Test d'ipotesi per il valore atteso di campioni gaussiani a varianza ignota

Test bilaterale

$X_1, \ldots, X_n$ un campione gaussiano di valore atteso μ e varianza σ^2 entrambi ignoti

- $H_0 : \qquad \mu = \mu_0 \qquad\qquad H_A : \qquad \mu \neq \mu_0$
- $\mathbb{E}\left[\frac{(\overline{X}-\mu_0)\sqrt{n}}{S}\right] = \mathbb{E}\left[\overline{X}-\mu_0\right]\sqrt{n}\,\mathbb{E}\left[\frac{1}{\sqrt{S^2}}\right] = 0$
  $\iff \mathbb{E}\left[\overline{X}\right] = \mu_0 \iff H_0$ è vera
- Considero $t := \frac{(\overline{x}-\mu_0)\sqrt{n}}{s}$:
- accetto H_0 se e solo se $|t| < \varepsilon$

- H_0 vera $\implies T := \dfrac{(\overline{X} - \mu_0)\sqrt{n}}{S}$ ha distribuzione $t(n-1)$.
- Livello di significatività

$$
\begin{aligned}
\alpha &= \mathbb{P}\left(|T| \geqslant \varepsilon\right) = \mathbb{P}\left(T \geqslant \varepsilon\right) + \mathbb{P}\left(T \leqslant -\varepsilon\right) \\
&= 1 - F_T\left(\varepsilon\right) + F_T\left(-\varepsilon\right) = 2\left(1 - F_T\left(\varepsilon\right)\right)
\end{aligned}
$$

$$
\iff F_T\left(\varepsilon\right) = 1 - \frac{\alpha}{2} \iff \varepsilon = t_{n-1,1-\frac{\alpha}{2}}
$$

Criterio di accettazione

Presi i dati $x_1, \ldots, x_n$, sia dunque $t_0 = \dfrac{(\overline{x} - \mu_0)\sqrt{n}}{s}$.
Accetto H_0 se e solo se $|t_0| < t_{n-1,1-\frac{\alpha}{2}}$ e la rifiuto altrimenti, ovvero

accetto H_0 se e solo se

$$\mu_0 - \frac{t_{n-1,1-\frac{\alpha}{2}}\, s}{\sqrt{n}} < \overline{x} < \mu_0 + \frac{t_{n-1,1-\frac{\alpha}{2}}\, s}{\sqrt{n}}$$

Test unilaterale inferiore con ipotesi nulla semplice

$X_1, \ldots, X_n$ un campione gaussiano di valore atteso μ e varianza σ^2 entrambi ignoti

- $H_0: \quad \mu = \mu_0, \qquad H_A: \quad \mu > \mu_0$
- Criterio di accettazione: accetto H_0 se e solo se $\dfrac{(\overline{x}-\mu_0)\sqrt{n}}{s} < \varepsilon$
- Livello di significatività:

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

  dove $\mathbb{P}_T = t(n-1)$.
- $\iff F_T(\varepsilon) = 1 - \alpha \iff \varepsilon = t_{n-1,1-\alpha}$

Criterio di accettazione

Presi i dati $x_1, \ldots, x_n$, sia dunque $t_0 = \frac{(\overline{x} - \mu_0)\sqrt{n}}{s}$. Accetto H_0 se e solo se $t_0 < t_{n-1,1-\alpha}$ ovvero

accetto H_0 se e solo se

$$\overline{x} \leqslant \mu_0 + \frac{t_{n-1,1-\alpha}\, s}{\sqrt{n}}$$

Test unilaterale inferiore con ipotesi nulla composta

$X_1, \ldots, X_n$ un campione gaussiano di valore atteso μ e varianza σ^2 entrambi incogniti

- $H_0: \quad \mu \leqslant \mu_0, \qquad H_A: \quad \mu > \mu_0$
- Criterio di accettazione: accetto H_0 se e solo se $\dfrac{(\overline{X}-\mu_0)\sqrt{n}}{S} < \varepsilon$.
- Livello di significatività

