STATISTICA E ANALISI DEI DATI

Capitolo 13 – Verifica delle Ipotesi

Test statistici per grandi campioni

Dott. Stefano Cirillo
Dott. Luigi Di Biasi

a.a. 2025-2026

Verifica Delle Ipotesi

- Quando l'ampiezza del campione è grande, per una popolazione descritta da una variabile aleatoria X con valore medio μ e varianza σ^2
  - Possiamo utilizzare il teorema centrale di convergenza ricordando che la variabile aleatoria

$$Z_n = \frac{\overline{X_n} - \mu}{\frac{\sigma_0}{\sqrt{n}}} \xrightarrow{\text{d}} Z$$

converge in distribuzione ad una variabile normale standard

Test Bilaterale Approssimato

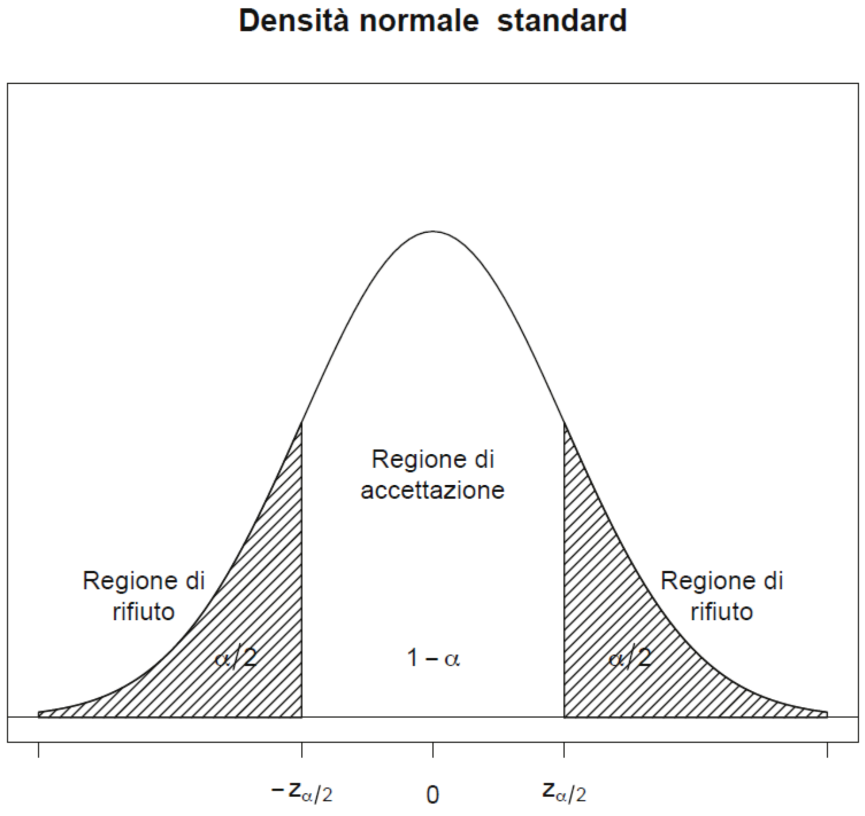
- **Test bilaterali approssimato:** Per campioni numerosi, il test bilaterale di misura α per le ipotesi
  - Consideriamo le ipotesi:

$$\mathbf{H_0}: \mu = \mu_0 \qquad \mathbf{H_1}: \mu \neq \mu_0$$

  considera come variabile aleatoria

$$\frac{\overline{X_n} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}}$$

  con σ_0 deviazione standard quando $\mu = \mu_0$



- L'ipotesi $\mathbf{H_0}$ si accetta se: $-z_{\frac{\alpha}{2}} < \frac{\overline{X_n} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} < z_{\frac{\alpha}{2}}$
- L'ipotesi $\mathbf{H_0}$ si rifiuta se se:

$$\frac{\overline{X_n} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} < -z_{\frac{\alpha}{2}} \; oppure \; \frac{\overline{X_n} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} > z_{\frac{\alpha}{2}}$$

