t-test appaiato

Vi sono casi in cui il confronto fra i responsi di due metodi di analisi non può essere effettuato replicando più volte l'analisi con ciascun metodo su uno stesso campione, ad esempio quando:

- ✓ la quantità disponibile di campione è sufficiente solo per una determinazione con ciascuno dei due metodi;
- ✓ sono disponibili soltanto piccole quantità di campioni provenienti da fonti diverse e/o contenenti quantità di analita leggermente diverse;
- ✓ i campioni per il confronto sono analizzati su un grande intervallo di tempo, per cui è difficile ottenere replicati affidabili.

In questo caso le eventuali differenze fra le medie ottenute dai due metodi possono essere mascherate da altri effetti, per cui non possono essere stimate con l'approccio dell'intervallo di fiducia.

Un test che consente di affrontare la problematica è il t-test appaiato.

- ✓ Si effettuano n misure accoppiate, una per ciascun campione analizzato con ciascuno dei due metodi

- ✓ si calcolano le differenze D_i fra i dati di ciascuna coppia di misure

- ✓ nell'ipotesi che non vi sia una differenza statisticamente significativa fra i due metodi i valori D_i saranno distribuiti intorno ad un valore medio 0.

Si possono dunque mettere a confronto le seguenti ipotesi, detta μ_D la media di popolazione delle differenze fra i dati:

$\mu_D = 0$ e $\mu_D \neq 0$

I valori delle differenze D_i possono essere usati come quelli derivanti da un campione di dimensioni n (di solito $n < 30$).

La situazione è analoga a quella del confronto fra una media ed un valore noto (pari a 0) nel caso 3: popolazione distribuita normalmente, con varianza ignota ed $n < 30$.

La variabile da considerare è, quindi:

$$\boxed{t = \frac{\bar{D} - 0}{s_D / \sqrt{n}} \sim t_{n-1}}$$

dove s_D è la deviazione standard relativa alle differenze D_i, ossia:

$$\boxed{s_D = \sqrt{\frac{\sum_{i=1}^{n}(D_i - \overline{D})^2}{n-1}}}$$

Scelto un livello di significatività α, l'ipotesi $\mu_D = 0$ viene rigettata se:

$$\boxed{|t| \geq t_{(1-\alpha/2),n-1}}$$

NOTA IMPORTANTE: Il t-test appaiato mantiene validità **SOLTANTO** se vale l'assunzione che l'errore sulle determinazioni messe a confronto non dipenda dalla concentrazione.

Se ciò non è vero il confronto fra i metodi va fatto con la regressione lineare, come verrà spiegato in seguito.

Esempio numerico di t-test appaiato

Le concentrazioni di glucosio in dieci campioni di plasma umano sono state determinate con i seguenti metodi analitici:

1) un test fotometrico-enzimatico di routine

2) un metodo di analisi per iniezione in flusso (FIA) basato sulla rivelazione amperometrica di H_2O_2 su un biosensore con *glucosio ossidasi* immobilizzata

I dati ottenuti (espressi in mg/100 mL) sono:

Campione	Metodo fotometrico	Metodo FIA	D_i
1	75	70	+5
2	100	103	-3
3	82	83	-1
4	85	82	+3
5	93	94	-1
6	78	77	+1
7	80	83	-3
8	90	88	+2
9	84	86	-2
10	95	94	+1

Le ipotesi a confronto sono:

$\mu_D = 0$
$\mu_D \neq 0$

Risulta inoltre: $n = 10$ $\overline{D} = 0.20$ $s_D = 2.66$

La realizzazione della statistica è:

$$\boxed{t = \frac{\overline{D}}{s_D / \sqrt{n}} = 0.24}$$

Al 5% di significatività ($1-\alpha/2 = 0.975$) risulta:

$t_{0.975,9} = 2.26$

Poiché $t = 0.24 < 2.26$ l'ipotesi $\mu_D = 0$ può essere accettata e quindi i due metodi non differiscono in modo significativo fra di loro.

Test sulle varianze o loro differenze

In analogia con quanto visto per la media e la differenza fra medie, è possibile sottoporre a confronto anche una varianza o una differenza fra varianze.