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

- H_0 vera $\implies \mathbb{E}\left[\overline{X}\right] = \mu \leqslant \mu_0 \implies$

$$\frac{(\overline{X}-\mu_0)\sqrt{n}}{S} \leqslant \frac{(\overline{X}-\mu)\sqrt{n}}{S} =: T, \qquad \mathbb{P}_T = t(n-1)$$

H_0 vera $(\mu \leqslant \mu_0) \implies$

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

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

$$\leqslant \mathbb{P}\left(\frac{(\overline{X} - \mu)\sqrt{n}}{S} > \varepsilon | \mathbb{E}\left[\overline{X}\right] = \mu \right) = \mathbb{P}\left(T > \varepsilon\right) = 1 - F_T(\varepsilon) = \alpha$$

- $\iff 1 - F_T(\varepsilon) = \alpha \iff \varepsilon = t_{n-1,1-\alpha}$

Criterio di accettazione

Presi i dati $x_1, \ldots, x_n$, sia $t_0 = \dfrac{(\overline{x} - \mu_0)\sqrt{n}}{s}$.
Accetto H_0 se e solo se $t_0 < t_{n-1,1-\alpha}$ ovvero

accetto H_0 se e solo se

$$\overline{x} \leqslant \mu_0 + \frac{t_{n-1,1-\alpha}\, s}{\sqrt{n}}$$

Test unilaterale superiore con ipotesi nulla semplice

$X_1, \ldots, X_n$ un campione gaussiano di valore atteso μ e varianza σ^2 entrambi incogniti

- $H_0: \quad \mu = \mu_0, \qquad H_A: \quad \mu < \mu_0$
- Criterio di accettazione: accetto H_0 se e solo se $\dfrac{(\bar{x} - \mu_0)\sqrt{n}}{s} > -\varepsilon$
- Livello di significatività

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

dove $\mathbb{P}_T = t(n-1)$

- $\iff F_T(-\varepsilon) = \alpha \iff \varepsilon = -t_{n-1,\alpha} = t_{n-1,1-\alpha}$

Criterio di accettazione

Presi i dati $x_1, \ldots, x_n$, sia dunque $t_0 = \dfrac{(\overline{x} - \mu_0)\sqrt{n}}{s}$.
Accetto H_0 se e solo se $t_0 > -t_{n-1,1-\alpha}$

ovvero accetto H_0 se e solo se

$$\overline{x} \geqslant \mu_0 - \frac{t_{n-1,1-\alpha}\, s}{\sqrt{n}}$$

Test unilaterale superiore con ipotesi nulla composta

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

- $H_0 : \quad \mu \geqslant \mu_0, \qquad H_A : \quad \mu < \mu_0$
- Criterio di accettazione: accetto H_0 se e solo se $\dfrac{(\bar{x} - \mu_0)\sqrt{n}}{s} \geqslant -\varepsilon.$
- Livello di significatività

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

- H_0 vera $\implies \mathbb{E}\left[\overline{X}\right] = \mu \geqslant \mu_0 \implies$

$$\frac{(\overline{X} - \mu_0)\sqrt{n}}{S} \geqslant \frac{(\overline{X} - \mu)\sqrt{n}}{S} =: T, \qquad \mathbb{P}_T = t(n-1)$$

UNIVERSITÀ DEGLI STUDI FIRENZE DIMAI Dipartimento di Matematica e Informatica "Ulisse Dini"

H_0 vera $\mu \geqslant \mu_0 \implies$

- $\left\{\frac{(\overline{X}-\mu_0)\sqrt{n}}{S} < -\varepsilon\right\} \subset \left\{\frac{(\overline{X}-\mu)\sqrt{n}}{S} < -\varepsilon\right\}$
- $$\begin{aligned}&\mathbb{P}\left(\frac{(\overline{X}-\mu_0)\sqrt{n}}{S} < -\varepsilon \middle| \mathbb{E}\left[\overline{X}\right] = \mu\right)\\ &\leqslant \mathbb{P}\left(\frac{(\overline{X}-\mu)\sqrt{n}}{S} < -\varepsilon \middle| \mathbb{E}\left[\overline{X}\right] = \mu\right)\\ &= \mathbb{P}\left(T < -\varepsilon\right) = F_T(-\varepsilon) = 1 - F_T(\varepsilon) = \alpha\end{aligned}$$
- $\iff 1 - F_T(\varepsilon) = \alpha \iff \varepsilon = t_{n-1,1-\alpha}$