- Per calcolare il valore $-z_{\frac{\alpha}{2}}$ si usa: $\text{qnorm}(1 - \alpha/2, \text{mean} = 0, \text{sd} = 1)$

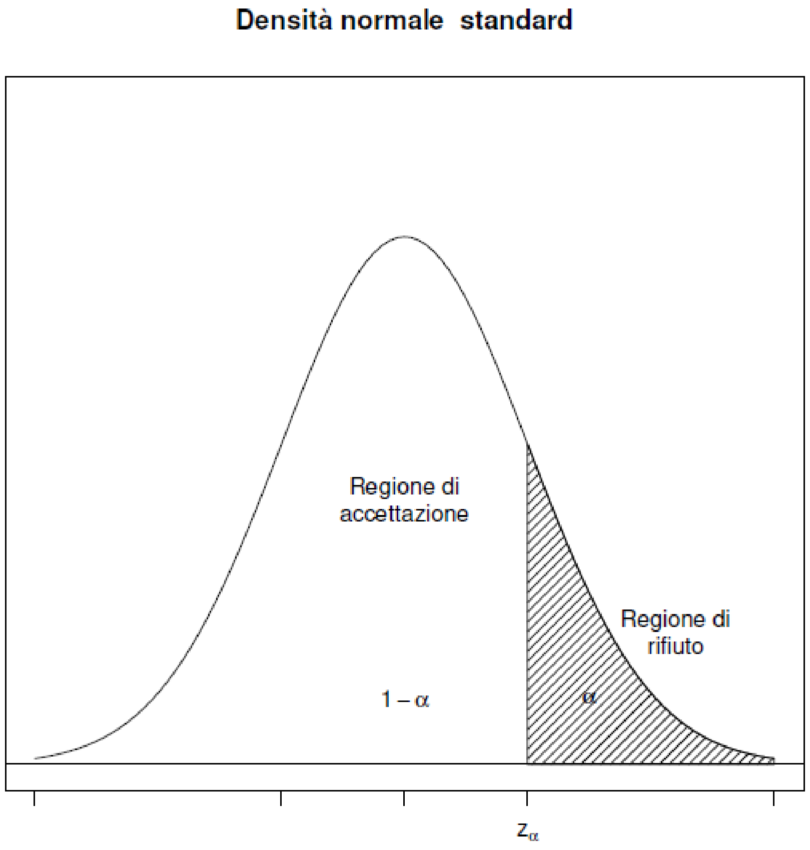
Test Unilaterale Sinistro

- **Test unilaterale sinistro**

  $$\mathbf{H_0}: \mu \leq \mu_0 \qquad \mathbf{H_1}: \mu > \mu_0$$

- L'ipotesi $\mathbf{H_0}$ si accetta se: $\frac{\overline{X_n}-\mu_0}{\frac{\sigma_0}{\sqrt{n}}} < z_\alpha$

- L'ipotesi $\mathbf{H_0}$ si rifiuta se se: $\frac{\overline{X_n}-\mu_0}{\frac{\sigma_0}{\sqrt{n}}} > z_\alpha$



