T-test per dati appaiati

Esempio

La guida in stato di ebbrezza è una delle cause principali degli incidenti automobilistici. Le interviste ai guidatori ubriachi che sono stati coinvolti in incidenti e sopravvissuti hanno rivelato che uno dei problemi principali è che i conducenti non si rendono conto del loro stato alterato, pensando "Ho bevuto solo 1-2 drink ... posso guidare".
È stato scelto un campione di 20 conducenti e i loro tempi di reazione in un percorso a ostacoli sono stati misurati prima e dopo aver bevuto due birre. Lo scopo di questo studio era di verificare se lo stato di attenzione dei conducenti è alterato dopo aver bevuto due birre.

y_B: reaction time

y_A: reaction time

Come valutereste se il tempo di reazione si è modificato dopo le due birre?

y_B Before	y_A After
6.25	6.85
2.96	4.78
4.95	5.57
3.94	4.01
4.85	5.91
4.81	5.34
6.6	6.09
5.33	5.84
5.19	4.19
4.88	5.75
5.75	6.25
5.26	7.23
3.16	4.55
6.65	6.42
5.49	5.25
4.05	5.59
4.42	3.96
4.99	5.93
5.01	6.03
4.69	3.72

$\bar{y}_B = 4.962$ $\bar{y}_A = 5.463$

Campioni dipendenti vs indipendenti

- Due campioni sono **indipendenti** se i valori campionari di una popolazione non sono associati - o in qualche modo appaiati/abbinati naturalmente- ai valori campionari dell'altro gruppo.
- Due campioni sono **dipendenti** (o sono costituiti da coppie appaiate) se i valori campionari sono in qualche modo abbinati/appaiati. (Vale a dire, ogni coppia di valori campionari consiste di due misurazioni dello stesso soggetto, come prima/dopo, oppure ogni coppia di valori consiste di coppie appaiate, come marito/moglie, dove l'abbinamento si basa su alcune relazioni significative.)

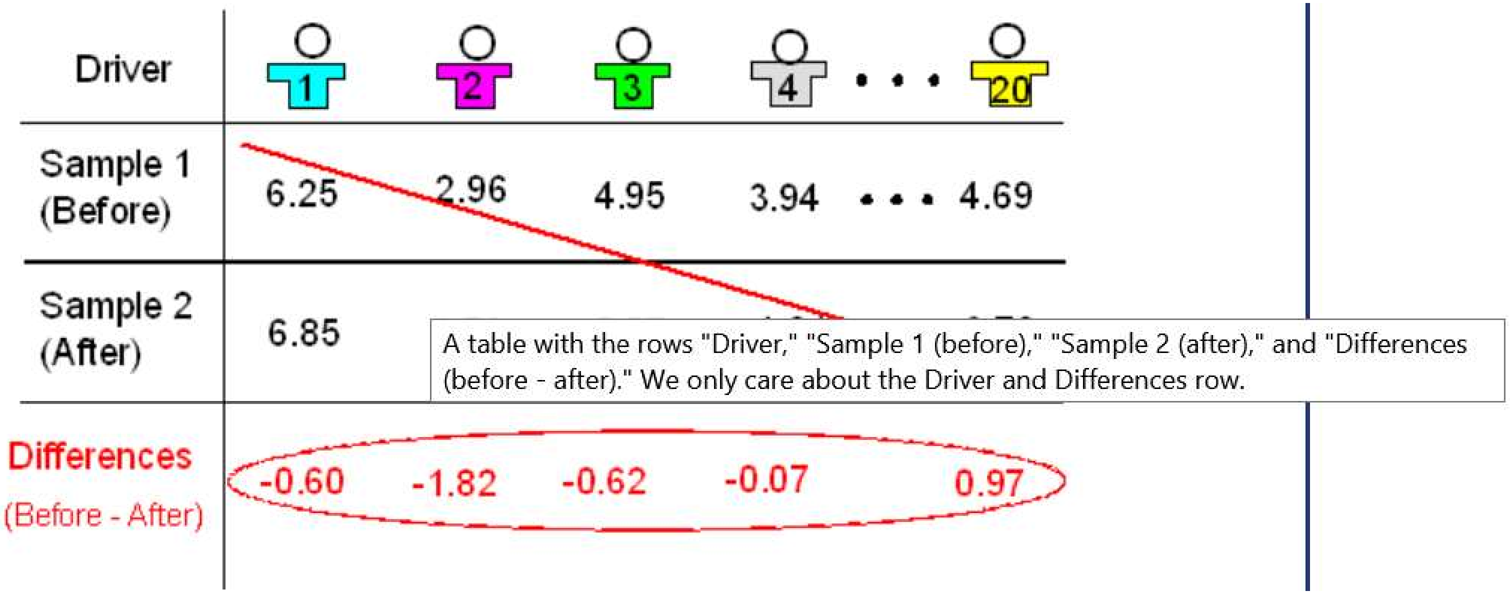
Attenzione: "Dipendenza" non richiede una relazione diretta causa/effetto.

