University of Antwerp
Faculty of Social Sciences

Objective rating method: Entropy

Speech intelligibility estimation

Jose Rivera

April 16, 2022

What are we going to talk about?

1. Preliminars

Research question

Research question

On two fronts:

1. Can comparative judgement (CJ) methods be used to assess speech intelligibility (SI)?,
   To investigate this wee need:
   - an objective measure of SI
2. where CJ stands versus absolute holistic judgement (HJ) methods?,
   In terms of:
   - validity
   - reliability
   - statistical efficiency
   - time efficiency

Objective measure of SI

the most objective measure of SI (we know of) comes from a transcription task:

1. transcribing children's utterances (made by multiple judges),
2. align transcriptions at the utterance level,
3. calculate an entropy measure (H), defined as

$$H = H(\mathbf{p}) = \frac{-\sum_{i=1}^{n} p_i \cdot \log_2(p_i)}{\log_2(N)}$$

4. characteristics of H [3, 7]
   - bounded in $[0, 1]$ space,
   - utterances with more agreement are more intelligible, and therefore $H \to 0$,
   - utterances with low agreement are less intelligible, and therefore $H \to 1$.

1. Preliminars

Research hypothesis production

A typical scientific lab[1]

What is needed?

1. Quality of theory
2. Quality of data
3. Reliable procedures and code
4. Quality of data analysis
5. Documentation
6. Reporting

What we will deal with:

1. Quality of theory
2. Quality of data
3. Reliable procedures and code
4. Quality of data analysis
5. Documentation
6. Reporting

[1] McElreath [13], lecture 20 and McElreath [14], chapter 17

Research hypothesis production[2]

Well known challenges

- Insufficient data
- Wrong population
- Measurement error
- Selection bias
- Confounding

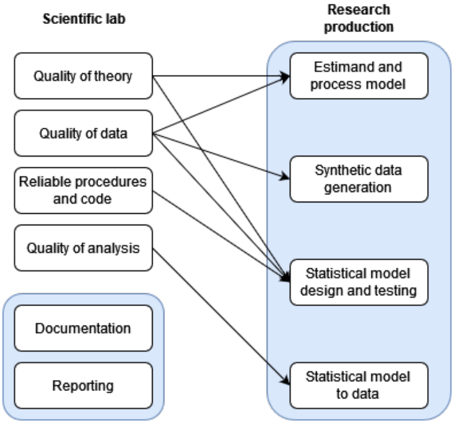
Known challenges in our research;

- Insufficient data (possibly)
- Wrong population
- Measurement error
- Selection bias
- Confounding

[2] Hernán [9], lesson 4

Research hypothesis schematics[3]

a. Estimand and process model
b. Synthetic data generation
c. Statistical model design and testing
d. Apply statistical model to data



[3] McElreath [14], lecture 20, Pearl [17]. Follow Fogarty et al. [8] on item (c).

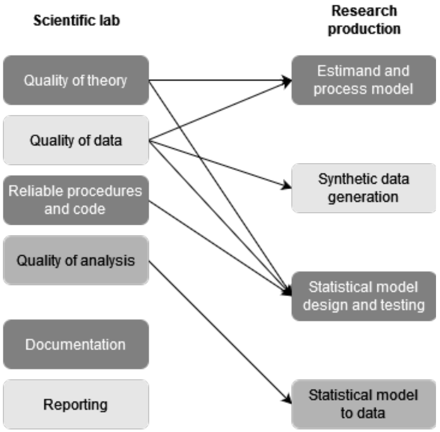
Why do we need to follow this?

Because the improvement of:

- A clear definition of the estimand and process model (assumptions).
- An improved the reliability of your procedures.
- As a documentation procedure.

leads to:

- A sound analysis, and sound results (even when we cannot answer our question).
- An improved planning to get data.

