Introduction to statistics

(Day 3)

Agenda

- **Where:**
  - Mar 4: ✅
  - Mar 5: ✅
  - Mar 6 : Online

- **When:**
  - 14-17
  - 1 coffee break

- **Who:**
  - Paola Dalmasso
    paola.dalmasso@unito.it
  - Alessia Visconti
    alessia.visconti@unito.it

- **How (to pass):**
  - Attend at least 2 lessons

Recap

Recap

- When can't study a population, we select a representative sample
- There are different sampling strategies

- There are different types of data
- Data are described with measures of centrality (mode, median, mean) and dispersion (range, IQR, standard deviation)
- Parameters (calculated on the population) *vs* statistics (calculated on the sample)

Recap

- Using the standard normal distribution we can calculate the probability of an observation
- Multiple phenomena and statistical distributions are normally distributed (CLT)

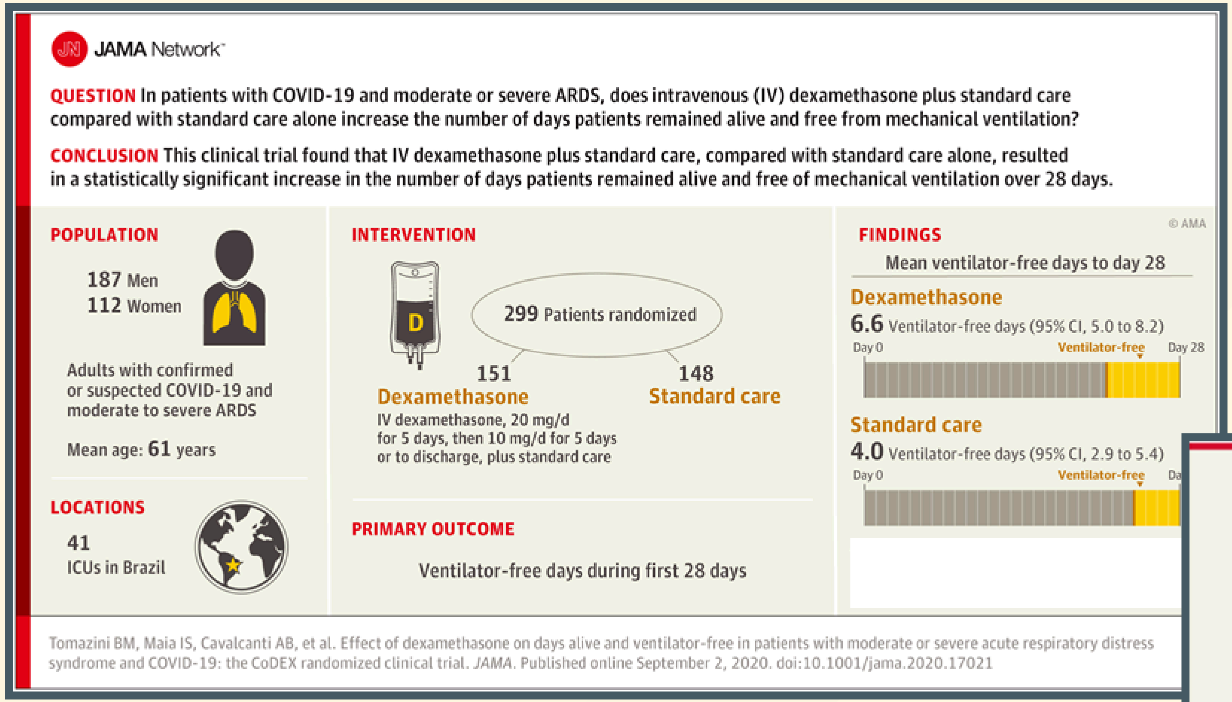
- We use statistics to estimate parameters (point estimates)
- We use interval estimates (confidence intervals) to estimate point estimates' uncertainty

Making decision with data

Learning objectives

- Make and test hypotheses
- Interpret P values
- Understand Type I and II errors
- Understand the power of a study

Making hypotheses

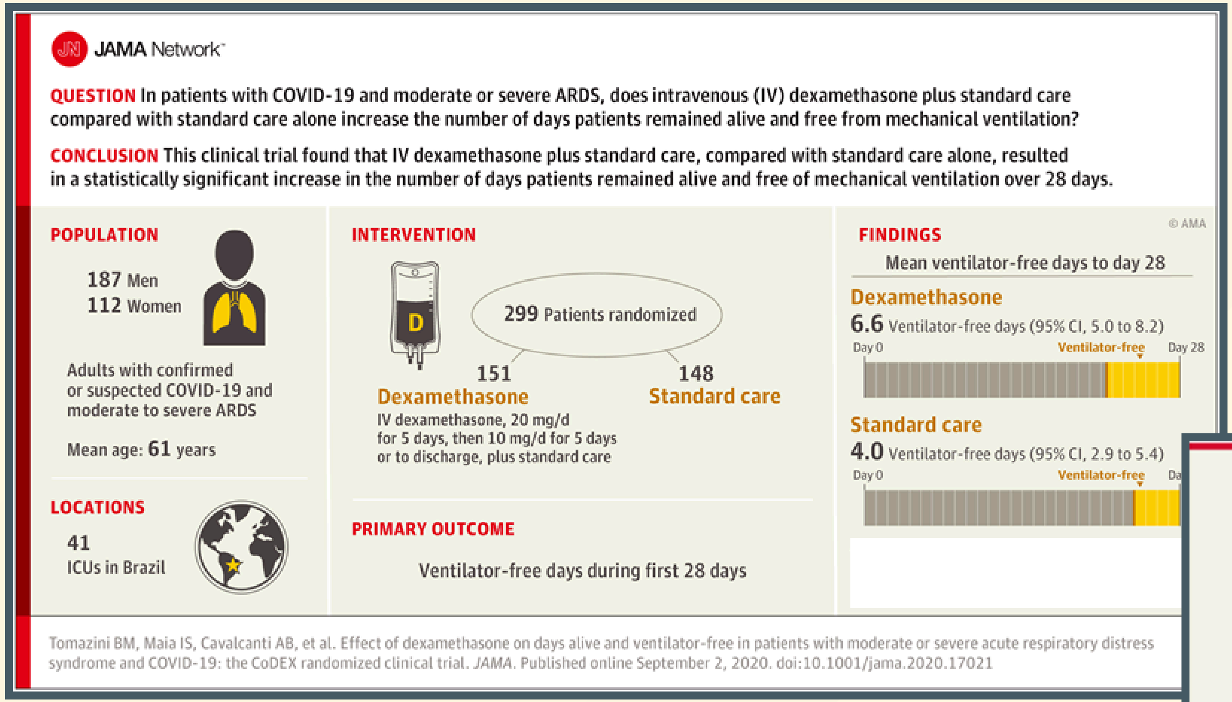


Key Points

Question In patients with coronavirus disease 2019 (COVID-19) and moderate or severe acute respiratory distress syndrome (ARDS), does intravenous dexamethasone plus standard care compared with standard care alone increase the number of days alive and free from mechanical ventilation?

📌 Is using Dexamethasone better than standard care?

Making hypotheses



Key Points

Question In patients with coronavirus disease 2019 (COVID-19) and moderate or severe acute respiratory distress syndrome (ARDS), does intravenous dexamethasone plus standard care compared with standard care alone increase the number of days alive and free from mechanical ventilation?

📌 Is Dexamethasone the same as standard care?

The Falsification Principle

The Falsification Principle

DINOSAUR EVOLUTION

A Jurassic ornithischian dinosaur from Siberia with both feathers and scales

Pascal Godefroit,[1*] Sofia M. Sinitsa,[2] Danielle Dhouailly,[3] Yuri L. Bolotsky,[4] Alexander V. Sizov,[5] Maria E. McNamara,[6,7] Michael J. Benton,[7] Paul Spagna[1]

Making hypotheses

📌 Is Dexamethasone the same as standard care?

$$n_{\mathrm{i}} = 151, \bar{x}_{\mathrm{i}} = 6.6, s_{\mathrm{i}} = 10.0$$
$$n_{\mathrm{c}} = 148, \bar{x}_{\mathrm{c}} = 4.0, s_{\mathrm{c}} = 8.7$$

$$\mu_{\mathrm{i}} - \mu_{\mathrm{c}} = 0$$
$$\rightarrow \text{Null hypothesis } (\mathcal{H}_0)$$

Key Points

Question In patients with coronavirus disease 2019 (COVID-19) and moderate or severe acute respiratory distress syndrome (ARDS), does intravenous dexamethasone plus standard care compared with standard care alone increase the number of days alive and free from mechanical ventilation?

