# Test per confronto fra medie 1/4

Sia $X_1, X_2, \ldots, X_{n_1}$ un campione casuale di dimensione $n_1$ estratto da una distribuzione normale con media $\mu_1$ e varianza $\sigma_1^2$; sia $Y_1, Y_2, \ldots, Y_{n_2}$ un altro campione casuale di dimensione $n_2$ estratto da una distribuzione normale con media $\mu_2$ e varianza $\sigma_2^2$. Si consideri il seguente test

$$\begin{aligned} H_0 &: \quad \mu_1 = \mu_2 \\ H_1 &: \quad \mu_1 \neq \mu_2 \end{aligned}$$

ad un livello di significatività $\alpha$, assegnati due campioni $(x_1, \ldots, x_{n_1})$ e $(y_1, \ldots, y_{n_2})$ aventi medie $\bar{x}$ e $\bar{y}$.

# Test per confronto fra medie 2/4

*Caso varianze popolazione $\sigma_1^2$ e $\sigma_2^2$ note.*

Sotto $H_0$ la statistica test e la distribuzione pivot sono:

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0, 1)\ .$$

Pertanto, utilizzando argomenti simili a quelli considerati in precedenza, si ottiene la regione di accettazione $\mathcal{A} = \{-z_{1-\alpha/2} \leq Z \leq z_{1-\alpha/2}\}$ o, equivalentemente

$$\mathcal{A} = \left\{ -z_{1-\alpha/2} \leq \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \leq +z_{1-\alpha/2} \right\}$$

# Test per confronto fra medie 3/4

*Caso varianze popolazione $\sigma_1^2$ e $\sigma_2^2$ incognite.*

Siano $S_1^2$ e $S_2^2$ le varianza campionarie rispettivamente di $X_1, X_2, \ldots, X_{n_1}$ e $Y_1, Y_2, \ldots, Y_{n_2}$.
Sotto l'ipotesi nulla $H_0$ si ha $\mu_A = \mu_B$; pertanto la statistica test e la distribuzione pivot sono:

$$T = \frac{(\bar{X} - \bar{Y})}{S\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

dove

$$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

e quindi, come in precedenza, si ottiene:

$$\mathcal{A} = \left\{-t_{1-/\alpha/2}(n_1 + n_2 - 2) \leq T \leq t_{1-/\alpha/2}(n_1 + n_2 - 2)\right\}.$$

# Test per confronto fra medie 4/4

Assegnati due campioni $(x_1, \ldots, x_{n_1})$ e $(y_1, \ldots, y_{n_2})$ aventi medie $\bar{x}$ e $\bar{y}$ e varianze rispettivamente $s_1^2$ e $s_2^2$, si calcola il valore della statistica test nel campione osservato:

$$t^* = \frac{(\bar{x} - \bar{y})}{s\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

dove

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

costituisce una stima puntuale (corretta) della varianza della popolazione, e si verifica se risulta $t^* \in \mathcal{A}$ o $t^* \in \mathcal{C}$.

# Esercizio D

Un'associazione di difesa dei consumatori, sostiene che il prezzo della benzina della compagnia *A* sia superiore a quello della compagnia *B*, prima di aver rilevato i dati. Si considerano due campioni, rispettivamente di $n_A = 35$ e $n_B = 40$ osservazioni ottenendo valori medi rispettivamente pari a $\bar{x}_A = 1.823$ euro e $\bar{x}_B = 1.752$ euro e varianze rispettivamente uguali a $s_A^2 = 0.19$ e $s_B^2 = 0.24$. In questo caso, $\mu_A$ rappresenta il prezzo medio alle stazioni di rifornimento *A* e $\mu_B$ è il prezzo medio nelle stazioni di rifornimento *B*.

Si vuole verificare l'ipotesi che la compagnia $A$ abbia prezzi alla pompa maggiori di quelli della compagnia $B$ a livello di significatività $\alpha = 0.05$. Si assumano distribuzioni normali con varianze uguali $\sigma_A^2 = \sigma_B^2$.

# Esercizio D - Soluzione 1/4

1. *Definire le ipotesi*. Il problema è quello di verificare l'affermazione $\mu_A > \mu_B$. Le ipotesi nulla ed alternativa si scrivono:

$$H_0 : \mu_A \leq \mu_B$$
$$H_1 : \mu_A > \mu_B$$

   Si tratta di un test a una coda.

2. *Definire la statistica test e la distribuzione pivot*. Poichè le varianze sono non note, la statistica test è:

$$T = \frac{[(\bar{X}_A - \bar{X}_B) - (\mu_A - \mu_B)]}{S\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \sim t(n_A + n_B - 2)$$

# Esercizio D - Soluzione 2/4

3. *Definire la regione di rifiuto*. L'ipotesi nulla $H_0 : \ \mu_A = \mu_B$ verrà rigettata se la statistica test $T$ assumerà valori grandi. Si ha pertanto:

$$\mathcal{C} = \left\{ \frac{(\bar{X}_A - \bar{X}_B)}{S\sqrt{\dfrac{1}{n_A} + \dfrac{1}{n_B}}} > t_{1-\alpha}(n_A + n_B - 2) \right\} .$$

Nel nostro caso la stima corretta $s^2$ della varianza della popolazione, sulla base delle stime nei due campioni, è data da

$$s^2 = \frac{(n_1 - 1)s_A^2 + (n_2 - 1)s_B^2}{n_1 + n_2 - 2} = \frac{(35 - 1) \cdot 0.19 + (40 - 1) \cdot 0.24}{35 + 40 - 2} = 0.2167$$

da cui $s = \sqrt{0.2167} = 0.4655$.

# Esercizio D - Soluzione 3/4

4. *Calcolare la statistica ed eseguire il test*. Sulla base dei dati raccolti, la distribuzione pivot è una distribuzione $t$ con $n_1 + n_2 - 2 = 35 + 40 - 2 = 73$. Poichè $n$ è abbastanza grande, la distribuzione $t$ può essere approssimata mediante una normale standard; pertanto il valore critico è $z_{0.95} = 1.645$. Calcolato:

$$t^* = \frac{1.823 - 1.752}{0.4655\sqrt{\frac{1}{35} + \frac{1}{40}}} = \frac{0.071}{0.108} = 0.66$$

la regola di decisione è:

- si accetta $H_0$ se $t^* \leq 1.645$;
- si rifiuta $H_0$ se $t^* > 1.645$;

Poichè $t^* = 0.66 < 1.645$, non c'è sufficiente evidenza per respingere l'ipotesi nulla $H_0$.

# Esercizio D - Soluzione 4/4

5. *Conclusione*. Il valore $\bar{x}_A$ non risulta significativamente maggiore di $\bar{x}_B$ e quindi non si può affermare che la compagnia $A$ abbia un costo mediamente superiore a quello della compagnia $B$.