Esempio VII.2 Due astronomi hanno raccolto osservazioni su una data stella. Le 12 osservazioni raccolte dal primo astronomo hanno una misura media di 1,20. Le 8 osservazioni ottenute dal secondo astronomo hanno una media di 1,15.

L'esperienza passata ha indicato che questi astronomi ottengono misure con una varianza di circa 0,40.

Ci si chiede se la differenza tra i due risultati è significativa con $\alpha = 0{,}01$.

H0: $\mu_x = \mu_y$; (σ^2=0,40; σ=0,63245)

Ha: $\mu_x \neq \mu_y$;

$$z_{\bar{X}-\bar{Y}} = \frac{\left(\bar{X}-\bar{Y}\right)-\left(\mu_x-\mu_y\right)}{\sqrt{\sigma_x^2/n_x+\sigma_y^2/n_y}} = \frac{\left(\bar{X}-\bar{Y}\right)-0}{\sigma\sqrt{1/n_x+1/n_y}} = \frac{1{,}20-1{,}15}{0{,}6325\sqrt{\frac{1}{12}+\frac{1}{8}}} = 0{,}17$$

Regola di decisione :
Si rigetta l'ipotesi H0 se $z < -2{,}58$ oppure se $z > 2{,}58$.

$z_{\bar{X}-\bar{Y}}$ è compreso tra -2,58 e +2,58

Accettiamo l'ipotesi H0 !

Dunque, la differenza tra le due medie non è significativa ad un livello $\alpha = 0{,}01$

Esempio VII.3 Due differenti tipi di razioni alimentari vengono somministrate ad un gruppo di maiali. Si vuole verificare quale dei due tipi sia migliore. Un campione di 12 maiali viene alimentato con la razione di tipo A ed un altro campione di 12 maiali riceve, invece, la razione di tipo B. I guadagni in peso registrati sono i seguenti:

Tipo A : **31, 34, 29, 26, 32, 35, 38, 34, 30, 29, 32, 31**

Tipo B: **26, 24, 28, 29, 30, 29, 32, 26, 31, 29, 32, 28**

α = 0,05.

H0: $\mu_x = \mu_y$;

Ha: $\mu_x \neq \mu_y$;

$$t_{\bar{X}-\bar{Y}} = \frac{\left(\bar{X}-\bar{Y}\right)-\left(\mu_x-\mu_y\right)}{\sqrt{\left(n_x-1\right)s_x^2+\left(n_y-1\right)s_y^2}}\sqrt{\frac{n_x n_y\left(n_x+n_y-2\right)}{n_x+n_y}}$$

Regola di decisione :

Si rigetta l'ipotesi H0 se t < - 2,07 oppure se t > 2,07.

Se le due popolazioni hanno distribuzioni normali con la stessa media e la stessa varianza, allora questa statistica ha una distribuzione t con 12+12-2 = 22 gradi di libertà:

$$t_{\bar{X}-\bar{Y}} = \frac{\left(\bar{X}-\bar{Y}\right)-\left(\mu_x-\mu_y\right)}{\sqrt{\left(n_x-1\right)s_x^2+\left(n_y-1\right)s_y^2}}\sqrt{\frac{n_x n_y\left(n_x+n_y-2\right)}{n_x+n_y}} =$$

$$\frac{\left(\bar{X}-\bar{Y}\right)-\left(\mu_x-\mu_y\right)}{\sqrt{\frac{\left(n_x-1\right)s_x^2+\left(n_y-1\right)s_y^2}{\left(n_x+n_y-2\right)}}}\sqrt{\frac{n_x n_y}{n_x+n_y}} = \frac{\left(\bar{X}-\bar{Y}\right)-\left(\mu_x-\mu_y\right)}{\sqrt{\frac{\left(n_x-1\right)s_x^2+\left(n_y-1\right)s_y^2}{\left(n_x+n_y-2\right)}\left(\frac{1}{n_x}+\frac{1}{n_y}\right)}}$$

Per il tipo A la media è $\mu_x = 31{,}7500$ mentre per la somma dei quadrati si ha:

$$\boxed{(N_x - 1).s_x^2 = \sum X_i^2 - \frac{\left(\sum X_i\right)^2}{N_x} = 112{,}25}$$

Per il tipo B la media è $\mu_y = 28{,}6667$ mentre per la somma dei quadrati si ha:

$$\boxed{(N_y - 1).s_y^2 = \sum Y_i^2 - \frac{\left(\sum Y_i\right)^2}{N_y} = 66{,}66}$$

Per cui:

$$\boxed{\mathbf{t = \frac{31{,}7500 - 28{,}6667}{2{,}85.\sqrt{\frac{1}{12}+\frac{1}{12}}} = \frac{3{,}0800.\sqrt{6}}{2{,}85} = 2{,}65}}$$

2,65> 2,07 si rigetta l'ipotesi H0!

