

Presi i dati $x_1, \dots, x_n$, sia dunque $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ la loro media.

Accetto H_0 se $\bar{x} > \mu_0 - \frac{\sigma}{\sqrt{n}} z_{1-\alpha}$ e la rifiuto altrimenti.

Test unilaterale superiore con H_0 composta

Sia $X_1, \dots, X_n$ un campione gaussiano di valore atteso μ incognito e varianza σ^2 nota. Vogliamo testare

$$H_0 : \mu \geq \mu_0 \quad H_A : \mu < \mu_0.$$

Accetto l'ipotesi nulla H_0 se la media campionaria è superiore a $\mu_0 - \varepsilon$ cioè se $\bar{x} > \mu_0 - \varepsilon$. La probabilità di commettere un errore di prima specie è allora

$$\mathbb{P}(\bar{X} \leq \mu_0 - \varepsilon | \mathbb{E}[\bar{X}] \geq \mu_0).$$

Poiché, se $\mathbb{P}_{X_i} = N(\mu, \sigma^2)$ si ha $\mathbb{P}_{\bar{X}} = N\left(\mu, \frac{\sigma^2}{n}\right)$, e $Z := \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ ha distribuzione $N(0, 1)$, abbiamo anche

$$\begin{aligned} \mathbb{P}(\bar{X} \leq \mu_0 - \varepsilon | \mathbb{E}[\bar{X}] = \mu \geq \mu_0) &= \mathbb{P}\left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq \frac{\mu_0 - \mu - \varepsilon}{\frac{\sigma}{\sqrt{n}}} | \mathbb{E}[\bar{X}] = \mu \geq \mu_0\right) = \\ &= \mathbb{P}\left(Z \leq \frac{(\mu_0 - \mu - \varepsilon)\sqrt{n}}{\sigma}\right) = \Phi\left(\frac{(\mu_0 - \mu - \varepsilon)\sqrt{n}}{\sigma}\right) \leq \Phi\left(\frac{-\varepsilon\sqrt{n}}{\sigma}\right) = 1 - \Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right). \end{aligned}$$

Se voglio limitare superiormente $\mathbb{P}(\bar{X} \leq \mu_0 - \varepsilon | \mu \geq \mu_0)$ cioè se voglio

$$\mathbb{P}(\bar{X} \leq \mu_0 - \varepsilon | \mathbb{E}[\bar{X}] = \mu \geq \mu_0) \leq \alpha \quad \forall \mu \geq \mu_0$$

scelgo ε in modo da avere $\Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) = 1 - \alpha$ cioè $\frac{\varepsilon\sqrt{n}}{\sigma} = z_{1-\alpha}$ e dunque scelgo

$$\varepsilon = \frac{\sigma}{\sqrt{n}} z_{1-\alpha}.$$

Presi i dati $x_1, \dots, x_n$, sia dunque $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ la loro media.

Accetto H_0 se $\bar{x} > \mu_0 - \frac{\sigma}{\sqrt{n}} z_{1-\alpha}$ e la rifiuto altrimenti.

6.2.2 Campione gaussiano di cui non è nota la varianza

Test bilaterale

Sia $X_1, \dots, X_n$ un campione gaussiano di valore atteso μ e varianza σ^2 entrambi ignoti. Vogliamo testare

$$H_0 : \mu = \mu_0 \quad H_A : \mu \neq \mu_0$$

H_0 è vera se e solo se $\mathbb{E}[\bar{X}] = \mu_0$ ovvero, per l'indipendenza di $\bar{X}$ e S^2 , se e solo se $\mathbb{E}\left[\frac{(\bar{X} - \mu_0)\sqrt{n}}{S}\right] = \mathbb{E}[\bar{X} - \mu_0] \sqrt{n} \mathbb{E}\left[\frac{1}{\sqrt{S^2}}\right] = 0$. Dunque considero $t := \frac{(\bar{x} - \mu_0)\sqrt{n}}{s}$ e accetto l'ipotesi nulla H_0 se $|t| \leq \varepsilon$.

Sappiamo che, se $\mu = \mu_0$, allora $T := \frac{(\bar{X} - \mu_0)\sqrt{n}}{S}$ ha distribuzione $t(n-1)$. Il livello di significatività è allora $\alpha = \mathbb{P}(|T| \geq \varepsilon)$ e si ha

$$\begin{aligned}\alpha &= \mathbb{P}(|T| \geq \varepsilon) = \mathbb{P}(T \geq \varepsilon) + \mathbb{P}(T \leq -\varepsilon) \\ &= 1 - F_T(\varepsilon) + F_T(-\varepsilon) = 2(1 - F_T(\varepsilon))\end{aligned}$$

Se voglio fissare a priori α , deve essere allora $F_T(\varepsilon) = 1 - \frac{\alpha}{2}$ dunque devo scegliere

$$\varepsilon = t_{n-1, 1-\frac{\alpha}{2}}.$$

Presi i dati $x_1, \dots, x_n$, dunque accetto H_0 se $|t| \leq t_{n-1, 1-\frac{\alpha}{2}}$ e la rifiuto altrimenti, ovvero

$$\text{accetto } H_0 \text{ se } \mu_0 - \frac{t_{n-1, 1-\frac{\alpha}{2}} s}{\sqrt{n}} \leq \bar{x} \leq \mu_0 + \frac{t_{n-1, 1-\frac{\alpha}{2}} s}{\sqrt{n}} \text{ e la rifiuto altrimenti.}$$

Test unilaterale inferiore con ipotesi nulla semplice

Sia $X_1, \dots, X_n$ un campione gaussiano di valore atteso μ e varianza σ^2 entrambi incogniti. Vogliamo testare

$$H_0 : \quad \mu = \mu_0, \quad H_A : \quad \mu > \mu_0$$

Diamo la seguente regola di accettazione: accettiamo H_0 se $\frac{(\bar{x} - \mu_0)\sqrt{n}}{s} \leq \varepsilon$.

