Correlation (Usually) Does Not Imply Causation

Probability and Statistics for Data Science

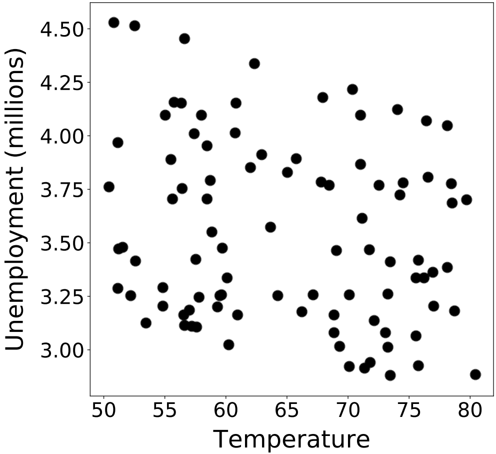
Carlos Fernandez-Granda

NYU | COURANT INSTITUTE OF MATHEMATICAL SCIENCES

NYU DATA SCIENCE

These slides are based on the book Probability and Statistics for Data Science by Carlos Fernandez-Granda, available for purchase here. A free preprint, videos, code, slides and solutions to exercises are available at https://www.ps4ds.net

Unemployment and temperature in Spain (2015-2022)



Correlation coefficient: -0.21

Would an increase in temperature decrease unemployment?

Causal inference

Key question: Does a treatment $\tilde{t}$ cause a certain outcome?

Potential outcome: $\widetilde{\mathrm{po}}_t$ (defined for observed and unobserved t)

Observed data:

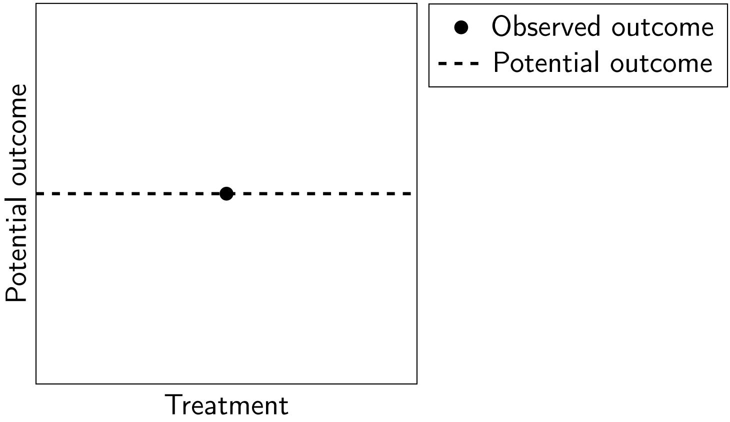
$$\tilde{y} := \widetilde{\mathrm{po}}_t \qquad \text{if} \qquad \tilde{t} = t$$

Fundamental problem of causal inference: Incomplete data

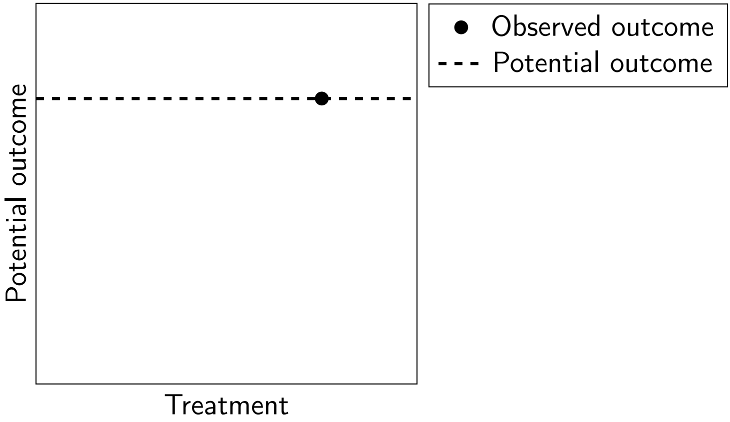
If temperature is 40°, we only observe $\tilde{y} := \widetilde{\mathrm{po}}_{40}$