Test su una singola varianza

Supponiamo di avere a disposizione i seguenti dati

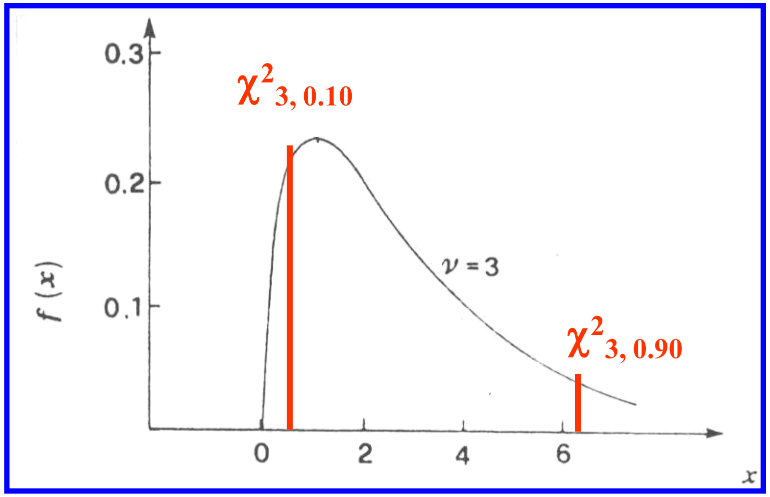
$x_1, x_2, \ldots, x_n$

estratti da una popolazione distribuita secondo $N(\mu, \sigma^2)$.

Per una delle proprietà della distribuzione χ^2 vale la relazione:

$$\Sigma_i (X_i - \overline{X})^2/\sigma^2 = (n-1)s^2/\sigma^2 \sim \chi^2_{n-1}$$

si può quindi effettuare il confronto fra una varianza ed un valore noto usando la distribuzione chi-quadro.



Poiché la distribuzione χ^2 non è simmetrica i criteri di rigetto dell'ipotesi di base (l'uguaglianza con un valore predefinito) sono diversi da quelli che riguardavano le distribuzioni normale e t di Student:

ipotesi a confronto	distribuzione dei dati tipo di test	criteri di rigetto dell'ipotesi di base
$\sigma^2 = \sigma_0^2$	$T = (n-1)s^2/\sigma_0^2 \sim \chi^2_{n-1}$	
$\sigma^2 > \sigma_0^2$	una coda	$t \geq \chi^2_{n-1\ (1-\alpha)}$
$\sigma^2 < \sigma_0^2$	una coda	$t \leq \chi^2_{n-1(\alpha)}$
$\sigma^2 \neq \sigma_0^2$	due code	$t \leq \chi^2_{n-1\ (\alpha/2)}$ oppure $t \geq \chi^2_{n-1\ (1-\alpha/2)}$

Test su una varianza: esempio numerico

Siano dati i seguenti numeri, ottenuti da un programma che genera numeri random:

+0.250	+1.620	-0.052	+0.014	-0.366
+0.756	+0.608	-2.150	+1.162	

Si vuole verificare se la varianza sia o meno uguale a 1.

In questo caso:

$n = 9$
$s^2 = 1.17$
$\sigma_0^2 = 1$

La realizzazione della statistica per il test della varianza (a due code) è:

$$\boxed{t = (n-1)s^2/\sigma_0^2 = 9.36}$$

Se si adotta un livello di significatività del 5 %, risulta:

$\chi^2_{n-1\ (\alpha/2)} = \chi^2_{8\ (0.025)} = 2.180$ e

$\chi^2_{n-1\ (1-\alpha/2)} = \chi^2_{8\ (0.975)} = 17.535$

Poiché $2.180 < t < 17.535$ possiamo accettare l'ipotesi di base e quindi affermare che la varianza della popolazione dei numeri random non è significativamente diversa da 1.

Test su due varianze (F test)

Supponiamo di avere a disposizione le seguenti serie di dati:

$x_1, x_2, \ldots, x_{n1}$ $\qquad\qquad$ $y_1, y_2, \ldots, y_{n2}$

derivanti da popolazioni distribuite normalmente:

$x \sim N(\mu_1, \sigma_1^2)$ e $y \sim N(\mu_2, \sigma_2^2)$

Può essere necessario stabilire su base statistica se le due varianze di popolazione siano significativamente diverse l'una dall'altra.

Ad esempio se le due serie di dati derivano dall'applicazione di metodi analitici diversi il confronto consente di stabilire se uno dei metodi è più preciso dell'altro o se essi sono equivalenti, in termini di precisione.

Per il test si può sfruttare la proprietà della distribuzione F:

$$\boxed{(s_1^2/\sigma_1^2)/(s_2^2/\sigma_2^2) \sim F_{n1-1,\, n2-1}}$$

Se $\sigma_1^2 = \sigma_2^2$ risulta: $T = (s_1^2/s_2^2) \sim F_{n1-1, n2-1}$

Per comodità si calcola la realizzazione di T in modo che al numeratore ci sia sempre la varianza campionaria più grande fra le due a confronto.