Criterio di accettazione

Presi i dati $x_1, \ldots, x_n$, sia dunque $t_0 = \dfrac{(\overline{x} - \mu_0)\sqrt{n}}{s}$.
Accetto H_0 se $t_0 > -t_{n-1,1-\alpha}$, ovvero

accetto H_0 se e solo se

$$\overline{x} \geqslant \mu_0 - \frac{t_{n-1,1-\alpha}\, s}{\sqrt{n}}$$

Test d'ipotesi per la varianza di campioni gaussiani

Test bilaterale

$X_1, \ldots, X_n$ campione gaussiano di valore atteso μ (noto o incognito) e varianza σ^2 incognita.

- $H_0: \quad \sigma^2 = \sigma_0^2 \qquad\qquad H_A: \quad \sigma^2 \neq \sigma_0^2$
- H_0 vera $\iff \mathbb{E}\left[S^2\right] = \sigma_0^2 \iff \mathbb{E}\left[\frac{S^2}{\sigma_0^2}\right] = 1$
- Criterio di accettazione: accetto H_0 se e solo se

$$1 - \varepsilon_1 < \frac{s^2}{\sigma_0^2} < 1 + \varepsilon_2, \qquad \varepsilon_1, \varepsilon_2 > 0$$

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

- Livello di significatività:

$$V := \frac{(n-1)S^2}{\sigma_0^2}, \qquad H_0 \text{ vera } \iff \mathbb{P}_V = \chi^2_{n-1}$$

H_0 vera $\Longrightarrow$

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

Possibile scelta:

$$\mathbb{P}\left(V > (n-1)(1+\varepsilon_2)\right) = \frac{\alpha}{2} \qquad \text{cioè} \qquad (n-1)(1+\varepsilon_2) = \chi^2_{n-1, 1-\frac{\alpha}{2}}$$

$$\mathbb{P}\left(V < (n-1)(1-\varepsilon_1)\right) = \frac{\alpha}{2} \qquad \text{cioè} \qquad (n-1)(1-\varepsilon_1) = \chi^2_{n-1, \frac{\alpha}{2}}$$

Criterio di accettazione

Accetto H_0 se e solo se $\chi^2_{n-1,\frac{\alpha}{2}} < \dfrac{(n-1)s^2}{\sigma_0^2} < \chi^2_{n-1,1-\frac{\alpha}{2}}$ ovvero

accetto H_0 se e solo se

$$\frac{\sigma_0^2}{n-1}\chi^2_{n-1,\frac{\alpha}{2}} < s^2 < \frac{\sigma_0^2}{n-1}\chi^2_{n-1,1-\frac{\alpha}{2}}$$

Test unilaterale inferiore con ipotesi semplice

$X_1, \ldots, X_n$ campione gaussiano di valore atteso μ (noto o incognito) e varianza σ^2 incognita.

- $H_0 : \qquad \sigma^2 = \sigma_0^2 \qquad\qquad H_A : \qquad \sigma^2 > \sigma_0^2$
- Criterio di accettazione: accetto H_0 se e solo se $\dfrac{s^2}{\sigma_0^2} \leqslant 1 + \varepsilon$
- Livello di significatività:

$$V := \frac{(n-1)S^2}{\sigma_0^2}, \qquad H_0 \text{ vera} \iff \mathbb{P}_V = \chi^2_{n-1}$$

H_0 vera $\Longrightarrow$

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

$$\Longleftrightarrow \; 1 - F_V\left((n-1)(1+\varepsilon)\right) = \alpha \; \Longleftrightarrow \; (n-1)(1+\varepsilon) = \chi^2_{n-1,1-\alpha}$$