Making hypotheses

📌 Is Dexamethasone the same as standard care?

$$n_{\mathrm{i}} = 151, \bar{x}_{\mathrm{i}} = 6.6, s_{\mathrm{i}} = 10.0$$
$$n_{\mathrm{c}} = 148, \bar{x}_{\mathrm{c}} = 4.0, s_{\mathrm{c}} = 8.7$$

$$\mu_{\mathrm{i}} - \mu_{\mathrm{c}} = 0$$
$$\rightarrow \text{Null hypothesis } (\mathcal{H}_0)$$

Key Points

Question In patients with coronavirus disease 2019 (COVID-19) and moderate or severe acute respiratory distress syndrome (ARDS), does intravenous dexamethasone plus standard care compared with standard care alone increase the number of days alive and free from mechanical ventilation?

$$\mu_{\mathrm{i}} - \mu_{\mathrm{c}} \neq 0$$
$$\rightarrow \text{Alternative hypothesis } (\mathcal{H}_1/\mathcal{H}_A)$$

Testing hypotheses

📌 Is Dexamethasone the same as standard care?

$$n_\mathrm{i} = 151, \bar{x}_\mathrm{i} = 6.6, s_\mathrm{i} = 10.0$$
$$n_\mathrm{c} = 148, \bar{x}_\mathrm{c} = 4.0, s_\mathrm{c} = 8.7$$

$$\mu_\mathrm{i} - \mu_\mathrm{c} = 0 \quad \leftarrow$$
$$\mu_\mathrm{i} - \mu_\mathrm{c} \neq 0$$

Testing hypotheses

📌 Is Dexamethasone the same as standard care?

$$n_{\mathrm{i}} = 151, \bar{x}_{\mathrm{i}} = 6.6, s_{\mathrm{i}} = 10.0$$
$$n_{\mathrm{c}} = 148, \bar{x}_{\mathrm{c}} = 4.0, s_{\mathrm{c}} = 8.7$$

$$\bar{X} = \bar{x}_i - \bar{x}_c = 2.6$$
$$\hat{\mathrm{SE}} = \sqrt{\frac{s_{\mathrm{i}}^2}{n_{\mathrm{i}}} + \frac{s_{\mathrm{i}}^2}{n_{\mathrm{i}}}} = 1.08$$

$$z = \frac{\bar{X} - \mu}{\hat{SE}} = \frac{2.6 - 0}{1.08} = 2.4$$

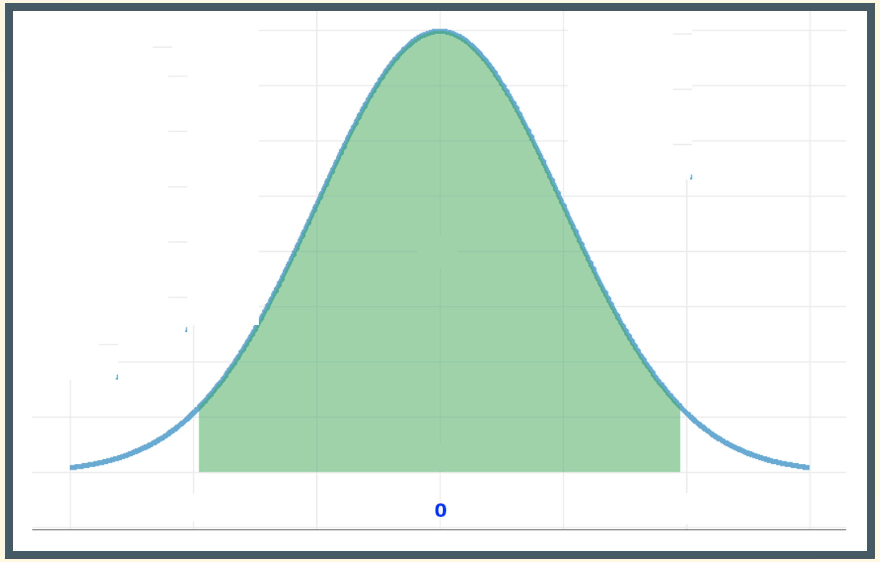
Testing hypotheses

📌 Is Dexamethasone the same as standard care?

$$n_{\mathrm{i}} = 151, \bar{x}_{\mathrm{i}} = 6.6, s_{\mathrm{i}} = 10.0$$
$$n_{\mathrm{c}} = 148, \bar{x}_{\mathrm{c}} = 4.0, s_{\mathrm{c}} = 8.7$$

$$\bar{X} = \bar{x}_i - \bar{x}_c = 2.6$$
$$\hat{\mathrm{SE}} = \sqrt{\frac{s_{\mathrm{i}}^2}{n_{\mathrm{i}}} + \frac{s_{\mathrm{i}}^2}{n_{\mathrm{i}}}} = 1.08$$

$$z = \frac{\bar{X} - \mu}{\hat{SE}} = \frac{2.6 - 0}{1.08} = 2.4$$



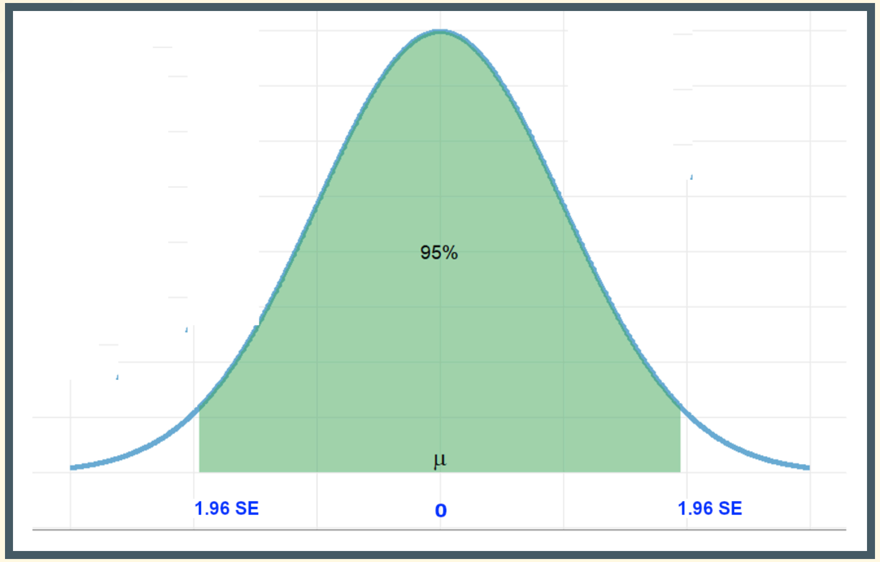
How far from $\mu = 0$ is too far to accept the null hypothesis with less than 5% chance?

Testing hypotheses

📌 Is Dexamethasone the same as standard care?

$$n_{\mathrm{i}} = 151, \bar{x}_{\mathrm{i}} = 6.6, s_{\mathrm{i}} = 10.0$$
$$n_{\mathrm{c}} = 148, \bar{x}_{\mathrm{c}} = 4.0, s_{\mathrm{c}} = 8.7$$

$$\bar{X} = \bar{x}_i - \bar{x}_c = 2.6$$
$$\hat{\mathrm{SE}} = \sqrt{\frac{s_{\mathrm{i}}^2}{n_{\mathrm{i}}} + \frac{s_{\mathrm{i}}^2}{n_{\mathrm{i}}}} = 1.08$$
$$z = \frac{\bar{X}-\mu}{\hat{SE}} = \frac{2.6-0}{1.08} = 2.4$$



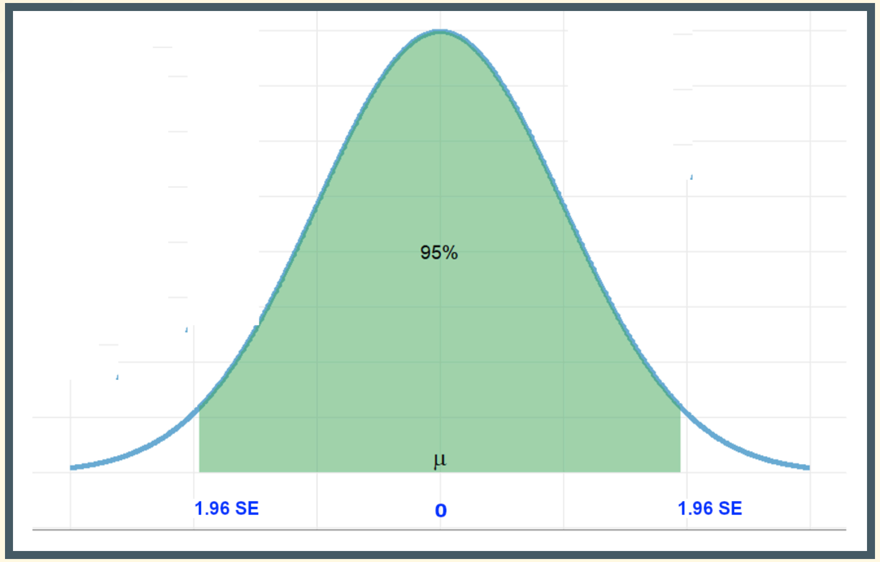
$$\alpha = 0.05$$
$$\rightarrow \pm 1.96 \times \hat{SE} = \pm 2.12$$

Testing hypotheses

📌 Is Dexamethasone the same as standard care?

$$n_{\mathrm{i}} = 151, \bar{x}_{\mathrm{i}} = 6.6, s_{\mathrm{i}} = 10.0$$
$$n_{\mathrm{c}} = 148, \bar{x}_{\mathrm{c}} = 4.0, s_{\mathrm{c}} = 8.7$$

$$\bar{X} = \bar{x}_i - \bar{x}_c = 2.6$$
$$\hat{\mathrm{SE}} = \sqrt{\frac{s_{\mathrm{i}}^2}{n_{\mathrm{i}}} + \frac{s_{\mathrm{i}}^2}{n_{\mathrm{i}}}} = 1.08$$
$$z = \frac{\bar{X}-\mu}{\hat{SE}} = \frac{2.6-0}{1.08} = 2.4$$



