10 Verifica delle ipotesi

10.1 Introduzione

La verifica delle ipotesi è strettamente connessa al problema della stima. Si consideri un c.c.$(X_1, ..., X_n)$ estratto da una popolazione con funzione di densità $f(\cdot; \theta)$. Ora non ci preoccupiamo più di stimare il parametro incognito θ, ma utilizziamo il campione per verificare qualche ipotesi che coinvolga il parametro.

Definizione: Un'ipotesi statistica è un'affermazione sul parametro della distribuzione della popolazione che non sappiamo se sia vera o falsa.

Definizione: Un test è una procedura per determinare se i valori del campione e l'ipotesi sono compatibili.

In molti problemi di verifica, due sono le ipotesi che vengono discusse:

IPOTESI NULLA : H_0

IPOTESI ALTERNATIVA : H_1

L'ipotesi nulla H_0 può essere:

1. SEMPLICE $\quad H_0 : \theta = 1$

(specifica completamente la distribuzione)

2. COMPOSTA $\quad H_0 : \theta \leq 1$ o $(\theta \geq 1)$

Il test è definito da una regione C detta regione critica del test:

accetta H_0 se $(X_1, ...X_n) \notin C$

rifiuta H_0 se $(X_1, ...X_n) \in C$.

In qualunque test per verificare un'ipotesi nulla il risultato può essere sbagliato in due modi.

Errore di 1^a specie : rifiutare H_0 quando è vera

Errore di 2^a specie : accettare H_0 quando è falsa

Quando H_0 è vera, la probabilità che venga rifiutata non deve superare un certo valore α detto livello di significatività del test (fissato in anticipo, p.es. 0.1, 0.05, 0.005).

α = P (errore di 1^a specie)

β = P (errore di 2^a specie)

Decisione	H_0 vera	H_0 falsa
Non rifiutare H_0	Decisione corretta Confidenza $= 1 - \alpha$	Errore di seconda specie P(err 2^a sp) $= \beta$
Rifiutare H_0	Errore di prima specie = livello di significatività P(err 1^a sp) $= \alpha$	Decisione corretta Potenza $= 1 - \beta$

Tra tutti i test che possiedono un errore di 1^a specie con lo stesso livello di significatività (ampiezza), si sceglie quello con ampiezza d'errore di 2^a specie più piccolo.

Ci sono test per ipotesi semplici e test per ipotesi composte e vari sono i metodi per la determinazione di un test. Noi esaminiamo quello chiamato metodo dell'intervallo di confidenza, che consiste nell'usare un intervallo di confidenza per ottenere un test.

Molti problema di verifica di ipotesi riguardano i parametri (media, varianza) di distribuzioni normali.

- TEST SULLA MEDIA
  1) se la varianza è nota (test Z)
  2) se la varianza non è nota (test t di Student)
- TEST SULLA VARIANZA
  1) se la media non è nota (test χ^2)
  2) se la media è nota (test χ^2)

10.2 Test sulla media di una popolazione normale quando la varianza è nota

10.2.1 Test bilaterali o a due code

Dato un c.c. $(X_1, ...X_n)$ estratto da una popolazione normale $N(\mu, \sigma^2 \text{ nota})$ e fissato un certo valore μ_0, si verifichi

$$\text{l'ipotesi nulla } H_0 : \mu = \mu_0$$

contro

$$\text{l'ipotesi alternativa } H_1 : \mu \neq \mu_0$$

La media campionaria $\overline{X}_n$ è lo stimatore puntuale per la media incognita μ della popolazione. Perciò H_0 si accetta quando $\overline{X}_n$ non si discosta troppo da μ_0. Allora definiamo la regione critica C

$$C = \{(X_1, ..., X_n) : |\overline{X}_n - \mu_0| > c\}$$

con c costante opportuna.

Se si desidera che α sia il livello di significatività del test, bisogna calcolare c in modo che

$$\alpha = P(\text{errore di } 1^{\text{a}} \text{ specie}) = P(|\overline{X}_n - \mu_0| > c)$$

