

## **Lezione 8**

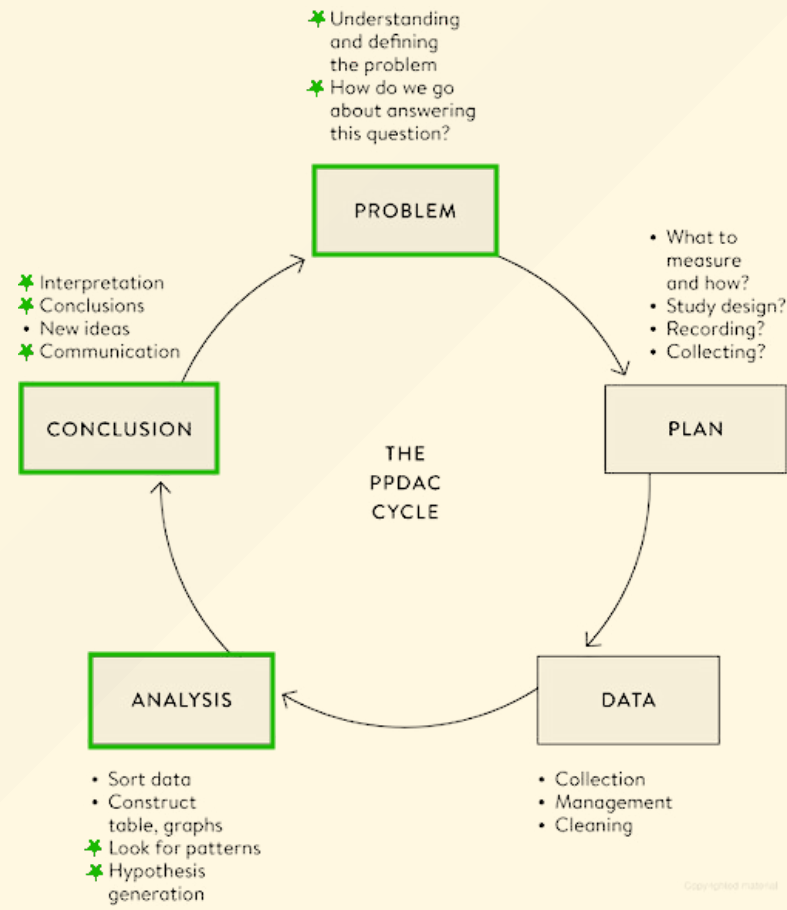
# **La statistica inferenziale**

## **(Parte II: Test di ipotesi)**

# Obiettivi di apprendimento

- Formulare e testare ipotesi
- Interpretare P-value (e la loro relazione con i CI)
- Saper distinguere tra errori del primo e del secondo tipo
- Interpretare la potenza di uno studio

# Le fasi della ricerca



## **Attenzione**

Questa parte continua a essere complessa, ma non demordete, siamo alla fine!

# Cos'è un'ipotesi?

“ Una possibile spiegazione per un fenomeno, che non rappresenta la verità assoluta, ma una congettura provvisoria ”

# Esempi di ipotesi

- Il peso alla nascita è diverso nei gemelli monozigoti e dizigoti inglesi?
- Il numero di bambini deceduti a seguito di interventi cardiocirurgici a Bristol tra il 1984 e il 1995 è diverso rispetto ad altri ospedali britannici?

# Il principio di falsificabilità e l'ipotesi nulla

- Il peso alla nascita è ~~diverso~~ *uguale* nei gemelli monozigoti e dizigoti inglesi?
- Il numero di bambini deceduti a seguito di interventi cardiocirurgici a Bristol tra il 1984 e il 1995 è ~~diverso~~ *uguale* rispetto ad altri ospedali britannici?

# Esercizio 1

? Qual è l'ipotesi nulla nei seguenti studi

Studio 1

**JAMA**  
Network | **Open**<sup>TM</sup>

**Original Investigation** | Geriatrics

**Social Isolation Changes and Long-Term Outcomes Among Older Adults**

Chen Lyu, MS, PhD; Katherine Siu, MS; Ian Xu, HS; Iman Osman, MD; Judy Zhong, PhD



**Key Points**  
**Question** Is social isolation change associated with long-term outcomes in older adults?



# Esercizio 1

? Qual è l'ipotesi nulla nei seguenti studi

Studio 1



---

**Original Investigation** | Geriatrics

**Social Isolation Changes and Long-Term Outcomes Among Older Adults**

Chen Lyu, MS, PhD; Katherine Siu, MS; Ian Xu, HS; Iman Osman, MD; Judy Zhong, PhD

**Key Points**  
**Question** Is social isolation change associated with long-term outcomes in older adults?

*Soluzione:* Social isolation changes are **NOT** associated with long-term outcomes

# Esercizio 1

? Qual è l'ipotesi nulla nei seguenti studi

Studio 2

JAMA Psychiatry | **Original Investigation**

## Virtual Reality and Transcranial Direct Current Stimulation for Posttraumatic Stress Disorder A Randomized Clinical Trial

Mascha van 't Wout-Frank, PhD; Amanda R. Arulpragasam, PhD;  
Christiana Faucher, BS; Emily Aiken, MA;  
M. Tracie Shea, PhD; Richard N. Jones, SciD; Benjamin D. Greenberg, MD, PhD;  
Noah S. Philip, MD

### Key Points

**Question** Can therapeutic exposure using virtual reality (VR) be augmented with simultaneously applied transcranial direct current stimulation (tDCS) to reduce symptoms of posttraumatic stress disorder (PTSD)?

# Esercizio 1

? Qual è l'ipotesi nulla nei seguenti studi

Studio 2

JAMA Psychiatry | **Original Investigation**

## Virtual Reality and Transcranial Direct Current Stimulation for Posttraumatic Stress Disorder A Randomized Clinical Trial

Mascha van 't Wout-Frank, PhD; Amanda R. Arulpragasam, PhD; Christiana Faucher, BS; Emily Aiken, MA; M. Tracie Shea, PhD; Richard N. Jones, ScID; Benjamin D. Greenberg, MD, PhD; Noah S. Philip, MD

**Key Points**

**Question** Can therapeutic exposure using virtual reality (VR) be augmented with simultaneously applied transcranial direct current stimulation (tDCS) to reduce symptoms of posttraumatic stress disorder (PTSD)?

*Soluzione:* Therapeutic exposure using VR **CANNOT** be augmented with simultaneously applied tDCS

# Esercizio 1

? Qual è l'ipotesi nulla nei seguenti studi

Studio 3

JAMA Psychiatry | **Original Investigation**

## Gratitude and Mortality Among Older US Female Nurses

Ying Chen, ScD; Olivia I. Okereke, MD, SM; Eric S. Kim, PhD; Henning Tiemeier, MD, PhD;  
Laura D. Kubzansky, PhD; Tyler J. VanderWeele, PhD

### Key Points

**Question** Do people who more frequently notice and feel grateful for positive experiences tend to live longer?

# Esercizio 1

? Qual è l'ipotesi nulla nei seguenti studi

Studio 3

JAMA Psychiatry | **Original Investigation**

## Gratitude and Mortality Among Older US Female Nurses

Ying Chen, ScD; Olivia I. Okereke, MD, SM; Eric S. Kim, PhD; Henning Tiemeier, MD, PhD;  
Laura D. Kubzansky, PhD; Tyler J. VanderWeele, PhD

### Key Points


**Question** Do people who more frequently notice and feel grateful for positive experiences tend to live longer?

*Soluzione:* People who more frequently notice and feel grateful for positive experiences **DO NOT** tend to live longer

## Esercizio 2

- ?
- Se non rifiuto l'ipotesi nulla significa che
- a) ho provato che l'ipotesi nulla sia vera
  - b) ho provato che l'ipotesi nulla sia falsa
  - c) le mie osservazioni sono compatibili con l'ipotesi nulla
  - d) le mie osservazioni non sono compatibili con l'ipotesi nulla
  - d) dipende dalla domanda di ricerca
  - e) nessuno dei precedenti

## Esercizio 2 -- Soluzione

- ?
- Se non rifiuto l'ipotesi nulla significa che
- a) ho provato che l'ipotesi nulla sia vera
  - b) ho provato che l'ipotesi nulla sia falsa
  - c) le mie osservazioni sono compatibili con l'ipotesi nulla 
  - d) le mie osservazioni non sono compatibili con l'ipotesi nulla
  - d) dipende dalla domanda di ricerca
  - e) nessuno dei precedenti

# $z$ -test

 Il peso alla nascita è ~~diverso~~ *uguale* nei gemelli monozigoti e dizigoti inglesi?

$$n_d = 3481, \bar{x}_d = 2462\text{g}, s_d = 577\text{g}$$

$$n_m = 3823, \bar{x}_m = 2350\text{g}, s_m = 579\text{g}$$

$$\mu_d - \mu_m = 0$$

→ Ipotesi nulla ( $\mathcal{H}_0$ )



# z-test

 Il peso alla nascita è ~~diverso~~ *uguale* nei gemelli monozigoti e dizigoti inglesi?

$$n_d = 3481, \bar{x}_d = 2462\text{g}, s_d = 577\text{g}$$

$$n_m = 3823, \bar{x}_m = 2350\text{g}, s_m = 579\text{g}$$

$$\mu_d - \mu_m = 0$$

→ Ipotesi nulla ( $\mathcal{H}_0$ )

$$\mu_d - \mu_m \neq 0$$

→ Ipotesi alternativa ( $\mathcal{H}_{1/A}$ )

# z-test

 Il peso alla nascita è ~~diverso~~ *uguale* nei gemelli monozigoti e dizigoti inglesi?

$$n_d = 3481, \bar{x}_d = 2462\text{g}, s_d = 577\text{g}$$

$$n_m = 3823, \bar{x}_m = 2350\text{g}, s_m = 579\text{g}$$

$$\mu_d - \mu_m = 0$$

→ Ipotesi nulla ( $\mathcal{H}_0$ )

$$\mu_d - \mu_m \neq 0$$

$$\mu_m - \mu_d \neq 0$$

$$\rightarrow |\mu_d - \mu_m| \neq 0$$

Ipotesi alternativa ( $\mathcal{H}_{1/A}$ )

# $z$ -test

 Il peso alla nascita è ~~diverso~~ *uguale* nei gemelli monozigoti e dizigoti inglesi?

$$n_d = 3481, \bar{x}_d = 2462\text{g}, s_d = 577\text{g}$$

$$n_m = 3823, \bar{x}_m = 2350\text{g}, s_m = 579\text{g}$$

$$\mu_d - \mu_m = 0 \quad \leftarrow$$

$$\bar{x}_d - \bar{x}_m = 2462 - 2350 = (\pm)112$$

# $z$ -test

📌 Il peso alla nascita è ~~diverso~~ *uguale* nei gemelli monozigoti e dizigoti inglesi?

