UNIVERSITÀ DEGLI STUDI DI TRIESTE

Dipartimento di Ingegneria e Architettura

Corso di misure meccaniche, termiche e collaudi

Prof. Lucia Parussini
Prof. Rodolfo Taccani
Prof. Ronelly De Souza

a.a. 2026-2027

Outline

- Introduzione
- Test d'ipotesi
- Interpretazione del risultato del test d'ipotesi
- Esempi di test di ipotesi

Introduzione

La statistica è una parte fondamentale di un efficiente processo di progettazione e produzione, di verifica, di collaudo.

Fasi:

- ❑ progettare esperimenti
- ❑ eseguire esperimenti
- ❑ raccogliere dati reali
- ❑ prendere decisioni basate sui dati

Questi esperimenti possono essere usati

- ❑ per verifiche di conformità e prestazioni
- ❑ per testare l'effetto del cambiamento di una dimensione, di un materiale o di una fase del processo produttivo

Introduzione

Questi esperimenti possono essere usati

- ❑ per verifiche di conformità e prestazioni
- ❑ per testare l'effetto del cambiamento di una dimensione, di un materiale o di una fase del processo produttivo

Ad esempio verifica delle prestazioni attraverso il confronto dei dati:

- ○ di due strumenti
- ○ di due metodi
- ○ di una procedura in periodi diversi
- ○ di due analisti o laboratori
- ○ con il valore di riferimento o assegnato a un campione di controllo, considerando anche la sua incertezza nel modello del confronto

Introduzione

Questi esperimenti possono essere usati

- ❑ per verifiche di conformità e prestazioni
- ❑ per testare l'effetto del cambiamento di una dimensione, di un materiale o di una fase del processo produttivo

Di solito si vuole cambiare una di queste variabili per migliorare la capacità del pezzo di essere realizzato entro le tolleranze o le prestazioni richieste per il prodotto.

Ad esempio:

- il pezzo continua a rompersi durante i test di affidabilità, e volete testare se cambiare il materiale riduce questi fallimenti
- dovete rilavorare una grande percentuale dei vostri pezzi, ma pensate che cambiare una dimensione renderà le rilavorazioni non necessarie e accelererà il processo

Hypothesis testing

Un test di ipotesi valuta la compatibilità dei dati osservati con un modello statistico e con un'ipotesi nulla.

Prima di osservare i risultati si definiscono l'ipotesi alternativa, il livello di significatività e la regola di decisione.

Il test può portare al rifiuto oppure al mancato rifiuto dell'ipotesi nulla. La decisione è soggetta a errori quantificati dal modello.

Hypothesis testing

Un'ipotesi statistica riguarda un parametro del modello della popolazione. Si distinguono il parametro e lo spazio dei suoi possibili valori.

$f(x;\theta), \qquad \theta \in \Theta$

L'ipotesi nulla e l'alternativa individuano insiemi disgiunti dello spazio parametrico, scelti prima dell'analisi.

$H_0: \theta \in \Theta_0 \quad H_1: \theta \in \Theta_1$

$\Theta_0 \cap \Theta_1 = \emptyset$

Il test cerca evidenza contro l'ipotesi nulla. Il mancato rifiuto non dimostra che l'ipotesi sia vera.

Hypothesis testing

Un'ipotesi è semplice se specifica completamente la distribuzione. Se lascia possibili più distribuzioni, è composta.

Esempio sulla durata media dei cuscinetti, espressa in ore

$H_0: \mu \leq 56000 \quad H_1: \mu > 56000$

In una famiglia normale, l'uguaglianza della sola media definisce un'ipotesi semplice soltanto se la varianza è nota o fissata.

$X \sim \mathrm{N}(\mu, \sigma^2), \qquad \sigma^2 = \sigma_0^2 \text{ nota}$

$\mu = 56000 : \text{ semplice}; \quad \mu > 56000 : \text{ composta}$

Con varianza incognita, anche l'ipotesi sulla media uguale al riferimento è composta.

Hypothesis testing

Nel modello i.i.d. il campione casuale è rappresentato da variabili aleatorie indipendenti e identicamente distribuite.

$X_1, X_2, \dots, X_n$

Dopo l'esperimento osserviamo le loro realizzazioni, cioè i valori numerici dei dati.

$x_1, x_2, \dots, x_n$

La regione critica contiene tutti i campioni osservabili che portano al rifiuto dell'ipotesi nulla. La si definisce prima di osservare il risultato.

$(x_1, \dots, x_n) \in \mathrm{G} \ : \ \text{si rifiuta} \ \ H_0$

Interpretazione del test di significatività

La stessa regola di test può essere espressa confrontando il p-value con il livello di significatività oppure la statistica con i valori critici.

I quantili delle statistiche continue seguono la convenzione cumulativa. La distribuzione di riferimento è quella prevista sotto l'ipotesi nulla.

$$P_{H_0}(T \leq q_p) = p$$

La scelta della coda destra, sinistra o di entrambe le code dipende dall'alternativa prefissata.

Interpretare il valore di p

Il p-value si confronta con il livello di significatività scelto prima di analizzare i dati. Un livello usuale è il 5%.

$p \leq \alpha$: si rifiuta H_0

$p > \alpha$: non si rifiuta H_0

Un livello più basso riduce il rischio di errore di tipo I. Non dimostra che l'ipotesi nulla sia più credibile o vera.

Un risultato significativo fornisce evidenza contro l'ipotesi nulla, nelle condizioni del test. Non misura da solo l'importanza pratica dell'effetto.

Interpretare il valore di p

Esempio di test di normalità

$p = 0.07 > \alpha = 0.05$

Al livello del 5% non si rifiuta l'ipotesi di normalità: il test non rileva evidenza sufficiente contro il modello normale.

Questo esito non dimostra che la popolazione sia normale e non assegna una probabilità del 95% alla normalità.

La relazione fra livello di test e livello di un intervallo di confidenza vale per test e intervalli costruiti in modo coerente. Non trasforma un test di normalità in una probabilità dell'ipotesi.

Interpretare il valore di p

Il p-value è la probabilità, calcolata sotto l'ipotesi nulla, di ottenere una statistica almeno altrettanto estrema di quella osservata, nella direzione dell'alternativa.

Esempio per una regione critica nella coda destra

$$p = P_{H_0}(T \geq t_{\text{obs}})$$

Un p-value piccolo indica scarsa compatibilità dei dati con il modello nullo. Non è la probabilità che l'ipotesi nulla sia vera o che il risultato sia dovuto al caso.

Il mancato rifiuto può dipendere da un effetto piccolo, da pochi dati o da una potenza insufficiente. Non dimostra la verità dell'ipotesi nulla.

Interpretare il valore di p

Scelte da fissare prima dell'analisi

Definire il quesito, l'ipotesi alternativa, il livello di significatività e le regole di raccolta e arresto dei dati.

Il p-value si calcola dai dati. Non si sceglie dopo aver osservato il risultato e non coincide con il livello di significatività.

Selezionare dati, test o regole di arresto fino a ottenere un risultato significativo è p-hacking. Questa pratica altera il controllo dell'errore e compromette l'interpretazione del test.

Le analisi esplorative sono utili, purché siano dichiarate e distinte dalle verifiche pianificate.

Interpretare i valori critici

I valori critici derivano dalla distribuzione della statistica sotto l'ipotesi nulla e dal livello di significatività.

La regione critica dipende dall'alternativa

Coda destra: $T > q_{1-\alpha}$

Coda sinistra: $T < q_{\alpha}$

Due code: $T < q_{\alpha/2}$ oppure $T > q_{1-\alpha/2}$

Queste regole, con i quantili appropriati, si usano nei test parametrici e non parametrici. Per statistiche discrete si sceglie una regione con probabilità non superiore al livello fissato.

Interpretare i valori critici

Esempio di test di normalità mediante valori critici

Se la statistica non appartiene alla regione critica al livello del 5%, non si rifiuta l'ipotesi di normalità.

Si conclude che il test non ha rilevato evidenza sufficiente contro la normalità. La decisione non dimostra che i dati provengano da una popolazione normale.

La stessa conclusione si ottiene con il p-value quando la soglia e la calibrazione del test sono coerenti.

Errori nei test di ipotesi statistici

Errore di tipo I: rifiutare un'ipotesi nulla vera. Il livello del test controlla questo rischio nelle condizioni del modello.

$$P_{H_0}(\text{rifiuto di } H_0) \leq \alpha$$

Errore di tipo II: non rifiutare l'ipotesi nulla quando l'alternativa è vera. Il suo rischio dipende dall'effetto reale, dalla numerosità e dal disegno.

$$\beta(\theta) = P_\theta(\text{non rifiuto di } H_0), \qquad \theta \in \Theta_1$$

$$\text{Potenza} = 1 - \beta(\theta)$$

Un falso negativo può essere frequente quando la potenza è bassa. Il livello del 5% non controlla entrambi gli errori.

Errori nei test di ipotesi statistici

La tabella distingue le due decisioni corrette dagli **errori di tipo I** e **II**. Il rischio di mancato rifiuto dipende dall'alternativa e dalla potenza del test.

