Elementi di Psicometria con Laboratorio di SPSS 1

11-Verifica di ipotesi fra due medie
(vers. 1.2, 10 aprile 2017)
versione per stampa

Germano Rossi[1]
germano.rossi@unimib.it

[1]Dipartimento di Psicologia, Università di Milano-Bicocca

10 aprile 2017

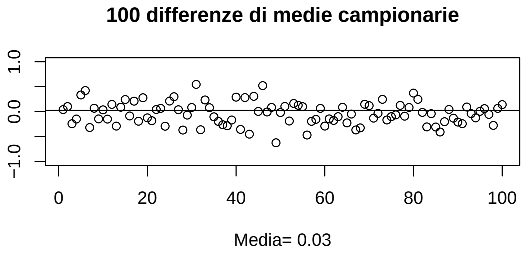
Differenza di due medie

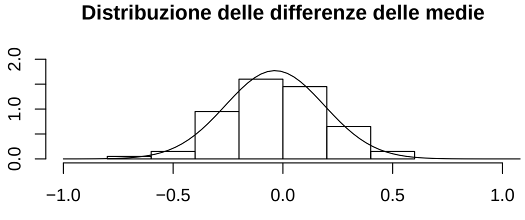
- Finora abbiamo visto la possibilità di verificare
  - la media di un gruppo rispetto ad una popolazione di cui conosciamo i parametri (μ e σ) [situazione solo teorica]
  - la media di un gruppo rispetto ad una popolazione di cui conosciamo solo il parametro della media (μ)
- Se non conoscessimo neppure la media della popolazione (situazione tipica in psicologia), non potremmo fare nessun tipo di inferenza
- Tuttavia, la maggior parte delle volte, noi non siamo interessati a sapere se un certo campione appartiene ad una certa popolazione, ma a sapere se due campioni provengono dalla stessa popolazione o da due popolazioni con parametri diversi
- In questo caso, l'ipotesi sarebbe che la differenza delle medie sia nulla ($\overline{X}_1 - \overline{X}_2 = 0$)

Differenza di due medie

- Con lo stesso ragionamento fatto per la distribuzione campionaria delle medie, si può fare la *distribuzione campionaria della differenza di due medie*.
- Se 2 campioni vengono estratti dalla stessa popolazione, la loro media dovrebbe tendere alla media della popolazione, qualunque essa sia.
- Se facciamo la differenza fra le due medie (ed entrambe tendono alla media della popolazione), la loro differenza *dovrebbe* tendere ad essere uguale a 0 ($\mu_1 - \mu_2 = 0$). *Dovrebbe*, ma non sempre è così.
- Tuttavia se estraiamo molte coppie di medie, la distribuzione della differenza di queste medie graviterà attorno allo 0.
- La stessa cosa dovrebbe capitare se i due campioni vengono da due popolazioni diverse che hanno, però, la stessa media ($\mu_1 = \mu_2$)

Differenza di due medie





Differenza di due medie

- Anche in questo caso abbiamo che per N $\geq$ 30, la distribuzione campionaria della differenza delle medie tenderà a distribuirsi normalmente
- e anche in questo caso potremo calcolare un *errore standard della differenza delle medie*
- e anche in questo caso, un valore piccolo dell'errore standard indica una piccola oscillazione delle differenze campionarie attorno allo 0
- e un valore grande indica una grossa oscillazione attorno allo 0
- Anche in questo caso, se potessimo estrarre un numero infinito di coppie di campioni, potremmo calcolare l'errore standard esatto
- Non potendo farlo, lo stimiamo a partire dalle deviazioni standard dei due campioni

Differenza di due medie

- Si pongono adesso due possibilità:
  - la varianza nelle due popolazioni è uguale
  - la varianza nelle due popolazioni è diversa
- se la varianza è uguale potremmo semplicemente sommare le singole varianze (in particolare se le numerosità dei due campioni è uguale)
- più spesso non possiamo ipotizzare che le varianze siano uguali e, ancora più spesso, i due campioni non hanno la stessa numerosità

Differenza di due medie

- Usiamo quindi una varianza combinata

$$s^2_{comb} = \frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{N_1 + N_2 - 2}$$

- e la stima dell'errore standard diventa

$$s_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{s^2_{comb}}{N_1} + \frac{s^2_{comb}}{N_2}}$$

- Mettendo assieme le due formule abbiamo

$$s_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{N_1 + N_2 - 2}\left(\frac{1}{N_1} + \frac{1}{N_2}\right)}$$

Differenza di due medie: test

- Il test sulla differenza delle medie si baserà sulla formula di z, ma produrrà una statistica t

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{s_{\overline{X}_1 - \overline{X}_2}}$$