ν_{num} e ν_{den} sono i gradi di libertà associati alle serie di dati a cui corrispondono rispettivamente la varianza al numeratore e quella al denominatore del rapporto T

I criteri di rigetto dell'ipotesi di base, ossia dell'uguaglianza fra le due varianze, sono, a seconda dei casi:

ipotesi a confronto	distribuzione dei dati tipo di test	criteri di rigetto dell'ipotesi di base
$\sigma_1^2 = \sigma_2^2$	$T = s^2_{num}/s^2_{den} \sim F_{\nu_{num}, \nu_{den}}$	
$\sigma_1^2 > \sigma_2^2$	una coda	$t \geq F_{\nu_{num}, \nu_{den}}(1-\alpha)$
$\sigma_1^2 < \sigma_2^2$	una coda	$t \geq F_{\nu_{num}, \nu_{den}}(1-\alpha)$
$\sigma_1^2 \neq \sigma_2^2$	due code	$t \geq F_{\nu_{num}, \nu_{den}}(1-\alpha/2)$

Si noti che il tipo di disequazione finale è sempre lo stesso proprio per il modo in cui è stata stabilita la statistica.

F test: esempio numerico

Supponiamo di avere a disposizione i seguenti set di dati, derivanti dall'applicazione di due diversi metodi analitici allo stesso campione:

Dato	Metodo A	Metodo B
1	5.11	5.18
2	5.14	5.13
3	5.13	5.27
4	5.17	5.12
5	5.12	5.27
6	5.08	5.29
7	5.15	5.17
8	5.20	5.28
9	5.16	5.14
10	5.14	5.18

<u>Metodo A</u>

$\overline{X} = 5.140$

$s_1^2 = 11.11 \times 10^{-4}$

$n_1 = 10$

<u>Metodo B</u>

$\overline{Y} = 5.203$

$s_2^2 = 45.34 \times 10^{-4}$

$n_2 = 10$

Si vuole verificare se esista una differenza significativa fra la precisione dei due metodi, e quindi fra le varianze delle due popolazioni.

Il test da utilizzare è tipicamente a due code, in questo caso.

Poiché la varianza campionaria più grande è quella derivante dal metodo B la realizzazione della statistica per l'F-test è:

$t = 45.34/11.11 = 4.08.$

Se si adotta un livello di significatività del 5% (e quindi $1-\alpha/2 = 0.975$) il valore critico è:

$F_{9,9\ (0.975)} = 4.03$

Poiché $t = 4.08 > 4.03$ l'ipotesi di base è rigettata, quindi le due varianze sono significativamente diverse l'una dall'altra.

In particolare il metodo B ha una varianza maggiore, quindi è meno preciso del metodo A.

Test per l'eliminazione dei *dati aberranti*

In termini statistici un dato aberrante (outlier) è un dato che appare chiaramente diverso dagli altri ottenuti da una serie di misure replicate.
Ad esempio può essere stato generato da un errore grossolano nella procedura di misura o da una fluttuazione accidentale nel segnale fornito dalla strumentazione.

Poiché eliminare un outlier prima di effettuare ulteriori calcoli sui dati a cui esso appartiene (ad esempio quelli dei parametri campionari) o test (come il test d'ipotesi) può avere conseguenze anche notevoli sull'esito di tali operazioni, è necessario effettuare dei test che giustifichino tale decisione.

Q-test di Dixon

Il Q-test di Dixon è uno dei possibili test per l'eliminazione dei dati aberranti e si basa sul confronto della differenza fra il valore sospetto e quello ad esso più vicino nella serie di dati con l'intero campo di variazione dei dati stessi. L'assunzione è che i dati siano distribuiti normalmente.

Si calcola la seguente statistica:

$$Q = |\text{valore sospetto} - \text{valore più vicino}| \, / \, |\text{valore massimo} - \text{valore minimo}|$$

La realizzazione di Q, q, viene confrontata con un valore critico al livello di fiducia desiderato, di solito 95%.

Se $q >$ valore critico $\Rightarrow$ il dato può essere rigettato

Se $q <$ valore critico $\Rightarrow$ il dato va conservato nella serie

Dimensioni del campione (n)	Valore critico a P = 0.95
4	0.831
5	0.717
6	0.621
7	0.570
8	0.524
9	0.492
10	0.464

Esempio numerico

Supponiamo di avere a disposizione i seguenti 4 dati di volume di titolante aggiunto al punto equivalente in una titolazione:

20.85 20.80 20.95 21.35

Si sospetta che l'ultimo dato sia un outlier.