$\widetilde{\mathrm{po}}_{30}$, $\widetilde{\mathrm{po}}_{45}$, $\widetilde{\mathrm{po}}_{63}$, $\widetilde{\mathrm{po}}_{75}$... are all counterfactuals

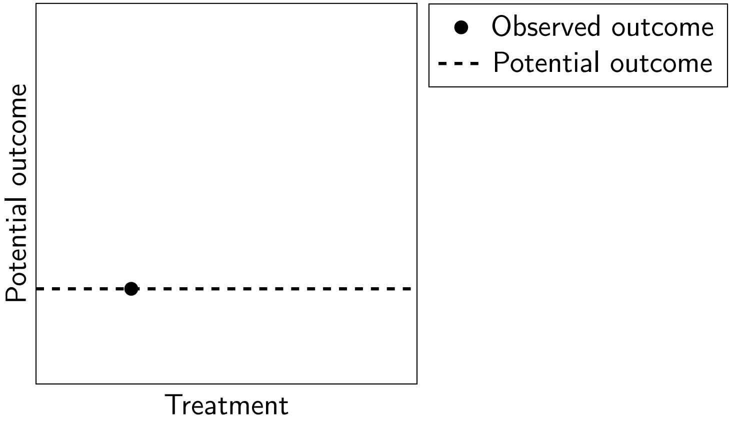
Potential outcomes



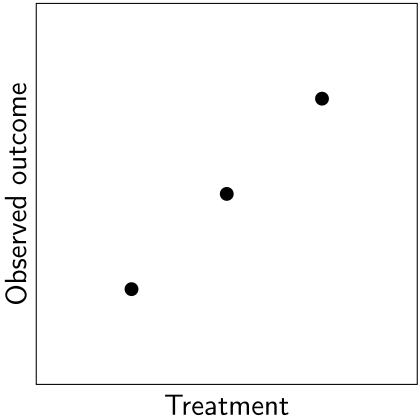
Potential outcomes



Potential outcomes



Observed data



Linear causal effect

For some constant $\beta \in \mathbb{R}$

$$\mathrm{E}\left[\widetilde{\mathsf{po}}_t\right] = \beta t$$

Key question: Can we estimate linear causal effects from data?

Idea

Use covariance between observed outcome $\tilde{y}$ and the treatment $\tilde{t}$

Assumption: $\widetilde{\mathrm{po}}_t$ and $\tilde{t}$ are independent for all t

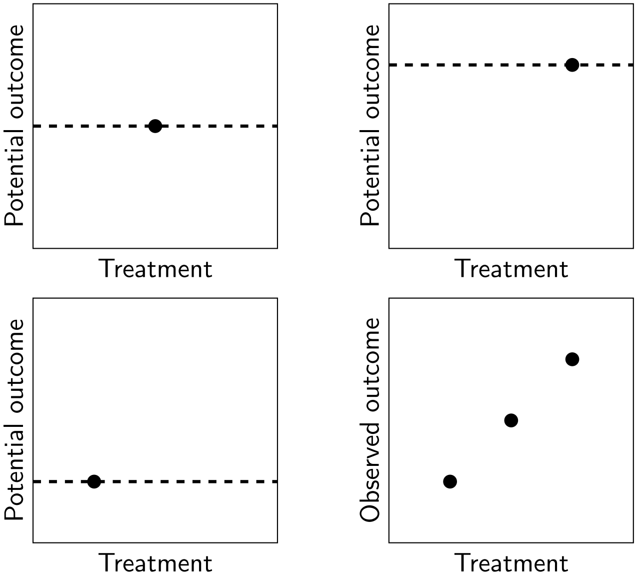
Iterated expectation

Assuming $\mathrm{E}[\tilde{t}] = 0$ and $\mathrm{E}[\tilde{t}^2] = 1$

$$\begin{aligned}
\operatorname{Cov}[\tilde{y}, \tilde{t}\,] = \mathrm{E}[\tilde{y}\tilde{t}\,] &= \mathrm{E}\left[\mu_{\tilde{y}\tilde{t}\,|\,\tilde{t}}(\tilde{t})\right] \\
&= \mathrm{E}\left[\beta \tilde{t}^{\,2}\right] \\
&= \beta \mathrm{E}\left[\tilde{t}^{\,2}\right] = \beta
\end{aligned}$$