Situazione reale	Non rifiuto della nulla	Rifiuto della nulla
Ipotesi nulla vera	Decisione corretta	Errore di tipo I
Ipotesi nulla falsa	Errore di tipo II	Decisione corretta

Un livello più piccolo riduce l'errore di tipo I. A parità di campione e disegno, in genere aumenta l'errore di tipo II.

$$\alpha = 3 \times 10^{-7} \approx \frac{1}{3333333}$$

Questo è il rischio massimo di un rifiuto erroneo quando la nulla è vera. Non è una probabilità a posteriori sulle ipotesi.

Esempi di test di ipotesi

Il quesito determina la famiglia di test

Quesito	Test principali	Esempio nelle misure
Una media rispetto a un riferimento	Z-test, t-test a un campione	Temperatura media rispetto al valore nominale
Confronto fra gruppi	t di Student, Welch, ANOVA Mann-Whitney, Wilcoxon, Kruskal-Wallis, Friedman	Prestazioni di materiali o processi diversi
Forma della distribuzione	Shapiro-Wilk, D'Agostino-Pearson, Anderson-Darling, chi-quadro, Kolmogorov-Smirnov	Distribuzione degli errori di misura
Variabilità e associazione	Chi-quadro e F per le varianze Pearson, Spearman e Kendall	Ripetibilità degli strumenti e legami fra grandezze

La scelta richiede anche le ipotesi del modello e il disegno sperimentale.

Guida all’uso dei test statistici

1 Obiettivo del confronto

Una media, la variabilità, una distribuzione o un’associazione?
Definire prima il parametro o la proprietà da verificare.

2 Tipo di variabile

Quantitativa, ordinale o categorica.
Le categorie di una macchina o di un materiale identificano i gruppi.

3 Disegno sperimentale

Uno, due o più gruppi. Campioni indipendenti oppure dati appaiati?
Lo stesso provino misurato prima e dopo produce una coppia.

4 Ipotesi di applicabilità

Verificare indipendenza, valori anomali e distribuzione pertinente.
Nel t-test appaiato conta la distribuzione delle differenze.

Un test di normalità, da solo, non determina quale test di confronto usare.

Guida all’uso dei test statistici

Confronto fra due gruppi

Disegno	Confronto delle medie	Test sui ranghi
Campioni indipendenti Esempio: due lotti di alberi	t di Student: varianze uguali Welch: varianze non assunte uguali	Mann-Whitney U (Wilcoxon rank-sum)
Dati appaiati Esempio: stesso sensore prima e dopo la taratura	t di Student appaiato Si analizzano le differenze di ogni coppia	Wilcoxon signed-rank Si usano segni e ranghi delle differenze

Ipotesi da verificare

Per i test t, normalità dei dati o delle differenze, oppure adeguata approssimazione.
Mann-Whitney riguarda le distribuzioni. L’interpretazione come confronto di posizione richiede ulteriori ipotesi. Wilcoxon richiede differenze simmetriche.

Guida all’uso dei test statistici

Confronto fra tre o più gruppi

Disegno	Confronto delle medie	Test sui ranghi
Campioni indipendenti Esempio: tre trattamenti superficiali	ANOVA a una via ANOVA di Welch per varianze disuguali	Kruskal-Wallis
Misure ripetute Esempio: stessi sensori a tre temperature	ANOVA per misure ripetute	Friedman

Ipotesi da verificare

L’ANOVA classica assume errori indipendenti, normali e con varianza comune.
Le misure ripetute richiedono un modello della dipendenza e, nel caso classico, la sfericità. Un esito significativo non identifica da solo i gruppi diversi.

Guida all’uso dei test statistici

Relazioni fra variabili quantitative

Obiettivo	Metodo	Esempio meccanico o termico
Associazione lineare	Correlazione di Pearson	Temperatura e dilatazione di una barra
Associazione monotona o dati ordinali	Spearman o Kendall	Livello di usura e intensità delle vibrazioni
Descrivere una risposta con un predittore	Regressione lineare	Dilatazione in funzione della temperatura
Descrivere una risposta con più predittori	Regressione lineare multipla	Vita dell’utensile in funzione di velocità e avanzamento

La correlazione non dimostra causalità. Per l’inferenza, verificare le ipotesi del metodo.

Guida all'uso dei test statistici

Altri quesiti frequenti

Quesito	Metodo	Condizione essenziale
Compatibilità con una distribuzione normale	Shapiro-Wilk D'Agostino-Pearson Anderson-Darling	Indipendenza dei dati. Il mancato rifiuto non prova la normalità.
Adattamento a un modello o confronto di due distribuzioni	Kolmogorov-Smirnov a uno o due campioni Chi-quadro di adattamento	K-S classico: modello continuo e, a un campione, completamente specificato.
Una o due varianze coincidono con i valori attesi?	Chi-quadro su una varianza F su due varianze; Levene	Chi-quadro e F classici richiedono normalità.
Le categorie sono associate?	Chi-quadro di indipendenza	Conteggi attesi adeguati. Esempio: macchina e conformità.
Quali fattori spiegano un esito binario?	Regressione logistica	Esempio: probabilità di rottura. È un modello parametrico.

Verifica di ipotesi sulla media con varianza nota (Z-Test)

Osservazioni indipendenti e identicamente distribuite, normali, con varianza nota.

$$H_0\colon \mu = \mu_0 \quad z = \frac{\overline{x} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}}$$

Si rifiuta l'ipotesi nulla secondo l'alternativa prefissata

$$H_1\colon \mu > \mu_0 \quad z > z_{1-\alpha}$$

$$H_1\colon \mu < \mu_0 \quad z < z_{\alpha}$$

$$H_1\colon \mu \neq \mu_0 \quad |z| > z_{1-\alpha/2}$$

Per dati non normali lo Z-test è approssimato solo se è adeguata l'approssimazione normale della media. La soglia di 30 dati non è una garanzia.

Verifica di ipotesi sulla media con varianza incognita (T-Test)

Osservazioni indipendenti e identicamente distribuite, normali, con varianza incognita.

$$H_0\colon \mu = \mu_0 \quad t = \frac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Si rifiuta l'ipotesi nulla secondo l'alternativa prefissata

$$H_1\colon \mu > \mu_0 \quad t > t_{1-\alpha,n-1}$$

$$H_1\colon \mu < \mu_0 \quad t < t_{\alpha,n-1}$$

$$H_1\colon \mu \neq \mu_0 \quad |t| > t_{1-\alpha/2,n-1}$$

$$\mathrm{SE} = \frac{s}{\sqrt{n}}, \qquad \nu = n - 1$$

Con dati non normali il test può essere approssimato; la numerosità, da sola, non garantisce la validità.

Verifica di ipotesi sulla media con varianza incognita (T-Test)

Esempio 1 - Test a destra - Durezza Brinell

25 pezzi di ferro dolce ricotto. Si assumono misure indipendenti e una popolazione normale. Il test riguarda la durezza media rispetto a 170.

170	167	174	179	179
187	179	183	179	156
163	156	187	156	167
156	174	170	183	179
174	179	170	159	187

$H_0\colon \mu \leq 170 \quad H_1\colon \mu > 170$

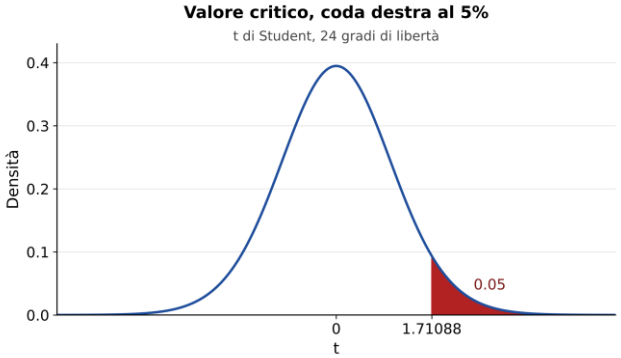
Numero	Media	Dev. std.	Errore std.	Limite inf. unilat. 95%
25	172.520	10.3123	2.06246	168.991

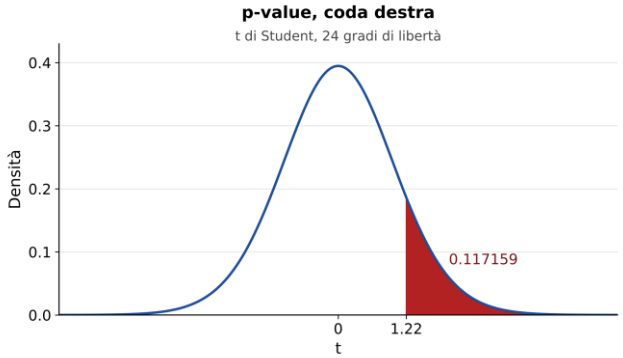
$t \approx 1.2218, \qquad \nu = 24, \qquad p \approx 0.1168$

Verifica di ipotesi sulla media con varianza incognita (T-Test)

Esempio 1 - Decisione al livello del 5%

Si confrontano la statistica osservata e il valore critico nella coda destra. Le aree dei grafici mostrano il livello del test e il p-value.





$$t \approx 1.2218 < t_{0.95,24} \approx 1.7109$$

$$p \approx 0.1168 > \alpha = 0.05$$

Non si rifiuta l'ipotesi nulla. Non emerge evidenza sufficiente che la durezza media dei pezzi di ferro dolce superi 170.

