

SCUOLA DI DOTTORATO DI RICERCA
SCIENZE STATISTICHE
Università degli Studi di Padova
Esame di ammissione al XXI ciclo
Prova scritta (10/11/2005)

Il candidato svolga, con il dettaglio appropriato, sia la parte di base della traccia, Statistica di Base, sia la parte più specializzata della traccia, per la quale sceglierà uno dei quattro quesiti corrispondenti, rispettivamente, agli ambiti disciplinari Statistica, Statistica Economica, Demografia, Statistica Sociale.

Tema n.1

Prima parte - Statistica di base

Valore atteso di una variabile casuale univariata: definizione, proprietà (comprese alcune ben note disequaglianze), problemi di inferenza su un valore atteso.

Seconda Parte

Statistica

Dopo aver svolto l'esercizio seguente, si fornisca per ciascun punto un breve inquadramento dei concetti e dei risultati pertinenti nella teoria dell'inferenza.

Sia $y = (y_1, \dots, y_n)$ realizzazione di una variabile casuale multivariata con componenti indipendenti $Y_i \sim N(\mu + c_i, \sigma_0^2)$, dove c_i ($i = 1, \dots, n$) e σ_0^2 sono costanti note, e $\mu \in \mathbb{R}$ è un parametro ignoto.

- (a) Si stabilisca se il modello statistico costituisce una famiglia esponenziale.
- (b) Di ciascuna delle seguenti statistiche si stabilisca se è sufficiente o sufficiente minimale

$$\sum_{i=1}^n y_i, \quad \sum_{i=1}^n (y_i - c_i), \quad \sum_{i=1}^n (y_i - c_i)^2, \quad \left(\sum_{i=1}^n y_i, \sum_{i=1}^n y_i^2 \right).$$

- (c) Si dica se per stimare μ è preferibile lo stimatore di massima verosimiglianza o lo stimatore

$$\tilde{\mu} = \frac{\sum_{i=1}^n i(y_i - c_i)}{\sum_{i=1}^n i}.$$

- (d) Si proponga un test con livello di significatività α per $H_0 : \mu = 0$ contro $H_1 : \mu < 0$.

Statistica economica

Il concetto di tendenza nell'analisi dei fenomeni economici. Il candidato illustri varie definizioni presenti in letteratura discutendo i problemi di identificazione e stima che sono ad esse collegati.

Demografia

Il censimento della popolazione è uno strumento ancora utile o obsoleto? Popolazione di censimento e popolazione a calcolo.

Statistica Sociale

Si intende studiare il fenomeno della conciliazione tra tempi di lavoro e tempi della famiglia in una determinata provincia italiana. Si discutano i seguenti aspetti:

- possibili obiettivi della ricerca;
- il disegno d'indagine conseguente;
- quali strumenti di raccolta di dati potrebbero essere utilizzati.

Tema n.2

Prima parte - Statistica di base

Leggi dei grandi numeri e teorema centrale del limite: enunciati (con commenti pertinenti), applicazioni nell'inferenza statistica.

Seconda Parte

Statistica

Dopo aver svolto l'esercizio seguente, si fornisca per ciascun punto un breve inquadramento dei concetti e dei risultati pertinenti nella teoria dell'inferenza.

Sia $y = (y_1, \dots, y_n)$ realizzazione della variabile casuale multivariata $Y = (Y_1, \dots, Y_n)$ con componenti indipendenti e identicamente distribuite in cui la funzione di ripartizione di Y_1 è

$$F_{Y_1}(y_1; \mu) = 1 - \exp\{-\exp\{y_1 - \mu\}\},$$

dove $\mu \in \mathbb{R}$ è ignoto.

1. Si scriva la funzione di log-verosimiglianza $l(\mu)$.
2. Si ottenga lo stimatore di massima verosimiglianza $\hat{\mu}$.
3. Si calcoli l'informazione osservata.
4. Si proponga un'approssimazione asintotica per la distribuzione di $\hat{\mu}$ quando μ è il vero valore del parametro.

Statistica economica

Il candidato illustri le caratteristiche di un processo ARIMA, soffermandosi sulle procedure diagnostiche a disposizione per la sua identificazione a partire da una serie storica osservata.

Demografia

Dopo decenni di diminuzione, a partire dal 1994 le nascite nell'Italia del Centro-Nord hanno ripreso ad aumentare. Il candidato valuti le cause plausibili di questa inversione di tendenza, illustrando anche le possibili modalità di verifica delle ipotesi stesse, non dimenticando le fonti necessarie per condurre tale verifica.

Statistica Sociale

Per meglio organizzare i servizi sociali dedicati agli anziani di un determinato territorio, gli organismi responsabili vogliono proporre un'analisi dei bisogni. Si descrivano quali possano essere gli strumenti conoscitivi da mettere in campo per svolgere tale attività, discutendone vantaggi e limiti.

Tema n.3

Prima parte - Statistica di base

Leggi normali univariate e multivariate: definizioni e principali proprietà, leggi di trasformazioni notevoli, applicazioni nell'inferenza statistica

Seconda Parte

Statistica

Dopo aver svolto l'esercizio seguente, si fornisca per ciascun punto un breve inquadramento dei concetti e dei risultati pertinenti nella teoria dell'inferenza.

Sia $y = (y_1, \dots, y_n)$ realizzazione della variabile casuale multivariata $Y = (Y_1, \dots, Y_n)$ con componenti indipendenti e $Y_i \sim N(\theta x_i, 1 + x_i^2)$, dove i valori x_i ($i = 1, \dots, n$) sono costanti note, non tutte nulle e tali che

$$\sum_{i=1}^n \frac{x_i^2}{1 + x_i^2} = 1,$$

e $\theta \in \mathbb{R}$ è ignoto. Si consideri il problema di verifica d'ipotesi

$$H_0 : \theta = 0 \quad \text{contr} \quad H_1 : \theta = 1.$$

Sia, per $c \in \mathbb{R}$,

$$R_c = \left\{ y \in \mathbb{R}^n : \sum_{i=1}^n \frac{x_i y_i}{1 + x_i^2} > c \right\}.$$

1. Si mostri che una regione critica del tipo R_c minimizza la probabilità di errore del secondo tipo per una fissata probabilità di errore del primo tipo.
2. Si determinino la distribuzione nulla e la distribuzione alternativa della statistica

$$t(y) = \sum_{i=1}^n \frac{x_i y_i}{1 + x_i^2}.$$

3. Si determinino il livello di significatività e la potenza del test con regione di rifiuto R_2 .
4. Si verifichi se il test con regione di rifiuto R_c è non distorto.

Statistica economica

La presenza di cambiamento strutturale nell'analisi dell'evoluzione dei fenomeni economici: il candidato ne dia esempi con riferimento alle sue esperienze di studio e di lavoro, ne illustri le conseguenze sulla stima dei parametri di un modello di regressione, e suggerisca metodi opportuni di diagnostica.

Demografia

Ormai da 20 anni, in Italia come in tutto l'Occidente, l'età media alla morte continua ad aumentare, grazie soprattutto ai guadagni di sopravvivenza in età anziana. Il candidato illustri le conseguenze demografiche e sociali di questo fenomeno.

Statistica Sociale

In molte indagini si raccolgono informazioni utilizzando diversi quesiti. Spesso il problema è quello della sintesi. Si inquadri il problema dell'analisi di dati multivariati dal punto di vista sostanziale (della ricerca sociale) e metodologico, soffermandosi sui principali strumenti statistici che il ricercatore ha a disposizione esplicitando limiti e condizioni di applicazione.