LA VERIFICA DELLE IPOTESI --- parte seconda:

LA VERIFICA DELLE IPOTESI NEL CASO DI UN CAMPIONE

Il test t di Student

Quando σ non è nota

Sino ad ora abbiamo considerato degli esempi in cui si assumeva che la deviazione standard della popolazione fosse nota

Spesso, tuttavia, non siamo in possesso di tale informazione

Se non conosciamo σ non possiamo calcolare $\sigma_{\bar{x}}$ (il denominatore del test statistico) perché quest'ultimo si ottiene dividendo la deviazione standard della popolazione per $\sqrt{n}$

Come calcolare il test statistico? Abbiamo bisogno di una formula che non dipende da σ

Quando σ non è nota

Se non conosciamo σ, possiamo stimarlo a partire dall'informazione contenuta nel campione

Analogamente a quanto avviene per la media:

s (la deviazione standard del campione) è la migliore stima puntuale di σ che possiamo ottenere, dato un campione di ampiezza n

Ricorda: la deviazione standard si calcola con la formula:

$$s = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}}$$

Poiché la stima viene effettuata sul campione, al denominatore n è sostituito da n-1

Quando σ non è nota

Quando non conosciamo la deviazione standard della popolazione la formula per il calcolo dell'errore standard della media è:

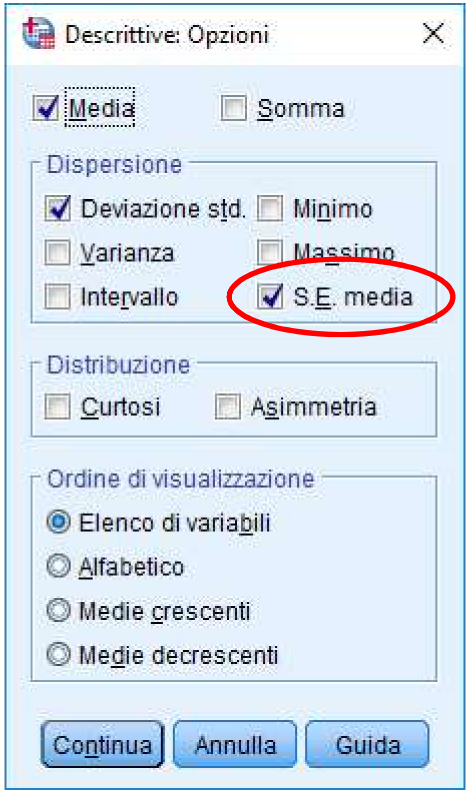
$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

Non possiamo però semplicemente sostituire $s_{\bar{x}}$ a $\sigma_{\bar{x}}$ nella formula per il calcolo del test statistico

Dobbiamo tenere in considerazione il fatto che s è una stima (imperfetta) di σ

L'errore standard in SPSS

È disponibile nella finestra di dialogo «Descrittive»
(selezionabile dal menu «Analizza», procedura «Statistiche Descrittive»)



Statistiche descrittive

	N	Media		Deviazione std.
	Statistica	Statistica	Errore std.	Statistica
Autoefficacia	40	7,8000	,35673	2,25616
Numero di casi validi (listwise)	40			

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{2,25616}{\sqrt{40}}$$

Quando σ non è nota

Per rendere conto di questa maggiore imprecisione, dobbiamo essere più conservativi (abbiamo bisogno di maggiori evidenze contro l'ipotesi nulla per rifiutare H_0)

Questa intuizione viene fatta risalire a William Gosset, un chimico noto con lo pseudonimo *Student:* a lui si deve il nome delle distribuzioni t di Student

Questa famiglia di distribuzioni è simile alla distribuzione normale (unimodale, simmetrica e asintotica), ma con code leggermente più "estese"

Quando σ non è nota

Di quanto la distribuzione t deve essere più «estesa» rispetto alla distribuzione normale?

Dipende da quanto la stima di σ è precisa, ovvero da quanto s si avvicina a σ

Se s è identico a σ, allora la distribuzione t di student deve essere identica alla distribuzione normale

Da cosa dipende l'accuratezza di s come stima di σ?

Dall'ampiezza del campione!

Quando σ non è nota

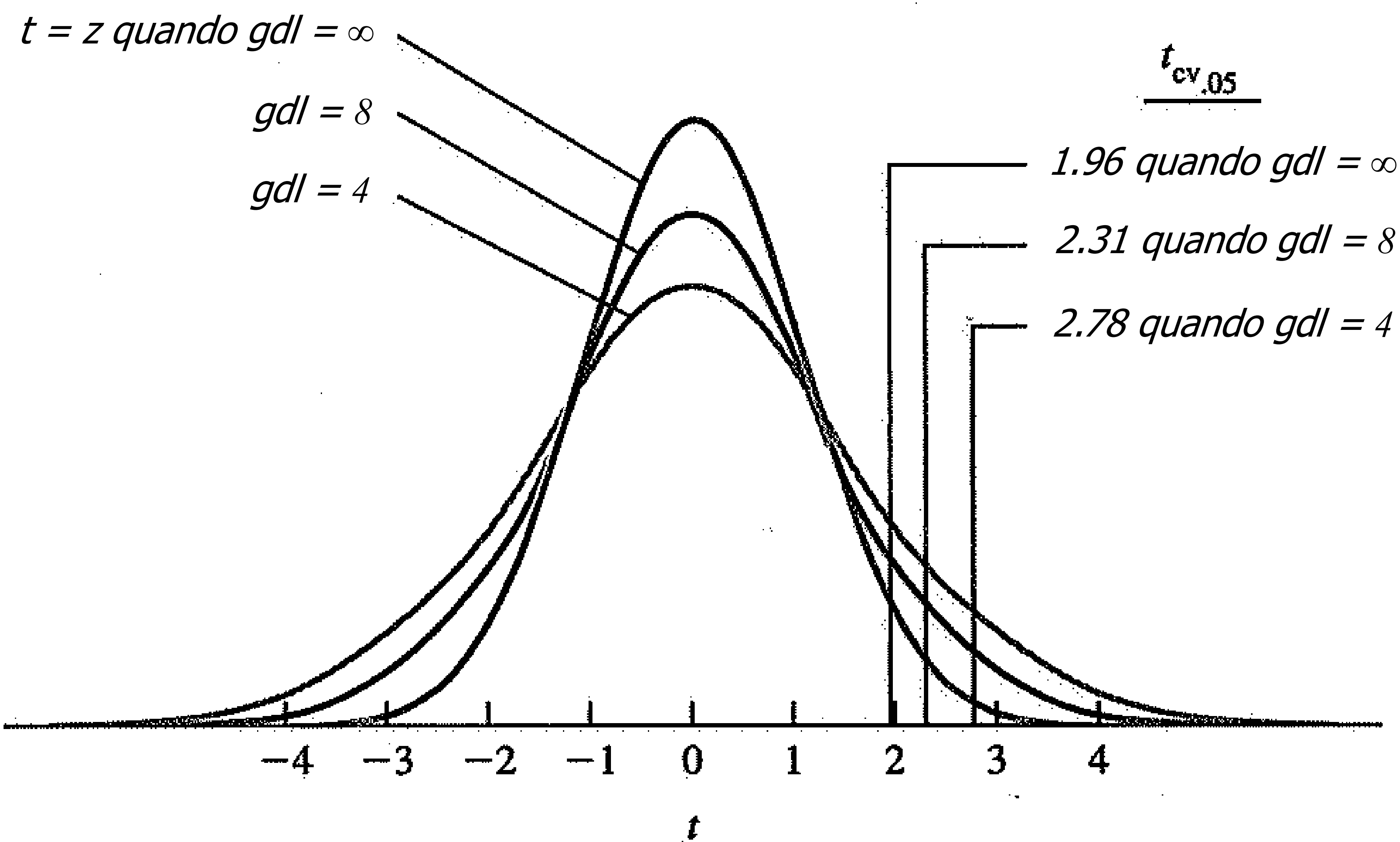
Minore è l'ampiezza del campione, meno la stima della deviazione standard della popolazione è precisa, più la distribuzione t è estesa rispetto alla distribuzione normale

Maggiore è l'ampiezza del campione, più la stima della deviazione standard della popolazione è precisa, più la distribuzione t è simile alla distribuzione normale

La distribuzione t dipende dunque da n (gradi di libertà)

I gradi di libertà sono pari a n -1

Verifica delle ipotesi



Verifica delle ipotesi

Quando n < 30

La distribuzione t di Student va usata anche quando si lavora su campioni di ampiezza inferiore a 30, estratti da una popolazione che non si distribuisce normalmente (o da popolazioni di cui non conosciamo la forma)

In questi casi, infatti, la distribuzione campionaria della media si discosta dalla distribuzione normale, risultando sempre più dispersa man mano che diminuisce la numerosità

Per questo bisogna considerare, come distribuzione di riferimento, la famiglia delle distribuzioni t di Student

Valori critici della distribuzione t di Student

gdl	Ipotesi monodirezionale .10	.05	.025	.01	.005	.0005
	Ipotesi bidirezionale .20	.10	.05	.02	.01	.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.598
3	1.638	2.353	3.182	4.541	5.841	12.941
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.859
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.405
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221
14	1.345	1.761	2.145	2.624	2.977	4.140
15	1.341	1.753	2.131	2.602	2.947	4.073
16	1.337	1.746	2.120	2.583	2.921	4.015
17	1.333	1.740	2.110	2.567	2.898	3.965
18	1.330	1.734	2.101	2.552	2.878	3.922
19	1.328	1.729	2.093	2.539	2.861	3.883
20	1.325	1.725	2.086	2.528	2.845	3.850
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.767
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	2.060	2.485	2.787	3.725
26	1.315	1.706	2.056	2.479	2.779	3.707
27	1.314	1.703	2.052	2.473	2.771	3.690
28	1.313	1.701	2.048	2.467	2.763	3.674
29	1.311	1.699	2.045	2.462	2.756	3.659
30	1.310	1.697	2.042	2.457	2.750	3.646
40	1.303	1.684	2.021	2.423	2.704	3.551
60	1.296	1.671	2.000	2.390	2.660	3.460
120	1.289	1.658	1.980	2.358	2.617	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.291

Verifica delle ipotesi

Come trovare il valore critico della t?