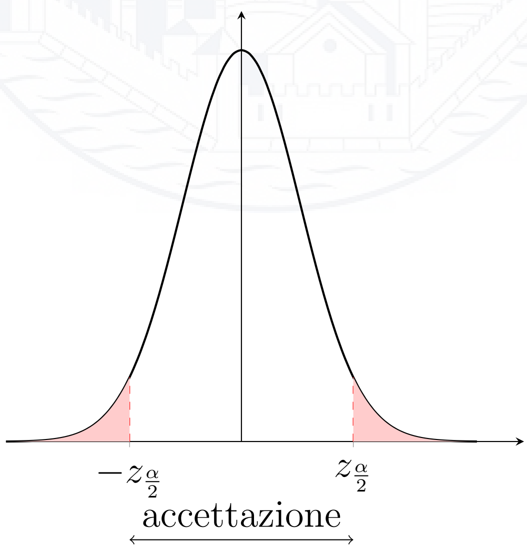
Poichè $\overline{X}_n \sim N(\mu, \frac{\sigma^2}{n}) \Rightarrow Z \sim N(0, 1)$ dove
$Z = \frac{\overline{X}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}}$
si ha:

$$\alpha = \mathrm{P}\left(\frac{|\overline{X}_n - \mu_0|}{\frac{\sigma}{\sqrt{n}}} > \frac{c\sqrt{n}}{\sigma}\right) = \mathrm{P}\left(|Z| > \frac{c\sqrt{n}}{\sigma}\right)$$
$$= 2\,\mathrm{P}\left(Z > \frac{c\sqrt{n}}{\sigma}\right)$$

e quindi $\mathrm{P}(Z > \frac{c\sqrt{n}}{\sigma}) = \frac{\alpha}{2}$
ma $\mathrm{P}(Z > z_{\frac{\alpha}{2}}) = \frac{\alpha}{2} \Rightarrow z_{\frac{\alpha}{2}} = \frac{c\sqrt{n}}{\sigma}$
cioè $c = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$
REGOLA:

si rifiuta H_0 se $|\frac{\overline{X}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}}| > z_{\frac{\alpha}{2}}$

si accetta H_0 se $|\frac{\overline{X}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}}| \leq z_{\frac{\alpha}{2}}$



Esempio

A Sorgente
segnale di valore μ

B Ricevente
segnale ricevuto con rumore $\sim N(\mu, 4)\,(\sigma = 2)$

Per ridurre il rumore il segnale viene inviato 5 volte. La media campionaria dei 5 segnali ricevuti è $\overline{X}_n = 9.5$.
B ha motivo di credere che il valore inviato sia $\mu = 8$. Verificare tale ipotesi:

- al livello del 5%

  $H_0 : \mu = \mu_0 = 8 \qquad H_1 : \mu \neq 8$

  $\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025$

  $\frac{|\overline{X}_n - \mu_0|}{\frac{\sigma}{\sqrt{n}}} = \frac{9.5-8}{\frac{2}{\sqrt{5}}} = \frac{\sqrt{5}}{2} \cdot (1.5) = \underline{1.68}$

  $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$

  $1.68 < 1.96 \Rightarrow$ si accetta H_0.

- al livello del 10%

  $\alpha = 0.1 \Rightarrow \frac{\alpha}{2} = 0.05$

$z_{\frac{\alpha}{2}} = z_{0.05} = 1.645$
$1.68 > 1.645 \Rightarrow$ si rifiuta H_0.

Problema: Qual è il livello "giusto" da scegliere?

Prima si calcola il valore della statistica del test cioè $\frac{|\overline{X}_n - \mu_0|}{\frac{\sigma}{\sqrt{n}}}$, poi la probabilità che una normale standard, in valore assoluto, superi tale quantità. Questa probabilità, detta il "p-dei-dati" del test fornisce il livello di significatività critico al di sotto del quale la decisione cambia da rifiuto ad accettazione.

$$\text{Se il p-dei-dati} \gg \alpha \text{ allora si accetta } H_0$$
$$\ll \alpha \text{ si rifiuta } H_0.$$

Esempio (precedente)

- se $\overline{X}_n = 8.5$
  $\frac{|\overline{X}_n - \mu_0|}{\frac{\sigma}{\sqrt{n}}} = \frac{8.5-8}{\frac{2}{\sqrt{5}}} = \frac{\sqrt{5}}{2} \cdot 0.5 \approx 0.559$
  $P(|Z| > 0.559) = 2\,P(Z > 0.559) = 2 \cdot 0.2877 = 0.5754$.

  Il valore critico è altissimo $\Rightarrow H_0$ si accetta.

- se $\overline{X}_n = 11.5$
  $\frac{|\overline{X}_n - \mu_0|}{\frac{\sigma}{\sqrt{n}}} \approx 3.91$

$\mathrm{P}(|Z| > 3.91) = 2\,\mathrm{P}(Z > 3.91) = 0.0001$
Il valore critico è piccolissimo $\Rightarrow$ $\underline{H_0 \text{ si rifiuta}}$.