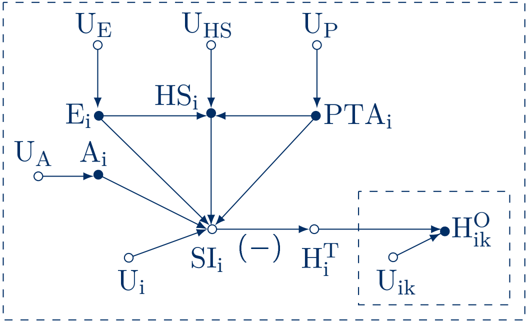


2. Estimand and Process model

The theory behind our research

- H_{ik} = (observed) entropy replicates
- H_{i} = (latent) child's entropy
- SI_{i} = (latent) child's SI score (inversely related to $\mathrm{H}_{\mathrm{i}}^{\mathrm{T}}$)
- A_{i} = child's "hearing" age
- E_{i} = child's etiology of disease
- HS_{i} = child's hearing status
- $\mathrm{PTA}_{\mathrm{i}}$ = child's pure tone average
- variables assumed independent, beyond the described relationships,

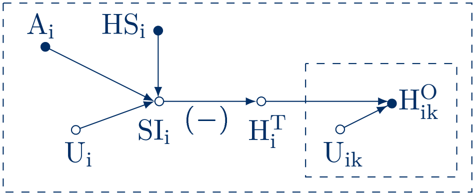
$$
\begin{aligned}
P(\mathbf{U}) &= P(U_{ik}, U_i, U_A, U_E, U_{HS}, U_P) \\
&= P(U_{ik})P(U_i)P(U_A)P(U_E)P(U_{HS})P(U_P)
\end{aligned}
$$



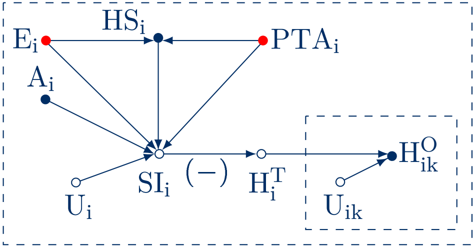
General causal diagram

Interested in two effects

1. total effects model inherits:
   - children's characteristics that lead to the fitting of specific apparatus,
   - the (convenience of) sample selection (fixed with post-stratification)
2. to do the last, we stratify for all variables that explain variability, ergo, use a direct effects model
3. We have two levels: replicates (k), child's level (i), denoted by squares
4. U_{ik} = replicates measurement error
   U_i = within child SI variability



(b) total effects



(a) direct effects

Causal and probabilistic model

$$
\begin{aligned}
H^{O}_{ik} &\sim \text{BetapProp}(H^{T}_{i}, df_{ik}) \\
H^{T}_{i} &= \text{inv\_logit}(-SI_{i}) \\
SI_{i} &\sim \text{Normal}(\mu_{SI}, \sigma_{Ui}) \\
&\quad \mu_{SI} = a_{i} + \alpha + \alpha_{HS[i]} + \alpha_{E[i]} \\
&\qquad + \beta_{A,HS[i]}(A_{i} - \bar{A}) + \beta_{P} PTA_{i} \\
HS_{i} &\sim \text{data} \\
A_{i} &\sim \text{data} \\
E_{i} &\sim \text{data} \\
PTA_{i} &\sim \text{data} \\
U &\sim \text{unobservable}
\end{aligned}
$$

(a) general probabilistic model

$$
\begin{aligned}
H^{O}_{ik} &\leftarrow f(H^{T}_{i}, U_{ik}) \\
H^{T}_{i} &\leftarrow f(SI_{i}) \\
SI_{i} &\leftarrow f(HS_{i}, A_{i}, E_{i}, PTA_{i}, U_{i}) \\
HS_{i} &\leftarrow f(U_{HS}) \\
A_{i} &\leftarrow f(U_{A}) \\
E_{i} &\leftarrow f(U_{E}) \\
PTA_{i} &\leftarrow f(U_{P}) \\
U &\sim P(\mathbf{U})
\end{aligned}
$$

(a) general structural model

University of Antwerp
Faculty of Social Sciences

3. Synthetic data generation

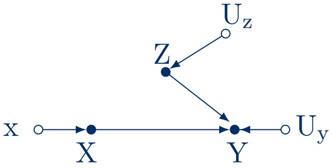
Intervention

- Purpose: to keep a control on all the factors responsible for the outcome's variation blue(understand the system).
- It is modeled by modifying the structural model (and causal diagram).
- remember: $\mathbf{V} = \{Z, X, Y\}$, $\mathbf{U} = \{U_z, U_x, U_y\}$, and $\mathbf{F} = \{f_z, f_x, f_y\}$.
- Intervention on X can be written in do-calculus[a] as: $P(\mathbf{V} \mid do(X = x))$.