Criterio di accettazione

Accetto H_0 se e solo se $\dfrac{(n-1)s^2}{\sigma_0^2} < \chi^2_{n-1,1-\alpha}$ ovvero

acccetto H_0 se e solo se

$$s^2 < \frac{\sigma_0^2}{n-1}\chi^2_{n-1,1-\alpha}$$

Test unilaterale inferiore con ipotesi composta

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

- $H_0: \quad \sigma^2 \leqslant \sigma_0^2 \qquad\qquad H_A: \quad \sigma^2 > \sigma_0^2$
- Criterio di accettazione: $\dfrac{s^2}{\sigma_0^2} \leqslant 1 + \varepsilon$
- Livello di significatività:

$$V := \frac{(n-1)S^2}{\sigma^2}, \qquad \text{se } \operatorname{Var}[X_i] = \sigma^2 \implies \mathbb{P}_V = \chi^2_{n-1}$$

H_0 ($\mathrm{Var}[X_i] = \sigma^2 \leqslant \sigma_0^2$) vera $\implies$

$$
\begin{aligned}
&\mathbb{P}\left(\frac{S^2}{\sigma_0^2} > 1+\varepsilon \middle|\, \mathrm{Var}[X_i] = \sigma^2 \leqslant \sigma_0\right)\\
&= \mathbb{P}\left(\frac{(n-1)S^2}{\sigma^2} > \frac{\sigma_0^2}{\sigma^2}(n-1)(1+\varepsilon) \middle|\, \mathrm{Var}[X_i] = \sigma^2 \leqslant \sigma_0\right)\\
&= \mathbb{P}\left(V > \frac{\sigma_0^2}{\sigma^2}(n-1)(1+\varepsilon)\right) = 1 - F_V\left(\frac{\sigma_0^2}{\sigma^2}(n-1)(1+\varepsilon)\right)\\
&\leqslant 1 - F_V\left((n-1)(1+\varepsilon)\right) = {\color{red} = \alpha}
\end{aligned}
$$

$$(\sigma^2 \leqslant \sigma_0^2 \implies \frac{\sigma^2}{\sigma_0^2} \leqslant 1)$$

$$\iff 1 - F_V\left((n-1)(1+\varepsilon)\right) = \alpha \iff {\color{red}(n-1)(1+\varepsilon) = \chi^2_{n-1,1-\alpha}}$$

Criterio di accettazione

accetto H_0 se e solo se $\dfrac{(n-1)s^2}{\sigma_0^2} < \chi^2_{n-1,1-\alpha}$ ovvero

accetto H_0 se e solo se

$$s^2 < \frac{\sigma_0^2}{n-1}\chi^2_{n-1,1-\alpha}$$

Test unilaterale superiore con ipotesi semplice

$X_1, \ldots, X_n$ un campione gaussiano di valore atteso μ (noto o incognito) e varianza σ^2 incognita

- $H_0 : \quad \sigma^2 = \sigma_0^2 \qquad H_A : \quad \sigma^2 < \sigma_0^2$
- Criterio di accettazione: accetto H_0 se e solo se $\dfrac{s^2}{\sigma_0^2} \geqslant 1 - \varepsilon$
- Livello di significatività:

$$V := \frac{(n-1)S^2}{\sigma_0^2}, \qquad H_0 \text{ vera} \iff \mathbb{P}_V = \chi^2_{n-1}$$

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

$$\iff (n-1)(1-\varepsilon) = \chi^2_{n-1,\alpha}$$

Criterio di accettazione

Accetto H_0 se e solo se $\dfrac{(n-1)s^2}{\sigma_0^2} > \chi^2_{n-1,\alpha}$ ovvero

accetto H_0 se e solo se

$$s^2 > \frac{\sigma_0^2}{n-1}\chi^2_{n-1,\alpha}$$

Test unilaterale superiore con ipotesi composta

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

- $H_0: \quad \sigma^2 \geqslant \sigma_0^2 \qquad\qquad H_A: \quad \sigma^2 < \sigma_0^2$
- Criterio di accettazione: accetto H_0 se e solo se $\dfrac{s^2}{\sigma_0^2} \geqslant 1 - \varepsilon$
- Livello di significatività:

$$V := \frac{(n-1)S^2}{\sigma^2}, \qquad \text{se } \operatorname{Var}[X_i] = \sigma^2 \implies \mathbb{P}_V = \chi^2_{n-1}$$