<u>Esempio VII.4</u> Due tipi di vernici devono essere confrontati. Di ciascun tipo di vernice vengono esaminati 5 campioni ai quali risultano attribuiti i punteggi seguenti:

Tipo I	**85,**	**87,**	**92,**	**80,**	**84**
Tipo II	**89,**	**89,**	**90,**	**84,**	**88**

H0: $\mu_x = \mu_y$

Ha: $\mu_x < \mu_y$;

α.= 0,05

$$t_{\bar{X}-\bar{Y}} = \frac{(\bar{X}-\bar{Y}) - (\mu_x - \mu_y)}{\sqrt{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}} \sqrt{\frac{n_x n_y (n_x + n_y - 2)}{n_x + n_y}}$$

Regola di decisione :
Si rigetta l'ipotesi H0 se t < $t_{0,05}$ e cioè se t < -1,86.

Si ha $\mu_x = 85{,}6$ ed $\mu_y = 88{,}0$; per la somma dei quadrati del tipo I si ha invece:

$$\boxed{(N_x - 1).s_x^2 = \sum X_i^2 - \frac{\left(\sum X_i\right)^2}{N_x} = 77{,}2}$$

e per il tipo II :

$$\boxed{(N_y - 1).s_y^2 = \sum Y_i^2 - \frac{\left(\sum Y_i\right)^2}{N_y} = 22}$$

Sulla base dei risultati suddetti si può calcolare:

$$\boxed{t = \frac{85{,}6 - 88{,}0}{3{,}52.\sqrt{\frac{1}{5}+\frac{1}{5}}} = \frac{-2{,}4.\sqrt{5/2}}{3{,}52} = -1{,}08}$$

DA VERIFICARE NUMERI

-1.08> - 1,86 si accetta l'ipotesi H0!

Accoppiamento delle osservazioni

Un piano degli esperimenti che talvolta permette di superare alcuni problemi del campionamento è quello che si basa sull'espediente di accoppiare le osservazioni, facendo in modo che le due componenti di ciascuna coppia siano simili sotto tutti gli aspetti tranne quello che si sta esaminando.

Questa situazione sarebbe l'ideale, naturalmente, ma nella generalità dei casi si è limitati dalla disponibilità di coppie che siano simili e dalla capacità che chi effettua l'esperimento ha di scegliere coppie simili.

Esempio I Se si vuol conoscere tra due diversi metodi di insegnamento quale sia migliore, si possono prendere due gruppi di studenti ed insegnare ad un gruppo con il primo metodo ed all'altro con il secondo metodo.

Se uno dei gruppi è formato da studenti migliori (o più maturi o meglio addestrati nelle materie di base, eccetera) di quelli dell'altro gruppo, i risultati dell'esperimento possono non riflettere l'efficacia dei metodi di insegnamento.

Con il criterio dell'accoppiamento, si dovrebbero prendere coppie di studenti di abilità approssimativamente uguale e, mentre uno dei componenti di ciascuna coppia dovrebbe essere assoggettato al primo metodo, l'altro dovrebbe essere sottoposto al secondo metodo.

Esempio II. Supponiamo che in un esperimento per verificare quale tra due tipi (A o B) di fertilizzanti sia migliore, in 10 stazioni sperimentali siano stati seminati a grano due appezzamenti di terreno per ciascuno.

Se le medie dei 10 appezzamenti fertilizzati con il tipo A fossero confrontate con le medie dei 10 appezzamenti che hanno ricevuto il tipo B, parte della differenza osservata (se ve ne fosse una) potrebbe essere dovuta al diverso tipo di suolo o a diverse condizioni meteorologiche invece che ai due diversi fertilizzanti.