[a] we are not delving into this (the usual suspects [15, 16, 18, 19])

$$M = \begin{cases} Z \leftarrow f_z(U_z) \\ X \leftarrow f_x(U_x) \\ Y \leftarrow f_y(X, Z, U_y) \\ U \sim P(\mathbf{U}) \end{cases}$$

(a) structural model



(b) causal diagram

Effects of interest

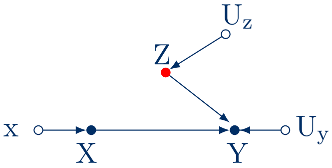
1. Average causal effect:
   $\mathrm{ACE}(x) = E[Y|do(x+1)] - E[Y|do(x)]$
2. Controlled direct effect:
   $\mathrm{CDE}(x, z) = E[Y|do(x+1), do(z)] - E[Y|do(x), do(z)]$

points to consider (more on part 3):

- CDE takes a particular relevance with observational data.
- There is also a distinction between total effect and direct effect.

$$M = \begin{cases} Z \leftarrow f_z(U_z) \\ X \leftarrow f_x(U_x) \\ Y \leftarrow f_y(X, Z, U_y) \\ U \sim P(\mathbf{U}) \end{cases}$$

(a) structural model



(b) causal diagram

3. Synthetic data generation

Fork bias: spurious relationships

Fork bias: spurious relationships[4]

also known as

- spurious association
- confounder

research question

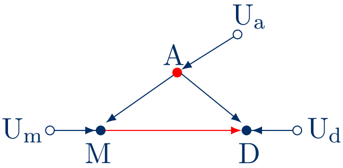
- Does M has a (direct) effect on D?

variables

- A, median age at marriage
- M, marriage rate
- D, divorce rate

$$\mathrm{M} = \begin{cases} \mathrm{A} \leftarrow \mathrm{f_a}(\mathrm{U_a}) \\ \mathrm{M} \leftarrow \mathrm{f_m}(\mathrm{A}, \mathrm{U_m}) \\ \mathrm{D} \leftarrow \mathrm{f_d}(\mathrm{A}, \mathrm{M}, \mathrm{U_d}) \\ \mathrm{U} \sim \mathrm{P}(\mathbf{U}) \end{cases}$$

(a) structural model



(b) causal diagram

[4] extracted from McElreath [13], chapter 05 (p. 125)

Simulation conventions

one way to defined it

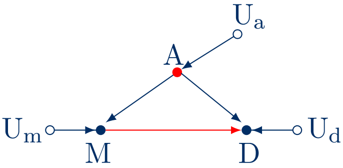
$$
\begin{aligned}
A &= U_a & U_a &\sim N(0, \sigma_a) \\
M &= \beta_A A + U_m & U_m &\sim N(0, \sigma_a) \\
D &= \beta_A A + \beta_M M + U_d & U_d &\sim N(0, \sigma_a)
\end{aligned}
$$

a more succinct way

$$
\begin{aligned}
A &\sim N(0, \sigma_a) \\
M &\sim N(\beta_A A, \sigma_m) \\
D &\sim N(\beta_A A + \beta_M M, \sigma_d)
\end{aligned}
$$

$$
M = \begin{cases} A \leftarrow f_a(U_a) \\ M \leftarrow f_m(A, U_m) \\ D \leftarrow f_d(A, M, U_d) \\ U \sim P(\mathbf{U}) \end{cases}
$$

(a) structural model



(b) causal diagram

Simulation setting

Implications

- $M \not\perp\!\!\!\perp D$
- $M \not\perp\!\!\!\perp D |\, A$

Specific definition

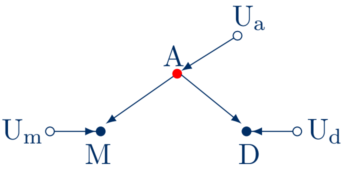
$$
\begin{aligned}
A &\sim N(0, 1) \\
M &\sim N(-A, 1) \\
D &\sim N(-A + 0 \times M, 1)
\end{aligned}
$$

R code

```
n = 100
A = rnorm( n )
M = rnorm( n , mean=-A )
D = rnorm( n , mean=-A + 0*M )
```

$$
M = \begin{cases} A \leftarrow f_a(U_a) \\ M \leftarrow f_m(A, U_m) \\ D \leftarrow f_d(A, U_d) \\ U \sim P(\mathbf{U}) \end{cases}
$$