Nelle righe in alto viene riportato il livello critico di significatività (alfa) nel caso di ipotesi alternative monodirezionali e bidirezionali

Nella colonna a sinistra sono riportati i gradi di libertà

gdl	Ipotesi monodirezionale .10	.05	.025	.01	.005	.0005
	Ipotesi bidirezionale .20	.10	.05	.02	.01	.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.598
3	1.638	2.353	3.182	4.541	5.841	12.941
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.859
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.405
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221
14	1.345	1.761	2.145	2.624	2.977	4.140
15	1.341	1.753	2.131	2.602	2.947	4.073
16	1.337	1.746	2.120	2.583	2.921	4.015
17	1.333	1.740	2.110	2.567	2.898	3.965
18	1.330	1.734	2.101	2.552	2.878	3.922
19	1.328	1.729	2.093	2.539	2.861	3.883
20	1.325	1.725	2.086	2.528	2.845	3.850
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.767
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	2.060	2.485	2.787	3.725
26	1.315	1.706	2.056	2.479	2.779	3.707
27	1.314	1.703	2.052	2.473	2.771	3.690
28	1.313	1.701	2.048	2.467	2.763	3.674
29	1.311	1.699	2.045	2.462	2.756	3.659
30	1.310	1.697	2.042	2.457	2.750	3.646
40	1.303	1.684	2.021	2.423	2.704	3.551
60	1.296	1.671	2.000	2.390	2.660	3.460
120	1.289	1.658	1.980	2.358	2.617	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.291

Esempio

Un test di lettura applicato a bambini di 3° elementare ha media $\mu = 100$. La deviazione standard della popolazione non è nota

Il ricercatore vuole verificare se un campione di 30 bambini, sottoposti ad un corso di lettura, differisce dalla popolazione generale

La media del campione è pari a 103, con una deviazione standard di 9

Esempio

1. Formuliamo le ipotesi statistiche:

$H_0: \mu = 100$

$H_1: \mu \neq 100$ → **L'IPOTESI ALTERNATIVA È BIDIREZIONALE**

2. Scegliamo un valore critico di .05

- L'errore standard della media è pari a:

$$s_{\overline{X}} = \frac{9}{\sqrt{30}} = \frac{9}{5.48} = 1.64$$

3. Quale è il valore t critico associato ad un livello α di .05, quando l'ipotesi nulla è bidirezionale?

Verifica delle ipotesi

Come trovare il valore critico della t?

I gradi di libertà sono pari a 29 (30 -1)

L'ipotesi alternativa è bidirezionale

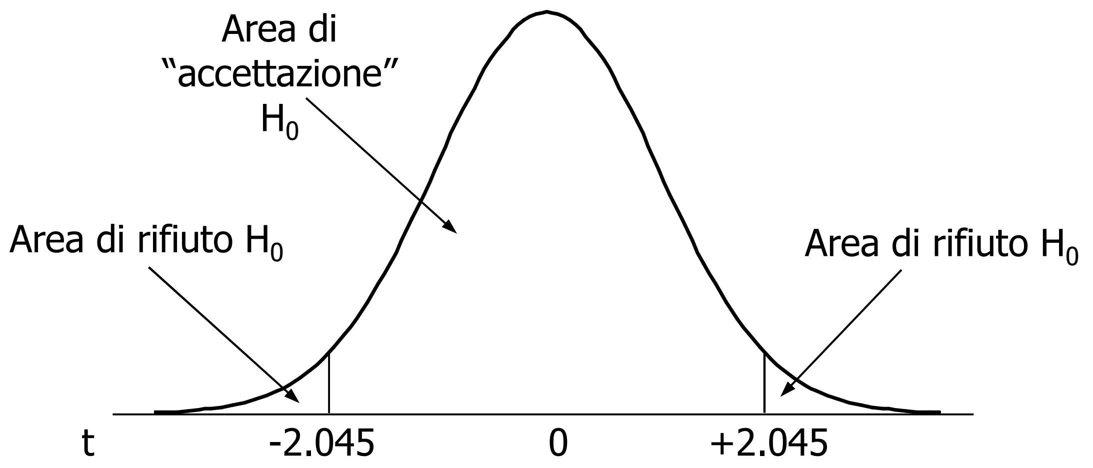
Il livello critico di significatività è del 5%

Il valore t critico è pari a 2.045

gdl	.10	.05	.025	.01	.005	.0005
	Ipotesi monodirezionale					
	Ipotesi bidirezionale					
	.20	.10	.05	.02	.01	.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.598
3	1.638	2.353	3.182	4.541	5.841	12.941
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.859
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.405
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221
14	1.345	1.761	2.145	2.624	2.977	4.140
15	1.341	1.753	2.131	2.602	2.947	4.073
16	1.337	1.746	2.120	2.583	2.921	4.015
17	1.333	1.740	2.110	2.567	2.898	3.965
18	1.330	1.734	2.101	2.552	2.878	3.922
19	1.328	1.729	2.093	2.539	2.861	3.883
20	1.325	1.725	2.086	2.528	2.845	3.850
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.767
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	2.060	2.485	2.787	3.725
26	1.315	1.706	2.056	2.479	2.779	3.707
27	1.314	1.703	2.052	2.473	2.771	3.690
28	1.313	1.701	2.048	2.467	2.763	3.674
29	1.311	1.699	$\boxed{2.045}$	2.462	2.756	3.659
30	1.310	1.697	2.042	2.457	2.750	3.646
40	1.303	1.684	2.021	2.423	2.704	3.551
60	1.296	1.671	2.000	2.390	2.660	3.460
120	1.289	1.658	1.980	2.358	2.617	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.291

NOTA. Nel test t l'ampiezza del campione (n) influenza la potenza del test non solo agendo sull'errore standard ma anche influenzando i gradi di libertà e dunque il valore t critico

Esempio



Dopo aver raccolto i dati:

4. Si calcola il valore del test statistico:

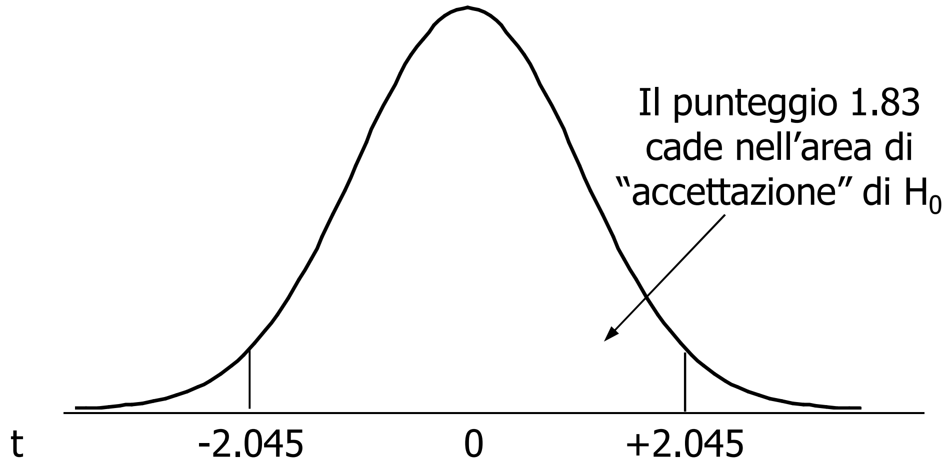
$$\frac{103-100}{1.64}=\frac{3}{1.64}=1.83$$

Esempio

5. Si confronta il valore del test statistico con il valore critico:

$$1.83 < 2.045$$

6. Poiché il test statistico è inferiore al valore critico (ovvero il test non è significativo), non possiamo rifiutare l'ipotesi nulla



Intervallo di confidenza

Quando σ non è noto, la formula per l'IC è:

$$\boldsymbol{limite\ inferiore}: \quad \bar{x} - (s_{\bar{x}} \times t_c) = \quad 103 - (1.64 \times 2.045) = \quad \mathbf{99.65}$$

$$\boldsymbol{limite\ superiore}: \quad \bar{x} + (s_{\bar{x}} \times t_c) = \quad 103 + (1.64 \times 2.045) = \quad \mathbf{106.35}$$

Si utilizza s come stimatore di σ nella formula dell'errore standard; i valori critici (t_c) si ricavano dalle tavole della distribuzione t di Student