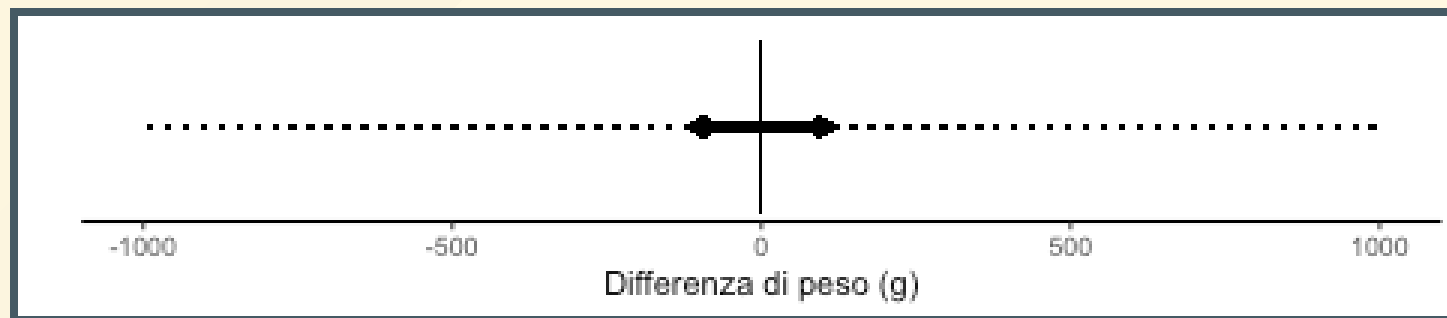
$$n_d = 3481, \bar{x}_d = 2462\text{g}, s_d = 577\text{g}$$

$$n_m = 3823, \bar{x}_m = 2350\text{g}, s_m = 579\text{g}$$

$$\mu_d - \mu_m = 0 \quad \leftarrow$$

$$\bar{x}_d - \bar{x}_m = 2462 - 2350 = (\pm)112$$

? Una differenza di 112 g è "abbastanza vicina" a zero per concludere  $\mu_d - \mu_m = 0$ ?



# $z$ -test

 Il peso alla nascita è ~~diverso~~ *uguale* nei gemelli monozigoti e dizigoti inglesi?

$$n_d = 3481, \bar{x}_d = 2462\text{g}, s_d = 577\text{g}$$

$$n_m = 3823, \bar{x}_m = 2350\text{g}, s_m = 579\text{g}$$

$$\mu_d - \mu_m = 0 \quad \leftarrow$$

$$\bar{x}_d - \bar{x}_m = 2462 - 2350 = (\pm)112$$

? Una differenza di 112 g è "abbastanza vicina" a zero per concludere  $\mu_d - \mu_m = 0$ ?

? Qual è la probabilità di osservare una differenza di 112 g se  $\mu_d - \mu_m = 0$ ?

# Facciamo un paio di passi indietro

1. La Normale è definita dalla sua media e deviazione standard e corrisponde a una distribuzione di probabilità  
→ Area sottesa a  $Z \equiv$  probabilità  $\mathcal{P}$
2. Il teorema del limite centrale ci dice che le distribuzioni campionarie (incluso la differenza delle medie) tendono alla Normale

# Facciamo un paio di passi indietro

1. La Normale è definita dalla sua media e deviazione standard e corrisponde a una distribuzione di probabilità  
→ Area sottesa a  $Z \equiv$  probabilità  $\mathcal{P}$
2. Il teorema del limite centrale ci dice che le distribuzioni campionarie (incluso la differenza delle medie) tendono alla Normale

Per la differenza tra due medie

$$\mathcal{N} = \left( \mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right) \text{ con } \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \rightarrow \text{standard error}$$

# z-test

📌 Il peso alla nascita è ~~diverso~~ *uguale* nei gemelli monozigoti e dizigoti inglesi?

$$n_d = 3481, \bar{x}_d = 2462\text{g}, s_d = 577\text{g}$$

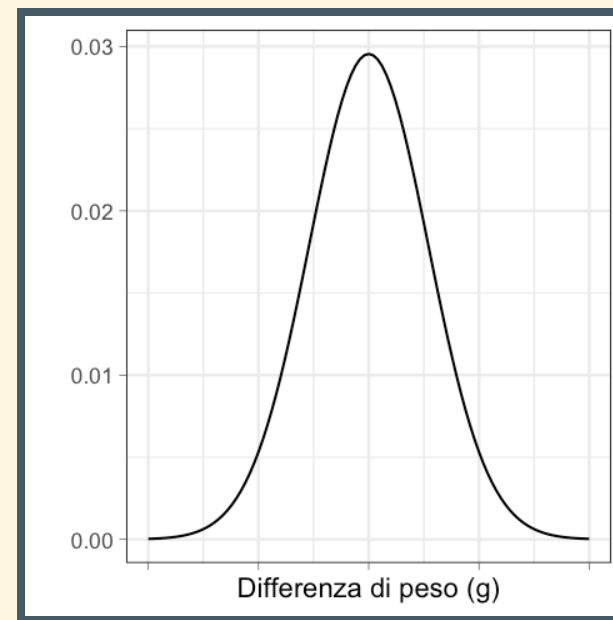
$$n_m = 3823, \bar{x}_m = 2350\text{g}, s_m = 579\text{g}$$

$$\mu_d - \mu_m = 0 \quad \leftarrow$$

$$\bar{x}_d - \bar{x}_m = 2462\text{g} - 2350\text{g} = 112\text{g}$$

$$\mathcal{N} = (\mu_d - \mu_m, \text{SE}), \text{ con } \mu_d - \mu_m = 0$$

$$\bar{x}_d - \bar{x}_m = 2462 - 2350 = (\pm)112$$



? Qual è la probabilità di osservare una differenza di 112g se  $\mu_d - \mu_m = 0$ ?  
Area sottesa a  $Z \equiv$  probabilità  $\mathcal{P}$



# z-test

📌 Il peso alla nascita è ~~diverso~~ *uguale* nei gemelli monozigoti e dizigoti inglesi?

$$n_d = 3481, \bar{x}_d = 2462\text{g}, s_d = 577\text{g}$$

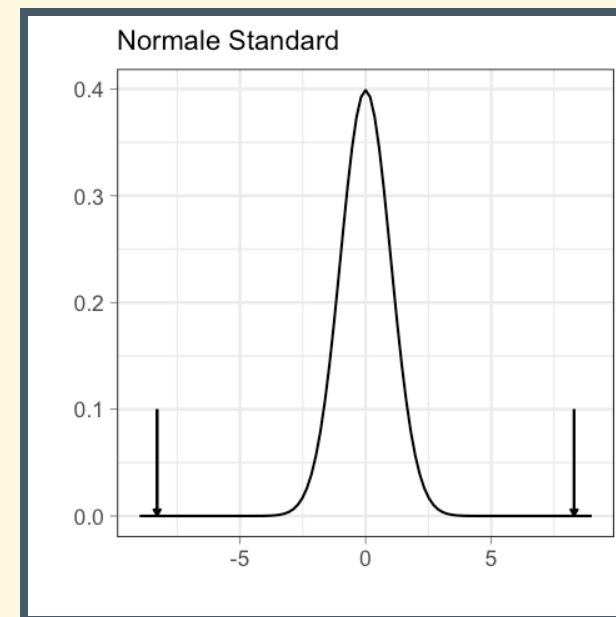
$$n_m = 3823, \bar{x}_m = 2350\text{g}, s_m = 579\text{g}$$

$$\mu_d - \mu_m = 0 \quad \leftarrow$$

$$\bar{x}_d - \bar{x}_m = 2462 - 2350 = (\pm)112$$

$$\mathcal{N} = (\mu_d - \mu_m, \text{SE}), \text{ con } \mu_d - \mu_m = 0$$


$$\text{e SE} = \sqrt{\frac{s_d^2}{n_d} + \frac{s_m^2}{n_m}} = 13.5\text{g}$$




? Qual è la probabilità di osservare una differenza di 112g se  $\mu_d - \mu_m = 0$ ?

$$z = \frac{\bar{X} - \mu}{\hat{SE}} = \frac{\pm 112 - 0}{13.5} = \pm 8.3 \rightarrow \mathcal{P} = 2 \times (7.0 \times 10^{-17}) = 1.4 \times 10^{-16}$$

# P-value

-  Il P-value misura la discrepanza tra i dati e  $\mathcal{H}_0$  e corrisponde alla probabilità di ottenere un risultato tanto estremo quanto quello ottenuto se l'ipotesi nulla fosse vera.

# P-value

 Il P-value misura la discrepanza tra i dati e  $\mathcal{H}_0$  e corrisponde alla probabilità di ottenere un risultato tanto estremo quanto quello ottenuto se l'ipotesi nulla fosse vera.

P-value = 0.5  $\rightarrow$  50%  $\rightarrow$  1 campione su 2

P-value = 0.1  $\rightarrow$  10%  $\rightarrow$  1 campione su 10

P-value = 0.05  $\rightarrow$  5%  $\rightarrow$  1 campione su 20

P-value = 0.01  $\rightarrow$  1%  $\rightarrow$  1 campione su 100

P-value = 0.005  $\rightarrow$  0.5%  $\rightarrow$  1 campione su 200

# P-value e significatività statistica

- 🎯 Il P-value misura la discrepanza tra i dati e  $\mathcal{H}_0$  e corrisponde alla probabilità di ottenere un risultato tanto estremo quanto quello ottenuto se l'ipotesi nulla fosse vera.

Se il P-value è minore di una soglia critica (o livello di significatività)  $\alpha$ , possiamo dire che il risultato è statisticamente significativo

$$\alpha = 0.05 \vee 0.01$$

# **Test di ipotesi, un passo alla volta**

# Test di ipotesi, un passo alla volta

## 1. Definisco la mia ipotesi nulla ( $\mathcal{H}_0$ )

Il peso alla nascita è ~~diverso~~ *uguale* nei gemelli monozigoti e dizigoti inglesi?

$$\mathcal{H}_0 : \mu_d - \mu_m = 0$$

# Test di ipotesi, un passo alla volta

1. Definisco la mia ipotesi nulla ( $\mathcal{H}_0$ )
2. Scelgo un test statistico che stimi qualcosa che, se abbastanza estremo, mi faccia dubitare di  $\mathcal{H}_0$

$z$ -test della differenza di due medie campionarie

# Test di ipotesi, un passo alla volta

1. Definisco la mia ipotesi nulla ( $\mathcal{H}_0$ )
2. Scelgo un test statistico che stimi qualcosa che, se abbastanza estremo, mi faccia dubitare di  $\mathcal{H}_0$
3. Genero la distribuzione campionaria del test scelto, assumendo  $\mathcal{H}_0$  vera

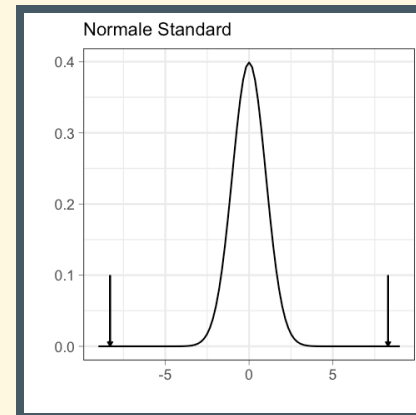
$$\mathcal{N} = (\mu_d - \mu_m, \text{SE}), \text{ con } \mu_d - \mu_m = 0$$



# Test di ipotesi, un passo alla volta

1. Definisco la mia ipotesi nulla ( $\mathcal{H}_0$ )
2. Scelgo un test statistico che stimi qualcosa che, se abbastanza estremo, mi faccia dubitare di  $\mathcal{H}_0$
3. Genero la distribuzione campionaria del test scelto, assumendo  $\mathcal{H}_0$  vera
4. Verifico se la statistica osservata si trovi sulla coda di questa distribuzione e assegno una probabilità (P-value) a questo evento

$$\text{P-value} = 1.4 \times 10^{-16}$$



# Test di ipotesi, un passo alla volta

1. Definisco la mia ipotesi nulla ( $\mathcal{H}_0$ )
2. Scelgo un test statistico che stimi qualcosa che, se abbastanza estremo, mi faccia dubitare di  $\mathcal{H}_0$
3. Genero la distribuzione campionaria del test scelto, assumendo  $\mathcal{H}_0$  vera
4. Verifico se la statistica osservata si trovi sulla coda di questa distribuzione e assegno una probabilità (P-value) a questo evento
5. Dichiaro il risultato come statisticamente significativo se il P-value è inferiore a una soglia critica  $\alpha$

$$\text{P-value} = 1.4 \times 10^{-16} < 0.05 \rightarrow \text{rifiuto } \mathcal{H}_0$$

# Esercizio #3

## RESULTS

Of the 355 children and adolescents who underwent screening, 290 were enrolled. A total of 146 participants were assigned to the oxytocin group and 144 to the placebo group; 139 and 138 participants, respectively, completed both the baseline and at least one postbaseline ABC-mSW assessments and were included in the modified intention-to-treat analyses. The least-squares mean change from baseline in the ABC-mSW score (primary outcome) was -3.7 in the oxytocin group and -3.5 in the placebo group (least-squares mean difference, -0.2; 95% confidence interval, -1.5 to 1.0;  $P=0.61$ ). Secondary outcomes generally did not differ between the trial groups. The incidence and severity of adverse events were similar in the two groups.