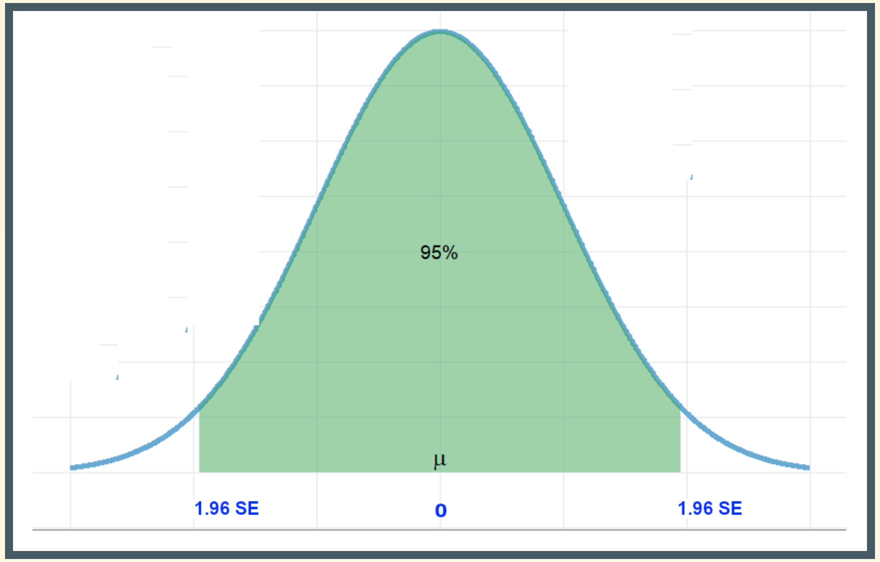
$$\alpha = 0.05$$
$$\rightarrow \pm 1.96 \times \hat{SE} = \pm 2.12$$

We refuse the null hypothesis

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

Significance

📌 How likely is that we made a mistake, *i.e.*, Dexamethasone differs from standard care, if we accepted $\mathcal{H}_0$?



Significance

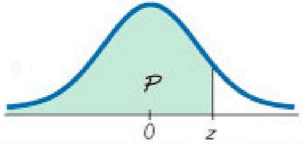
📌 How likely is that we made a mistake, *i.e.*, Dexamethasone differs from standard care, if we accepted $\mathcal{H}_0$?

$$\bar{X} = \bar{x}_i - \bar{x}_c = 2.6$$

$$\hat{\mathrm{SE}} = \sqrt{\frac{s_{\mathrm{i}}^2}{n_{\mathrm{i}}} + \frac{s_{\mathrm{i}}^2}{n_{\mathrm{i}}}} = 1.08$$

$$z = \frac{\bar{X} - \mu}{\hat{SE}} = \frac{2.6}{1.08} = 2.4$$

$$\mathcal{P}(\bar{X} \geq 2.4) = 1 - 0.9918 = \\ = 0.0082$$

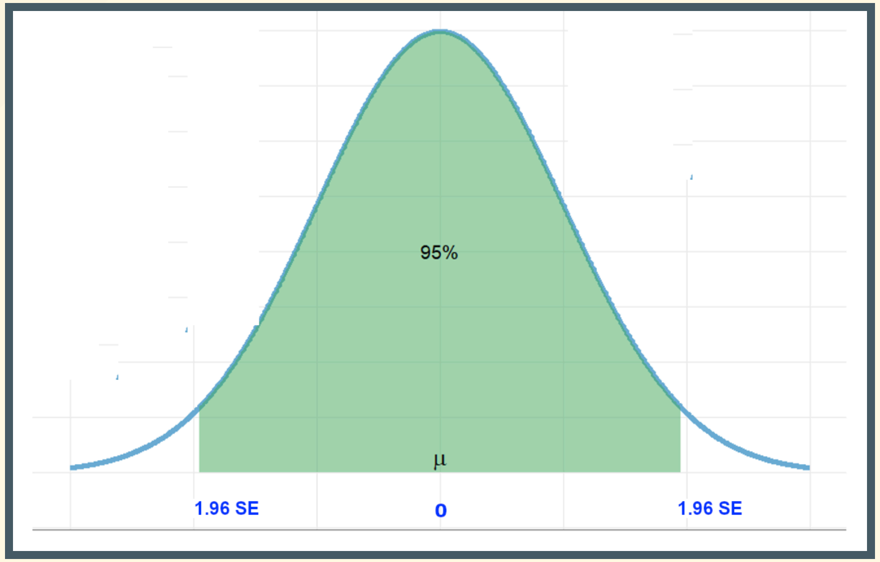


z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Significance

📌 How likely is that we made a mistake, *i.e.*, Dexamethasone differs from standard care, if we accepted $\mathcal{H}_0$?

$$\mu_{\mathrm{i}} - \mu_{\mathrm{c}} = 0 \;\rightarrow\; \mathcal{P}(|\bar{X}| \geq 2.4)$$



Significance

📌 How likely is that we made a mistake, *i.e.*, Dexamethasone differs from standard care, if we accepted $\mathcal{H}_0$?

$$\mathcal{P}(|\bar{X}| \geq 2.4) = 2 \times (1 - 0.9918) = 2 \times 0.0082 = 0.0164 \quad \leftarrow \quad \text{P value}$$

One- and two-tailed tests

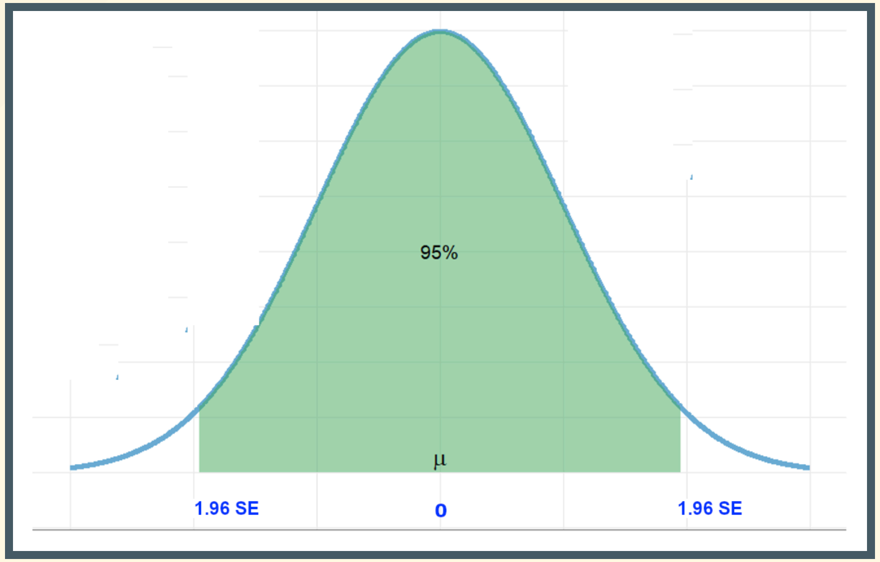
🎯 $\mathcal{H}_1\colon \quad \mu_\mathrm{i} - \mu_\mathrm{c} \neq 0$

$\rightarrow$ two-tailed test

$\mathcal{H}_1\colon \quad \mu_\mathrm{i} - \mu_\mathrm{c} > 0$

$\mu_\mathrm{i} - \mu_\mathrm{c} < 0$

$\rightarrow$ one-tailed test



Hypothesis testing (in steps)

1. Set $\mathcal{H}_0$ and $\mathcal{H}_1$

2. Define α

3. Calculate the test statistics and the P value

4. Make a decision about $\mathcal{H}_0$

Exercise 15

? Does the birth weight of babies born from smoking mothers differ from that of babies born from non-smoking mothers?

$$n_{\mathrm{s}} = 5065, \bar{x}_{\mathrm{s}} = 3241.6\ \mathrm{kg}, s_{\mathrm{s}} = 476.5\ \mathrm{kg}$$
$$n_{\mathrm{c}} = 8143, \bar{x}_{\mathrm{c}} = 3424.1\ \mathrm{kg}, s_{\mathrm{c}} = 474.6\ \mathrm{kg}$$

Exercise 15 -- Solution

? Does the birth weight of babies born from smoking mothers differ from that of babies born from non-smoking mothers?

$$n_{\mathrm{s}} = 5065, \bar{x}_{\mathrm{s}} = 3241.6 \text{ kg}, s_{\mathrm{s}} = 476.5 \text{ kg}$$
$$n_{\mathrm{c}} = 8143, \bar{x}_{\mathrm{c}} = 3424.1 \text{ kg}, s_{\mathrm{c}} = 474.6 \text{ kg}$$

1. Set $\mathcal{H}_0$ and $\mathcal{H}_1$ $\;\rightarrow\;$ $\mathcal{H}_0 : \mu_s - \mu_c = 0, \mathcal{H}_1 : \mu_s - \mu_c \neq 0$