In questo caso possiamo concludere che, con il 95% di probabilità, la media della popolazione da cui proviene il campione è compresa tra 99.65 e 106.35

Poiché la media ipotizzata in base ad H0 (μ = 100) ricade all'interno di questo intervallo, non possiamo rifiutare l'ipotesi nulla

Intervallo di confidenza

Come già detto in precedenza per il test z, l'IC si può anche centrare attorno alla differenza tra la media osservata sul campione e la media attesa in base ad H0:

$$\textit{limite inferiore}: \quad diff - (s_{\bar{x}} \times t_c) = \quad 3 - (1.64 \times 2.045) = \quad \mathbf{-0.35}$$

$$\textit{limite superiore}: \quad diff + (s_{\bar{x}} \times t_c) = \quad 3 + (1.64 \times 2.045) = \quad \mathbf{6.35}$$

Se l'intervallo <u>non</u> contiene lo zero rifiutiamo H0: con il 95% di probabilità, la media della popolazione da cui proviene il campione è diversa dalla media ipotizzata in base ad H0 («la differenza è significativa»)

Se l'intervallo include lo zero, come in questo caso, <u>non</u> possiamo rifiutare H0 («la differenza non è significativa)

Dimensione dell'effetto

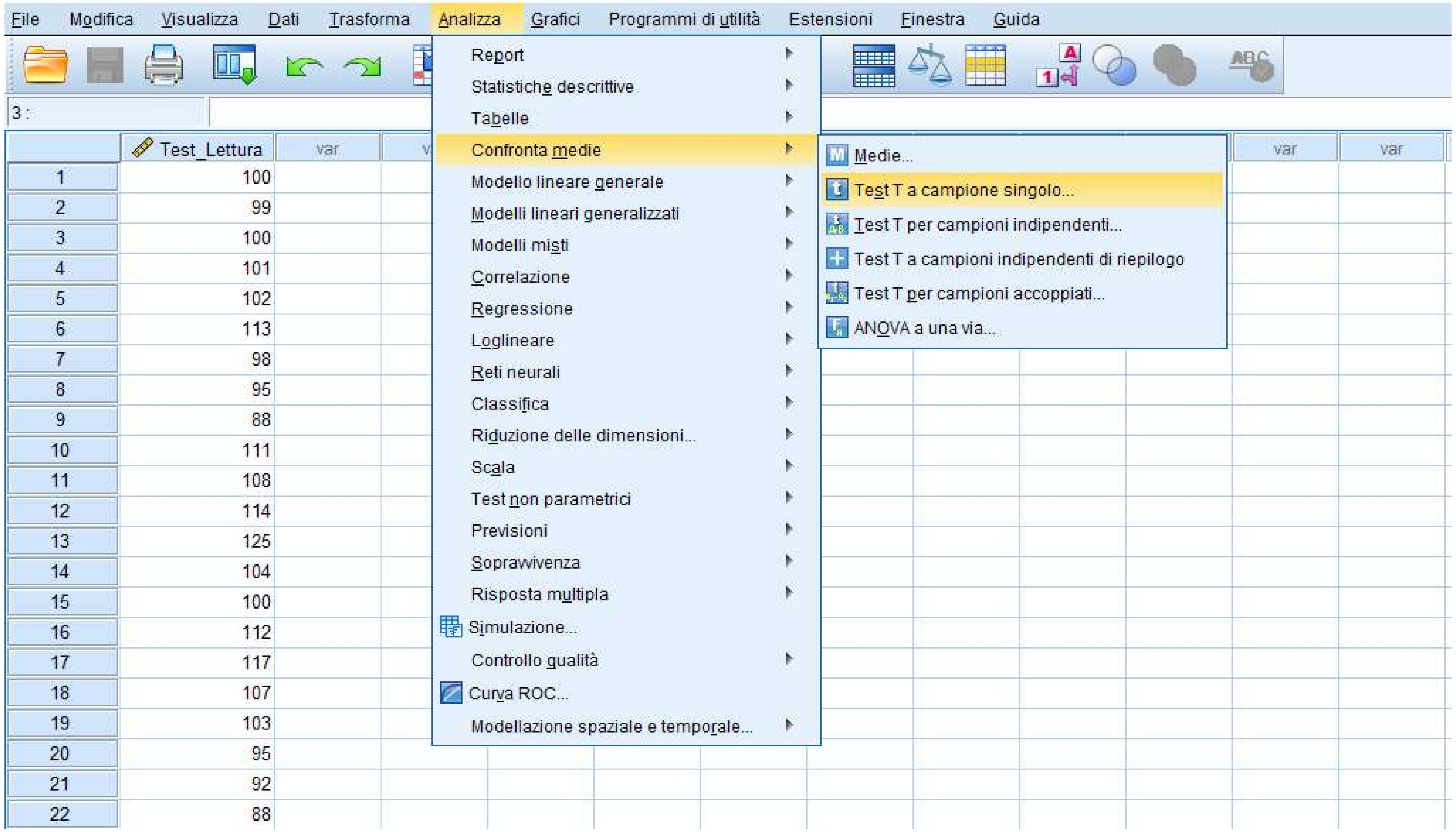
Nel test t su un campione, la formula per il calcolo dell'ampiezza dell'effetto (d di Cohen) è:

$$d = \frac{\overline{X} - \mu}{s}$$

Le linee guida per l'interpretazione della d sono uguali a quelle del test z su un campione:

<.20: trascurabile
.20 - .50: piccolo
.50 - .80: moderato
> .80: grande

Il test t per un campione in SPSS

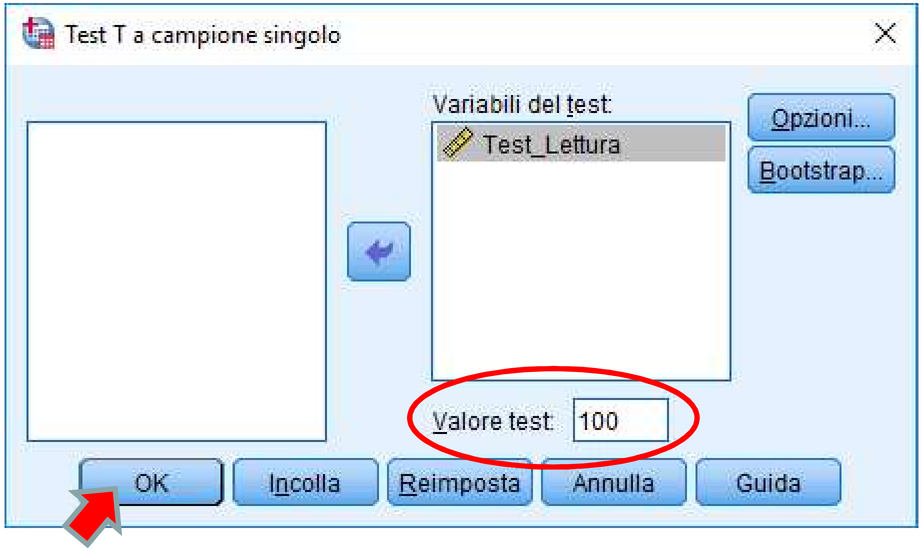


È disponibile nella finestra di dialogo «Test T a campione singolo»

(selezionabile dal menu «Analizza», procedura «Confronta medie»)

Il test t per un campione in SPSS

1. Selezionare la variabile (es. «Test_Lettura») e spostarla nel menu delle variabili attive

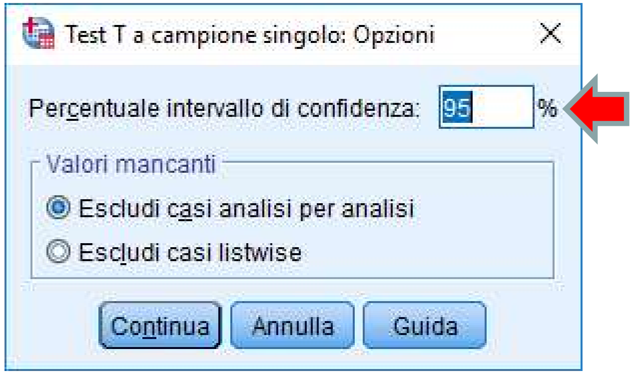


3. Cliccare su «**OK**»

2. Inserire, nel riquadro «Valore test», la media della popolazione da cui, secondo l'ipotesi nulla, proviene il campione (es. «100»)
Questo valore rappresenta il **valore atteso in base ad H0**

Il test t per un campione in SPSS

Cliccando su «Opzioni», è possibile modificare il livello di fiducia dell'intervallo di confidenza



Il valore di default è 95%

Il test t per un campione in SPSS

Errore standard della media

Statistiche campione singolo

	N	Media	Deviazione std.	Media errore standard
Test_Lettura	30	103,00	9,002	1,644

Media osservata sul campione

Valore atteso in base ad H0

Test a campione singolo

	Valore di test = 100					
	t	gl	Sign. (a due code)	Differenza della media	Intervallo di confidenza della differenza di 95% Inferiore	Superiore
Test_Lettura	1,825	29	,078	3,000	-,36	6,36

t di Student (t «empirica»)

gradi di libertà (N -1)

Differenza tra media osservata sul campione e media attesa in base ad H0: **numeratore del test t**