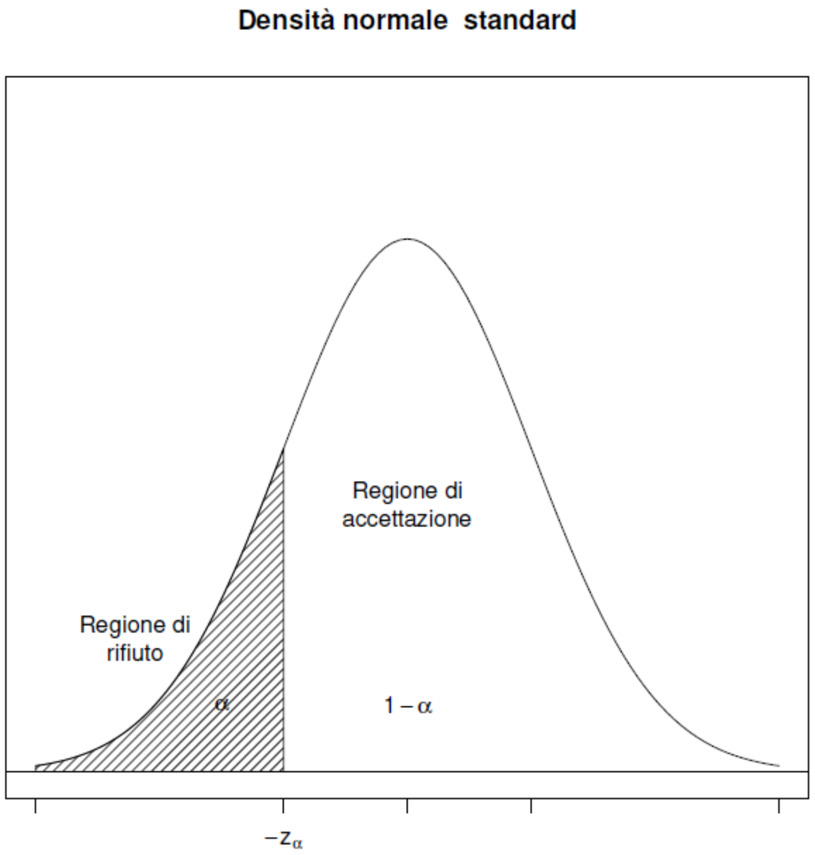
Test Unilaterale Destro

- **Test unilaterale destro**

$$\mathbf{H_0}: \mu \geq \mu_0 \qquad \mathbf{H_1}: \mu < \mu_0$$

- L'ipotesi $\mathbf{H_0}$ si accetta se: $\frac{\overline{X}_n - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} > -z_\alpha$

- L'ipotesi $\mathbf{H_0}$ si rifiuta se se: $\frac{\overline{X}_n - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} < -z_\alpha$



STATISTICA E ANALISI DEI DATI

Popolazione di Bernoulli

Verifica Delle Ipotesi

- Consideriamo una popolazione di Bernoulli descritta dalla variabile aleatoria $X \sim B(p)$
- Siamo interessati a costruire dei test unilaterali e bilaterali per il valore medio $E(X) = p$
- Per campioni numerosi, fissato a priori un livello di significatività α, si ha che:
  - **Test Bilaterale:**

$$\mathbf{H_0}: p = p_0 \qquad \mathbf{H_1}: p \neq p_0$$

  - **Test Unilaterale Sinistro**

$$\mathbf{H_0}: p \leq p_0 \qquad \mathbf{H_1}: p > p_0$$

  - **Test Unilaterale Destro:**

$$\mathbf{H_0}: p \geq p_0 \qquad \mathbf{H_1}: p < p_0$$

- Essendo $\mu_0 = p_0$ e $\sigma_0^2 = p_0(1 - p_0)$ nei test unilaterali e bilaterali occorre considerare:

$$z_{os} = \frac{\overline{X_n} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} = \frac{\overline{X_n} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

Esempio

- Una ditta farmaceutica è interessata a verificare l'efficacia di un nuovo farmaco per curare una data malattia
- Da un'indagine condotta su 900 pazienti affetti da questa malattia trova che il farmaco è efficace in 740 casi
- Possiamo supporre che la popolazione sia distribuita secondo Bernoulli, con p che denota la probabilità che il farmaco sia efficace
  - L'intervallo di confidenza di grado $1 - \alpha = 0.95$ per p è $(0.796, 0.846)$
- Si desidera verificare l'ipotesi $\mathbf{H_0}: p \geq 0.8$ in alternativa $\mathbf{H_1}: p < 0.8$ con un livello di significatività $\alpha = 0.05$
- Occorre considerare un test unilaterale destro

```
> p0<-0.8
> alpha<-0.05
> qnorm(alpha,mean=0,sd=1)
[1] -1.644854
> n<-900
> meancamp<-740/900
> (meancamp-p0)/sqrt(p0*(1-p0)/n)
[1] 1.666667
```