Buon disegno sperimentale

- Supponiamo di voler testare l'efficacia di un farmaco progettato per abbassare la pressione sanguigna. Sarebbe meglio usare le misurazioni prima/dopo di un singolo gruppo di soggetti trattati con il farmaco piuttosto che utilizzare le misurazioni da due gruppi di soggetti separati trattati e non trattati con il farmaco. Il vantaggio dell'uso di coppie appaiate (misurazioni prima/dopo) è che riduciamo la variabilità esterna, che potrebbe verificarsi con i due campioni indipendenti. Questa strategia per la progettazione di un esperimento può essere generalizzata secondo il seguente principio di progettazione:
- Quando si progetta un esperimento o si pianifica uno studio osservazionale, l'uso di campioni dipendenti con coppie appaiate («matched») è generalmente migliore (più efficiente) rispetto all'utilizzo di due campioni indipendenti.

Le birre hanno influenzato il tempo di reazione?

Considerare la differenza entro soggetto consente di controllare per la variabilità del risultato non dovuta all'esposizione.



Driver	1	2	3	4	• • •	20
Sample 1 (Before)	6.25	2.96	4.95	3.94	• • •	4.69
Sample 2 (After)	6.85	[illegible]	[illegible]	[illegible]		[illegible]
Differences (Before - After)	-0.60	-1.82	-0.62	-0.07		0.97

A table with the rows "Driver," "Sample 1 (before)," "Sample 2 (after)," and "Differences (before - after)." We only care about the Driver and Differences row.

Quindi, invece di considerare μ_B e μ_A separatamente, possiamo concentrarci sulla media delle differenze tra i soggetti $\bar{d}$ = -0.5015, che è una stima della differenza (sconosciuta) $\delta_{before-after}$ nella popolazione prima e dopo l’assunzione di birra

Le birre hanno influenzato il tempo di reazione?

$H_0 : \delta = 0$

$H_1 : \delta \neq 0$

$\bar{d} = -0.5015$

$s_d = 0.8686$

$n = 20$

$$t = \frac{-0.5015}{0.8686/\sqrt{20}} = -2.5821$$

α=0.05 Valore critico: $t_{df=20-1,0.025} = \pm 2.093$

Rifiutiamo l'ipotesi nulla: due birre influenzano il tempo di reazione dei conducenti

P-value=0.0183 (dalla tabella t-student 0.01<p-value<0.02)

Le birre hanno influenzato il tempo di reazione?

$H_0: \delta = 0$

$H_1: \delta \neq 0$

$\bar{d} = -0.5015$

$s_d = 0.8686$

$n = 20$

$$95\%CI: -0.5015 \pm 2.093 * \frac{0.8686}{\sqrt{20}}$$

$$95\%CI: [-0.095; -0.908]$$

Abbiamo una confidenza del 95% nel dire che le due birre hanno aumentato il tempo di reazione degli autisti da un minimo di circa 0.1 a un massimo di 0.9 secondi.

