

STATISTICA CORSO BASE. Prova scritta del 10-1-2012. Tempo: 2 ore - Compito A

Cognome e Nome: Matricola:

Attenzione: Prima di affrontare la prova si consiglia di leggere attentamente le note che seguono.

Indicare cognome e nome (in stampatello e senza cancellature)

Verrà corretto solo l'elaborato contenuto in questi fogli. Alla consegna, allegare comunque la brutta copia.

Non è consentita la consultazione di libri o appunti personali ma solo del formulario fornito dal docente

Tenere spenti i telefoni cellulari. Il loro utilizzo non è consentito pena l'esclusione dalla prova.

Per ogni domanda a risposta multipla sono previste tre possibili risposte di cui una sola è corretta. Per fornire la risposta ad una domanda barrare la casella corrispondente alla risposta prescelta. Nel caso di correzioni, fare in modo che sia chiaro quale sia la risposta finale. **Il punteggio assegnato ad ogni quesito è :**

1 se si indica la risposta corretta, -0.5 se si indica la risposta errata, 0 se non si fornisce risposta.

Se un candidato ritiene che ci sia un errore ed in particolare che nessuna delle tre alternative disponibili risponda al quesito posto, può aggiungere la risposta corretta opportunamente motivata.

Nelle altre domande, quelle aperte, fornire, oltre al risultato, anche la spiegazione della procedura adottata.

Esercizio 1. 8 punti

- La tabella seguente contiene la distribuzione di frequenza in classi del costo in euro dei libri economici stampati in Italia nel 2009 (*La produzione libraria, Istat*):

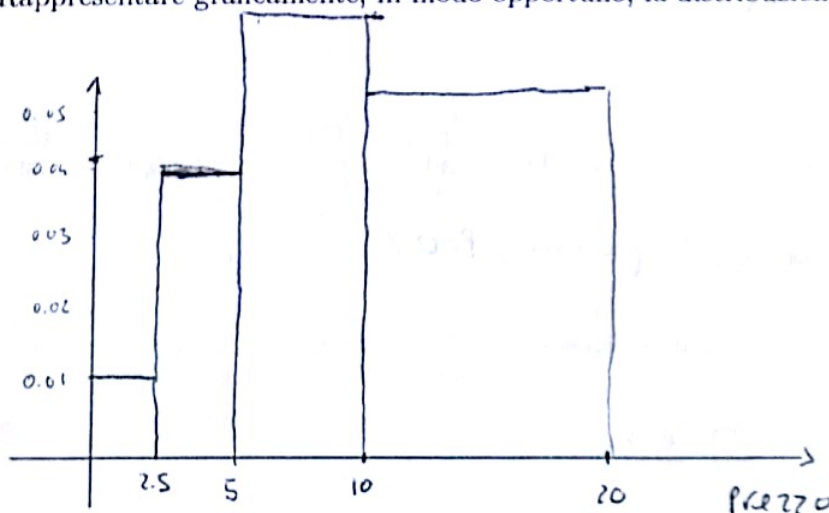
classi di prezzo (in euro)	frequenze relative
(0, 2.5]	0.026
(2.5, 5]	0.095
(5, 10]	0.347
(10, 20]	0.532
totale	1

- ★ Ciascun libro rappresenta
 - una modalità
 - una frequenza
 - una unità statistica

- ★ Il valore "0.026" nella tabella è
 - una modalità
 - una frequenza
 - una unità statistica

- ★ Il valore "20" nella tabella è
 - una modalità
 - una frequenza
 - una unità statistica

- ★ Rappresentare graficamente, in modo opportuno, la distribuzione.



★ Qual è la classe modale?

(2.5, 5]

(5, 10]

(10, 20]

★ Qual è il costo medio dei libri economici stampati in Italia nel 2009?

$$\bar{x} = 10.972$$

★ Quale tra i seguenti ritieni sia il valore della varianza della distribuzione (in euro²)?

176

0

-176

★ In quale classe ritieni possa cadere la mediana della distribuzione?

(2.5, 5]

(5, 10]

(10, 20]

Esercizio 2. 7 punti

- Il 40% dei clienti di una banca on line ha un Conto Deposito. L'accredito dello stipendio sul conto è stato attivato dal 70% dei clienti titolari di un Conto Deposito e dal 20% dei titolari di un Conto Corrente. Ogni cliente della banca può accedere ad una sola delle due tipologie di conto. Le cinque domande che seguono sono riferite a questa situazione.

2 punti: ★ Se scegliamo una persona a caso tra i clienti della banca, qual è la probabilità che abbia l'accredito dello stipendio sul conto?

RISPOSTA: 0.28

★ Se il cliente scelto a caso ha l'accredito dello stipendio sul conto, qual è la probabilità che non sia titolare di un Conto Deposito?