$$\begin{aligned}
\mu_{\tilde{y}\tilde{t}\,|\,\tilde{t}}(t) &= \int_{y=-\infty}^{\infty} yt\, f_{\tilde{y}\,|\,\tilde{t}}(y \,|\, t)\,\mathrm{d}y \\
&= \int_{y=-\infty}^{\infty} yt\, f_{\widetilde{\mathrm{po}}_t\,|\,\tilde{t}}(y \,|\, t)\,\mathrm{d}y \\
&= t \int_{y=-\infty}^{\infty} y\, f_{\widetilde{\mathrm{po}}_t}(y)\,\mathrm{d}y \\
&= t\mathrm{E}\left[\widetilde{\mathrm{po}}_t\right] \\
&= \beta t^2
\end{aligned}$$

Why do we need independence between $\widetilde{\text{po}}_t$ and $\tilde{t}$?



Guinea-pig rescue

Goal: Fatten the guinea pigs

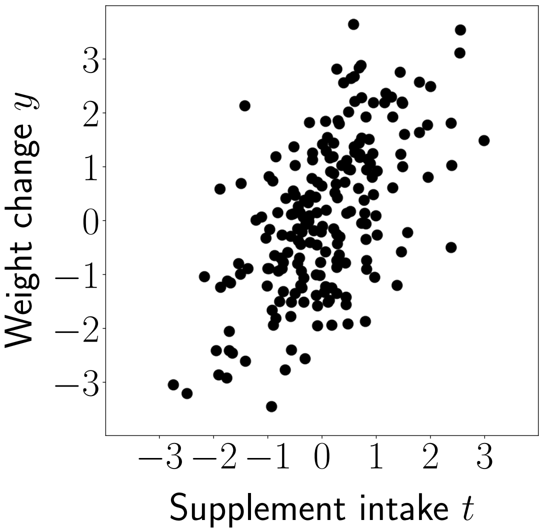
Question: Does a nutritional supplement increase weight?

Week 1

Supplement mixed with food

Supplement mixed with food

Covariance = 0.8

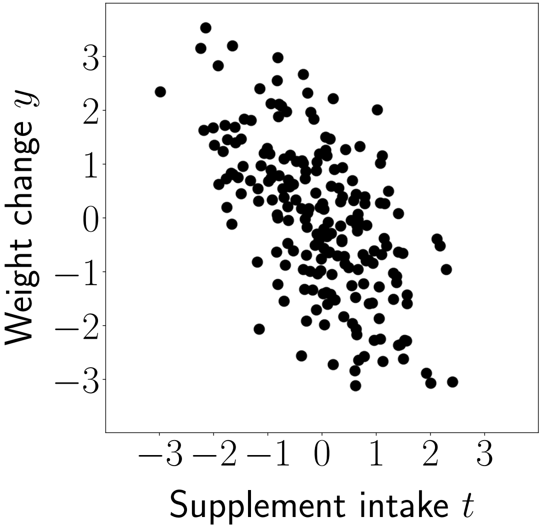


Week 2

Supplement provided after food

Supplement after the food

Covariance = -0.8



What's going on?

Treatment is dependent on food intake

Potential outcomes are also dependent on food intake

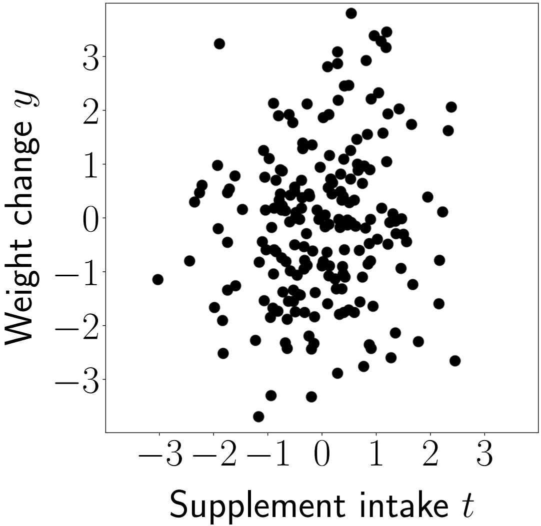
Treatment is dependent on potential outcomes!