La probabilità di commettere un errore di prima specie è allora

$$\alpha = \mathbb{P}\left(\frac{(\bar{X} - \mu_0)\sqrt{n}}{S} > \varepsilon \mid \mu = \mu_0\right) = \mathbb{P}(T > \varepsilon) = 1 - F_T(\varepsilon).$$

dove $\mathbb{P}_T = t(n-1)$. Se vogliamo stabilire il livello di significatività α dovremmo scegliere ε in modo che

$$1 - F_T(\varepsilon) = \alpha$$

cioè $\varepsilon = t_{n-1, 1-\alpha}$.

Presi i dati $x_1, \dots, x_n$, sia dunque $t_0 = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s}$. Accetto H_0 se $t_0 \leq t_{n-1, 1-\alpha}$ ovvero

$$\text{accetto } H_0 \text{ se } \bar{x} \leq \mu_0 + \frac{t_{n-1, 1-\alpha} s}{\sqrt{n}} \text{ e la rifiuto altrimenti.}$$

Test unilaterale inferiore con ipotesi nulla composta

Sia $X_1, \dots, X_n$ un campione gaussiano di valore atteso μ e varianza σ^2 entrambi incogniti. Vogliamo testare

$$H_0 : \quad \mu \leq \mu_0, \quad H_A : \quad \mu > \mu_0$$

Diamo la seguente regola di accettazione: accettiamo H_0 se $\frac{(\bar{X} - \mu_0)\sqrt{n}}{S} \leq \varepsilon$.

La probabilità di commettere un errore di prima specie è allora

$$\mathbb{P} \left(\frac{(\bar{X} - \mu_0)\sqrt{n}}{S} > \varepsilon | \mathbb{E}[\bar{X}] = \mu \leq \mu_0 \right).$$

Se H_0 è vera, allora $\mathbb{E}[\bar{X}] = \mu \leq \mu_0$ e dunque

$$\frac{(\bar{X} - \mu_0)\sqrt{n}}{S} \leq \frac{(\bar{X} - \mu)\sqrt{n}}{S} =: T, \quad \mathbb{P}_T = t(n-1).$$

Di conseguenza

$$\left\{ \frac{(\bar{X} - \mu_0)\sqrt{n}}{S} > \varepsilon \right\} \subset \left\{ \frac{(\bar{X} - \mu)\sqrt{n}}{S} > \varepsilon \right\}$$

Dunque, per ogni $\mu \leq \mu_0$ si ha

$$\begin{aligned} \mathbb{P} \left(\frac{(\bar{X} - \mu_0)\sqrt{n}}{S} > \varepsilon | \mathbb{E}[\bar{X}] = \mu \right) &\leq \\ &\leq \mathbb{P} \left(\frac{(\bar{X} - \mu)\sqrt{n}}{S} > \varepsilon | \mathbb{E}[\bar{X}] = \mu \right) = \mathbb{P}(T > \varepsilon) = 1 - F_T(\varepsilon). \end{aligned}$$

Se vogliamo controllare il livello di significatività α dovremmo scegliere ε in modo che

$$1 - F_T(\varepsilon) = \alpha$$

cioè $\varepsilon = t_{n-1, 1-\alpha}$.

Presi i dati $x_1, \dots, x_n$, sia dunque $t_0 = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s}$. Accetto H_0 se $t_0 \leq t_{n-1, 1-\alpha}$ ovvero

$$\text{accetto } H_0 \text{ se } \bar{x} \leq \mu_0 + \frac{t_{n-1, 1-\alpha} s}{\sqrt{n}} \text{ e la rifiuto altrimenti.}$$

Test unilaterale superiore con ipotesi nulla semplice

Sia $X_1, \dots, X_n$ un campione gaussiano di valore atteso μ e varianza σ^2 entrambi incogniti. Vogliamo testare

$$H_0: \mu = \mu_0, \quad H_A: \mu < \mu_0.$$

Diamo la seguente regola di accettazione: accettiamo H_0 se $\frac{(\bar{x} - \mu_0)\sqrt{n}}{s} \geq -\varepsilon$.

La probabilità di commettere un errore di prima specie è allora

$$\alpha = \mathbb{P} \left(\frac{(\bar{X} - \mu_0)\sqrt{n}}{S} < -\varepsilon | \mu = \mu_0 \right) = \mathbb{P}(T < -\varepsilon) = F_T(-\varepsilon)$$

dove $\mathbb{P}_T = t(n-1)$. Se vogliamo stabilire il livello di significatività α dovremmo scegliere ε in modo che

$$F_T(-\varepsilon) = \alpha$$

cioè $\varepsilon = -t_{n-1, \alpha} = t_{n-1, 1-\alpha}$.

Presi i dati $x_1, \dots, x_n$, sia dunque $t_0 = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s}$. Accetto H_0 se $t_0 \geq -t_{n-1, 1-\alpha}$ e la rifiuto altrimenti, ovvero accetto H_0 se

$$\bar{x} \geq \mu_0 - \frac{t_{n-1, 1-\alpha} s}{\sqrt{n}}$$

e la rifiuto altrimenti.

Test unilaterale superiore con ipotesi nulla composta

Sia $X_1, \dots, X_n$ un campione gaussiano di valore atteso μ e varianza σ^2 entrambi incogniti. Vogliamo testare l'ipotesi

$$H_0 : \mu \geq \mu_0, \quad H_A : \mu < \mu_0.$$

Diamo la seguente regola di accettazione: accettiamo H_0 se $\frac{(\bar{x} - \mu_0)\sqrt{n}}{s} \geq -\varepsilon$.