Con l'accoppiamento, ogni coppia di appezzamenti dovrebbe approssimativamente avere lo stesso tipo di suolo, essere esposto a condizioni meteorologiche identiche, eccetera.

Esempio VII.6 In uno studio sulla capacità di apprendimento, da una classe di matricole sono stati scelti a caso 10 ragazzi e 10 ragazze. Ad essi fu attribuito un punteggio, misurando la loro abilità ad apprendere sillabe senza senso. Poiché il conduttore dell'esperimento sospettava che la varianza sarebbe stata diversa

per ragazzi e ragazze (σ_x^2 diverso da σ_y^2), prese la decisione di accoppiare i dati.

L'analisi che viene mostrata è appropriata per un accoppiamento casuale dei soggetti.

Numero d'ordine della coppia

	1	2	3	4	5	6	7	8	9	10
Ragazzi	**28**	**18**	**22**	**27**	**25**	**30**	**21**	**21**	**20**	**27**
Ragazze	**19**	**38**	**42**	**25**	**15**	**31**	**22**	**37**	**30**	**24**
Differenza	**9**	**-20**	**-20**	**2**	**10**	**-1**	**-1**	**-16**	**-10**	**3**

$$\sum d_i = -44 \qquad \sum d_i^2 = 1352$$

$$S^2(d) = \left(1352 - \frac{44^2}{10}\right)/9 = 128{,}7$$

Se x e y hanno distribuzioni normali, allora questa statistica ha una distribuzione t con 10-1 = 9 gradi di libertà.

$$\boxed{t = \frac{(\mu_d - 0)}{s_d/\sqrt{N}} = \frac{-4.4}{\sqrt{128{,}7}/\sqrt{10}} = -1{,}2.}$$

H0: $\mu_x = \mu_y$ quindi : $\mu_d = 0$

Ha: $\mu_x \neq \mu_y$; quindi : $\mu_d \neq 0$

α = 0,01

Regola di decisione :
Si rigetta l'ipotesi H0 se t < - 3,25 o t>3,25

-1.2> -3.25 si accetta l'ipotesi H0!

Esempio VII.8 Un certo stimolo deve essere verificato per i suoi effetti sulla pressione del sangue. Su 12 uomini viene misurata la pressione sanguigna prima e dopo lo stimolo. I risultati sono i seguenti.

Soggetti esaminati	Prima	Dopo	Differenza d_i	d^2_i
1	120	128	+8	64
2	124	131	+7	49
3	130	131	+1	1
4	118	127	+9	81
5	140	132	-8	64
6	128	125	-3	9
7	140	141	+1	1
8	135	137	+2	4
9	125	118	-8	64
10	130	132	+2	4
11	126	129	+3	9
12	127	135	+8	64
			22	414

Si può ritenere che lo stimolo elevi la pressione sanguigna in media di 5 punti?

H0: $\mu_x + 5 = \mu_y$ quindi : $\mu_d = 5$

Ha: $\mu_x + 5 \geq \mu_y$; quindi : $\mu_d \leq 5$

α = 0,05.

Si ha :

$$S^2(d) = \left(414 \text{ - } \frac{22^2}{12} \right) / 11 = 33.97 \qquad S(d) = 5,83$$

Si utilizza la statistica:

$$\boxed{t = \frac{(\mu_d - 5)}{s_d / \sqrt{N}} = \frac{22/12 - 5}{5.83/\sqrt{12}} = -1.88}$$

d) Se le due popolazioni hanno distribuzioni normali allora questa statistica ha una distribuzione t con (12 - 1) = 11 gradi di libertà.

Regola di decisione :

Si rigetta l'ipotesi H0 se t < - 1,8

-1,8 > -1,88 si rifiuta l'ipotesi H0!

Significatività della differenza tra percentuali nei g. c.

In alcuni problemi spesso si devono usare misure percentuali, invece di valori assoluti.

Se indichiamo con p'1 e p'2 le percentuali relative a due campioni basate su n1 ed n2 prove e se p1 e p2 rappresentano le percentuali delle corrispondenti popolazioni, allora la media e la deviazione standard delle differenze tra percentuali sono date dalle seguenti relazioni:

$$M_{p'_1 - p'_2} = p_1 - p_2$$

e

$$\sigma_{p'_1 - p'_2} = \sqrt{\frac{p_1 \cdot q_1}{n_1} + \frac{p_2 \cdot q_2}{n_2}}$$

in cui q1 = (1 - p'1) e q2 = (1 - p'2).