How can we break dependence between $\widetilde{\text{po}}_t$ and $\tilde{t}$?

Randomizing supplement provided to each pig

Randomized supplement

Covariance = 0



Idealized model

Potential outcome $\widetilde{\mathrm{po}}_t$ depends linearly on treatment t, but also on unobserved confounder $\tilde{c}$

$$\widetilde{\mathrm{po}}_t := \beta_{\mathrm{treat}} t + \beta_{\mathrm{conf}} \tilde{c} + \tilde{z}$$

$$\tilde{y} = \beta_{\mathrm{treat}} \tilde{t} + \beta_{\mathrm{conf}} \tilde{c} + \tilde{z}$$

Assumptions:

- $\tilde{t}$ and $\tilde{c}$ are standardized
- $\tilde{z}$ is independent from $\tilde{t}$ and $\tilde{c}$

Iterated expectation

$$
\begin{aligned}
\mathrm{Cov}\left[\tilde{y}, \tilde{t}\right] &= \mathrm{E}\left[\tilde{y}\tilde{t}\right] \\
&= \mathrm{E}\left[\beta_{\text{treat}}\tilde{t}^{2} + \beta_{\text{conf}}\tilde{c}\tilde{t} + \tilde{z}\tilde{t}\right] \\
&= \beta_{\text{treat}}\mathrm{E}\left[\tilde{t}^{2}\right] + \beta_{\text{conf}}\mathrm{E}\left[\tilde{c}\tilde{t}\right] + \mathrm{E}\left[\tilde{z}\right]\mathrm{E}\left[\tilde{t}\right] \\
&= \beta_{\text{treat}} + \beta_{\text{conf}}\sigma_{\tilde{t},\tilde{c}}
\end{aligned}
$$

Covariance is distorted by confounder!

Guinea pigs

Potential outcome $\widetilde{\mathrm{po}}_t$: Weight change

$$\widetilde{\mathrm{po}}_t := \tilde{c} + \tilde{z}$$

$$\tilde{y} = \tilde{c} + \tilde{z}$$

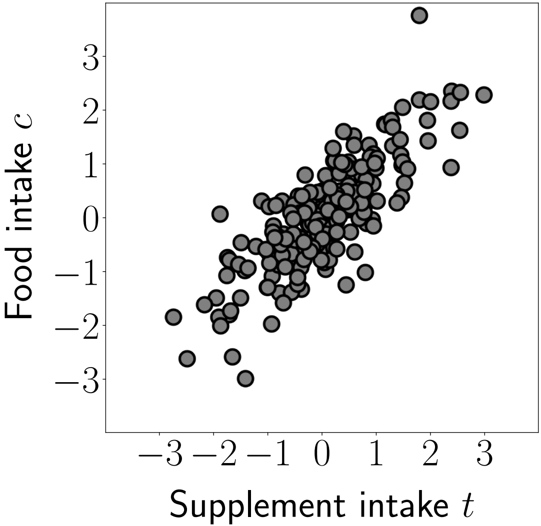
Treatment $\tilde{t}$: Supplement intake, $\beta_{\text{treat}} := 0$

