

Stat 202 2015S - W7 - Wed

§2.6 The Question of Causation

New

In many studies the goal is to establish that changes to the explanatory variable Cause changes to the response variable

Famous saying: Correlation does not imply Causation

More basic question: What causes association?

Consider the following observed associations between x and y

1. x : mother's body mass index
 y : daughter's body mass index
2. x : amount of artificial Sweetener (saccharin) in ^{rats} diet.
 y : count of tumors in rats bladder
3. x : a student's SAT score as high school senior
 y : a student's first year GPA
4. x : monthly flow of money into stock mutual funds
 y : monthly rate of return for the stock market
5. x : whether a person regularly attends religious services
 y : how long a person lives
6. x : number of years of education a worker has
 y : the worker's income

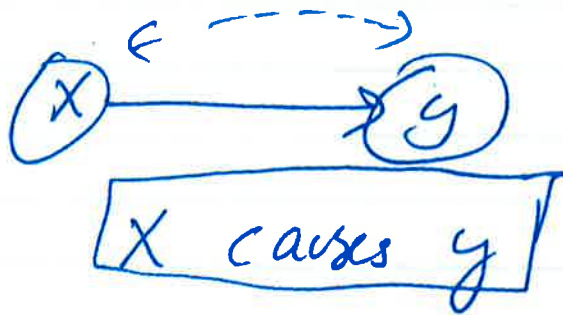
Links between variables that
can cause association

—————→ solid arrow denotes
direct causation

←-----→ dashed arrow denotes
association

Possibility
1

One possibility is direct causation



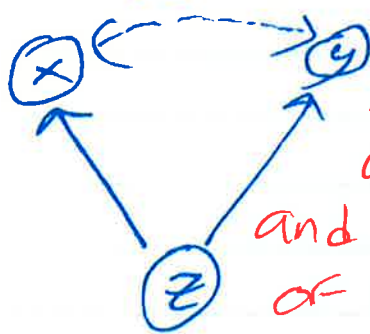
According to book items 1 and 2 are
examples of direct causation

But even when direct causation is present
very often it is not the complete
explanation of an association

To show X causes Y best evidence is to
change X hold all other factors fixed and
observe Y . If Y changes we have good
reason to think X causes change in Y .

Common responsepossibility
2

X and y ~~can~~ show
a common response
to a third (often lurking)
variable Z



A lurking variable is a variable that is not among the explanatory or response variables in the study and yet may influence the interpretation of the relationships among those variables.

Items 3 and 4 are examples of a common response.

Item 3: Student's aptitude
Item 4: Economy

No ~~direct~~ causal link between X and Y

but association none the less

Men are more likely to get aggressive treatment (e.g. by pass surgery) for complaints of chest pain than women. Is this association between gender and treatment due to discrimination?

Maybe but not necessarily, women tend to get problems when they are much older so doctors may hesitate to prescribe aggressive treatments because of age. Correlation is real, but causation suspects.

possibility
3

Confounded

Pg 7

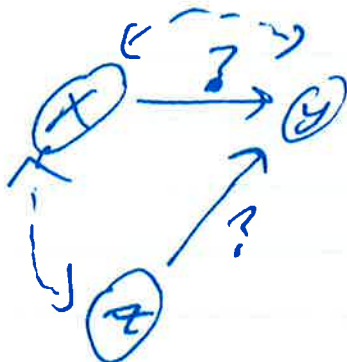
Two variables are confounded when their effects on a response variable cannot be distinguished from each other

The confounded variables may be explanatory variables or may be lurking variables.

Eg Mother / Daughter BMI
Nature vs. Nurture.

When many uncontrolled variables are related to a response variables ask whether confounding prevents you from drawing conclusions about causation

Items 5 and 6 are confounding



Even a very strong association between two variables is not by itself good evidence that there is a cause-and-effect link between the variables

How to establish a causal connection between x and y

Only compelling method - carefully designed experiment in which effects of all possible lurking variables are controlled

Needless to say that it is not always possible to do such an experiment

EG: For ethical reasons it is not possible to run an experiment where we force people to smoke.

That would give the best evidence however that smoking causes cancer.

Nevertheless it has been agreed upon that smoking does cause cancer

How? By what criteria

What criteria allow us to conclude smoking causes cancer absent controlled studies

- * The association is strong between smoking and lung cancer
- * The association is consistent. Across country and different groups (reduces chance that a lurking which might be specific to one group causes association)
- * Higher dose associated with strong response
- X Alleged cause precedes effect in time
- X Alleged cause is plausible - experiment in rats show smoke causes cancer,