(a) structural model



(b) causal diagram

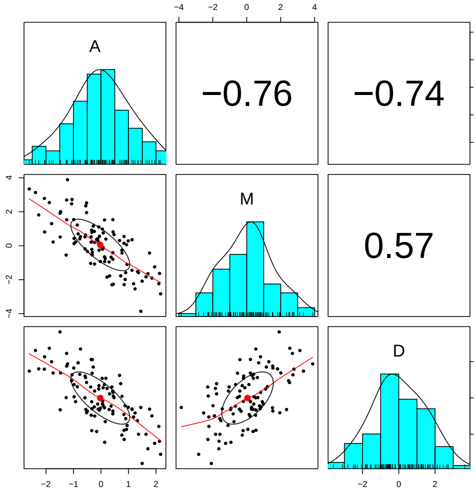
What we see?

and what we might decide

Two regressions, one result

```
summary(lm(D \  M, data=d))

> Coefficients:
> Estimate Std. Error t value Pr(>|t|)
> (Intercept) -0.03364 0.11863 -0.284 0.777
> M 0.54561 0.07848 6.952 4.03e-10 ***
```



Result

Frishc-Waugh-Lovell Backdoor criterion on bad controls paper (p. 18)

What is going on?

4. Anecdotal cases

Experimental design: the panacea

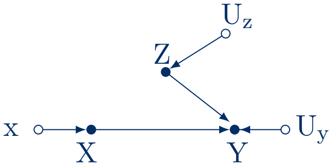
Intervention

- Purpose: to keep a control on all the factors responsible for the outcome's variation blue(understand the system).
- It is modeled by modifying the structural model (and causal diagram).
- remember: $\mathbf{V} = \{Z, X, Y\}$, $\mathbf{U} = \{U_z, U_x, U_y\}$, and $\mathbf{F} = \{f_z, f_x, f_y\}$.
- Intervention on X can be written in do-calculus[a] as: $P(\mathbf{V} \mid do(X = x))$.

[a] we are not delving into this (the usual suspects [15, 16, 18, 19])

$$M = \begin{cases} Z \leftarrow f_z(U_z) \\ X \leftarrow f_x(U_x) \\ Y \leftarrow f_y(X, Z, U_y) \\ U \sim P(\mathbf{U}) \end{cases}$$

(a) structural model



(b) causal diagram

Effects of interest

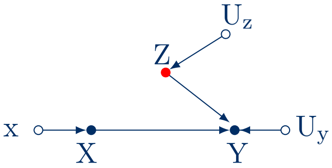
1. Average causal effect:
   $ACE(x) = E[Y|do(x+1)] - E[Y|do(x)]$
2. Controlled direct effect:
   $CDE(x, z) = E[Y|do(x+1), do(z)] - E[Y|do(x), do(z)]$

points to consider (more on part 3):

- CDE takes a particular relevance with observational data.
- There is also a distinction between total effect and direct effect.

$$M = \begin{cases} Z \leftarrow f_z(U_z) \\ X \leftarrow f_x(U_x) \\ Y \leftarrow f_y(X, Z, U_y) \\ U \sim P(\mathbf{U}) \end{cases}$$

(a) structural model



(b) causal diagram

4. Anecdotal cases

Fork bias: spurious relationships

Fork bias: spurious relationships[5]

also known as

- spurious association
- confounder

research question

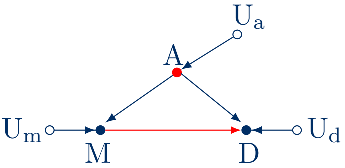
- Does M has a (direct) effect on D?

variables

- A, median age at marriage
- M, marriage rate
- D, divorce rate

$$M = \begin{cases} A \leftarrow f_a(U_a) \\ M \leftarrow f_m(A, U_m) \\ D \leftarrow f_d(A, M, U_d) \\ U \sim P(\mathbf{U}) \end{cases}$$

(a) structural model



(b) causal diagram

[5] extracted from McElreath [13], chapter 05 (p. 125)

Simulation conventions

one way to defined it

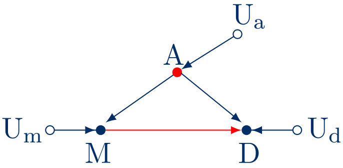
$$
\begin{aligned}
A &= U_a & U_a &\sim N(0, \sigma_a) \\
M &= \beta_A A + U_m & U_m &\sim N(0, \sigma_a) \\
D &= \beta_A A + \beta_M M + U_d & U_d &\sim N(0, \sigma_a)
\end{aligned}
$$

a more succinct way

$$
\begin{aligned}
A &\sim N(0, \sigma_a) \\
M &\sim N(\beta_A A, \sigma_m) \\
D &\sim N(\beta_A A + \beta_M M, \sigma_d)
\end{aligned}
$$

$$
M = \begin{cases}
A \leftarrow f_a(U_a) \\
M \leftarrow f_m(A, U_m) \\
D \leftarrow f_d(A, M, U_d) \\
U \sim P(\mathbf{U})
\end{cases}
$$

