Statistics
Association
Causation

STATISTICAL LITERACY

3/99
1

Wollongong University

March 1999

Milo Schield

Augsburg College Minneapolis Minnesota USA

www.augsburg.edu/ppages/schield
schield@augsburg.edu

Statistics
Association
Causation

Definitions

3/99
2

Statistics studies variation in data.
What are the natures and causes of variation?

Statistical inference studies the results of chance: sampling distributions, confidence intervals and hypothesis tests.

Statistical literacy studies the use of statistics as evidence in arguments.

Statistics
Association
Causation

Natures and Causes of Variation

3/99
3

This variation is expected if due to chance.

This variation is unlikely if due to chance.

This variation is unlikely due to chance.

This variation is unlikely to be due to chance.

This variation is likely to be due to a determinate cause.

Statistics
Association
Causation

Analyzing the influence of chance

3/99
4

Probability: The higher the probability, the more reason one has to believe that the outcome is true (or will occur).

Confidence Intervals. The higher the level of confidence, the more reason one has to believe that the fixed interval contains the fixed parameter.

Hypothesis tests: The smaller the p-value, the more reason one has to believe that the alternate is true.

Statistics
Association
Causation

Analyzing the Influence of Bias

3/99
5

The death rate in Washington DC is twice as high as that in Alaska.

Suppose a randomly selected group of Alaskans are moved from Alaska to Washington DC.

Are they twice as likely to die?

Yes because ….

No because ….

Statistics
Association
Causation

Distinguishing Association From Causation

3/99
6

In regressing the value of houses on the number of baths, we find that we can expect an 25,000 increase in the price of the house

- **in houses having an additional bathroom**
- **for each additional bathroom**
- **when adding an additional bathroom.**

Statistics Association Causation

Predicting Reversal: % of Murders - Death Penalty

3/99 7

The death penalty was given to
- 11.9% of white murderers and
- 10.5% of black murders.

The death penalty was given in
- 14.0% of the cases with a white victim and
- 5.4% of the cases with a black victim.

We could have a Simpson's Paradox reversal.

Statistics Association Causation

From Statistical Association To Policy Prescription

3/99 8

Statistics are used to support policy prescriptions.

Data from an observational study is transformed into supporting a public policy as follows:

1. **Association or Correlation**
2. **Causation [this is often implicit]**
3. **Prediction or Prescription**

Examples: Accident rates and car phones
Death rates and radon levels

Statistics Association Causation

Statistical Literacy and General Thinking

3/99 9

Statistical Literacy must be related to general literacy.

Statistical Literacy should focus on general tools and techniques that students will use again in a variety of courses.

Statistics Association Causation

General Literacy: Basic Tools

3/99 10

All students should be able to
- **recognize a claim: a prediction, an explanation, an evaluation, a prescription, a generalization, etc..**
- **evaluate the disputability of a claim.**
- **identify an argument used to support a conclusion.**
- **distinguish premises and conclusion in an argument.**
- **analyze the support given for the truth of a claim.**
- **evaluate the strength of an argument.**

Statistics Association Causation

Statistical Literacy: Basic

3/99 11

Students should be able to
- **read and express statistics (counts, percents, rates and statistical measures) in both tables and graphs.**
- **distinguish association from causation.**
- **recognize that association is not causation.**
- **distinguish chance from a determinate cause.**
- **distinguish common causes from a direct cause(s).**
- **distinguish experiments from observational studies.**

Statistics Association Causation

Conclusion

3/99 12

Statistical Literacy focuses on the use of numerical statistics to identify causes and to recommending actions and decisions.

Students should be able to read and evaluate broad arguments involving statistics:
- **The Bell Curve by Herrnstein and Murray**
- **Population by Julian Simon**
- **Books by Thomas Sowell**

STATISTICAL LITERACY

Wollongong University

March 1999

Milo Schield

Augsburg College Minneapolis Minnesota USA

www.augsburg.edu/ppages/schield

schield@augsburg.edu

Definitions

Statistics studies variation in data.
What are the natures and causes of variation?

Statistical inference studies the results of chance: sampling distributions, confidence intervals and hypothesis tests.

Statistical literacy studies the use of statistics as evidence in arguments.

Natures and Causes of Variation

This variation is expected if due to chance.

This variation is unlikely if due to chance.

This variation is unlikely due to chance.

This variation is unlikely to be due to chance.

This variation is likely to be due to a determinate cause.

Analyzing the influence of chance

Probability: The higher the probability, the more reason one has to believe that the outcome is true (or will occur).

Confidence Intervals. The higher the level of confidence, the more reason one has to believe that the fixed interval contains the fixed parameter.

Hypothesis tests: The smaller the p-value, the more reason one has to believe that the alternate is true.

Analyzing the Influence of Bias

The death rate in Washington DC is twice as high as that in Alaska.

Suppose a randomly selected group of Alaskans are moved from Alaska to Washington DC.

Are they twice as likely to die?

Yes because ….

No because ….

Distinguishing Association From Causation

In regressing the value of houses on the number of baths, we find that we can expect an 25,000 increase in the price of the house

- **in houses having an additional bathroom**
- **for each additional bathroom**
- **when adding an additional bathroom.**

Predicting Reversal: % of Murders - Death Penalty

The death penalty was given to

- 11.9% of white murderers and
- 10.5% of black murders.

The death penalty was given in

- 14.0% of the cases with a white victim and
- 5.4% of the cases with a black victim.

We could have a Simpson's Paradox reversal.

From Statistical Association To Policy Prescription

Statistics are used to support policy prescriptions.

Data from an observational study is transformed into supporting a public policy as follows:

1. **Association or Correlation**
2. **Causation [this is often implicit]**
3. **Prediction or Prescription**

Examples: Accident rates and car phones
Death rates and radon levels

Statistical Literacy and General Thinking

Statistical Literacy must be related to general literacy.

Statistical Literacy should focus on general tools and techniques that students will use again in a variety of courses.

General Literacy: Basic Tools

All students should be able to

- **recognize a claim: a prediction, an explanation, an evaluation, a prescription, a generalization, etc..**
- **evaluate the disputability of a claim.**
- **identify an argument used to support a conclusion.**
- **distinguish premises and conclusion in an argument.**
- **analyze the support given for the truth of a claim.**
- **evaluate the strength of an argument.**

Statistical Literacy: Basic

Students should be able to

- **read and express statistics (counts, percents, rates and statistical measures) in both tables and graphs.**
- **distinguish association from causation.**
- **recognize that association is not causation.**
- **distinguish chance from a determinate cause.**
- **distinguish common causes from a direct cause(s).**
- **distinguish experiments from observational studies.**

Conclusion

Statistical Literacy focuses on the use of numerical statistics to identify causes and to recommending actions and decisions.

Students should be able to read and evaluate broad arguments involving statistics:

- **The Bell Curve by Herrnstein and Murray**
- **Population by Julian Simon**
- **Books by Thomas Sowell**