Esempio

- Una ditta farmaceutica è interessata a verificare l'efficacia di un nuovo farmaco per curare una data malattia
- Da un'indagine condotta su 900 pazienti affetti da questa malattia trova che il farmaco è efficace in 740 casi
- Possiamo supporre che la popolazione sia distribuita secondo Bernoulli, con p che denota la probabilità che il farmaco sia efficace
  - L'intervallo di confidenza di grado $1 - \alpha = 0.95$ per p è $(0.796, 0.846)$
- Si desidera verificare l'ipotesi $\mathbf{H_0}: p \geq 0.8$ in alternativa $\mathbf{H_1}: p < 0.8$ con un livello di significatività $\alpha = 0.05$
- Occorre considerare un test unilaterale destro

```
> p0<-0.8
> alpha<-0.05
> qnorm(alpha,mean=0,sd=1)
[1] -1.644854
> n<-900
> meancamp<-740/900
> (meancamp-p0)/sqrt(p0*(1-p0)/n)
[1] 1.666667
```

$-z_\alpha$

z_{os}

Poiché cade nella regione di accettazione, occorre quindi accettare l'ipotesi nulla con un livello di significatività del 5% $(\alpha = 0.05)$

STATISTICA E ANALISI DEI DATI

Popolazione di Poisson

Verifica Delle Ipotesi

- Consideriamo una popolazione di Poisson descritta dalla variabile aleatoria $X \sim P(\lambda)$
- Siamo interessati a costruire dei test unilaterali e bilaterali per il valore medio $E(X) = \lambda$
- Per campioni numerosi, fissato a priori un livello di significatività α, si ha che:
  - **Test bilaterale:**

$$\mathbf{H_0}: \lambda = \lambda_0 \qquad \mathbf{H_1}: \lambda \neq \lambda_0$$

  - **Test unilaterale sinistro**

$$\mathbf{H_0}: \lambda \leq \lambda_0 \qquad \mathbf{H_1}: \lambda > \lambda_0$$

  - **Test Unilaterale Destro:**

$$\mathbf{H_0}: \lambda \geq \lambda_0 \qquad \mathbf{H_1}: \lambda < \lambda_0$$

- Essendo $\mu_0 = \lambda_0$ e $\sigma_0^2 = \lambda$ nei test unilaterali e bilaterali occorre considerare:

$$z_{os} = \frac{\overline{x_n} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} = \frac{\overline{x_n} - \lambda_0}{\sqrt{\frac{\lambda_0}{n}}} = \sqrt{n}\frac{\overline{x_n} - \lambda_0}{\sqrt{\lambda_0}}$$

Esempio

- Si supponga che il numero $N(t)$ di chiamate che arrivano ad un centralino telefonico nell'intervallo $(0, t)$ sia distribuito secondo Poisson con valore medio $E[N(t)] = \lambda t$
- In 100 osservazioni effettuate in intervalli di tempo di $t = 10$ minuti si riscontra che in media sono state effettuate 4 chiamate
  - L'intervallo di confidenza di grado $1 - \alpha = 0.95$ per λ è $(0.3627, 0.4412)$
- Si desidera verificare l'ipotesi $\mathbf{H_0}$: $10\lambda \leq 3.5$ in alternativa $\mathbf{H_1}$: $10\lambda > 3.5$ con un livello di significatività $\alpha = 0.05$
- Occorre considerare un test unilaterale sinistro

```
> lambda0<-3.5
> alpha<-0.05
> qnorm(1-alpha,mean=0,sd=1)
[1] 1.644854
> n<-100
> meancamp<-4
> (meancamp-lambda0)/sqrt(lambda0/n)
[1] 2.672612
```

Esempio

- Si supponga che il numero $N(t)$ di chiamate che arrivano ad un centralino telefonico nell'intervallo $(0, t)$ sia distribuito secondo Poisson con valore medio $E[N(t)] = \lambda t$
- In 100 osservazioni effettuate in intervalli di tempo di $t = 10$ minuti si riscontra che in media sono state effettuate 4 chiamate
  - L'intervallo di confidenza di grado $1 - \alpha = 0.95$ per λ è $(0.3627, 0.4412)$
- Si desidera verificare l'ipotesi $\mathbf{H_0}: 10\lambda \leq 3.5$ in alternativa $\mathbf{H_1}: 10\lambda > 3.5$ con un livello di significatività $\alpha = 0.05$
- Occorre considerare un test unilaterale sinistro

```
> lambda0<-3.5
> alpha<-0.05
> qnorm(1-alpha,mean=0,sd=1)
[1] 1.644854
```