- Anche se è teoricamente possibile ipotizzare che la differenza di due medie corrisponda ad un certo valore (ad es. 5), la maggior parte delle volte si ipotizzerà che la differenza delle medie sia nulla; in questo caso $\mu_1 - \mu_2 = 0$ e la formula si riduce alla sola differenze delle medie dei campioni
- In ogni caso, la statistica t calcolata avrà gradi di libertà pari a

$$(N_1 - 1) + (N_2 - 1) = N_1 + N_2 - 2$$

Test della differenza delle medie

- Conosciamo: M_1, M_2, s_1 e s_2 allora poniamo:
- $H_0 : \mu_1 = \mu_2$
- $H_1 : \mu_1 \neq \mu_2$ (oppure > oppure <)
- scegliamo α e troviamo il t critico
- applichiamo la formula

$$t_{M_1 - M_2} = \frac{M_1 - M_2}{\sqrt{\frac{(N_1-1)s_1^2+(N_2-1)s_2^2}{N_1+N_2-2}\left(\frac{1}{N_1}+\frac{1}{N_2}\right)}}$$

con $gdl = N_1 + N_2 - 2$

- se $t_{M_1 - M_2} < t_c$ (in valore assoluto) accetto H_0
- se maggiore (in valore assoluto), rifiuto H_0 (e accetto quindi H_1)

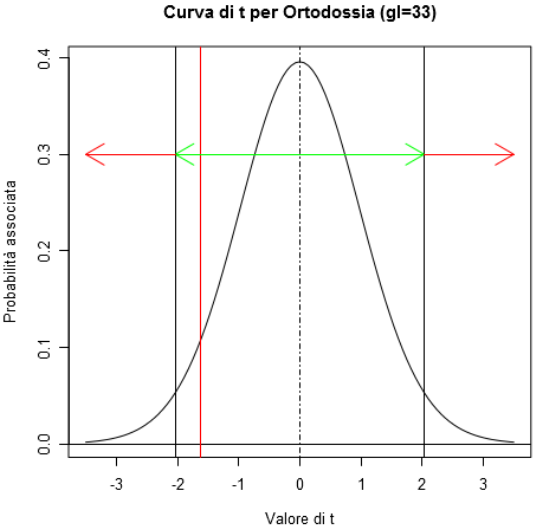
Test della differenza delle medie

- Abbiamo misurato l'ortodossia in un campione di Testimoni di Geova e ci chiediamo se vi è differenza fra maschi (N=23) e femmine (N=12)
- $M_m = 14.17$, $M_f = 15.5$, $s_m = 2.27$ e $s_f = 2.32$ allora:
- facciamo le ipotesi nulla e alternativa $H_1 : \mu_m \neq \mu_f$
- con $\alpha = .05$ bidirezionale il t critico con $gdl = 23 + 12 - 2 = 33$ è $t_c = 2.04$
- applichiamo la formula

$$t_{M_1 - M_2} = \frac{14.17 - 15.5}{\sqrt{\left(\frac{2.27^2(23-1)+2.32^2(12-1)}{23+12-2}\right)\left(\frac{1}{23}+\frac{1}{12}\right)}} = -1.63$$

- siccome $t_{M_1-M_2} < t_c$ accetto H_0

Test della differenza delle medie



- $t = -1.63$ (rosso)
- α =5% (bi)
  $t_c = |2.04|$ (nero)
- Accetto H_0 perché
- $|t| < |t_c|$
- $|t|$ cade nell'area di accettazione di H_0 (verde)

Due approcci all'inferenza puntuale

1. Uso delle tavole per trovare un valore critico
   - Si applicano le formule e si trova la statistica (`z` o `t`)
   - Si decide un livello α e si cerca un valore critico sulle tavole (eventualmente usando i gdl)
   - Confrontando la statistica trovata con il valore critico accettiamo o rifiutiamo l'ipotesi nulla

   $v_t < v_c \Rightarrow$ accetto H_0

2. Calcolo diretto della probabilità
   - Si applicano le formule e si trova la statistica (`z` o `t`)
   - Usando la distribuzione di probabilità di quella statistica si calcola direttamente la probabilità associata
   - Confrontando la probabilità trovata con l'α scelto, accettiamo o rifiutiamo l'ipotesi nulla

   $P(v_t) > \alpha \Rightarrow$ accetto H_0

Due approcci all'inferenza puntuale

Per trovare un valore critico

- Con Excel, `zc=INV.NORM.ST(1-alfa/2)` `zc=INV.NORM.ST(1-alfa)` bi e mono
- `tc=INV.T(alfa;gl)`
- Con R, `zc=qnorm(1-alfa)` (bi) o `zc=qnorm(1-alfa/2)` (mono)
- `tc=qt(1-alfa, gl)` (bi) o `tc=qt(1-alfa/2, gl)`