Verifica di ipotesi sulla media con varianza incognita (T-Test)

Esempio 2 - Test a sinistra - Dissipatori termici

Un nuovo dissipatore deve ridurre l'incremento di temperatura rispetto al valore nominale di 15.7 °C. Si provano 33 esemplari indipendenti alla stessa potenza termica e nelle stesse condizioni ambientali.

Incrementi di temperatura misurati (°C)

11.5	11.8	15.7	16.1	14.1	10.5	9.3	15	11.1	15.2	19
12.8	12.4	19.2	13.5	12.2	13.3	16.5	13.5	14.4	16.7	10.9
13	10.3	15.8	15.1	17.1	13.3	12.4	8.5	14.3	12.9	13.5

$H_0: \mu = 15.7\,°\mathrm{C}$ $\quad H_1: \mu < 15.7\,°\mathrm{C}$

Numero	Media (°C)	Dev. std. (°C)	Errore std. (°C)	Limite sup. 95% (°C)
33	13.664	2.544	0.443	14.414

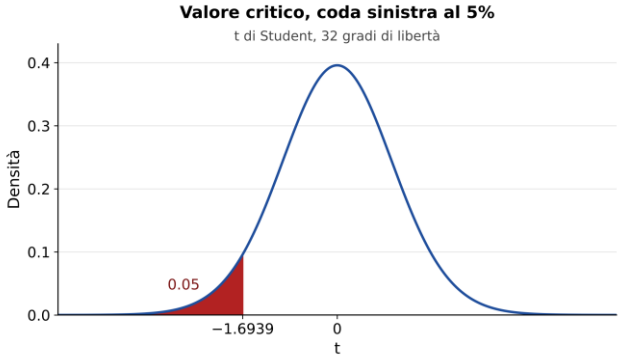
$$t = \frac{(\overline{x} - 15.7)\sqrt{33}}{s} \approx -4.5990, \qquad \nu = 32, \qquad p = 0.000032$$

Assunta l'approssimazione normale. Limite superiore unilaterale 95%.

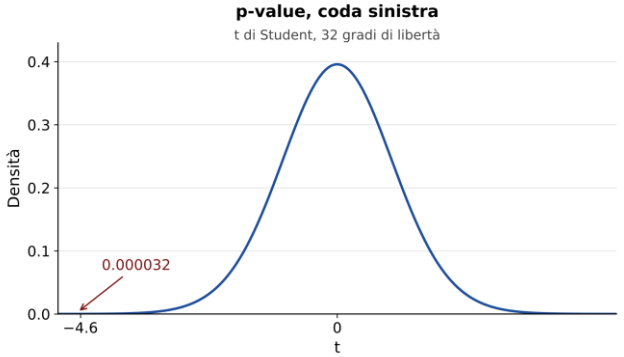
Verifica di ipotesi sulla media con varianza incognita (T-Test)

Esempio 2 - Test a sinistra - Dissipatori termici

Decisione al livello di significatività del 5%



$$t = -4.60 < t_{0.05,32} = -1.6939$$



$$p = 0.000032 < \alpha = 0.05$$

Si rifiuta l'ipotesi nulla. I dati forniscono evidenza che l'incremento medio di temperatura dei dissipatori è inferiore a 15.7 °C.

Verifica di ipotesi sulla media con varianza incognita (T-Test)

Esempio 3 - Test bilaterale - Spessore di guarnizioni

Un produttore dichiara uno spessore medio di 7.5 mm per le guarnizioni elastomeriche. Il controllo qualità preleva casualmente 10 guarnizioni da un lotto e ne misura lo spessore.

Spessori misurati (mm)

7.65	7.60	7.65	7.70	7.55
7.55	7.40	7.40	7.50	7.50

$H_0: \mu = 7.5\ \text{mm} \qquad H_1: \mu \neq 7.5\ \text{mm}$

Numero	Media (mm)	Dev. std. (mm)	Errore std. (mm)	IC 95% della media (mm)
10	7.550	0.1027	0.0325	(7.4765, 7.6235)

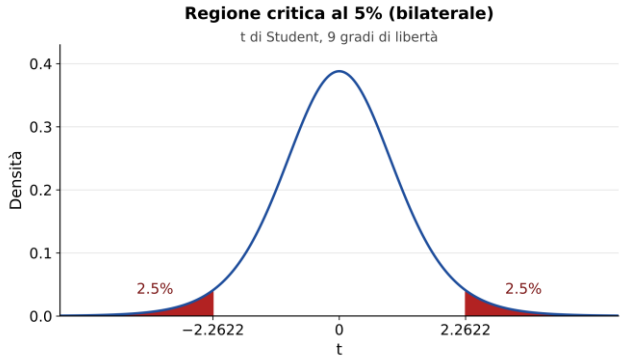
$$t = \frac{(\overline{x} - 7.5)\sqrt{10}}{s} \approx 1.5390, \qquad \nu = 9, \qquad p = 0.158$$

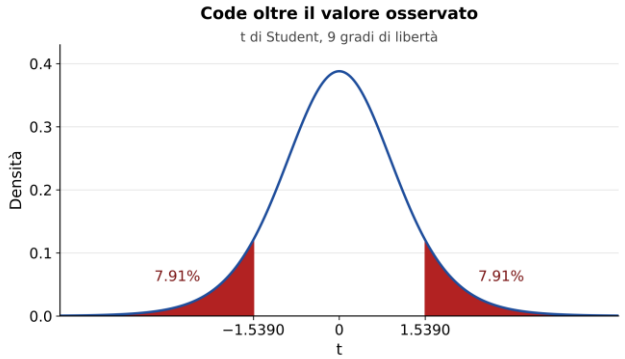
Si assumono indipendenza e normalità degli spessori nel lotto.

Verifica di ipotesi sulla media con varianza incognita (T-Test)

Esempio 3 - Test bilaterale - Spessore di guarnizioni

Decisione al livello di significatività del 5%





$$t_{0.025,9} \approx -2.2622 < t \approx 1.5390 < t_{0.975,9} \approx 2.2622$$

$$p = 0.158 > \alpha = 0.05$$

Non si rifiuta l'ipotesi nulla. Al livello del 5%, non emerge una differenza significativa dello spessore medio delle guarnizioni rispetto a 7.5 mm.

Test chi-quadro sulla varianza – caso a media nota

Campione i.i.d. normale. La media è nota e fissata.

$$H_0\colon \sigma^2 = \sigma_0^2 \quad s_{\mu_0}^2 = \frac{\sum_{i=1}^{n}(x_i - \mu_0)^2}{n}$$

$$Q = \frac{n s_{\mu_0}^2}{\sigma_0^2} \sim \chi_n^2, \qquad \nu = n$$

Distribuzione sotto l'ipotesi nulla. Regioni di rifiuto

$$H_1\colon \sigma^2 > \sigma_0^2 \quad Q > \chi_{1-\alpha,\nu}^2$$

$$H_1\colon \sigma^2 < \sigma_0^2 \quad Q < \chi_{\alpha,\nu}^2$$

$$H_1\colon \sigma^2 \neq \sigma_0^2 \quad Q < \chi_{\alpha/2,\nu}^2 \text{ oppure } Q > \chi_{1-\alpha/2,\nu}^2$$

La normalità è essenziale per la calibrazione chi-quadro esatta, anche con campioni grandi. Tutti i quantili della distribuzione sono non negativi.

Test chi-quadro sulla varianza – caso a media incognita

Campione i.i.d. normale. La media è stimata dal campione.

$$H_0\colon \sigma^2 = \sigma_0^2 \quad s^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}$$

$$Q = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi^2_{n-1}, \qquad \nu = n-1$$

Distribuzione sotto l'ipotesi nulla. Regioni di rifiuto

$$H_1\colon \sigma^2 > \sigma_0^2 \quad Q > \chi^2_{1-\alpha,\nu}$$

$$H_1\colon \sigma^2 < \sigma_0^2 \quad Q < \chi^2_{\alpha,\nu}$$

$$H_1\colon \sigma^2 \neq \sigma_0^2 \quad Q < \chi^2_{\alpha/2,\nu} \text{ oppure } Q > \chi^2_{1-\alpha/2,\nu}$$

La normalità è essenziale per la calibrazione chi-quadro esatta, anche con campioni grandi. Tutti i quantili della distribuzione sono non negativi.

F-Test sulla varianza

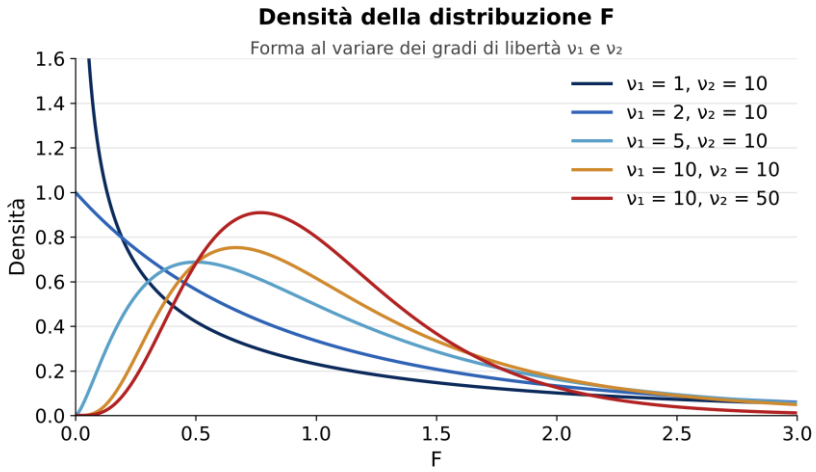
Da due campioni indipendenti provenienti da popolazioni normali si calcolano due varianze campionarie.

$$F = \frac{s_1^2}{s_2^2}$$

Sotto l'ipotesi di varianze uguali il rapporto confronta due stime della stessa varianza. Non deve però essere necessariamente vicino a uno: può variare molto, soprattutto con pochi dati.

