

Rapporto tecnico

Silvia Cambiago – mat. 879382
Probabilità e Statistica per l'Informatica
CdL Informatica @UNIMIB

21/06/2023

Introduzione

Il seguente rapporto tratterà l'analisi dei dati eseguita in seguito alla richiesta di valutare l'ampliamento della sezione "Assistenza e Riparazioni" della sezione ciclismo di Decathlon Milano. È fornito un prospetto delle vendite della sezione ciclismo relativo all'anno 2022.

Sviluppo tecnico

Per stabilire se sia conveniente ampliare l'area di assistenza e riparazioni, è opportuno effettuare un confronto sulla differenza delle medie delle componenti vendute tramite servizio di assistenza e quelle vendute direttamente al pubblico.

Le colonne considerate (7 e 8 del prospetto) contano 52 elementi ciascuna, è quindi possibile considerarle campioni numerosi. Dato che i dati relativi alle vendite sono correlati in quanto raccolti nello stesso periodo di tempo, si procede ad effettuare un t -test a campioni appaiati sulla differenza delle medie dei due campioni.

Si parte innanzitutto dalla definizione dell'ipotesi nulla e, di conseguenza, di quella alternativa:

- H_0 : la differenza tra la media delle componenti vendute al pubblico e la media delle componenti vendute tramite servizio di assistenza è nulla;
- H_1 : la differenza tra la media delle componenti vendute al pubblico e la media delle componenti vendute tramite servizio di assistenza è diversa da 0.

Il livello di significatività scelto è 95%, quindi $\alpha = 0,05$.

Poste queste basi, si procede ora ad effettuare il t -test a campioni appaiati tramite R.

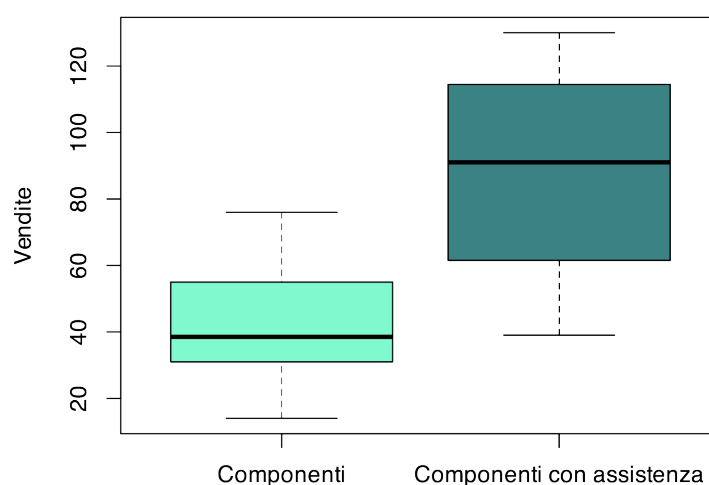
Risultati

In seguito all'esecuzione del test, si ottiene che la differenza media tra i valori delle due colonne è compresa nell'intervallo $(-49.83985, -41.16015)$ a livello 5%.

Il p -value risulta essere inferiore a 2.2×10^{-16} , quindi i dati sono in netto contrasto con l'ipotesi nulla. Pertanto, H_0 viene rifiutata, accettando di conseguenza l'ipotesi alternativa H_1 .

Di seguito è riportato un boxplot, sempre ottenuto tramite R, che evidenzia le distribuzioni delle vendite senza e con assistenza ed i rispettivi valori medi.

Confronto vendite per l'anno 2022



Conclusioni

L'esecuzione del test ha portato al rifiuto di H_0 e quindi alla conclusione che la differenza tra la media delle componenti vendute al pubblico e quella delle componenti vendute tramite servizio di assistenza discordano.

In particolare, un p-value così basso denota che l'ipotesi nulla ed i dati sono in netto contrasto, il che vuol dire che le medie dei campioni differiscono notevolmente. Essendo questa differenza a favore delle componenti vendute tramite servizio di assistenza (e che alle stesse, essendo vendute in quantità maggiore, sarà correlata una porzione maggiore in termini di ricavi), è sensato concludere che un ampliamento della sezione "Assistenza e Riparazioni" possa portare ad un incremento dei profitti.