Calcolo diretto della probabilità

- Con Excel, `P(z)= 1-DISTRIB.NORM.ST(z)`
- `tc=DISTRIB.T(t;gl;2)` `tc=DISTRIB.T(t;gl;1)` bi e mono
- Con R, `P(z)=pnorm(z)`
- `P(t)=1-pt(t,gl)`

zc= z critico; tc=t critico; alfa = valore di α; gl= gradi di libertà
z=punti z

Stima intervallare

- Anche per la differenza delle medie, possiamo calcolare un'intervallo di confidenza, sempre usando il valore critico di t al 5% o all'1% per avere intervalli di fiducia pari al 95% o al 99%

$$(\overline{X}_1 - \overline{X}_2) \pm t_c s_{\overline{X}_1 - \overline{X}_2}]$$

- Applicandolo all'esempio dei Testimoni di Geova:

$$(\overline{X}_1 - \overline{X}_2) = -1.33 \quad t_{95\%} = 2.04 \quad s_{\overline{X}_1 - \overline{X}_2} = 2.87$$

$$-1.33 + 2.04 * 2.87 \text{ e } -1.33 - 2.04 * 2.87$$

- ovvero l'intervallo di confidenza oscilla fra -5.99 e 3.34
- Poiché l'intervallo include anche il valore 0 ($H_0 : \mu_1 - \mu_2 = 0$) corrispondente alla nostra ipotesi nulla, dobbiamo accettarla come vera.

Assunti del t-test

1. Le misure della variabile dipendente di ciascun gruppo
   a) sono indipendenti tra di loro
   b) sono indipendenti dall'altro gruppo
2. La variabile dipendente si distribuisce normalmente
3. Le varianze dei due gruppi sono uguali

- La condizione 2 può essere ignorata, perché il test t non è molto sensibile alle violazioni di normalità
- La condizione 3 può essere ignorata se i due campioni hanno uguale numerosità; il test t, in questo caso, non distorce troppo e si può usare la distribuzione di t senza problemi

Assunti del t-test

- Se i due campioni non hanno uguale numerosità (accettabile fino a 1.5 circa), si pongono diverse condizioni:
  1. la soluzione migliore è quella di ridurre il campione più numeroso ed equiparare le numerosità (basta fare una selezione casuale del campione)
  2. provare a sottoporre le variabile a trasformazioni che mantengono la linearità, ma cambiano la distribuzione (le vedremo nel *Trattamento dei dati*)
  3. si può usare la stima di varianza separata, ovvero il test t diventa

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

Inferenza con programmi statistici

- Nell'uso normale, sono i computer a fare i conti per il confronto di medie: **t di Student (o T-Test)**
- stessa variabile su due campioni indipendenti. Le formule usate sono quelle indicate qui e prevedono
  - una formula che ipotizza che i campioni abbiano varianza uguale (varianza combinata)
  - un'altra formula che ipotizza che abbiano varianza diversa (Varianze separate)
- Viene applicato un test per l'*omogeneità delle varianze* (test di Levene) e in base ai risultati si sceglie il test appropriato
- Il test di Levene è basato sul rapporto delle due varianze: se sono simili tenderà a 1; al test è associato un valore di probabilità

Test campioni indipendenti

		Test di Levene per l'eguaglianza delle varianze		Test t per l'eguaglianza delle medie				
		F	Sign.	t	gl	Sign. (a due code)	Differenza della media	Differenza errore standard
Ortodossia	Varianze uguali presunte	1,326	,258	-1,630	33	,113	-1,32609	,81371
	Varianze uguali non presunte			-1,619	22,019	,120	-1,32609	,81908

SPSS: 2 campioni indipendenti

- `Analizza | Confronta medie | Test T: campioni indipendenti`
- Trascinare una o più variabile dipendente (quantitativa) in `Variabili oggetto del test`
- Trascinare una variabile qualitativa (con 2 o più valori possibili) in `Variabile di raggruppamento`
- Premere `Definisci gruppi` e inserire i due valori della qualitativa da usare. Quindi `Continua`
- Infine `OK`

Esempio in SPSS

Medie e deviazioni standard suddivise per genere

Statistiche gruppo

	Genere	N	Media	Deviazione std.	Media errore standard
Intrins	1	158	23,0805	5,68417	,45221
	2	181	23,5057	5,58231	,41493
Estr_soc	1	158	5,8165	3,09888	,24653
	2	181	6,0829	2,97000	,22076
Estr_pers	1	158	9,4494	3,54327	,28189
	2	181	10,8619	2,97914	,22144

- Dalle medie possiamo vedere (“a naso”) che non ci sono grosse differenze fra i sessi nelle prime due variabili.
- Forse, nell’ultima c’è una differenza.