Exercise 15 -- Solution

? Does the birth weight of babies born from smoking mothers differ from that of babies born from non-smoking mothers?

$$n_{\mathrm{s}} = 5065, \bar{x}_{\mathrm{s}} = 3241.6, s_{\mathrm{s}} = 476.5$$
$$n_{\mathrm{c}} = 8143, \bar{x}_{\mathrm{c}} = 3424.1 \text{ kg}, s_{\mathrm{c}} = 474.6 \text{ kg}$$

1. $\mathcal{H}_0 : \mu_s - \mu_c = 0$

2. Define $\alpha \quad \rightarrow \quad \alpha = 0.05$

Exercise 15 -- Solution

? $n_{\mathrm{s}} = 5065, \bar{x}_{\mathrm{s}} = 3241.6\text{ kg}, s_{\mathrm{s}} = 476.5\text{ kg}$
$n_{\mathrm{c}} = 8143, \bar{x}_{\mathrm{c}} = 3424.1\text{ kg}, s_{\mathrm{c}} = 474.6\text{ kg}$

1. $\mathcal{H}_0 : \mu_s - \mu_c = 0$
2. $\alpha = 0.05$

3. Calculate the test statistics $\downarrow$

$$\bar{X} = \bar{x}_s - \bar{x}_c = 3241.6 - 3424.1 = -182.5$$

$$\hat{\mathrm{SE}} = \sqrt{\frac{s_{\mathrm{s}}^2}{n_{\mathrm{s}}} + \frac{s_{\mathrm{c}}^2}{n_{\mathrm{c}}}} = \sqrt{\frac{476.5^2}{5065} + \frac{474.6^2}{8143}} = 8.51$$

$$\rightarrow \pm 1.96 \times \hat{SE} = \pm 1.96 \times 8.51 = \pm 16.68$$

Exercise 15 -- Solution

? $n_{\mathrm{s}} = 5065, \bar{x}_{\mathrm{s}} = 3241.6\text{ kg}, s_{\mathrm{s}} = 476.5\text{ kg}$
$n_{\mathrm{c}} = 8143, \bar{x}_{\mathrm{c}} = 3424.1\text{ kg}, s_{\mathrm{c}} = 474.6\text{ kg}$

1. $\mathcal{H}_0 : \mu_s - \mu_c = 0$
2. $\alpha = 0.05$

3. Calculate the test statistics $\downarrow$

$$\bar{X} = \bar{x}_s - \bar{x}_c = 3241.6 - 3424.1 = -182.5$$

$$\hat{\mathrm{SE}} = \sqrt{\frac{s_{\mathrm{s}}^2}{n_{\mathrm{s}}} + \frac{s_{\mathrm{c}}^2}{n_{\mathrm{c}}}} = \sqrt{\frac{476.5^2}{5065} + \frac{474.6^2}{8143}} = 8.51$$

$$\rightarrow \pm 1.96 \times \hat{SE} = \pm 1.96 \times 8.51 = \pm 16.68$$

$$z = \frac{\bar{X} - \mu}{\hat{\mathrm{SE}}} = \frac{-182.5 - 0}{8.51} = -21.44$$

$$\mathcal{P}(|\bar{X}| \geq 21.44) = 0 \times 2 = 0$$

Exercise 15 -- Solution

? $n_{\mathrm{s}} = 5065, \bar{x}_{\mathrm{s}} = 3241.6 \text{ kg}, s_{\mathrm{s}} = 476.5 \text{ kg}$
$n_{\mathrm{c}} = 8143, \bar{x}_{\mathrm{c}} = 3424.1 \text{ kg}, s_{\mathrm{c}} = 474.6 \text{ kg}$

1. $\mathcal{H}_0 : \mu_s - \mu_c = 0$
2. $\alpha = 0.05$
3. $\pm 1.96 \times \hat{SE} = 1.96 \times 8.51 = \pm 16.68$
   $z = \frac{\bar{X}-\mu}{\hat{\mathrm{SE}}} = \frac{-182.5-0}{8.51} = -21.44$
   $\mathcal{P}(|\bar{X}| \geq 21.44) = 0 \times 2 = 0$

4. Make a decision about $\mathcal{H}_0 \rightarrow$ We refuse $\mathcal{H}_0$

Making decision

🎯 $p < \alpha \rightarrow \text{reject } \mathcal{H}_0$
$p \geq \alpha \rightarrow \text{fail to reject } \mathcal{H}_0$

$$\alpha = 0.10 \rightarrow P < 0.10 \rightarrow \pm\, 1.65 \times \hat{SE}$$
$$\alpha = 0.05 \rightarrow P < 0.05 \rightarrow \pm\, 1.96 \times \hat{SE}$$
$$\alpha = 0.01 \rightarrow P < 0.01 \rightarrow \pm\, 2.58 \times \hat{SE}$$

Multiple testing correction

📌 We tested the mean ventilator-free days to day 28, but what if we tested M multiple outcomes?

Multiple testing correction

📌 We tested the mean ventilator-free days to day 28, but what if we tested M multiple outcomes?

- Bonferroni correction: $\alpha = \alpha / M$
- False discovery rate (FDR)

Testing differences in proportion

- 📌 Is zidovudine the same as standard care?

$$n_\mathrm{i} = 180, m_\mathrm{i} = 13$$
$$n_\mathrm{c} = 183, m_\mathrm{c} = 40$$

The New England
Journal of Medicine

Volume 331 NOVEMBER 3, 1994 Number 18

REDUCTION OF MATERNAL–INFANT TRANSMISSION OF HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 WITH ZIDOVUDINE TREATMENT

Edward M. Connor, M.D., Rhoda S. Sperling, M.D., Richard Gelber, Ph.D., Pavel Kiselev, Ph.D., Gwendolyn Scott, M.D., Mary Jo O'Sullivan, M.D., Russell VanDyke, M.D., Mohammed Bey, M.D., William Shearer, M.D., Ph.D., Robert L. Jacobson, M.D., Eleanor Jimenez, M.D., Edward O'Neill, M.D., Brigitte Bazin, M.D., Jean-François Delfraissy, M.D., Mary Culnane, M.S., Robert Coombs, M.D., Ph.D., Mary Elkins, M.S., Jack Moye, M.D., Pamela Stratton, M.D., and James Balsley, M.D., Ph.D.,
for the Pediatric AIDS Clinical Trials Group Protocol 076 Study Group*

Testing differences in proportion

Is zidovudine the same as standard care?

$$n_\mathrm{i} = 180, m_\mathrm{i} = 13$$
$$n_\mathrm{c} = 183, m_\mathrm{c} = 40$$

$$\mathcal{H}_0 : \pi_\mathrm{i} - \pi_\mathrm{c} = 0$$
$$\mathcal{H}_1 : \pi_\mathrm{i} - \pi_\mathrm{c} \neq 0$$

$$\alpha = 0.05$$

The New England
Journal of Medicine

Volume 331 NOVEMBER 3, 1994 Number 18

REDUCTION OF MATERNAL–INFANT TRANSMISSION OF HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 WITH ZIDOVUDINE TREATMENT

Edward M. Connor, M.D., Rhoda S. Sperling, M.D., Richard Gelber, Ph.D., Pavel Kiselev, Ph.D., Gwendolyn Scott, M.D., Mary Jo O'Sullivan, M.D., Russell VanDyke, M.D., Mohammed Bey, M.D., William Shearer, M.D., Ph.D., Robert L. Jacobson, M.D., Eleanor Jimenez, M.D., Edward O'Neill, M.D., Brigitte Bazin, M.D., Jean-François Delfraissy, M.D., Mary Culnane, M.S., Robert Coombs, M.D., Ph.D., Mary Elkins, M.S., Jack Moye, M.D., Pamela Stratton, M.D., and James Balsley, M.D., Ph.D.,
for the Pediatric AIDS Clinical Trials Group Protocol 076 Study Group*

Testing differences in proportion

📌 Is zidovudine the same as standard care?

$$n_{\mathrm{i}} = 180, m_{\mathrm{i}} = 13$$
$$n_{\mathrm{c}} = 183, m_{\mathrm{c}} = 40$$

$$\mathcal{H}_0 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} = 0, \quad \alpha = 0.05$$

$$\bar{P} = \hat{p}_{\mathrm{i}} - \hat{p}_{\mathrm{c}} = \frac{13}{180} - \frac{40}{183} = 0.07 - 0.22 = -0.15 = -15\%$$

$$\hat{\mathrm{SE}} = \sqrt{\frac{\bar{P}\times(1-\bar{P})}{n_i} + \frac{\bar{P}\times(1-\bar{P})}{n_c}} = \sqrt{\frac{0.15\times(1-0.15)}{180} + \frac{0.15\times(1-0.15)}{183}} = 0.037$$
$$\rightarrow \pm 1.96 \times \hat{SE} = \pm 1.96 \times 0.037 = \pm 0.073$$

Testing differences in proportion

📌 Is zidovudine the same as standard care?

