Help! Statistics!

Unravelling statistical paradoxes using causal graphs

Sacha la Bastide-van Gemert

Medical Statistics and Decision Making

Department of Epidemiology

UMCG

Help! Statistics! Lunch time lectures

What? Frequently used statistical methods and questions in a manageable timeframe for all researchers at the UMCG.
No knowledge of advanced statistics is required.

When? Lectures take place every 2nd Tuesday of the month, 12.00-13.00 hrs.

Who? Unit for Medical Statistics and Decision Making

>>> lectures will be announced on the UMGC Intranet Agenda <<<

When?	**Where?**	**What?**	**Who?**
Nov 13, 2018	Room 16	Unravelling statistical paradoxes using causal graphs	S la Bastide
Dec 11, 2018	Room 16	Non-parametrical tests	Douwe Postmus
		*** *winter break* ***	
Feb 12, 2019	....	....	....

Slides can be downloaded from:
https://www.rug.nl/research/epidemiology/download-area

Unravelling statistical paradoxes using causal graphs

- Introducing Lord's original paradox (1967)
- More examples of statistical paradoxes
- Explaining the paradoxal statistical effects
- The real underlying issue: confounding
- Introducing causal graphs
- Concluding remarks

Lord's paradox (1967)

Psychological Bulletin
1967, Vol. 68, No. 5, 304–305

A PARADOX IN THE INTERPRETATION OF GROUP COMPARISONS

FREDERIC M. LORD

Educational Testing Service

Attention is called to a basic source of confusion in the interpretation of certain types of group comparison data.

It is common practice in behavioral research, and in other areas, to apply the analysis of covariance in the investigation of preexisting natural groups. The research worker is usually interested in some criterion variable (y) and would like to make allowances for the fact that his groups are not matched on some important independent variable or "control" variable (x). The situation is such that observed differences in the dependent variable might logically be caused by differences in the independent variable, and the research worker wishes to rule out this possibility.

parison of the educational achievements of different racial groups. The research worker usually uses analysis of covariance regardless, or he may try to resort to a simple and direct interpretation of group means.

The present note points out a type of problem that arises in interpreting data on preexisting groups. The difficulty can most easily be pointed out with the help of a hypothetical illustrative example.

A large university is interested in investigating the effects on the students of the diet provided in the university dining halls and any sex difference in these effects. Various

The question:

is total weight change (after six months) among students eating in the same university dining hall different for boys and girls?

Same data, two answers:

Statistician 1:
no difference in weight gain between boys and girls

?

Statistician 2:
For identical starting weight, there **is** a difference in weight gain between boys and girls

Lord’s paradox (1967)

Statistician 1:
Initial & final weight: no systematic changes (similar means & frequency distributions for each group)

Ergo: no effect of diet on student weight

Statistician 2:
Similar slope, but difference in initial weight

Ergo: for same initial weight, on average, boys gain more than girls

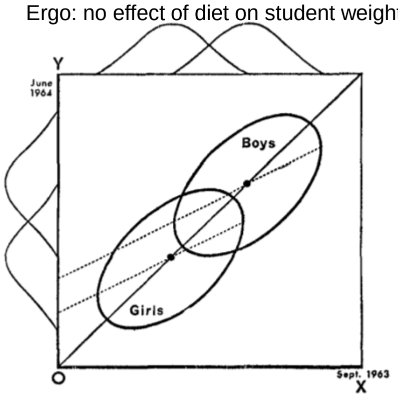


FIG. 1. Hypothetical scatterplots showing initial and final weight for boys and for girls.

The paradox:
two ``equally valid’’ analyses,
yet opposite conclusions!

Lord’s conclusion (1967):

``*There simply is no logical of statistical procedure that can be counted on to make proper allowances for uncontrolled preexisting differences between groups*’’

Spoiler: there is anno 2018!

