

**CORSO DI STATISTICA DI BASE (Prof. GIORGIO ALLEVA)**

Anno Accademico 2019-2020

Prova scritta dell'8 gennaio 2020

A

**ESERCIZI (9 PUNTI ciascuno)**

**Esercizio 1.** Si disponga della seguente distribuzione di frequenza delle variabili X e Y, osservate su un campione di 38 unità.

| X\Y | 1  | 3 | 8  | Tot |
|-----|----|---|----|-----|
| 1   | 2  | 4 | 6  | 12  |
| 5   | -  | 2 | 6  | 8   |
| 11  | 10 | - | 8  | 18  |
| Tot | 12 | 6 | 20 | 38  |

- Si misuri l'eterogeneità di X e si valuti se questa sia o meno elevata;
- si determini l'intervallo che contiene l'80% dei dati campionari centrali di X;
- si rappresenti la concentrazione di X attraverso il diagramma di Lorenz;
- si confronti la variabilità di X e di  $Z=2X$  attraverso: d1) la varianza; d2) il coefficiente di variazione.

**Esercizio 2.** Sulla base dei dati campionari contenuti nella precedente tabella, una volta specificate le assunzioni necessarie:

- si determini l'intervallo di confidenza della media di X per  $\alpha = 0,02$  e si esponga il suo significato;
- si indichi come varierebbe l'ampiezza di tale intervallo qualora il campione fosse di 61 unità (a parità delle altre caratteristiche);
- si indichi sotto quali condizioni lo stimatore della media sia consistente;
- si verifichi attraverso un test se la varianza di X possa essere considerata nella popolazione uguale a 15 per  $\alpha = 0,02$ ;
- si indichi cosa si intenda per livello di significatività  $\alpha$  e per potenza del test.

**Esercizio 3.** Da una rilevazione di una popolazione si conoscano le informazioni seguenti:

$$\mu(X) = 12 \quad \mu(Y) = 15 \quad \mu(X^2) = 225 \quad \mu(Y^2) = 369 \quad \mu(XY) = 85$$

- Determinare l'equazione della retta di regressione di Y su X e indicare il significato dei parametri;
- rappresentare graficamente la retta e stimare Y per  $X = 18$ ;
- misurare la bontà dell'adattamento della retta di regressione e indicare in che intervallo di valori sia compresa in generale tale misura;
- come varia il precedente intervallo qualora si conosca che  $\eta^2_{Y|X} = 0,85$  e  $\eta^2_{X|Y} = 0,9$  ?
- aggiungendo nel modello una seconda variabile esplicativa Z si indichi l'equazione del piano di regressione e l'espressione della devianza residua di Y;
- in quale caso la scelta della variabile Z comporta che il piano di regressione sia indeterminato?

**QUESITI (barrare la risposta ritenuta esatta)**

**(PUNTI 2 per risposta corretta, PUNTI -1 per risposta sbagliata, PUNTI 0 per assenza di risposta)**

Uno stimatore è corretto se:

- $E(\Theta) = T$
- $E(T) = \Theta$
- $MSE(T) = 0$

La curva di regressione di Y su X (o interpolatrice ottima):

- passa per le medie condizionate di X
- la residua di Y è la devianza esterna
- nessuna delle precedenti

L'indipendenza assoluta implica:

- l'indipendenza in media
- l'indipendenza lineare
- entrambe

L'indice dei prezzi di Laspeyres al tempo 1 in base 0 è una media aritmetica degli indici elementari dei prezzi con pesi:

- $p_0 q_0$
- $p_0 q_1$
- nessuno dei precedenti

La v.c. di Poisson con parametro  $\lambda$ :

- ha media e varianza entrambe uguali a  $\lambda$
- è una v.c. discreta con  $0 \leq X \leq \infty$
- entrambe

**CORSO DI STATISTICA DI BASE (Prof. GIORGIO ALLEVA)**

**Anno Accademico 2019-2020**

**Prova scritta dell'8 gennaio 2020**

**B**

**ESERCIZI (9 PUNTI ciascuno)**

**Esercizio 1.** Si disponga della seguente distribuzione di frequenza delle due variabili X e Y, osservate su un campione di 38 unità.

| X\Y | 1  | 3 | 8  | Tot |
|-----|----|---|----|-----|
| 1   | 2  | 4 | 6  | 12  |
| 5   | -  | 2 | 6  | 8   |
| 11  | 10 | - | 8  | 18  |
| Tot | 12 | 6 | 20 | 38  |

- Si determinino i quartili di Y e si misurino, sulla base di questi, la variabilità e l'asimmetria della variabile Y;
- si calcoli il coefficiente di variazione di Y;
- si rappresenti la concentrazione di Y attraverso il diagramma di Lorenz Y.

**Esercizio 2.** Sulla base dei dati campionari contenuti nella precedente tabella, una volta specificate le assunzioni necessarie:

- si determini l'intervallo di confidenza della varianza di Y per  $\alpha = 0,02$  e si esponga il suo significato;
- si verifichi attraverso un test se la media di Y possa essere considerata maggiore di 4 con un livello di significatività dell'1%;
- si mostri se la decisione sarebbe la stessa qualora si conoscesse che la varianza della popolazione fosse pari a 5;
- indicare cosa si intenda per livello di significatività  $\alpha$  e per potenza del test;
- Indicare sotto quale condizioni lo stimatore della media sia corretto e pienamente efficiente.

**Esercizio 3.** Da una rilevazione condotta su una popolazione di 120 unità si conoscano le informazioni seguenti:

$$\mu(X) = 8 \quad \mu(Y) = 10 \quad \mu(X^2) = 80 \quad \mu(Y^2) = 200 \quad \mu(XY) = 44$$

- Determinare l'equazione della retta di regressione di Y su X e indicare il significato dei parametri;
- rappresentare graficamente la retta e stimare Y per  $X = 12$ ;
- misurare la devianza spiegata e la devianza residua di Y dalla retta di regressione, valutando la bontà di adattamento della retta di regressione;
- aggiungendo nel modello una seconda variabile esplicativa Z si indichi l'equazione del piano di regressione e l'espressione della devianza residua di Y;
- in quale caso la scelta della variabile Z comporta che il piano di regressione sia indeterminato?

**QUESITI (barrare la risposta ritenuta esatta)**

**(PUNTI 2 per risposta corretta, PUNTI -1 per risposta sbagliata, PUNTI 0 per assenza di risposta)**

Uno stimatore  $T_1$  è più efficiente di uno stimatore  $T_2$  se:

- $\sigma^2(T_1)$  tende a zero al crescere della dimensione del campione
- $\sigma^2(T_1) < \sigma^2(T_2)$
- $MSE(T_1) < MSE(T_2)$

La curva di regressione di Y su (o interpolatrice ottima):

- passa per le medie condizionate di X
- rende minima la devianza esterna
- nessuna delle precedenti

$\chi^2 = 0$  implica:

- $\eta^2_{X|Y} = 0$
- $r^2 = 0$
- $\eta^2_{X|Y} = r^2 = 0$

L'indice dei prezzi di Paasche al tempo 1 in base 0 è una media aritmetica degli indici elementari delle quantità con pesi:

- $p_0 q_0$
- $p_1 q_1$
- nessuno dei precedenti

La v.c. binomiale con parametro p:

- ha media pari a np e varianza pari a npq
- è una v.c. discreta con  $0 \leq X \leq n$
- entrambe