due popolazioni normali X_1 e X_2 rispettivamente con media μ_1 e μ_2 e con varianze non note, ma uguali. Si fornisca l'espressione della regione di rifiuto per provare:

$$H_0 : \mu_1 - \mu_2 = 0 \quad \text{contro} \quad H_1 : \mu_1 - \mu_2 \neq 0$$

con una dimensione del test fissata al livello $\alpha=10\%$

c2) Supponendo di aver estratto da X_1 e X_2 due campioni rispettivamente di ampiezza $n_1=10$ e $n_2=15$, con media pari a 20,1 e 24,6 e varianze campionarie corrette $S_1^2=30$ e $S_2^2=35$, si dica, giustificando la risposta, se accettereste o rifiutereste l'ipotesi nulla.

d) Si enunci il teorema centrale del limite.