Esempio in SPSS

Test campioni indipendenti

		Test di Levene per l'eguaglianza delle varianze					
		F	Sign.	t	gl	Sign. (a due code)	D
Intrins	Presumi varianze uguali	,230	,632	-,694	337	,488	
	Non presumere varianze uguali			-,693	329,150	,489	
Estr_soc	Presumi varianze uguali	,923	,337	-,807	337	,420	
	Non presumere varianze uguali			-,805	326,577	,421	
Estr_pers	Presumi varianze uguali	9,477	,002	-3,987	337	,000	
	Non presumere varianze uguali			-3,940	308,190	,000	

- Il test di Levene ci dice se i due gruppi hanno la stessa varianza (F vicina a 1, con `Sig` superiore ad α) oppure no (F molto grande, con `Sig` inferiore ad α).
- A questo punto possiamo leggere e interpretare l'esatto `t` e confrontare la probabilità associata (`Sig`) a quel `t`, con quei gradi di libertà (`df`) direttamente con l'α scelto.

Esempio in SPSS

Test t per l'eguaglianza delle medie				
Sign. (a due code)	Differenza media	Differenza errore standard	Intervallo di confidenza della differenza 95%	
			Inferiore	Superiore
,488	-,42519	,61297	-1,63092	,78054
,489	-,42519	,61373	-1,63251	,78213
,420	-,26642	,32997	-,91548	,38265
,421	-,26642	,33093	-,91744	,38460
,000	-1,41251	,35430	-2,10943	-,71560
,000	-1,41251	,35846	-2,11785	-,70717

- Gli intervalli di confidenza ci dicono da quali popolazioni può essere estratto un campione come il nostro sotto la condizione dell'ipotesi nulla.
- Se gli intervalli includono lo 0 allora accettiamo l'ipotesi nulla ($H_0 = 0$), altrimenti accettiamo quella alternativa.
- Stima puntuale e stima intervallare devono coincidere.

Riportare i risultati

Esempio

Per quanto riguarda l'orientamento religioso, per la religiosità intrinseca e l'estrinseca sociale non sembrano esserci effetti del genere, in quanto è abbastanza probabile che le medie siano uguali fra loro. Al contrario, per quanto riguarda la religiosità estrinseca personale, il gruppo femminile mostra una media (10.86) più alta del gruppo maschile (9.45). La differenza è statisticamente significativa $t(308,19)=-3.94$, $p < .001$.

Applicabilità

- Per confrontare la media di una variabile fra due gruppi
- Cosa si usa
  - 1 variabile qualitativa (indipendente) che viene usata per suddividere il campione in 2 gruppi
  - 1 variabile quantitativa (dipendente) su cui vengono calcolate le media (una per ciascun gruppo)

Effect size (Ampiezza dell'effetto)

- Tramite il test della differenza delle medie (t-test) abbiamo visto se due medie sono diverse ovvero se i campioni su cui le medie sono state calcolate provengono da popolazioni con parametri statistici uguali o diversi
- questa informazione però è solamente vero/falso: o c'è differenza oppure no
- e l'eventuale differenza delle medie va interpretata come relativa
- se la religiosità estrinseca personale è diversa in base al genere e vediamo che la media dei maschi è minore di quella delle femmine non possiamo interpretare nulla di più della diversità.

Effect size (Ampiezza dell'effetto)

- L'effect size indica invece di quanto due medie differiscono fra loro e se l'ampiezza dell'effetto trovato è piccolo, medio o grande
- e si "potrebbe" calcolare con una formula (teorica) simile ad una z, proposta da Cohen (formula generica):

$$d = \frac{\mu_1 - \mu_2}{\sigma}$$

- Una possibile formula per il t-test (per campioni indipendenti) è stata proposta da Hedge, usando la varianza combinata:

$$g = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{(N_1-1)s_1^2+(N_2-1)s_2^2}{N_1+N_2-2}}}$$

Effect size (Ampiezza dell'effetto)

- Dal momento che sia g sia t usano la varianza combinata, conoscendo le numerosità e il valore di t, si può calcolare g anche come:

$$g = \frac{t}{\sqrt{\frac{N_1 N_2}{N_1 + N_2}}} = t\sqrt{\frac{N_1 + N_2}{N_1 N_2}}$$

- Se i due gruppi hanno la stessa numerosità si può usare una formula più veloce:

$$g = \frac{t}{\sqrt{\frac{N}{2}}} = t\sqrt{\frac{2}{N}}$$

Effect size (Ampiezza dell'effetto)