Qual è l'ipotesi nulla dello studio?

- a) L'intervento riduce l'interazione sociale
- b) Nel gruppo di controllo non c'è aumento di interazione sociale
- c) Non c'è differenza in aumento di interazione sociale tra il gruppo di intervento e quello di controllo
- d) Nei due gruppi c'è una differenza significativa nell'aumento della interazione sociale

Sikich, L. *et al.*, *Intranasal Oxytocin in Children and Adolescents with Autism Spectrum Disorder*, NEJM, 2021

02:00


# Esercizio #3 -- Soluzione

## RESULTS

Of the 355 children and adolescents who underwent screening, 290 were enrolled. A total of 146 participants were assigned to the oxytocin group and 144 to the placebo group; 139 and 138 participants, respectively, completed both the baseline and at least one postbaseline ABC-mSW assessments and were included in the modified intention-to-treat analyses. The least-squares mean change from baseline in the ABC-mSW score (primary outcome) was -3.7 in the oxytocin group and -3.5 in the placebo group (least-squares mean difference, -0.2; 95% confidence interval, -1.5 to 1.0;  $P=0.61$ ). Secondary outcomes generally did not differ between the trial groups. The incidence and severity of adverse events were similar in the two groups.



Qual è l'ipotesi nulla dello studio?

- a) L'intervento riduce l'interazione sociale
- b) Nel gruppo di controllo non c'è aumento di interazione sociale
- c) Non c'è differenza in aumento di interazione sociale tra il gruppo di intervento e quello di controllo 
- d) Nei due gruppi c'è una differenza significativa nell'aumento della interazione sociale

Sikich, L. *et al.*, *Intranasal Oxytocin in Children and Adolescents with Autism Spectrum Disorder*, NEJM, 2021

# Esercizio #4

## RESULTS

Of the 355 children and adolescents who underwent screening, 290 were enrolled. A total of 146 participants were assigned to the oxytocin group and 144 to the placebo group; 139 and 138 participants, respectively, completed both the baseline and at least one postbaseline ABC-mSW assessments and were included in the modified intention-to-treat analyses. The least-squares mean change from baseline in the ABC-mSW score (primary outcome) was -3.7 in the oxytocin group and -3.5 in the placebo group (least-squares mean difference, -0.2; 95% confidence interval, -1.5 to 1.0;  $P=0.61$ ). Secondary outcomes generally did not differ between the trial groups. The incidence and severity of adverse events were similar in the two groups.



Formulare operativamente l'ipotesi nulla

a)  $\mu_i - \mu_c = 0$

b)  $\mu_i - \mu_c \neq 0$

c)  $\bar{x}_i - \bar{x}_c = 0$

d)  $\bar{x}_i - \bar{x}_c \neq 0$

02:00

Sikich, L. *et al.*, *Intranasal Oxytocin in Children and Adolescents with Autism Spectrum Disorder*, NEJM, 2021

# Esercizio #4 -- Soluzione

## RESULTS

Of the 355 children and adolescents who underwent screening, 290 were enrolled. A total of 146 participants were assigned to the oxytocin group and 144 to the placebo group; 139 and 138 participants, respectively, completed both the baseline and at least one postbaseline ABC-mSW assessments and were included in the modified intention-to-treat analyses. The least-squares mean change from baseline in the ABC-mSW score (primary outcome) was -3.7 in the oxytocin group and -3.5 in the placebo group (least-squares mean difference, -0.2; 95% confidence interval, -1.5 to 1.0;  $P=0.61$ ). Secondary outcomes generally did not differ between the trial groups. The incidence and severity of adverse events were similar in the two groups.



Formulare operativamente l'ipotesi nulla

a)  $\mu_i - \mu_c = 0$  

b)  $\mu_i - \mu_c \neq 0$

c)  $\bar{x}_i - \bar{x}_c = 0$


d)  $\bar{x}_i - \bar{x}_c \neq 0$

Sikich, L. *et al.*, *Intranasal Oxytocin in Children and Adolescents with Autism Spectrum Disorder*, NEJM, 2021

## Esercizio #5

- ?
- Il P-value associato alla variabile “Sex” è pari a 0.84. Con un livello di significatività del 5%, ci sono differenze statisticamente significative nella distribuzione maschi/femmine nei due gruppi?
- a) sì, perché si tratta di uno studio clinico randomizzato
  - b) sì, perché il P-value è minore del livello di significatività
  - c) no, perché il P-value è maggiore del livello di significatività
  - d) non ci sono informazioni sufficienti per stabilirlo

## Esercizio #5 -- Soluzione

- ? Il P-value associato alla variabile “Sex” è pari a 0.84. Con un livello di significatività del 5%, ci sono differenze statisticamente significative nella distribuzione maschi/femmine nei due gruppi?
- a) sì, perché si tratta di uno studio clinico randomizzato
  - b) sì, perché il P-value è minore del livello di significatività
  - c) no, perché il P-value è maggiore del livello di significatività 
  - d) non ci sono informazioni sufficienti per stabilirlo



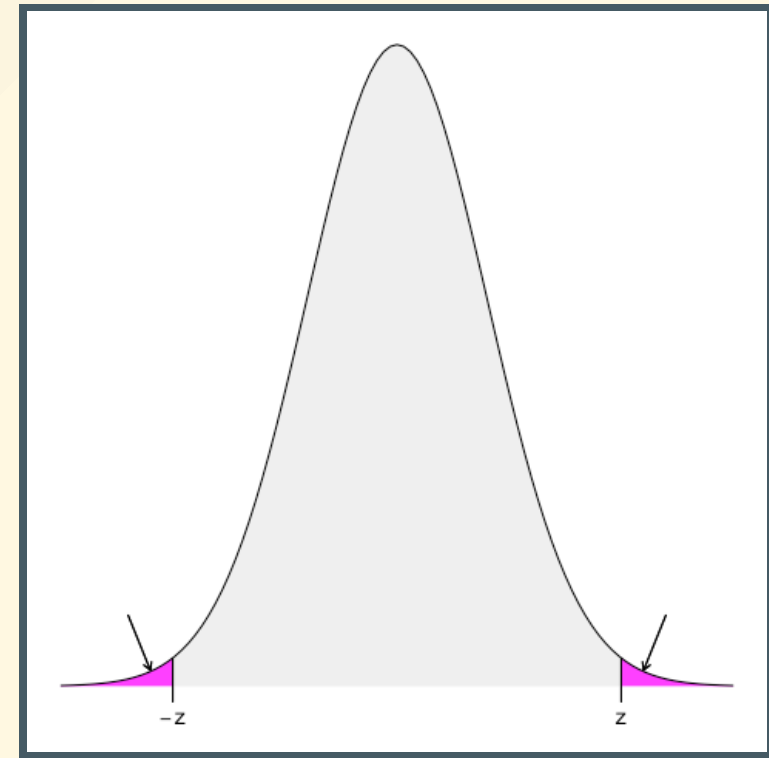
# Uguale, diverso, maggiore, minore?



$$\mathcal{H}_1: \mu_i - \mu_c \neq 0$$

$$\mathcal{H}_0: \mu_i - \mu_c = 0$$

→ test a due code



# Uguale, diverso, maggiore, minore?



$$\mathcal{H}_1: \mu_i - \mu_c \neq 0$$

$$\mathcal{H}_0: \mu_i - \mu_c = 0$$

→ test a due code

$$\mathcal{H}_1: \mu_i - \mu_c < 0$$

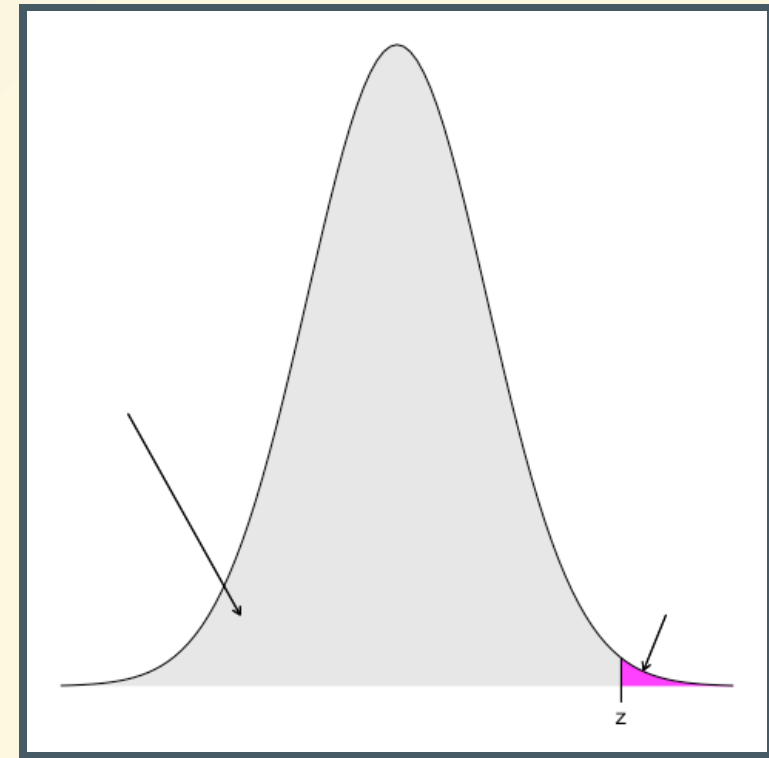
$$\mathcal{H}_0: \mu_i - \mu_c \geq 0$$

oppure

$$\mathcal{H}_1: \mu_i - \mu_c > 0$$

$$\mathcal{H}_0: \mu_i - \mu_c \leq 0$$

→ test a una coda



## Esercizio #6

- ? Nei metodi è stato riportato che “All significance tests were 2-sided, and  $P < 0.05$  was used to indicate significance”. Qual è il livello di significatività delle analisi?
- a) 0.05
  - b) 0.025
  - c) il P-value
  - d) non ci sono informazioni sufficienti per stabilirlo

## Esercizio #6 -- Soluzione

- ? Nei metodi è stato riportato che “All significance tests were 2-sided, and  $P < 0.05$  was used to indicate significance”. Qual è il livello di significatività delle analisi?
- a) 0.05 ☒
  - b) 0.025
  - c) il P-value
  - d) non ci sono informazioni sufficienti per stabilirlo

## Esercizio #7

? Dopo aver analizzato i nostri dati, abbiamo ottenuto  $z = 1.9$   
Quali delle seguenti ipotesi nulle andrò a rifiutare con  $\alpha = 0.05$ ?

a)  $\mu_i - \mu_c \neq 0$

b)  $\mu_i - \mu_c \geq 0$

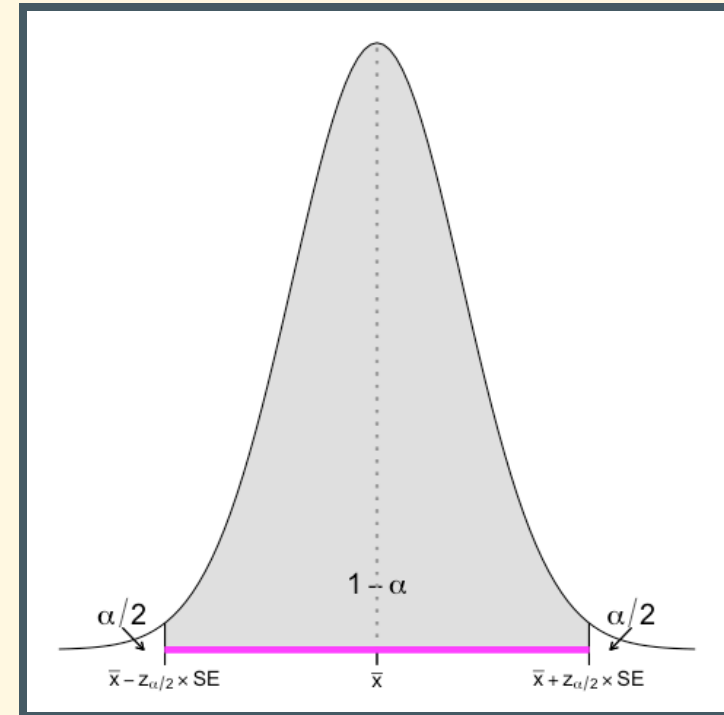
c)  $\mu_i - \mu_c \leq 0$

## Esercizio #7 -- Soluzione

? Dopo aver analizzato i nostri dati, abbiamo ottenuto  $z = 1.9$   
Quali delle seguenti ipotesi nulle andrò a rifiutare con  $\alpha = 0.05$ ?

