

5 August 2009 ASA 1

Association vs. Causation in Titles of News Stories

Milo Schield
W. M. Keck Statistical Literacy Project
Augsburg College

Robert Raymond
Emeritus, University of St. Thomas

ASA JSM 2009
Washington DC 5 August 2009
www.StatLit.org/pdf/2009SchieldRaymondASA6up.pdf

5 August 2009 ASA 2

Causation-Association Important Distinction

In a survey, statistical educators were asked to rank statistics topics in importance.

- Causation-association was second
- Variation was first.

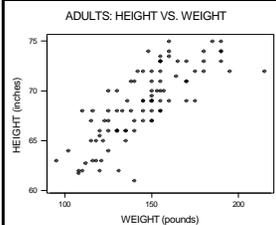
See McKenzie (2004).

If statistical literacy is to be empirically based, the causation-association distinction must be analyzed based on how it is conveyed in the news.

5 August 2009 ASA 3

Problems Distinguishing: % who said "Yes"

- 77%: Adults who weigh more tend to be taller than those who weigh less. (A)
- 59%: As weight increases, height increases (B)
- 31%: As an adult's weight increases, their height tends to increase. (B)
- 15%: If weight increases, height will tend to increase. (B)



5 August 2009 ASA 4

Selected 1,904 News Stories

Selected media reports using numbers if they:

- had the word 'study' or 'survey' in the title,
- involved samples (surveys) or medical tests,
- involved longitudinal data or random assignment,
- involved taking into account confounders, OR
- used statistics as evidence for causation.

Excluded articles on sports, weather and finance.

Primary source: Yahoo Health: 2005 – present.

5 August 2009 ASA 5

Computer Match: From Syntax to Semantics

Pro

- Decreases human error, readily verifiable
- Decreases time to analyze
- Allows for large number of search terms (448)

Con

- Can not identify part-of-speech
- Can not identify role of word in title
- Generates false positives (gives upper limits)

5 August 2009 ASA 6

What Does "Between" Mean?

Words that mean "association" to statisticians, but may imply "causation" to many laymen.

Examples:

- Autism often found **after** vaccination.
- TV **ups** kids' risk of attention problems.
- Fat toddlers (**are**) at risk for iron deficiency.
- Scientists find gene **link** to Alzheimer's.

5 August 2009 ASA 7

**Search Terms (448):
General Categories**

General categories:

- **Context keywords (10):** study, report, survey, poll
- **Believe keywords(6):** say, find, hope, believe, think,
- **Ratios (17):** percent, rate, chance, likely
- **Modals (8):** can, may, should, will
- **Auxiliaries (5):** have, had, has, help, helps
- **Comparatives (14):** more/less, bigger/smaller, higher/lower, major/minor

5 August 2009 ASA 8

**Search Terms (448):
A-B-C Categories:**

Categories are based on presence of keywords:

A: Association (18): associate, relate, correlate, predict

C: Causation (13): cause, effect, result, prevent

B: Between:

- **Temporal (15):** before/after, precede/follow
- **Logical (4):** because, due to, responsible
- **Action Verbs (267):** ‘affects’ to ‘ups’
- **Action-Nouns (55):** ‘accelerator’ to ‘suppressor’

5 August 2009 ASA 9

**Content Analysis of Titles:
General Results**

Percentage of titles where a keyword indicates:

- Association (associate, relate, correlate): 5%
- Causation (cause, effect, result, prevent): 6%
- **Between (factor, follows, links, ups): 42%**
- More/less comparatives (bigger/smaller): 9%

Percentage of titles that contain:

- Context indicators (studies, surveys, polls): 21%
- Modals (can, may): 19%
- Ratios (percent, rate, chance): 16%

5 August 2009 ASA 10

**Content Analysis of Titles:
“Between” Titles**

Percentage of “Between” titles that involve:

- Longitudinal (before/after): 8%
- Logical (because, due to): <1%
- Action nouns (factor, booster, reducer): 1%
- **Action verbs (links, ups, cuts, boosts): 91%**

5 August 2009 ASA 11

**“Between” Action Verbs
% of Titles with Action Verbs**

12% links. 6% ups. 5% cut, boost, raise. 4% get.
3%: increase, do, affect, change, lower, reduce.
2%: fight, make, drop, improve, protect, ease, grow.
1%: double, stop, trigger, delay, end, hit, keep, kill,
slow, extend, curb, speed, beat, hurt, fuel,
influence.
< 1%: spur, impact, produce, decrease, prolong,
bend, contribute, thwart.

5 August 2009 ASA 12

Conclusion: Action Needed

- The real problem here occurs when an association is mistaken for a cause-effect relationship.
- Informal polls find most statistics instructors already say “Association is not causation”; it seems just saying this is not enough.
- An activity may help. We recommend showing students real article titles and asking them to classify the type of relationship.