- Ovviamente, usando le forme inverse, possiamo calcolare t usando g
- Se i due gruppi hanno la stessa numerosità:

$$t = g\sqrt{\frac{N}{2}}$$

- Se i due gruppi non hanno la stessa numerosità:

$$t = g\sqrt{\frac{N_1 N_2}{N_1 + N_2}}$$

- Ma è più facile avere t che non g, in quanto molti software calcolano t, ma non g

Effect size (Ampiezza dell'effetto)

- g (di Hedge) o **d** (di Cohen) dovrebbe oscillare fra 0 e 1, ma spesso supera il valore 1
- Cohen propose una interpretazione generica di **d** che viene tutt'ora utilizzata
- $\mathbf{d} = .20$, effetto piccolo
- $\mathbf{d} = .50$, effetto medio
- $\mathbf{d} = .80$, effetto grande
- in letteratura si trovano **d** superiori a 1 che vengono generalmente considerati giganti (*huge*)

Esempio di uso di **d**

	Italiano		Originale		Test t	
	M	s	M	s	p	d
Fondamentalismo	75,35	32,03	84,6	33	.003	-.28
Intrinseco	22,25	4,57	37,2	5,8	<.001	-2.64
Estrinseco	15,94	5,29	25,6	5,7	<.001	-1.71

Esempio

“Confrontando il nostro campione con quello originale, osserviamo che ci sono alcune differenze. I dati sembrano evidenziare che la popolazione americana è più fondamentalista di quella italiana, anche se osservando l’*effect size* possiamo affermare che la differenza non è così profonda come potrebbe apparentemente sembrare. La situazione cambia per quanto riguarda l’orientamento religioso: ci sono differenze più grandi tra le 2 popolazioni sia per quanto riguarda la religiosità intrinseca che per quella estrinseca. Infatti i valori della **d** di Cohen sono decisamente giganti e testimoniano una sostanziale differenza tra le due popolazioni sulla base di probabili differenze culturali, storiche e sociali che caratterizzano i differenti paesi.”

Esempio di uso di **d**

	Primo		Secondo		Test t	
	M	s	M	s	p	d
Fondamentalismo	75,35	32,03	81,7	31,59	.034	.20
Intrinseco	21,37	5,41	23,4	5,66	<.001	-.36
Estrinseco	15,94	5,29	16,1	5,21	n.s.	
Estrinseco Personale	9,45	3,12	10,21	3,32	.013	-.23
Estrinseco Sociale	6,45	3,13	5,94	3,01	n.s.	.17

Esempio

"I risultati del nostro campione sono stati confrontati con quelli di un campione italiano precedente. Attraverso il test t osserviamo che vi sono tre variabili statisticamente diverse nei due campioni. Tuttavia le differenze riscontrate sono minime, infatti l'*effect size* più elevato è di solo .36, quindi l'ampiezza degli effetti sono tutte piccole. Questo ci permette affermare che potrebbe esiste comunque una somiglianza tra i 2 campioni italiani, la cui diversità potrebbe anche essere casuale."

Confronti dipendenti o appaiati

- Alcune volte può capitare di confrontare fra loro due variabili misurate sullo stesso campione (ad es. situazioni prima/dopo)
- Oppure capita di lavorare con campioni appaiati (due campioni in cui i casi sono simili a coppie [moglie-marito] oppure sono stati appaiati a posteriori)
- Si può usare una versione del t-test che si basa sulle differenze fra i punteggi dei singoli casi
- La differenza sostanziale è che si usa la media delle differenze

$$\overline{D} = X_1 - X_2$$

anziché la differenza delle medie

$$\overline{X_1} - \overline{X_2}$$

- si fa poi riferimento alla **distribuzione campionaria della media delle differenze**

Confronti dipendenti o appaiati

- La distribuzione campionaria della media delle differenze si approssima a una t con gl=N-1 dove N equivale al numero di coppie appaiate
- per la statistica si usa $\overline{D} = X_{i1} - X_{i2}$
- l'ipotesi nulla è che $\mu_{\overline{D}} = 0$
- la statistica puntuale si calcola con

$$t = \frac{\overline{D} - \mu_{\overline{D}}}{\sqrt{\frac{s^2_D}{N}}}$$

con $df = N - 1$

- quella intervallare con

$$\overline{D} - ts_{\overline{D}} < \mu_{\overline{D}} < \overline{D} + ts_{\overline{D}}$$

Procedimento manuale

- Usando i dati del file `Welk_c11eser6.sav` (N=14) facciamo la differenza fra X e Y: $d = X - Y$
- Facciamo la media dei 14 valori d trovati: $3,3571$
- Calcoliamo la deviazione standard: $6,912721$
- Calcoliamo `t`:

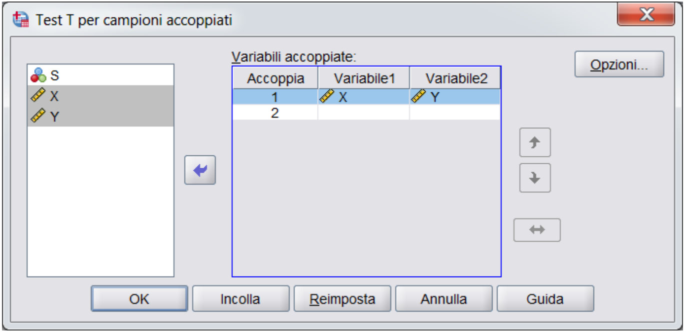
$$t = \frac{\bar{d}}{\sqrt{\frac{s_d^2}{N}}} = \frac{3,3571}{\sqrt{\frac{6{,}912721^2}{14}}} = 1,817$$

- Cerchiamo sulla Tavola B (p.474) $N - 1 = 14 - 1 = 13$ gradi di libertà, il valore critico al 5% bidirezionale (pari a ± 2.16)
- Dal momento che 1.817 è minore di 2.16, il nostro `t` cade nell'area del 95% e non in una delle due code (del 5%)
- Il nostro `t` non è significativo

SPSS: Campioni appaiati

Usando `Welk_c11eser6.sav` in SPSS

- `Analizza | Confronta medie | Test T: campioni appaiati`
- In `Variabili appaiate` bisogna inserire due variabili quantitative che verranno confrontate fra loro a coppie
- Infine `OK`



SPSS: Campioni appaiati

Usando `Welk_c1leser6.sav` in SPSS

Test campioni accoppiati

	Differenze accoppiate							
	Media	Deviazione std.	Media errore standard	Intervallo di confidenza della differenza 95% Inferiore	Intervallo di confidenza della differenza 95% Superiore	t	gl	Sign. (a due code)
Coppia 1 X - Y	3,357	6,913	1,848	-,634	7,348	1,817	13	,092

- La stima puntuale ci dà lo stesso t calcolato a mano (1.817)
- La probabilità associata alla stima puntuale è p=.092 (non significativo)
- La stima intervallare ci da gli estremi delle popolazioni possibili
- Il valore 0 è incluso nell'intervallo (non significativo)

SPSS: Campioni appaiati

Statistiche campioni accoppiati

		Media	N	Deviazione std.	Media errore standard
Coppia 1	ans_b	18,430	100	4,3304	,4330
	ans_p	19,580	100	6,4325	,6433
Coppia 2	ans_b	18,430	100	4,3304	,4330
	ans_d	19,400	100	4,7694	,4769

- Per ogni coppia di variabili c'è una sola riga di statistiche

Test campioni accoppiati

		Differenze accoppiate					t	gl	Sign. (a due code)
		Media	Deviazione std.	Media errore standard	Intervallo di confidenza della differenza di 95% Inferiore	Superiore			
Coppia 1	ans_b - ans_p	-1,1500	5,3737	,5374	-2,2163	-,0837	-2,140	99	,035
Coppia 2	ans_b - ans_d	-,9700	3,5175	,3518	-1,6679	-,2721	-2,758	99	,007

Applicabilità

- Per confrontare la media di due variabili misurate in uno stesso gruppo
- Cosa si usa
  - 2 variabili quantitative (entrambi dipendenti) su cui vengono calcolate le media (una per ciascun gruppo)
  - il motivo per cui la variabile è stata misurata due volte, è la variabile indipendente

Più confronti di medie

- Se devo confrontare 3 medie?
  Faccio 3 confronti: M1 con M2, poi M1 con M3 e quindi M2 con M3
  Equivale a una una combinazione di 3 elementi presi 2 a 2 senza ripetizione.
- Con 4 medie?
  Faccio 6 confronti!
- Con 5 medie?
  Faccio 10 confronti
- Ma...