Statistical paradoxes

Paradox:
``*A statement or proposition which, despite (apparently) sound reasoning from acceptable premises, leads to a conclusion that seems logically unacceptable or self-contradictory*’’ (Oxford Dictionary)

Lord’s paradox: the effect of a variable on another *in subgroups* changes direction (or disappears) when compared to the effect of the variable in the whole group

Related paradoxes: Simpson’s, Berkson’s paradox, suppression effect, ...

More general: after statistically ``correcting for’’ an extra background variable the effect of a third variable on outcome changes direction or disappears

Example I (Tu et al. 2008)

birth-weight (bw) → ? *blood pressure (bp)*

? current weight (cw) ?

	normal bp	high bp	total
low bw	354	132	486
high bw	328	186	514
total	682	318	1000

Example I (Tu et al. 2008)

birth-weight (bw) → ? *blood pressure (bp)*

? current weight (cw) ?

	normal bp	high bp	total	%high bp
low bw	354	132	486	27.2%
high bw	328	186	514	36.2%
total	682	318	1000	31.8%

→ low birth-weight seems to have a positive effect on blood pressure

Example I (continued)

Divided in subgroups (current weight categorical):

	normal bp	high bp	total	%high bp
low bw	354	132	486	27.2%
low cw	329	99	428	
high cw	25	33	58	
high bw	328	186	514	36.2%
low cw	221	55	276	
high cw	107	131	238	
total	682	318	1000	31.8%
low cw	550	154	704	
high cw	132	164	296	

Example I (continued)

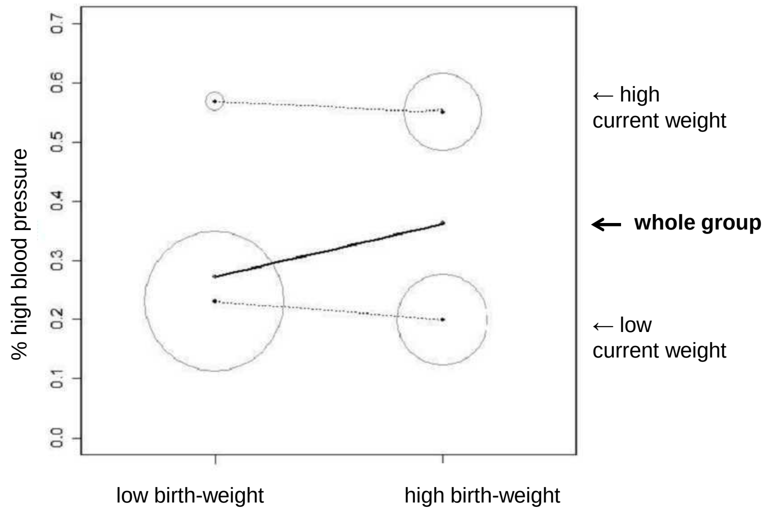
Divided in subgroups (current weight categorical):

	normal bp	high bp	total	%high bp
low bw	354	132	486	27.2%
low cw	329	99	428	23.1%
high cw	25	33	58	56.9%
high bw	328	186	514	36.2%
low cw	221	55	276	19.9%
high cw	107	131	238	55.0%
total	682	318	1000	31.8%
low cw	550	154	704	21.9%
high cw	132	164	296	55.4%

→ in both subgroups having low birth-weight seems to increase risk for having high blood pressure

... what's happening?

Example I: Simpson's paradox



Example II

Now: same data, *cw* and *bp* continuous variables

birth weight (bw) —?→ *blood pressure (bp)*

current weight (cw) ?