H_0 ($\mathrm{Var}[X_i] = \sigma^2 \geqslant \sigma_0^2$) vera $\Longrightarrow$

$$
\begin{aligned}
&\mathbb{P}\left(\frac{S^2}{\sigma_0^2} < 1-\varepsilon \,\middle|\, \mathrm{Var}[X_i] = \sigma^2 \geqslant \sigma_0^2\right) \\
&= \mathbb{P}\left(\frac{(n-1)S^2}{\sigma^2} < \frac{\sigma_0^2}{\sigma^2}(n-1)(1-\varepsilon) \,\middle|\, \mathrm{Var}[X_i] = \sigma^2 \geqslant \sigma_0^2\right) \\
&= F_V\left(\frac{\sigma_0^2}{\sigma^2}(n-1)(1-\varepsilon)\right) \leqslant F_V\left((n-1)(1-\varepsilon)\right) = \alpha
\end{aligned}
$$

$$
\Longleftrightarrow \; F_V\left((n-1)(1-\varepsilon)\right) = \alpha \iff (n-1)(1-\varepsilon) = \chi^2_{n-1,\alpha}
$$

Criterio di accettazione

Accetto H_0 se e solo se $\dfrac{(n-1)s^2}{\sigma_0^2} > \chi^2_{n-1,\alpha}$ ovvero

accetto H_0 se se e solo se

$$s^2 > \frac{\sigma_0^2}{n-1}\chi^2_{n-1,\alpha}$$

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

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

Due campioni, entrambi gaussiani e tra di loro indipendenti

$$
\begin{aligned}
X\colon X_1, \ldots, X_n \qquad & \mathbb{P}_{X_i} = N\left(\mu_X, \sigma_X^2\right), \\
Y\colon Y_1, \ldots, Y_k \qquad & \mathbb{P}_{Y_j} = N\left(\mu_Y, \sigma_Y^2\right).
\end{aligned}
$$

Scopo

Costruire un intervallo di confidenza per la differenza $\mu_X - \mu_Y$.

Remark

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

Consideriamo due casi distinti (ma non esaustivi)

1: Le varianze σ_X^2 e σ_Y^2 sono note

- $\mathbb{P}_{X_i} = N\left(\mu_X, \sigma_X^2\right)$, $\quad \mathbb{P}_{Y_j} = N\left(\mu_Y, \sigma_Y^2\right) \implies$

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

- $Z := \dfrac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{k}}} \implies \mathbb{P}_Z = N(0,1)$

- $1 - \alpha = \mathbb{P}\left(z_{\frac{\alpha}{2}} < \dfrac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{k}}} < z_{1-\frac{\alpha}{2}} \right) =$

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

Intervallo di confidenza bilaterale di livello $1-\alpha$

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

Intervalli unilaterali

Analogamente si dimostra che

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

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

σ^2 := il comune valore delle due varianze.

- $S_X^2 = \dfrac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2, \qquad S_Y^2 = \dfrac{1}{k-1} \sum_{j=1}^{k} (Y_j - \overline{Y})^2$
- $V_X := \dfrac{(n-1)S_X^2}{\sigma^2}, \quad V_Y := \dfrac{(k-1)S_Y^2}{\sigma^2},$
- $\implies \mathbb{P}_{V_X} = \chi^2_{n-1}, \qquad \mathbb{P}_{V_Y} = \chi^2_{k-1}.$ V_X e V_Y sono indipendenti
- $\implies \mathbb{P}_{V_X+V_Y} = \chi^2_{n+k-2}$
- $V_X + V_Y = \dfrac{(n-1)S_X^2 + (k-1)S_Y^2}{\sigma^2}$

