Stat 1600

Loren Heun

WMU

July 7, 2018

Statistics and Data Analysis

Lecture 2 Knowledge and data

Outline

Data Presentation # 1

- Statistics and Data
- Variable Types
- Summarizing Categorical Data

Knowledge and data

- Step-by-step knowledge building
- Some fallacies in interpreting evidence

Building knowledge step by step

1. Conceptualize the problem
   - This is the problem of interest. State it broadly.

Building knowledge step by step

2. Operationalize the problem
   - The investigator formulates specific questions to answer
   - What do you want to measure? These will become our dependent variables.

Building knowledge step by step

3. Design the Study
   - How will you select your sample?
   - How many groups will you compare?

Building knowledge step by step

4. Collect the Data

- What instrument or technique will you use to collect data?
- Will you use a survey, questionnaire, interview, observation?
- Are you measuring a variable that will require special equipment/technology?

Building knowledge step by step

5. Analyze the Data
   - Are you comparing means? Percentages?
   - Are differences statistically significant?

Building knowledge step by step

6. Conclusions
   - Do the results generalize to a larger population?
   - Did you show cause-and-effect or just associations?

Building knowledge step-by-step

7. Disseminate results
   - How are you sharing your results?
   - News reports?
   - Scientific journals?
   - Etc. . .

Step-by-step Knowledge Building

- Conceptualize the problem – broad wording
- Operationalize the problem – specific questions
- Design the Study – how to select samples
- Collect the Data – measurement instrument
- Analyze the Data – comparing what
- Conclusions – repeatability and generalization
- Disseminate results – presentation of results

Example – comparing wt loss program

- Zone
  - Balances carbohydrates, protein, fat
- Atkins
  - Low carbohydrate, high fat, unrestricted calories
- LEARN
  - Low fat, and based on national guidelines
- Ornish
  - Low fat, high carbohydrate, unrestricted calories

How are we going to design a study to compare these?

Building knowledge step-by-step

1. Conceptualize the problem
   - Which weight loss program is most effective?
   - Which one is most healthy?

Building knowledge step-by-step

2. Operationalize the problem
   - How do we measure 'effective' and 'healthy?'
   - At what time point are we interested in measuring? In 2 weeks, 2 months, 2 years?
   - Are we comparing average weight loss or perhaps the percentage of people who lost 15 pounds or more?
   - How do we measure healthy? LDL cholesterol reduction, BP reduction, Glucose levels?

Building knowledge step-by-step

3. Design the Study
   - Where are we recruiting our subjects?
   - How long will the study last?
   - Do they choose the diet or do we randomly assign them to it?
   - How do we ensure they stay on a diet?
   - What do we do with participants who go off the diet, do we eliminate them from the study?

Building knowledge step-by-step

4. Collect the Data
   - How many times will we measure their weights?
   - Are we taking blood samples? Urine samples? Are we sending samples to the lab?

Building knowledge step-by-step

5. Analyze the Data
   - Are there significant differences in average weight loss between the diet groups?
   - Are there differences in cholesterol, blood pressure, glucose levels or other biochemistry measures relating to health?
   - Are there differences in how well participants adhere to each diet plan?

Building knowledge step-by-step

6. Conclusions

- After analyzing the results what do we conclude is the best diet? Why?
- Can we generalize results to the larger population?
- Are we sure weight loss can be attributed to the diet?

Building knowledge step-by-step

7. Disseminate results
   - How are we going to present the results?
   - What tables and graphs would make the study easy to read and understand?

Questioning results of a study

If we are reading the results of a study we need to be able to ask ourselves some questions:

- What is the long-term result (perhaps the results will differ if measurements are taken at longer time points)?
- What was the sample and to what population are we trying to generalize the results (males, females, age, ethnic differences)? We want to make sure we can generalize to the population outside of the study sample.
- Was the sample size large enough to allow for generalizing to the outside population?

Questioning results of a study

There is variation in study design, and some studies are designed better than others. We need to be able to judge the validity and reliability of a study.

Fallacies in interpreting evidence

1. Lack of evidence
   - "No proof that the drug is unsafe."
   - This is flawed as a lack of evidence does not mean the contrary is true and that the drug is safe.
2. Anecdotal evidence
   - "Testimonies of real people this worked for . . ."
   - Infomercials.
   - Existence does not mean prevalence. Perhaps the drug or supplement worked for some people, but does that mean it is effective for the broader population?

Fallacies in interpreting evidence

- Correlation equals causation
  - "married people are happier than single people."
  - Did marriage cause the 'happier' outcome? Maybe happy people are the ones who tend to get married.
  - Two things happening at the same time does not mean one causes the other.

Examples of Wrong Reasoning Leading to Wrong Conclusions

- Lack of evidence fallacy. The fallacy lies in the reasoning that lack of evidence means the contrary is true.
- Anecdotal evidence fallacy. The fallacy lies in the reasoning that existence means prevalence.
- Correlation equals causation fallacy. The fallacy lies in the reasoning that "two things happening together" must mean one causes the other.