La realizzazione di Q è:

$$\boxed{q = (21.35-20.95)/(21.35-20.80) = 0.727}$$

Per n = 4, al 95% di fiducia, il valore critico di Q è 0.831;

poiché $0.727 < 0.831$ il dato 21.35 va conservato nella serie.

Supponiamo ora di aggiungere altri dati alla serie, effettuando tre nuove misure:

20.85 20.80 20.95 21.35 20.70 20.90 20.82

Il nuovo valore della realizzazione di Q è dato da:

$$\boxed{q = (21.35-20.95)/(21.35-20.70) = 0.615}$$

Per n = 7, al 95% di fiducia, il valore critico di Q è 0.570;

poiché $0.615 > 0.570$ il dato 21.35 va eliminato dalla serie.

E' evidente che un numero maggiore di misure consente di prendere la decisione sull'outlier in modo più appropriato.

Va sottolineata la differenza fra i parametri campionari prima e dopo l'eliminazione dell'outlier:

$\bar{x}$ passa da 20.91 a 20.84, ma soprattutto s passa da 0.21 a 0.09!

Va sottolineato che il test finora descritto è efficace nel caso in cui sia presente un solo dato sospetto nella serie, ad un'estremità o all'altra.

Se sono presenti due o più dati anomali in corrispondenza della stessa estremità della serie occorre applicare delle varianti del test, che prevedono l'esclusione dai calcoli dei dati più esterni, ad una o ad entrambe le estremità del campo di variazione.

Test chi-quadro

Il test chi-quadro fa parte dei cosiddetti *goodness of fit* test, ossia test che verificano quanto una particolare funzione descriva la distribuzione delle frequenze relativa ad una serie di dati ottenuti sperimentalmente.

Si tratta di un test che, in generale, confronta le funzioni di distribuzione sperimentale e teorica.

La statistica da considerare per il test è:

$$T = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

k è il numero complessivo delle classi di frequenza in cui i dati grezzi vengono preliminarmente divisi

O_i sono le frequenze osservate (O = observed) per le varie classi

E_i sono le frequenze previste (E = expected) se fosse valida la funzione F(x) sotto test.

Purché:

✓ il numero complessivo di dati, N, sia sufficientemente elevato ($N > 50$)

✓ la frequenza minima in una classe sia 5

risulta:

$T \sim \chi^2_\nu$

dove $\nu = k-1-h$, con h = numero dei parametri che caratterizzano la funzione di distribuzione sottoposta al test.

Scelto un livello di significatività α, si determina un valore critico $\chi^2_{\nu\,(1-\alpha)}$

Se la realizzazione della statistica T, t, è inferiore al valore critico l'ipotesi secondo cui la funzione di distribuzione presa in esame descriva i dati sperimentali può essere accettata.

Esempio numerico 1

Riconsideriamo i 65 dati (espressi in kJ/mole) di uno dei set di misure del ΔH di neutralizzazione dell'HCl con NaOH.

L'ipotesi più semplice che si possa avanzare è che i dati siano distribuiti normalmente.

Per la media e la varianza da usare nell'espressione della distribuzione normale $N(\mu,\sigma^2)$ si possono considerare i rispettivi valori campionari:

$\mu \Rightarrow \overline{X} = 57.32$ e $\sigma^2 \Rightarrow s^2 = 2.731$

Raggruppiamo i dati sperimentali in 9 classi, ciascuna di ampiezza unitaria:

classe	intervallo	frequenze osservate O_i
1	53.05- 54.05	1
2	54.05- 55.05	4
3	55.05- 56.05	11
4	56.05- 57.05	11
5	57.05- 58.05	19
6	58.05- 59.05	8
7	59.05- 60.05	8
8	60.05- 61.05	2
9	61.05 - 62.05	1

Come si può notare, ci sono quattro classi con una frequenza inferiore a 5, per cui occorre raggruppare diversamente i dati sperimentali per poter applicare il test chi-quadro.

Occorre quindi scegliere gli intervalli delle classi in modo diverso (si noti che le ampiezze degli intervalli non devono essere necessariamente uguali):