Se gli stessi dati fossero provenienti da due campioni indipendenti?

Si sono reclutati ***40 conducenti*** che sono stati divisi casualmente in due gruppi in cui sono misurati i tempi di reazione in un percorso ad ostacoli dopo l'assunzione di bevande. Il primo gruppo ha bevuto due bicchieri d'acqua e il secondo due birre.
Lo scopo di questo studio era di verificare se lo stato di attenzione dei conducenti fosse alterato dopo aver bevuto due birre.

y_1:

y_2:

y_1 No beer	y_2 beer
6.25	6.85
2.96	4.78
4.95	5.57
3.94	4.01
4.85	5.91
4.81	5.34
6.6	6.09
5.33	5.84
5.19	4.19
4.88	5.75
5.75	6.25
5.26	7.23
3.16	4.55
6.65	6.42
5.49	5.25
4.05	5.59
4.42	3.96
4.99	5.93
5.01	6.03
4.69	3.72

Se gli stessi dati fossero provenienti da due campioni indipendenti?

$H_0: \delta = 0$ $\qquad \bar{x}_1 = 4.9615 \qquad \bar{x}_2 = 5.4630$

$H_1: \delta \neq 0$ $\qquad s_1 = 0.9709 \qquad s_2 = 0.9829$

$n_1 = 20 \qquad n_2 = 20$

$$s = \sqrt{\frac{19*0.9709^2+19*0.9829^2}{38}} = 0.9769$$

$$t = \frac{-0.5015}{0.9769 * \sqrt{\frac{1}{20}+\frac{1}{20}}} = -1.62$$

α=0.05 Valore critico: $t_{df=40-2,0.025}=\pm 2.021$

Non rifiuto l'ipotesi nulla: non ho sufficiente evidenza che due birre influenzino il tempo di reazione dei conducenti

P-value=0.113 (dalla tabella t-student 0.1<p-value<0.2)

Valutazione della potenza

1- β = P(rif H_0| H_1 è vero: $\delta_{prima-dopo} \neq 0$)

Assumiamo che le due birre aumentino il tempo di reazione di 1 secondo:

$$H_1: \delta_{prima\,-dopo} = \Delta = -1$$

La distribuzione della differenza campionaria sotto H_1:

- Con dati appaiati:

$$\bar{d}_{b-a} \sim N\left(\Delta\,;\ \sigma_d\sqrt{\frac{1}{n}}\right)$$

- Con dati indipendenti:

$$\bar{d}_{b-a} \sim N\left(\Delta\,;\ \sqrt{\frac{\sigma^2}{n_b} + \frac{\sigma^2}{n_a}}\right)$$

Valutazione della potenza – con dati appaiati

1- β = P(accettare H_1 | H_1 è vero: $\delta_{prima-dopo} \neq 0$)

