

Nell'epoca dei BIG DATA, la conoscenza degli strumenti statistici per l'analisi dei dati è diventata fondamentale. L'insegnamento si propone di introdurre gli studenti alle prime nozioni di analisi descrittive e inferenziali. Per la descrizione dettagliata degli obiettivi formativi visitare la pagina web:

<https://www.unipg.it/didattica/corsi-di-laurea-e-laurea-magistrale/archivio/offerta-formativa-2022-23?idcorso=170&annoregolamento=2021&idins=208496>

L'insegnamento è suddiviso in due parti, denominate Statistica Descrittiva e Statistica Inferenziale

PROGRAMMA DI STATISTICA DESCRITTIVA

Nozioni introduttive: cenni storici sullo sviluppo della statistica; terminologia essenziale; misurazione dei caratteri; genesi dei dati statistici; raccolta dei dati; matrice dei dati; rapporti statistici.

Distribuzioni statistiche: distribuzioni statistiche disaggregate; distribuzioni di frequenza; frequenze relative; frequenze cumulate; distribuzioni di frequenza con dati raggruppati in classi; classi reali; densità di frequenza; uniforme distribuzione delle unità nelle classi; distribuzioni doppie e multiple; distribuzioni di quantità; serie storiche; serie territoriali.

Rappresentazioni grafiche: grafici per distribuzioni relative a caratteri quantitativi: diagramma ad aste; istogramma di frequenza; rappresentazione grafica delle serie sconnesse: grafici a nastri; grafici a settori circolari; grafici tridimensionali; rappresentazione grafica delle serie storiche e delle serie territoriali; il problema della scala.

Medie: media aritmetica; media geometrica; media quadratica; il caso delle distribuzioni di frequenza; il caso di dati raggruppati in classi; medie ponderate; mediana; quartili e quantili; valore centrale; moda.

Variabilità e concentrazione: il fenomeno della variabilità; scostamento semplice medio; deviazione standard; formula alternativa della deviazione standard; campo di variazione; differenza interquartile; indici di variabilità percentuali; concentrazione; indici di concentrazione G e R; interpretazione geometrica degli indici di concentrazione.

Indici di asimmetria: simmetria e asimmetria; indici di asimmetria α_1 e α_2 .

Uno sguardo d'insieme alle costanti caratteristiche: grafici e costanti caratteristiche (lo studio del paragrafo è utile per migliorare la comprensione delle costanti caratteristiche); diagramma a scatola.

Numeri indici: numeri indici a base fissa e a base mobile; variazioni percentuali medie; numero indice di Laspeyres.

Analisi della dipendenza: distribuzioni doppie disaggregate e di frequenza; distribuzioni marginali e distribuzioni condizionate; rappresentazioni grafiche delle distribuzioni doppie; dipendenza statistica; indice chi-quadro per la misurazione della dipendenza.

Analisi di regressione: relazioni statistiche; regressione lineare semplice; il metodo dei minimi quadrati per la determinazione dei parametri della retta di regressione; l'adattamento ai dati della retta di regressione; indice r^2 e sue proprietà. Il caso delle serie storiche. Errore medio di predizione.

Correlazione: nozione di correlazione; coefficiente di correlazione di Bravais e relative proprietà.

Testo consigliato

G. Cicchitelli, P. D'Urso e M. Minozzo, *Statistica - Principi e metodi (terza edizione)*, Pearson, Milano, 2017.

PROGRAMMA DI STATISTICA INFERENZIALE

Probabilità: esperimenti casuali; spazio campionario ed eventi; operazioni su insiemi; probabilità; interpretazione della probabilità; calcolo delle probabilità; probabilità condizionata; indipendenza, formula di Bayes.

Variabili casuali: variabili casuali discrete; media e deviazione standard; variabili casuali standardizzate; variabili casuali continue; media e deviazione standard; quantili; variabili casuali doppie discrete; funzione di probabilità congiunta e funzioni di probabilità marginali; covarianza; variabili casuali discrete indipendenti; variabili casuali doppie continue (cenni); variabili casuali multiple (cenni); valore atteso e varianza di combinazioni lineari di variabili casuali.

Alcuni particolari modelli probabilistici: distribuzione di Bernoulli; distribuzione binomiale; distribuzione di Poisson; distribuzione normale; normale standardizzata e relative tavole; approssimazione della distribuzione binomiale con la normale; distribuzione chi-quadrato.

Distribuzioni campionarie: campione casuale; parametro; inferenza statistica: stima dei parametri e verifica delle ipotesi; statistiche campionarie; distribuzione campionaria della media per popolazioni generatrici normali e per grandi campioni (teorema del limite centrale); distribuzione campionaria della varianza; distribuzione campionaria della media quando la varianza della popolazione non è nota; distribuzione t di Student e relative tavole.

Stima puntuale dei parametri: stimatore; proprietà degli stimatori; non distorsione; errore quadratico medio; proprietà asintotiche.

Stima per intervallo: stimatore per intervallo e stima per intervallo; stima per intervallo della media di popolazioni generatrici normali; ampiezza dell'intervallo di confidenza; il caso in cui la varianza non è nota; stima per intervallo della media nel caso di grandi campioni (incluso il caso di una popolazione Bernoulliana); stima per intervallo della varianza di una popolazione normale.

Verifica delle ipotesi: ipotesi statistiche; verifica di ipotesi sulla media di una popolazione normale; il test Z ; il p -value o livello di significatività osservato; il test t ; verifica di ipotesi sulla media nel caso di grandi campioni (incluso il caso di una popolazione Bernoulliana); verifica di ipotesi sul parametro p di una popolazione bernoulliana; verifica di ipotesi sulla varianza di una popolazione normale; criteri di ottimizzazione nella verifica delle ipotesi (solo cenni); errori di prima e di seconda specie e relative probabilità.

Verifica dell'ipotesi di indipendenza mediante il test chi-quadrato.

Inferenza per il modello di regressione semplice: Inferenza per il modello di regressione lineare: I residui come variabili casuali. Distribuzione degli stimatori ottenuti con il metodo dei minimi quadrati in caso di normalità e per grandi campioni: intervallo di confidenza. Il test di significatività e il p -value.

Testo consigliato

G. Cicchitelli, P. D'Urso e M. Minozzo, *Statistica - Principi e metodi (terza edizione)*, Pearson, Milano, 2017.