

# Documentazione Progetto Python:

## Implementazione del Test del Chi-Quadro e/o del Test ANOVA

Michela Rossi

IN400 - Modulo A

Il presente progetto ha lo scopo di effettuare il test del  $\chi^2$  o il test di ANOVA sulla base di un file csv fornito in input dall'utente.

Fornisce una visualizzazione grafica della distribuzione associata al test, evidenzia la regione critica, il quantile associato alla significatività del test scelta dall'utente e il valore della statistica calcolato sui dati forniti; inoltre permette di salvare tale grafico come PDF o PNG

Ho strutturato il mio progetto in un package e un main.

Nel package ho creato due moduli: stat e I.O

- **stat** contiene le funzioni che calcolano i valori delle statistiche e le funzioni che effettuano i test. Include :
  - **Chi\_Quadro** -> funzione che prende in input il file formato csv contenente i dati e calcola il valore della statistica  $\chi^2$  associata
  - **Chi\_Quadro\_test** -> funzione che prende in input il valore della statistica sopra calcolata, la significatività del test e confronta la statistica calcolata con il quantile relativo alla significatività scelta
  - **Fisher** -> funzione che prende in input il file formato csv contenente i dati e calcola il valore della statistica di Fisher associata
  - **Anova\_test** -> funzione che prende in input il valore della statistica sopra calcolata, la significatività del test e confronta la statistica calcolata con il quantile relativo alla significatività scelta

- `valore_statistica` -> funzione che prende in input il file in formato csv contenente i dati, il metodo selezionato dall'utente e richiama la funzione adeguata in base al metodo, restituendo il valore della statistica
- `test` -> funzione che prende in input il valore della statistica, la significatività, chiama le funzioni di test in base alla scelta dell'utente fornendo l'esito finale del test
- **I\_O** è un modulo di input output, consente di far interagire l'utente con il programma fornendo anche un'interfaccia grafica. Include
  - `leggi_csv` -> funzione che prende in input il file csv e lo salva su una variabile di tipo *dataframe*.
  - `stampa_chi_quadro` -> funzione che prende in input il valore della statistica prodotta dal test  $\chi^2$ , i gradi di libertà, la significatività e produce un plot con la distribuzione, il valore del quantile associato alla significatività, il valore della statistica e colora la regione di rifiuto
  - `stampa_fisher` -> funzione che prende in input il valore della statistica prodotta dal test ANOVA, i due gradi di libertà, la significatività e produce un plot con la distribuzione, il valore del quantile associato alla significatività, il valore della statistica e colora la regione di rifiuto
  - `stampa_statistica` -> funzione che prende in input il valore della statistica, la significatività e, richiamando la funzione di stampa adeguata, stampa la statistica in base al test che si esegue
  - `scegli_metodo` -> consente di scegliere quale test effettuare e ritorna il metodo scelto (nome del test)
  - `menu_popup` -> apre un menù popup tramite cui l'utente interagisce con il programma scegliendo il metodo, inserendo il nome del file, scegliendo la significatività del test

- `scegli_formato` → funzione che mostra una finestra con due pulsanti: PDF e PNG e consente di scegliere in che formato salvare il file e ritorna una stringa che indica il formato

- `salva_come` → funzione che chiede all'utente se vuole salvare il plot che prende in input come classe `Figure` di `matplotlib`

- **main** è un file Jupyter formato da tre celle:

- (1) importazione del package

- (2) chiamata al menù popup e salvataggio dell'output su tre variabili: `file_csv`, `metodo_scelto`, `alpha`

- (3) calcolo del valore della statistica tramite l'omonima funzione, chiamata alla funzione di stampa e chiamata alla funzione che restituisce il risultato del test

- **Possibili miglioramenti**

- le funzioni di I/O `stampa_statistica` e `test` scelgono quale funzione di stampa richiamare e/o quale test effettuare sulla base della dimensione della tupla restituita dalla funzione di stat `valore_statistica`. Si potrebbe migliorare facendo ritornare alla suddetta funzione un flag "F" o "CQ" in base al tipo di test e effettuare la scelta in `stampa_statistica` e in `test` tramite questo flag.

- è presente un bug nella funzione `salva_come`: quando si apre la finestra che chiede se si vuole salvare il file, cliccando sulla x per chiudere tale finestra, non accade nulla e rimane aperta.