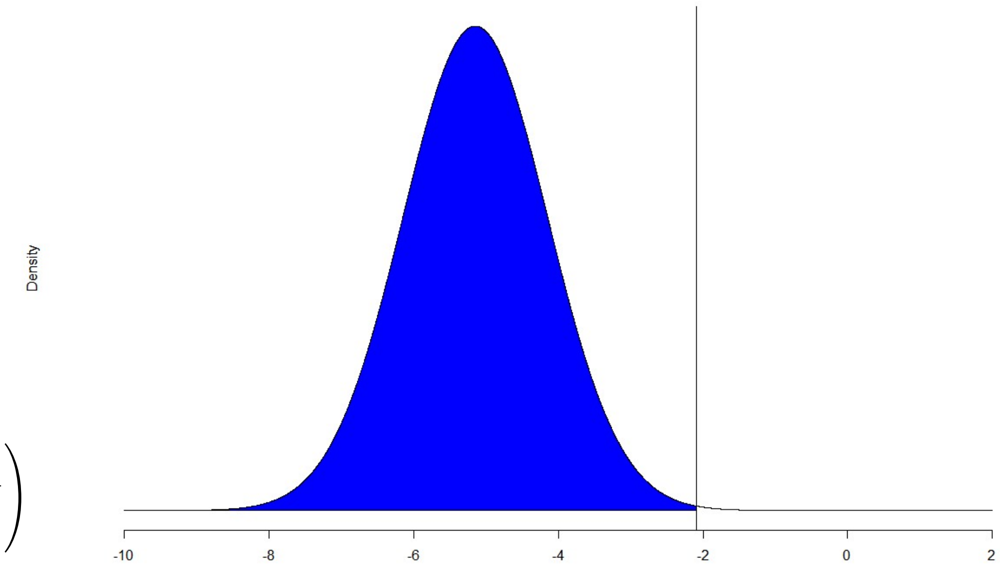
Assumiamo che le due birre aumentino il tempo di reazione di 1 secondo: $H_1 : \delta_{prima-dopo} = \Delta = -1$

$$\bar{D}_{b-a} \sim N\left(\Delta; \sigma_d \sqrt{\frac{1}{n}}\right)$$

$$Z = \frac{\bar{d}_{b-a}}{\sigma_d * \sqrt{\frac{1}{n}}} \sim N\left(\frac{\Delta}{\sigma_d * \sqrt{\frac{1}{n}}}; 1\right)$$

$$N\left(-5.15 = \frac{-1}{0.8686 * \sqrt{\frac{1}{20}}}; 1\right)$$

standardised mean difference



L'area blu rappresenta la probabilità(99.8%) di rigettare H_0 se H_1 è vera.

σ=0.8686 n=20

Valutazione della potenza – senza dati appaiati

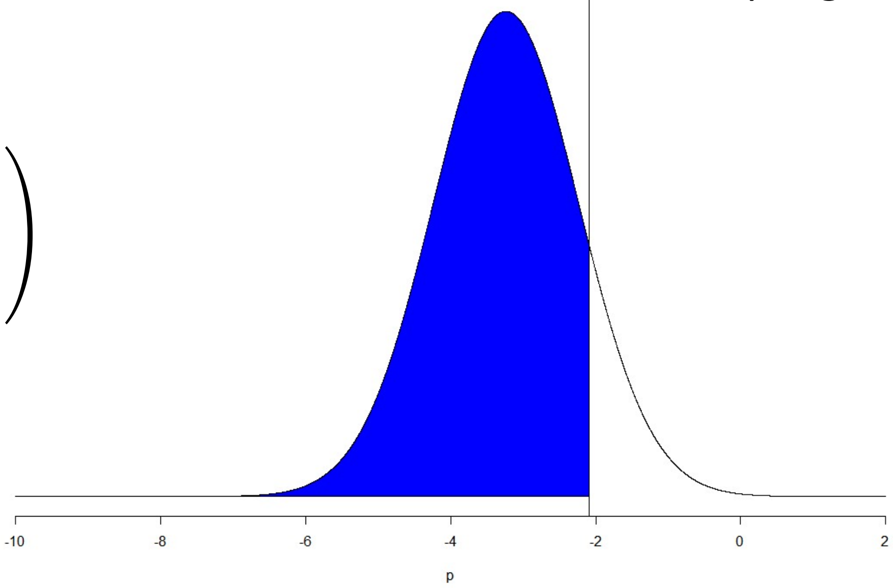
1- β = P(accettare H_1| H_1 è vero: $\delta_{prima-dopo} \neq 0$)