Ripetendo il campionamento si ottiene la distribuzione F, con gradi di libertà determinati dalle numerosità dei due campioni.

F-Test sulla varianza



La distribuzione F dipende dai gradi di libertà

$\nu_1 = n_1 - 1$

$\nu_2 = n_2 - 1$

Si è fin qui parlato di una popolazione normalmente distribuita con parametri μ e σ^2; in realtà questa condizione è un po' più restrittiva del necessario: in realtà il campionamento può avvenire anche da due differenti popolazioni, sempre distribuite normalmente, con medie μ_1 e μ_2 differenti ma con la stessa varianza σ^2. La condizione di omogeneità delle varianze è nota col termine di *omoschedasticità*.

In sintesi, la statistica F è il rapporto di due stime indipendenti della comune varianza di due popolazioni normalmente distribuite e omoschedastiche.

F-Test sulla varianza

Campioni indipendenti da popolazioni normali. L'ordine del rapporto è fissato prima di osservare le varianze.

H_0: $\sigma_1^2 = \sigma_2^2$ $\quad F = \dfrac{s_1^2}{s_2^2}$

$\nu_1 = n_1 - 1, \qquad \nu_2 = n_2 - 1$

H_1: $\sigma_1^2 > \sigma_2^2$ $\quad F > F_{1-\alpha,\nu_1,\nu_2}$

H_1: $\sigma_1^2 < \sigma_2^2$ $\quad F < F_{\alpha,\nu_1,\nu_2}$

H_1: $\sigma_1^2 \neq \sigma_2^2$ $\quad F < F_{\alpha/2,\nu_1,\nu_2}$ oppure $F > F_{1-\alpha/2,\nu_1,\nu_2}$

Nel test bilaterale si può porre la varianza maggiore al numeratore, mantenendo coerenti i gradi di libertà e il livello di ciascuna coda.

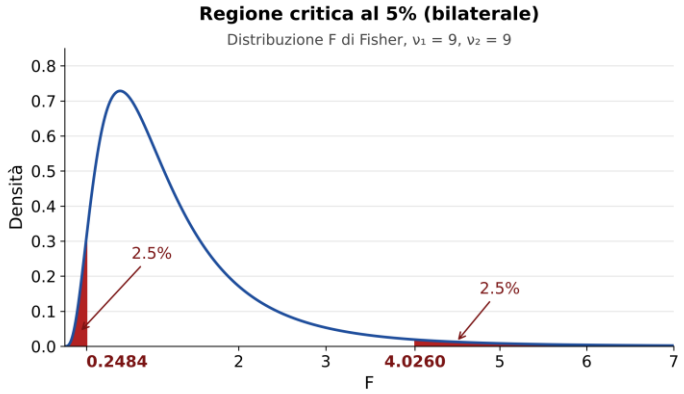
F-Test sulla varianza

Confronto della dispersione delle misure di due analisti. Si assumono ripetizioni indipendenti e popolazioni normali.

$H_0: \sigma_A^2 = \sigma_B^2 \quad H_1: \sigma_A^2 \neq \sigma_B^2$

$$F = \frac{0.7448889}{0.4601111} \approx 1.6189$$

$$\nu_A = \nu_B = 9$$



A	B
10.2	9.7
10.7	9
10.5	10.2
9.9	10.3
9	10.8
11.2	11.1
11.5	9.4
10.9	9.2
8.9	9.8
10.6	10.2

$$F_{0.025,9,9} \approx 0.2484 < F < F_{0.975,9,9} \approx 4.0260$$

Al 5% non si rifiuta l'uguaglianza delle varianze. Non emerge una differenza significativa nella dispersione; questo esito non prova l'uguaglianza delle precisioni.

F-Test sulla varianza

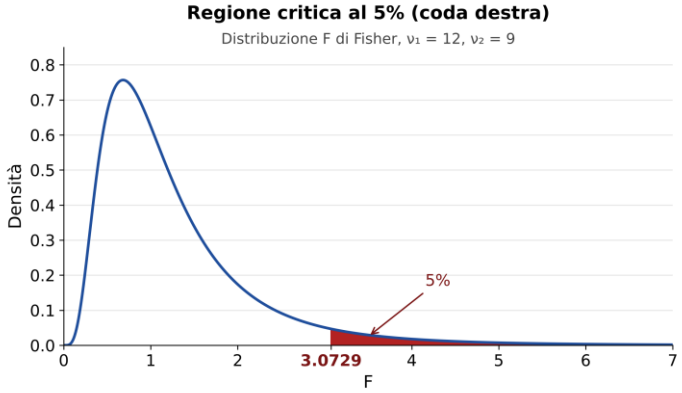
Si assumono ripetizioni indipendenti e misure normalmente distribuite sullo stesso materiale. A: metodo Scheibler. B: neutralizzazione dell'acido. Si verifica se B ha maggiore dispersione.

$$H_0: \sigma_B^2 \leq \sigma_A^2 \quad H_1: \sigma_B^2 > \sigma_A^2$$

$$F = \frac{0.1557692}{0.00988889} \approx 15.7519$$

$$\nu_B = 12, \qquad \nu_A = 9$$

$$F > F_{0.95,12,9} \approx 3.0729$$



A	B
2.5	1.7
2.4	1.9
2.5	2.3
2.6	2.3
2.5	2.8
2.5	2.5
2.4	1.6
2.6	1.9
2.7	2.6
2.4	1.7
	2.4
	2.2
	2.6

Al 5% si rifiuta l'ipotesi nulla: B mostra una dispersione maggiore nelle condizioni della prova. Il test confronta la precisione, non l'accuratezza complessiva dei metodi.

Test t di Student per dati appaiati

Il test t appaiato analizza le differenze fra due misure sullo stesso oggetto o individuo. Le coppie devono essere indipendenti fra loro; i due valori nella coppia possono essere dipendenti.

$$d_i = x_i - y_i, \qquad i = 1, \dots, n$$

$$\overline{d} = \frac{\sum_{i=1}^{n} d_i}{n} \qquad s_d = \sqrt{\frac{\sum_{i=1}^{n} \left(d_i - \overline{d}\right)^2}{n-1}}$$

$$\mathrm{SE}_d = \frac{s_d}{\sqrt{n}}$$

Per il test esatto conta la normalità delle differenze, non la normalità delle due serie considerate separatamente.

Test t di Student per dati appaiati

La differenza ipotizzata è un valore fissato prima dell'analisi, spesso zero. Si verifica la media della popolazione delle differenze.

$$H_0\colon \mu_d = \Delta_0 \quad t = \frac{\overline{d} - \Delta_0}{\mathrm{SE}_d} \quad \nu = n - 1$$

$$H_1\colon \mu_d > \Delta_0 \quad t > t_{1-\alpha,\nu}$$

$$H_1\colon \mu_d < \Delta_0 \quad t < t_{\alpha,\nu}$$

$$H_1\colon \mu_d \neq \Delta_0 \quad |t| > t_{1-\alpha/2,\nu}$$

$$\overline{d} < \Delta_0 + t_{\alpha/2,\nu}\mathrm{SE}_d \text{ oppure } \overline{d} > \Delta_0 + t_{1-\alpha/2,\nu}\mathrm{SE}_d$$

L'ultima riga è la forma equivalente per il test bilaterale. Con molti dati il test può essere approssimato se è adeguata la distribuzione della media delle differenze.

Test t di Student per campioni indipendenti

Due campioni indipendenti da popolazioni normali con varianza comune. La deviazione standard pooled stima la radice della varianza comune.

$$H_0\colon \mu_a = \mu_b \quad t = \frac{\overline{x_a} - \overline{x_b}}{\mathrm{SE}} \qquad {s_p}^2 = \frac{(n_a - 1)s_a^2 + (n_b - 1)s_b^2}{n_a + n_b - 2} \qquad \mathrm{SE} = s_p\sqrt{\frac{1}{n_a} + \frac{1}{n_b}}$$

$$\nu = n_a + n_b - 2$$

$$H_1\colon \mu_a > \mu_b \quad t > t_{1-\alpha,\nu}$$

$$H_1\colon \mu_a < \mu_b \quad t < t_{\alpha,\nu}$$

$$H_1\colon \mu_a \neq \mu_b \quad |t| > t_{1-\alpha/2,\nu}$$

La statistica conserva il segno. Il valore assoluto si usa nella regola bilaterale.