$$n_{\mathrm{i}} = 180, m_{\mathrm{i}} = 13$$
$$n_{\mathrm{c}} = 183, m_{\mathrm{c}} = 40$$

$$\mathcal{H}_0 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} = 0, \quad \alpha = 0.05$$

$$\bar{P} = -0.15, \quad \hat{\mathrm{SE}} = 0.037$$
$$\rightarrow \pm 1.96 \times \hat{SE} = \pm 1.96 \times 0.037 = \pm 0.073$$

$$z = \frac{\bar{P} - 0}{\hat{SE}} = \frac{0.15}{0.037} = 4.06 > 0.073 \;\rightarrow\; \text{we reject } \mathcal{H}_0$$

Testing differences in proportion

📌 Is zidovudine the same as standard care?

$$n_\mathrm{i} = 180, m_\mathrm{i} = 13$$
$$n_\mathrm{c} = 183, m_\mathrm{c} = 40$$

$$\mathcal{H}_0 : \pi_\mathrm{i} - \pi_\mathrm{c} = 0, \quad \alpha = 0.05$$

$$\rightarrow \pm 1.96 \times \hat{SE} = \pm 1.96 \times 0.037 = \pm 0.073$$
$$z = \frac{\bar{P} - 0}{\hat{SE}} = \frac{0.15}{0.037} = 4.06 > 0.073 \;\rightarrow\; \text{we reject } \mathcal{H}_0$$

$$\mathcal{P}(|\bar{P}| \geq 4.06) = 2 \times (1 - 0.999975) = 0.00005 = 5 \times 10^{-5} \leftarrow \text{P value}$$

Pearson's χ^2 test

📌 Is zidovudine the same as standard care?

$$n_{\mathrm{i}} = 180, m_{\mathrm{i}} = 13$$
$$n_{\mathrm{c}} = 183, m_{\mathrm{c}} = 40$$

$$\mathcal{H}_0 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} = 0$$
$$\mathcal{H}_1 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} \neq 0$$

$$\alpha = 0.05$$

Pearson's χ^2 test

📌 Is zidovudine the same as standard care?

$$n_{\mathrm{i}} = 180, m_{\mathrm{i}} = 13$$
$$n_{\mathrm{c}} = 183, m_{\mathrm{c}} = 40$$

$$\mathcal{H}_0 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} = 0$$
$$\mathcal{H}_1 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} \neq 0$$

$$\alpha = 0.05$$

	Int	Cnt	Tot
HIV+			
HIV-			
Total	180	183	363

Pearson's χ^2 test

📌 Is zidovudine the same as standard care?

$$n_{\mathrm{i}} = 180, m_{\mathrm{i}} = 13$$
$$n_{\mathrm{c}} = 183, m_{\mathrm{c}} = 40$$

$$\mathcal{H}_0 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} = 0$$
$$\mathcal{H}_1 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} \neq 0$$

$$\alpha = 0.05$$

	Int	Cnt	Tot
HIV+	13	40	53
HIV-			
Total	180	183	363

Pearson's χ^2 test

📌 Is zidovudine the same as standard care?

$$n_{\mathrm{i}} = 180, m_{\mathrm{i}} = 13$$
$$n_{\mathrm{c}} = 183, m_{\mathrm{c}} = 40$$

$$\mathcal{H}_0 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} = 0$$
$$\mathcal{H}_1 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} \neq 0$$

$$\alpha = 0.05$$

	Int	Cnt	Tot
HIV+	13	40	53
HIV-	167	143	310
Total	180	183	363

Pearson's χ^2 test

📌 Is zidovudine the same as standard care?

$$n_\mathrm{i} = 180, m_\mathrm{i} = 13$$
$$n_\mathrm{c} = 183, m_\mathrm{c} = 40$$

$$\mathcal{H}_0 : \pi_\mathrm{i} - \pi_\mathrm{c} = 0$$
$$\mathcal{H}_1 : \pi_\mathrm{i} - \pi_\mathrm{c} \neq 0$$

$$\alpha = 0.05$$

$$\Pi = \frac{\mathrm{tot}^+}{tot} = \frac{53}{363} = 0.146$$

	Int	Cnt	Tot
HIV+	13	40	53
HIV-	167	143	310
Total	180	183	363

Pearson's χ^2 test

📌 Is zidovudine the same as standard care?

$$n_\mathrm{i} = 180, m_\mathrm{i} = 13$$
$$n_\mathrm{c} = 183, m_\mathrm{c} = 40$$

$$\mathcal{H}_0 : \pi_\mathrm{i} - \pi_\mathrm{c} = 0$$
$$\mathcal{H}_1 : \pi_\mathrm{i} - \pi_\mathrm{c} \neq 0$$

$$\alpha = 0.05$$

$$\Pi = \frac{\mathrm{tot}^+}{tot} = \frac{53}{363} = 0.146$$

	Int	Cnt	Tot
HIV+	13	40	53
HIV-	167	143	310
Total	180	183	363

	Int	Cnt	Tot
HIV+	180 $\times$ 0.146	183 $\times$ 0.146	53
HIV-			310
Total	180	183	363

Pearson's χ^2 test

📌 Is zidovudine the same as standard care?

$$n_{\mathrm{i}} = 180, m_{\mathrm{i}} = 13$$
$$n_{\mathrm{c}} = 183, m_{\mathrm{c}} = 40$$

$$\mathcal{H}_0 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} = 0$$
$$\mathcal{H}_1 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} \neq 0$$

$$\alpha = 0.05$$

$$\Pi = \frac{\mathrm{tot}^+}{tot} = \frac{53}{363} = 0.146$$

	Int	Cnt	Tot
HIV+	13	40	53
HIV-	167	143	310
Total	180	183	363

	Int	Cnt	Tot
HIV+	26.28	27.72	53
HIV-			
Total	180	183	363

Pearson's χ^2 test

📌 Is zidovudine the same as standard care?

$$n_{\mathrm{i}} = 180, m_{\mathrm{i}} = 13$$
$$n_{\mathrm{c}} = 183, m_{\mathrm{c}} = 40$$

$$\mathcal{H}_0 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} = 0$$
$$\mathcal{H}_1 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} \neq 0$$

$$\alpha = 0.05$$

$$\Pi = \frac{\mathrm{tot}^{+}}{tot} = \frac{53}{363} = 0.146$$

	Int	Cnt	Tot
HIV+	13	40	53
HIV-	167	143	310
Total	180	183	363

	Int	Cnt	Tot
HIV+	26.28	27.72	53
HIV-	153.72	155.28	310
Total	180	183	363

Pearson's χ^2 test

📌 Is zidovudine the same as standard care?

$$n_{\mathrm{i}} = 180, m_{\mathrm{i}} = 13$$
$$n_{\mathrm{c}} = 183, m_{\mathrm{c}} = 40$$

$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected} =$$

$$= \frac{(13-16.28)^2}{26.28} + \frac{(167-153.72)^2}{153.72} + \frac{(40-27.2)^2}{27.2} + \frac{(143-155.26)^2}{155.26} = $$
$$= 14.85$$

	Int	Cnt	Tot
HIV+	13	40	53
HIV-	167	143	310
Total	180	183	363

	Int	Cnt	Tot
HIV+	26.28	27.72	53
HIV-	153.72	155.28	310
Total	180	183	363

Pearson's χ^2 test

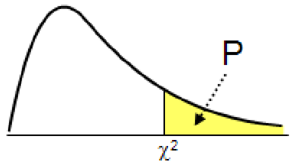
📌 Is zidovudine the same as standard care?

$$n_{\mathrm{i}} = 180, m_{\mathrm{i}} = 13$$
$$n_{\mathrm{c}} = 183, m_{\mathrm{c}} = 40$$

$$\mathcal{H}_0 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} = 0$$
$$\mathcal{H}_1 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} \neq 0$$

$$\alpha = 0.05$$

$$\chi^2 = 14.85 \quad \mathrm{df} = 1$$



	P										
DF	**0.995**	**0.975**	**0.20**	**0.10**	**0.05**	**0.025**	**0.02**	**0.01**	**0.005**	**0.002**	**0.001**
1	0.0000393	0.000982	1.642	2.706	3.841	5.024	5.412	6.635	7.879	9.550	10.828
2	0.0100	0.0506	3.219	4.605	5.991	7.378	7.824	9.210	10.597	12.429	13.816
3	0.0717	0.216	4.642	6.251	7.815	9.348	9.837	11.345	12.838	14.796	16.266
4	0.207	0.484	5.989	7.779	9.488	11.143	11.668	13.277	14.860	16.924	18.467
5	0.412	0.831	7.289	9.236	11.070	12.833	13.388	15.086	16.750	18.907	20.515
6	0.676	1.237	8.558	10.645	12.592	14.449	15.033	16.812	18.548	20.791	22.458

Pearson's χ^2 test

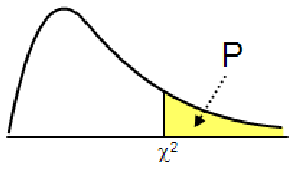
Is zidovudine the same as standard care?

$$n_{\mathrm{i}} = 180, m_{\mathrm{i}} = 13$$
$$n_{\mathrm{c}} = 183, m_{\mathrm{c}} = 40$$

$$\mathcal{H}_0 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} = 0$$
$$\mathcal{H}_1 : \pi_{\mathrm{i}} - \pi_{\mathrm{c}} \neq 0$$

$$\alpha = 0.05$$

$$\chi^2 = 14.85 \quad \mathrm{df} = 1$$
$$\chi^2_\alpha = 3.84 < 14.85$$
$$\rightarrow \text{reject } \mathcal{H}_0$$