Livello di significatività:
probabilità di osservare un valore del test statistico pari ad almeno 1.825, assumendo che l'ipotesi nulla sia vera

non è il valore critico di probabilità!
Quest'ultimo è fissato a priori dal ricercatore, in genere a .05

Se «Sign.» < .05, si rifiuta H0

Il test t per un campione in SPSS

Statistiche campione singolo

	N	Media	Deviazione std.	Media errore standard
Test_Lettura	30	103,00	9,002	1,644

Test a campione singolo

	Valore di test = 100					
	t	gl	Sign. (a due code)	Differenza della media	Intervallo di confidenza della differenza di 95% Inferiore	Superiore
Test_Lettura	1,825	29	,078	3,000	-,36	6,36

$$\lim \inf = 3 - (2.045 \times 1.644) = -0.36$$

$$\lim \sup = 3 + (2.045 \times 1.644) = +6.36$$

Differenza tra media osservata e media attesa in base a H0

t-critico per alfa = .05 (con H_1 bidirezionale) e 29 gdl

errore standard della media

Nel t-test su un campione, l'intervallo di confidenza è centrato attorno alla differenza tra la media osservata sul campione e la media attesa in base ad H_0

Con il 95% di probabilità, questa differenza ricade tra i limiti definiti dall'intervallo

Se l'intervallo **non** contiene lo zero, possiamo **rifiutare H0**: con il 95% di probabilità, la differenza tra questi due valori è diversa da zero

Il test t per un campione in SPSS

Statistiche campione singolo

	N	Media	Deviazione std.	Media errore standard
Test_Lettura	30	103,00	9,002	1,644

Test a campione singolo

	Valore di test = 100					
	t	gl	Sign. (a due code)*	Differenza della media	Intervallo di confidenza della differenza di 95% Inferiore	Superiore
Test_Lettura	1,825	29	,078	3,000	-,36	6,36

*** Nota: l'ipotesi alternativa è bidirezionale. In SPSS non è possibile definire H1 come unidirezionale**

Se vogliamo utilizzare un'ipotesi alternativa unidirezionale possiamo:

1. **confrontare il valore del test statistico (t = 1.825) con il valore critico ricavabile dalle tavole della distribuzione t di Student (con gdl = 29, alfa = .05 e H1 unidirezionale, la t «critica» è pari a 1.699: poiché 1.825 > 1.669, possiamo rifiutare H0)**

oppure

2. **dividere per 2 la probabilità del test (Sign./2 → .078/2= .039) e confrontarla con il livello critico di probabilità prescelto (.05): poiché .039 < .05, possiamo rifiutare H0**

Il test t per un campione in SPSS

Statistiche campione singolo

	N	Media $\bar{x}$	Deviazione std. s	Media errore standard
Test_Lettura	30	103,00	9,002	1,644

Test a campione singolo

Valore di test = 100 μ

	t	gl	Sign. (a due code)	Differenza della media	Intervallo di confidenza della differenza di 95%	
					Inferiore	Superiore
Test_Lettura	1,825	29	,078	3,000	-,36	6,36

Nell'output non viene riportata l'ampiezza dell'effetto.

La d di Cohen si può comunque calcolare facilmente, utilizzando i dati riportati nelle tabelle:

$$d = \frac{\bar{X} - \mu}{s} =$$

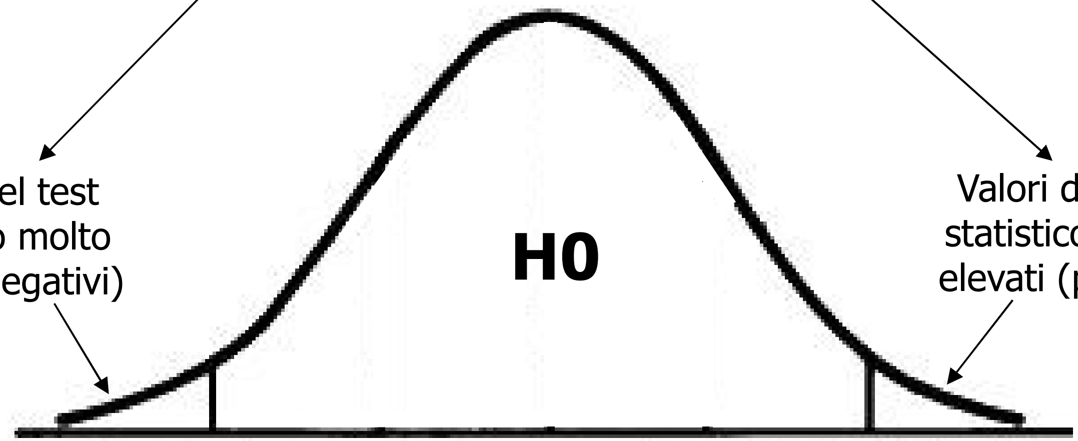
$$\frac{103 - 100}{9} = \frac{3}{9} = .33$$

Per riassumere …

valori del test statistico <u>molto elevati</u> sono estremamente improbabili

più i valori del test statistico sono elevati, minore è la probabilità che si verifichino (nella distribuzione definita da H0)

Valori del test statistico molto elevati (negativi)



Valori del test statistico molto elevati (positivi)

<u>Punteggi</u>	*z*	-3	-1.96	-1	0	+1	+1.96	+3
<u>Probabilità</u>	*p*	.01	.05				.05	.01

valore critico

valore critico

Per riassumere ...

Per decidere se respingere o meno l'ipotesi nulla possiamo procedere in due modi:

1. Calcolare il valore del test statistico (es. t «empirica») e confrontarlo con il valore critico della distribuzione (t «critica»)

Rifiuto H0 se: t «empirica» > t «critica»

è un valore poco probabile sotto H0 (la probabilità di osservare questo valore - o valori superiori - è del 5%)

Per riassumere …

2. Calcolare la probabilità che si verifichi un valore del test statistico pari ad almeno quello osservato nel campione, assumendo che H0 sia vera (livello di significatività) e confrontare questa probabilità con il livello <u>critico</u> di significatività (es. .05).

Rifiuto H0 se: **livello di significatività < .05**

probabilità della t «empirica» osservata (assumendo che H0 sia vera)

livello critico di significatività (alfa). È un valore soglia definito a priori

Esercizio 5

In una prova di orientamento spaziale, si vuole verificare se un gruppo di 15 soggetti, per i quali si sospetta un disturbo neurofisiologico, presenta un punteggio inferiore rispetto alla popolazione generale. I soggetti ottengono un punteggio medio pari 97.3 (s=12.51)

Il punteggio medio alla prova nella popolazione generale, quando le funzioni cognitive sono integre, è pari a 100

È possibile sostenere che i soggetti in questione presentano effettivamente un disturbo neurofisiologico che ne limita la prestazione al test? Considera un livello di alfa pari a .05

Verifica delle ipotesi

Esercizio 6

Uno psicologo sostiene che il training autogeno ha un qualche effetto sull'ipertensione. Per avvalorare questa affermazione, egli analizza i dati relativi alla pressione diastolica di 32 suoi pazienti dopo 4 settimane di training autogeno. I risultati sono i seguenti:

Statistiche campione singolo

	N	Media	Deviazione std.	Media errore standard
Pressione	32	106,97	11,765	2,080

Test a campione singolo

	Valore di test = 115					
					Intervallo di confidenza della differenza di 95%	
	t	gl	Sign. (a due code)	Differenza della media	Inferiore	Superiore
Pressione	-3,862	31	,001	-8,031	-12,27	-3,79

1. La media della pressione diastolica osservata nel campione di pazienti è pari a:
2. Il valore atteso in base ad H0 è pari a:
3. La deviazione standard della distribuzione campionaria delle medie è pari a:
4. Il t test è significativo (con alfa=.05)? Possiamo rifiutare H0? Cosa è possibile concludere?
5. Commentare il risultato relativo all'intervallo di confidenza:
6. Cosa cambia se consideriamo un livello di alfa= .01?
7. Cosa cambia se consideriamo H1 come monodirezionale (alfa= .05)?

Esercizio 7

Un campione di 17 adolescenti a rischio di abbandono scolastico viene sottoposto ad un test di disadattamento sociale, ottenendo un punteggio medio pari a ____ e una dev.st. pari a ____ . Nella popolazione di adolescenti la variabile ha una media pari a ____. È stata sottoposta a verifica l'ipotesi che gli adolescenti a rischio di abbandono scolastico presentino livelli più elevati di disadattamento sociale rispetto alla popolazione di riferimento, dato un livello di alfa del 5%. I risultati sono i seguenti:

Statistiche campione singolo

	N	Media	Deviazione std.	Media errore standard
Disad_Soc	17	29,29	6,008	1,457

Test a campione singolo

	Valore di test = 25					
					Intervallo di confidenza della differenza di 95%	
	t	gl	Sign. (a due code)	Differenza della media	Inferiore	Superiore
Disad_Soc	2,947	16	,009	4,294	1,21	7,38

1. Completa la traccia con i valori mancanti (ricavabili dall'output di SPSS):
2. In base a quanto riportato nella traccia, formula H0 e H1: ________________________ ;
3. L'errore standard della media è pari a: ____________ ;
4. Il t test è significativo? Possiamo rifiutare H0? Cosa è possibile concludere?
5. Commenta il risultato relativo all'intervallo di confidenza:
6. Calcola l'ampiezza dell'effetto: cosa si può concludere?
7. Cosa cambia se consideriamo un livello di alfa= .01?