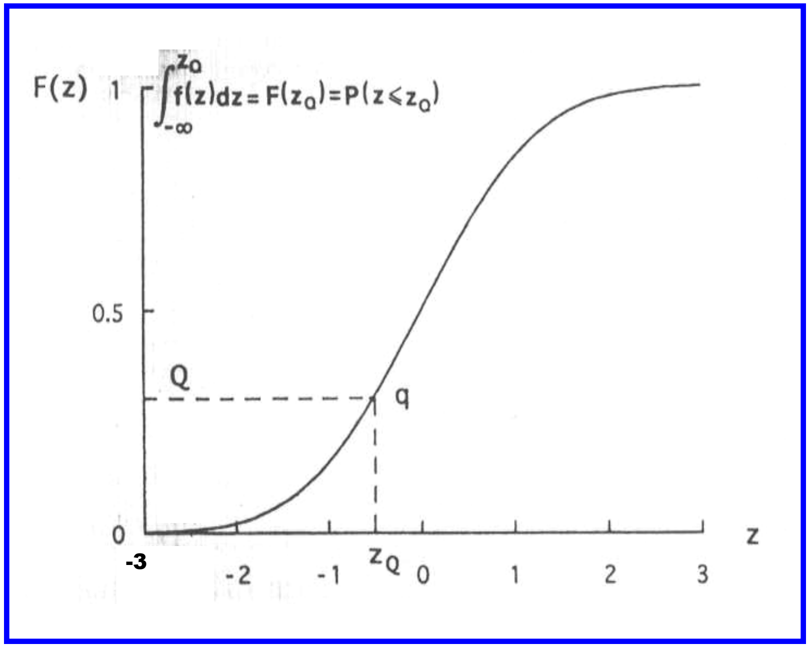
classe	intervallo	O_i	E_i	O_i^2 / E_i
1	$-\infty$, 55.05	5	5.51	4.54
2	55.05 , 56.05	11	8.87	13.64
3	56.05 , 57.05	11	13.91	8.70
4	57.05 , 58.05	19	15.31	23.58
5	58.05 , 59.05	8	11.81	5.41
6	59.05 , $+\infty$	11	9.59	12.62

I valori di E_i sono stati ottenuti standardizzando i valori limite delle diverse classi in modo da trasformarli in valori z per la distribuzione normale standard.

Detto a il valore di uno degli estremi delle classi il corrispondente valore z è dato da :

$$z = \frac{a - 57.32}{\sqrt{2.731}}$$

Ad esempio per la classe 55.05 ÷ 56.05 i valori di z corrispondenti sono -1.374 e -0.768, per cui il valore di E_i della classe si ricava dalla funzione di distribuzione normale standardizzata F(z):



$$E_{(55.05 \div 56.05)} = [F(-0.768) - F(-1.374)] \times N = 0.1365 \times 65 = 8.87$$

Per il calcolo della realizzazione della statistica T possono essere impiegati i valori O_i^2/E_i in virtù della seguente serie di eguaglianze:

$$t = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = \sum_{i=1}^{k} \frac{O_i^{\,2}}{E_i} + \sum_{i=1}^{k} \frac{E_i^{\,2}}{E_i} - 2\sum_{i=1}^{k} \frac{O_i E_i}{E_i} = \sum_{i=1}^{k} \frac{O_i^{\,2}}{E_i} - N$$

l'ultima deriva dalla considerazione che sia $\sum_{i=1}^{k} O_i$ che $\sum_{i=1}^{k} E_i$ sono uguali ad N.

Calcolando dunque la somma dei termini O_i^2/E_i si può risalire alla realizzazione della statistica T:

$$t = \sum_{i=1}^{k} \frac{O_i^{\,2}}{E_i} - N = 68.50 - 65 = 3.50$$

Il valore critico con cui confrontare tale realizzazione si ottiene dalla distribuzione chi-quadro a $\nu = k-1-h = 6-1-2 = 3$, poiché i parametri (h) che caratterizzano la distribuzione normale sono 2 (media e varianza).

Ad un livello di significatività del 5% risulta $1-\alpha = 0.95$ e quindi:

$$\chi^2_{\nu\,(1-\alpha)} = \chi^2_{3\,(0.95)} = 7.81$$

Essendo $t = 3.50 < 7.81$ l'ipotesi di base può essere accettata, il che significa che i dati di ΔH sono effettivamente distribuiti in modo normale.

Esempio numerico 2

Supponiamo di considerare il numero di pezzi di vetreria rotti da quattro operatori di laboratorio in un determinato periodo di tempo:

24 17 11 9

e di voler verificare in modo statistico se c'e' una differenza nella loro affidabilità.

L'ipotesi di base del test è che gli operatori non differiscano in affidabilità.

Esso si può effettuare considerando il confronto fra la distribuzione reale dei pezzi rotti dai quattro operatori e quella ipotetica sotto l'ipotesi di base, corrispondente al caso in cui ciascuno dei quattro operatori avesse rotto lo stesso numero di pezzi di vetreria, ossia (24+17+11+9)/4 = 15.25

In questo caso il raggruppamento dei dati è:

classe	O_i	E_i	O_i^2 / E_i
1	24	15.25	37.77
2	17	15.25	18.95
3	11	15.25	7.93
4	9	15.25	5.31

Il valore della realizzazione della statistica da usare nel test è:

$$\boxed{t = \sum_{i=1}^{k} \frac{O_i^{\,2}}{E_i} - N = 69.96 - 61 = 8.96}$$

Il numero di gradi di libertà da adottare è: $\nu = k-1-h = 4-1-0 = 3$ perché la distribuzione teorica scelta non dipende da parametri.