	P										
DF	**0.995**	**0.975**	**0.20**	**0.10**	**0.05**	**0.025**	**0.02**	**0.01**	**0.005**	**0.002**	**0.001**
1	0.0000393	0.000982	1.642	2.706	3.841	5.024	5.412	6.635	7.879	9.550	10.828
2	0.0100	0.0506	3.219	4.605	5.991	7.378	7.824	9.210	10.597	12.429	13.816
3	0.0717	0.216	4.642	6.251	7.815	9.348	9.837	11.345	12.838	14.796	16.266
4	0.207	0.484	5.989	7.779	9.488	11.143	11.668	13.277	14.860	16.924	18.467
5	0.412	0.831	7.289	9.236	11.070	12.833	13.388	15.086	16.750	18.907	20.515
6	0.676	1.237	8.558	10.645	12.592	14.449	15.033	16.812	18.548	20.791	22.458

Pearson's χ^2 test

	No Exercise	Sporadic Exercise	Regular Exercise	Total
Primary education				
Lower secondary education				
Upper secondary education				
Bachelor/Master				
Doctorate				
Total				

Exercise 16

? Does using seatbelt when driving changes the chance of death?

$$n_{\mathrm{s}} = 250, m_{\mathrm{s}} = 3$$
$$n_{\mathrm{c}} = 290, m_{\mathrm{c}} = 13$$

Use the Pearson's χ^2 test, $\alpha = 0.05\ (\chi^2_\alpha = 3.84)$

Exercise 16 -- Solution

? Does using seatbelt when driving changes the chance of death?

$$n_{\mathrm{s}} = 250, m_{\mathrm{s}} = 3$$
$$n_{\mathrm{c}} = 290, m_{\mathrm{c}} = 13$$

1. Set $\mathcal{H}_0$ and $\mathcal{H}_1$ $\;\rightarrow\;$ $\mathcal{H}_0 : \pi_s - \pi_c = 0, \mathcal{H}_1 : \pi_s - \pi_c \neq 0$

Exercise 16 -- Solution

❓ Does using seatbelt when driving changes the chance of death?

$$n_\mathrm{s} = 250, m_\mathrm{s} = 3$$
$$n_\mathrm{c} = 290, m_\mathrm{c} = 13$$

1. $\mathcal{H}_0 : \pi_s - \pi_c = 0, \mathcal{H}_1 : \pi_s - \pi_c \neq 0$

2. Define $\alpha \quad \rightarrow \quad \alpha = 0.05$

Exercise 16 -- Solution

$$n_{\mathrm{s}} = 250, m_{\mathrm{s}} = 3$$
$$n_{\mathrm{c}} = 290, m_{\mathrm{c}} = 13$$

1. $\mathcal{H}_0 : \pi_s - \pi_c = 0, \mathcal{H}_1 : \pi_s - \pi_c \neq 0$
2. $\alpha = 0.05$

3. Calculate the test statistics $\downarrow$

	Seatbelt	No seatbelt	Total
Death	3	13	16
Survived	247	277	524
Total	250	290	540

$$\Pi = \frac{tot_{\mathrm{death}}}{tot} = \frac{15}{540} = 0.03$$

Exercise 16 -- Solution

$$n_\mathrm{s} = 250, m_\mathrm{s} = 3$$
$$n_\mathrm{c} = 290, m_\mathrm{c} = 13$$

1. $\mathcal{H}_0 : \pi_s - \pi_c = 0, \mathcal{H}_1 : \pi_s - \pi_c \neq 0$
2. $\alpha = 0.05$

3. Calculate the test statistics $\downarrow$

	Seatbelt	No seatbelt	Total
Death	3	13	16
Survived	247	277	524
Total	250	290	540

$$\Pi = \frac{tot_\mathrm{death}}{tot} = \frac{15}{540} = 0.03$$

	Seatbelt	No seatbelt	Total
Death	7.5	8.7	16
Survived	242.5	281.3	524
Total	250	290	540

$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$

Exercise 16 -- Solution

$$n_s = 250, m_s = 3$$
$$n_c = 290, m_c = 13$$

1. $\mathcal{H}_0 : \pi_s - \pi_c = 0, \mathcal{H}_1 : \pi_s - \pi_c \neq 0$
2. $\alpha = 0.05$

3. Calculate the test statistics $\downarrow$

	Seatbelt	No seatbelt	Total
Death	3	13	16
Survived	247	277	524
Total	250	290	540

	Seatbelt	No seatbelt	Total
Death	7.5	8.7	16
Survived	242.5	281.3	524
Total	250	290	540

$$\chi^2 = \frac{(3-7.5)^2}{7.5} + \frac{(13-8.7)^2}{8.7} + \frac{(247-242.5)^2}{242.5} + \frac{(277-281.3)^2}{281.3} = 4.98$$

Exercise 16 -- Solution

$$n_\mathrm{s} = 250, m_\mathrm{s} = 3$$
$$n_\mathrm{c} = 290, m_\mathrm{c} = 13$$

1. $\mathcal{H}_0 : \pi_s - \pi_c = 0, \mathcal{H}_1 : \pi_s - \pi_c \neq 0$
2. $\alpha = 0.05$

3. Calculate the test statistics $\downarrow$

$$\chi^2 = 4.98 \quad \mathrm{df} = 1$$
$$\chi^2_\alpha = 3.84 < 4.98$$
$$\mathcal{P}(\chi^2 \geq 4.98) = 0.03$$

Exercise 16 -- Solution

$$n_{\mathrm{s}} = 250, m_{\mathrm{s}} = 3$$
$$n_{\mathrm{c}} = 290, m_{\mathrm{c}} = 13$$

1. $\mathcal{H}_0 : \pi_s - \pi_c = 0, \mathcal{H}_1 : \pi_s - \pi_c \neq 0$
2. $\alpha = 0.05$

3. Calculate the test statistics $\downarrow$

$$\chi^2 = 4.98 \quad \mathrm{df} = 1$$
$$\chi^2_\alpha = 3.84 < 4.98$$
$$\mathcal{P}(\chi^2 \geq 4.98) = 0.03$$

4. Make a decision about $\mathcal{H}_0$ $\rightarrow$ We refuse $\mathcal{H}_0$

Pearson's χ^2 test -- Yates' correction

🎯 $$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$

$$\downarrow$$

$$\chi^2 = \sum \frac{(|Observed - Expected| - 0.5)^2}{Expected}$$

Errors in decision making

🎯 $p < \alpha \rightarrow \text{reject } \mathcal{H}_0$
$p \geq \alpha \rightarrow \text{fail to reject } \mathcal{H}_0$

Errors in decision making

🎯 $p < \alpha \rightarrow \text{reject } \mathcal{H}_0$
$p \geq \alpha \rightarrow \text{fail to reject } \mathcal{H}_0$

$\mathcal{H}_0$ is	TRUE	FALSE
Rejected		
Non rejected		

Errors in decision making

$\mathcal{H}_0$ is	TRUE	FALSE
Rejected	false positive	
Non rejected		false negative

Errors in decision making

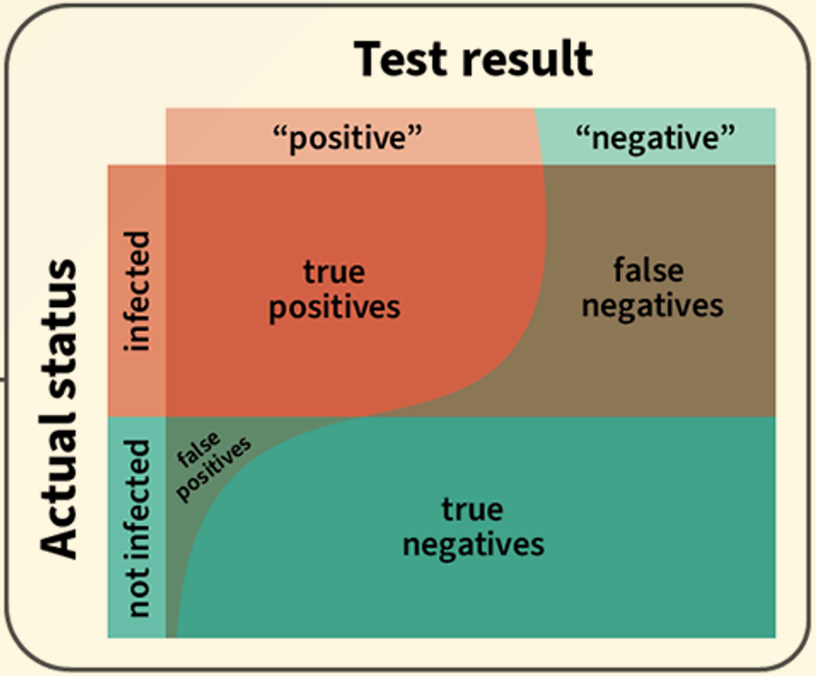
🎯 α is the level of significance, or Type I error
β is the Type II error

$\mathcal{H}_0$ is	TRUE	FALSE
Rejected	α	
Non rejected		β

Specificity *vs* sensitivity

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.