Linear regression models:

1) Model 1

$$bp = \beta_{c1} + \beta_{bw1} \cdot bw \longrightarrow \beta_{bw1} = 2.49$$

(unconditional effect of *bw* on *bp*)

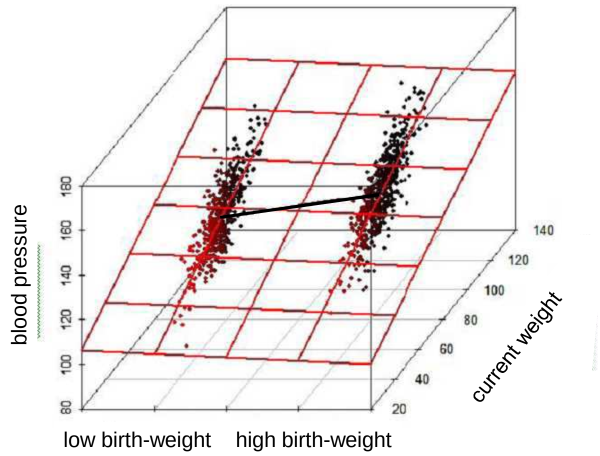
2) Model 2

$$bp = \beta_{c2} + \beta_{cw2} \cdot cw + \beta_{bw2} \cdot bw \longrightarrow \beta_{bw2} = -2.94$$

(conditional effect of *bw* on *bp*, given *cw*)

... what should be our conclusion regarding the effect of *bw* on *bp*?
Should we control for current weight?

Example II: Lord’s paradox



The problem in a nut-shell

The issue: which analysis is the correct one?
(conditioning for current weight yes/no?)

The real question: is the effect we are interested in influenced by other variables? *(confounding)*

If so, for which variables do we need to correct to obtain the correct estimate of the effect? *((set of) confounders*)

Solution: using causal graphs (before performing the analysis!) can help unravel this problem

Note: when building prediction models, this is not *relevant (nor a problem)*

The problem only occurs as soon as we want to explore/interpret relationships between variables (causally)!

Causal graphs

Causal graphs: a graphical representation of causal relationships between variables

A ⟶ B

- force the researcher to be explicit about (often implicit) causal assumptions
- a non-parametric (gives general results), conceptual tool,
  to be used in addition to empirical analysis and presentation of numerical results
- linked to the theory of structural equation modelling and path analysis

The use of causal graphs

Given the causal assumptions, causal graphs and the associated theory enables the researcher to:

- identify confounding
- identify which variables need be controlled for to obtain unconfounded effect estimates
- identify which variables need NOT be controlled for to obtain unconfounded effect estimates
- *help unravel statistical paradoxes*

Detecting confounding
… in simple situations

A variable is a confounder when:

- it is associated with the exposure of interest

  cw is a confounder → it is associated with bw...

- it is independently associated with the outcome

  … cw is (independently of bw) associated with bp and...

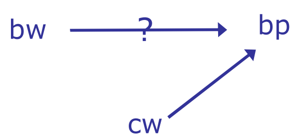
- it is **not** on the causal pathway

  … the following does ***not*** *hold:*

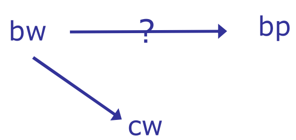
  bw → cw → bp

Possible situations

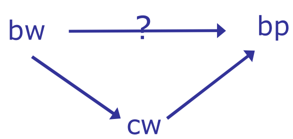
Is *cw* a confounder?



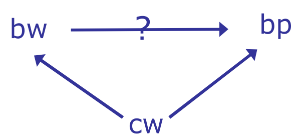
1) *cw* is not a confounder



2) *cw* is not a confounder



3) *cw* is not a confounder

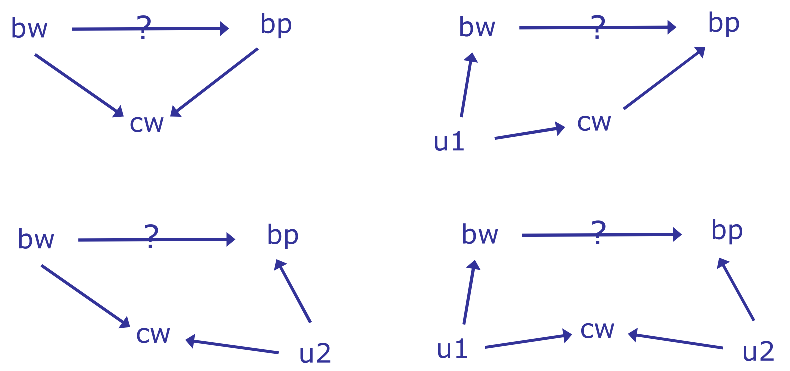


4) *cw* is a confounder

But what about more complex situations? When does confounding occur?