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

- $\overline{S}^2 := \dfrac{(n-1)S_X^2 + (k-1)S_Y^2}{n+k-2}$
- $V_X + V_Y = \dfrac{(n+k-2)\overline{S}^2}{\sigma^2}$
- $\mathbb{P}_{\overline{X}-\overline{Y}} = N\left(\mu_X - \mu_Y, \sigma^2\left(\dfrac{1}{n} + \dfrac{1}{k}\right)\right)$
- $Z := \dfrac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sigma\sqrt{\frac{1}{n} + \frac{1}{k}}} \implies \mathbb{P}_Z = N(0,1)$
- $T := \dfrac{Z\sqrt{n+k-2}}{\sqrt{V_X + V_Y}} = \dfrac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{1}{n} + \frac{1}{k}}} \dfrac{\sqrt{n+k-2}}{\sqrt{(n-1)S_X^2 + (k-1)S_Y^2}}$

$$= \frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{1}{n} + \frac{1}{k}}} \frac{1}{\overline{S}}.$$

- Campioni sono gaussiani e indipendenti
- $\Longrightarrow$ $\overline{X}$, S_X^2, $\overline{Y}$ e S_Y^2 indipendenti
- $\Longrightarrow$ $\overline{X}-\overline{Y}$ e V_X+V_Y indipendenti
- $\Longrightarrow$ $\Longrightarrow \mathbb{P}_T = t_{n+k-2}$

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

$$= \mathbb{P}\Big(\overline{X}-\overline{Y}-t_{n+k-2,1-\frac{\alpha}{2}}\overline{S}\sqrt{\frac{1}{n}+\frac{1}{k}} < \mu_X-\mu_Y <$$

$$< \overline{X}-\overline{Y}+t_{n+k-2,1-\frac{\alpha}{2}}\overline{S}\sqrt{\frac{1}{n}+\frac{1}{k}}\Big)$$

Intervallo di confidenza bilaterale di livello $1-\alpha$

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

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

Due campioni, entrambi gaussiani e tra di loro indipendenti

$$X\colon X_1, \ldots, X_n \qquad \mathbb{P}_{X_i} = N\left(\mu_X, \sigma_X^2\right),$$
$$Y\colon Y_1, \ldots, Y_k \qquad \mathbb{P}_{Y_j} = N\left(\mu_Y, \sigma_Y^2\right).$$

- $H_0: \qquad \mu_X - \mu_Y = d \qquad\qquad H_A: \qquad \mu_X - \mu_Y \neq d.$
- $\mu_X - \mu_Y = d \iff \mathbb{E}\left[\overline{X} - \overline{Y}\right] = d.$
- Distinguiamo due diversi casi

Le varianze σ_X^2 e σ_Y^2 sono note

- $\mathbb{P}_{\overline{X}} = N\left(\mu_X, \frac{\sigma_X^2}{n}\right)$, $\mathbb{P}_{\overline{Y}} = N\left(\mu_Y, \frac{\sigma_Y^2}{k}\right)$
- $W := \overline{X} - \overline{Y} \implies \mathbb{P}_W = N\left(\mu_X - \mu_Y, \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}\right)$
- H_0 vera $\iff \mathbb{P}_W = N\left(d, \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}\right)$
- Criterio di accettazione: accetto H_0 iff $|w - d| = |\overline{x} - \overline{y} - d| < \varepsilon$.
- Livello di significatività:

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

- $Z := \dfrac{W - d}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{k}}}$
- H_0 vera, $\implies \mathbb{P}_Z = N(0,1)$
- Livello di significatività α se $\dfrac{\varepsilon}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{k}}} = z_{1-\frac{\alpha}{2}} \iff$

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

Criterio di accettazione

Accetto l'ipotesi H_0 se e solo se

$$|\overline{x} - \overline{y} - d| < z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}}$$

Remark

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

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

$\sigma^2 :=$ il comune valore di σ_X^2 e σ_Y^2.