Test t di Welch per campioni indipendenti

Due campioni indipendenti, senza assumere varianze uguali. Il test t di Welch usa i gradi di libertà di Welch–Satterthwaite.

$$H_0\colon \mu_a = \mu_b \qquad t = \frac{\overline{x_a} - \overline{x_b}}{\mathrm{SE}} \qquad \mathrm{SE} = \sqrt{\frac{s_a^2}{n_a} + \frac{s_b^2}{n_b}} \qquad \nu = \frac{\mathrm{SE}^4}{\frac{s_a^4}{{n_a}^2(n_a - 1)} + \frac{s_b^4}{{n_b}^2(n_b - 1)}}$$

$H_1\colon \mu_a > \mu_b \quad t > t_{1-\alpha,\nu}$

$H_1\colon \mu_a < \mu_b \quad t < t_{\alpha,\nu}$

$H_1\colon \mu_a \neq \mu_b \quad |t| > t_{1-\alpha/2,\nu}$

I gradi di libertà possono essere reali: non si arrotondano nei software. Con campioni piccoli si assume normalità delle popolazioni; la calibrazione t di Welch è approssimata.

Confronto delle medie con campioni grandi

Con campioni indipendenti abbastanza grandi si può usare un'approssimazione normale per la differenza delle medie, senza assumere varianze uguali.

$$H_0\colon \mu_a = \mu_b \qquad z = \frac{\overline{x_a} - \overline{x_b}}{\mathrm{SE}} \qquad \mathrm{SE} = \sqrt{\frac{s_a^2}{n_a} + \frac{s_b^2}{n_b}}$$

$$H_1\colon \mu_a > \mu_b \quad z > z_{1-\alpha}$$

$$H_1\colon \mu_a < \mu_b \quad z < z_{\alpha}$$

$$H_1\colon \mu_a \neq \mu_b \quad |z| > z_{1-\alpha/2}$$

La sola soglia di 30 dati per gruppo non garantisce l'approssimazione. Contano dipendenza, asimmetria, code e valori estremi.

In alternativa si usa Welch, con il suo errore standard e i suoi gradi di libertà.

Test t per dati appaiati

Esempio – Taratura di sensori di temperatura

Dati didattici di 10 sensori: errori di indicazione (°C) rispetto a un riferimento di 100 °C, misurati sullo stesso sensore prima e dopo la taratura.

Sensore	Prima (°C)	Dopo (°C)	Differenza (°C)
1	0.42	0.27	0.15
2	0.38	0.33	0.05
3	0.55	0.33	0.22
4	0.30	0.34	-0.04
5	0.47	0.38	0.09
6	0.60	0.42	0.18
7	0.35	0.34	0.01
8	0.51	0.37	0.14
9	0.33	0.36	-0.03
10	0.46	0.35	0.11

$$d_i = e_{\text{prima,i}} - e_{\text{dopo,i}} \quad H_0: \mu_d \leq 0 \quad H_1: \mu_d > 0$$

Test t per dati appaiati

Calcolo sulla media delle differenze

Si assumono normali le differenze e indipendenti i sensori. Il livello di significatività scelto prima della prova è il 5%.

Numero coppie	Media diff. (°C)	Dev. std. diff. (°C)	Errore std. (°C)
10	0.0880	0.08867	0.02804

$$t = \frac{\bar{d} - 0}{\frac{s_d}{\sqrt{n}}} \approx 3.1384 \qquad \nu = n - 1 = 9$$

$$t \approx 3.1384 > t_{0.95,9} \approx 1.8331 \qquad p \approx 0.00598 < \alpha = 0.05$$

Intervallo di confidenza al 95% per riduzione media dell'errore di indicazione

$$\left[\bar{d} - t_{0.95,9}\,\mathrm{SE}_d, +\infty\right) \text{°C} = [0.03660, +\infty) \text{°C}$$

La taratura riduce l'errore medio positivo al punto di 100 °C. La conclusione riguarda la media, non tutti i sensori né l'intera scala.

Test t per campioni indipendenti

Esempio - Sedi per cuscinetti di due lotti di alberi

Si misurano a 20 °C le sedi per cuscinetti di 8 alberi del lotto A e 10 del lotto B, con diametro nominale di 50 mm. Osservazioni indipendenti.

Pezzo	Lotto A (mm)	Lotto B (mm)
1	49.94	49.99
2	49.99	50.03
3	50.01	50.04
4	49.97	50.01
5	50.03	50.06
6	50.00	50.02
7	49.96	50.00
8	50.02	50.05
9	—	50.01
10	—	50.04

$$H_0: \mu_A = \mu_B \qquad H_1: \mu_A \neq \mu_B$$

Test t per campioni indipendenti

Modello: popolazioni normali, campioni indipendenti e varianza comune.

Lotto	Numero	Media (mm)	Dev. std. (mm)
A	8	49.9900	0.03117
B	10	50.0250	0.02273

$$s_p{}^2 = \frac{(n_A - 1)s_A{}^2 + (n_B - 1)s_B{}^2}{n_A + n_B - 2} \approx 0.0007156\ \text{mm}^2$$

$$t = \frac{\overline{x_A} - \overline{x_B}}{s_p\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \approx -2.7583 \qquad \nu = 8 + 10 - 2 = 16$$

$$|t| \approx 2.7583 > t_{0.975,16} \approx 2.1199 \qquad p \approx 0.01399 < \alpha = 0.05$$

$$IC_{95\%}(\mu_A - \mu_B) = \overline{x_A} - \overline{x_B} \pm t_{0.975,16}\, SE = [-0.06190, -0.00810]\ mm$$

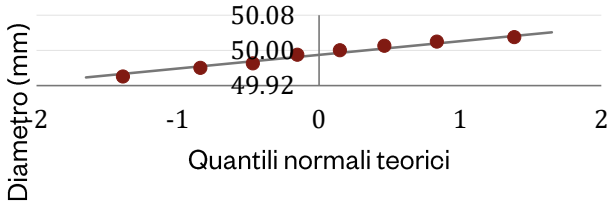
Al 5% si rifiuta l'ipotesi nulla. Il lotto A ha un diametro medio inferiore a B.

Controllo di normalità e uguaglianza delle varianze

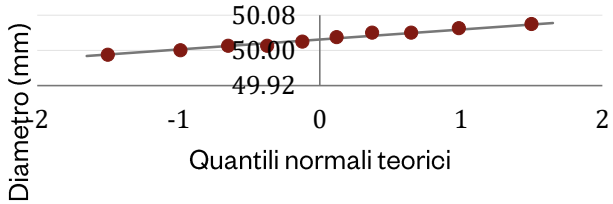
Esempio: diametri delle sedi per cuscinetti dei lotti A (n = 8) e B (n = 10)

Q–Q plot: un andamento quasi rettilineo è compatibile con la normale

Lotto A (n = 8)



Lotto B (n = 10)



Normalità: Shapiro–Wilk

H_0: popolazione normale, $\alpha = 0.05$

Lotto	n	W	p
A	8	0.9649	0.8555
B	10	0.9689	0.8809

Non si rifiuta la normalità in entrambi i lotti.

Varianze: F di Fisher, bilaterale

$$H_0: \sigma_A^2 = \sigma_B^2 \qquad H_1: \sigma_A^2 \neq \sigma_B^2$$

$$F = s_A^2 / s_B^2 = 1.8802$$

$$\nu_A = 7, \quad \nu_B = 9 \quad p = 0.3719 \quad \alpha = 0.05$$

Non si rifiuta l'ipotesi di varianze uguali.

Test F valido sotto normalità. Con pochi dati, p > 0.05 non dimostra le ipotesi.

Test t di Welch – varianze diverse

Esempio – Risposta al gradino termico

Si misura la costante di tempo (s) di 6 sensori A e 8 sensori B, provati separatamente con lo stesso gradino termico.

Sensore	Tipo A (s)	Tipo B (s)
1	10.1	10.0
2	9.8	12.8
3	10.4	11.2
4	10.0	13.1
5	9.9	9.5
6	10.2	12.0
7	—	10.7
8	—	12.4

Si confrontano le costanti di tempo medie senza assumere varianze uguali.

$$H_0: \mu_A = \mu_B \qquad H_1: \mu_A \neq \mu_B$$

Test t di Welch – varianze diverse

Si assumono gaussiane le costanti di tempo e indipendenti i due campioni.

Tipo	Numero	Media (s)	Dev. std. (s)
A	6	10.0667	0.21602
B	8	11.4625	1.32443

$$\mathrm{SE} = \sqrt{\frac{{s_A}^2}{n_A} + \frac{{s_B}^2}{n_B}} \approx 0.47649\ \mathrm{s} \qquad t = \frac{\overline{x_A} - \overline{x_B}}{\mathrm{SE}} \approx -2.9294$$

$$\nu = \frac{\mathrm{SE}^4}{\frac{{s_A}^4}{{n_A}^2(n_A - 1)} + \frac{{s_B}^4}{{n_B}^2(n_B - 1)}} \approx 7.4922$$

$$|t| \approx 2.9294 > t_{0.975,7.4922} \approx 2.3335 \qquad p \approx 0.02043 < \alpha = 0.05$$

$$IC_{95\%}(\mu_A - \mu_B) = \overline{x_A} - \overline{x_B} \pm t_{0.975,7.4922} SE = [-2.50772, -2.28395\ ]\ s$$

Si rifiuta l'uguaglianza delle medie al 5%: i sensori A hanno una risposta media più rapida.