La probabilità di commettere un errore di prima specie è allora

$$\mathbb{P} \left(\frac{(\bar{X} - \mu_0)\sqrt{n}}{S} < -\varepsilon | \mathbb{E}[\bar{X}] = \mu \geq \mu_0 \right).$$

Se H_0 è vera, allora $\mathbb{E}[\bar{X}] = \mu \geq \mu_0$ e dunque

$$\frac{(\bar{X} - \mu_0)\sqrt{n}}{S} \geq \frac{(\bar{X} - \mu)\sqrt{n}}{S} =: T, \quad \mathbb{P}_T = t(n-1).$$

Di conseguenza

$$\left\{ \frac{(\bar{X} - \mu_0)\sqrt{n}}{S} < -\varepsilon \right\} \subset \left\{ \frac{(\bar{X} - \mu)\sqrt{n}}{S} < -\varepsilon \right\}$$

Dunque per ogni $\mu \geq \mu_0$ si ha

$$\begin{aligned} \mathbb{P} \left(\frac{(\bar{X} - \mu_0)\sqrt{n}}{S} < -\varepsilon | \mathbb{E}[\bar{X}] = \mu \right) &\leq \mathbb{P} \left(\frac{(\bar{X} - \mu)\sqrt{n}}{S} < -\varepsilon | \mathbb{E}[\bar{X}] = \mu \right) \\ &= \mathbb{P}(T < -\varepsilon) = F_T(-\varepsilon) = 1 - F_T(\varepsilon). \end{aligned}$$

Se vogliamo controllare il livello di significatività α dovremmo scegliere ε in modo che

$$1 - F_T(\varepsilon) = \alpha$$

cioè $\varepsilon = t_{n-1, 1-\alpha}$.

Presi i dati $x_1, \dots, x_n$, sia dunque $t_0 = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s}$. Accetto H_0 se $t_0 \geq -t_{n-1, 1-\alpha}$ e la rifiuto altrimenti, ovvero

$$\text{accetto } H_0 \text{ se } \bar{x} \geq \mu_0 - \frac{t_{n-1, 1-\alpha} s}{\sqrt{n}} \text{ e la rifiuto altrimenti.}$$

6.3 Test d'ipotesi per la varianza di campioni gaussiani

Test bilaterale

Sia $X_1, \dots, X_n$ un campione gaussiano di valore atteso μ (noto o incognito) e varianza σ^2 incognita. Vogliamo testare

$$H_0 : \sigma^2 = \sigma_0^2 \quad H_A : \sigma^2 \neq \sigma_0^2$$

H_0 è vera se e solo se $\mathbb{E}[S^2] = \sigma_0^2$ ovvero se e solo se $\mathbb{E}\left[\frac{S^2}{\sigma_0^2}\right] = 1$. Dunque accetto H_0 se $1 - \varepsilon_1 < \frac{s^2}{\sigma_0^2} < 1 + \varepsilon_2$, $\varepsilon_1, \varepsilon_2$ positivi, cioè se e solo se

$$(n-1)(1 - \varepsilon_1) < \frac{(n-1)s^2}{\sigma^2} < (n-1)(1 + \varepsilon_2).$$

Devo scegliere ε_1 e ε_2 in modo da ottenere il livello di significatività α desiderato. Sappiamo che se H_0 è vera, allora la v.a. $V := \frac{(n-1)S^2}{\sigma^2}$ ha distribuzione χ_{n-1}^2 .

$$\begin{aligned} \alpha &= \mathbb{P}\left(\frac{S^2}{\sigma_0^2} > 1 + \varepsilon_2 \mid \sigma^2 = \sigma_0^2\right) + \mathbb{P}\left(\frac{S^2}{\sigma_0^2} < 1 - \varepsilon_1 \mid \sigma^2 = \sigma_0^2\right) \\ &= \mathbb{P}\left(\frac{(n-1)S^2}{\sigma_0^2} > (n-1)(1 + \varepsilon_2) \mid \sigma^2 = \sigma_0^2\right) + \mathbb{P}\left(\frac{(n-1)S^2}{\sigma_0^2} < (n-1)(1 - \varepsilon_1) \mid \sigma^2 = \sigma_0^2\right) \\ &= \mathbb{P}(V > (n-1)(1 + \varepsilon_2)) + \mathbb{P}(V < (n-1)(1 - \varepsilon_1)). \end{aligned}$$

Una possibile scelta è allora

$$\begin{aligned} \mathbb{P}(V > (n-1)(1 + \varepsilon_2)) &= \frac{\alpha}{2} & \text{cioè} & \quad (n-1)(1 + \varepsilon_2) = \chi_{n-1, 1-\frac{\alpha}{2}}^2 \\ \mathbb{P}(V < (n-1)(1 - \varepsilon_1)) &= \frac{\alpha}{2} & \text{cioè} & \quad (n-1)(1 - \varepsilon_1) = \chi_{n-1, \frac{\alpha}{2}}^2. \end{aligned}$$

Dunque accetto H_0 se $\chi_{n-1, \frac{\alpha}{2}}^2 < \frac{(n-1)s^2}{\sigma_0^2} < \chi_{n-1, 1-\frac{\alpha}{2}}^2$ ovvero

$$\text{accetto } H_0 \text{ se } \frac{\sigma_0^2}{n-1} \chi_{n-1, \frac{\alpha}{2}}^2 < s^2 < \frac{\sigma_0^2}{n-1} \chi_{n-1, 1-\frac{\alpha}{2}}^2 \text{ e la rifiuto altrimenti.}$$

Test unilaterale inferiore con ipotesi semplice

Sia $X_1, \dots, X_n$ un campione gaussiano di valore atteso μ (noto o incognito) e varianza σ^2 incognita. Vogliamo testare

$$H_0 : \quad \sigma^2 = \sigma_0^2 \qquad H_A : \quad \sigma^2 > \sigma_0^2.$$

Accetto l'ipotesi nulla se $\frac{s^2}{\sigma_0^2} \leq 1 + \varepsilon$.