- a)  $\mu_i - \mu_c \neq 0$
- b)  $\mu_i - \mu_c \geq 0$
- c)  $\mu_i - \mu_c \leq 0$

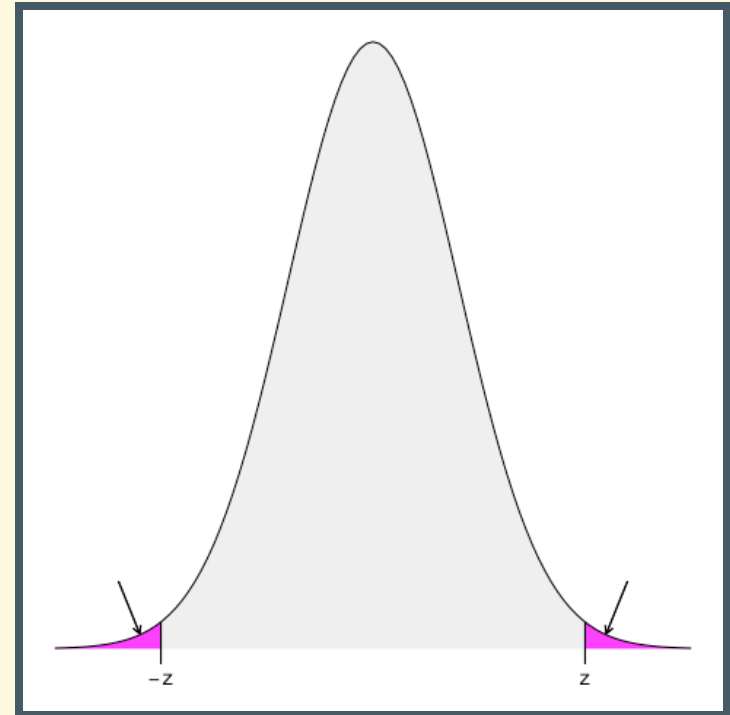
Confidence Level	$\alpha$	$\alpha/2$	$z_{\alpha/2}$
95%	5%	2.5%	1.96



## Esercizio #7 -- Soluzione

? Dopo aver analizzato i nostri dati, abbiamo ottenuto  $z = 1.9$   
Quali delle seguenti ipotesi nulle andrò a rifiutare con  $\alpha = 0.05$ ?

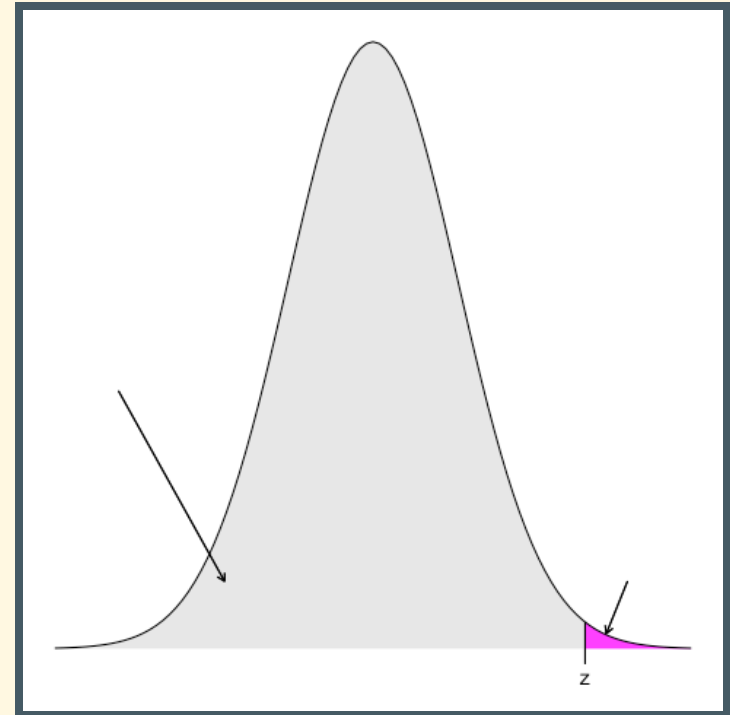
- a)  $\mu_i - \mu_c \neq 0$   
 $\rightarrow P = 2 \times 0.29 = 0.56$
- b)  $\mu_i - \mu_c \geq 0$
- c)  $\mu_i - \mu_c \leq 0$



## Esercizio #7 -- Soluzione

? Dopo aver analizzato i nostri dati, abbiamo ottenuto  $z = 1.9$   
Quali delle seguenti ipotesi nulle andrò a rifiutare con  $\alpha = 0.05$ ?

- a)  $\mu_i - \mu_c \neq 0$   
 $\rightarrow P = 2 \times 0.029 = 0.056$
- b)  $\mu_i - \mu_c \geq 0$   
 $\rightarrow P = 0.029$
- c)  $\mu_i - \mu_c \leq 0$   
 $\rightarrow P = 1 - 0.029 = 0.971$

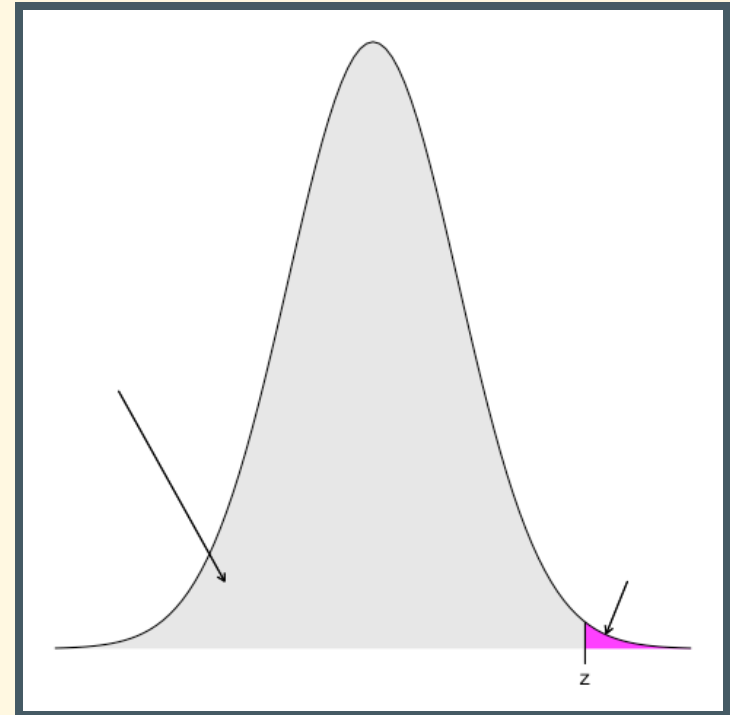




## Esercizio #7 -- Soluzione

? Dopo aver analizzato i nostri dati, abbiamo ottenuto  $z = 1.9$   
Quali delle seguenti ipotesi nulle andrò a rifiutare con  $\alpha = 0.05$ ?

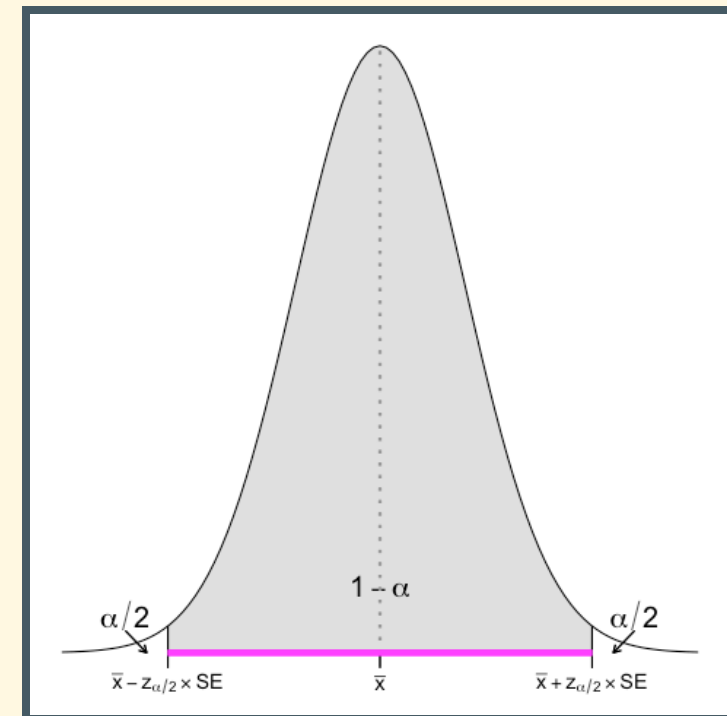
- a)  $\mu_i - \mu_c \neq 0$   
 $\rightarrow P = 2 \times 0.029 = 0.056$
- b)  $\mu_i - \mu_c \geq 0$   
 $\rightarrow P = 0.029$  ✓
- c)  $\mu_i - \mu_c \leq 0$   
 $\rightarrow P = 1 - 0.029 = 0.971$



# Test di ipotesi & intervallo di confidenza

- 🎯 Il 95% confidence interval è l'insieme delle ipotesi nulle che non sono rifiutate con  $\alpha = 0.05$

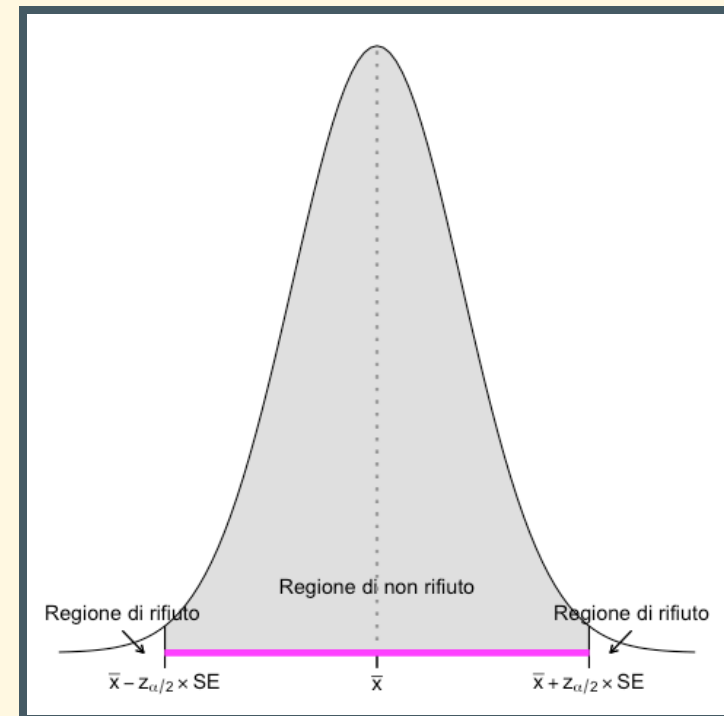
Confidence Level	$\alpha$	$\alpha/2$	$z_{\alpha/2}$
95%	5%	2.5%	1.96



# Test di ipotesi & intervallo di confidenza

🎯 Il 95% confidence interval è l'insieme delle ipotesi nulle che non sono rifiutate con  $\alpha = 0.05$

In un test a due code,  $P < 0.05$   
se il 95% CI non include l'ipotesi nulla (solitamente zero)



# Pearson's $\chi^2$ test



Il numero di bambini deceduti a seguito di interventi cardiocirurgici a Bristol tra il 1984 e il 1995 è ~~diverso~~ *uguale* rispetto ad altri ospedali britannici?