Ad un livello di significatività del 5% risulta $1-\alpha = 0.95$ e quindi:

$$\chi^2_{\nu\,(1-\alpha)} = \chi^2_{3\,(0.95)} = 7.81$$

Poiché $t = 8.96 > 7.81$ l'ipotesi H_0 è rigettata, il che significa che c'è una differenza significativa nell'affidabilità degli operatori, come prevedibile considerando il dato relativo al primo operatore.

Per verificare ciò su base statistica il test può essere ripetuto limitando la considerazione agli altri tre operatori.

In questo caso N = 37, k = 3 ed E_i = 37/3 = 12.33:

classe	O_i	E_i	O_i^2 / E_i
2	17	12.33	23.44
3	11	12.33	9.81
4	9	12.33	6.57

Il valore critico da usare in questo nuovo test è:

$$t = \sum_{i=1}^{k} \frac{O_i^2}{E_i} - N = 39.82 - 37 = 2.82$$

Il numero di gradi di libertà da adottare è: $\nu = k-1-h = 3-1-0 = 2$

Ad un livello di significatività del 5 % risulta:

$$\chi^2_{\nu\,(1-\alpha)} = \chi^2_{2\,(0.95)} = 5.99$$

Poiché t = 2.82 < 5.99 si conferma che, su base statistica, non c'è differenza significativa fra i restanti tre operatori, il che vuol dire che la variabilità comunque esistente fra loro è compatibile con quella accettabile.

Test di Kolmogorov - Smirnov

Il test di Kolmogorov-Smirnov rappresenta un altro approccio alla risoluzione del problema statistico della *goodness of fit*; rispetto al test chi-quadro ha il vantaggio di poter essere applicato anche a set di dati più piccoli ($N < 50$) e di essere efficace su variabili continue.

Il test si basa sul confronto fra la spezzata delle frequenze cumulative relativa ai dati sperimentali e la curva prevista dalla distribuzione teorica. Nel caso specifico non è richiesta la suddivisione in classi dei dati *grezzi*: ogni dato ottenuto rappresenta una classe.

Il test può essere effettuato in due modi:

test ad una coda) si considera la massima differenza positiva fra la curva delle frequenze cumulative sperimentali e quella teorica o viceversa;

test a due code) si considera il massimo dei valori assoluti delle differenze fra la curva delle frequenze cumulative sperimentali e quella teorica;

I valori ora indicati vengono confrontati con un valore critico, diverso a seconda della modalità del test e della funzione di distribuzione usata.

Test di Kolmogorov-Smirnov per la casualità: un esempio numerico

Supponiamo di aver annotato su un certo intervallo di tempo le ultime cifre registrate da un tecnico di laboratorio nel fare letture con una buretta, ad esempio nel valutare il volume di titolante aggiunto in titolazioni relative a soluzioni diverse fra loro.

La distribuzione delle cifre su 50 letture è la seguente:

Cifra:	0	1	2	3	4	5	6	7	8	9
Frequenza:	1	6	4	5	3	11	2	8	3	7

Ci chiediamo se le letture siano effettivamente distribuite in modo casuale.

Il test ha lo scopo di valutare se l'operatore prediliga indicare un particolare numero, laddove sia in dubbio, come ultima cifra.

Naturalmente la valutazione è possibile soltanto se le 50 titolazioni da cui derivano le letture sono totalmente indipendenti una dall'altra.

