

Corso di

Statistica

La verifica delle ipotesi sulla varianza

Lezione quattordici

Cristina Davino

cdavino@unimc.it



Test per la differenza tra due medie

Esempio (1° caso)

Un automobilista decide di cambiare gli pneumatici della propria automobile, potendo scegliere tra due diverse marche, **X** e **Y**, di uguale prezzo e per i quali le case costruttrici garantiscono una durata media di 35mila km e una deviazione standard di 2000 km, uguale per le due marche.

Prima di effettuare l'acquisto, l'automobilista decide di sottoporre a verifica le ipotesi:

$$H_0: \mu_x - \mu_y = 0$$

$$H_1: \mu_x - \mu_y \neq 0$$

Si informa quindi presso amici e conoscenti che utilizzano i due tipi di pneumatici, trovando 14 persone che utilizzano la marca **X** e 9 che utilizzano la marca **Y**. I risultati sono:

$$\bar{x} = 33.500$$

$$\bar{y} = 36.000$$

Supponendo che la durata degli pneumatici si distribuisca secondo una legge Normale, si verifichi l'ipotesi H_0 con un livello di significatività $\alpha=0,05$.



La stima della differenza fra medie

$$X \sim N(\mu_x; \sigma_x^2) \quad Y \sim N(\mu_y; \sigma_y^2)$$

(varianze note)

$$H_0: \mu_x = \mu_y \quad H_1: \mu_x > \mu_y$$

$$H_1: \mu_x < \mu_y$$

$$H_1: \mu_x \neq \mu_y$$

$$n_x; n_y$$

(campioni indipendenti)

STIMATORE

$$\begin{aligned} E(\bar{X} - \bar{Y}) &= \mu_x - \mu_y \\ \text{Var}(\bar{X} - \bar{Y}) &= \frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y} \end{aligned}$$

Sotto H_0

$$\bar{X} - \bar{Y} \sim N\left(0; \frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}\right)$$

Per **campioni grandi** oppure quando **sia X che Y** si distribuiscono Normalmente

Test per la differenza tra due medie

Esempio (1° caso)

(sulla differenza fra medie)

Un automobilista decide di cambiare gli pneumatici della propria automobile, potendo scegliere tra due diverse marche, **X** e **Y**, di uguale prezzo e per i quali le case costruttrici garantiscono una durata media di 35mila km ed una deviazione standard di 2000 km. Si informa quindi presso amici e conoscenti che utilizzano i due tipi di pneumatici, trovando 14 persone che utilizzano la marca **X** e 9 che utilizzano la marca **Y**. Supponendo che la durata degli pneumatici si distribuisca secondo una legge Normale, si verifichi l'ipotesi H_0 con un livello di significatività $\alpha=0,05$.

$$\mu_x = \mu_y = 35.000$$

$$\sigma_x = \sigma_y = 2.000 \text{ km}$$

$$X \sim N$$

$$km$$

$$n_x = 14$$

$$n_y = 9$$

$$\bar{x} = 33.500$$

$$\bar{y} = 36.000$$

$$z_{\alpha/2} = 1,96$$

○ **Ipotesi**

$$H_0: \mu_x - \mu_y = 0$$

$$H_1: \mu_x - \mu_y \neq 0$$

○ **Significatività**

$$\alpha = 0,05$$

○ **Statistica**

$$(\bar{X} - \bar{Y}) \sim N$$

○ **Regola di decisione**

$$\text{Rifiuto } H_0 \text{ se: } \left| \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} \right| > z_{\alpha/2}$$

$$\frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sigma \cdot \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} = \frac{\bar{x} - \bar{y}}{2000 \cdot \sqrt{\frac{1}{14} + \frac{1}{9}}} = 2,93$$

1.

2.

Valori critici non standardizzati: $(\mu_x - \mu_y) \mp 1,96 \cdot \sigma \sqrt{\frac{1}{n_x} + \frac{1}{n_y}} = \begin{matrix} -1675 \\ +1675 \end{matrix}$

$$-2500 < -1675 \Rightarrow \text{Rifiuto } H_0$$

$$2,93 > 1,96 \Rightarrow \text{Rifiuto } H_0$$



Test per la differenza tra due medie

Esempio (2° caso) (sulla differenza fra medie)

Un automobilista decide di cambiare gli pneumatici della propria automobile, potendo scegliere tra due diverse marche, **X** e **Y**, di uguale prezzo e per i quali le case costruttrici garantiscono una durata media di 35mila km. Si può ipotizzare che la variabilità della durata sia identica (ma incognita).