Confounding in more complex situations

Is there confounding? (*u1*, *u2* unmeasured variables)
If so, for which variables should we correct in our analysis?

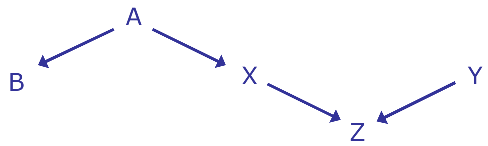


→ **the theory of causal graphs can help!**

Introducing causal graphs (1)

Directed acyclic graphs (DAGs)

Causal graphs: a graphical representation of causal relationships between variables
Parents, children, ancestors and *descendants*



A path on a causal graph does not need to follow the directions of the arrows:

Z – X – A , B – A – X

Collider: a particular node on a path such that both the preceding and subsequent nodes on the path have directed edges going into that node

X → Z ← Y

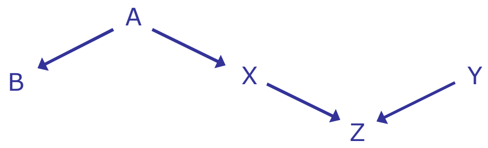
Any path which contains a collider, is called a ***blocked path***

A – X – Z – Y (Z is a collider on this path) ;

otherwise ***unblocked***

B – A – X – Z

Introducing causal graphs (2)
Statistical associations



Two variables will only be statistically associated in the population as a whole if:

EITHER one is a cause of the other
X causes Z; A causes B

OR they share a common cause or ancestor
B and X are caused by A, as are B and Z

Introducing causal graphs (3)
Conditioning on parents

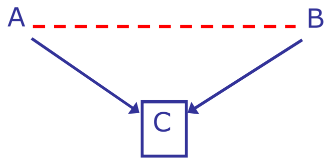
Conditioning on a variable is graphically represented by placing a box around that variable

Conditional on its parents, a variable C will be independent of all variables which are not descendants of C

Introducing causal graphs (4)
Conditioning on children

Conditioning on children influences the associations between parents/ancestors of that variable.

A: the battery is low
B: the gas tank is empty
C: the car does not start

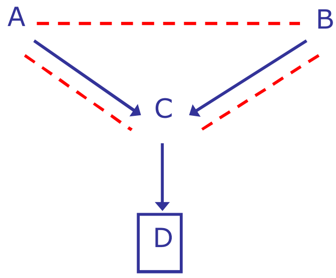


Introducing causal graphs (5)
Conditioning on descendants

Conditioning on D can

- introduce an association between A and B (within strata of D)
- also change the magnitude of the associations between A-C and B-C

A: the battery is low
B: the gas tank is empty
C: the car does not start
D: biking to work



Introducing causal graphs (6)
The effect of conditioning

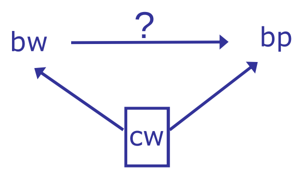
Let's summarize:

Conditioning can:
- remove marginal dependencies
- introduce new (conditional) dependencies
- alter the magnitude of already existing dependencies

Back to our problem:
When does confounding occur? When should you condition and on what variables?