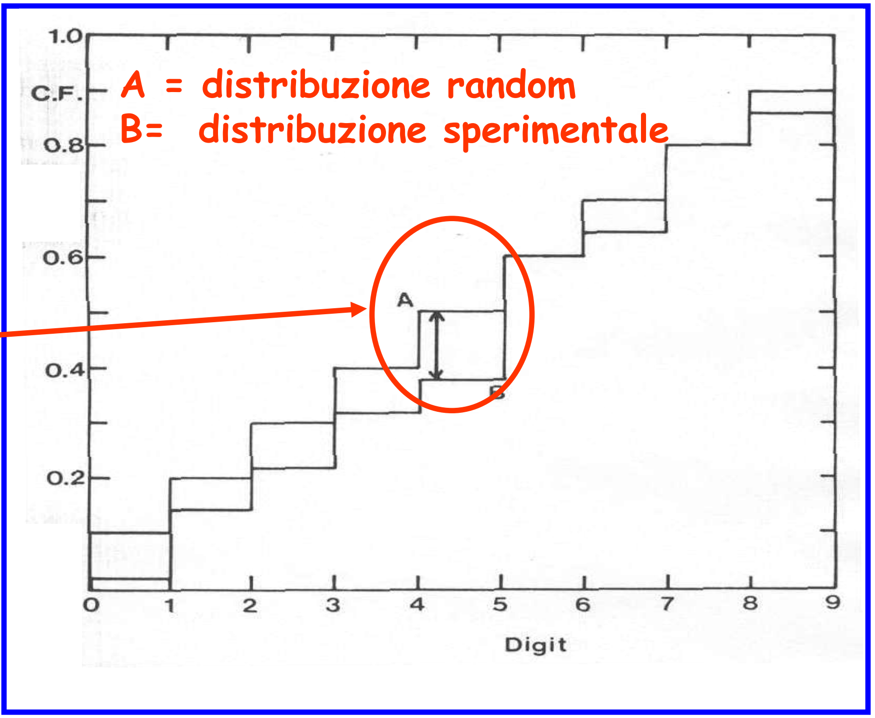
Usando le frequenze cumulative relative, i valori sperimentali per il test sono:

Cifra:	0	1	2	3	4	5
Frequenza:	0.02	0.14	0.22	0.32	0.38	0.6
Cifra:	6	7	8	9		
Frequenza:	0.64	0.80	0.86	1.00		

D'altra parte i valori teorici, in caso di distribuzione effettivamente casuale, sarebbero: 0.1, 0.2, …, 0.9, 1.

Dal confronto delle frequenze cumulative relative si nota che:

- ✓ la massima differenza si verifica per la cifra 4 ed è pari a 0.12
- ✓ non ci sono differenze negative (al minimo sono pari a 0, per le cifre 5 e 7).



Applicando il test a due code il valore da testare è 0.12 e va confrontato con quello critico per la casualità con N = 50, ossia 0.188.

Poiché 0.12 < 0.188 si può dire che la distribuzione delle frequenze delle cifre finali nelle letture sia effettivamente random.

Le condizioni del problema sono tali da consentire anche l'applicazione del test chi-quadro, che porta allo stesso risultato.

Test di Kolmogorov-Smirnov per la normalità

Prima dell'applicazione del test di Kolmogorov-Smirnov per la verifica della normalità di una distribuzione è necessario trasformare i dati iniziali in valori della variabile normale standard z:

$$Z = (x-\mu)/ \sigma$$

dove μ e σ sono la media e la varianza:

- di popolazione, se si vuole fare il confronto con una particolare distribuzione gaussiana, di cui si conoscono media e varianza;
- campionarie, ossia ricavate dai dati a disposizione, se interessa solo stabilire se i dati siano distribuiti secondo una gaussiana.

Esempio numerico

Sono state effettuate otto titolazioni della stessa soluzione, ottenendo i seguenti risultati:

25.13	25.02	25.11	25.07
25.03	24.97	25.14	25.09

Ci si chiede se:

1) i dati siano distribuiti secondo una distribuzione normale con media 25.00 mL e deviazione standard 0.05 mL;

oppure se:

2) i dati siano distribuiti secondo una distribuzione normale.