Esercizio 8

In un campione di 250 studenti al terzo anno di scuola media il punteggio nella scala dell'apertura mentale del *Big Five Questionnaire* (*BFQ*) ha media $\overline{X}$ = 84.2 e dev. standard s = 12.5

Calcola l'intervallo di confidenza della media campionaria, con un livello di fiducia del 95%. Limite inferiore e superiore dell'intervallo di confidenza sono pari rispettivamente a:

a) 82.6 e 86.2 b) 71.7 e 96.7 c) 68.4 e 100.0 d) 82.6 e 85.8

Cosa cambia nella stima intervallare se viene scelto un livello di fiducia del 99% (invece del 95%)?

a) l'intervallo di fiducia diventa più ristretto
b) l'intervallo di fiducia diventa più ampio
c) l'intervallo di fiducia rimane uguale

Cosa cambia nella stima intervallare se decidiamo di aumentare l'ampiezza del campione?

a) l'intervallo di fiducia diventa più ristretto
b) l'intervallo di fiducia diventa più ampio
c) l'intervallo di fiducia rimane uguale

LA VERIFICA DELLE IPOTESI NEL CASO DI DUE CAMPIONI INDIPENDENTI

La verifica delle ipotesi nel caso un campione

Negli applicazioni precedenti il ricercatore era interessato a valutare delle ipotesi riferite ad una popolazione, utilizzando un singolo campione

L'ipotesi nulla, in questi casi, ha la seguente forma:

$$H_0: \mu = ?$$

È il valore definito dall'ipotesi nulla

Nella verifica delle ipotesi su un campione, il ricercatore è interessato a verificare un'ipotesi sulla media della popolazione, potendo o meno conoscerne la deviazione standard

La verifica delle ipotesi nel caso un campione

Come abbiamo detto, nel caso di un singolo campione la forma del test statistico è:

Quale STATISTICA? La media osservata sul campione

$$\frac{\text{Statistica} - \text{Parametro}}{\text{Errore standard}}$$

Quale PARAMETRO? Quello ipotizzato nella popolazione in base ad H_0: uno specifico valore definito a priori

Quale ERRORE STANDARD? Quello della distribuzione campionaria delle medie

La verifica delle ipotesi nel caso di due campioni

Il caso più frequente nella ricerca empirica, tuttavia, riguarda l'analisi della differenze tra due campioni

In questo caso siamo interessati a verificare se le popolazioni da cui provengono i campioni differiscono per la caratteristica oggetto di studio

Esempi:

- **I bambini piccoli sono più creativi di quelli grandi?**
- **Le femmine sono più socievoli dei maschi?**
- **Gli elettori di destra hanno maggiori livelli di autoritarismo rispetto a quelli di sinistra?**

La verifica delle ipotesi nel caso di due campioni

La logica del procedimento alla base della verifica delle ipotesi non cambia

L'ipotesi nulla e l'ipotesi alternativa sono del tipo:

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$ → **IPOTESI ALTERNATIVA BIDIREZIONALE**

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 > \mu_2$ → **IPOTESI ALTERNATIVA MONODIREZIONALE "DESTRA"**

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 < \mu_2$ → **IPOTESI ALTERNATIVA MONODIREZIONALE "SINISTRA"**

La verifica delle ipotesi nel caso di due campioni

Cambia invece la distribuzione di riferimento

La distribuzione campionaria delle medie non è utilizzabile in questi casi, perché è riferita ad un singolo campione

È necessario fare riferimento ad un'altra distribuzione: la distribuzione campionaria della differenza tra due medie

La verifica delle ipotesi nel caso di due campioni

Per costruire la distribuzione campionaria della differenza tra due medie occorre:

1→estrarre dalla prima popolazione tutti i possibili campioni di ampiezza n_1 e calcolare la media di ciascun campione

2→estrarre dalla seconda popolazione tutti i possibili campioni di ampiezza n_2 e calcolare la media di ciascun campione

3→calcolare le differenze tra le medie di tutte le possibili coppie di campioni provenienti dalle due popolazioni

4→calcolare la media e la deviazione standard di tutte queste differenze [distribuzione campionaria della differenza tra due medie]

La verifica delle ipotesi nel caso di due campioni

La media della distribuzione campionaria della differenza tra due medie viene indicata con il simbolo:

$$\mu_{\bar{x}1-\bar{x}2}$$

È possibile dimostrare che la media della distribuzione campionaria della differenza tra due medie è uguale alla differenza tra le medie delle due popolazioni

$$\mu_{\bar{x}1-\bar{x}2} = \mu_1 - \mu_2$$

La verifica delle ipotesi nel caso di due campioni

La varianza della distribuzione campionaria della differenza tra due medie viene indicata con il simbolo:

$$\sigma^2_{\bar{x}1-\bar{x}2}$$

È possibile dimostrare che tale varianza è uguale alla somma delle varianze delle distribuzioni campionarie delle medie provenienti dalle due popolazioni

$$\sigma^2_{\bar{x}1-\bar{x}2} = \sigma^2_{\bar{x}1} + \sigma^2_{\bar{x}2}$$

sono gli errori standard della media al quadrato

La verifica delle ipotesi nel caso di due campioni

La deviazione standard della distribuzione campionaria della differenza tra due medie viene indicata con il simbolo:

$$\sigma_{\bar{x}1-\bar{x}2}$$

Viene chiamata anche: "errore standard della differenze tra le medie"

È pari alla radice quadrata della varianza della distribuzione campionaria della differenza tra due medie:

$$\sigma_{\bar{x}1-\bar{x}2} = \sqrt{\sigma^2{}_{\bar{x}1-\bar{x}2}}$$

La verifica delle ipotesi nel caso di due campioni

La distribuzione campionaria della differenza tra le medie di due campioni di numerosità n_1 e n_2 ha forma normale se:

$$n_1 \text{ e } n_2 > 30$$

La verifica delle ipotesi nel caso di due campioni

Anche nel caso di due campioni, la formula del test statistico è:

Quale STATISTICA?
La differenza osservata tra le medie dei due campioni

$$\frac{\text{Statistica} - \text{Parametro}}{\text{Errore standard}}$$

Quale PARAMETRO?
La differenza tra le medie delle due popolazioni, ipotizzata in base ad H_0

Quale ERRORE STANDARD?
Quello della distribuzione campionaria della differenza tra due medie

La verifica delle ipotesi nel caso di due campioni

Anche nel caso di due campioni, la formula del test statistico è:

Quale STATISTICA?
La differenza osservata tra le medie dei due campioni

$$\frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sigma_{\bar{X}1-\bar{X}2}}$$

Quale PARAMETRO?
La differenza tra le medie delle due popolazioni, ipotizzata in base ad H_0

Quale ERRORE STANDARD?
Quello della distribuzione campionaria della differenza tra due medie

In base all'ipotesi nulla, questo valore è uguale a zero

La verifica delle ipotesi nel caso di due campioni

La formula del «test z su due campioni indipendenti» si può semplificare in:

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{X}1-\bar{X}2}}$$

$$\sqrt{\sigma^2_{\bar{X}1} + \sigma^2_{\bar{X}2}}$$

Varianza della distribuzione campionaria delle medie (il quadrato dell'errore standard) della prima popolazione

$$\frac{\sigma_1^{\ 2}}{n_1}$$

$$\frac{\sigma_2^{\ 2}}{n_2}$$

Varianza della distribuzione campionaria delle medie (il quadrato dell'errore standard) della seconda popolazione

La verifica delle ipotesi nel caso di due campioni

«test t»: se la varianza delle popolazioni non è nota si utilizza la varianza dei due campioni per stimarla:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}1-\bar{X}2}}$$

$$s_{\bar{X}1-\bar{X}2} = \sqrt{\frac{s_1^{\;2}}{n_1} + \frac{s_2^{\;2}}{n_2}}$$

Esempio

Uno psicologo è interessato a verificare se le femmine adolescenti sono socievoli quanto i maschi coetanei

Per verificare l'ipotesi, lo psicologo somministra una scala per la misura della socievolezza a due campioni, composti da 50 maschi e 50 femmine

I risultati sono i seguenti:

$$\overline{X}_F = 32.7 \qquad \overline{X}_M = 30.3$$

$$s^2_F = 68.03 \qquad s^2_M = 82.95$$

Fissiamo il livello critico di probabilità a .05

Cosa è possibile concludere?