Sensitivity = number of **true positives** / number of those tested who **really are infected** = **“how many of the infections did we find?”**

Specificity = number of **true negatives** / number of those tested who **really are not infected** = **“how many of the healthy people did we clear?”**

Errors in decision making

$\mathcal{H}_0$ is	TRUE	FALSE
Rejected		true positive
Non rejected	true negative	

Power of a study

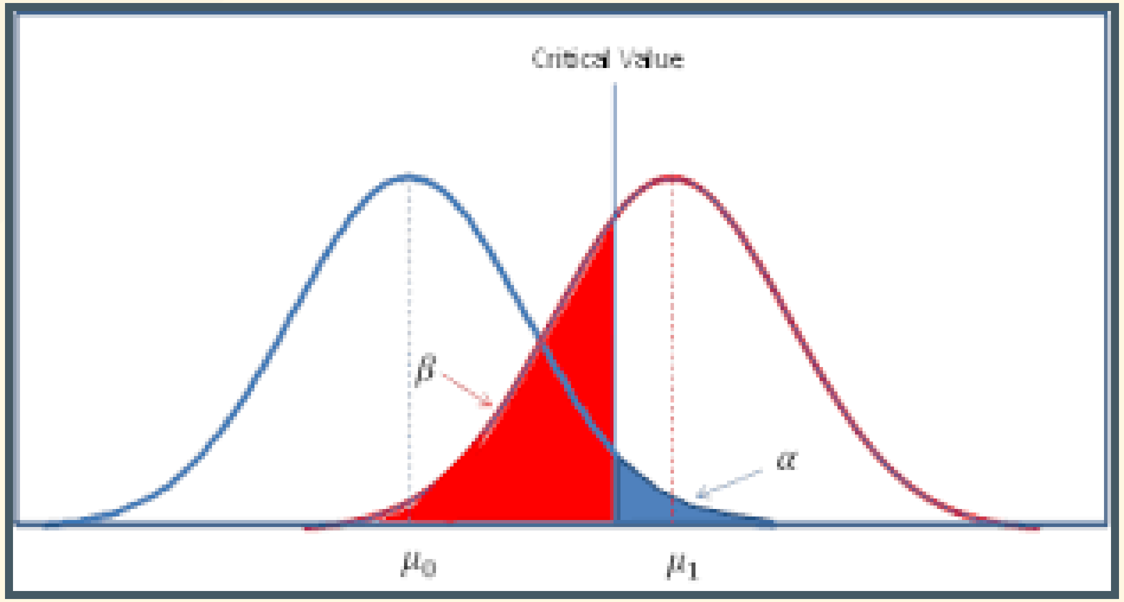
🎯 $1 - \beta$ is the power of a statistical test
(Acceptable power: $\geq 80\%$)

$\mathcal{H}_0$ is	TRUE	FALSE
Rejected	α	$1 - \beta$
Non rejected	$1 - \alpha$	β

Power of a study

🎯 The power is increased by:

- larger α
- larger $\mu_i - \mu_c$
- smaller σ^2
- larger sample size n



Exercise #17

? There was a shepherd boy who repeatedly cried wolf when there was no wolf. Yet, each time, villagers went to help him. Then, the wolf arrived, but, when the boy cried wolf, no villager helped.

Which kind of errors are the villagers making?

a) Type I error, then Type II error
b) Type II error, then Type I error
c) Null error, then alternative error
d) None of the above

Exercise #17 -- Solution

❓ There was a shepherd boy who repeatedly cried wolf when there was no wolf. Yet, each time, villagers went to help him. Then, the wolf arrived, but, when the boy cried wolf, no villager helped.

Which kind of errors are the villagers making?

a) Type I error, then Type II error ✅
b) Type II error, then Type I error
c) Null error, then alternative error
d) None of the above

Exercise #18

❓ I want to increase the power of my study, what factors are under my control?

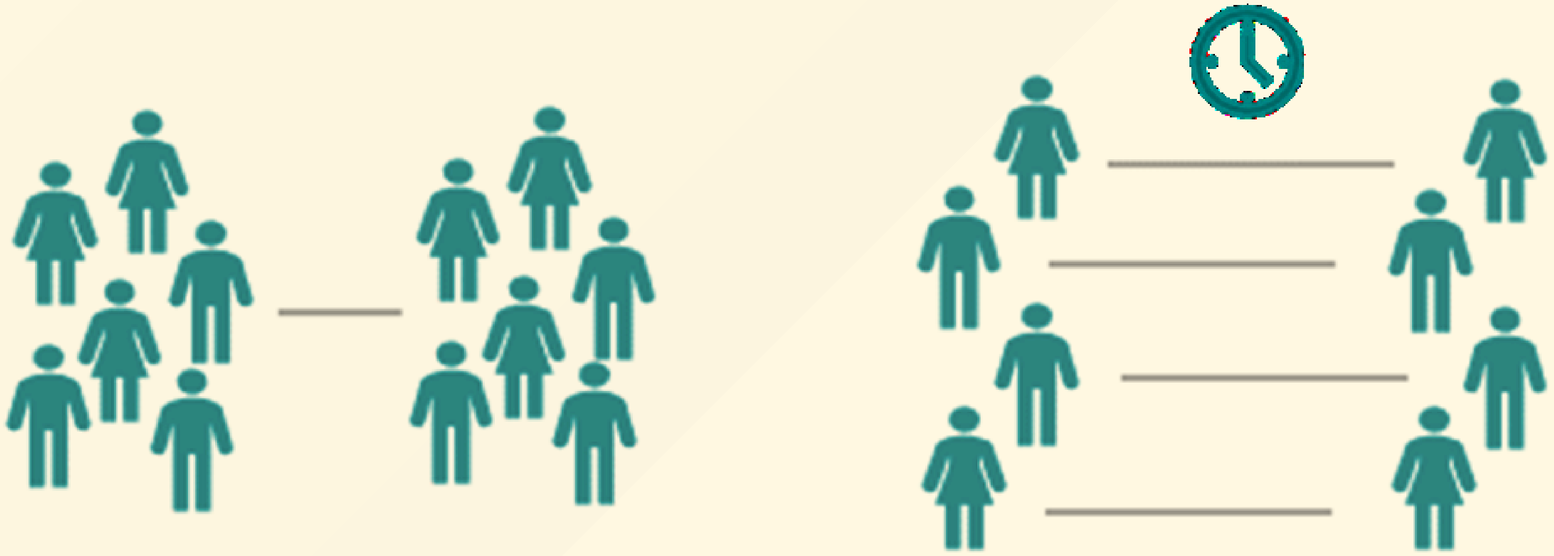
a) the level of significance α
b) the difference $\mu_i - \mu_c$
c) the samples' σ^2
d) the samples' size n
e) Both a) and d)
f) Both a) and c)

Exercise #18 -- Solution

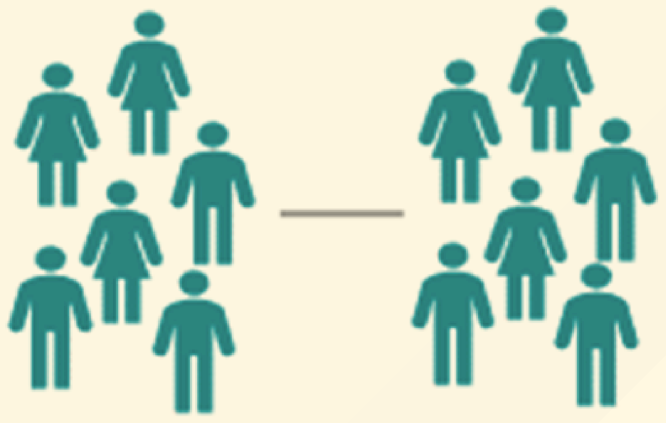
❓ I want to increase the power of my study, what factors are under my control?