Confronto delle medie – campioni grandi

Esempio – Temperatura a regime di cuscinetti

Si confrontano due produzioni di cuscinetti, provati allo stesso carico e nelle stesse condizioni ambientali. Dati: temperature a regime.

Produzione	Numero	Media (°C)	Dev. std. (°C)
A	80	72.6	4.2
B	100	71.4	5.1

$$H_0: \mu_A = \mu_B \qquad H_1: \mu_A \neq \mu_B$$

Campioni indipendenti. Si assume adeguata l'approssimazione normale delle medie, con varianze finite e senza forte asimmetria o valori estremi dominanti.

Confronto delle medie – campioni grandi

Approssimazione normale: calcolo e decisione

$$\mathrm{SE} = \sqrt{\frac{4.2^2}{80} + \frac{5.1^2}{100}} \approx 0.69325\ ^\circ\mathrm{C}$$

$$z = \frac{72.6 - 71.4}{\mathrm{SE}} \approx 1.7310$$

$$|z| \approx 1.7310 < z_{0.975} \approx 1.9600$$

$$p \approx 0.08346 > \alpha = 0.05$$

Intervallo di confidenza approssimato al 95% per la differenza delle medie

$$[1.2 \pm 1.959964 \times 0.693253]\ ^\circ\mathrm{C} \rightarrow [-0.1588, 2.5588]\ ^\circ\mathrm{C}$$

Al 5% non emerge una differenza significativa tra le temperature medie delle due produzioni.

Analisi della varianza (ANOVA)

Spesso si presenta la necessità di prendere in considerazione esperimenti od osservazioni relative **a più di due gruppi** individuati sulla base di un fattore di interesse.

L'**analisi della varianza** è una tecnica che consente di confrontare le medie di più di due gruppi (popolazioni):

- quando i gruppi sono definiti sulla base di un singolo fattore si parla di **analisi della varianza** a un fattore o **a una via** (*one way Anova*)
- nell'**analisi della varianza a più vie** si studiano più fattori che possono creare delle differenze tra le medie

È un'estensione **a più gruppi del test t-Student.**

Analisi della varianza (ANOVA)

$$\mathrm{SST} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \overline{y})^2$$

$$\mathrm{SSA} = \sum_{i=1}^{k} n_i (\overline{y_i} - \overline{y})^2$$

$$\mathrm{SSE} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \overline{y_i})^2$$

$$\mathrm{SST} = \mathrm{SSA} + \mathrm{SSE}$$

SSA, *l*a devianza fra gruppi, contiene sia l'eventuale effetto del fattore sia variabilità casuale.
SSE, la devianza interna, descrive la variabilità residua entro i gruppi.

Il numero di gruppi è indicato con k. L'ipotesi nulla è l'uguaglianza delle medie; l'alternativa ammette almeno una media diversa.

Analisi della varianza (ANOVA)

Le devianze si dividono per i rispettivi gradi di libertà per ottenere le medie dei quadrati.

$$\nu_{\text{tot}} = n_t - 1, \qquad \nu_A = k - 1, \qquad \nu_E = n_t - k$$

$$\text{MSA} = \frac{\text{SSA}}{k-1} \quad \text{MSE} = \frac{\text{SSE}}{n_t - k}$$

Nel modello ANOVA classico, MSE stima la varianza residua comune. MSA stima la stessa varianza soltanto sotto l'ipotesi di medie uguali.

Per il test F esatto si assumono errori indipendenti, normalmente distribuiti e con varianza comune.

Analisi della varianza (ANOVA)

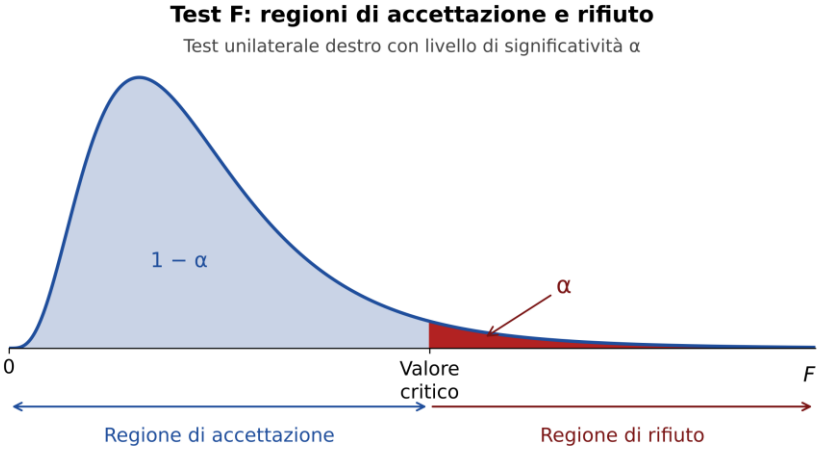
Per confrontare le medie di k gruppi si usa il rapporto fra le medie dei quadrati, nel modello con errori indipendenti, normali e omoschedastici.

$$H_0\colon\ \mu_1 = \cdots = \mu_k$$

$$F = \frac{\mathrm{MSA}}{\mathrm{MSE}} \sim F_{k-1,n_t-k}$$

Distribuzione sotto l'ipotesi nulla. Si rifiuta per valori grandi del rapporto, nella sola coda destra.

$$F > F_{1-\alpha,k-1,n_t-k}$$



Analisi della varianza (ANOVA)

Esempio - ANOVA a una via - Usura di componenti

Si misura la perdita di massa in mg dopo una prova di usura identica su provini metallici. Il controllo comprende 7 provini senza trattamento, A e B comprendono 4 e 5 provini con due trattamenti superficiali. Le perdite medie differiscono fra i tre gruppi?

Perdita di massa dopo la prova (mg)

Controllo	A	B
52	22	31
30	25	28
27	17	43
42	32	40
25		36
38		
29		

Si assume il modello classico con errori indipendenti, normali e con varianza comune.

Analisi della varianza (ANOVA)

Esempio di usura con tre gruppi

$$k = 3, \qquad n_C = 7, \qquad n_A = 4, \qquad n_B = 5, \qquad n_t = 16$$

$$\overline{y} = 32.3125, \qquad \overline{y_C} \approx 34.7143, \qquad \overline{y_A} = 24.0, \qquad \overline{y_B} = 35.6$$

$$\mathrm{SSA} \approx 370.8089 \qquad \mathrm{SSE} \approx 842.6286$$

$$\nu_A = k - 1 = 2, \qquad \nu_E = n_t - k = 13$$

$$\mathrm{MSA} = \frac{\mathrm{SSA}}{2} \approx 185.4045 \quad \mathrm{MSE} = \frac{\mathrm{SSE}}{13} \approx 64.8176$$

Le medie sono in mg, devianze e medie dei quadrati in mg^2. I calcoli usano i dati non arrotondati.

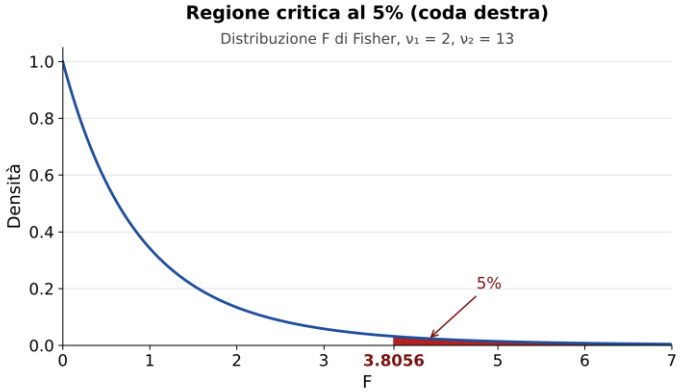
Analisi della varianza (ANOVA)

Esempio di usura - Decisione al 5%

$$F = \frac{\text{MSA}}{\text{MSE}} \approx 2.8604$$

$$F_{0.95,2,13} \approx 3.8056$$

$$p = \int_{2.8604}^{\infty} f_F(x; 2,13)dx \approx 0.09344$$



Non si rifiuta l'uguaglianza delle medie al livello del 5%. Non emerge evidenza sufficiente di differenze tra le perdite medie dei tre gruppi.

Analisi della varianza (ANOVA)

Una ANOVA significativa indica almeno una differenza fra le medie, ma non identifica le coppie diverse.

Per più confronti si definisce la famiglia delle ipotesi e si controlla la probabilità di almeno un errore di tipo I.

La procedura di Bonferroni è una correzione standard che garantisce un errore familiare non superiore al livello fissato.

$$p_j \leq \frac{\alpha}{c}, \qquad j = 1, \dots, c$$

Esempio con cinque confronti pianificati e livello familiare del 5%

$$c = 5, \qquad \alpha = 0.05, \qquad \frac{\alpha}{c} = 0.01$$

Chi-Square Goodness-of-Fit Test

Ipotesi nulla: la popolazione segue una distribuzione normale, con la densità correttamente normalizzata.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$$

$$E_j = np_j \quad Q = \sum_{j=1}^{m} \frac{(O_j - E_j)^2}{E_j}$$

Le classi coprono tutta la distribuzione, comprese le code. Si controllano le frequenze attese e si usano classi e criteri di raggruppamento prefissati.

$$\nu = m - 1 - r, \qquad Q > \chi^2_{1-\alpha,\nu}$$

r è il numero di parametri stimati dai conteggi raggruppati. Per la normale si stimano due parametri; se entrambi sono noti non si sottraggono parametri stimati.