$\rightarrow -z_\alpha$

```
> n<-100
> meancamp<-4
> (meancamp-lambda0)/sqrt(lambda0/n)
[1] 2.672612
```

$\rightarrow z_{os}$

Poiché cade nella regione di rifiuto, occorre quindi rifiutare l'ipotesi nulla con un livello di significatività del 5% $(\alpha = 0.05)$

STATISTICA E ANALISI DEI DATI

Popolazione Esponenziale

Verifica Delle Ipotesi

- Consideriamo una popolazione di Esponenziale descritta dalla variabile aleatoria $X \sim \varepsilon(\lambda)$
- Siamo interessati a costruire dei test unilaterali e bilaterali per il valore medio $E(X) = \frac{1}{\lambda}$
- Per campioni numerosi, fissato a priori un livello di significatività α, si ha che:
  - **Test bilaterale:**

$$\mathbf{H_0}: \frac{1}{\lambda} = \frac{1}{\lambda_0} \qquad \mathbf{H_1}: \frac{1}{\lambda} \neq \frac{1}{\lambda_0}$$

  - **Test unilaterale sinistro**

$$\mathbf{H_0}: \frac{1}{\lambda} \leq \frac{1}{\lambda_0} \qquad \mathbf{H_1}: \frac{1}{\lambda} > \frac{1}{\lambda_0}$$

  - **Test Unilaterale Destro:**

$$\mathbf{H_0}: \frac{1}{\lambda} \geq \frac{1}{\lambda_0} \qquad \mathbf{H_1}: \frac{1}{\lambda} < \frac{1}{\lambda_0}$$

- Essendo $\mu_0 = \frac{1}{\lambda_0}$ e $\sigma_0^2 = \frac{1}{\lambda_0^2}$ nei test unilaterali e bilaterali occorre considerare:

$$z_{os} = \frac{\overline{X_n} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} = \frac{\overline{X_n} - \frac{1}{\lambda_0}}{\sqrt{\frac{1}{n\lambda_0^2}}} = \sqrt{n}(\lambda_0 \overline{X_n} - 1)$$