Assumiamo che le due birre aumentino il tempo di reazione di 1 secondo: $H_1 : \delta_{prima-dopo} = \Delta = -1$

$$Z = \frac{\bar{Y}_b - \bar{Y}_a}{\sigma * \sqrt{\frac{1}{n_b} + \frac{1}{n_a}}} \sim N\left(\frac{\Delta}{\sigma * \sqrt{\frac{1}{n_b} + \frac{1}{n_a}}}; 1\right)$$

$$N\left(-3.23 = \frac{-1}{0.98 * \sqrt{\frac{1}{20} + \frac{1}{20}}}; 1\right)$$

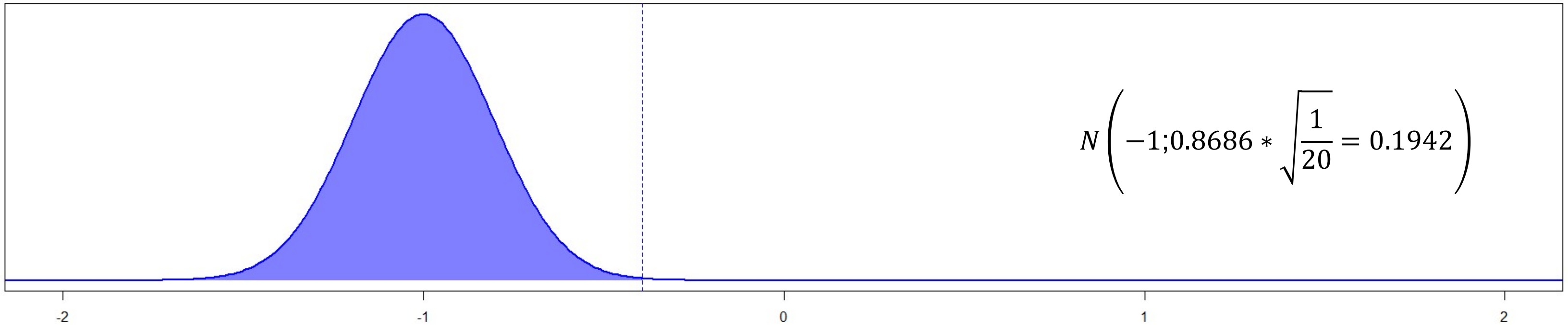
standardised difference between sampling means

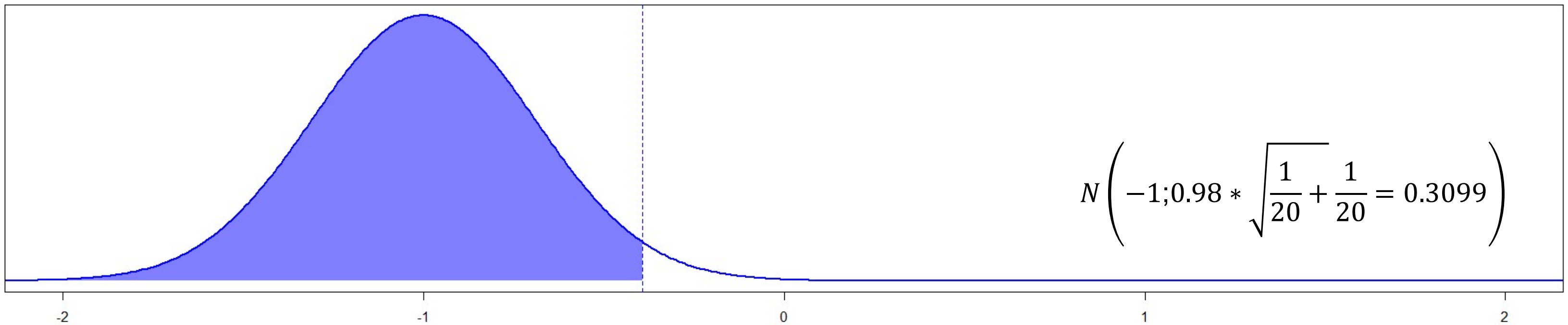


L'area blu rappresenta la probabilità(88%) di rigettare H_0 se H_1 è vera. Questa probabilità si riduce se ignoriamo l'appaiamento!