Più confronti di medie

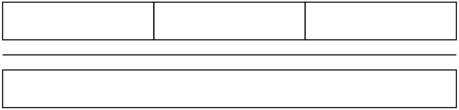
- Se per ogni confronto, ho una possibilità di sbagliare pari al livello α che scelgo, per 10 confronti avrò una possibilità pari a 10α
- Ovvero, se $\alpha = .05, 10 \times 0.05 = 0.50$
- Ciò significherebbe che su 10 confronti almeno la metà potrebbero essere inaffidabili
- È chiaro che non posso correre un rischio così elevato.
- Ho due possibili soluzioni
  - Analisi della varianza (preferibile)
  - Criterio di Bonferroni: dividere l'alfa per il numero di confronti e usare il risultato come nuovo alfa (es. $\alpha = .05$ con 10 confronti, $\alpha/10 \Rightarrow .05/10 = .005$

Analisi della varianza

- Il test dell'*Analisi della varianza*, risolve questo problema (anche conosciuto con i nomi Anova o AOV, da *Analysis of variance*)
- Serve per confrontare fra loro tre o più gruppi e decidere se vengono dalla stessa popolazione di riferimento
- Le situazioni più comuni sono:
  - 1 variabile dipendente (I/R) suddivisa in base ad 1 variabile indipendente che ha 3 o più categorie (N/O): Anova ad 1 via (o fattore)
  - 1 variabile dipendente (I/R) suddivisa in base a 2 o più variabili indipendenti (N/O): Anova a 2 o più vie (o fattori)

Analisi della varianza

- Con sintesi semplicistica, possiamo dire che l'AOV confronta la varianza calcolata in due modi diversi:
- la varianza tra i singoli gruppi (ogni gruppo è considerato separato)
- la varianza entro i singoli gruppi (un solo gruppo totale ottenuto ignorando i singoli gruppi)
- la statistica che risulta è il rapporto fra le due varianze



= F

Analisi della varianza

- La statistica dell'anova è indicata con F (da Fischer che l'ha pensata e studiata)
- Se le due varianze sono uguali, F si avvicinerà ad 1
- Se le due varianze sono diverse, F sarà tanto più grande quanto maggiore è la diversità
- I gradi di libertà sono 2 (uno per ogni varianza utilizzata) e dipendono dai gruppi confrontati ($k-1$) e dalle loro numerosità (in genere $N_x - 1$, in base ai confronti)

Tipi di anova

- L'anova ad 1 via (ad un fattore) implica una variabile quantitativa e una variabile categoriale con almeno 3 categorie (3 o più grupppi)
- L'anova a due vie (a 2 fattori) implica una variabile quantitativa e due variabili categoriali ciascuna con almeno 2 categorie (4 o più grupppi)
- Ci possono essere più variabili categoriale (in base al disegno sperimentale) e l'unico limite è la capacità di interpretazione dei risultati
- Esistono forme di anova per misure ripetute

Esempio SPSS: anova 1 via

Orientamento politico	Scala di ortodossia RFS N	Mean	Std. Deviation
Nessuno	79	5,63	8,11
Sinistra	120	-1,26	9,34
Centro	104	6,38	8,21
Destra	34	5,24	10,14
Total	337	3,37	9,43

4.032,535 / 3 = 1.344,178
4.032,535 / 3 = 1.344,178
1.344,178 / 77,651 = 17,310

	SQ=Somme quadrati	df	MQ=Medie quadrati	F	Sig.
Fra	4.032,535	3	1.344,178	17,310	0,000
Errore	25.857,839	333	77,651		
Total	29.890,374	336			

SQ è il numeratore della varianza; MQ=SQ/df ed è la varianza; F è il rapporto fra le due MQ

Esempio SPSS: anova e post-hoc

- Dal momento che l'anova è statisticamente significativa (ho rifiutato H_0) almeno un gruppo è statisticamente diverso dagli altri
- per sapere qual è, usa l'analisi a *post-hoc* che confronta ciascun gruppo con tutti gli altri
- uno dei metodi a *post-hoc* crea gruppi statisticamente omogenei

	Orientamento politico	N	Subset for alpha = .05 1	2
Student-Newman-Keuls	Sinistra	120	-1,2583	
	Destra	34		5,2353
	Nessuno	79		5,6329
	Centro	104		6,375
	Sig.		1	0,736

Esempio SPSS: anova 1 via

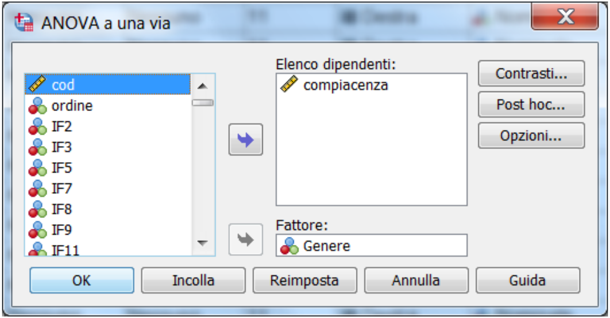
Riportare i dati

Esempio

"I risultati mostrano che la variabile indipendente *Orientamento politico* influisce significativamente sulla variabile dipendente *Ortodossia religiosa*, $F(3, 333) = 17,31$, $p < 0.001$. Dall'osservazione delle medie dei gruppi emerge che il gruppo di sinistra presenta punteggi di ortodossia più bassi (M = -1,26) rispetto a quelli relativi agli altri gruppi di destra (M = 5,24), centro (M = 6,38) e senza orientamento politico (M=5,63). L'analisi a post-hoc (con SNK) dimostra che, in effetti, i gruppi centro, destra e senza orientamento sono omogenei fra loro."