Chi-Square Goodness-of-Fit Test

Esempio Si pesano 100 componenti torniti con una bilancia. Le masse, in grammi, sono riportate sotto. Si applica il test chi-quadro per valutare l'adattamento a una distribuzione gaussiana.

487	504	501	515	491	496	482	502	508	494	505	501	485	503	507	494	489	501	510	491
503	492	483	501	500	493	505	501	517	500	494	503	500	488	496	500	519	499	495	490
503	500	497	492	510	506	497	499	489	506	502	484	495	498	502	496	512	504	490	497
488	503	512	497	480	509	496	513	499	502	487	499	505	493	498	508	492	498	486	511
499	504	495	500	484	513	509	497	505	510	516	499	495	507	498	514	506	500	508	494

Classi dell'istogramma secondo la regola di Sturgis

$$m = 1 + 3.3 \log_{10}(n) = 7.6 \rightarrow 8$$

Le 8 classi descrivono l'istogramma. Per il test si usa la partizione di 6 classi con code estese della slide successiva. Sturgis non garantisce l'applicabilità del test.

Classe	1	2	3	4	5	6	7	8
Intervallo (g)	479.5–484.5	484.5–489.5	489.5–494.5	494.5–499.5	499.5–504.5	504.5–509.5	509.5–514.5	514.5–519.5
Frequenza	5	8	13	23	24	14	9	4

Chi-Square Goodness-of-Fit Test

Per l'esempio si assume prefissata la partizione seguente, con code fino a meno e più infinito. Tutte le frequenze attese sono almeno 5.

Stima di massima verosimiglianza dai conteggi raggruppati

$\hat{\mu} \approx 499.6043$ g, $\hat{\sigma} \approx 8.6894$ g

Classe di massa (g)	Osservata	Attesa	Contributo a Q
Fino a 489.5	13	12.24492	0.04656
489.5–494.5	13	15.60123	0.43371
494.5–499.5	23	21.67491	0.08101
499.5–504.5	24	21.82105	0.21758
504.5–509.5	14	15.91897	0.23132
Oltre 509.5	13	12.73893	0.00535

$Q \approx 1.01554$, $\nu = 6 - 1 - 2 = 3$

Chi-Square Goodness-of-Fit Test

Decisione per la partizione di sei classi

$$Q \approx 1.01554 < \chi^2_{0.95,3} \approx 7.81473$$

$$p \approx 0.79749 > \alpha = 0.05$$

Non si rifiuta l'adattamento al modello normale al livello del 5%. Il test non dimostra la normalità della popolazione.

La calibrazione chi-quadro è asintotica. Qui la partizione si tratta come fissata prima del test e i parametri si stimano dai conteggi delle sei classi.

Se confini o raggruppamenti vengono scelti dai dati, occorre verificare la calibrazione simulando l'intera procedura di scelta e stima.

Test di Kolmogorov-Smirnov (buon adattamento)

Il test K-S classico a un campione verifica un modello continuo completamente specificato, con osservazioni i.i.d.

$H_0: F = F_0 \qquad H_1: F \neq F_0$

Per i dati ordinati la ripartizione empirica è

$$F_n(x) = 0, \qquad x < x_{(1)}$$

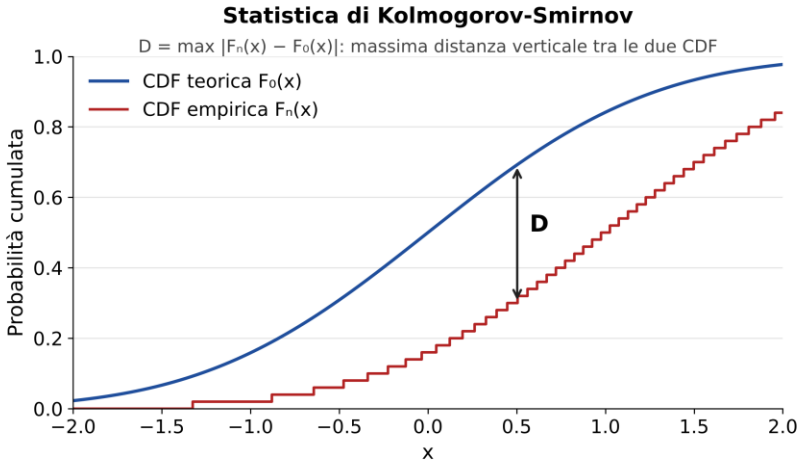
$$F_n(x) = \frac{k}{n}, \qquad x_{(k)} \leq x < x_{(k+1)}$$

$$F_n(x) = 1, \qquad x \geq x_{(n)}$$

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{x_i \leq x}$$

La forma con le indicatrici gestisce anche i valori ripetuti. Le soglie del K-S classico richiedono però un modello nullo continuo senza parametri stimati dagli stessi dati.

Test di Kolmogorov-Smirnov (buon adattamento)



La freccia mostra una distanza verticale illustrativa.
La statistica del test è il massimo delle distanze su tutto l'asse, non una distanza scelta in un punto qualsiasi.

Test di Kolmogorov-Smirnov (buon adattamento)

Se una prova ha probabilità di successo p, la somma di n indicatori indipendenti segue una distribuzione binomiale.

$$q = 1 - p, \qquad K \sim Bin(n, p)$$

$$P(K = k) = \frac{n!}{k!\,(n-k)!} p^k q^{n-k}, \qquad k = 0, \dots, n$$

$$E(K) = np, \qquad Var(K) = npq$$

La probabilità è zero per gli altri valori interi. Un singolo indicatore è Bernoulli; la somma degli indicatori è binomiale.

Test di Kolmogorov-Smirnov (buon adattamento)

Per un punto x fissato e un campione i.i.d., ogni indicatore assume i valori zero e uno. La probabilità di successo è la vera ripartizione della popolazione.

$$I_{X_i \leq x} \sim \text{Bernoulli}(F(x)) \qquad \sum_{i=1}^{n} I_{X_i \leq x} \sim \text{Bin}(n, F(x))$$

$$E(F_n(x)) = F(x) \qquad \text{Var}(F_n(x)) = \frac{F(x)[1 - F(x)]}{n}$$

La ripartizione empirica è uno stimatore puntualmente non distorto e consistente in media quadratica della vera F. Una sua realizzazione è la stima ottenuta dai dati osservati.

Test di Kolmogorov-Smirnov (buon adattamento)

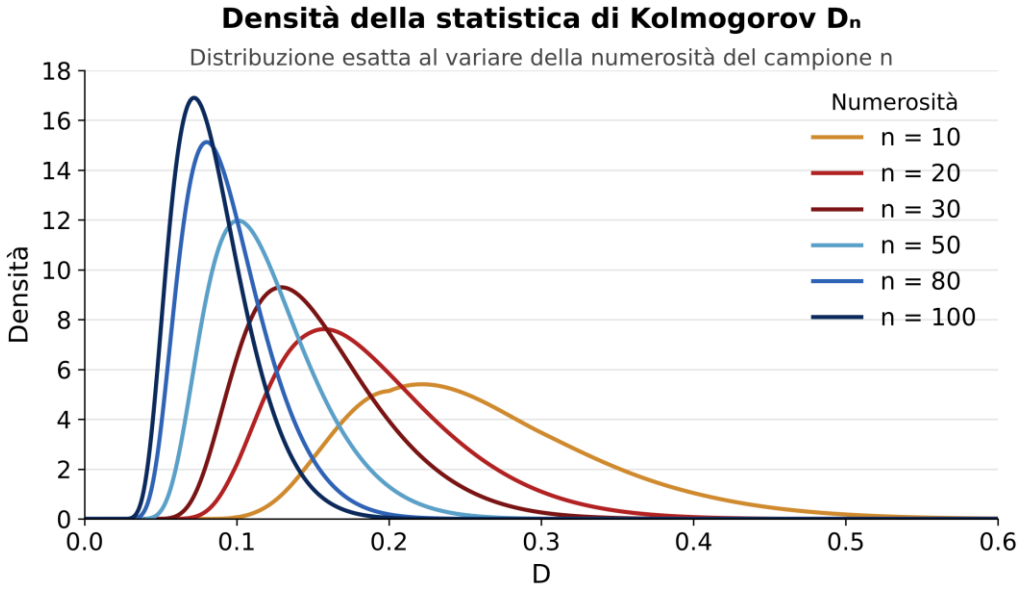
Il K-S misura la massima distanza fra due funzioni di ripartizione.

$$D_n = \sup_{x \in \mathbb{R}} |F_n(x) - F_0(x)|$$

$$D_n > q_{KS,1-\alpha,n}$$

Si rifiuta per distanze oltre la soglia.

Le densità in figura mostrano la dipendenza dalla numerosità. Si considera un modello continuo completamente specificato.



Test K-S con parametri stimati e bootstrap

Esempio Primo campione di 100 masse in g. Per l'esempio si assume arrotondamento al grammo più vicino. Il bootstrap parametrico riproduce l'arrotondamento e ristima i parametri in ogni replica.

