The Logic of Research

Soc 357

Fall 2006

General Form of a Proposition

- ***Conceptual***: For population (P) in condition (C), independent variable (X) causes dependent variable (Y)
- ***Operational:*** For sample (p) in condition (c), independent variable (x) has a statistical association with dependent variable (y)

Induction

- ***Induction:*** reasoning from the specific to the general = Empirical generalization
- No logical proof of induction: future cases may be different from those you have seen
- ***Sampling theory*** tells us how we can use the observations we have make probabilistic statements about future cases

Induction Examples - 1

- This ice is cold

  All ice is cold

- Teenagers get lots of speeding tickets

  All teenagers speed

Induction Examples - 2

- 62 percent of voters in a random sample of 400 registered voters (polled on February 20, 2004) said that they favor John Kerry over George W. Bush for President in the 2004 Presidential election.
- This supports the generalization that between 57 percent and 67 percent of **all** registered voters favor Kerry over Bush for President (at or around the time the poll was taken)

Use of Induction in Social Science

- Field observation, exploratory research uses induction in order to ***build*** theories
- Use induction in sample studies, where we generalize about a population based on a smaller group

Deduction

- ***Deduction:*** reasoning from the general to the specific, following the rules of logic
- Used to ***test*** theories. But we have to be careful about what conclusions we draw.

  Ex. 1: *When it rains, the ground gets wet*
  It is raining
  Therefore the ground is wet.

  Ex. 2: *When it rains, the ground gets wet*
  The ground is wet
  Therefore it is raining.

Example of Illogic of "Proof"

- Theory: Students who participate in extra-curricular activities are more likely to vote
- Data: A survey shows student athletes are more likely to vote than non-athletes
- Does this mean that your theory is true?

Illogic: Data “Proves” Theory

If theory is correct, then X is true.
X is true.
Therefore, theory is correct.

INVALID LOGIC: ***X might be true for another reason***

Logic of Falsification

If theory is correct, then X is true.
X is not true.
Therefore, theory is not correct.

VALID LOGIC

Example with logic of falsification

- Theory: Students whose parents talk politics at the dinner table are more likely to vote
- Data: A survey shows students whose parents talk politics at dinner are no more likely to vote than students whose parents talk about non-political stuff at the dinner table
- Can we conclude that the theory is false? YES (assuming reliable & valid measures)

Falsification

- We cannot prove theories to be correct
- We CAN prove theories to be INCORRECT
- Research proceeds on a logic of falsification
  - We subject theories to tests which could falsify them
  - If a theory avoids falsification, we say it is “confirmed” (not proven)
  - If a theory repeatedly avoids falsification, we build our confidence that it is correct, but it could still be proven wrong later

The Logic of Research

- We try to falsify other theories while confirming our own
  - Theory A: Participating in activities → higher voting rates among youth
  - Theory B: Parental Socialization → higher voting rates youth
  - Gather data that confirms Theory A (student athletes & non-athletes data) while falsifying Theory B (parents who talk politics vs. parents who don't talk politics)

Causation

- It is impossible directly to observe causation
- Criteria for theorizing about causation:
  - **Statistical association**: two things vary together
  - **Direction of influence**: Cause precedes effect in time
  - **Elimination of rival hypotheses:** Other variables are ruled out as possible explanations for the relationship
  - **We can identify the mechanism** for the cause-effect relationship, we know how it works

Full Logic of Hypothesis Testing

Research Syllogism:

If A causes B {theory}

And if X measures/indicates A

{measurement assumption}

And if Y measures/indicates B

{measurement assumption}

Then X will be statistically associated with Y

{prediction}

Confirmation of Theory

Research Syllogism:
If A causes B **{theory}**
And if X measures/indicates A **{measurement assumption}**
And if Y measures/indicates B **{measurement assumption}**
Then X will be statistically associated with Y **{prediction}**

Data 1:
X is statistically associated with Y **{prediction is correct}**

We cannot prove that A is associated with B, but we can say the data *confirms* or *supports* our theory that A causes B (also confirms measurement assumptions)