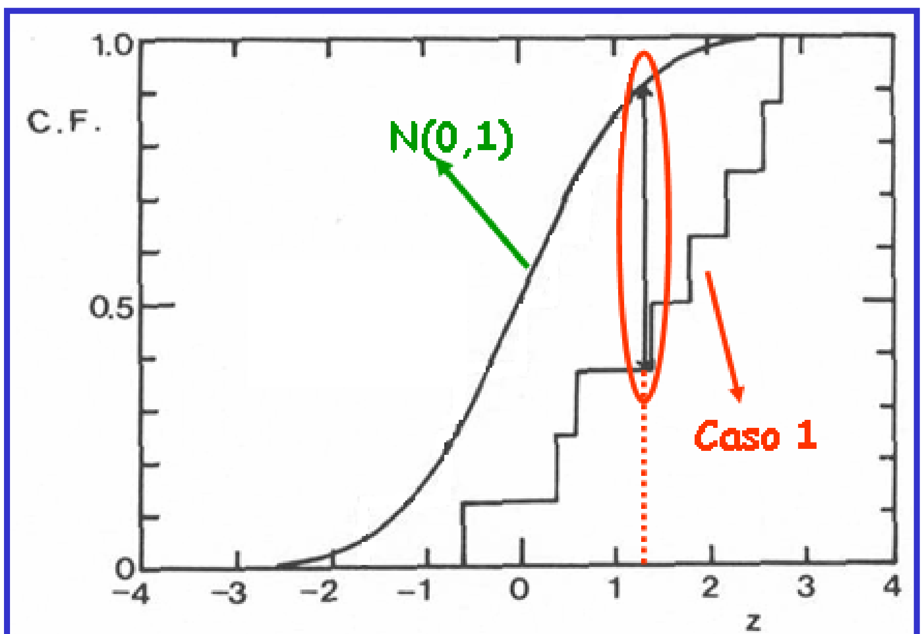
Caso 1

Essendo μ = 25 e σ = 0.05, i valori di z corrispondenti agli otto dati sono:

2.6 0.4 2.2 1.4 0.6 -0.6 2.8 1.8

Essi possono essere inseriti nel grafico delle frequenze cumulative relative considerando che la frequenza relativa di ciascuno è 1/8 = 0.125.

Nello stesso grafico si può inserire la curva continua delle frequenze cumulative relative per la distribuzione N(0,1):



La massima differenza si registra in prossimità del dato 1.4 ed è pari a 0.545. Non ci sono differenze negative.
Il valore critico al 5 % di significatività, e con n = 8, si ricava dalle tavole apposite del test di Kolmogorov (a due code) per la normalità ed è pari a 0.288.

Poiché 0.545 > 0.288 l'ipotesi 1 va rigettata.

Caso 2

In questo caso si devono stimare i parametri campionari relativi ai dati a disposizione:

$\overline{X} = 25.07$ e $s = 0.059$

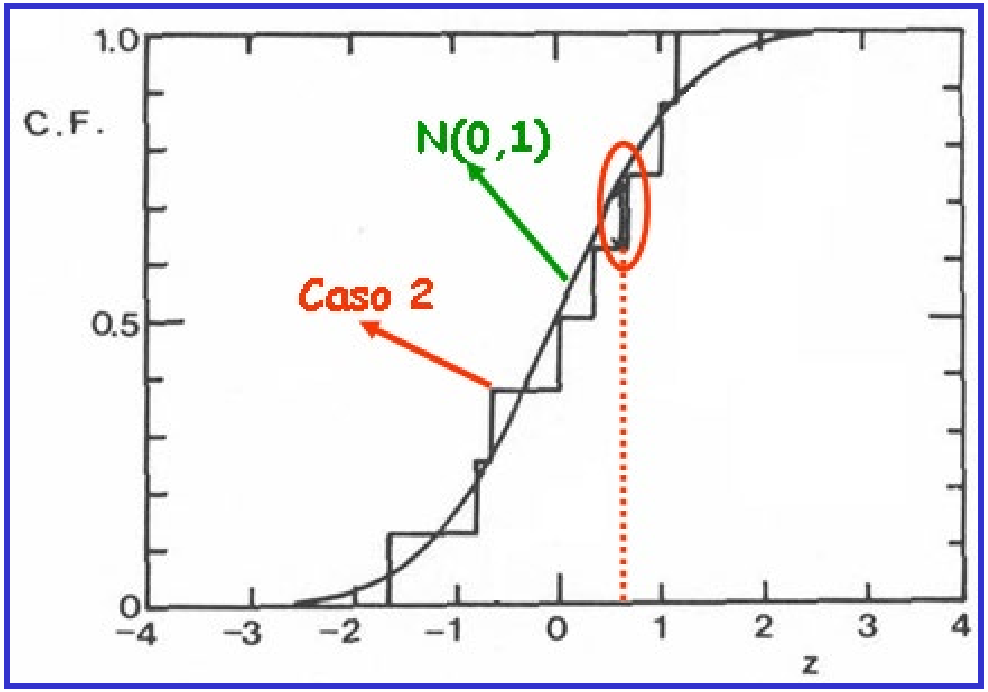
I valori di z si ricaveranno quindi dalla relazione:

$z = (x-25.07)/0.059$

e sono:

1.02	-0.85	0.68	0
-0.68	-1.69	1.19	0.34

Inserendo questi dati nel grafico contenente la curva derivante da N(0,1) si nota che la massima differenza fra essa e la distribuzione sperimentale è 0.125. Essa è maggiore del valore assoluto della massima differenza negativa.



Essendo 0.125 < 0.288 l'ipotesi che i dati delle titolazioni siano distribuiti normalmente può essere accettata al 5 % di significatività.

In questa situazione la media e la deviazione standard campionarie sono le stime migliori per i rispettivi parametri di popolazione.