Se la varianza è σ_0^2 , allora $V := \frac{(n-1)S^2}{\sigma_0^2}$ ha distribuzione χ_{n-1}^2 e la probabilità di commettere errore di prima specie è

$$\mathbb{P}\left(\frac{S^2}{\sigma_0^2} > 1 + \varepsilon \mid \sigma^2 = \sigma_0^2\right) = \mathbb{P}\left(\frac{(n-1)S^2}{\sigma_0^2} > (n-1)(1 + \varepsilon) \mid \sigma^2 = \sigma_0^2\right) = 1 - F_V((n-1)(1 + \varepsilon)).$$

Posso allora limitare superiormente con α la probabilità di commettere errore di prima specie imponendo

$$1 - F_V((n-1)(1 + \varepsilon)) = \alpha$$

cioè scegliendo ε in modo che

$$(n-1)(1+\varepsilon) = \chi_{n-1,1-\alpha}^2.$$

Dunque accetto l'ipotesi nulla H_0 se $\frac{(n-1)s^2}{\sigma_0^2} < \chi_{n-1,1-\alpha}^2$ ovvero

$$\text{accetto } H_0 \text{ se } s^2 < \frac{\sigma_0^2}{n-1} \chi_{n-1,1-\alpha}^2 \text{ e la rifiuto altrimenti.}$$

Test unilaterale inferiore con ipotesi composta

Sia $X_1, \dots, X_n$ un campione gaussiano di valore atteso μ (noto o incognito) e varianza σ^2 incognita. Vogliamo testare

$$H_0: \quad \sigma^2 \leq \sigma_0^2 \quad \quad H_A: \quad \sigma^2 > \sigma_0^2.$$

Accetto l'ipotesi nulla se $\frac{s^2}{\sigma_0^2} \leq 1 + \varepsilon$.

Se la varianza è $\sigma^2 \leq \sigma_0^2$, allora $V := \frac{(n-1)S^2}{\sigma^2}$ ha distribuzione χ_{n-1}^2 e la probabilità di commettere errore di prima specie è

$$\begin{aligned} \mathbb{P} \left(\frac{S^2}{\sigma_0^2} > 1 + \varepsilon | \text{Var}[X_i] = \sigma^2 \leq \sigma_0^2 \right) &= \mathbb{P} \left(\frac{(n-1)S^2}{\sigma^2} > \frac{\sigma_0^2}{\sigma^2} (n-1)(1+\varepsilon) | \text{Var}[X_i] = \sigma^2 \leq \sigma_0^2 \right) \\ &= \mathbb{P} \left(V > \frac{\sigma_0^2}{\sigma^2} (n-1)(1+\varepsilon) \right) = 1 - F_V \left(\frac{\sigma_0^2}{\sigma^2} (n-1)(1+\varepsilon) \right) \\ &\leq 1 - F_V((n-1)(1+\varepsilon)) \end{aligned}$$

dove abbiamo usato la monotonia di F_V e il fatto che $\sigma^2 \leq \sigma_0^2$ implica $\frac{\sigma^2}{\sigma_0^2} \leq 1$.

Posso allora limitare superiormente con α la probabilità di commettere errore di prima specie imponendo

$$1 - F_V((n-1)(1+\varepsilon)) = \alpha$$

cioè scegliendo ε in modo che

$$(n-1)(1+\varepsilon) = \chi_{n-1,1-\alpha}^2.$$

Dunque accetto l'ipotesi nulla H_0 se $\frac{(n-1)s^2}{\sigma_0^2} < \chi_{n-1,1-\alpha}^2$ ovvero

$$\text{accetto } H_0 \text{ se } s^2 < \frac{\sigma_0^2}{n-1} \chi_{n-1,1-\alpha}^2 \text{ e la rifiuto altrimenti.}$$

Test unilaterale superiore con ipotesi semplice

Sia $X_1, \dots, X_n$ un campione gaussiano di valore atteso μ (nota o incognita) e varianza σ^2 incognita. Vogliamo testare

$$H_0 : \quad \sigma^2 = \sigma_0^2 \qquad H_A : \quad \sigma^2 < \sigma_0^2.$$

Accetto l'ipotesi nulla se $\frac{s^2}{\sigma_0^2} \geq 1 - \varepsilon$.

Se H_0 è vera, allora $V := \frac{(n-1)S^2}{\sigma_0^2}$ ha distribuzione χ_{n-1}^2 e la probabilità di commettere errore di prima specie è

$$\alpha = \mathbb{P} \left(\frac{S^2}{\sigma_0^2} < 1 - \varepsilon \mid \sigma^2 = \sigma_0^2 \right) = F_V((n-1)(1-\varepsilon)).$$

Deve quindi essere

$$(n-1)(1-\varepsilon) = \chi_{n-1,\alpha}^2.$$

Dunque accetto l'ipotesi nulla H_0 se $\frac{(n-1)s^2}{\sigma_0^2} > \chi_{n-1,\alpha}^2$ ovvero

$$\text{accetto } H_0 \text{ se } s^2 > \frac{\sigma_0^2}{n-1} \chi_{n-1,\alpha}^2 \text{ e la rifiuto altrimenti.}$$

Test unilaterale superiore con ipotesi composta

Sia $X_1, \dots, X_n$ un campione gaussiano di valore atteso μ (nota o incognita) e varianza σ^2 incognita. Vogliamo testare

$$H_0 : \quad \sigma^2 \geq \sigma_0^2 \qquad H_A : \quad \sigma^2 < \sigma_0^2.$$

Accetto l'ipotesi nulla se $\frac{s^2}{\sigma_0^2} \geq 1 - \varepsilon$.

Se la varianza è $\sigma^2 \geq \sigma_0^2$, allora $V := \frac{(n-1)S^2}{\sigma^2}$ ha distribuzione χ_{n-1}^2 e la probabilità di commettere errore di prima specie è

$$\begin{aligned} \mathbb{P} \left(\frac{S^2}{\sigma_0^2} < 1 - \varepsilon \mid \text{Var}[X_i] = \sigma^2 \geq \sigma_0^2 \right) \\ = \mathbb{P} \left(\frac{(n-1)S^2}{\sigma^2} < \frac{\sigma_0^2}{\sigma^2} (n-1)(1-\varepsilon) \mid \text{Var}[X_i] = \sigma^2 \geq \sigma_0^2 \right) \\ = F_V \left(\frac{\sigma_0^2}{\sigma^2} (n-1)(1-\varepsilon) \right) \leq F_V((n-1)(1-\varepsilon)). \end{aligned}$$

Posso allora limitare superiormente con α la probabilità di commettere errore di prima specie imponendo

$$F_V((n-1)(1-\varepsilon)) = \alpha$$

cioè scegliendo ε in modo che

$$(n-1)(1-\varepsilon) = \chi_{n-1,\alpha}^2.$$

Dunque accetto l'ipotesi nulla H_0 se $\frac{(n-1)s^2}{\sigma_0^2} > \chi_{n-1,\alpha}^2$ ovvero

accetto H_0 se $s^2 > \frac{\sigma_0^2}{n-1} \chi_{n-1, \alpha}^2$ e la rifiuto altrimenti.

