5. Rapporto tra varianze

Il test F applicato al rapporto tra due varianze campionarie, consente di verificare se le varianze di due campioni sono tra di loro omogenee (cioè sostanzialmente uguali).

Indicati con x_1 i dati del primo campione, con n_1 il loro numero e con $\bar{x}_1$ la loro media, ed indicati con x_2 i dati del secondo campione, con n_2 il loro numero e con $\bar{x}_2$ la loro media, la varianza del primo campione viene calcolato come

$$s_1^2 = \Sigma (x_1 - \bar{x}_1)^2 / (n_1 - 1)$$

con gradi di libertà pari a

$$n_1 - 1$$

la varianza del secondo campione viene calcolata come

$$s_2^2 = \Sigma (x_2 - \bar{x}_2)^2 / (n_2 - 1)$$

con gradi di libertà pari a

$$n_2 - 1$$

E' possibile semplificare il calcolo della varianza s_1^2 e della varianza s_2^2 ricordando che

$$\Sigma (x_1 - \bar{x}_1)^2 = \Sigma x_1^2 - (\Sigma x_1)^2 / n_1$$
$$\Sigma (x_2 - \bar{x}_2)^2 = \Sigma x_2^2 - (\Sigma x_2)^2 / n_2$$

Il valore del test F di omogeneità fra le varianze viene calcolato allora come rapporto fra la varianza maggiore e la varianza minore, cioè

$$F = s_1^2 / s_2^2 \qquad se\ \ s_1^2 > s_2^2$$

ovvero

$$F = s_2^2 / s_1^2 \qquad se\ \ s_2^2 > s_1^2$$

Il valore di p corrispondente alla statistica F rappresenta la probabilità di osservare per caso un rapporto fra varianze della grandezza di quello effettivamente osservato: se tale probabilità è sufficientemente piccola, si conclude che le varianze sono significativamente diverse. Varianze diverse stanno ad indicare che i due campioni sono stati tratti da popolazioni diverse, aventi medie diverse. In questo senso il rapporto tra varianze corrisponde a un confronto tra medie.