Test bilaterale per una normale

Test bilaterale	Regione di accettazione di $\mathbf{H}_0$
Test su μ con varianza σ^2 nota $\mathbf{H}_0 : \mu = \mu_0, \quad \mathbf{H}_1 : \mu \neq \mu_0$	$-z_{\alpha/2} < \dfrac{\overline{x}_n - \mu_0}{\sigma/\sqrt{n}} < z_{\alpha/2}$ $z_{\alpha/2}$ si calcola con $\mathsf{qnorm}(1 - \mathsf{alpha}/2, \mathsf{mean} = 0, \mathsf{sd} = 1)$ . . . . $pvalue = P(Z_n < -\|z_{os}\|) + P(Z_n > \|z_{os}\|) = 2\Big[1 - P(Z_n \leq \|z_{os}\|)\Big]$ il p-value si calcola con $2 * (1 - \mathsf{pnorm}(\mathsf{abs}(\mathsf{z_{os}}), \mathsf{mean} = 0, \mathsf{sd} = 1))$
Test su μ con varianza non nota $\mathbf{H}_0 : \mu = \mu_0, \quad \mathbf{H}_1 : \mu \neq \mu_0$	$-t_{\alpha/2,n-1} < \dfrac{\overline{x}_n - \mu_0}{s_n/\sqrt{n}} < t_{\alpha/2,n-1}$ $t_{\alpha/2,n-1}$ si calcola con $\mathsf{qt}(1 - \mathsf{alpha}/2, \mathsf{df} = \mathsf{n} - 1))$ . . . . $pvalue = P(T_n < -\|t_{os}\|) + P(T_n > \|t_{os}\|) = 2\Big[1 - P(T_n \leq \|t_{os}\|)\Big]$ il p-value si calcola con $2 * (1 - \mathsf{pt}(\mathsf{abs}(\mathsf{t_{os}}), \mathsf{df} = \mathsf{n} - 1))$
Test su σ^2 con valore medio μ noto $\mathbf{H}_0 : \sigma^2 = \sigma_0^2, \quad \mathbf{H}_1 : \sigma^2 \neq \sigma_0^2$	$\chi^2_{1-\alpha/2,n} < \displaystyle\sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma_0}\right)^2 < \chi^2_{\alpha/2,n}$ $\chi^2_{1-\alpha/2,n}$ si calcola con $\mathsf{qchisq}(\mathsf{alpha}/2, \mathsf{df} = \mathsf{n})$ $\chi^2_{\alpha/2,n}$ si calcola con $\mathsf{qchisq}(1 - \mathsf{alpha}/2, \mathsf{df} = \mathsf{n})$
Test su σ^2 con valore medio non noto $\mathbf{H}_0 : \sigma^2 = \sigma_0^2, \quad \mathbf{H}_1 : \sigma^2 \neq \sigma_0^2$	$\chi^2_{1-\alpha/2,n-1} < \dfrac{(n-1)\,s_n^2}{\sigma_0^2} < \chi^2_{\alpha/2,n-1}$ $\chi^2_{1-\alpha/2,n-1}$ si calcola con $\mathsf{qchisq}(\mathsf{alpha}/2, \mathsf{df} = \mathsf{n} - 1)$ $\chi^2_{\alpha/2,n-1}$ si calcola con $\mathsf{qchisq}(1 - \mathsf{alpha}/2, \mathsf{df} = \mathsf{n} - 1)$

Test unilaterale sinistro per una normale

Test unilaterale sinistro	Regione di accettazione di $\mathbf{H}_0$
Test su μ con varianza σ^2 nota $\mathbf{H}_0 : \mu \leq \mu_0, \quad \mathbf{H}_1 : \mu > \mu_0$	$\dfrac{\overline{x}_n - \mu_0}{\sigma/\sqrt{n}} < z_\alpha$ z_α si calcola con $\mathsf{qnorm}(1 - \mathsf{alpha}, \mathsf{mean} = 0, \mathsf{sd} = 1)$ . . . . $pvalue = P(Z_n > z_{os}) = 1 - P(Z_n \leq z_{os})$ il p-value si calcola con $1 - \mathsf{pnorm}(\mathsf{z_{os}}, \mathsf{mean} = 0, \mathsf{sd} = 1)$
Test su μ con varianza non nota $\mathbf{H}_0 : \mu \leq \mu_0, \quad \mathbf{H}_1 : \mu > \mu_0$	$\dfrac{\overline{x}_n - \mu_0}{s_n/\sqrt{n}} < t_{\alpha,n-1}$ $t_{\alpha,n-1}$ si calcola con $\mathsf{qt}(1 - \mathsf{alpha}, \mathsf{df} = \mathsf{n} - 1))$ . . . . $pvalue = P(T_n > t_{os}) = 1 - P(T_n \leq t_{os})$ il p-value si calcola con $1 - \mathsf{pt}(\mathsf{t_{os}}, \mathsf{df} = \mathsf{n} - 1)$
Test su σ^2 con valore medio μ noto $\mathbf{H}_0 : \sigma^2 \leq \sigma_0^2, \quad \mathbf{H}_1 : \sigma^2 > \sigma_0^2$	$\displaystyle\sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma_0}\right)^2 < \chi^2_{\alpha,n}$ $\chi^2_{\alpha,n}$ si calcola con $\mathsf{qchisq}(1 - \mathsf{alpha}, \mathsf{df} = \mathsf{n})$
Test su σ^2 con valore medio non noto $\mathbf{H}_0 : \sigma^2 \leq \sigma_0^2, \quad \mathbf{H}_1 : \sigma^2 > \sigma_0^2$	$\dfrac{(n-1)\, s_n^2}{\sigma_0^2} < \chi^2_{\alpha,n-1}$ $\chi^2_{\alpha,n-1}$ si calcola con $\mathsf{qchisq}(1 - \mathsf{alpha}, \mathsf{df} = \mathsf{n} - 1)$