487	504	501	515	491	496	482	502	508	494	505	501	485	503	507	494	489	501	510	491
503	492	483	501	500	493	505	501	517	500	494	503	500	488	496	500	519	499	495	490
503	500	497	492	510	506	497	499	489	506	502	484	495	498	502	496	512	504	490	497
488	503	512	497	480	509	496	513	499	502	487	499	505	493	498	508	492	498	486	511
499	504	495	500	484	513	509	497	505	510	516	499	495	507	498	514	506	500	508	494

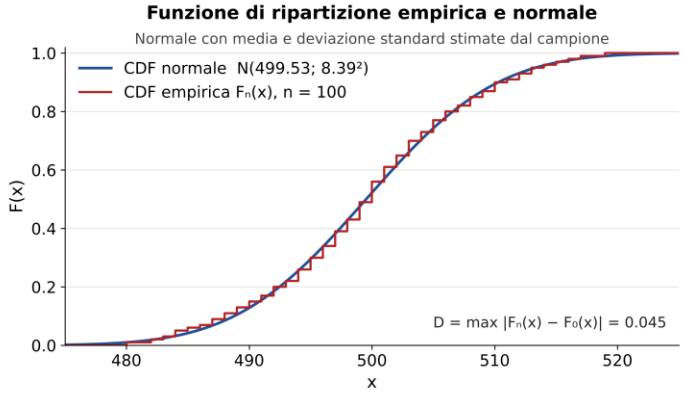
Stima del modello normale latente da dati arrotondati

$\hat{\mu} \approx 499.53\ \text{g}, \qquad \hat{\sigma} \approx 8.39\ \text{g}$

$D_n \approx 0.0447, \qquad p_{\text{MC}} \approx 0.99300$

$q_{\text{boot},0.95} \approx 0.10177$

19 999 repliche. Al 5% non si rifiuta il modello normale.
Nota che il risultato non dimostra la normalità.



Test K-S con parametri stimati e bootstrap

Secondo campione di 100 masse (g). Si assume arrotondamento al grammo più vicino e si applica lo stesso bootstrap del modello normale latente, ristimando i parametri.

480	480	480	480	480	480	480	480	480	480	480	480	480	480	480	480	480	480	480	480
480	480	480	480	480	480	480	480	480	480	480	480	480	480	480	480	480	480	480	480
480	480	480	480	485	485	485	485	485	485	485	485	485	505	505	505	505	505	505	505
505	505	505	505	519	519	519	519	519	519	519	519	519	519	519	519	519	519	519	519
519	519	519	519	519	519	519	519	519	519	519	519	519	519	519	519	519	519	519	519

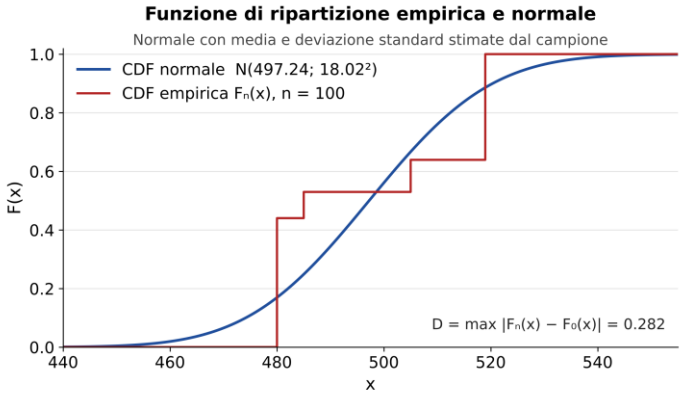
Stima del modello normale latente da dati arrotondati

$\hat{\mu} \approx 497.24$ g, $\hat{\sigma} \approx 18.02$ g

$D_n \approx 0.282$ $p_{\mathrm{MC}} = \dfrac{1}{20000} = 0.00005$

$q_{\mathrm{boot},0.95} \approx 0.09285$

Su 19 999 repliche, nessuna supera la distanza osservata.
Al 5% si rifiuta il modello.



Correlazione lineare e indipendenza

Il **coefficiente di Pearson** descrive l'associazione lineare tra variabili con varianze finite e positive.

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sigma_x \sigma_y}, \qquad |\rho| \leq 1$$

$$|\rho| = 1 \quad \text{implica} \quad Y = aX + b \text{ quasi certamente}$$

Con correlazione nulla le variabili sono scorrelate. L'indipendenza implica correlazione nulla, ma in generale il contrario non vale.

Nel caso congiuntamente gaussiano, la correlazione nulla equivale all'indipendenza. Per altri modelli Pearson verifica l'associazione lineare.

Normalità congiunta e indipendenza

Coppie i.i.d. da una normale bidimensionale non degenere.

$$\sigma_x > 0, \qquad \sigma_y > 0, \qquad |\rho| < 1, \qquad u = \frac{x - \mu_x}{\sigma_x}, \qquad v = \frac{y - \mu_y}{\sigma_y}$$

$$f_{XY}(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left[\frac{-(u^2 + v^2 - 2\rho uv)}{2(1-\rho^2)}\right]$$

Con correlazione nulla l'esponenziale si separa nel prodotto di due densità normali.

$$\rho = 0: \quad f_{XY}(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[\frac{-(u^2 + v^2)}{2}\right] = f_X(x) f_Y(y)$$

Nel modello congiuntamente gaussiano, indipendenza e correlazione nulla sono equivalenti. La normalità delle sole marginali non basta.

Test di correlazione nel modello gaussiano congiunto

Coppie i.i.d. da una normale bidimensionale non degenere. Sotto l'ipotesi di correlazione nulla la statistica ha distribuzione t centrale.

$$H_0\colon \rho = 0, \qquad R = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}}$$

$$\mathrm{ST} = \frac{R\sqrt{n-2}}{\sqrt{1-R^2}} \sim t_{n-2}, \qquad n \geq 3$$

$H_1\colon \rho > 0 \quad \mathrm{ST} > t_{1-\alpha,n-2}$

$H_1\colon \rho < 0 \quad \mathrm{ST} < t_{\alpha,n-2}$

$H_1\colon \rho \neq 0 \quad |\mathrm{ST}| > t_{1-\alpha/2,n-2}$

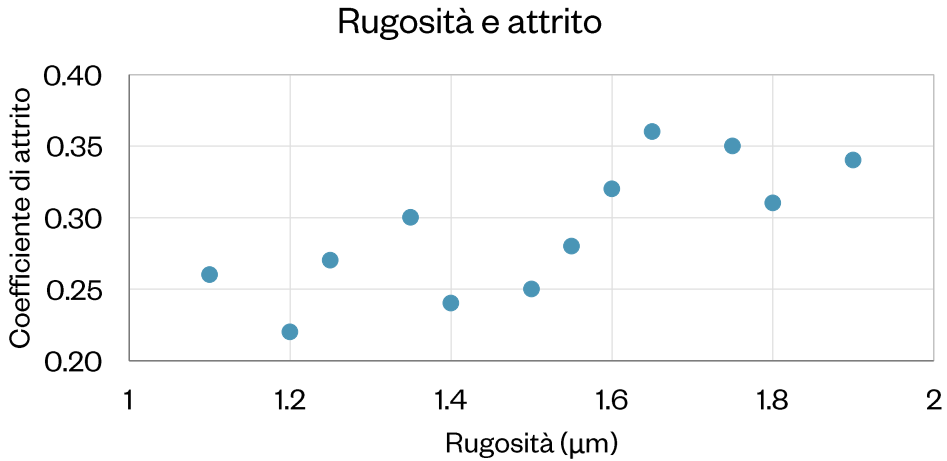
R è lo stimatore campionario di ρ. Fuori dal modello gaussiano congiunto il test riguarda la correlazione lineare.

Esempio di correlazione nel modello gaussiano congiunto

Esempio – Rugosità superficiale e attrito

Dati su 12 provini: rugosità (µm) e coefficiente di attrito, a parità di lubrificazione.

Provino	Rugosità (µm)	Attrito
1	1.10	0.26
2	1.55	0.28
3	1.20	0.22
4	1.75	0.35
5	1.35	0.30
6	1.60	0.32
7	1.90	0.34
8	1.40	0.24
9	1.80	0.31
10	1.25	0.27
11	1.50	0.25
12	1.65	0.36



$H_0: \rho = 0 \quad H_1: \rho \neq 0$

Esempio di correlazione nel modello gaussiano congiunto

Coppie indipendenti e normalità congiunta. Test bilaterale con livello di significatività del 5%.

Numero coppie	Media rugosità (µm)	Media attrito	Correlazione R
12	1.50417	0.29167	0.76003

$$R = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \approx 0.76003$$

$$\mathrm{ST} = R\sqrt{\frac{n-2}{1-R^2}} \approx 3.6982 \qquad \nu = n - 2 = 10$$

$$|\mathrm{ST}| \approx 3.6982 > t_{0.975,10} \approx 2.2281$$

$$p \approx 0.00412 < \alpha = 0.05$$

Nel modello gaussiano congiunto si rifiuta l'indipendenza: l'associazione lineare è positiva.

UNIVERSITÀ
DEGLI STUDI
DI TRIESTE

Dipartimento di
Ingegneria
e Architettura