- $S_X^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_i - \overline{X})^2, \qquad S_Y^2 = \dfrac{1}{k-1}\displaystyle\sum_{j=1}^{k}(Y_j - \overline{Y})^2$
- $V_X := \dfrac{(n-1)S_X^2}{\sigma^2}$, $V_Y := \dfrac{(k-1)S_Y^2}{\sigma^2}$
- $\mathbb{P}_{V_X} = \chi^2_{n-1}, \quad \mathbb{P}_{V_Y} = \chi^2_{k-1}$, V_X e V_Y sono idipendenti
- $\mathbb{P}_{V_X+V_Y} = \chi^2_{n-1+k-1} = \chi^2_{n+k-2}$
- $V_X + V_Y = \dfrac{(n-1)S_X^2 + (k-1)S_Y^2}{\sigma^2} =$
  $\dfrac{n+k-2}{\sigma^2}\,\dfrac{(n-1)S_X^2 + (k-1)S_Y^2}{n+k-2}$

- $\overline{S}^2 := \dfrac{(n-1)S_X^2 + (k-1)S_Y^2}{n+k-2}$
- $\implies V_X + V_Y = \dfrac{(n+k-2)\overline{S}^2}{\sigma^2}$
- $\mathbb{P}_{\overline{X}-\overline{Y}} = N\left(\mu_X - \mu_Y, \sigma^2\left(\dfrac{1}{n} + \dfrac{1}{k}\right)\right)$
- $Z := \dfrac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sigma\sqrt{\frac{1}{n} + \frac{1}{k}}}, \implies \mathbb{P}_Z = N(0,1)$
- $T := \dfrac{\overline{X} - \overline{Y} - d}{\sigma\sqrt{\frac{1}{n} + \frac{1}{k}}} \dfrac{\sqrt{n+k-2}}{\sqrt{V_X + V_Y}} =$
  $\dfrac{\overline{X} - \overline{Y} - d}{\sqrt{\frac{1}{n} + \frac{1}{k}}} \dfrac{\sqrt{n+k-2}}{\sqrt{(n-1)S_X^2 + (k-1)S_Y^2}}.$

- $\overline{X}$, S_X^2, $\overline{Y}$ e S_Y^2 sono indipendenti
- $\overline{X} - \overline{Y}$ e $V_X + V_Y$ sono indipendenti
- $\mu_X - \mu_Y = d$ se e solo se e $\mathbb{E}\left[T\right] = 0$
  Dim: Per l'indipendenza abbiamo

$$\mathbb{E}\left[T\right] = \frac{\mathbb{E}\left[\overline{X} - \overline{Y} - d\right]}{\sigma\sqrt{\frac{1}{n} + \frac{1}{k}}} \sqrt{n+k-2}\,\mathbb{E}\left[\frac{1}{\sqrt{V_X + V_Y}}\right] = 0$$

  se e solo se $\mathbb{E}\left[\overline{X} - \overline{Y} - d\right] = 0$
- Criterio di accettazione: accetto H_0 se e solo se $|t| < \varepsilon$

- Livello di significatività: H_0 vera $\implies \mathbb{P}_T = t(n+k-2)$
- $\alpha = \mathbb{P}\left(|T| \geqslant \varepsilon\right) \iff \varepsilon = t_{n+k-2,1-\frac{\alpha}{2}}$

Criterio di accettazione

x: $x_1, \ldots, x_n$ e y: $y_1, \ldots, y_k$ dati, $\overline{x}$ e $\overline{y}$ le rispettive medie, s_x^2 e s_y^2 le rispettive varianze:

accetto H_0 se e solo se

$$\frac{|\overline{x}-\overline{y}-d|}{\sqrt{\dfrac{1}{n}+\dfrac{1}{k}}}\,\frac{\sqrt{n+k-2}}{\sqrt{(n-1)s_X^2+(k-1)s_Y^2}} < t_{n+k-2,1-\frac{\alpha}{2}}$$

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

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

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**

F variabile aleatoria con questa distribuzione,

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

Theorem

Siano U e V v.a. indipendenti con distribuzioni $\mathbb{P}_U = \chi^2_k$, $\mathbb{P}_V = \chi^2_n$. Allora la v.a. $F := \dfrac{U/k}{V/n}$ segue la distribuzione di Fisher-Snedecor con k ed n gradi di libertà.