7. Intervalli di confidenza e Test di ipotesi per il confronto di campioni gaussiani

7.1 Intervalli di confidenza per la differenza dei valori attesi di campioni gaussiani

Supponiamo di avere due campioni, entrambi gaussiani e tra di loro indipendenti

$$\begin{aligned} X: X_1, \dots, X_n & \quad \mathbb{P}_{X_i} = N(\mu_X, \sigma_X^2), \\ Y: Y_1, \dots, Y_k & \quad \mathbb{P}_{Y_j} = N(\mu_Y, \sigma_Y^2). \end{aligned}$$

Vogliamo costruire un intervallo di confidenza per la differenza $\mu_X - \mu_Y$. Consideriamo due casi distinti (ma non esaustivi)

7.1.1 Le varianze σ_X^2 e σ_Y^2 sono note

Osservazione 7.1.1. Si può dimostrare che $\bar{X} - \bar{Y}$ è stimatore di massima verosimiglianza per $\mu_X - \mu_Y$.

Sappiamo che $\mathbb{P}_{X_i} = N(\mu_X, \sigma_X^2)$, $\mathbb{P}_{Y_j} = N(\mu_Y, \sigma_Y^2)$, dunque

$$\mathbb{P}_{\bar{X}} = N\left(\mu_X, \frac{\sigma_X^2}{n}\right), \quad \mathbb{P}_{\bar{Y}} = N\left(\mu_Y, \frac{\sigma_Y^2}{k}\right).$$

Di conseguenza la v.a. $\frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}}}$ ha distribuzione gaussiana standard. Abbiamo dunque

$$\begin{aligned} 1 - \alpha &= \mathbb{P}\left(z_{\frac{\alpha}{2}} < \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}}} < z_{1-\frac{\alpha}{2}}\right) \\ &= \mathbb{P}\left(\bar{X} - \bar{Y} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}} < \mu_X - \mu_Y < \bar{X} - \bar{Y} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}}\right). \end{aligned}$$

Abbiamo dunque determinato l'intervallo di confidenza bilaterale di livello $1 - \alpha$

$$(L_i, L_s) = \left(\bar{X} - \bar{Y} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}}, \bar{X} - \bar{Y} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}}\right)$$

Esercizio 7.1.1. Dimostrare che

- $(-\infty, L_s) = \left(-\infty, \bar{X} - \bar{Y} + z_{1-\alpha} \sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}}\right)$ è un intervallo di confidenza inferiore di livello $1 - \alpha$;
- $(L_i, +\infty) = \left(\bar{X} - \bar{Y} - z_{1-\alpha} \sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}}, +\infty\right)$ è un intervallo di confidenza superiore di livello $1 - \alpha$.

7.1.2 Le varianze σ_X^2 e σ_Y^2 sono ignote ma si possono ritenere uguali

Indico con σ^2 il comune valore delle due varianze. Considero le varianze campionarie dei due campioni

$$S_X^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2, \quad S_Y^2 = \frac{1}{k-1} \sum_{j=1}^k (Y_j - \bar{Y})^2.$$

Sappiamo che $V_X := \frac{(n-1)S_X^2}{\sigma^2}$ segue la distribuzione χ_{n-1}^2 , e che $V_Y := \frac{(k-1)S_Y^2}{\sigma^2}$ segue la distribuzione χ_{k-1}^2 . Inoltre, poiché i due campioni sono indipendenti, anche V_X e V_Y sono indipendenti. Dunque, per il Teorema 3.3.2, $V_X + V_Y$ segue la distribuzione $\chi_{n+k-2}^2 = \chi_{n+k-2}^2$: $\mathbb{P}_{V_X+V_Y} = \chi_{n+k-2}^2$.

D'altra parte

$$V_X + V_Y = \frac{(n-1)S_X^2 + (k-1)S_Y^2}{\sigma^2} = \frac{n+k-2}{\sigma^2} \frac{(n-1)S_X^2 + (k-1)S_Y^2}{n+k-2}.$$

Se definiamo la statistica:

$$\bar{S}^2 := \frac{(n-1)S_X^2 + (k-1)S_Y^2}{n+k-2}.$$

abbiamo

$$V_X + V_Y = \frac{(n+k-2)\bar{S}^2}{\sigma^2}.$$

Inoltre sappiamo che $\bar{X} - \bar{Y}$ ha distribuzione $N\left(\mu_X - \mu_Y, \sigma^2 \left(\frac{1}{n} + \frac{1}{k}\right)\right)$, quindi

$$Z := \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sigma \sqrt{\frac{1}{n} + \frac{1}{k}}}$$

ha distribuzione gaussiana standard $N(0, 1)$. Considero

$$\begin{aligned} T &:= \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sigma \sqrt{\frac{1}{n} + \frac{1}{k}}} \frac{\sqrt{n+k-2}}{\sqrt{V_X + V_Y}} \\ &= \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{1}{n} + \frac{1}{k}}} \frac{\sqrt{n+k-2}}{\sqrt{(n-1)S_X^2 + (k-1)S_Y^2}} = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{1}{n} + \frac{1}{k}}} \frac{1}{\bar{S}}. \end{aligned}$$

Poiché i due campioni sono gaussiani e indipendenti le v.a. $\bar{X}$, S_X^2 , $\bar{Y}$ e S_Y^2 sono indipendenti, quindi $\bar{X} - \bar{Y}$ e $V_X + V_Y$ sono indipendenti, e dunque $\mathbb{P}_T = t_{n+k-2}$, vedi Teorema 3.3.8.

