Il calcolo del p-value varia a seconda del test di verifica delle ipotesi che viene utilizzato. Infatti, possiamo scegliere di adottare un test statistico diverso in base ai dati che abbiamo e in base a quali obiettivi si vogliono raggiungere.

Per calcolare il p-value bisogna prima definire il test d’ipotesi, il quale ci consente di trarre conclusioni su una poplazione partendo dai dati osservati in un campione. Questo test valuta due ipotesi opposte tra loro riguardo la popolazione:

- ipotesi nula

- ipotesi alternativa

L’ipotesi nulla è rappresentata di solito da un’affermazione del tipo “non si riscontra alcun effetto” oppure “non ci sono differenze” ed è l’ipotesi che si cerca di confutare con il test d’ipotesi. L’ipotesi alternativa, al contrario rappresenta un’affermazione del tipo “si riscontra un effetto” oppure “ci sono differenze”.

L’ipotesi nulla ha una distribuzione di probabilità definita. Anche l’ipotesi alternativa però ne ha una. Se conducendo il test d’ipotesi emerge che si tratta della stessa distribuzione di probabilità allora accettiamo l’ipotesi nulla, altrimenti quella alternativa.

Il p-value è quel valore che ci aiuta a capire proprio questo e quindi ci aiuta a decidere se accettare o rifiutare l’ipotesi nulla. Quando rifiutiamo l’ipotesi nulla diciamo che i risultati sono statisticamente significativi.

Il p-value si basa su questo presupposto. Considerando la distribuzione di probabilità dell’ipotesi nulla, se il risultato che abbiamo osservato si verifica molto raramente, ovvero se si trova agli estremi della distibuzione di probabilità, allora è altamente verosimile che l’ipotesi nulla in realtà è falsa. Per essere più precisi nel descrivere il punto esatto dal quale in poi consideriamo l’ipotesi nulla falsa, è stto introdotto il concetto di p-value. Esso indica un valore di probabilità, compreso ovviamente tra 0 ed 1. Il p-value ci dice quanto è probabile oservare dei dati come quelli del campione che stiamo analizzando quando stimao considerndo la distribuzione di probabilità dell’ipotesi nulla.

Quado il p-value è basso (cioè vicino a 0) si dice che il risultato **è statisticamente significativo**.

Abbiamo capito quindi che più vicino allo 0 è, e meglio è. Vogliamo però adesso introdurre un altro concetto, ovvero quello della **soglia di accettazione** chiamata **alpha.** Se il p-value è minore di alpha allora rifiutiamo l’ipotesi nulla, altrimenti falliamo nel rigettare l’ipotesi nulla e di conseguenza il risultato non è statisticamente significativo. Sostanzialmente stiamo imponendo che da un certo valore di p-value in poi, non rifiutiamo più l’ipotesi nulla.

Alpha, quindi, condiziona la probabilità di ottenere un errore del tipo 1, ovvero il rifiuto dell’ipotesi nulla quando invece è vera. Se scegliamo un alpha molto piccolo abbiamo meno probabilità di rifiutare l’ipotesi nulla rispetto a sé scegliamo un alpha grande.

**One-Tailed test:**

quando la nostra ipotesi testa soltanto per un lato del valore. Es:

- assumando come ipotesi nulla “la media è uguale a k”

- scegliamo come ipotesi alternativa “la media è maggiore di k” **oppure** “la media è minore di k”

**Two-Tailed test:**

quando la nostra ipotesi testa entrambi i lati di un valore. Es:

- assumando come ipotesi nulla “la media è uguale a k”

- scegliamo come ipotesi alternativa “la media è maggiore di k” **e** “la media è minore di k”. Quindi dobbiamo controllare entrambi i lati.

**Alpha:**

Quanto vicino agli estremi devono essere i dati per rigettare l’ipotesi nulla. E’ un treshold per il p-value. Di solito viene preso come 0.01 o 0.05 oppure 0.1.

**P-value:**

Quanto probabile è vedere dei dati così come quelli estratti, assumendo vera l’ipotesi nulla, cioè considerando la distribuzione di probabilità dell’ipotesi nulla. Dato che alpha è un treshold per p-value, se p-value è <= alpha allora rigettiamo l’ipotesi nulla e diciamo che i dati sono statisticamente signifiativi, altrimenti accettimo l’ipotesi nulla.

**T-Test:**

Usato per determinare se c’è della differenza significativa tra le medie di due variabili e ci permette di capire se appartengno alla stessa distribuzione. E’ un tipo di test two-tailed.