Nell'ospedale di Bristol, sono stati effettuati 143 interventi cardiocirurgici e sono stati registrati 41 decessi (27.8%). Negli altri ospedali britannici, a fronte di 3176 operazioni sono stati registrati 356 decessi (10.7%).

# Pearson's $\chi^2$ test



Il numero di bambini deceduti a seguito di interventi cardiocirurgici a Bristol tra il 1984 e il 1995 è ~~diverso~~ *uguale* rispetto ad altri ospedali britannici?

Nell'ospedale di Bristol, sono stati effettuati 143 interventi cardiocirurgici e sono stati registrati 41 decessi (27.8%). Negli altri ospedali britannici, a fronte di 3176 operazioni sono stati registrati 356 decessi (10.7%).

1. Definisco la mia ipotesi nulla ( $\mathcal{H}_0$ )

$$\mathcal{H}_0 : \pi_B - \pi_H = 0$$

$$\mathcal{H}_1 : \pi_B - \pi_H \neq 0$$

# Pearson's $\chi^2$ test

- 📌 Il numero di bambini deceduti a seguito di interventi cardiocirurgici a Bristol tra il 1984 e il 1995 è ~~diverso~~ *uguale* rispetto ad altri ospedali britannici?

Nell'ospedale di Bristol, sono stati effettuati 143 interventi cardiocirurgici e sono stati registrati 41 decessi (27.8%). Negli altri ospedali britannici, a fronte di 3176 operazioni sono stati registrati 356 decessi (10.7%).

2. Scelgo un test statistico che stimi qualcosa che, se abbastanza estremo, mi faccia dubitare di  $\mathcal{H}_0$

Pearson's  $\chi^2$  test per dati categorici

# Pearson's $\chi^2$ test



Il numero di bambini deceduti a seguito di interventi cardiocirurgici a Bristol tra il 1984 e il 1995 è ~~diverso~~ *uguale* rispetto ad altri ospedali britannici?

Nell'ospedale di Bristol, sono stati effettuati 143 interventi cardiocirurgici e sono stati registrati 41 decessi (27.8%). Negli altri ospedali britannici, a fronte di 3176 operazioni sono stati registrati 356 decessi (10.7%).

3. Genero la distribuzione campionaria del test scelto, assumendo  $\mathcal{H}_0$  vera

# Pearson's $\chi^2$ test

- Il numero di bambini deceduti a seguito di interventi cardiocirurgici a Bristol tra il 1984 e il 1995 è ~~diverso~~ *uguale* rispetto ad altri ospedali britannici?

Nell'ospedale di Bristol, sono stati effettuati 143 interventi cardiocirurgici e sono stati registrati 41 decessi (27.8%). Negli altri ospedali britannici, a fronte di 3176 operazioni sono stati registrati 356 decessi (10.7%).

- ? Come completiamo questa tabella di contingenza?

Ospedale/Deceduti	Si	No	Totale
Bristol			
Altri			
Totale			



# Pearson's $\chi^2$ test



Il numero di bambini deceduti a seguito di interventi cardiocirurgici a Bristol tra il 1984 e il 1995 è ~~diverso~~ *uguale* rispetto ad altri ospedali britannici?

Nell'ospedale di Bristol, sono stati effettuati 143 interventi cardiocirurgici e sono stati registrati 41 decessi (27.8%). Negli altri ospedali britannici, a fronte di 3176 operazioni sono stati registrati 356 decessi (10.7%).

Valori osservati

Ospedale/Deceduti	Si	No	Totale
Bristol	41	102	143
Altri	356	2820	3176
Totale	397	2922	3319

# Pearson's $\chi^2$ test



Il numero di bambini deceduti a seguito di interventi cardiocirurgici a Bristol tra il 1984 e il 1995 è ~~diverso~~ *uguale* rispetto ad altri ospedali britannici?

$$\Pi = \frac{tot_{decessi}}{tot_{interventi}} = \frac{397}{3319} = 0.1196$$

Valori osservati

Ospedale/Deceduti	Si	No	Totale
Bristol	41	102	143
Altri	356	2820	3176
Totale	397	2922	3319

# Pearson's $\chi^2$ test

- 📌 Il numero di bambini deceduti a seguito di interventi cardiocirurgici a Bristol tra il 1984 e il 1995 è ~~diverso~~ *uguale* rispetto ad altri ospedali britannici?

$$\Pi = \frac{tot_{decessi}}{tot_{interventi}} = \frac{397}{3319} = 0.1196$$

Valori osservati

Ospedale/Deceduti	Si	No	Totale
Bristol	41	102	143
Altri	356	2820	3176
Totale	397	2922	3319

Valori attesi

Ospedale/Deceduti	Si	No	Totale
Bristol	$143 \cdot 0.1196$		143
Altri	$3176 \cdot 0.1196$		3176
Totale	397	2922	3319

# Pearson's $\chi^2$ test



Il numero di bambini deceduti a seguito di interventi cardiocirurgici a Bristol tra il 1984 e il 1995 è ~~diverso~~ *uguale* rispetto ad altri ospedali britannici?

$$\Pi = \frac{tot_{decessi}}{tot_{interventi}} = \frac{397}{3319} = 0.1196$$

Valori osservati

Ospedale/Deceduti	Si	No	Totale
Bristol	41	102	143
Altri	356	2820	3176
Totale	397	2922	3319

Valori attesi

Ospedale/Deceduti	Si	No	Totale
Bristol	17.1		143
Altri	379.9		3176
Totale	397	2922	3319

# Pearson's $\chi^2$ test



Il numero di bambini deceduti a seguito di interventi cardiocirurgici a Bristol tra il 1984 e il 1995 è ~~diverso~~ *uguale* rispetto ad altri ospedali britannici?

$$\Pi = \frac{tot_{decessi}}{tot_{interventi}} = \frac{397}{3319} = 0.1196$$

Valori osservati

Ospedale/Deceduti	Si	No	Totale
Bristol	41	102	143
Altri	356	2820	3176
Totale	397	2922	3319

Valori attesi

Ospedale/Deceduti	Si	No	Totale
Bristol	17.1	125.9	143
Altri	379.9	2796.1	3176
Totale	397	2922	3319

# Pearson's $\chi^2$ test



Il numero di bambini deceduti a seguito di interventi cardiocirurgici a Bristol tra il 1984 e il 1995 è ~~diverso~~ *uguale* rispetto ad altri ospedali britannici?

Valori osservati

Ospedale/Deceduti	Si	No	Totale
Bristol	41	102	143
Altri	356	2820	3176
Totale	397	2922	3319

Valori attesi

Ospedale/Deceduti	Si	No	Totale
Bristol	17.1	125.9	143
Altri	379.9	2796.1	3176
Totale	397	2922	3319

$$\chi^2 = \sum \frac{(Osservati - Attesi)^2}{Attesi} = \frac{(41 - 17.1)^2}{17.1} + \frac{(102 - 125.9)^2}{125.9} + \frac{(356 - 379.9)^2}{379.9} + \frac{(2820 - 2796.1)^2}{2796.1} = 39.65$$

# Pearson's $\chi^2$ test



Il numero di bambini deceduti a seguito di interventi cardiocirurgici a Bristol tra il 1984 e il 1995 è ~~diverso~~ *uguale* rispetto ad altri ospedali britannici?

Valori osservati

Ospedale/Deceduti	Si	No	Totale
Bristol	41	102	143
Altri	356	2820	3176
Totale	397	2922	3319

Valori attesi

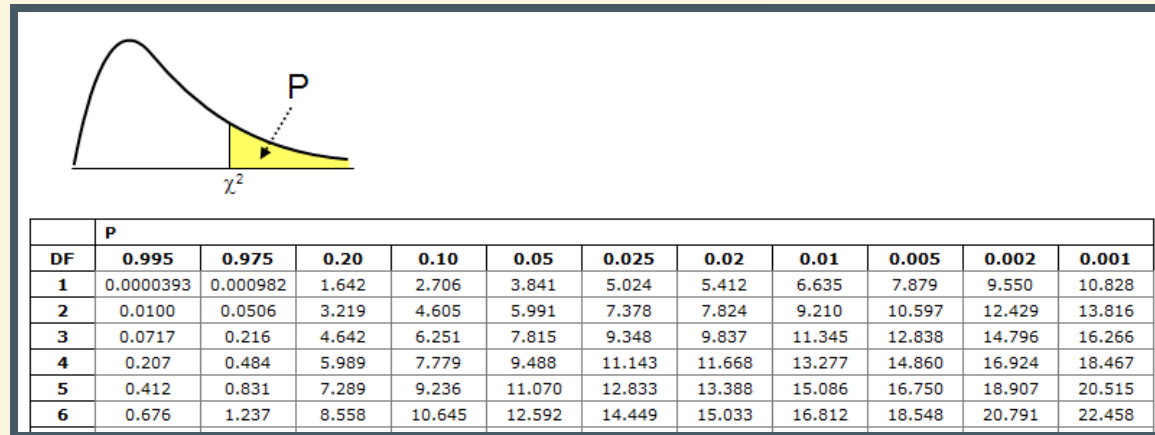
Ospedale/Deceduti	Si	No	Totale
Bristol	17.1	125.9	143
Altri	379.9	2796.1	3176
Totale	397	2922	3319

$$\chi^2 = \sum \frac{(Osservati - Attesi)^2}{Attesi} = \frac{(41 - 17.1)^2}{17.1} + \frac{(102 - 125.9)^2}{125.9} + \frac{(356 - 379.9)^2}{379.9} + \frac{(2820 - 2796.1)^2}{2796.1} = 39.65$$

$$df = (n_{\text{righe}} - 1) \times (n_{\text{colonne}} - 1) = 1$$

# Pearson's $\chi^2$ test

- 📌 Il numero di bambini deceduti a seguito di interventi cardiocirurgici a Bristol tra il 1984 e il 1995 è ~~diverso~~ *uguale* rispetto ad altri ospedali britannici?