Confounder $\tilde{c}$: Food intake, $\beta_{\text{conf}} := 1$

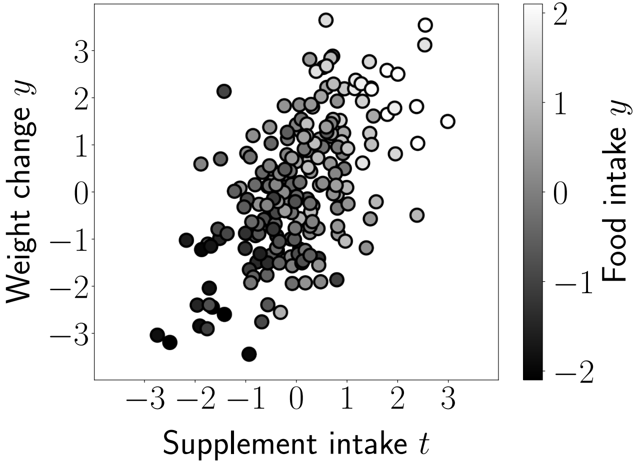
Covariance between observed weight change and supplement?

$$\begin{aligned} \operatorname{Cov}[\tilde{y}, \tilde{t}\,] &= \beta_{\text{treat}} + \beta_{\text{conf}} \sigma_{\tilde{t},\tilde{c}} \\ &= \sigma_{\tilde{t},\tilde{c}} \end{aligned}$$

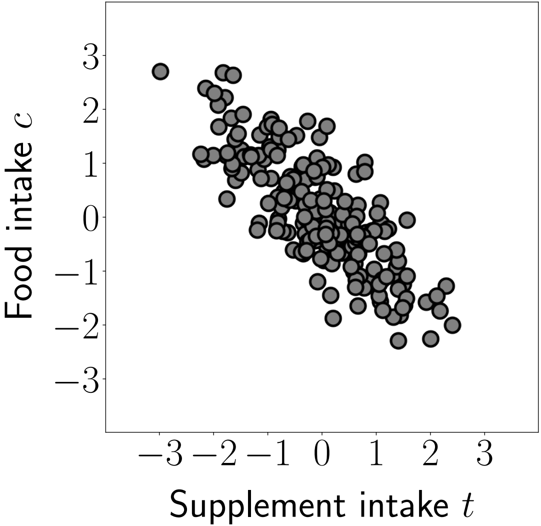
Supplement mixed with food: $\sigma_{\tilde{t},\tilde{c}} := 0.8$



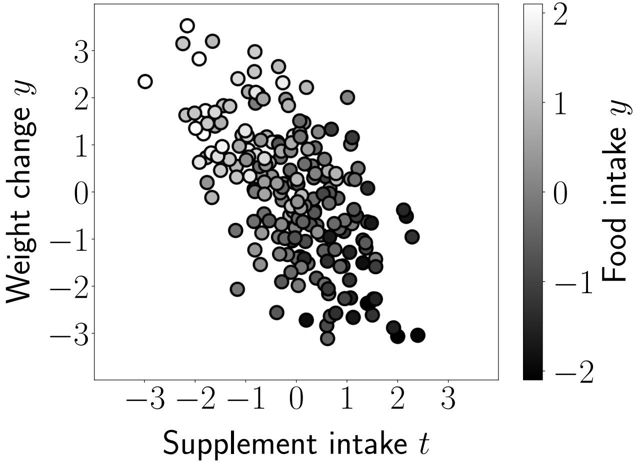
Supplement mixed with food: $\mathrm{Cov}[\tilde{y}, \tilde{t}\,] = 0.8$



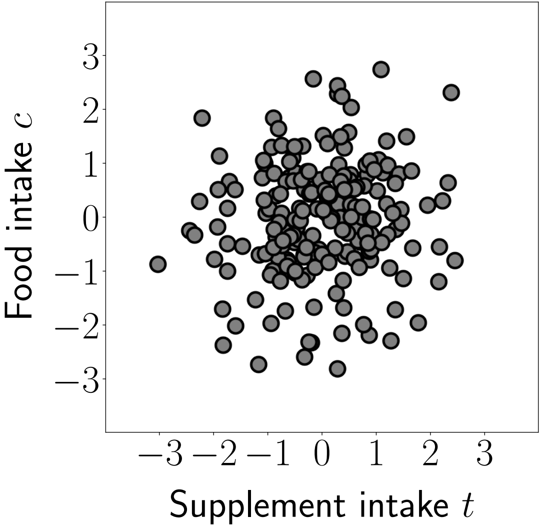
Supplement after the food: $\sigma_{\tilde{t},\tilde{c}} := -0.8$