Esempio VII.8 **Due diversi tipi di siero sono somministrati a 134 mucche colpite da una malattia.**

- **Il siero A è iniettato a 50 mucche e 5 di queste muoiono.**
- **Il siero B è iniettato alle restanti 84 mucche e 6 di esse muoiono.**

Si vuol sapere se la differenza di effetto tra i due tipi di siero sulla mortalità delle mucche per la malattia in questione è significativa.

Il tasso di mortalità corrispondente al siero del tipo A è 5/50 = 1/10 ossia il 10% ovvero p'1 = 0,10.

Invece, il tasso di mortalità corrispondente al siero del tipo B è 6/84=1/14=0,0714 ossia il 7,14% ovvero p'2=0,0714.

Il numero totale di mucche malate è 134, mentre complessivamente le mucche morte sono 11.

Quindi, avremo p=11/134=0,082 cioè 8,2% e q=1-0,082=0,918.

In questo problema p1 e p2 non sono noti.

Ipotesi H0

Se si suppone che p1 = p2 = p, sarà anche q1 = q2 = q e si avrà sotto H0:

$$M_{p_1^{'}-p_2^{'}} = p_1 - p_2 = 0$$

$$\sigma_{p_1^{'}-p_2^{'}} = \sqrt{\frac{p_1 \cdot q_1}{n_1} + \frac{p_2 \cdot q_2}{n_2}} = \sqrt{0,082x0,918x\left(\frac{1}{50} + \frac{1}{84}\right)}$$

cioè

$$\sigma_{p_1^{'}-p_2^{'}} = \sqrt{\frac{0,082x0,918x134}{4200}} = \sqrt{0,002401663} = 0,049$$

Con questi risultati, si ottiene, quindi:

$$t = \frac{0,10 - 0,0714}{0,049} = \frac{0,0286}{0,049} = 0,58367$$

per cui:

$$\boxed{\frac{1}{\sqrt{2\pi}} \int_0^{0,58367} e^{-\frac{t^2}{2}} . dt = 0,2202.}$$

Quindi la probabilità che si abbia una differenza maggiore di (0,10-0,07)=0,03 è pari a (1-2.0,2202)=(1-0,4404)=0,5596.

Quindi la probabilità che si abbia una differenza maggiore di (0,10-0,07)=0,03 è pari a (1-2.0,2202)=(1-0,4404)=0,5596.
Si conclude, dunque, che circa 56 coppie di campioni casuali su 100 dovrebbero dare una differenza più grande di 0,03 cosicché la differenza tra i due sieri non è affatto significativa.

α **= 0,1**
Regola di decisione :
Si rigetta l'ipotesi H0 se $t < -1{,}64$ o $t>1{,}64$

$-1{,}64< t=0.58 <1{,}64$ **si accetta l'ipotesi H0!**

Distribuzione della deviazione standard nei g.c.

$$\boxed{M_s = \sigma} \quad \text{e} \quad \boxed{\sigma_s = \frac{\sigma}{\sqrt{2n}}}$$

Inoltre, per grandi campioni la distribuzione della deviazione standard diviene approssimativamente normale.

<u>Esempio VII.9</u>
Si supponga di avere due grandi campioni per i quali si ha:

$n_1=329$; $n_2=302$;
$m_1=32{,}75$; $m_2=30{,}6$;
$s_1=8{,}05$; $s_2=6{,}95$.

Si supponga inoltre che sia $\sigma_1 = \sigma_2$.

Con queste posizioni applicando le formule indicate, si avrà:

$$z = \frac{(s_1 - s_2) - M_{s_1 - s_2}}{\sigma_{s_1 - s_2}} =$$

$$= \frac{(s_1 - s_2) - 0}{\sqrt{\sigma^2_{s_1} + \sigma^2_{s_2}}} = \frac{(s_1 - s_2)}{\sqrt{\frac{s_1^2}{2n_1} + \frac{s_2^2}{2n_2}}}$$