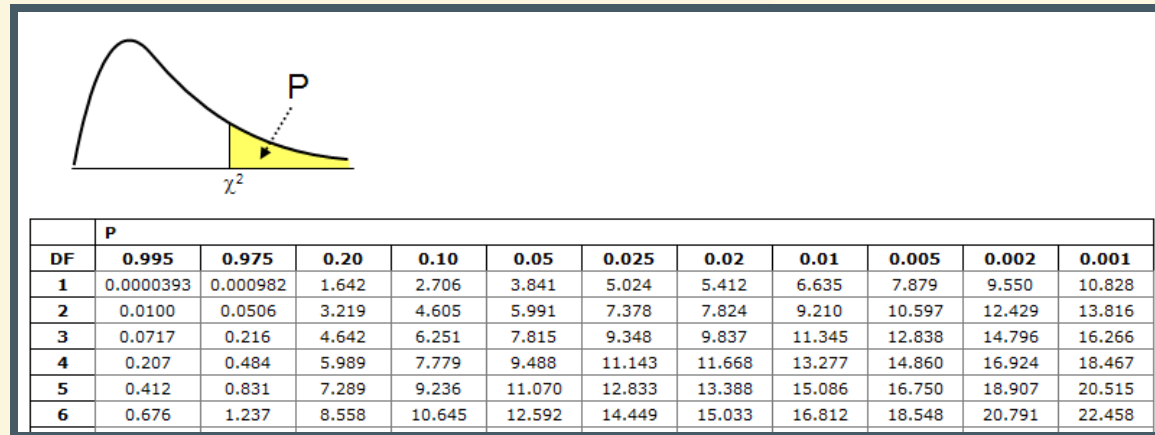
4. Verifico se la statistica osservata si trovi sulla coda di questa distribuzione e assegno una probabilità (P-value) a questo evento

$$\chi^2 = 39.65 \quad df = 1$$



# Pearson's $\chi^2$ test

- 📌 Il numero di bambini deceduti a seguito di interventi cardiocirurgici a Bristol tra il 1984 e il 1995 è ~~diverso~~ *uguale* rispetto ad altri ospedali britannici?



5. Dichiaro il risultato come statisticamente significativo se il P-value è inferiore a una soglia critica  $\alpha$

$$\chi^2 = 39.65 \quad df = 1 \quad \rightarrow \quad P < 0.001 = 3 \times 10^{-10} < \alpha = 0.05$$

# Pearson's $\chi^2$ test

 Il livello di istruzione influenza la frequenza dell'esercizio fisico?

Valori osservati

	Nessuno	Sporadico	Regolare	Totale
Licenza elementare	$f_{1,1}$	$f_{1,2}$	$f_{1,3}$	$\Sigma\text{Riga}_1$
Licenza media	$f_{1,2}$	...	...	$\Sigma\text{Riga}_2$
Diploma	$f_{1,3}$	...	...	$\Sigma\text{Riga}_3$
Laurea	$f_{1,4}$	...	...	$\Sigma\text{Riga}_4$
Totale	$\Sigma\text{Colonna}_1$	$\Sigma\text{Colonna}_2$	$\Sigma\text{Colonna}_3$	Totale

# Pearson's $\chi^2$ test

📌 Il livello di istruzione influenza la frequenza dell'esercizio fisico?

Valori attesi

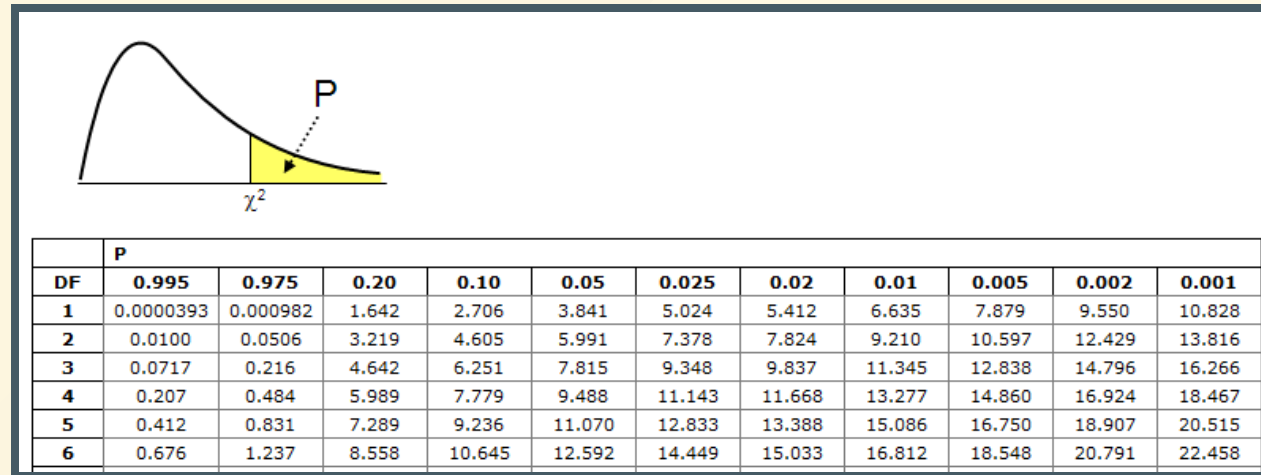
	Nessuno	Sporadico	Regolare	Totale
Licenza elementare	$\frac{\Sigma \text{Riga}_1 \times \Sigma \text{Colonna}_1}{\text{Totale}}$	$\frac{\Sigma \text{Riga}_1 \times \Sigma \text{Colonna}_2}{\text{Totale}}$	$\frac{\Sigma \text{Riga}_1 \times \Sigma \text{Colonna}_3}{\text{Totale}}$	$\Sigma \text{Riga}_1$
Licenza media	$\frac{\Sigma \text{Riga}_2 \times \Sigma \text{Colonna}_1}{\text{Totale}}$	...	...	$\Sigma \text{Riga}_2$
Diploma	$\frac{\Sigma \text{Riga}_3 \times \Sigma \text{Colonna}_1}{\text{Totale}}$	...	...	$\Sigma \text{Riga}_3$
Laurea	$\frac{\Sigma \text{Riga}_4 \times \Sigma \text{Colonna}_1}{\text{Totale}}$	...	...	$\Sigma \text{Riga}_4$
Totale	$\Sigma \text{Colonna}_1$	$\Sigma \text{Colonna}_2$	$\Sigma \text{Colonna}_3$	Totale

? df = ?

# Esercizio 8

? Le donne hanno un diverso atteggiamento verso l'uso di anticoncezionali rispetto agli uomini?

Ad un questionario a cui hanno risposto 42 donne, 21 si sono dichiarate favorevoli, 6 contrarie e 15 incerte, all'uso di concezionali. Dei 58 uomini, i favorevoli, contrari e incerti sono stati invece 11, 24 e 23.



## Esercizio 8 -- Soluzione

? Le donne **NON** hanno un diverso atteggiamento verso l'uso di anticoncezionali rispetto agli uomini?

Ad un questionario a cui hanno risposto 42 donne, 21 si sono dichiarate favorevoli, 6 contrarie e 15 incerte, all'uso di concezionali. Dei 58 uomini, i favorevoli, contrari e incerti sono stati invece 11, 24 e 23.

Valori osservati

Risposta/Sesso	Donne	Uomini	Totale
Favorevole	21	11	32
Contrario	6	24	30
Incerto	15	23	38
Totale	42	58	100

## Esercizio 8 -- Soluzione

- ? Le donne **NON** hanno un diverso atteggiamento verso l'uso di anticoncezionali rispetto agli uomini?

Valori osservati

Risposta/Sesso	Donne	Uomini	Totale
Favorevole	21	11	32
Contrario	6	24	30
Incerto	15	23	38
Totale	42	58	100

## Esercizio 8 -- Soluzione

- ? Le donne **NON** hanno un diverso atteggiamento verso l'uso di anticoncezionali rispetto agli uomini?

Valori osservati

Risposta/Sesso	Donne	Uomini	Totale
Favorevole	21	11	32
Contrario	6	24	30
Incerto	15	23	38
Totale	42	58	100

Valori attesi

Risposta/Sesso	Donne	Uomini	Totale
Favorevole	$\frac{32 \times 42}{100}$	$\frac{32 \times 58}{100}$	32
Contrario	$\frac{30 \times 42}{100}$	$\frac{30 \times 58}{100}$	30
Incerto	$\frac{38 \times 42}{100}$	$\frac{38 \times 58}{100}$	38
Totale	42	58	100

## Esercizio 8 -- Soluzione

- ? Le donne **NON** hanno un diverso atteggiamento verso l'uso di anticoncezionali rispetto agli uomini?

Valori osservati

Risposta/Sesso	Donne	Uomini	Totale
Favorevole	21	11	32
Contrario	6	24	30
Incerto	15	23	38
Totale	42	58	100

Valori attesi

Risposta/Sesso	Donne	Uomini	Totale
Favorevole	13.44	18.56	32
Contrario	12.60	17.40	30
Incerto	15.96	22.04	38
Totale	42	58	100

$$\chi^2 = \sum \frac{(Osservati - Attesi)^2}{Attesi}$$



# Esercizio 8 -- Soluzione

? Le donne **NON** hanno un diverso atteggiamento verso l'uso di anticoncezionali rispetto agli uomini?

Valori osservati

Risposta/Sesso	Donne	Uomini	Totale
Favorevole	21	11	32
Contrario	6	24	30
Incerto	15	23	38
Totale	42	58	100

Valori attesi

Risposta/Sesso	Donne	Uomini	Totale
Favorevole	13.44	18.56	32
Contrario	12.60	17.40	30
Incerto	15.96	22.04	38
Totale	42	58	100

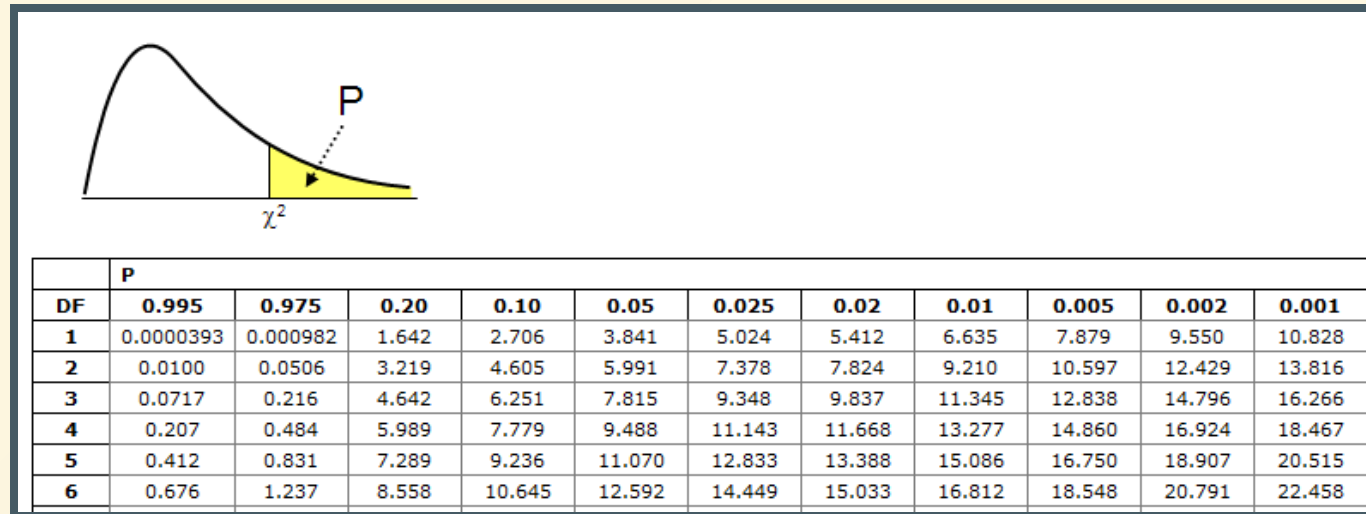
$$\chi^2 = \frac{(21-13.44)^2}{13.44} + \frac{(11-18.56)^2}{18.56} + \frac{(6-12.6)^2}{12.6} + \frac{(24-17.4)^2}{17.4} + \frac{(15-15.96)^2}{15.96} + \frac{(23-22.04)^2}{22.04} = 13.39$$

$$df = (n_{\text{righe}} - 1) \times (n_{\text{colonne}} - 1) = 2$$

# Esercizio 8 -- Soluzione

? Le donne **NON** hanno un diverso atteggiamento verso l'uso di anticoncezionali rispetto agli uomini?