La probabilità di un errore di 2^a specie (accetto H_0 quando è falsa) dipende da μ.

$$\begin{aligned}
\beta(\mu) &= \mathrm{P}(\text{accettare } H_0) \text{ quando la media reale è } \mu. \\
&= \mathrm{P}\Big(\frac{|\overline{X}_n - \mu_0|}{\frac{\sigma}{\sqrt{n}}} \leq z_{\frac{\alpha}{2}}\Big) \\
&= \mathrm{P}\Big(-z_{\frac{\alpha}{2}} \leq \frac{\overline{X}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}} \leq z_{\frac{\alpha}{2}}\Big) \\
&= \mathrm{P}\Big(\frac{\mu_0 - \mu}{\frac{\sigma}{\sqrt{n}}} - z_{\frac{\alpha}{2}} \leq Z \leq \frac{\mu_0 - \mu}{\frac{\sigma}{\sqrt{n}}} + z_{\frac{\alpha}{2}}\Big) \\
&= \Phi\Big(\frac{\mu_0 - \mu}{\frac{\sigma}{\sqrt{n}}} + z_{\frac{\alpha}{2}}\Big) - \Phi\Big(\frac{\mu_0 - \mu}{\frac{\sigma}{\sqrt{n}}} - z_{\frac{\alpha}{2}}\Big)
\end{aligned}$$

dove $Z = \frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$.

<u>Esempio</u> (precedente)
Quanto vale la probabilità di accettare $\mu = 8\,(= \mu_0)$ quando in realtà $\mu = 10$ con un l.d.s. $\alpha = 0.05$?

$z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$

$$\begin{aligned}\beta(10) &= \Phi(-\sqrt{5} + 1.96) - \Phi(-\sqrt{5} - 1.96) \\ &= \Phi(-0.276) - \Phi(-4.196) \\ &= \cancel{1} - \Phi(0.276) - \cancel{1} + \Phi(4.196) \approx 1 - 0.60835 \\ &\simeq 0.39165\end{aligned}$$

La funzione $1 - \beta(\mu)$ è detta funzione di potenza del test.

Per μ fissato la potenza del test è la probabilità di rifiutare (correttamente) l'ipotesi nulla H_0 quando μ è il valore vero.

Quesito

Calcolare il valore di n con il quale la probabilità di accettare $H_0 : \mu = \mu_0$ quando invece il valore vero è μ_1, sia pari ad un valore fissato $\bar{\beta}$ cioè:

$$\beta(\mu_1) = \bar{\beta}$$

Devo risolvere:

$$\Phi(\frac{\mu_0 - \mu_1}{\frac{\sigma}{\sqrt{n}}} + z_{\frac{\alpha}{2}}) - \Phi(\frac{\mu_0 - \mu_1}{\frac{\sigma}{\sqrt{n}}} - z_{\frac{\alpha}{2}}) = \bar{\beta}$$

Da calcoli (laboriosi) si ottiene

$$\boxed{n \approx \left[\frac{(z_{\frac{\alpha}{2}} + z_{\bar{\beta}})\sigma}{\mu_1 - \mu_0}\right]^2}$$

Esempio (precedente)
Quante volte bisogna inviare il segnale perchè la verifica di $H_0 : \mu = 8$ con $\alpha = 0.05$ abbia il 75% di probabilità di rifiutare H_0 quando $\mu = 9.2$?

$\alpha = 0.05 \quad \frac{\alpha}{2} = 0.025\,, \quad z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$

$\underline{1 - \beta(\mu) = 0.75}$

$\bar{\beta} = 0.25\,, \quad z_{\bar{\beta}} = z_{0.25} = 0.67$

$\mu_1 = 9.2 \quad \mu_0 = 8$

$$\Rightarrow n \approx \left(\frac{1.96 + 0.67}{1.2}\right)^2 \cdot 4 \approx 19.21$$

$$\Rightarrow \underline{n = 20}$$

Per $n = 20$

$$\begin{aligned}\beta &= \Phi\left(-\frac{1.2}{\frac{2}{\sqrt{20}}} + 1.96\right) - \Phi\left(-\frac{1.2}{\frac{2}{\sqrt{20}}} - 1.96\right) \\ &= \Phi(-0.723) - \Phi(-4.623) \\ &\approx 0.235\end{aligned}$$

$$\Rightarrow 1 - \beta(9.2) \approx 0.765\,, \text{ cioè } 76.5\%.$$

10.2.2 Test unilaterali o ad una coda

- $H_0 : \mu = \mu_0$ contro $H_1 : \mu > \mu_0$
  $(\mu \leq \mu_0)$

Si rifiuta l'ipotesi nulla quando $\overline{X}_n$, stimatore di μ è molto più grande di μ_0.
La regione critica è così definita:

$$C = (X_1, ..., X_n) : \overline{X}_n - \mu_0 > c$$

con c tale che $\alpha = \mathrm{P}(\overline{X}_n - \mu_0 > c)$
Se $\mu = \mu_0$, $\overline{X}_n \sim N(\mu_0, \frac{\sigma^2}{n})$ e $Z = \frac{\overline{X}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$
perciò:

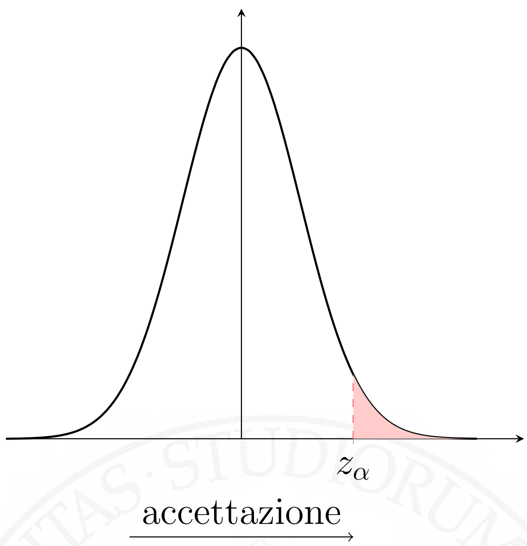
$$\alpha = \mathrm{P}(Z > c\frac{\sqrt{n}}{\sigma}), \text{ ma } \mathrm{P}(Z > z_\alpha) = \alpha$$

$$\Rightarrow \boxed{c = z_\alpha \frac{\sigma}{\sqrt{n}}}$$

Per il test con livello di significatività α vale REGOLA:

$$\text{si rifiuta } H_0 \text{ se } \frac{\overline{X}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}} > z_\alpha$$

$$\text{si accetta } H_0 \text{ se } \frac{\overline{X}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}} \leq z_\alpha$$



$$\beta(\mu) = \mathrm{P}(\text{accettare } H_0) = \mathrm{P}\left(\frac{\overline{X}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}} \leq z_\alpha\right)$$

$$= \mathrm{P}\left(\frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \leq \frac{\mu_0 - \mu}{\frac{\sigma}{\sqrt{n}}} + z_\alpha\right)$$

$$= \mathrm{P}\left(Z \leq \frac{\mu_0 - \mu}{\frac{\sigma}{\sqrt{n}}} + z_\alpha\right)$$

$$= \Phi\left(\frac{\mu_0 - \mu}{\frac{\sigma}{\sqrt{n}}} + z_\alpha\right)$$

Poichè $\Phi(z_\alpha) = 1 - \alpha \Rightarrow \beta(\mu_0) = 1 - \alpha$.

Esempio (precedente)

Si sa che il segnale inviato è superiore a 8. I dati sono compatibili con l'ipotesi che la media è 8?

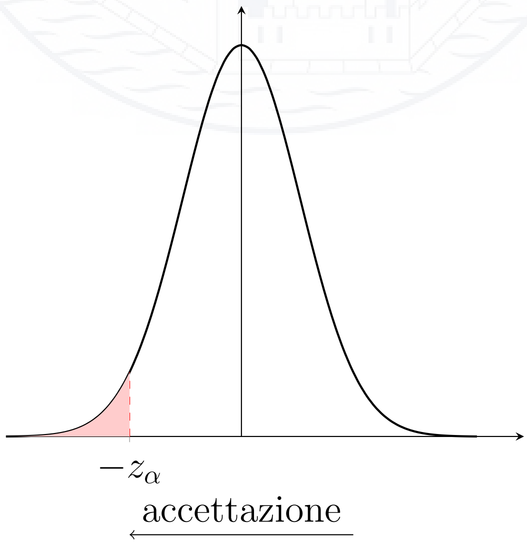
$H_0 : \mu = 8$ contro $H_1 : \mu > 8$ (alternativa a una coda)
Se $\overline{X}_n = 9.5\,,\ \frac{\overline{X}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}} = 1.68$
$\alpha = 0.05 \quad z_\alpha = z_{0.05} = 1.645$
poichè $1.68 > 1.645 \Rightarrow H_0$ va rifiutata.

- $H_0 : \mu = \mu_0$ $(\mu \geq \mu_0)$ contro $H_1 : \mu < \mu_0$

Al livello di significatività α si ha
REGOLA

$$\text{si rifiuta } H_0 \text{ se } \frac{\overline{X}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}} < -z_\alpha$$

$$\text{si accetta } H_0 \text{ se } \frac{\overline{X}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}} \geq -z_\alpha$$



Osservazione: l'analogia tra la stima di parametri attraverso gli intervalli di confidenza e la verifica delle ipotesi è evidente.

10.3 Test sulla media di una popolazione normale quando la varianza non è nota

10.3.1 Test bilaterali o a due code

Sia dato un c.c. $X_1, ..., X_n$ estratto da un popolazione normale N con μ e σ^2 incognite.