cioè

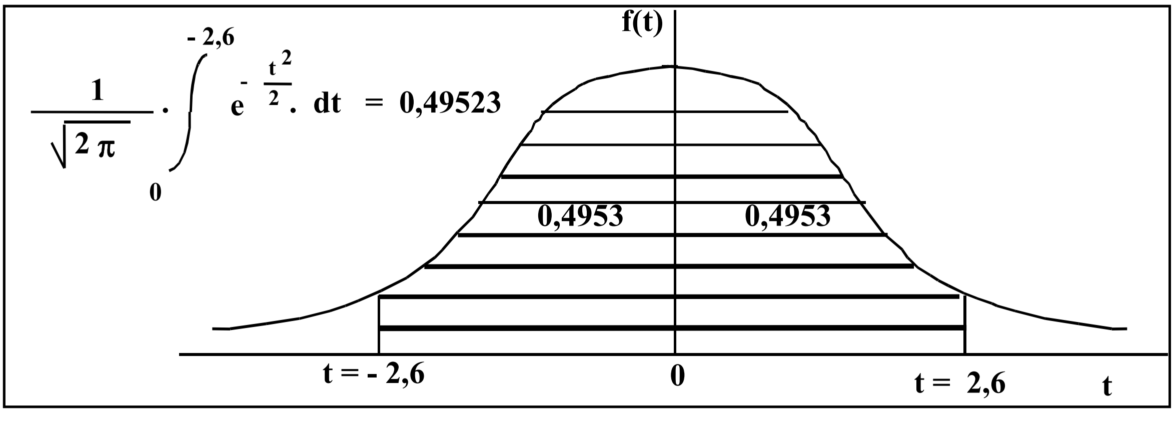
$$t = \frac{8,05 - 6,95}{\sqrt{\frac{8,05^2}{2x329} + \frac{6,95^2}{2x302}}} = \frac{1,10}{0,422439} = 2,6039.$$

Ora, dalla tavola della curva normale standardizzata si ha che per t = 2,6 l'integrale tra zero e t vale:

$$\boxed{\frac{1}{\sqrt{2\pi}} \int_0^{2,6} e^{-\frac{t^2}{2}} . dt = 0,4953}$$

e quindi:

$$\boxed{1 - \frac{1}{\sqrt{2\pi}} \int_{-2,6}^{2,6} e^{-\frac{t^2}{2}} . dt = 1 - 2x0,4953 = 0,0092.}$$



Questa probabilità è più piccola del 5% ed infatti è 0,0094 < 0,05, per cui l'ipotesi che sia $\sigma_1 = \sigma_2$ non è giustificata.

In altre parole, esiste una differenza significativa tra le deviazioni standard delle due popolazioni.

<u>Esempio</u>

Si sa che un certo tipo di fiore ha 30 petali. Un campione di 10 fiori viene sottoposto ad un trattamento molto costoso ed i petali in questo campione risultano essere 33, con una deviazione standard di 2. *L'aumento del numero dei petali è significativo?*

In questo problema n = 10, M = 30, m=33 ed s = 2. Applicando la formula adatta nel caso dei piccoli campioni, si ha:

$$\boxed{t = \frac{m - M}{\frac{s}{\sqrt{n-1}}} = \frac{(33-30)x\sqrt{9}}{2} = 4,5}$$

In questa formula è ν = n-1 = 10-1 = 9. Quindi cercando sulla tavola il valore corrispondente a t = 4,5 e ν = 9 si trova che la probabilità di una deviazione maggiore di t=4,5 è minore di 0,005. Questo valore è più piccolo del livello di significatività del 5% ed anche dell' 1%.

Risulta, quindi, che l'aumento nel numero di petali è molto significativo e si può attribuire all'effetto del trattamento costoso.

Stima nei piccoli campioni

16. Limiti di confidenza della media nei p. c.

Per un campione di n < 30 (o 50) elementi si abbiano la media m, la sua deviazione standard s ed un numero di gradi di libertà pari a ν = (n-1).