Disconfirmation of Theory

Research Syllogism:
If A causes B {theory}
And if X measures/indicates A {measurement assumption}
And if Y measures/indicates B {measurement assumption}
Then X will be statistically associated with Y {prediction}

Data 2:
X is NOT statistically associated with Y {prediction is wrong}

Then at least one assumption must be wrong. Either
A does not cause B {theoretical assumption is wrong} AND/OR
X does not measure A {measurement assumption is wrong} AND/OR
Y does not measure B {measurement assumption is wrong}

But: Falsification may show up due to sampling error or extraneous variables – more on this later

Statistical Association

- A change in one variable is associated with a change in another variable, in a way that is not likely to have occurred just by chance

Statistical Association

- There are "tests of significance" that we will not do
- Rule of thumb: association doesn't change direction if one or two people changed responses
- Instead, we'll focus on three general outcomes
  - Confirms the hypothesis
  - Disconfirms the hypothesis
  - Indeterminate

*** The most common mistake is to call a zero relationship "indeterminate"

Types of Statistical Association (Bivariate)

Independent	Dependent	Statistic
Qualitative	**Qualitative**	Difference of Conditional Percentages
Qualitative	**Quantitative**	Difference of Means
Quantitative	**Quantitative**	Correlation or regression

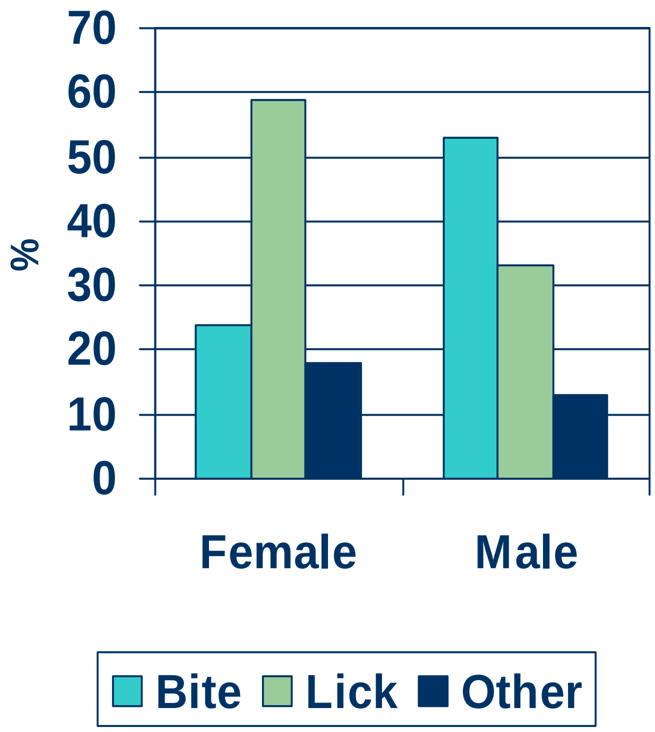
Difference of Conditional Percentages: Sex and Ice Cream Cone Eating