Quantili

$f_{k,n,\alpha} :=$ quantile di livello α associato alla distribuzione di Fisher di parametri k ed n.
U e V v.a. indipendenti con $\mathbb{P}_U = \chi^2_k$, $\mathbb{P}_V = \chi^2_n$
$\alpha \in (0, 1)$

$$\alpha = \mathbb{P}\left(\frac{U/k}{V/n} \leqslant f_{k,n,\alpha}\right) = \mathbb{P}\left(\left(\frac{U/k}{V/n}\right)^{-1} \geqslant \frac{1}{f_{k,n,\alpha}}\right)$$

$$= \mathbb{P}\left(\frac{V/n}{U/k} \geqslant \frac{1}{f_{k,n,\alpha}}\right) = 1 - \mathbb{P}\left(\frac{V/n}{U/k} \leqslant \frac{1}{f_{k,n,\alpha}}\right)$$

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

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

Due campioni gaussiani e tra di loro indipendenti

$$X\colon X_1, \ldots, X_k \qquad \mathbb{P}_{X_i} = N\left(\mu_X, \sigma_X^2\right),$$
$$Y\colon Y_1, \ldots, Y_n \qquad \mathbb{P}_{Y_j} = N\left(\mu_Y, \sigma_Y^2\right).$$

- $H_0: \qquad \sigma_X^2 = \sigma_Y^2 \qquad\qquad H_A: \qquad \sigma_X^2 \neq \sigma_Y^2$
- $\mathbb{E}\left[S_X^2\right] = \sigma_X^2 \quad \mathbb{E}\left[S_Y^2\right] = \sigma_Y^2$
- Criterio di accettazione: accetto H_0 se e solo se $1 - \varepsilon_1 < \dfrac{s_X^2}{s_Y^2} < 1 + \varepsilon_2$
- Livello di significatività:

$$V_X = \frac{(k-1)S_X^2}{\sigma_X^2} \quad \mathbb{P}_{V_X} = \chi^2_{k-1} \qquad V_Y = \frac{(n-1)S_Y^2}{\sigma_Y^2} \quad \mathbb{P}_{V_Y} = \chi^2_{n-1}$$

- $\dfrac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2}$ ha distribuzione di Fisher con $k-1$ ed $n-1$ gradi di libertà
- $\implies H_0$ vera $\iff F := \dfrac{S_X^2}{S_Y^2}$ ha distribuzione di Fisher con $k-1$ ed $n-1$ gradi di libertà.
- $\alpha = \mathbb{P}\left(\dfrac{S_X^2}{S_Y^2} \leqslant 1-\varepsilon_1 | \sigma_X^2 = \sigma_Y^2\right) + \mathbb{P}\left(\dfrac{S_X^2}{S_Y^2} \geqslant 1+\varepsilon_2 | \sigma_X^2 = \sigma_Y^2\right)$.
- Scegliamo di *distribuire equamente l'errore* imponendo

$$\frac{\alpha}{2} = \mathbb{P}\left(\frac{S_X^2}{S_Y^2} \leqslant 1-\varepsilon_1 | \sigma_X^2 = \sigma_Y^2\right) = \mathbb{P}\left(F \leqslant 1-\varepsilon_1\right)$$

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

$\iff 1-\varepsilon_1 = f_{k-1,n-1,\frac{\alpha}{2}},\ 1+\varepsilon_2 = f_{k-1,n-1,1-\frac{\alpha}{2}}$

Criterio di accettazione

Accetto H_0 se e solo se

$$f_{k-1,n-1,\frac{\alpha}{2}} < \frac{s_X^2}{s_Y^2} < f_{k-1,n-1,1-\frac{\alpha}{2}}$$