(a) structural model



(b) causal diagram

Simulation setting

Implications

- $M \not\perp\!\!\!\perp D$
- $M \not\perp\!\!\!\perp D |\ A$

Specific definition

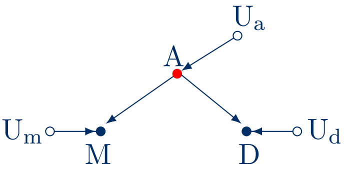
$$\begin{aligned} A &\sim N(0,1) \\ M &\sim N(-A,1) \\ D &\sim N(-A + 0 \times M, 1) \end{aligned}$$

R code

```
n = 100
A = rnorm( n )
M = rnorm( n , mean=-A )
D = rnorm( n , mean=-A + 0*M )
```

$$M = \begin{cases} A \leftarrow f_a(U_a) \\ M \leftarrow f_m(A, U_m) \\ D \leftarrow f_d(A, U_d) \\ U \sim P(\mathbf{U}) \end{cases}$$

(a) structural model



(b) causal diagram

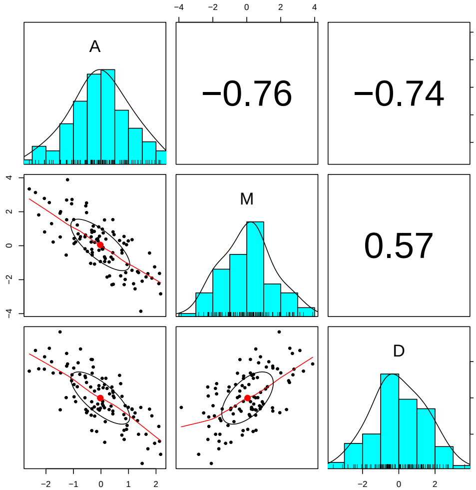
What we see?

and what we might decide

Two regressions, one result

```
summary(lm(D \  M, data=d))

> Coefficients:
> Estimate Std. Error t value Pr(>|t|)
> (Intercept) -0.03364 0.11863 -0.284 0.777
> M 0.54561 0.07848 6.952 4.03e-10 ***
```



Result

Frishc-Waugh-Lovell Backdoor criterion on bad controls paper (p. 18)

What is going on?

5. Anecdotal cases

Experimental design: the panacea

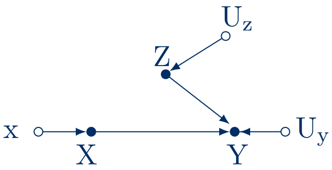
Intervention

- Purpose: to keep a control on all the factors responsible for the outcome's variation blue(understand the system).
- It is modeled by modifying the structural model (and causal diagram).
- remember: $\mathbf{V} = \{Z, X, Y\}$, $\mathbf{U} = \{U_z, U_x, U_y\}$, and $\mathbf{F} = \{f_z, f_x, f_y\}$.
- Intervention on X can be written in do-calculus[a] as: $P(\mathbf{V} \mid do(X = x))$.

[a] we are not delving into this (the usual suspects [15, 16, 18, 19])

$$M = \begin{cases} Z \leftarrow f_z(U_z) \\ X \leftarrow f_x(U_x) \\ Y \leftarrow f_y(X, Z, U_y) \\ U \sim P(\mathbf{U}) \end{cases}$$

(a) structural model



(b) causal diagram

Effects of interest

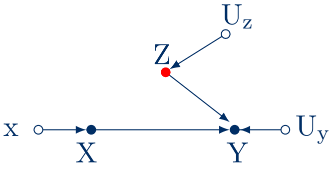
1. Average causal effect:
   $\mathrm{ACE}(x) = E[Y|do(x+1)] - E[Y|do(x)]$
2. Controlled direct effect:
   $\mathrm{CDE}(x, z) = E[Y|do(x+1), do(z)] - E[Y|do(x), do(z)]$

points to consider (more on part 3):