L'intervallo di confidenza della media ad un livello di accettabilità p e cioè con probabilità di errore inferiore ad (1-p), tenendo presente la relazione:

$$\boxed{-t_{1-p} < \frac{m - M}{\frac{s}{\sqrt{n-1}}} < t_{1-p}}$$

sarà determinato dalla doppia disuguaglianza:

$$\boxed{m - t_{1-p}\frac{s}{\sqrt{n-1}} < M < m + t_{1-p}\frac{s}{\sqrt{n-1}}}$$

Stima nei piccoli campioni

<u>Esempio</u>
In una certa provincia sia stata eseguita un'indagine su un campione di 26 campi coltivati a una certa coltura. La resa media per ettaro ottenuta risulta di 300 Kg con una deviazione standard di 20 Kg. ***Trovare l'intervallo di confidenza al 95% per il valore medio della resa in tutta la provincia.***

In questo esempio M non è conosciuto; inoltre, si hanno i seguenti dati iniziali: n = 26, m = 300, s = 20 ed infine ν = 25. Per cui sarà σ_m = s/√(n-1) = 20/√25 = 20/5 = 4.

Si sa inoltre che, avendo un piccolo campione, la t segue una distribuzione avente un numero di gradi di libertà pari a ν = n-1 = 25.

Con una probabilità del 95% si avrà:

$$\boxed{-t_{0,05} < \frac{m-M}{\frac{s}{\sqrt{n-1}}} < t_{0,05}}$$

In questa espressione $t_{0,05}$ indica il valore di t, con (n-1) gradi di libertà, tale che la probabilità di avere un valore $t > t_{0,05}$ è di appena il 5%.

Allora, l'intervallo di confidenza è determinato dalla seguente relazione:

$$\boxed{m - t_{0,05}\frac{s}{\sqrt{n-1}} < M < m + t_{0,05}\frac{s}{\sqrt{n-1}}}$$

Sostituendo i valori relativi all'esempio considerato nelle disuguaglianze suddette si ottengono i limiti ricercati. Infatti, dalla

tavola della t di Student per $\nu = 25$ gradi di libertà si trova $t_{0,05} = 1,708$ e quindi i limiti cercati saranno:

$$\boxed{300 - 1,708x4 < M < 300 + 1,708x4}$$

$$\boxed{293,2 < M < 306,8.}$$

Stima nei piccoli campioni

17. Significatività della differenza tra due medie nei p. c.

Si dimostra che per la distribuzione della differenza tra le medie di due piccoli campioni, la distribuzione della quantità:

$$\boxed{t = \frac{(m_1 - m_2) - (M_1 - M_2)}{\sqrt{n_1 . s_1^2 + n_2 . s_2^2}} \sqrt{\frac{n_1 . n_2 . (n_1 + n_2 - 2)}{n_1 + n_2}}}$$

è una t di Student con $\nu=(n_1+n_2-2)$ gradi di libertà.

Per usare questa espressione *è necessario supporre che le varianze delle due popolazioni siano uguali*.

Nell'applicare questa formula per verificare la significatività della differenza tra le medie dei due campioni si ammette l'ipotesi che sia $M_1 = M_2$.

Cioè, si pone $(M_1-M_2)=0$ nel calcolo del valore di t della formula precedente e quindi si cerca, come al solito, la probabilità corrispondente nella tavola della t, con $\nu=(n_1+n_2-2)$ gradi di libertà.

Stima nei piccoli campioni

17. Significatività della differenza tra due medie nei p. c.

Evidentemente sfruttando la stessa formula data per la t è possibile anche trovare i limiti di confidenza per la differenza tra le medie delle popolazioni da cui i due campioni sono estratti.

A questo scopo, è sufficiente ricavare dalla tavola della t relativa all'intervallo di confidenza al livello del 5% il valore di t che corrisponde ai gradi di libertà del campione considerato e quindi avremo:

$$\boxed{-t(\nu) < \frac{\dfrac{(m_1 - m_2) - (M_1 - M_2)}{\sqrt{n_1 . s_1^2 + n_2 . s_2^2}}}{\sqrt{\dfrac{n_1 . n_2 . (n_1 + n_2 - 2)}{n_1 + n_2}}} < t(\nu)}$$

Stima nei piccoli campioni

17. Significatività della differenza tra due medie nei p. c.

Ne segue che dovrà essere soddisfatta la doppia disuguaglianza:

$$\left(m_1 - m_2\right) - t(\nu).\frac{\sqrt{n_1.s_1^2 + n_2.s_2^2}}{\sqrt{\frac{n_1.n_2.\left(n_1 + n_2 - 2\right)}{n_1 + n_2}}} < \left(M_1 - M_2\right)$$

$$\left(M_1 - M_2\right) < \left(m_1 - m_2\right) + t(\nu).\frac{\sqrt{n_1.s_1^2 + n_2.s_2^2}}{\sqrt{\frac{n_1.n_2.\left(n_1 + n_2 - 2\right)}{n_1 + n_2}}}$$

che definisce, appunto, l'intervallo nel quale dovrebbe risultare compresa la differenza tra le medie delle due popolazioni.