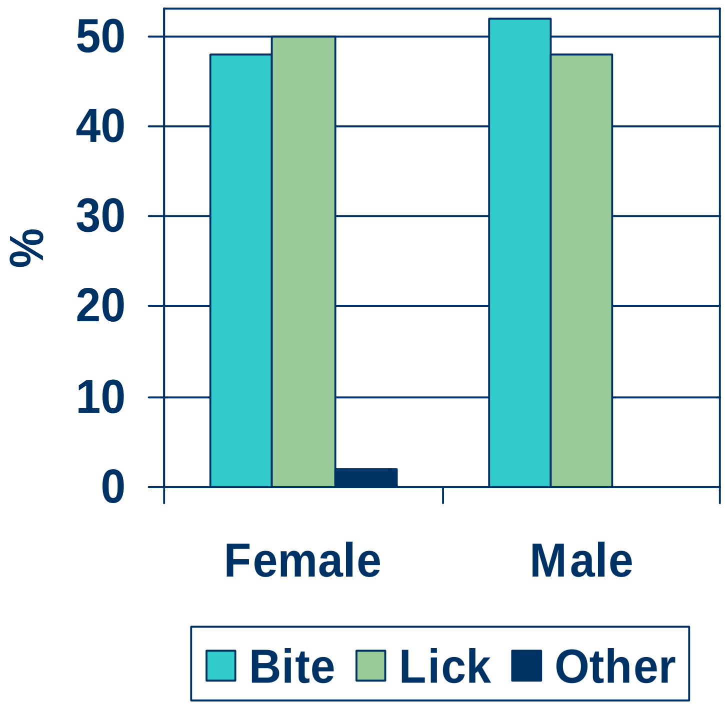
	Male	Female
Bite	53%	24%
Lick	33%	59%
Other	13%	18%
Total	99%*	101%*
(N)	(15)	(17)

Statistical Association:

- **Males bit 53% of the time compared to 24% of the women (a percentage difference of 29%)**
- **Females licked 59% of the time compared to 33% for males (a percentage difference of 26%)**
- **"Other" was only slightly different for men and women.**

Qualitative relationships: Association & Non-Association





Difference of Conditional Means: Sex and Time to Complete Sales Transactions

	Men	Women
Mean Seconds for Transaction	27.1	40.5
(N)	(20)	(27)

Interpretation: Women took 13.4 seconds longer than men, on average, to complete their sales transactions.

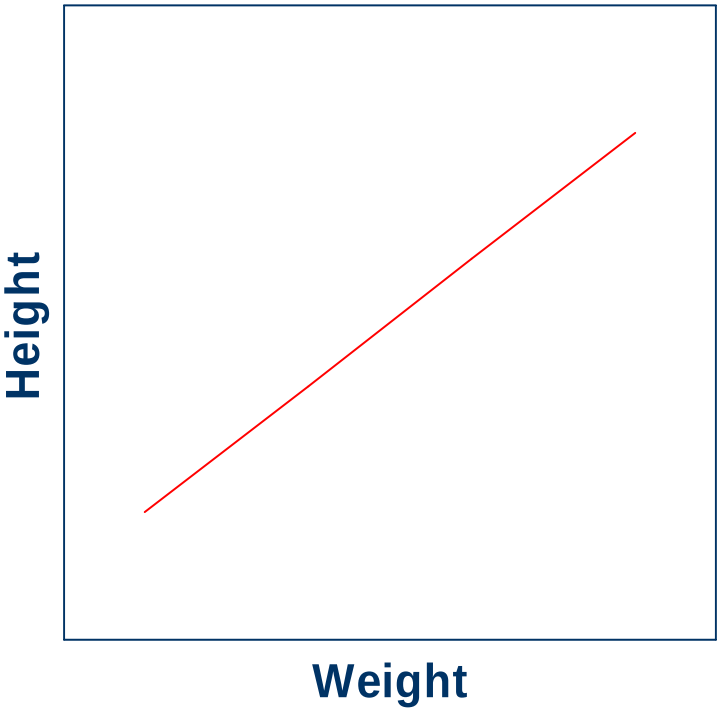
Correlations

- ***Correlations:*** calculating if a change in one variable is associated with a change in another variable
- Range between –1 (perfect negative correlation) to +1 (perfect positive correlation).