Prima di effettuare l'acquisto, l'automobilista decide di sottoporre a verifica le ipotesi:

$$H_0: \mu_x - \mu_y = 0 \quad H_1: \mu_x - \mu_y \neq 0$$

Si informa quindi presso amici e conoscenti che utilizzano i due tipi di pneumatici, trovando 14 persone che utilizzano la marca **X** e 9 che utilizzano la marca **Y**. I risultati sono:

$$\bar{X} = 33.500 \quad \bar{Y} = 36.000 \quad \tilde{S}_x^2 = 4.326.400 \quad \tilde{S}_y^2 = 3.880.900$$

Supponendo che la durata degli pneumatici si distribuisca secondo una legge Normale, si verifichi l'ipotesi H_0 con un livello di significatività $\alpha=0,05$.

La stima della differenza fra medie

Esempio (2° caso) (sulla differenza fra medie)

Un automobilista decide di cambiare gli pneumatici della propria automobile, potendo scegliere tra due diverse marche, **X** e **Y**, di uguale prezzo e per i quali le case costruttrici garantiscono una durata media di 35mila km, potendo ipotizzare che la variabilità della durata sia identica.

L'automobilista si informa quindi presso amici e conoscenti che utilizzano i due tipi di pneumatici, trovando 14 persone che utilizzano la marca **X** e 9 che utilizzano la marca **Y**. Supponendo che la durata degli pneumatici si distribuisca secondo una legge Normale, si verifichi l'ipotesi H_0 con un livello di significatività $\alpha=0,05$.



$$\mu_x = \mu_y = 35.000 \quad \sigma_x = \sigma_y = ? \quad X \sim N$$

$$n_x = 14 \quad n_y = 9 \quad \bar{x} = 33.500 \quad \bar{y} = 36.000 \quad \tilde{S}_x^2 = 4.326.400 \quad \tilde{S}_y^2 = 3.880.900$$

○ **Ipotesi** ○ **Significatività** ○ **Statistica** ○ **Regola di decisione**

$$H_0: \mu_x - \mu_y = 0$$

$$\alpha = 0,05$$

$$(\bar{X} - \bar{Y}) \sim N$$

$$\text{Rifiuto } H_0 \text{ se: } \left| \frac{\bar{x} - \bar{y}}{\tilde{S} \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \right| > t_{\frac{\alpha}{2}, (n+m-2)}$$

Occorre stimare la varianza comune

$$\tilde{S} = \sqrt{\frac{(n_x - 1) \cdot \tilde{S}_x^2 + (n_y - 1) \cdot \tilde{S}_y^2}{n_x + n_y - 2}} = \sqrt{\frac{13 \cdot 4.326.400 + 8 \cdot 3.880.900}{14 + 9 - 2}} = 2038,8$$

$$\frac{\bar{x} - \bar{y}}{\tilde{S} \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} = \frac{33.500 - 36.000}{2038,8 \cdot \sqrt{\frac{1}{14} + \frac{1}{9}}} = \frac{-2500}{2038,8 \cdot \sqrt{\frac{1}{14} + \frac{1}{9}}} = -2,87$$

$$t_{0,025;21} = 2,08$$

$$|-2,87| > 2,08 \Rightarrow \text{Rifiuto } H_0$$

La stima della differenza fra medie

$$X \sim N(\mu_x; ?) \quad Y \sim N(\mu_y; ?)$$

(varianze ignote ma uguali)

$$H_0: \mu_x = \mu_y \quad H_1: \mu_x > \mu_y$$

$$n_x; n_y$$

$$H_1: \mu_x < \mu_y$$

(campioni indipendenti)

$$H_1: \mu_x \neq \mu_y$$

$$S_x; S_y$$

STIMATORE

$$\text{Sotto } H_0 \quad T = \frac{\bar{X} - \bar{Y}}{\sqrt{\tilde{S}^2 \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}} \sim T_{n_x + n_y - 2}$$

Per **campioni grandi** oppure quando **sia X che Y** si distribuiscono Normalmente

$$\tilde{S} = \sqrt{\frac{(n_x - 1) \cdot \tilde{S}_x^2 + (n_y - 1) \cdot \tilde{S}_y^2}{n_x + n_y - 2}}$$

La stima della differenza fra medie

Eteroschedasticità: $\sigma_x^2 \neq \sigma_y^2$

Omoschedasticità: $\sigma_x^2 = \sigma_y^2$

Riepilogo sulla stima della differenza fra medie