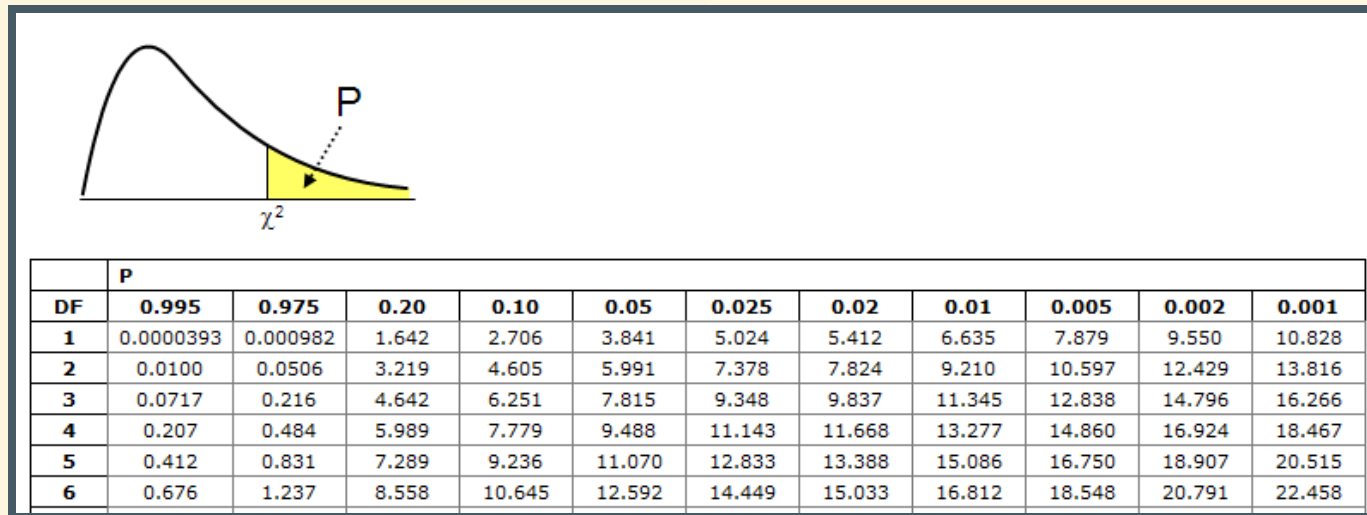
$$\chi^2 = 13.39, \quad df = 2$$



# Esercizio 8 -- Soluzione

? Le donne **NON** hanno un diverso atteggiamento verso l'uso di anticoncezionali rispetto agli uomini?

$$\chi^2 = 13.39, \quad df = 2 \rightarrow P < 0.002 = 0.0012 < \alpha = 0.05$$



# Rischio relativo, rischio assoluto, odds ratio

## Could you quantify the risk of eating red meat and processed meat?

The consumption of processed meat was associated with small increases in the risk of cancer in the studies reviewed. In those studies, the risk generally increased with the amount of meat consumed. An analysis of data from 10 studies estimated that every 50 gram portion of processed meat eaten daily increases the risk of colorectal cancer by about 18%.

<https://www.who.int/news-room/questions-and-answers/item/cancer-carcinogenicity-of-the-consumption-of-red-meat-and-processed-meat>

# Rischio relativo, rischio assoluto, odds ratio

## Could you quantify the risk of eating red meat and processed meat?

The consumption of processed meat was associated with small increases in the risk of cancer in the studies reviewed. In those studies, the risk generally increased with the amount of meat consumed. An analysis of data from 10 studies estimated that every 50 gram portion of processed meat eaten daily increases the risk of colorectal cancer by about 18%.

Rischio relativo (RR) = 18%

# Rischio relativo, rischio assoluto, odds ratio

## Could you quantify the risk of eating red meat and processed meat?

The consumption of processed meat was associated with small increases in the risk of cancer in the studies reviewed. In those studies, the risk generally increased with the amount of meat consumed. An analysis of data from 10 studies estimated that every 50 gram portion of processed meat eaten daily increases the risk of colorectal cancer by about 18%.

Rischio relativo (RR) = 18%

Rischio assoluto (baseline) = 6% o 1/16  $\rightarrow Odds = 6/94$

Rischio assoluto (+50g/giorno) = 7% o 1/14  $\rightarrow Odds = 7/93$

# Rischio relativo, rischio assoluto, odds ratio

## Could you quantify the risk of eating red meat and processed meat?

The consumption of processed meat was associated with small increases in the risk of cancer in the studies reviewed. In those studies, the risk generally increased with the amount of meat consumed. An analysis of data from 10 studies estimated that every 50 gram portion of processed meat eaten daily increases the risk of colorectal cancer by about 18%.

Rischio relativo (RR) = 18%

Rischio assoluto (baseline) = 6% o 1/16  $\rightarrow Odds = 6/94$

Rischio assoluto (+50g/giorno) = 7% o 1/14  $\rightarrow Odds = 7/93$

$$\text{Odds ratio (OR)} = \frac{\text{Odds}_{+50\text{g/giorno}}}{\text{Odds}_{\text{baseline}}} = \frac{7/93}{6/94} = 1.18$$

# Rischio relativo, rischio assoluto, odds ratio

Approccio	Baseline	+50g/giorno
Event rate	6%	7%
Expected frequency	6 su 100 1 in 16	7 su 100 1 in 14
Odds	6/94	7/93

Comparazione	
Odds ratio	$\frac{(7/93)}{(6/94)} = 1.18$
Relative risk	1.18 aumento del 18%
Absolute risk difference	1% o 1 su 100
Number Needed to Treat	100



# Esercizio #9

? Uno studio di popolazione<sup>1</sup> ha osservato che nella popolazione Svedese, in circa 3000 uomini con il livello di educazione più basso sono stati diagnosticati 5 tumori al cervello, mentre in altrettanti uomini con il livello di educatione più alto, ne sono stati diagnosticati 6.

Calcolate le seguenti quantità:

- a) Frequenza nei due gruppi
- b) Odds ratio (baseline: livello di istruzione basso)
- c) Rischio relativo

05:00

<sup>1</sup> Khanolkar et al., *Socioeconomic Position and the Risk of Brain Tumour: A Swedish National Population-Based Cohort Study*, Journal of Epidemiology and Community Health, 2016, doi:10.1136/jech-2015-207002

## Esercizio #9 -- Soluzione

- ? Uno studio di popolazione<sup>1</sup> ha osservato che nella popolazione Svedese, in circa 3000 uomini con il livello di educazione più basso sono stati diagnosticati 5 tumori al cervello, mentre in altrettanti uomini con il livello di educatione più alto, ne sono stati diagnosticati 6.

Calcolate le seguenti quantità:

a) Frequenza nei due gruppi

$$f_{basso} = 5 \text{ su } 3000 = 1 \text{ su } 600$$

$$f_{alto} = 6 \text{ su } 3000 = 1 \text{ su } 500$$

## Esercizio #9 -- Soluzione

- ? Uno studio di popolazione<sup>1</sup> ha osservato che nella popolazione Svedese, in circa 3000 uomini con il livello di educazione più basso sono stati diagnosticati 5 tumori al cervello, mentre in altrettanti uomini con il livello di educatione più alto, ne sono stati diagnosticati 6.

Calcolate le seguenti quantità:

a) Frequenza nei due gruppi

$$f_{basso} = 5 \text{ su } 3000 = 1 \text{ su } 600 \quad f_{alto} = 6 \text{ su } 3000 = 1 \text{ su } 500$$

b) Odds ratio (baseline: livello di istruzione basso)

$$\frac{(1/499)}{(1/599)} = 1.20$$

## Esercizio #9 -- Soluzione

- ? Uno studio di popolazione<sup>1</sup> ha osservato che nella popolazione Svedese, in circa 3000 uomini con il livello di educazione più basso sono stati diagnosticati 5 tumori al cervello, mentre in altrettanti uomini con il livello di educatione più alto, ne sono stati diagnosticati 6.

Calcolate le seguenti quantità:

a) Frequenza nei due gruppi

$$f_{basso} = 5 \text{ su } 3000 = 1 \text{ su } 600 \quad f_{alto} = 6 \text{ su } 3000 = 1 \text{ su } 500$$

b) Odds ratio (baseline: livello di istruzione basso)

$$\frac{(1/499)}{(1/599)} = 1.20$$

c) Rischio relativo

1.20 o un aumento del 20%

# Esercizio #10

? Il tasso di mortalità nella Marina Militare americana durante la guerra ispano-americana (1898) era del nove per mille. Per i civili nella città di New York, nello stesso periodo, era il sedici per mille.

Queste cifre ci dicono che era più sicuro essere nella Marina che fuori.


- a) vero
- b) falso
- c) non ho abbastanza elementi per decidere
- d) non è possibile rispondere

02:00

# Esercizio #10 -- Soluzione

- ? Il tasso di mortalità nella Marina Militare americana durante la guerra ispano-americana (1898) era del nove per mille. Per i civili nella città di New York nello stesso periodo era il sedici per mille.

Queste cifre ci dicono che era più sicuro essere nella Marina che fuori.

- a) vero
- b) falso
- c) non ho abbastanza elementi per decidere
- d) non è possibile rispondere 

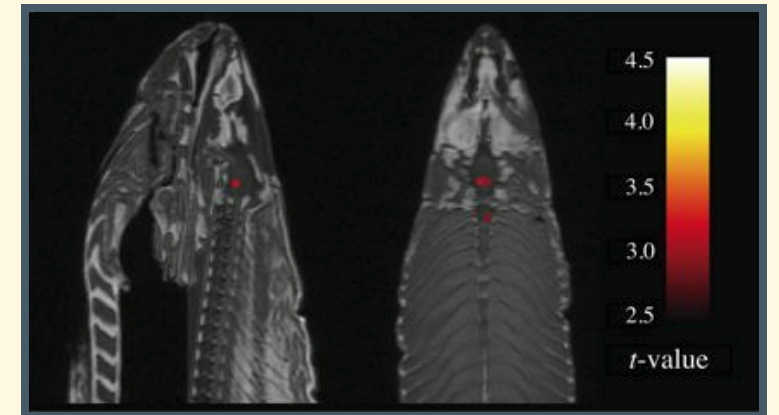
# Quando facciamo più di un test alla volta?

- 📌 Un gruppo di ricerca ha effettuato fMRI su un singolo soggetto (\*) mentre gli venivano mostrate delle fotografie in cui le persone fotografate esprimevano diverse emozioni. Sedici regioni cerebrali risultano statisticamente significative con  $P < 0.001$ .

# Quando facciamo più di un test alla volta?

- 📌 Un gruppo di ricerca ha effettuato fMRI su un singolo soggetto (\*) mentre gli venivano mostrate delle fotografie in cui le persone fotografate esprimevano diverse emozioni. Sedici regioni cerebrali risultano statisticamente significative con  $P < 0.001$ .

