

Domande poste alla professoressa M. G. Ottaviani

I Domanda

Rapporto probabilità statistica: cosa differenzia la statistica dalla probabilità? Cosa le lega?

Come gestire al meglio gli strumenti statistici?

Rapporto probabilità statistica: cosa differenzia la statistica dalla probabilità? Cosa le lega?

Probabilità e statistica sono due discipline diverse, nate in tempi diversi per dare risposta a problemi diversi.

Il calcolo delle probabilità, che si può definire con Quetelet “Le calcul qui enseigne à trouver le degré de probabilité d’un événement” (A. Quetelet, *Instructions populaires sur le calcul des probabilités*, Tarlier et Hayez, Bruxelles, MDCCCXXVIII, p. 4), nasce legato ai giochi di sorte e viene successivamente applicato anche alle scienze sperimentali. Nel tempo si sviluppa secondo impostazioni diverse: classica (P.S. Laplace, 1812), frequentista (R.von Mises 1920), logica, soggettivista, assiomatica. E’ una branca della matematica.

“La statistica è il mezzo che consente la conoscenza quantitativa dei fenomeni collettivi, ossia la tecnica appropriata o il metodo appropriato per indagare tali fenomeni” (G. Leti, L. Cerbara, *Elementi di statistica descrittiva*, il Mulino, Bologna, 2009, p. 16).

Nasce nel 1660 all’Università di Helmstedt, dove Hermann Conring, nel suo corso di lezioni , comincia ad esaminare e descrivere le cose notevoli di uno Stato anche se senza numeri né tabelle (Statistica descrittiva tedesca). In quello stesso periodo in Inghilterra si fa largo una corrente di pensiero secondo la quale i fenomeni demografici e sociali devono essere studiati sulla base di dati numerici tramite la matematica (Aritmetica politica). A partire dal 1790 i due filoni si fondono, assumendo la denominazione di “Statistica”, dall’italiano “ragione di stato”.

I fenomeni collettivi si conoscono attraverso una pluralità di osservazioni. Se le osservazioni riguardano tutta la collettività di individui che forma il collettivo di interesse (popolazione) si opera nell’ambito della statistica descrittiva. Se le osservazioni riguardano solo una parte della popolazione (campione) con l’intento di risalire induttivamente alla conoscenza della popolazione ignota dalla quale il campione si presume estratto, si opera nell’ambito della statistica inferenziale.

La statistica inferenziale richiede l’utilizzo delle distribuzioni teoriche di probabilità (in particolare: la distribuzione binomiale e la distribuzione normale).

La distribuzione teorica degli errori accidentali (o distribuzione normale o legge di Gauss) fornisce alla statistica un utile modello teorico per interpolare con una unica funzione analitica un collettivo

di misure affette da errori e più in generale una distribuzione empirica di un carattere quantitativo rappresentabile graficamente con una distribuzione campanulare simmetrica.

L'insegnamento di statistica e probabilità, insieme a scuola nella disciplina matematica, richiede di saper rispettare la peculiarità delle due discipline nella consapevolezza che, pur nella loro diversità, concorrono insieme ad affrontare l'ambito dell'incertezza. (M. G.Ottaviani, Statistica e matematica a scuola: due discipline e un solo insegnamento. Confronto culturale e opportunità interdisciplinare, Induzioni, vol 36)

Come gestire al meglio gli strumenti statistici?

Per gestire gli strumenti statistici ai quali fanno riferimento le Indicazioni ministeriali delle scuole dei diversi ordini è fondamentale comprendere alcuni concetti base della statistica (cfr. file allegati), in particolare la classificazione dei caratteri, ossia la classificazione del tipo di informazione raccolta sulle unità, poiché essa è la guida alle elaborazioni possibili sui dati. In particolare la rappresentazione grafica dei dati e le sintesi da utilizzare dipendono strettamente dal tipo di carattere.

Classificazione dei caratteri

Carattere			
qualitativo		quantitativo	
a modalità non ordinate o sconnesso	a modalità ordinate	discreto	continuo

Queste considerazioni consentono di dare risposta anche alla II Domanda

II Domanda

Premessa: si è costruito un questionario di gradimento distribuito su un campione di 1000 individui.

Per ogni domanda è prevista una risposta chiusa con 4 possibili scelte: A, B, C, D (A minor gradimento, D maggior gradimento) Tabulo i dati e trovo le percentuali relative ad ogni risposta.

Domanda : *quale indice di sintesi è opportuno utilizzare: per ogni domanda oppure per gruppi di domande?*

Ogni domanda fornisce una distribuzione statistica delle 1000 risposte classificate rispetto alle modalità A, B, C, D del carattere gradimento. Queste modalità sono ordinate, ciò significa

operativamente che per ogni domanda è possibile individuare la moda e la mediana, il primo e il terzo quartile della distribuzione delle risposte . Graficamente si può costruire un box plot, a patto che si ipotizzi che le scelte A, B, C, D formano una scala dove i gradini hanno la stessa altezza, il che significa che A, B, C, D sono 4 punti ordinati su un segmento orizzontale e sono equidistanziati.

Può essere utile creare un grafico, sovrapponendo i box-plot di ciascuna domanda, per rendere più agevole la discussione sugli esiti della rilevazione . Se le domande sono molte e di natura diversa, per facilitare l'interpretazione dei risultati, si possono raggruppare i box-plot delle domande che hanno la stessa natura.

Risorse:

Per la scuola primaria: file <<statistica- a- scuola-basi-concettuali-mg-ottaviani PRIMARIA>

da risorse della proposta “L'animale preferito”

Per la scuola secondaria

file: <concetti_chiave AAA>, <modaB1>, <medianaB2>, <media aritmetica>, <altri valori mediB3>, <proprietà_ mediaB4>, <misure_ di_ variabilità C>

Dalle proposte “Come ci alimentiamo” e “Di media non ce n'è una solai”