- CDE takes a particular relevance with observational data.
- There is also a distinction between total effect and direct effect.

$$M = \begin{cases} Z \leftarrow f_z(U_z) \\ X \leftarrow f_x(U_x) \\ Y \leftarrow f_y(X, Z, U_y) \\ U \sim P(\mathbf{U}) \end{cases}$$

(a) structural model



(b) causal diagram

5. Anecdotal cases

Fork bias: spurious relationships

Fork bias: spurious relationships[6]

also known as

- spurious association
- confounder

research question

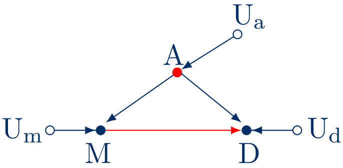
- Does M has a (direct) effect on D?

variables

- A, median age at marriage
- M, marriage rate
- D, divorce rate

$$M = \begin{cases} A \leftarrow f_a(U_a) \\ M \leftarrow f_m(A, U_m) \\ D \leftarrow f_d(A, M, U_d) \\ U \sim P(\mathbf{U}) \end{cases}$$

(a) structural model



(b) causal diagram

[6] extracted from McElreath [13], chapter 05 (p. 125)

Simulation conventions

one way to defined it

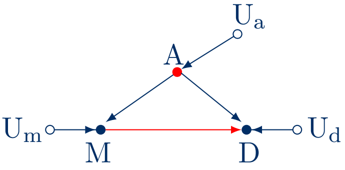
$$
\begin{aligned}
A &= U_a & U_a &\sim N(0, \sigma_a) \\
M &= \beta_A A + U_m & U_m &\sim N(0, \sigma_a) \\
D &= \beta_A A + \beta_M M + U_d & U_d &\sim N(0, \sigma_a)
\end{aligned}
$$

a more succinct way

$$
\begin{aligned}
A &\sim N(0, \sigma_a) \\
M &\sim N(\beta_A A, \sigma_m) \\
D &\sim N(\beta_A A + \beta_M M, \sigma_d)
\end{aligned}
$$

$$
M = \begin{cases} A \leftarrow f_a(U_a) \\ M \leftarrow f_m(A, U_m) \\ D \leftarrow f_d(A, M, U_d) \\ U \sim P(\mathbf{U}) \end{cases}
$$

(a) structural model



(b) causal diagram

Simulation setting

Implications

- $M \not\perp\!\!\!\perp D$
- $M \not\perp\!\!\!\perp D \mid A$

Specific definition

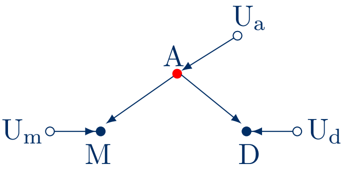
$$
\begin{aligned}
A &\sim N(0, 1) \\
M &\sim N(-A, 1) \\
D &\sim N(-A + 0 \times M, 1)
\end{aligned}
$$

R code

```
n = 100
A = rnorm( n )
M = rnorm( n , mean=-A )
D = rnorm( n , mean=-A + 0*M )
```

$$
\mathcal{M} = \begin{cases} A \leftarrow f_a(U_a) \\ M \leftarrow f_m(A, U_m) \\ D \leftarrow f_d(A, U_d) \\ U \sim P(\mathbf{U}) \end{cases}
$$

(a) structural model



(b) causal diagram

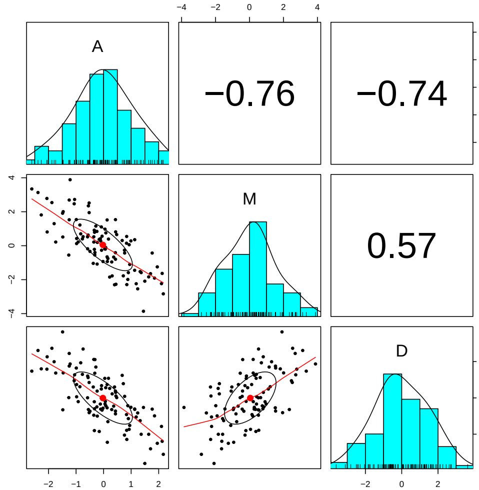
What we see?

and what we might decide

Two regressions, one result

```
summary(lm(D \ M, data=d))

> Coefficients:
> Estimate Std. Error t value Pr(>|t|)
> (Intercept) -0.03364 0.11863 -0.284 0.777
> M 0.54561 0.07848 6.952 4.03e-10 ***
```



Result

Frishc-Waugh-Lovell Backdoor criterion on bad controls paper (p. 18)

What is going on?