(\*) Atlantic salmon, '*not alive at the time of scanning*'





# Quando facciamo più di un test alla volta?



$P = 0.05 \rightarrow 5\%$  rifiutiamo  $\mathcal{H}_0$  anche se è vera

$$\mathcal{P} = 1 - 0.95 = 0.05$$

# Quando facciamo più di un test alla volta?



$P = 0.05 \rightarrow 5\%$  rifiutiamo  $\mathcal{H}_0$  anche se è vera

$$\mathcal{P} = 1 - 0.95 = 0.05$$

Con 2 test, averne almeno uno con  $P < 0.05$  è:

$$\mathcal{P} = 1 - 0.95 \times 0.95 = 1 - 0.95^2 = 0.0975 \rightarrow \approx 10\%$$

Con 3 test, averne almeno uno con  $P < 0.05$  è:

$$\mathcal{P} = 1 - 0.95^3 = 0.145 \rightarrow \approx 14\%$$

Con 10 test, averne almeno uno con  $P < 0.05$  è:

$$\mathcal{P} = 1 - 0.95^{10} = 0.40 \rightarrow \approx 40\%$$

# Correzione per test multipli



Quando si fanno più test, si richiede un  $\alpha$  inferiore

**Bonferroni-correction:**  $\alpha = \frac{0.05}{N_{\text{test}}}$

Con 10 test, averne almeno uno con  $P < \frac{0.05}{10} = 0.005$  è

$$\mathcal{P} = 1 - 0.995^{10} = 0.049 \rightarrow \approx 5\%$$

# Correzione per test multipli



Quando si fanno più test, si richiede un  $\alpha$  inferiore

Quando si fanno più test, si fissa il numero di 'scoperte che sono false

**False discovery rate (FDR, Benjamini-Hochberg procedure):**

1. ordino i risultati per P value crescente
2. Rifiuto  $\mathcal{H}_0$  sino a che  $P_{(k)} > \alpha \times \frac{k}{N_{\text{test}}}$

# Errori dei test statistici

$\mathcal{H}_0$ è	Non rifiutata	Rifiutata
Vera		
Falsa		

# Errori dei test statistici

$\mathcal{H}_0$ è	Non rifiutata	Rifiutata
Vera		Falso Positivo
Falsa	Falso negativo	

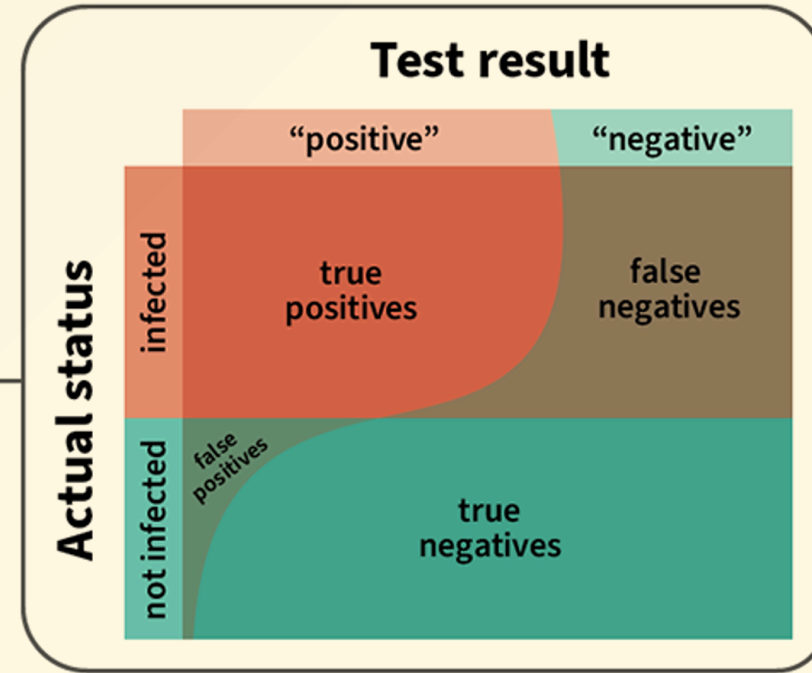
# Errori dei test statistici

Sospetto è	Assolto	Condannato
Innocente		Condanno un innocente
Colpevole	Assolvo un colpevole	

# Errori dei test statistici

The COVID-19 swab test is highly **specific** but not as **sensitive**.

That means a positive result is almost always true, but a negative result is sometimes false.



$$\text{Sensitivity} = \frac{\text{number of true positives}}{\text{number of those tested who really are infected}} = \text{“how many of the infections did we find?”}$$

$$\text{Specificity} = \frac{\text{number of true negatives}}{\text{number of those tested who really are not infected}} = \text{“how many of the healthy people did we clear?”}$$



# Esercizio #11

? In un villaggio, c'era un pastorello che faceva la guardia alle pecore. Annoiandosi, per diverse notti, si mise ad urlare "Al lupo! Al lupo!", così tutti accorrevano per aiutarlo. Una notte, un lupo venne veramente. Il pastorello cominciò a gridare: "Al lupo, al lupo!", ma nessuno venne perché tutti pensarono che fosse uno scherzo.

Che tipo di errore stanno facendo gli abitanti del villaggio?

- a) Errore del primo tipo, poi del secondo tipo
- b) Errore del secondo tipo, poi del primo tipo
- c) Errore nullo, poi errore alternativo
- d) Errore alternativo, poi errore nullo
- e) Nessuno dei precedenti

# Esercizio #11 -- Soluzione

? In un villaggio, c'era un pastorello che faceva la guardia alle pecore. Annoiandosi, per diverse notti, si mise ad urlare "Al lupo! Al lupo!", così tutti accorrevano per aiutarlo. Una notte, un lupo venne veramente. Il pastorello cominciò a gridare: "Al lupo, al lupo!", ma nessuno venne perché tutti pensarono che fosse uno scherzo.

Che tipo di errore stanno facendo gli abitanti del villaggio?

- a) Errore del primo tipo, poi del secondo tipo ☒
- b) Errore del secondo tipo, poi del primo tipo
- c) Errore nullo, poi errore alternativo
- d) Errore alternativo, poi errore nullo
- e) Nessuno dei precedenti

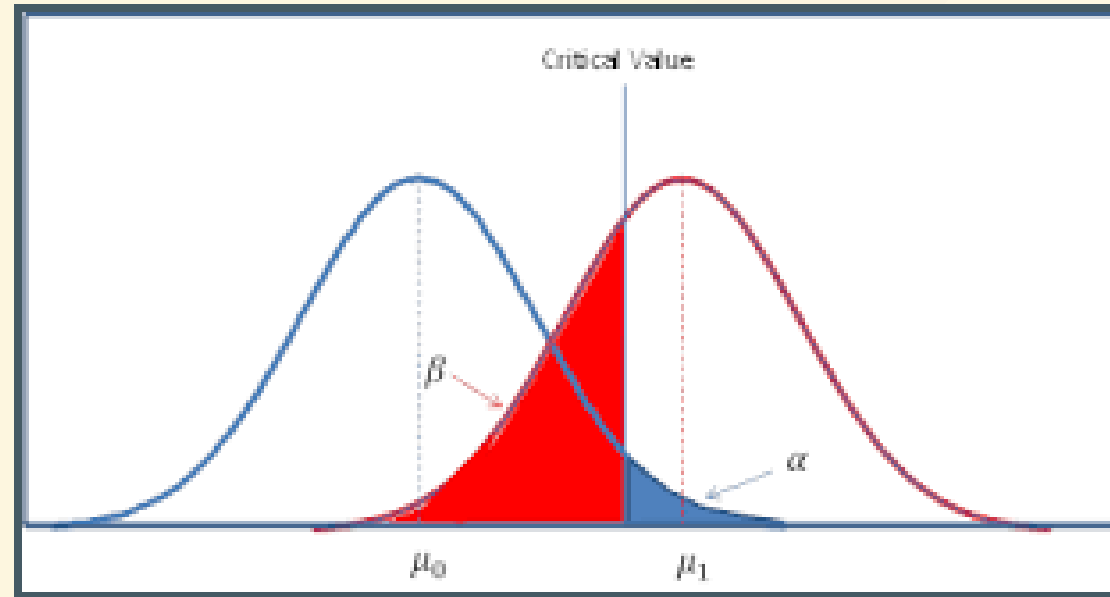
# Errori dei test statistici

$\mathcal{H}_0$ è	Non rifiutata	Rifiutata
Vera		$\alpha$
Falsa	$\beta$	

# La potenza di un test

$\mathcal{H}_0$ è	Non rifiutata	Rifiutata
Vera		$\alpha$
Falsa	$\beta$	$1 - \beta$

# La potenza di un test



- $\alpha = 0.05$
- $1 - \beta = 0.8$



# Esercizio #12

? Voglio aumentare la potenza del mio studio. Quali fattori posso modificare?

- a) il livello di significatività  $\alpha$
- b) la differenza  $\mu_i - \mu_c$
- c) la deviazione standard ( $\sigma^2$ ) dei due campioni
- d) la dimensione  $n$  dei due campioni
- e) nessuna delle precedenti

## Esercizio #12 -- Soluzione

? Voglio aumentare la potenza del mio studio. Quali fattori posso modificare?

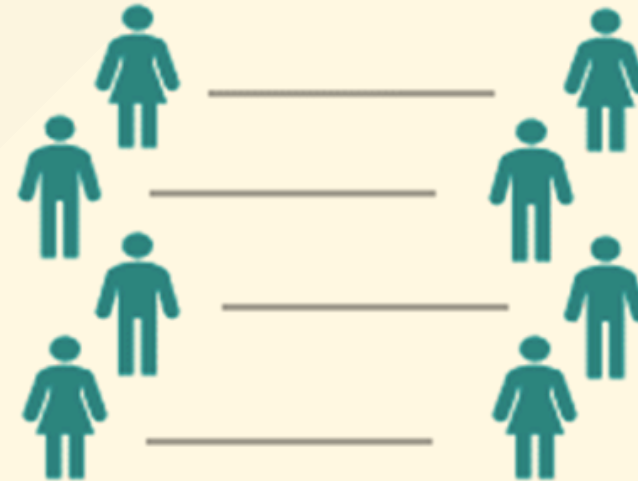
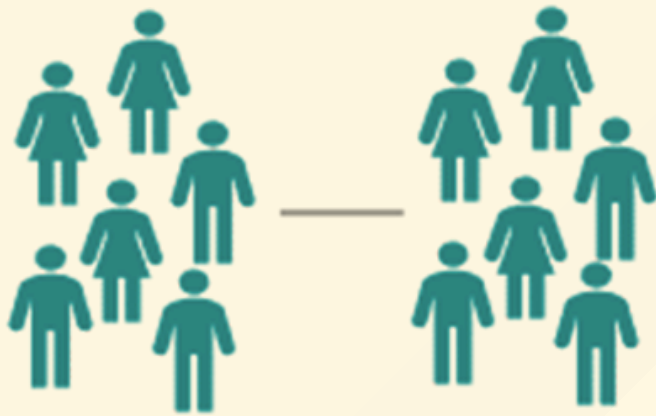
- a) il livello di significatività  $\alpha$  
- b) la differenza  $\mu_i - \mu_c$
- c) la deviazione standard ( $\sigma^2$ ) dei due campioni
- d) la dimensione  $n$  dei due campioni 
- e) nessuna delle precedenti

# Campioni indipendenti & dipendenti





# Campioni indipendenti & dipendenti



# Test non-parametrici

Campione	Tipo del dato	$\mathcal{H}_0$	Test non parametrico
Indipendenti	Numerici	$\mu_1 = \mu_2$	Mann-Whitney's test
Dipendenti	Numerici	$\mu_1 = \mu_2$	Wilcoxon's test
Indipendenti	Categoriche	$\pi_1 = \pi_2$	Fisher's test
Dipendenti	Categoriche	$\pi_1 = \pi_2$	McNemar's test

# Cosa abbiamo imparato in questa lezione?

- P-value misura l'incompatibilità tra i dati e la nostra ipotesi (probabilità di osservare valori così estremi se  $\mathcal{H}_0$  è vera)
- Tradizionalmente,  $P < 0.05$  o  $< 0.01$  sono considerati statisticamente significativi, ma queste soglie devono essere corrette per il numero di test
- C'è una corrispondenza tra CI e P-value, e se il 95% CI non include lo zero, possiamo rifiutare  $\mathcal{H}_0$  a un livello di significatività  $\alpha = 0.05$
- Errori del primo tipo dipendono dalla soglia di significatività  $\alpha$
- Esiste un legame tra errori del secondo tipo  $\beta$  e potenza di uno studio
- Per dati con distribuzioni non-normali possiamo usare test non parametrici