Chapter 2: Observational studies

Marius Ionescu

09/01/2011

Observational study

Definition

In observational studies the patients assign themselves to different groups.

Observational study

Definition

In observational studies the patients assign themselves to different groups.

Observational study

Definition

In observational studies the patients assign themselves to different groups.

Examples

Observational study

Definition

In observational studies the patients assign themselves to different groups.

Examples

- smoker vs non-smoker

Observational study

Definition

In observational studies the patients assign themselves to different groups.

Examples

- smoker vs non-smoker
- male vs female

Observational study

Definition

In observational studies the patients assign themselves to different groups.

Examples

- smoker vs non-smoker
- male vs female
- sick vs healthy

Observational study

Definition

In observational studies the patients assign themselves to different groups.

Examples

- smoker vs non-smoker
- male vs female
- sick vs healthy

Observational study

Definition

In observational studies the patients assign themselves to different groups.

Examples

- smoker vs non-smoker
- male vs female
- sick vs healthy

They are not experiments due to the prevalence of hidden confounding variables.

Examples: smoking

Example

What types of confounding variables do you think are involved when studying the effect of smoking on the health?

- Many good statisticians did not believe the evidence against ciggaretes!

Example

Example (Ultrasound and low birthweight)

Investigators ran an observational study to find out if the ultrasound examinations can cause a low birthweight at the John Hopkins Hospital in Baltimore. Babies exposed to ultrasound in the womb had lower birthweight, on average, than babies who were not exposed. Is this evidence that ultrasound causes lower birthweight?

Example

Example (Vitamin C and Treatment of Advanced Cancer)

In 1976, Cameron and Pauling performed the following study: They gave vitamin C to 100 patients believed to be terminally ill from advanced cancer and studied subsequent survival. For each such patient, 10 historical controls were selected of the same age and gender, and other similar characteristics (*matched* sampling). They found that, as a group, patients receiving vitamin C survived four times longer than the controls. However, few scientists believe today that vitamin C has any effect on cancer. What do you think went wrong in this experiment?

Controlling for the confounding variable

Fact

Controlling for the confounding variable

Fact

- *One source of trouble in observational studies is that subjects differ among themselves in crucial ways besides the treatment.*

Controlling for the confounding variable

Fact

- *One source of trouble in observational studies is that subjects differ among themselves in crucial ways besides the treatment.*
- *Sometimes these differences can be adjusted for, by comparing smaller and smaller and more homogeneous subgroups.*

Controlling for the confounding variable

Fact

- *One source of trouble in observational studies is that subjects differ among themselves in crucial ways besides the treatment.*
- *Sometimes these differences can be adjusted for, by comparing smaller and smaller and more homogeneous subgroups.*
- *This technique is called controlling for the confounding factor.*

Berkeley Discrimination Case

Example

An observational study on sex bias in admissions was done at UC Berkeley.

Berkeley Discrimination Case

Example

An observational study on sex bias in admissions was done at UC Berkeley.

- There were 8442 men who applied for admission to graduate school and 4321 women.

Berkeley Discrimination Case

Example

An observational study on sex bias in admissions was done at UC Berkeley.

- There were 8442 men who applied for admission to graduate school and 4321 women.
- About 44% of the men and 30% of the women were admitted.

Berkeley Discrimination Case (cont'd)

Example

Let's find the responsible major(s); here is the list of the 6 largest majors:

Major	Males	Females	#applicants M	#applicants F
A	62%	82%	825	108
B	63%	68%	560	25
C	37%	34%	325	593
D	33%	35%	417	375
E	28%	24%	191	393
F	6%	7%	373	341

Simpson's Paradox

Fact (Simpon's paradox)

The relationship between percentages in subgroups ***can*** *be reversed when the subgroups are combined.*

Example

Example

Two baseball players are up for the batting trophy for the highest year-long batting percentage. Player X had a higher batting average than Player Y in both the first and second halves of the season. Furthermore, each player had 1000 at-bats for the season. However, Player Y won the award. Is this possible, or was a mistake made?

Example

Suppose that we have the following admission data for two departments and two groups (by sex)

	Male	Female	% admission
Dept 1	800	300	60% each group
Dept 2	400	600	30% each group

Is there any discrimination at this school?