See: http://noppa5.pc.helsinki.fi/koe/corr/cor7.html

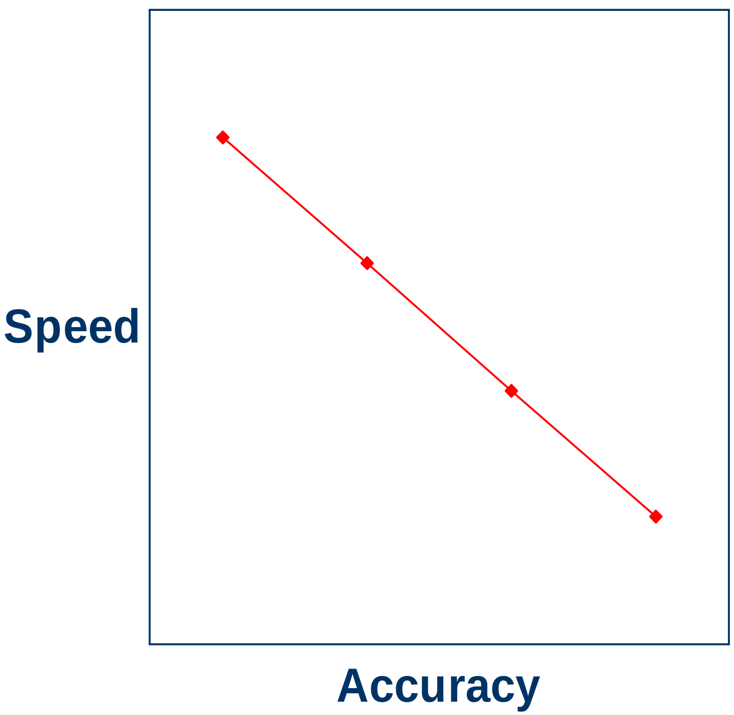
Quantitative relationships – 1

- **Positive**: when one variable is greater, the other tends to be greater
- Eg. Height is *positively* associated with weight. The taller you are, the more you are likely to weigh.



Quantitative relationships – 2

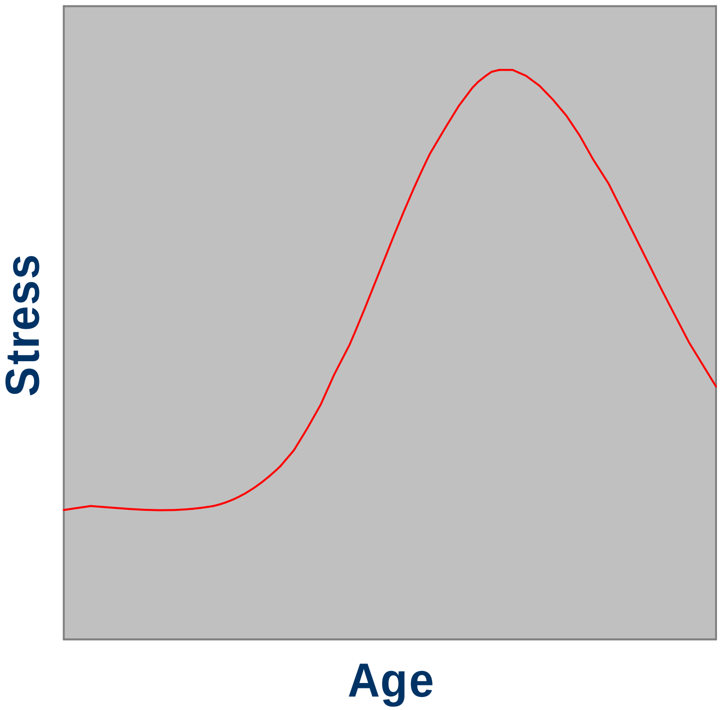
- **Negative**: When one variable is greater, the other tends to be smaller
- Eg. Speed is *negatively* associated with accuracy. The more you rush, the worse your accuracy is.



Quantitative relationships - 3

- ***Curvilinear***: Any non-linear relation, but especially one that is first positive and then negative, or vice versa

- Eg. Stress is related *curvilinearly* to age. Middle aged people feel the most stress, while young & old report less stress.



Spuriousness

- Problem: A statistical relationship between X and Y may be the result of some other variable, Z.
- Solution: hold extraneous variables “constant”
  - We check that the relationship still exists for different groups of variable z
  - E.g. Smoking, gender, and lung cancer
- Problem: Can’t control for everything