Esempio

Le ipotesi statistiche sono:

$$H_0: \mu_F = \mu_M$$
$$H_1: \mu_F \neq \mu_M$$

Esempio

Poiché lo psicologo non conosce la deviazione standard della popolazione, per calcolare l'errore standard della differenza tra le medie occorre:

- Utilizzare s come stimatore di σ

- Calcolare la varianza della distribuzione campionaria delle medie nel campione delle femmine:

$$s^2_{\bar{X}F} = \frac{68.03}{50} = 1.37$$

- Calcolare la varianza della distribuzione campionaria delle medie nel campione dei maschi:

$$s^2_{\bar{X}M} = \frac{82.95}{50} = 1.66$$

Esempio

- Calcolare la varianza della distribuzione campionaria della differenza tra le medie:

$$s^2_{\bar{X}_F - \bar{X}_M} = 1.37 + 1.66 = 3.03$$

- Calcolare la deviazione standard della distribuzione campionaria della differenza tra le medie (o errore standard della differenza tra le medie):

$$\sqrt{3.03} = 1.74$$

È il denominatore del test statistico

Verifica delle ipotesi

Dobbiamo ora trovare il valore t critico, utilizzando la tavola della t

Nel caso di due campioni, i gradi di libertà sono pari a: n1 + n2 – 2:

50 + 50 – 2 = 98

Il valore critico, per α .05 e ipotesi alternativa bidirezionale, è pari a 1.980

gdl	Ipotesi monodirezionale .10	.05	.025	.01	.005	.0005
	Ipotesi bidirezionale .20	.10	.05	.02	.01	.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.598
3	1.638	2.353	3.182	4.541	5.841	12.941
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.859
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.405
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221
14	1.345	1.761	2.145	2.624	2.977	4.140
15	1.341	1.753	2.131	2.602	2.947	4.073
16	1.337	1.746	2.120	2.583	2.921	4.015
17	1.333	1.740	2.110	2.567	2.898	3.965
18	1.330	1.734	2.101	2.552	2.878	3.922
19	1.328	1.729	2.093	2.539	2.861	3.883
20	1.325	1.725	2.086	2.528	2.845	3.850
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.767
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	2.060	2.485	2.787	3.725
26	1.315	1.706	2.056	2.479	2.779	3.707
27	1.314	1.703	2.052	2.473	2.771	3.690
28	1.313	1.701	2.048	2.467	2.763	3.674
29	1.311	1.699	2.045	2.462	2.756	3.659
30	1.310	1.697	2.042	2.457	2.750	3.646
40	1.303	1.684	2.021	2.423	2.704	3.551
60	1.296	1.671	2.000	2.390	2.660	3.460
120	1.289	1.658	$\boxed{1.980}$	2.358	2.617	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.291

Esempio

Dobbiamo infine calcolare il valore del test statistico e confrontarlo con il valore critico

Il test statistico è pari a:

$$\frac{32.7-30.3}{1.74}=\frac{2.4}{1.74}=1.38$$

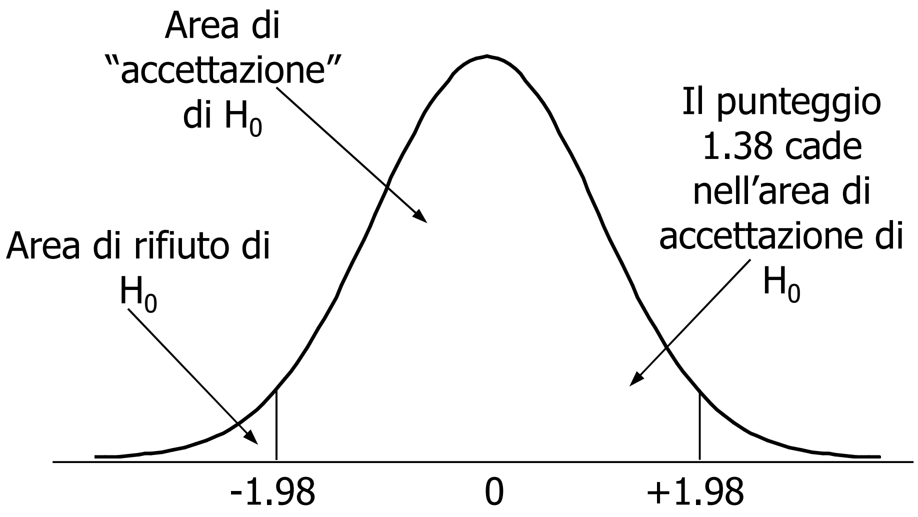
Confronto tra il test statistico e il valore critico:

$$1.38<1.98$$

Poiché il test statistico è inferiore al valore critico (ovvero il test non è significativo), non possiamo rifiutare l'ipotesi nulla

Verifica delle ipotesi

Esempio



Possiamo concludere che non vi sono differenze di genere nella socievolezza

La differenza tra le medie dei due campioni NON è significativa, ovvero i due campioni provengono dalla stessa popolazione (o da due popolazioni che hanno la stessa media)

Alcune precisazioni …

Nel caso di due campioni, l'intervallo di confidenza è centrato attorno alla differenza tra le due medie:

$$\begin{array}{l}\textit{limite inferiore}:\\ diff - (s_{\bar{x}1-\bar{x}2} \times t_c) =\\ 2.4 - (1.74 \times 1.98) =\\ \mathbf{-1.05}\end{array} \qquad \begin{array}{l}\textit{limite superiore}:\\ diff + (s_{\bar{x}1-\bar{x}2} \times t_c) =\\ 2.4 + (1.74 \times 1.98) =\\ \mathbf{5.85}\end{array}$$

Se l'intervallo non contiene lo zero rifiutiamo H0: con il 95% di probabilità, i due campioni provengono da popolazioni con medie differenti («la differenza è significativa»)

Se l'intervallo include lo zero, come in questo caso, non rifiutiamo H0 («la differenza non è significativa): i due campioni provengono dalla stessa popolazione (o da popolazioni che hanno la stessa media)

Alcune precisazioni ...

La formula per il calcolo dell'errore standard della differenza tra le medie che abbiamo utilizzato è:

$$s_{\bar{x}1-\bar{x}2} = \sqrt{\frac{s_1^{\,2}}{n_1} + \frac{s_2^{\,2}}{n_2}}$$

Questa formula può essere utilizzata solo se n è uguale nei due gruppi (i due gruppi sono omogenei)

Verifica delle ipotesi

Alcune precisazioni ...

Se la numerosità dei due gruppi è differente ($n_1 \neq n_2$), la formula da utilizzare al denominatore del test è (pag. 111 del libro):

$$s_{\bar{x}1-\bar{x}2} = \sqrt{\frac{n_1 - 1\, s_1^2 + n_2 - 1\, s_2^2}{n_1 + n_2 - 2}\left(\frac{n_1 + n_2}{n_1 n_2}\right)}$$

Questo termine viene definito «stimatore congiunto della varianza (*pooled variance*)

È una media ponderata delle due varianze

Quando la n è identica nei due gruppi, la formula coincide con: $\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

Alcune precisazioni …

L'utilizzo dello stimatore congiunto assume che le varianze delle popolazioni da cui provengono i due campioni siano omogenee

Questa assunzione viene definita omoschedasticità

Per esaminare se le varianze sono omogenee si può applicare la verifica delle ipotesi

In questo caso il parametro di interesse è la varianza: vogliamo verificare delle ipotesi su σ^2, a partire dalla nostra conoscenza di s^2

La distribuzione campionaria di riferimento è la distribuzione campionaria del rapporto tra due varianze

Verifica delle ipotesi

Alcune precisazioni ...

La distribuzione campionaria del rapporto tra due varianze approssima una distribuzione nota e tabulata, la distribuzione F di Fisher

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

Il test per verificare l'omoschedasticità delle varianze (test di Levene) sottopone a verifica le seguenti ipotesi statistiche:

$$H0: \sigma^2_1 = \sigma^2_2$$

$$H1: \sigma^2_1 \neq \sigma^2_2$$

Torneremo su questi aspetti nelle prossime lezioni, quando affronteremo l'analisi della varianza

Verifica delle ipotesi

Alcune precisazioni …

Per sottoporre a verifica queste ipotesi si procede nel modo consueto (analogamente a quanto avviene per il test t):

(1) si calcola il test statistico (test F): se le due varianze sono simili, il loro rapporto è vicino a 1;

(2) si confronta il test F con il valore critico della distribuzione

(3) se F empirico > F critico, si rifiuta H0: le varianze non sono omogenee (l'assunzione di omoschedasticità non viene rispettata)

Torneremo su questi aspetti nelle prossime lezioni, quando affronteremo l'analisi della varianza

Verifica delle ipotesi

Alcune precisazioni ...