RISPOSTA: 0.3

★ Gli eventi {Cliente con Conto Corrente} e {Cliente con accredito stipendio in banca} sono

incompatibili

dipendenti

indipendenti

★ Consideriamo 10 clienti tra coloro che non sono titolari di un Conto Deposito ed indichiamo con X il numero di quelli che hanno l'accredito dello stipendio sul conto. Qual è la distribuzione della variabile casuale X ?

$$X \sim \text{BINOMIALE } (n=10, p=0.2)$$

★ Qual è la probabilità che al più uno dei 10 clienti abbia l'accredito dello stipendio sul conto?

$$P(X \leq 1) = 0.38$$

Esercizio 3. 3 punti

- ★ Si è appena conclusa una campagna pubblicitaria per incentivare l'accredito dello stipendio sul conto tra i clienti della banca titolari di un conto corrente. Per verificare l'efficacia della campagna pubblicitaria, i responsabili del servizio clienti decidono di confrontare le due ipotesi:

$$H_0 : p = 0.2 \qquad H_a : p > 0.2$$

in cui p indica la proporzione di clienti con accredito dello stipendio tra tutti i titolari di conto corrente. Si scelgono a caso $n = 100$ clienti titolari di conto corrente: 25 di loro hanno attivato l'accredito dello stipendio sul conto. Quale risulta il valore della statistica test?

- 1.25 0.25 0.2

- ★ Con riferimento al punto precedente, quanto vale il *valore P*?

- 0.5987 0.1056 0.8944

- ★ Sulla base delle informazioni raccolte, cosa concludono i responsabili del servizio clienti della banca al livello di significatività del 5%?

- non c'è evidenza sufficiente per rifiutare l'ipotesi nulla
 i dati forniscono evidenza sufficiente per il rifiuto dell'ipotesi nulla
 le informazioni a disposizione non sono sufficienti a prendere una decisione

Esercizio 4. 6 punti

- Una ditta produce pacchetti di patatine da 100 grammi. Il responsabile del controllo di qualità deve valutare se il contenuto medio μ delle confezioni corrisponde a quanto dichiarato sull'etichetta o se risulti diverso, a causa di malfunzionamenti durante l'impacchettamento. Viene scelto a caso un campione di $n = 9$ confezioni di patatine. A questa situazione si riferiscono le seguenti cinque domande

- (a) Formulare il problema come una verifica di ipotesi, specificando H_0 e H_a .

$$H_0 : \mu = 100 \qquad H_a : \mu \neq 100$$

- (b) Sulla base delle esperienze passate, si sa che la variabile casuale "peso delle patatine in ciascuna confezione" ha distribuzione normale con varianza 36 gr^2 . I dati osservati forniscono una media campionaria pari a 104 gr. Ricavare un intervallo di confidenza al 99% per μ .

$$(98.84 ; 109.16)$$

- (c) Se fissiamo il livello di confidenza per μ al 95%, l'ampiezza dell'intervallo aumenta o diminuisce?

DIMINUISCE

- 2 punti (d) Riprendendo il problema del punto (a), cosa concludiamo al livello di significatività $\alpha = 0.01$?

NON RIFIUTIAMO H_0

- (e) Sempre al livello di significatività $\alpha = 0.01$, avremmo raggiunto le stesse conclusioni del punto precedente se l'inferenza fosse stata basata su un campione di numerosità $n = 4$ e che fornisce lo stesso valore della media campionaria?

SI.

Esercizio 5. 9 punti

- Sulla base di un campione di $n = 20$ lavoratori a tempo pieno laureati di età compresa tra 25 e 65 anni è stata ricavata la seguente retta dei minimi quadrati che spiega la retribuzione media settimanale y (in euro) in funzione dell'età x (in anni):

$$\hat{y} = a + bx$$

con $a = 696.7$ e $b = 9.6$. Inoltre è stato ricavato che il coefficiente di correlazione tra x e y risulta pari a $r = 0.15$. A questi dati si riferiscono gli otto quesiti che seguono.

- ★ Quale percentuale della variabilità di y è spiegata dal modello di regressione lineare?

2.3%

- ★ Sapendo che un euro vale 0.7 dollari statunitensi, quale sarebbe stato il valore del coefficiente di correlazione se i salari fossero stati misurati in dollari?

✓ rimane INVARIATO

- ★ Sulla base del modello stimato, a quanto ci aspettiamo che ammonti il salario settimanale di un dipendente di 30 anni d'età?

986.7 euro

- ★ Sempre sulla base del modello stimato, a quanto ci aspettiamo che ammonti l'incremento della retribuzione settimanale (in euro) se aumentiamo di uno l'età (in anni)?

9.6

- ★ Ritieni che il modello stimato possa fornire una previsione affidabile del salario settimanale di una persona di 90 anni, che continua a lavorare? Giustificare la risposta.

NO

- 0.5 punti
- ★ Sulla base dei dati disponibili, vogliamo confrontare le seguenti ipotesi riguardanti il valore del coefficiente angolare del modello di regressione

$$H_0 : \beta = 0$$

$$H_a : \beta > 0$$

Sapendo che l'errore standard di b , la stima del coefficiente angolare, è pari a $ES_b = 624.1$, qual è il valore della statistica test?

65

9.6

0.02

- 1.5 punti
- ★ Quale tra i seguenti è il valore P corrispondente al valore della statistica test calcolata al punto precedente?

0.49

0.5

1

- ★ Con riferimento al test di significatività su β , cosa concludiamo ad un livello $\alpha = 0.05$?

NON POSSIAMO RIFIUTARE H_0