$H_0 : \mu = \mu_0$ contro $H_1 : \mu \neq \mu_0$
$\downarrow$
Non è ipotesi semplice (non fornisce il valore di σ^2). Poichè σ^2 non è nota, si usa come statistica la deviazione standard campionaria

$$S = \sqrt{\frac{1}{n-1} \sum_i (X_i - \overline{X}_n)^2}$$

e sappiamo che $T_{n-1} = \frac{\overline{X}_n - \mu_0}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$

Si rifiuta l'ipotesi nulla H_0 quando $|\frac{\overline{X}_n - \mu_0}{\frac{S}{\sqrt{n}}}| > c$ con c da determinare in funzione di α.

$$\begin{aligned}\alpha &= \mathrm{P}(\text{errore di } 1^a \text{ specie}) \\ &= \mathrm{P}(|T_{n-1}| > c) = \mathrm{P}(T_{n-1} < -c) + \mathrm{P}(T_{n-1} > c) \\ &= 2\,\mathrm{P}(T_{n-1} > c)\end{aligned}$$

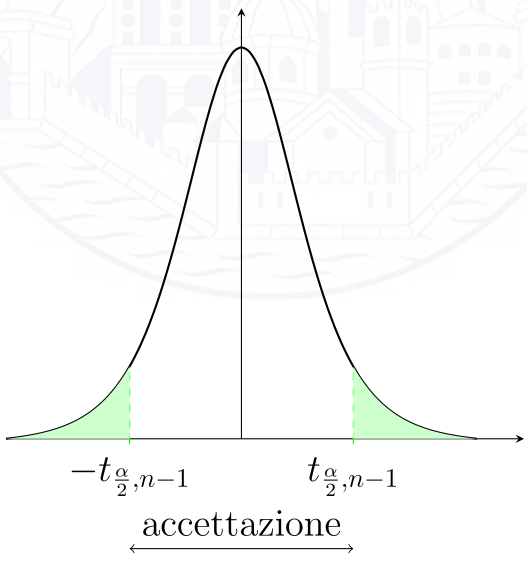
$\Rightarrow \mathrm{P}(T_{n-1} > c) = \frac{\alpha}{2}$, ma $\mathrm{P}(T_{n-1} > t_{\frac{\alpha}{2},n-1}) = \frac{\alpha}{2}$

perciò: $\boxed{c = t_{\frac{\alpha}{2},n-1}}$

REGOLA:

si rifiuta H_0 se $\left|\frac{\overline{X}_n-\mu_0}{\frac{S}{\sqrt{n}}}\right| > t_{\frac{\alpha}{2},n-1}$

si accetta H_0 se $\left|\frac{\overline{X}_n-\mu_0}{\frac{S}{\sqrt{n}}}\right| \leq t_{\frac{\alpha}{2},n-1}$



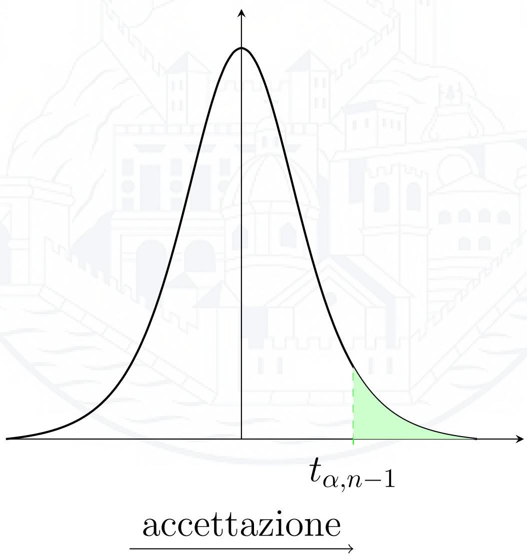
10.3.2 Test unilaterali o ad una coda

- $H_0 : \mu = \mu_0$ contro $H_1 : \mu > \mu_0$
  $(\mu \leq \mu_0)$

al livello di significatività α.
REGOLA:

$$\text{si rifiuta } H_0 \text{ se } \frac{\overline{X}_n - \mu_0}{\frac{S}{\sqrt{n}}} > t_{\alpha,\, n-1}$$

$$\text{si accetta } H_0 \text{ se } \frac{\overline{X}_n - \mu_0}{\frac{S}{\sqrt{n}}} \leq t_{\alpha,\, n-1}$$