Se l'omoschedasticità non si verifica, il denominatore del test t di Student va calcolato con la formula:

$$\sqrt{\frac{s_1^{\,2}}{n_1}+\frac{s_2^{\,2}}{n_2}}$$

(è la formula vista in precedenza, che abbiamo usato quando n è uguale nei due gruppi: t di Welch)

I gradi di libertà si calcolano con una formula più complessa (non più $n_1 + n_2 - 2$) che tiene conto della differenza tra le varianze (Satterthwaite approximation)

Dimensione dell'effetto

Nel test t su due campioni indipendenti si utilizzano le seguenti formule per il calcolo della d di Cohen:

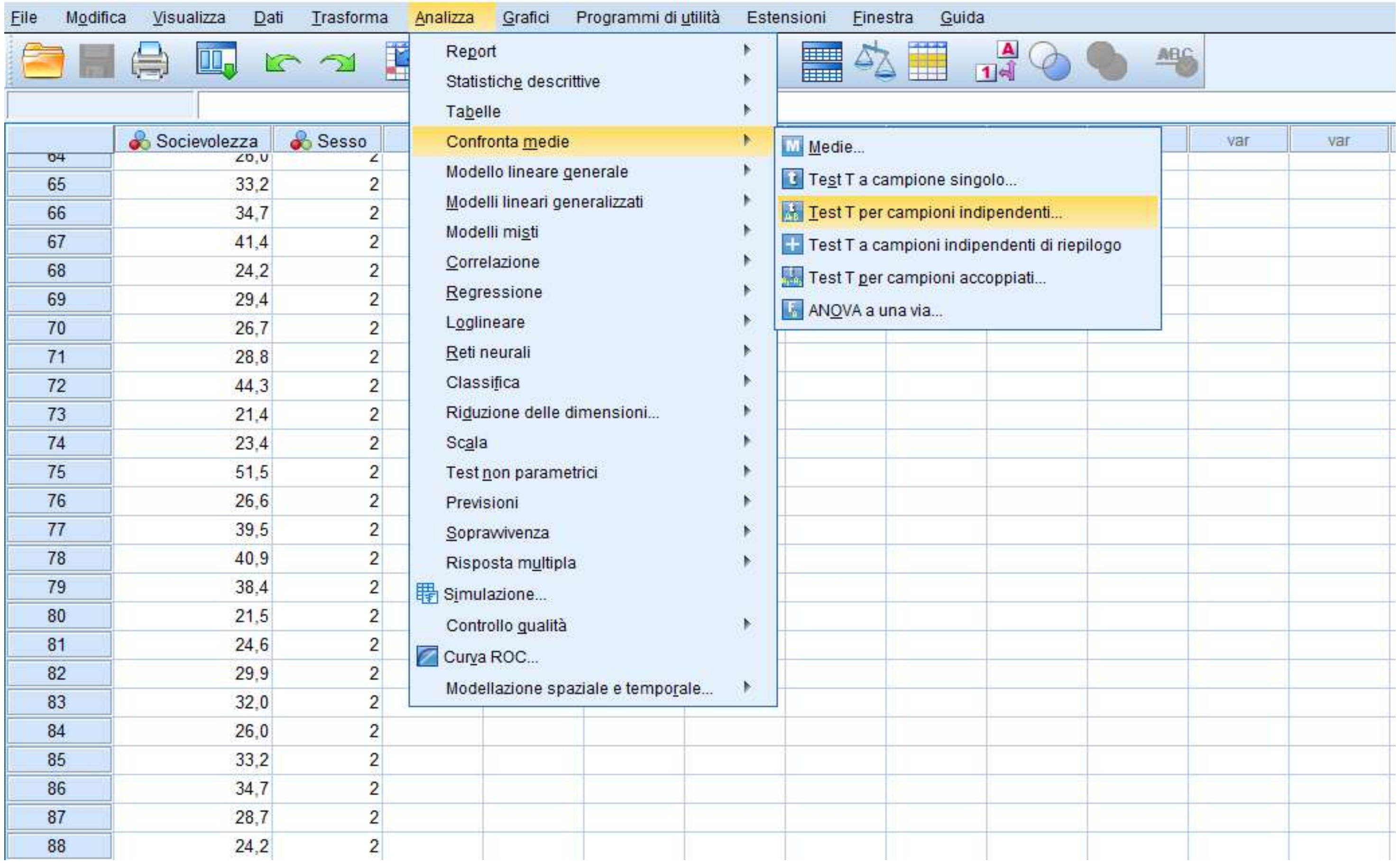
$$d = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}}$$

Quando n è uguale nei due gruppi, il denominatore si semplifica in:

$$\sqrt{\frac{s_1^2 + s_2^2}{2}}$$

Le linee guida per l'interpretazione della d sono uguali a quelle dei test z e t su un campione

Il test t per due campioni indipendenti in SPSS

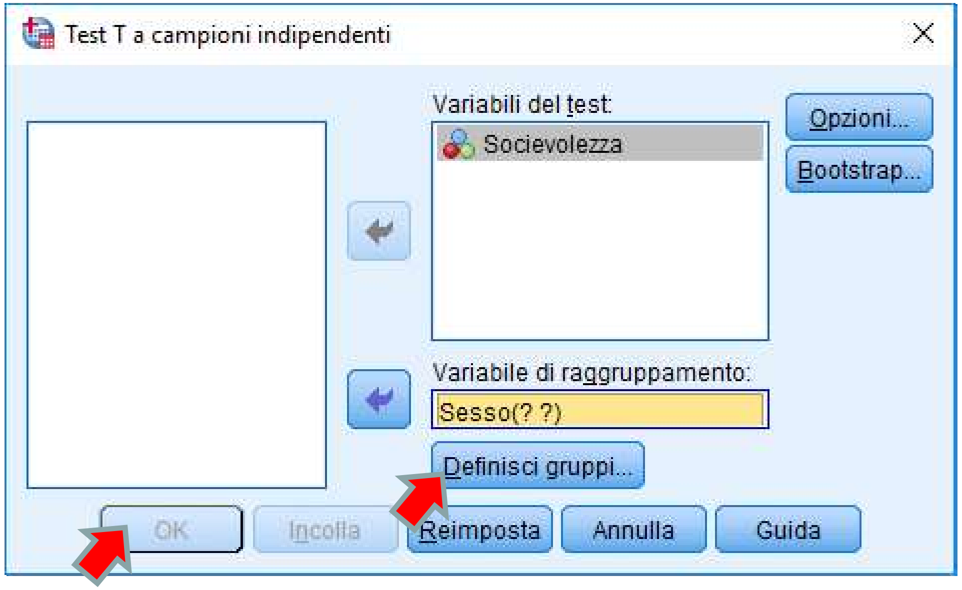


È disponibile nella finestra di dialogo «Test T per campioni indipendenti»

(selezionabile dal menu «Analizza», procedura «Confronta medie»)

Il test t per due campioni indipendenti in SPSS

1. Selezionare la variabile metrica, quella di cui si vuole calcolare la media (es. «Socievolezza») e spostarla nel menu delle variabili attive

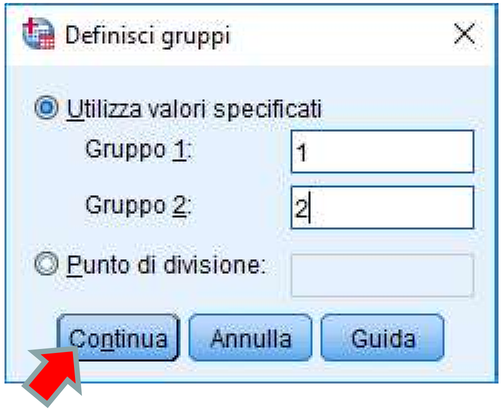


3. Cliccare su «**OK**»

2. Inserire, nel riquadro «Variabile di raggruppamento», la variabile che definisce l'appartenenza dei soggetti ai due campioni (es. «Sesso»)

3. Cliccare su «Definisci gruppi»

4. Definire i valori utilizzati per codificare i due gruppi (es. 1= «maschi», 2= «femmine»)



5. Cliccare su «Continua»

Il test t per due campioni indipendenti in SPSS

Errore standard della media ($\sigma_{\bar{x}}$) calcolato separatamente nei due gruppi

Esempio I

Statistiche gruppo

	Sesso	N	Media	Deviazione std.	Media errore standard
Socievolezza	1	50	32,700	8,2477	1,1664
	2	50	30,300	9,1077	1,2880

Test campioni indipendenti

		Test di Levene per l'eguaglianza delle varianze		Test t per l'eguaglianza delle medie						
									Intervallo di confidenza della differenza di 95%	
		F	Sign.	t	gl	Sign. (a due code)	Differenza della media	Differenza errore standard	Inferiore	Superiore
Socievolezza	Varianze uguali presunte	,847	,360	1,381	98	,170	2,4000	1,7377	-1,0484	5,8484
	Varianze uguali non presunte			1,381	97,051	,170	2,4000	1,7377	-1,0488	5,8488

Test di Levene

.36 > .05:
possiamo concludere che…
… le varianze delle popolazioni da cui provengono i campioni sono uguali:
non si rifiuta H0: $\boldsymbol{\sigma^2_1 = \sigma^2_2}$

In questo caso vanno interpretati i valori nel riquadro: "**Varianze uguali presunte**"

L'assunzione di omoschedasticità viene rispettata

Il test t per due campioni indipendenti in SPSS

Esempio I

Statistiche gruppo

	Sesso	N	Media	Deviazione std.	Media errore standard
Socievolezza	1	50	32,700	8,2477	1,1664
	2	50	30,300	9,1077	1,2880

Test campioni indipendenti

		Test di Levene per l'eguaglianza delle varianze		Test t per l'eguaglianza delle medie						
									Intervallo di confidenza della differenza di 95%	
		F	Sign.	t	gl	Sign. (a due code)	Differenza della media	Differenza errore standard	Inferiore	Superiore
Socievolezza	Varianze uguali presunte	,847	,360	1,381	98	,170	2,4000	1,7377	-1,0484	5,8484
	Varianze uguali non presunte			1,381	97,051	,170	2,4000	1,7377	-1,0488	5,8488

t di Student (t «empirica»)

gradi di libertà: n1 + n2 -2

.170 > .05
Possiamo concludere che non ci sono differenze di genere nella socievolezza: **non** si rifiuta H0

differenza tra le due medie (32.7 – 30.3) **è il numeratore del test t:**
$\bar{X}_1 - \bar{X}_2$

errore standard della differenza tra due medie **è il denominatore del test t:** $\hat{\sigma}_{\bar{x}1-\bar{x}2}$

$$\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}\left(\frac{n_1+n_2}{n_1 n_2}\right)}$$