Abbiamo dunque

$$\begin{aligned} 1 - \alpha &= \mathbb{P} \left(t_{n+k-2, \frac{\alpha}{2}} < \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y) \frac{1}{\bar{S}}}{\sqrt{\frac{1}{n} + \frac{1}{k}}} < t_{n+k-2, 1-\frac{\alpha}{2}} \right) \\ &= \mathbb{P} \left(\bar{X} - \bar{Y} - t_{n+k-2, 1-\frac{\alpha}{2}} \bar{S} \sqrt{\frac{1}{n} + \frac{1}{k}} < \mu_X - \mu_Y < \bar{X} - \bar{Y} + t_{n+k-2, 1-\frac{\alpha}{2}} \bar{S} \sqrt{\frac{1}{n} + \frac{1}{k}} \right). \end{aligned}$$

Abbiamo dunque l'intervallo di confidenza di livello $1 - \alpha$

$$(L_i, L_s) = \left(\bar{X} - \bar{Y} - t_{n+k-2, 1-\frac{\alpha}{2}} \bar{S} \sqrt{\frac{1}{n} + \frac{1}{k}}, \bar{X} - \bar{Y} + t_{n+k-2, 1-\frac{\alpha}{2}} \bar{S} \sqrt{\frac{1}{n} + \frac{1}{k}} \right)$$

7.2 Test d'ipotesi per la differenza dei valori attesi di campioni gaussiani

Supponiamo di avere due campioni, entrambi gaussiani e tra di loro indipendenti

$$\begin{aligned} X: X_1, \dots, X_n & \quad \mathbb{P}_{X_i} = N(\mu_X, \sigma_X^2), \\ Y: Y_1, \dots, Y_k & \quad \mathbb{P}_{Y_j} = N(\mu_Y, \sigma_Y^2). \end{aligned}$$

Vogliamo testare

$$H_0: \quad \mu_X - \mu_Y = d \quad \quad H_A: \quad \mu_X - \mu_Y \neq d.$$

Osserviamo che $\mu_X - \mu_Y = d$ se e solo se $\mathbb{E}[\bar{X} - \bar{Y}] = d$.

Distinguiamo due diversi casi

7.2.1 Le varianze σ_X^2 e σ_Y^2 sono note

Sappiamo che $\mathbb{P}_{\bar{X}} = N\left(\mu_X, \frac{\sigma_X^2}{n}\right)$, $\mathbb{P}_{\bar{Y}} = N\left(\mu_Y, \frac{\sigma_Y^2}{k}\right)$. Considero la v.a. $W := \bar{X} - \bar{Y}$. Poiché i due campioni sono indipendenti, anche $\bar{X}$ e $\bar{Y}$ sono indipendenti, abbiamo che

$$\mathbb{P}_W = N\left(\mu_X - \mu_Y, \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}\right).$$

Dunque H_0 è vera se e solo se $\mathbb{P}_W = N\left(d, \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}\right)$. Stabilisco quindi il seguente criterio di accettazione:

$$\text{Accetto } H_0 \text{ se e solo se } |w - d| = |\bar{x} - \bar{y} - d| < \varepsilon.$$

La probabilità di commettere errore di prima specie vale allora

$$\alpha = \mathbb{P}(|W - d| \geq \varepsilon | \mu_X - \mu_Y = d) = \mathbb{P}\left(\frac{|W - d|}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}}} \geq \frac{\varepsilon}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}}} | \mu_X - \mu_Y = d\right)$$

D'altra parte, se H_0 è vera, allora $Z := \frac{W - d}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}}}$ ha distribuzione gaussiana standard $N(0, 1)$, e dunque dovremo scegliere $\frac{\varepsilon}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}}} = z_{1-\frac{\alpha}{2}}$ ovvero

$$\varepsilon = z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}}.$$

Dunque

accetto l'ipotesi H_0 se $|\bar{x} - \bar{y} - d| < z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}}$ e la rifiuto altrimenti.

Osservazione 7.2.1. Se $\sigma_X^2 = \sigma_Y^2 = \sigma_0^2$ e $k = n$, allora $\varepsilon = z_{1-\frac{\alpha}{2}} \sigma_0 \sqrt{\frac{2}{n}}$.

7.2.2 Le varianze σ_X^2 e σ_Y^2 sono ignote ma si possono ritenere uguali

Consideriamo le due varianze campionarie

$$S_X^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2, \quad S_Y^2 = \frac{1}{k-1} \sum_{j=1}^k (Y_j - \bar{Y})^2.$$

Indico con σ^2 il comune valore di σ_X^2 e σ_Y^2 . Sappiamo che $V_X := \frac{(n-1)S_X^2}{\sigma^2}$ segue la distribuzione χ_{n-1}^2 , e che $V_Y := \frac{(k-1)S_Y^2}{\sigma^2}$ segue la distribuzione χ_{k-1}^2 . Inoltre, poiché i due campioni sono indipendenti, anche V_X e V_Y sono indipendenti. Dunque, per il Teorema 3.3.2, $V_X + V_Y$ segue la distribuzione $\chi_{n-1+k-1}^2 = \chi_{n+k-2}^2$.

D'altra parte

$$V_X + V_Y = \frac{(n-1)S_X^2 + (k-1)S_Y^2}{\sigma^2} = \frac{n+k-2}{\sigma^2} \frac{(n-1)S_X^2 + (k-1)S_Y^2}{n+k-2}.$$

Se definiamo la statistica:

$$\bar{S}^2 := \frac{(n-1)S_X^2 + (k-1)S_Y^2}{n+k-2}.$$

abbiamo

$$V_X + V_Y = \frac{(n+k-2)\bar{S}^2}{\sigma^2}.$$

Inoltre sappiamo che $\bar{X} - \bar{Y}$ ha distribuzione $N\left(\mu_X - \mu_Y, \sigma^2 \left(\frac{1}{n} + \frac{1}{k}\right)\right)$, quindi

$$Z := \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sigma \sqrt{\frac{1}{n} + \frac{1}{k}}}$$

ha distribuzione gaussiana standard $N(0, 1)$. Considero

$$T := \frac{\bar{X} - \bar{Y} - d}{\sigma \sqrt{\frac{1}{n} + \frac{1}{k}}} \frac{\sqrt{n+k-2}}{\sqrt{V_X + V_Y}} = \frac{\bar{X} - \bar{Y} - d}{\sqrt{\frac{1}{n} + \frac{1}{k}}} \frac{\sqrt{n+k-2}}{\sqrt{(n-1)S_X^2 + (k-1)S_Y^2}}.$$

Poiché i due campioni sono gaussiani e indipendenti le v.a. $\bar{X}$, S_X^2 , $\bar{Y}$ e S_Y^2 sono indipendenti, quindi $\bar{X} - \bar{Y}$ e $V_X + V_Y$ sono indipendenti, e dunque $\mu_X - \mu_Y = d$ se e solo se $\mathbb{E}[T] = 0$. Infatti, per l'indipendenza, si ha

$$\mathbb{E}[T] = \frac{\mathbb{E}[\bar{X} - \bar{Y} - d]}{\sigma \sqrt{\frac{1}{n} + \frac{1}{k}}} \sqrt{n+k-2} \mathbb{E}\left[\frac{1}{\sqrt{V_X + V_Y}}\right].$$

Come criterio di accettazione per l'ipotesi nulla H_0 scelgo pertanto $|t| < \varepsilon$.

