

9. Test di Wilcoxon per dati appaiati

Il test di Wilcoxon per dati appaiati è l'equivalente non parametrico del test t di Student per dati appaiati, e va utilizzato in luogo di questo quando i dati non siano distribuiti in modo gaussiano.

Per i calcoli si procede in questo modo :

- determinare le differenze (con il segno) fra le n coppie di valori;
- stabilire, per ciascuna differenza, il numero di posizione nella lista delle differenze ordinate (questa volta ignorando il segno) in ordine numerico crescente: la più piccola differenza osservata avrà numero di posizione 1, e via dicendo;
- quando due o più differenze sono uguali, assegnare a ciascuna di esse la media dei numeri di posizione che esse dovrebbero avere; così, per esempio, se la quinta e la sesta differenza sono uguali, assegnare come numero di posizione nella lista il valore 5.5 a entrambe;
- riassegnare il segno ai numeri di posizione nella lista (se la differenza aveva quel numero di posizione era negativa, assegnare il segno meno, se era positiva assegnare il segno più);
- calcolare il totale per i numeri di posizione con segno negativo e per i numeri di posizione con segno positivo, e chiamare T il più piccolo di questi due totali;
- calcolare la deviata normale standardizzata Z come

$$Z = (\mu - T - 0,5) / s$$

essendo

$$\begin{aligned}\mu &= n \cdot (n + 1) / 4 \\ s &= \sqrt{(2n + 1) \cdot \mu / 6}\end{aligned}$$

La statistica Z così calcolata corrisponde a sottoporre al test la mediana delle differenze.

Il valore di p corrispondente alla statistica Z rappresenta la probabilità di osservare per caso una differenza della grandezza di quella effettivamente osservata: se tale probabilità è sufficientemente piccola, si conclude per una significatività della differenza fra le mediane. Questa soluzione è sufficientemente accurata per $n > 16$.