Poiché n1 = n2, si può calcolare anche con: $\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$

Il test t per due campioni indipendenti in SPSS

Esempio I

Statistiche gruppo

	Sesso	N	Media	Deviazione std.	Media errore standard
Socievolezza	1	50	32,700	8,2477	1,1664
	2	50	30,300	9,1077	1,2880

Test campioni indipendenti

		Test di Levene per l'eguaglianza delle varianze		Test t per l'eguaglianza delle medie						
									Intervallo di confidenza della differenza di 95%	
		F	Sign.	t	gl	Sign. (a due code)	Differenza della media	Differenza errore standard	Inferiore	Superiore
Socievolezza	Varianze uguali presunte	,847	,360	1,381	98	,170	2,4000	1,7377	-1,0484	5,8484
	Varianze uguali non presunte			1,381	97,051	,170	2,4000	1,7377	-1,0488	5,8488

$$\lim \inf = 2.4 - (1.984 \times 1.7377) = -1.048$$

$$\lim \sup = 2.4 + (1.984 \times 1.7377) = +1.048$$

differenza tra le due medie

t-critico per alfa = .05 (con H_1 bidirezionale) e 98 gdl

errore standard della differenza tra due medie

Intervallo di confidenza della differenza tra le due medie

Se l'intervallo **non** contiene lo zero, **rifiutiamo H0**: con il 95% di probabilità, le due medie sono diverse (provengono da popolazioni differenti)

Il test t per due campioni indipendenti in SPSS

Esempio II

Statistiche gruppo

	Sesso	N	Media	Deviazione std.	Media errore standard
Apertura_M	1	28	19,18	5,063	,957
	2	32	21,31	8,372	1,480

Test campioni indipendenti

		Test di Levene per l'eguaglianza delle varianze		Test t per l'eguaglianza delle medie						
									Intervallo di confidenza della differenza di 95%	
		F	Sign.	t	gl	Sign. (a due code)	Differenza della media	Differenza errore standard	Inferiore	Superiore
Apertura_M	Varianze uguali presunte	6,243	,015	-1,173	58	,245	-2,134	1,819	-5,775	1,507
	Varianze uguali non presunte			-1,211	51,916	,231	-2,134	1,762	-5,670	1,403

.015 < .05
le varianze delle popolazioni da cui provengono i campioni sono diverse (**$\sigma^2_1 \neq \sigma^2_2$**):
si rifiuta H0: $\sigma^2_1 = \sigma^2_2$

L'assunzione di omoschedasticità **non** viene rispettata

I gradi di libertà sono calcolati con una formula complessa (*Satterthwaite approximation*)

Conclusione:
.231 > .05
non vi sono differenze di genere nella variabile Apertura_M

$$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

I limiti dell'intervallo di confidenza cambiano rispetto a quando le varianze sono uguali, perché cambiano l'"errore standard e i gradi di libertà

In questo caso vanno interpretati i valori nel riquadro: "**Varianze uguali non presunte**"

Esercizio 8

Due gruppi di elettori, rispettivamente di destra e di sinistra, hanno compilato un test per la misura dell'autoritarismo

Entrambi i gruppi sono composti da 10 soggetti

Elettori di destra: $\overline{X}_1 = 6.24, \; s_1^2 = 1.94$

Elettori di sinistra: $\overline{X}_2 = 4.64, \; s_2^2 = 2.26$

- **La differenza tra i due gruppi è significativa? In altri termini, i due campioni provengono da popolazioni con medie differenti?**

Utilizza un livello di alfa = .05

Verifica delle ipotesi

- In quale campione la dispersione è più elevata? ____________________

- In questo caso il ricercatore deve utilizzare:

a) test z su un campione b) test t su due campioni indipendenti
c) test t su un campione b) test z su due campioni indipendenti

- L'errore standard della differenza tra le medie è pari a: _________

- Il valore del test statistico è: ___________________________________

- La probabilità di commettere un errore di I tipo è pari a:

a) 95% b) 5% c) non si può dire c) 1%

- La probabilità di commettere un errore di II tipo è pari a:

a) 95% b) 5% c) non si può dire c) 1%

- Il valore critico della distribuzione è pari a: _________

- I gradi di libertà sono: _________

Verifica delle ipotesi

- L'ipotesi alternativa è: a) monodirezionale destra b) monodirezionale sinistra c) bidirezionale

- Cosa si può dire a proposito della forma della distribuzione campionaria della differenza tra due medie? ______________________________

- È possibile sostenere che la differenza osservata tra le medie dei due campioni sia dovuta esclusivamente al caso? ____________________

  ____

  ____

- Utilizza ora un livello di alfa dell'1%. Cosa è possibile concludere?

  ____

- Utilizzando un livello di alfa dell'1% (invece del 5%):
  a) la potenza statistica aumenta
  b) la potenza statistica diminuisce
  c) la potenza statistica rimane invariata
  d) la probabilità di commettere un errore di I tipo aumenta

Esercizio 9

A due campioni, composti rispettivamente da 28 maschi adulti e 28 femmine adulte, è stata somministrata una scala per la misura della coscienziosità

Si sono ottenuti i seguenti risultati:

Maschi: $\overline{X}_1 = 45, \;\; s^2{}_1 = 48.00$

Femmine: $\overline{X}_2 = 49, \;\; s^2{}_2 = 35.93$

È possibile sostenere che vi sono differenze di genere nella coscienziosità?

Utilizza un livello critico di significatività dell'1%

Esercizio 10

Durante un consiglio del corso di laurea triennale sono stati presentati i dati relativi ai voti ottenuti durante l'anno accademico 2012-2013 dagli studenti iscritti a psicologia e salute (PS) e psicologia e processi sociali (PPS). Una t di Student per campioni indipendenti è stata utilizzata per verificare eventuali differenze tra i gruppi nel voto medio, dato un livello di alfa = .05. I risultati sono i seguenti:

Statistiche gruppo

	Corso	N	Media	Deviazione std.	Media errore standard
MEDIA	1 PS	481	25,58001	2,464715	,112381
	2 PPS	477	25,03780	2,488497	,113940

Test campioni indipendenti

		Test di Levene per l'eguaglianza delle varianze		Test t per l'eguaglianza delle medie						
									Intervallo di confidenza della differenza di 95%	
		F	Sign.	t	gl	Sign. (a due code)	Differenza della media	Differenza errore standard	Inferiore	Superiore
MEDIA	Varianze uguali presunte	,081	,776	3,388	956	,001	,542214	,160031	,228161	,856267
	Varianze uguali non presunte			3,388	955,692	,001	,542214	,160037	,228149	,856280

1. Formula ipotesi nulla e ipotesi alternativa: H0: ________; H1: _________;
2. L'assunzione di omoschedasticità è rispettata?
3. L'errore standard della differenza tra le medie è pari a:
4. Il t test è significativo? Possiamo rifiutare H0? Cosa è possibile concludere?
5. Commenta il risultato relativo all'intervallo di confidenza:
6. Calcola l'ampiezza dell'effetto: cosa si può concludere?

Esercizio 11

Lo stesso confronto è stato effettuato in relazione al numero di esami superati al termine del secondo anno di corso. I risultati sono i seguenti:

Statistiche gruppo

	Corso	N	Media	Deviazione std.	Media errore standard
ESAMI	1 PS	481	12,193	8,398	,383
	2 PPS	477	11,797	8,157	,374

Test campioni indipendenti

		Test di Levene per l'eguaglianza delle varianze		Test t per l'eguaglianza delle medie						
									Intervallo di confidenza della differenza di 95%	
		F	Sign.	t	gl	Sign. (a due code)	Differenza della media	Differenza errore standard	Inferiore	Superiore
ESAMI	Varianze uguali presunte	,475	,491	?	?	,459	?	,535	-,653	1,447
	Varianze uguali non presunte					,459		,535	-,653	1,446

1. Completa la tabella, inserendo i valori mancanti al posto dei punti interrogativi

2. Le varianze delle popolazioni da cui provengono i gruppi sono uguali?

3. Cosa è possibile concludere rispetto alla differenza tra le medie dei due gruppi nel numero di esami sostenuti?

Esercizio 12

Un test per la misura del comportamento prosociale è stato somministrato a due campioni, composti rispettivamente da 40 maschi e 36 femmine. Il test t di Student per campioni indipendenti ha evidenziato i seguenti risultati:

Statistiche gruppo

	Sesso	N	Media	Deviazione std.	Media errore standard
Prosociale	1 Maschi	40	4,03	2,106	,333
	2 Femmine	36	5,00	1,434	,239

Test campioni indipendenti

		Test di Levene per l'eguaglianza delle varianze		Test t per l'eguaglianza delle medie						
									Intervallo di confidenza della differenza di 95%	
		F	Sign.	t	gl	Sign. (a due code)	Differenza della media	Differenza errore standard	Inferiore	Superiore
Prosociale	Varianze uguali presunte	10,617	,002	-2,333	74	,022	-,975	,418	-1,808	-,142
	Varianze uguali non presunte			-2,379	69,105	,020	-,975	,410	-1,793	-,157

1. Interpreta i risultati: cosa è possibile concludere?