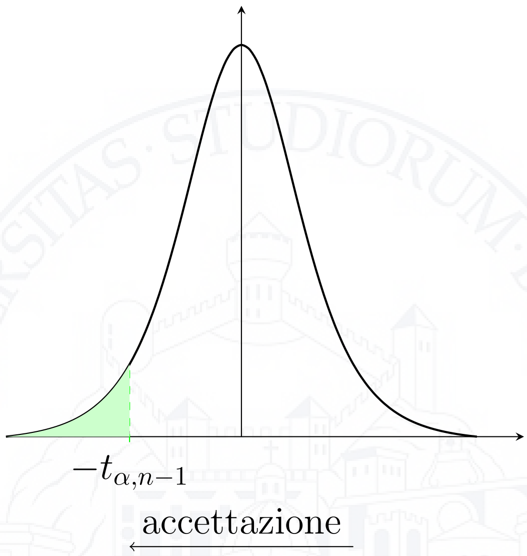


- $H_0 : \mu = \mu_0$ contro $H_1 : \mu < \mu_0$
  $(\mu \geq \mu_0)$

REGOLA:

$$\text{si rifiuta } H_0 \text{ se } \frac{\overline{X}_n - \mu_0}{\frac{S}{\sqrt{n}}} < -t_{\alpha,\, n-1}$$

$$\text{si accetta } H_0 \text{ se } \frac{\overline{X}_n - \mu_0}{\frac{S}{\sqrt{n}}} \geq -t_{\alpha,\, n-1}$$



Esempio

Il produttore di un pneumatico afferma che la vita media del suo prodotto è di almeno 40000 miglia. Preso un campione di 12 pneumatici si trovano i seguenti dati (1000 = unità)

Gomma	1	2	3	4	5	6	7	8	9	10	11	12
Vita	36.1	40.2	33.8	38.5	42	35.8	37	41	36.8	37.2	33	36

Fissiamo $\alpha = 5\%$. Verifichiamo $H_0 : \mu \geq 40$ contro $H_1 : \mu < 40$.

I calcoli forniscono:
$\bar{X}_n \simeq 37.2833$
$S \simeq 2.7319 \quad (S = \sqrt{S^2} \quad S^2 = \frac{1}{11}\sum_{i=1}^{12}(X_i - \overline{X}_n)^2)$

$$T_{n-1} = \frac{(37.2833 - 40)}{\frac{2.7319}{\sqrt{12}}} \approx -3.445$$

$-t_{0.05,11} = -1.796$
$-3.445 < -1.796 \Rightarrow H_0$ va rifiutata.

Esempio

Si vuole verificare l'ipotesi che il consumo medio di H_2O per abitazione sia di 350 galloni al dì. Per un campione di 20 abitazioni si ha:

340	356	332	362	318	344	386	402	322	360
362	354	340	372	338	375	364	355	324	370

Cosa si conclude?

Dobbiamo verificare
$H_0 : \mu = 350$ contro $H_1 : \mu \neq 350$
$\overline{X}_n = 353.8$
$S \approx 21.85$
$T_{n-1} = \frac{3.8}{\frac{21.85}{\sqrt{20}}} \approx \underline{0.778}$

- se $\alpha = 0.10 \quad \frac{\alpha}{2} = 0.05 \quad t_{\frac{\alpha}{2},19} = t_{0.05,19} = 1.729$

$0.778 < 1.729$ l'ipotesi nulla è accettata ad un livello del 10%.

- se $\alpha = 0.05 \;\; \frac{\alpha}{2} = 0.025 \;\; t_{\frac{\alpha}{2},19} = t_{0.025,19} = 2.093$

$0.778 < 2.093$ l'ipotesi H_0 è accettata ad un livello del 5%.

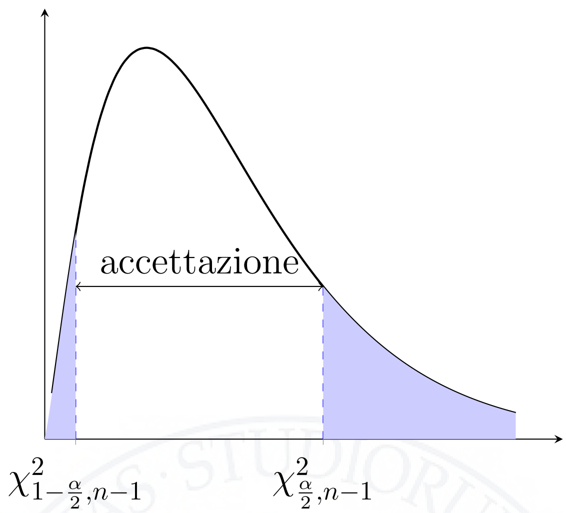
10.4 Test sulla varianza di una popolazione normale quando la media non è nota

10.4.1 Test bilaterali

Sia dato un c.c. $X_1, ..., X_n$ estratto da una una popolazione N con μ e σ^2 incognite
Si vuole verificare per un valore σ_0^2 fissato:
$H_0 : \sigma^2 = \sigma_0^2$ contro $H_1 : \sigma^2 \neq \sigma_0^2$
Ricordiamo che la statistica
$V := \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$
Perciò si adotta la seguente:
REGOLA:

si accetta H_0 se $\chi^2_{1-\frac{\alpha}{2},n-1} \leq \frac{(n-1)S^2}{\sigma_0^2} \leq \chi^2_{\frac{\alpha}{2},n-1}$

si rifiuta H_0 negli altri casi.