σ=0.98 n_a=20 n_b=20

Sulla scale della differenza ..





Riassunto sul test d'ipotesi

	VARIABILE DI RISPOSTA	
	CONTINUA (es. BMI, pressione sanguigna)	CATEGORICA (es. sì/no)
Confronto di un campione con un valore di riferimento	Test per il confronto di una media con un valore di riferimento (distribuzione t se varianza non nota)	Test per il confronto di una proporzione con un valore di riferimento (se np e n(1-p) $\geq$5)
Confronto di due campioni indipendenti (es. std/trt)	Test per il confronto di due medie (distribuzione t se varianza non nota)	Test per il confronto di due proporzioni (se np e n(1-p) $\geq$5)/test chi-quadrato
Confronto di due campioni appaiati/dipendenti (es.prima/dopo)	t-test per dati appaiati	McNeamer test
Confronto di tre o più campioni indipendenti (es. 3 classi di età)	Analisi della varianza/ regressione lineare	Test Chi-quadrato

Esercizio

L’effetto soporifero di un nuovo farmaco, F_2, è sperimentato rispetto al medesimo effetto di un farmaco già noto, F_1, su un gruppo di 8 volontari. È stato somministrato prima F_1 e nella notte successiva F_2. I risultati, valutati in termini di numero addizionale di ore di sonno rispetto ad un valore medio per soggetto, già noto, sono i seguenti:

ID	F_1	F_2	d_i	$(d_i-\bar{d})^2$
1	+0.4	+0.6		
2	+0.3	+0.5		
3	+0.9	+0.7		
4	+0.4	+0.6		
5	+1	+0.9		
6	+1	+1.1		
7	+1	+1.5		
8	+1	+2.1		

Verificare se i farmaci hanno lo stesso effetto con un opportuno test (α=0.05) e costruire un intervallo di confidenza per la differenza tra i due effetti

Risposte

$$\begin{cases} H_0 : \delta = 0 \\ H_1 : \delta \neq 0 \end{cases}$$

$$\bar{d} = \frac{2}{8} = 0.25$$

$$s_d^2 = \frac{1.14}{7} = 0.163 \Rightarrow s_d = 0.4$$

$$t = \frac{0.25}{0.4/\sqrt{8}} = \frac{0.25}{0.14} = 1.77$$

ID	F_1	F_2	$d_i = F_2 - F_1$	$(d_i - đ)^2$
1	+0.4	+0.6	+0.2	0.0025
2	+0.3	+0.5	+0.2	0.0025
3	+0.9	+0.7	-0.2	0.2025
4	+0.4	+0.6	+0.2	0.0025
5	+1	+0.9	-0.1	0.1225
6	+1	+1.1	+0.1	0.0225
7	+1	+1.5	+0.5	0.0625
8	+1	+2.1	+1.1	0.7225
			2	1.14

Poiché $t_{7;0.05}=2.365 \Rightarrow$
$|t|<t_0$ e non rifiuto H_0

$$I.C._{95\%} = 0.25 \pm 2.365 \cdot 0.14 = [-0.08; 0.58]$$

L'IC comprende il valore 0 quindi con confidenza del 95% concorda con il test

T-Test per dati appaiati («paired t-test»)

- Attenzione: le misurazioni fatte con F1 e F2 non sono indipendenti, perché sono effettuate sugli stessi soggetti!
- Non posso pertanto applicare il t-test per due campioni.
- Questo è un disegno di studio APPAIATO
- Un disegno appaiato garantisce che le osservazioni siano a parità di variabilità biologica.
- Si isola in maniera più **precisa** il possibile effetto del trattamento (che ha anch'esso una sua variabilità) rispetto a un disegno non appaiato.
- La variabilità biologica viene «eliminata» grazie all'appaiamento!