6. Concluding remarks

Concluding remarks

- Research is filled with challenges, some are obvious, some are not (you: Duh!!)
- Statistical models are not theory (you: so obvious again!!)
- You should not trust your statistical model, when no DAG is involved (me: how about that?!)
- For explanation, no amount of data can save you, when no DAG is involved (me: booya?!)
- For prediction, sometimes a DAG can help (me: did you expect this one?!)

7. Do you wanna know more???

7. Do you wanna know more???

[1] Anderson, D. [2008]. Model Based Inference in the Life Sciences: A Primer on Evidence, Springer.

[2] Bareinboim, E. and Pearl, J. [2016]. Causal inference and the data-fusion problem, Proceedings of the National Academy of Sciences 113(27): 7345–7352. doi: https://doi.org/10.1073/pnas.1510507113.

[3] Boonen, N., Kloots, H., Nurzia, P. and Gillis, S. [2021]. Spontaneous speech intelligibility: early cochlear implanted children versus their normally hearing peers at seven years of age, Journal of Child Language pp. 1–26. doi: https://doi.org/10.1017/S0305000921000714.

[4] Chamberlain, T. [1965]. The method of multiple working hypotheses, Science 148(3671): 754–759. url: https://www.jstor.org/stable/1716334.

[5] Cinelli, C., Forney, A. and Pearl, J. [2021]. A crash course in good and bad controls, Technical report.

[6] Cunningham, S. [2022]. Causal inference: The mixtape. url: https://mixtape.scunning.com/index.html.

[7] Faes, J., De Maeyer, S. and Gillis, S. [2021]. Speech intelligibility of children with an auditory brainstem implant: a triple-case study, pp. 1–50. (submitted).

[8] Fogarty, L., Madeleine, A., Holding, T., Powell, A. and Kandler, A. [2022]. Ten simple rules for principled simulation modelling, PLOS Computational Biology 18(3): 1–8.
doi: https://doi.org/10.1371/journal.pcbi.1009917.

[9] Hernán, M. [2020]. Causal diagrams: Draw your assumptions before your conclusions.
url: https://www.edx.org/course/causal-diagrams-draw-your-assumptions-before-your.

[10] Hernán, M. and Robins, J. [2020]. Causal Inference: What If, 1 edn, Chapman and Hall/CRC.
url: https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book.

[11] Jaynes, E. [2003]. Probability Theory: The Logic of Science, Cambridge University Press.

[12] McElreath, R. [2019]. Statistical rethinking, 2019 course.
url: https://github.com/rmcelreath/statrethinking_winter2019.

[13] McElreath, R. [2020]. Statistical Rethinking: A Bayesian Course with Examples in R and STAN, Chapman and Hall/CRC.

[14] McElreath, R. [2022]. Statistical rethinking, 2022 course.
url: https://github.com/rmcelreath/stat_rethinking_2022.

[15] Pearl, J. [1988]. Probabilistic reasoning in intelligent systems: Networks of plausible inference, The Journal of Philosophy 88(8): 434–437.
doi: https://doi.org/10.2307/2026705.
url: https://www.jstor.org/stable/2026705.

[16] Pearl, J. [2009]. Causality: Models, Reasoning and Inference, Cambrige University Press.

[17] Pearl, J. [2019]. The seven tools of causal inference, with reflections on machine learning, Communications of the ACM 62(3): 54–60.
doi: https://doi.org/10.1177/096228021558601O.

[18] Pearl, J., Glymour, M. and Jewell, N. [2016]. Causal Inference in Statistics: A Primer, John Wiley Sons, Inc.

[19] Pearl, J. and Mackenzie, D. [2018]. The Book of Why: The New Science of Cause and Effect, 1st edn, Basic Books, Inc.

[20] Spirtes, P., Glymour, C. and Scheines, R. [1991]. From probability to causality, Philosophical Studies 64(1): 1–36.
url: https://www.jstor.org/stable/4320244.

[21] Textor, J., van der Zander, B., Gilthorpe, M., Liskiewicz, M. and Ellison, G. [2016]. Robust causal inference using directed acyclic graphs: the r package 'dagitty', Int J Epidemiol 45(6): 1887–1894. doi: https://doi.org/10.1093/ije/dyw341.