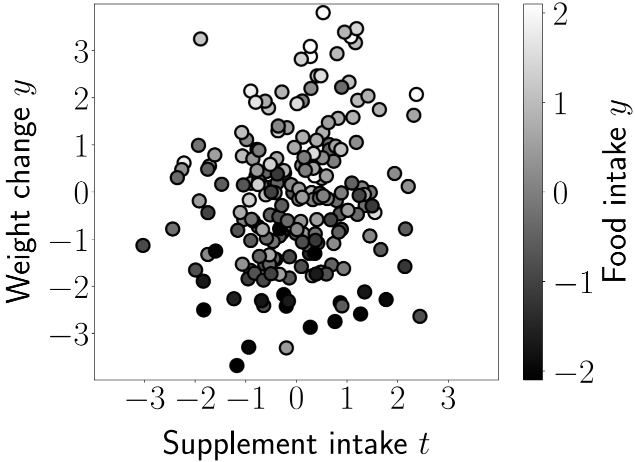
Supplement after the food: $\mathrm{Cov}\left[\tilde{y}, \tilde{t}\,\right] = -0.8$



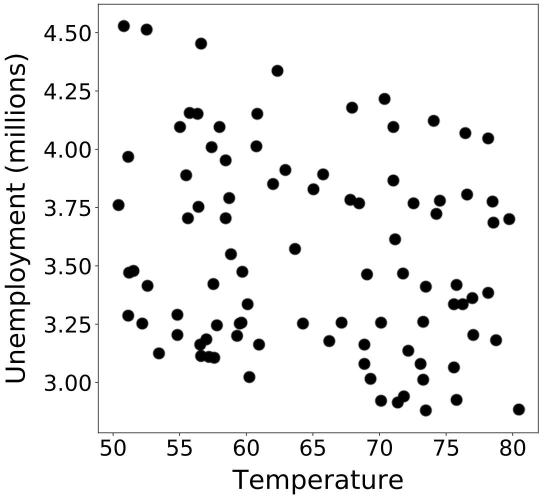
Randomized supplement: $\sigma_{\tilde{t},\tilde{c}} = 0$



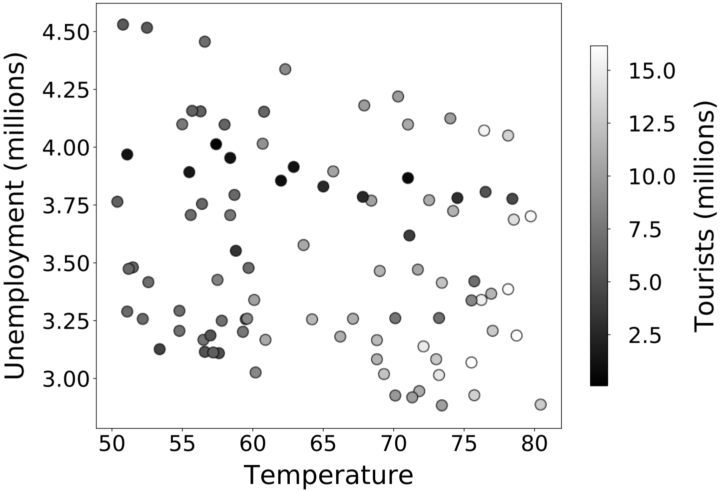
Randomized supplement: $\mathrm{Cov}[\tilde{y}, \tilde{t}\,] = 0$



Confounder?


Unemployment (millions)
4.50
4.25
4.00
3.75
3.50
3.25
3.00
50
55
60
65
70
75
80
Temperature

Tourists!



Unemployment (millions)
4.50
4.25
4.00
3.75
3.50
3.25
3.00
50
55
60
65
70
75
80
Temperature
15.0
12.5
10.0
7.5
5.0
2.5
Tourists (millions)

What have we learned

Correlation does not imply causation

However, it does if the treatment is randomized

Otherwise, unobserved confounders can produce spurious correlation