Inoltre, se H_0 è vera, allora per il Teorema 3.3.8 la v.a. T segue la distribuzione $t(n+k-2)$. La probabilità di commettere errore di prima specie è quindi $\alpha = \mathbb{P}(|T| \geq \varepsilon)$. Fissato il livello di significatività α , devo dunque scegliere $\varepsilon = t_{n+k-2, 1-\frac{\alpha}{2}}$.

Siano $x: x_1, \dots, x_n$ e $y: y_1, \dots, y_k$ i dati, $\bar{x}$ e $\bar{y}$ le rispettive medie, s_x^2 e s_y^2 le rispettive varianze:

$$\text{accetto } H_0 \text{ se } \frac{|\bar{x} - \bar{y} - d|}{\sqrt{\frac{1}{n} + \frac{1}{k}}} \frac{\sqrt{n+k-2}}{\sqrt{(n-1)s_x^2 + (k-1)s_y^2}} < t_{n+k-2, 1-\frac{\alpha}{2}}, \text{ e la rifiuto altrimenti.}$$

7.3 Test d'ipotesi per l'uguaglianza delle varianze di campioni gaussiani

Introduciamo prima una nuova distribuzione.

7.3.1 Distribuzione di Fisher-Snedecor a k e n gradi di libertà

Si può dimostrare che la funzione

$$f(x) = \begin{cases} \frac{\Gamma\left(\frac{k+n}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{n}{2}\right)} \left(\frac{k}{n}\right)^{\frac{k}{2}} \frac{x^{\frac{k}{2}-1}}{\left(1 + \frac{kx}{n}\right)^{\frac{k+n}{2}}} & x > 0, \\ 0 & x \leq 0. \end{cases}$$

è una densità di probabilità. La distribuzione assolutamente continua ad essa associata si dice *distribuzione di Fisher-Snedecor a k ed n gradi di libertà*, o semplicemente *distribuzione di Fisher a k ed n gradi di libertà*.

Si può dimostrare che se F è una variabile aleatoria con questa distribuzione, allora

$$\mathbb{E}[F] = \begin{cases} \frac{n}{n-2} & n > 2, \\ +\infty & n = 1, 2, \end{cases} \quad \text{Var}[F] = \begin{cases} \frac{2n^2(k+n-2)}{k(n-2)^2(n-4)} & n > 4, \\ +\infty & n = 3, 4, \\ \text{non esiste} & n = 1, 2. \end{cases}$$

Teorema 7.3.1. *Siano U e V variabili aleatorie indipendenti con distribuzioni $\mathbb{P}_U = \chi_k^2$, $\mathbb{P}_V = \chi_n^2$. Allora la v.a. $F := \frac{U/k}{V/n}$ segue la distribuzione di Fisher-Snedecor con k ed n gradi di libertà.*

Dimostrazione. Sappiamo che $\mathbb{P}_U = f(u)du$, $\mathbb{P}_V = g(v)dv$ dove

$$f(u) = \begin{cases} \frac{1}{\Gamma\left(\frac{k}{2}\right)} \left(\frac{1}{2}\right)^{\frac{k}{2}} u^{\frac{k}{2}-1} \exp\left(-\frac{u}{2}\right) & u > 0, \\ 0 & u \leq 0, \end{cases}$$

$$g(v) = \begin{cases} \frac{1}{\Gamma\left(\frac{n}{2}\right)} \left(\frac{1}{2}\right)^{\frac{n}{2}} v^{\frac{n}{2}-1} \exp\left(-\frac{v}{2}\right) & v > 0, \\ 0 & v \leq 0. \end{cases}$$

Possiamo scrivere $F = \varphi \circ (U, V)$ dove

$$\varphi: (u, v) \in \mathbb{R}^2 \mapsto \begin{cases} \frac{un}{kv} & v \neq 0, \\ 0 & v = 0. \end{cases}$$

Sia $\psi: \mathbb{R} \rightarrow \mathbb{R}$ una funzione di Borel non negativa. Abbiamo

$$\begin{aligned} \int_{\mathbb{R}} \psi(t) dt &= \int_{\mathbb{R}^2} \psi(\varphi(u, v)) \mathbb{P}_{U,V}(dudv) \\ &= \int_{(0,+\infty)^2} \psi\left(\frac{nu}{kv}\right) \frac{1}{\Gamma\left(\frac{k}{2}\right) \Gamma\left(\frac{n}{2}\right)} \left(\frac{1}{2}\right)^{\frac{k+n}{2}} u^{\frac{k}{2}-1} v^{\frac{n}{2}-1} \exp\left(-\frac{(u+v)}{2}\right) dudv \end{aligned}$$

sostituiamo $t = \frac{nu}{kv}$, $u = \frac{kv}{n}t$, $du = \frac{kv}{n}dt$

$$= \int_0^{+\infty} \psi(t) \frac{1}{\Gamma\left(\frac{k}{2}\right) \Gamma\left(\frac{n}{2}\right)} \left(\frac{1}{2}\right)^{\frac{k+n}{2}} \left(\frac{k}{n}\right)^{\frac{k}{2}} t^{\frac{k}{2}-1} \left(\int_0^{+\infty} y^{\frac{k+n}{2}-1} \exp\left(-\frac{v}{2} \left(1 + \frac{kt}{n}\right)\right) dv\right) dt$$