Detecting confounding
in relation to a particular effect

General recipe for detecting confounding:

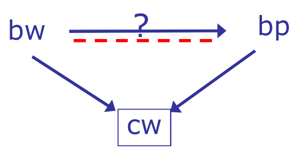


1) Draw the causal graph
2) Eliminate all effects from treatment/exposure
3) In the resulting graph: are there any unblocked paths leading from treatment to outcome?
   - Yes → confounding (*conditioning/correcting is needed)*
   - No → no confounding

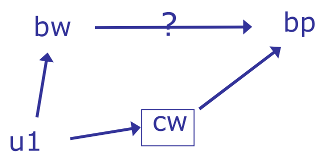
→ *Confounding can be eliminated by blocking these unblocked paths by conditioning on appropriate variables (backdoor criterium/d-separation)*

More complex situations (cont.)

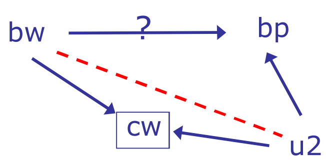
Is there confounding?



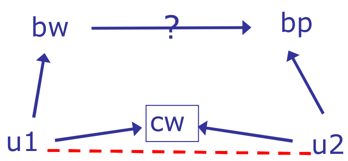
1) No confounding (correcting may even introduce confounding/bias!)



2) Confounding → correcting for *cw* (or *u1*) to eliminate confounding



3) No confounding (correcting may even introduce confounding/bias!)



4) No confounding (correcting may even introduce confounding/bias!)

Back to Lord's paradox (Pearl 2016)

Is there an effect of gender on weight gain?

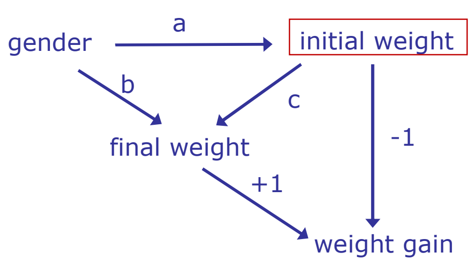
Two possible answers:

Statistician 1

total effect *gender* → *weight gain* (uncorrected → no gender effect)

direct effect *gender* → *weight gain* (corrected for initial weight → gender effect)

Statistician 2



Conclusion: there is no contradiction.

Each analysis estimates a different effect and answers a different research question

Linear model, standardized variables:

total effect	= b*1+a*(c*1-1)	= b - a(1-c)	(combines all three paths)
direct effect	= b*1	= b	(``skips'' paths through initial weight)

Concluding remarks

- The paradoxes are but a symptom of a larger, underlying problem:
  correctly identifying and correcting for confounding
  (*not always as straightforward as it may seem...*)
- This problem cannot be solved by applying a statistical test!
- Causal graphs help identify confounding and identify variables that must be measured and controlled for in the statistical analysis to obtain unbiased effect estimates (*with sometimes surprising results*)
- Correcting for (all) co-variates/background variables `just to be sure' is dangerous and may even introduce bias!
- Beware: in one situation, different causal models can be equally plausible, but can have different consequences for your analysis

A selection of literature

- Arah, OA, `The role of causal reasoning in understanding Simpson's paradox, Lord's paradox, and the suppression effect: covariate selection in the analysis of observational studies', *Emerging Themes in Epidemiology* 5 (2008)
- Greenland, S, Pearl, J, Robins, JM, `Causal diagrams for epidemiologic research', *Epidemiology* 10 (1999) 37-48
- Hernán, MA, Hernández-Diaz, S, Robins, JM, `A structural approach to selection bias', *Epidemiology* 15 (2004) 615-625
- Lord, FM, `A paradox in the interpretation of group comparisons', *Psychological Bulletin* 68 (1967) 304-305
- Pearl, J, *Causality. Models, reasoning and inference* (Cambridge 2000)
- Pearl, J, `Lord's paradox revisited – (Oh Lord! Kumbaya!)', *Journal of causal inference* 4 (2016)
- Tu, Y-K, Gunnel, D, Gilthorpe, MS, `Simpson's paradox, Lord's paradox, and Suppression effect are the same phenomenon – the reversal paradox', *Emerging Themes in Epidemiology* 5 (2008)

Next Help! Statistics! lecture:

December 11, 2018

Room 16

Non-parametrical tests

Douwe Postmus