



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Corsi di teologia

**METODOLOGIE STATISTICHE UTILIZZATE PER
L'ANALISI DELL'OPINIONE DEGLI STUDENTI**

A.A. 2020-2021

NUCLEO DI VALUTAZIONE DI ATENEIO
MILANO • PIACENZA - CREMONA • ROMA • BRESCIA

APRILE 2022

Indice

INTRODUZIONE pag. 3

- I)** Nota al § 3.1.1 e al § 4.1.1 Tabella 4 “*Giudizio complessivo positivo dell’insegnamento*” ripartito per raggruppamenti di studenti frequentanti e per anno di corso pag. 4
- II)** Nota al § 3.1.2 e al § 4.1.2 Tabella 5 “*Distribuzione di frequenza percentuale delle domande più rilevanti*” pag. 4
- III)** Nota al § 3.2.1 e al § 4.2.1 Tabella 9 “*Confronto tra la soddisfazione complessiva per gli insegnamenti e l’interesse agli argomenti indicati nel programma negli aa.aa. 2019/20 e 2020/21*” pag. 4
- IV)** Nota al § 3.2.1 e al § 4.2.1 Tabella 10 “*Correlazione di Spearman in ordine decrescente di importanza*” pag. 5

INTRODUZIONE

Il testo funge da corredo statistico alla relazione sulla valutazione della soddisfazione degli studenti frequentanti e ha lo scopo di illustrare le metodologie utilizzate per l'analisi dei dati.

In particolare viene descritto il contenuto e vengono esposti i criteri con i quali sono stati calcolati i dati delle tabelle contenute nella relazione. Inoltre viene presentato il coefficiente ρ di *Spearman* utilizzato per il calcolo della dipendenza tra le domande.

La relazione e questo fascicolo sulle metodologie statistiche utilizzate sono anche disponibili sulle pagine web dedicate alla "Valutazione della didattica".

Note sull'approfondimento dei risultati emersi

I) Nota al § 3.1.1 e al § 4.1.1 Tabella 4 *“Giudizio complessivo positivo dell'insegnamento” ripartito per raggruppamenti di studenti frequentanti e per anno di corso*

Nella tabella *“Giudizio complessivo positivo dell'insegnamento” ripartito per raggruppamenti di studenti frequentanti e per anno di corso* viene riportato nell'ordine:

- il *raggruppamento* che è il range del numero di studenti che hanno compilato il questionario per insegnamento
- il *numero di insegnamenti* di ciascun raggruppamento
- il *totale rispondenti alla domanda* che è il numero di questionari compilati per ciascun raggruppamento
- *di cui con giudizio positivo* è il numero di risposte positive (“Decisamente SI” e “Più Sì che no”) alla domanda *“È complessivamente soddisfatto delle lezioni che sono state tenute dal docente che sta valutando?”*
- la *percentuale ed il relativo numero di studenti* che per ogni anno di corso hanno risposto positivamente
- il *valore medio ponderato* delle precedenti percentuali, ovvero la percentuale di giudizi positivi di ciascun raggruppamento

II) Nota al § 3.1.2 e al § 4.1.2 Tabella 5 *“Distribuzione di frequenza percentuale delle domande più rilevanti”*

Nella tabella *“Distribuzione di frequenza percentuale delle domande più rilevanti”* viene indicato, in termini percentuali, il giudizio espresso dagli studenti rispetto alle domande del questionario maggiormente inerenti all'attività didattica, per ciascuna Sede. Per ciascuna domanda è stato inoltre calcolato l'“Indice di sintesi” che varia tra 2 e 9 ed è ottenuto come media aritmetica delle risposte alle quali sono stati attribuiti i punteggi: 2 alla modalità “Decisamente NO”, 4 alla modalità “più NO che sì”, 7 alla modalità “Più SI che no” e 9 alla modalità “Decisamente SI”.