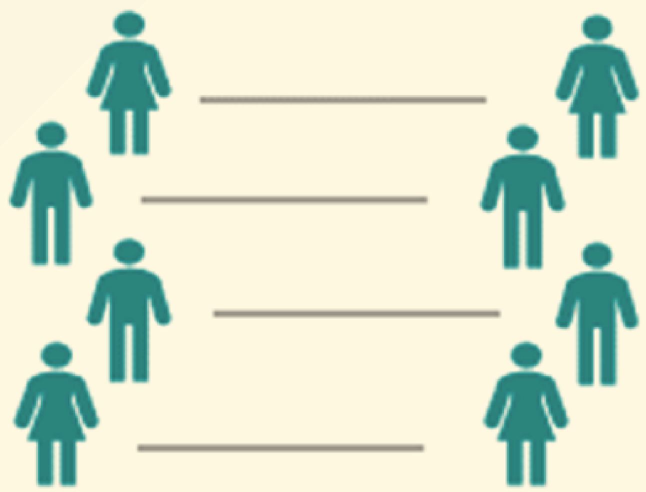
a) the level of significance α
b) the difference $\mu_i - \mu_c$
c) the samples' σ^2
d) the samples' size n
e) Both a) and d) ✅
f) Both a) and c)

Independent and paired samples

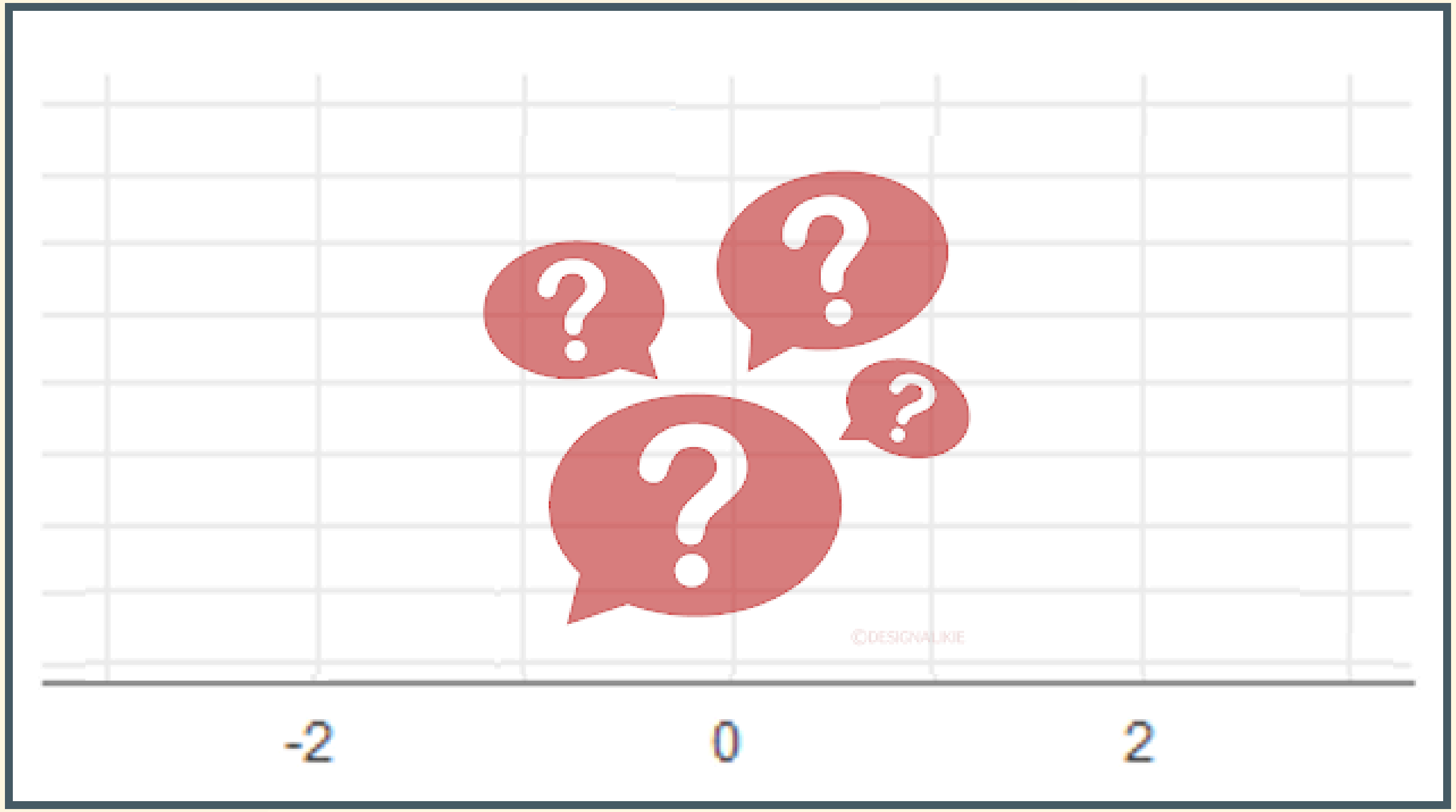


Independent and paired samples





Non-parametric tests

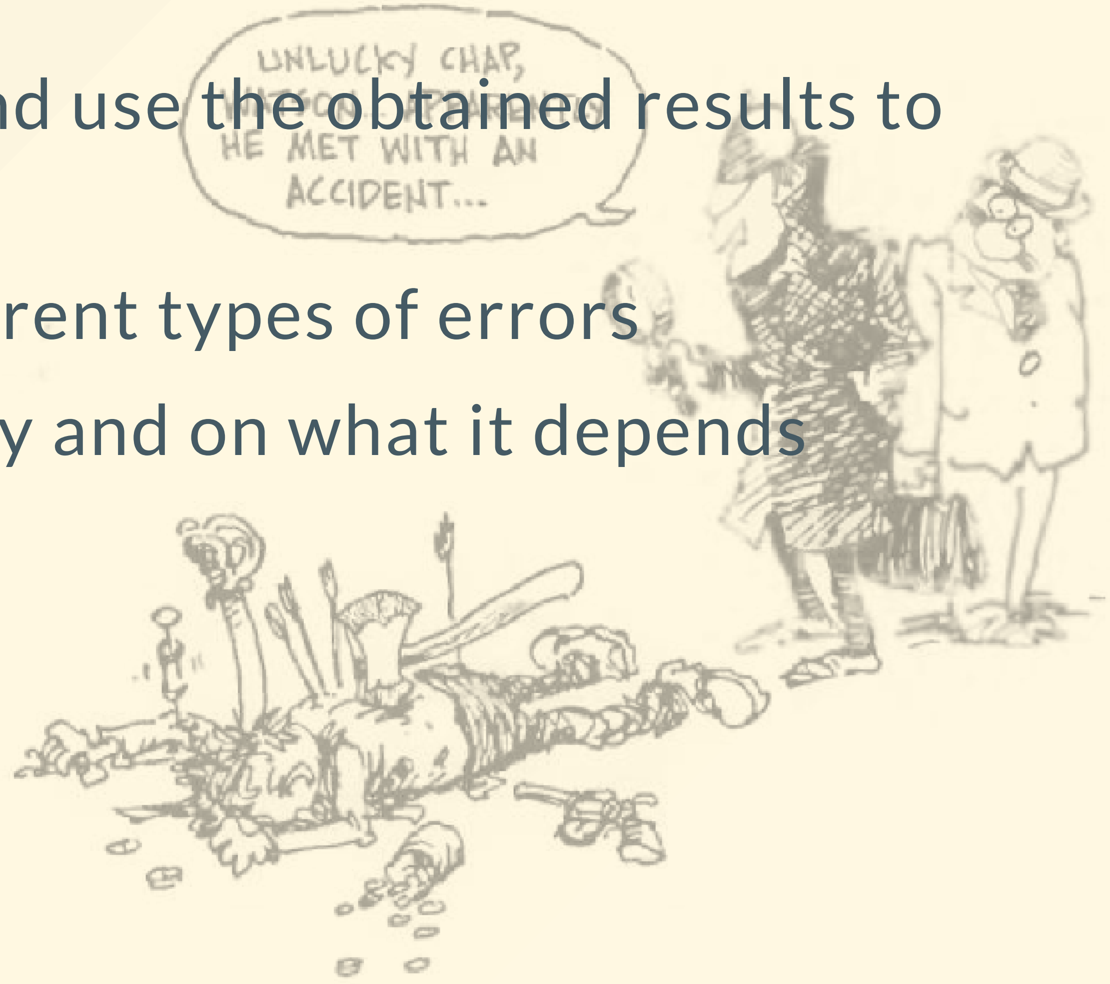


Non-parametric tests

Sample	Data type	$\mathcal{H}_0$	Non-parametric test
Independent	Numerical	$\mu_1 = \mu_2$	Mann-Whitney's test
Paired	Numerical	$\mu_1 = \mu_2$	Wilcoxon's test
Independent	Categorical	$\pi_1 = \pi_2$	Fisher's test
Paired	Categorical	$\pi_1 = \pi_2$	McNemar's test

Summary

- We can make and test hypotheses, and use the obtained results to make decision
- We are aware that we can make different types of errors
- We know what is the power of a study and on what it depends



Wrap up

The PARACHUTE trial

RESEARCH

Parachute use to prevent death and major trauma when jumping from aircraft: randomized controlled trial

Robert W Yeh,[1] Linda R Valsdottir,[1] Michael W Yeh,[2] Changyu Shen,[1] Daniel B Kramer,[1] Jordan B Strom,[1] Eric A Secemsky,[1] Joanne L Healy,[1] Robert M Domeier,[3] Dhruv S Kazi,[1] Brahmajee K Nallamothu[4] On behalf of the PARACHUTE Investigators

WHAT IS ALREADY KNOWN ON THIS TOPIC

Parachutes are routinely used to prevent death or major traumatic injury among individuals jumping from aircraft, but their efficacy is based primarily on biological plausibility and expert opinion

No randomized controlled trials of parachute use have yet been attempted, presumably owing to a lack of equipoise

WHAT THIS STUDY ADDS

This randomized trial of parachute use found no reduction in death or major injury compared with individuals jumping from aircraft with an empty backpack

Lack of enrolment of individuals at high risk could have influenced the results of the trial

The PARACHUTE trial

Closing remarks

> “ *To consult the statistician after an experiment is finished is often merely to ask him to conduct a post mortem examination. He can perhaps say what the experiment died of.* ”

R. Fisher

Thank you