Test unilaterale destro per una normale

Test unilaterale destro	Regione di accettazione di $\mathbf{H}_0$
Test su μ con varianza σ^2 nota $\mathbf{H}_0 : \mu \geq \mu_0, \quad \mathbf{H}_1 : \mu < \mu_0$	$\dfrac{\overline{x}_n - \mu_0}{\sigma/\sqrt{n}} > -z_\alpha$ $-z_\alpha$ si calcola con $\mathsf{qnorm(alpha, mean = 0, sd = 1)}$ . . . . $pvalue = P(Z_n \leq z_{os})$ il p-value si calcola con $\mathsf{pnorm(z_{os}, mean = 0, sd = 1)}$
Test su μ con varianza non nota $\mathbf{H}_0 : \mu \geq \mu_0, \quad \mathbf{H}_1 : \mu < \mu_0$	$\dfrac{\overline{x}_n - \mu_0}{s_n/\sqrt{n}} > -t_{\alpha,n-1}$ $-t_{\alpha,n-1}$ si calcola con $\mathsf{qt(alpha, df = n - 1))}$ . . . . $pvalue = P(T_n \leq t_{os})$ il p-value si calcola con $\mathsf{pt(t_{os}, df = n - 1)}$
Test su σ^2 con valore medio μ noto $\mathbf{H}_0 : \sigma^2 \geq \sigma_0^2, \quad \mathbf{H}_1 : \sigma^2 < \sigma_0^2$	$\sum_{i=1}^{n} \left(\dfrac{x_i - \mu}{\sigma_0} \right)^2 > \chi^2_{1-\alpha,n}$ $\chi^2_{1-\alpha,n}$ si calcola con $\mathsf{qchisq(alpha, df = n)}$
Test su σ^2 con valore medio non noto $\mathbf{H}_0 : \sigma^2 \geq \sigma_0^2, \quad \mathbf{H}_1 : \sigma^2 < \sigma_0^2$	$\dfrac{(n-1)\, s_n^2}{\sigma_0^2} > \chi^2_{1-\alpha,n-1}$ $\chi^2_{1-\alpha,n-1}$ si calcola con $\mathsf{qchisq(alpha, df = n - 1)}$

Test sulla media per campioni numerosi

	Regione di accettazione di $\mathbf{H_0}$
Test bilaterale $\mathbf{H}_0 : \mu = \mu_0, \quad \mathbf{H}_1 : \mu \neq \mu_0$	$-z_{\alpha/2} < \dfrac{\overline{x}_n - \mu_0}{\sigma_0/\sqrt{n}} < z_{\alpha/2}$ $z_{\alpha/2}$ si calcola con $\mathsf{qnorm}(1 - \mathsf{alpha}/2, \mathsf{mean} = 0, \mathsf{sd} = 1)$
Test unilaterale sinistro $\mathbf{H}_0 : \mu \leq \mu_0, \quad \mathbf{H}_1 : \mu > \mu_0$	$\dfrac{\overline{x}_n - \mu_0}{\sigma_0/\sqrt{n}} < z_{\alpha}$ z_{α} si calcola con $\mathsf{qnorm}(1 - \mathsf{alpha}, \mathsf{mean} = 0, \mathsf{sd} = 1)$
Test unilaterale destro $\mathbf{H}_0 : \mu \geq \mu_0, \quad \mathbf{H}_1 : \mu < \mu_0$	$\dfrac{\overline{x}_n - \mu_0}{\sigma_0/\sqrt{n}} > -z_{\alpha}$ $-z_{\alpha}$ si calcola con $\mathsf{qnorm}(\mathsf{alpha}, \mathsf{mean} = 0, \mathsf{sd} = 1)$