III) Nota al § 3.2.1 e al § 4.2.1 Tabella 9 *“Confronto tra la soddisfazione complessiva per gli insegnamenti e l'interesse agli argomenti indicati nel programma negli aa. aa. 2019/20 e 2020/21”*

La tabella *“Confronto tra la soddisfazione complessiva per gli insegnamenti e l'interesse agli argomenti indicati nel programma”* propone un confronto tra la distribuzione di frequenza assoluta e percentuale delle risposte date alle domande: *“È complessivamente soddisfatto delle lezioni che sono state tenute dal docente che sta valutando?”* e *“All'inizio dell'insegnamento, era interessato agli argomenti indicati nel programma?”* negli aa.aa. 2019/2020 e 2020/2021. Il dato è dettagliato per sede e per anno di corso.

IV) Nota al § 3.2.1 e al § 4.2.1 Tabella 10 “Correlazione di Spearman in ordine decrescente di importanza”

Per determinare il grado di correlazione tra le variabili è stato utilizzato il coefficiente di correlazione di Spearman indicato con il simbolo greco ρ (rho) (proposto con l'articolo *The proof and measurement of association between two things* su American Journal of Psychology vol. 15, pp.72 – 101 e con l'articolo *A footnote for measuring correlation*, pubblicato nel 1906 su British Journal of Psychology).

Si tratta di un coefficiente di correlazione lineare, applicato a variabili ordinali, tra posizioni (*ranghi*), in grado di quantificare il grado di concordanza tra due variabili; dove per *rango* si intende la posizione assunta da ogni osservazione della variabile, dopo avere ordinato l'intera distribuzione. Ad esempio, in una variabile che assume le seguenti modalità: 2, 12, 7, 4, l'osservazione "2" ha rango 1, l'osservazione "4" ha rango 2, l'osservazione "7" ha rango 3, l'osservazione "12" ha rango 4.

Il metodo consiste nel calcolare la correlazione

$$r = \frac{\text{cov}(xy)}{\sqrt{\text{var}(x) * \text{var}(y)}}$$

non sui punteggi originali ma su quelli trasformati in ranghi, assegnando a ciascuna osservazione una posizione nella serie di dati preventivamente ordinati, pertanto la formula diviene:

$$\rho = \frac{\sum_{i=1}^N \left(Rx_i - \frac{N+1}{2} \right) * \left(Ry_i - \frac{N+1}{2} \right)}{N(N^2 - 1)}$$

12

dove:

Rx_i e Ry_i sono i ranghi delle 2 variabili

N è il numero complessivo delle osservazioni

Quando due o più valori delle variabili sono identici e quindi hanno lo stesso rango, l'attribuzione dei punteggi medi riduce il valore della varianza ($\text{var}(x)$, $\text{var}(y)$), in tal caso è opportuno calcolare un fattore di correzione T per entrambe le variabili x e y .

Dove:

$$T = \sum_{i=1}^g (t_i^3 - t_i)$$

g è il numero di valori con punteggi identici

t è il numero di ranghi identici per ogni valore

Con questa correzione la formula del ρ di *Spearman* diventa:

$$\rho = \frac{N^3 - N - 6 * d - (T_x + T_y) / 2}{\sqrt{(N^3 - N)^2 - (T_x + T_y) * (N^3 - N) + T_x * T_y}}$$

dove:

$$d = \sum_{i=1}^N (R_{x_i} - R_{y_i})^2$$

Tale coefficiente assume valori compresi tra -1 e +1, dove gli estremi indicano la perfetta dipendenza (inversa o diretta) tra le variabili, anche se ciò è difficilmente riscontrabile nella realtà.

Se il coefficiente assume valori positivi (il valore che esprime la massima correlazione possibile è pari a 1) c'è dipendenza diretta tra i giudizi dati alle singole domande e il giudizio complessivo (al crescere di un giudizio, cresce anche l'altro), se invece assume valori negativi si manifesta una dipendenza inversa. Un coefficiente pari a zero indica infine l'assoluta indipendenza delle due variabili.