sostituiamo $x = \frac{v}{2} \left(1 + \frac{kt}{n}\right) = \frac{v}{2} \frac{n+kt}{n}$, $v = \frac{2nx}{n+kt}$, $dv = \frac{2n}{n+kt} dx$

$$\begin{aligned} &= \int_0^{+\infty} \psi(t) \frac{1}{\Gamma\left(\frac{k}{2}\right) \Gamma\left(\frac{n}{2}\right)} \left(\frac{1}{2}\right)^{\frac{k+n}{2}} \left(\frac{k}{n}\right)^{\frac{k}{2}} t^{\frac{k}{2}-1} \left(\int_0^{+\infty} \left(\frac{2n}{n+kt}\right)^{\frac{k+n}{2}} x^{\frac{k+n}{2}-1} e^{-x} dx\right) dt \\ &= \int_0^{+\infty} \psi(t) \frac{\Gamma\left(\frac{k+n}{2}\right)}{\Gamma\left(\frac{k}{2}\right) \Gamma\left(\frac{n}{2}\right)} \left(\frac{n}{n+kt}\right)^{\frac{k+n}{2}} \left(\frac{k}{n}\right)^{\frac{k}{2}} t^{\frac{k}{2}-1} dt \end{aligned}$$

da cui la tesi. □

Osservazione 7.3.1. Indichiamo con $f_{k,n,\alpha}$ il quantile di livello α associato alla distribuzione di Fisher di parametri k ed n . Siano U e V sono come nel Teorema 7.3.1: U e V variabili aleatorie indipendenti con distribuzioni $\mathbb{P}_U = \chi_k^2$, $\mathbb{P}_V = \chi_n^2$ e sia $\alpha \in (0, 1)$. Si ha

$$\begin{aligned}\alpha = \mathbb{P} \left(\frac{U/k}{V/n} \leq f_{k,n,\alpha} \right) &= \mathbb{P} \left(\left(\frac{U/k}{V/n} \right)^{-1} \geq \frac{1}{f_{k,n,\alpha}} \right) \\ &= \mathbb{P} \left(\frac{V/n}{U/k} \geq \frac{1}{f_{k,n,\alpha}} \right) = 1 - \mathbb{P} \left(\frac{V/n}{U/k} \leq \frac{1}{f_{k,n,\alpha}} \right)\end{aligned}$$

ovvero $\mathbb{P} \left(\frac{V/n}{U/k} \leq \frac{1}{f_{k,n,\alpha}} \right) = 1 - \alpha$ cioè $\frac{1}{f_{k,n,\alpha}} = f_{n,k,1-\alpha}$.

7.4 Test d'ipotesi per l'uguaglianza delle varianze di campioni gaussiani

Supponiamo di avere due campioni, entrambi gaussiani e tra di loro indipendenti

$$\begin{aligned}X: X_1, \dots, X_k & \quad \mathbb{P}_{X_i} = N(\mu_X, \sigma_X^2), \\ Y: Y_1, \dots, Y_n & \quad \mathbb{P}_{Y_j} = N(\mu_Y, \sigma_Y^2).\end{aligned}$$

Vogliamo testare

$$H_0: \quad \sigma_X^2 = \sigma_Y^2 \quad \quad H_A: \quad \sigma_X^2 \neq \sigma_Y^2.$$

Sappiamo che S_X^2 e S_Y^2 sono stimatori non distorti di σ_X^2 e σ_Y^2 , rispettivamente. Dunque:

accettiamo H_0 se $1 - \varepsilon_1 < \frac{S_X^2}{S_Y^2} < 1 + \varepsilon_2$, rifiutiamo altrimenti.

Per scegliere ε_1 ed ε_2 in base al livello di significatività desiderato, consideriamo le v.a.

$$V_X = \frac{(k-1)S_X^2}{\sigma_X^2}, \quad V_Y = \frac{(n-1)S_Y^2}{\sigma_Y^2}.$$

Sappiamo che $\mathbb{P}_{V_X} = \chi_{k-1}^2$, $\mathbb{P}_{V_Y} = \chi_{n-1}^2$. Dunque, la v.a. $\frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2}$ segue la distribuzione di

Fisher con $k-1$ ed $n-1$ gradi di libertà. In particolare H_0 è vera se e solo se $F := \frac{S_X^2}{S_Y^2}$ segue la distribuzione di Fisher con $k-1$ ed $n-1$ gradi di libertà.

Abbiamo dunque

$$\alpha = \mathbb{P} \left(\frac{S_X^2}{S_Y^2} \leq 1 - \varepsilon_1 | \sigma_X^2 = \sigma_Y^2 \right) + \mathbb{P} \left(\frac{S_X^2}{S_Y^2} \geq 1 + \varepsilon_2 | \sigma_X^2 = \sigma_Y^2 \right).$$

Scegliamo di *distribuire equamente l'errore* imponendo

$$\begin{aligned}\frac{\alpha}{2} &= \mathbb{P} \left(\frac{S_X^2}{S_Y^2} \leq 1 - \varepsilon_1 | \sigma_X^2 = \sigma_Y^2 \right) = \mathbb{P} (F \leq 1 - \varepsilon_1) \\ \frac{\alpha}{2} &= \mathbb{P} \left(\frac{S_X^2}{S_Y^2} \geq 1 + \varepsilon_2 | \sigma_X^2 = \sigma_Y^2 \right) = \mathbb{P} (F \geq 1 + \varepsilon_2) = 1 - \mathbb{P} (F \leq 1 + \varepsilon_2).\end{aligned}$$

Dovrà dunque essere $1 - \varepsilon_1 = f_{k-1, n-1, \frac{\alpha}{2}}$, $1 + \varepsilon_2 = f_{k-1, n-1, 1-\frac{\alpha}{2}}$. In definitiva:

accetto H_0 se $f_{k-1, n-1, \frac{\alpha}{2}} < \frac{S_X^2}{S_Y^2} < f_{k-1, n-1, 1-\frac{\alpha}{2}}$. Rifiuto altrimenti.