Stima nei piccoli campioni

Esempio

Nel corso di un esperimento furono selezionati a caso due insiemi di piante di grano, il primo contenente 17 piante ed il secondo 13. Le altezze medie della spiga per i due campioni sono risultate uguali a 7,5 e 6,0 centimetri e le loro varianze uguali a 0,75 e 0,9 cm rispettivamente.

Trovare al livello del 5% di confidenza se la differenza tra i due campioni di altezze delle spighe di grano è significativa.

Si ponga (M_1 - M_2) = 0 e si ricordi che si hanno i seguenti dati sperimentali: n_1 = 17; n_2 = 13; m_1 = 7,5; m_2 =6,0; s_1 = 0,75; s_2 = 0,9. Applicando la formula per il calcolo della t nei piccoli campioni si avrà:

$$t = \frac{7,5 - 6,0}{\sqrt{17x0,75 + 13x0,9}} \cdot \sqrt{\frac{17x13x(17 + 13 - 2)}{17 + 13}} =$$

$$= \frac{1,5}{\sqrt{12,75+11,70}} \cdot \sqrt{206,267} = \frac{1,5x14,36}{4,94469} = 4,356.$$

Cercando nella tavola della t il valore trovato t = 4,356 con v = 28 gradi di libertà, si ha che la probabilità relativa è troppo piccola, poiché una probabilità uguale a 0,01 corrisponde solo ad un valore di t = 2,76. Di conseguenza *questa differenza tra le medie è fortemente significativa e l'ipotesi che sia $M_1=M_2$ non è giustificata*.

E' anche possibile trovare i limiti di confidenza per la differenza tra le medie delle popolazioni da cui i due campioni sono estratti. Usando la stessa formula della t, si ha:

$$\boxed{t = \frac{1,5-(M_1-M_2)}{4,9} . 14,3 = \frac{1,5-(M_1-M_2)}{0,3426}}$$

Dalla tavola della t l'intervallo di confidenza al livello del 5% con 28 gradi di libertà corrisponde ad un valore di t = 2,048 e quindi si avrà:

$$\boxed{-2,048 < \frac{1,5-(M_1-M_2)}{0,3426} < 2,048}$$

ossia

$$\boxed{0,804 < (M_1-M_2) < 2,196}.$$

Stima nei piccoli campioni

18. L'errore standard di s nei p. c.

Nel caso dei piccoli campioni il valore di σ nella formula:

$$\boxed{\sigma_s = \frac{\sigma}{\sqrt{2n}}}$$

può essere sostituito dal valore della sua stima definito dalla relazione:

$$\boxed{\hat{\sigma} = s\sqrt{\frac{n}{n-1}}}$$

quindi si ha:

$$\boxed{\sigma_s = \frac{\hat{\sigma}}{\sqrt{2n}} = \frac{s\sqrt{n}}{\sqrt{2n}\sqrt{n-1}} = \frac{s}{\sqrt{2(n-1)}}}$$

Il valore di t definito come segue:

$$\boxed{t = \frac{s-\sigma}{\frac{s}{\sqrt{2(n-1)}}}}$$

ha una distribuzione t di Student con ν=(n-1) gradi di libertà.

Stima nei piccoli campioni

Esempio

Si consideri un campione di 9 elementi per il quale si è trovato che la deviazione standard è uguale a s = 2,4. Si avrà:

$$\boxed{\sigma_s = \frac{s}{\sqrt{2(n-1)}} = \frac{2,4}{\sqrt{2(9-1)}} = \frac{2,4}{4} = 0,6.}$$

Nella tavola della t di Student, a ν = 8 gradi di libertà e con limiti di confidenza al 5% corrisponde un valore di t = 2,306. Quindi è possibile dire che la deviazione standard della popolazione σ cadrà con il 95% di probabilità tra i limiti seguenti:

$$\boxed{2,4 - 2,306 x 0,6 < \sigma < 2,4 + 2,306 x 0,6}$$

$$\boxed{1,014 < \sigma < 3,786.}$$