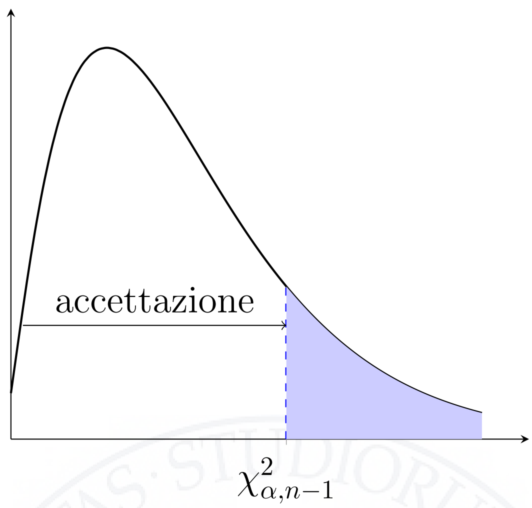
10.4.2 Test unilaterali

- $H_0 : \sigma^2 = \sigma_0^2$ contro $H_1 : \sigma > \sigma_0^2$ $(\sigma^2 \leq \sigma_0^2)$

REGOLA:

$$\text{si accetta } H_0 \text{ se } \frac{(n-1)S^2}{\sigma_0^2} \leq \chi_{\alpha,n-1}$$

$$\text{si rifiuta } H_0 \text{ se } \frac{(n-1)S^2}{\sigma_0^2} > \chi_{\alpha,n-1}$$

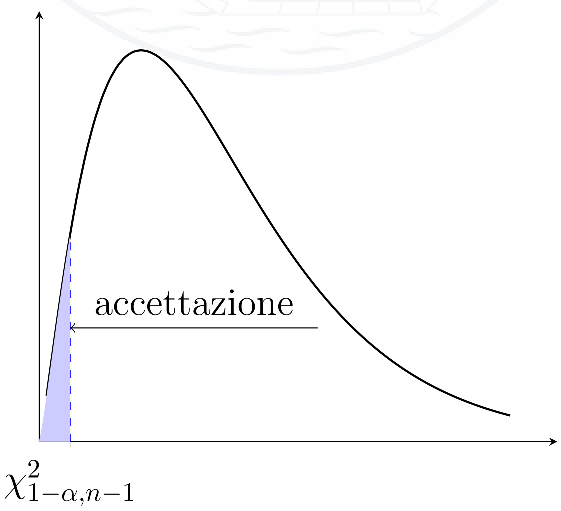


- $H_0 : \sigma^2 = \sigma_0^2$ contro $H_1 : \sigma < \sigma_0^2$ $(\sigma^2 \geq \sigma_0^2)$

REGOLA:

$$\text{si accetta } H_0 \text{ se } \frac{(n-1)S^2}{\sigma_0^2} \geq \chi_{1-\alpha,n-1}$$

$$\text{si rifiuta } H_0 \text{ se } \frac{(n-1)S^2}{\sigma_0^2} < \chi_{1-\alpha,n-1}$$



Esempio

Una macchina deve controllare la quantità di filo su un rocchetto. Va considerata efficiente se la deviazione standard della quantità di nastro selezionata non supera i 0.15cm.
Un campione di 20 pezzi fornisce una varianza campionaria $S^2 = 0.025\text{cm}^2$.
Si può concludere che la macchina non è efficiente?
$H_0 =$ macchina efficiente.
$\sigma_0 = 0.15$, $\sigma_0^2 = 0.0225$

Le due ipotesi sono:
$H_0 : \sigma^2 \leq 0.0225$ contro $H_1 : \sigma^2 > 0.0225$

$$V = \frac{(n-1)S^2}{\sigma_0^2} = \frac{19 \times 0.025}{0.0225} \approx 21.11$$

se $\alpha = 5\% = 0.05$, $\chi^2_{0.05,19} \cong 30.144$.
e poichè $21.11 < 30.144$, H_0 va accettata.

10.5 Test sulla varianza di una popolazione normale quando la media è nota

È possibile costruire test per la varianza anche quando μ è nota.