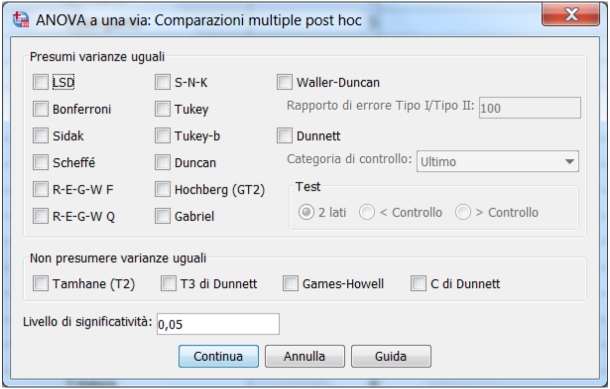
Anova 1 via in SPSS

- `Analizza | Confronta medie | Anova a 1 via`
- Inserire in `Elenco dipendenti` la/le variabili quantitative [comunque un t-test separato]
- Inserire in `Fattore` la variabile dipendente
- Infine `OK`



Anova 1 via in SPSS

- Con il bottone `Post-Hoc` si possono scegliere i metodi di confronto a post-hoc
- Generalmente si fa solo per risultati significativi



Anova e t-test

- Se si fa un'anova a 1 via con una variabile indipendente che ha solo 2 gruppi
- si vede che F è il quadrato di t
- e t è la radice quadrata di F

	t	gl	Sign.
compiacenza	-,243	58	,809

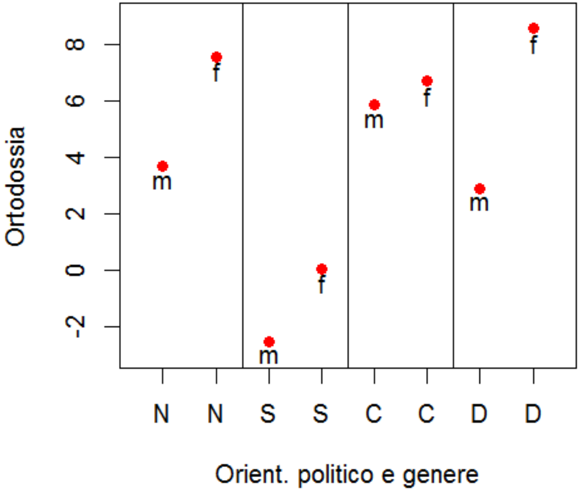
$$t = -0,243 = \sqrt{0,059}$$
$$F = 0,059 = -0,243^2$$

	Somma quadrati	gl	Media quadratica	F	Sign.
Tra gruppi	8,929	1	8,929	,059	,809
Entro i gruppi	8787,804	58	151,514		
Totale	8796,733	59			

Esempio SPSS: Anova a 2 vie

		Scala di ortodossia		
Orientamento politico	Genere	Mean	Std. Deviation	N
Nessuno	Maschi	3,67	9,07	39
	Femmine	7,55	6,63	40
	Total	5,63	8,11	79
Sinistra	Maschi	-2,55	9,56	60
	Femmine	0,03	9,01	60
	Total	-1,26	9,34	120
Centro	Maschi	5,85	8,62	40
	Femmine	6,70	8,00	64
	Total	6,37	8,21	104
Destra	Maschi	2,90	11,47	20
	Femmine	8,57	6,96	14
	Total	5,24	10,14	34
Total	Maschi	1,77	10,02	159
	Femmine	4,79	8,65	178
	Total	3,37	9,43	337

Esempio SPSS: Anova a 2 vie



N=Nessun orientamento; S=Sinistra; C=Centro; D=Destra

m=maschi; f=femmine

Esempio SPSS: Anova a 2 vie

Source Source	Type III Sum of Sq.	df df	Mean Square	F	Sig.
Corrected Model	4813,336	7	687,619	9,021	0,000
Intercept	4352,743	1	4352,743	57,106	0,000
Fasce pol	3981,851	3	1327,284	17,413	0,000
Genere	685,986	1	685,986	9,000	0,003
Fasce pol * Genere	184,718	3	61,573	0,808	0,49
Error	25077,04	329	76,222		
Total	33713	337			
Corrected Total	29890,37	336			