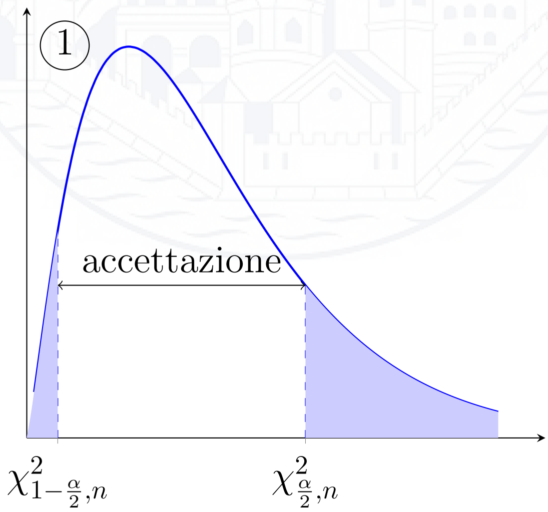
In analogia a quanto detto per gli intervalli di confidenza la statistica da usare è:

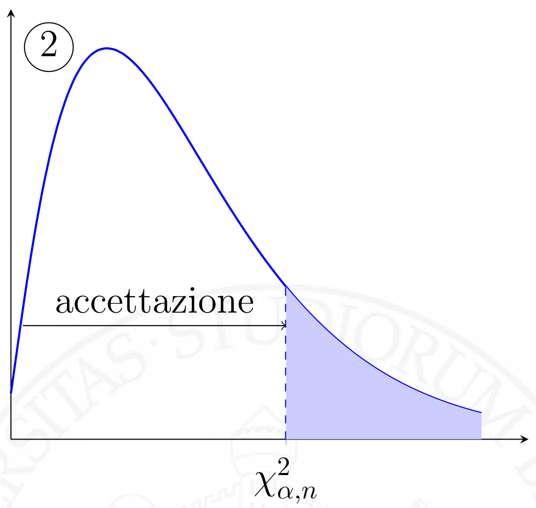
$$U = \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi^2_n$$

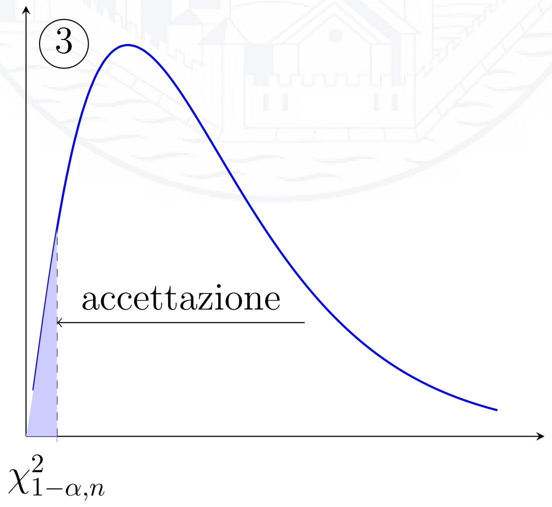
1. bilaterali $H_0 : \sigma^2 = \sigma_0^2 \quad H_1 : \sigma^2 \neq \sigma_0^2$

2. unilaterali $H_0 : \sigma^2 \leq \sigma_0^2 \quad H_1 : \sigma^2 > \sigma_0^2$

3. unilaterali $H_0 : \sigma^2 \geq \sigma_0^2 \quad H_1 : \sigma^2 < \sigma_0^2$




2
accettazione
$\chi^2_{\alpha,n}$


3
accettazione
$\chi^2_{1-\alpha,n}$

<u>Osservazioni</u>

1) test sulla media con σ^2 nota

- se c.c. è estratto da $N \Rightarrow Z = \frac{\overline{X}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$

- se c.c. è estratto da $f(\mu, \sigma^2 \text{nota})$ qlq.
  basta che $n \geq 30$ che per TLC $\Rightarrow Z \dot{\sim} N(0,1)$

2) test sulla media con σ^2 incognita

- se c.c. è estratto da N <u>ed n è grande</u> si usa ancora la statistica

$$Z = \frac{\overline{X}_n - \mu_0}{\frac{S}{\sqrt{n}}} \sim N(0,1)$$

  dove si sostituisce σ incognita con S deviazione standard campionaria.

- se c.c. è estratto da N ed <u>n è piccolo</u> si usa

$$T_{n-1} = \frac{\overline{X}_n - \mu_0}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$$

3) test sulla varianza con μ incognita

- se c.c. è estratto da N ed n è grande o piccolo si usa

$$V = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

Nella pratica però si usa solo per n piccolo. Infatti se n è grande si usa la statistica

$$Z = \frac{S - \sigma_0}{\frac{\sigma_0}{\sqrt{2n}}} \dot{\sim} N(0,1)$$

4) test sulla varianza con μ nota

- se c.c. è estratto da N ed n è grande o piccolo si usa

$$U = \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi^2_n$$