7/10/2008

Binary Confounders as Mathematical Objects

MILO SCHIELD
Augsburg College Dept of Business Administration
Director, W. M. Keck Statistical Literacy Project

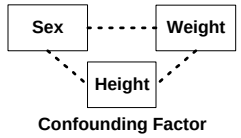
THOMAS V.V. BURNHAM
Cognitive Consulting

European Conference of Methodology
Oviedo, Spain
10 July 2008
Paper 2006SchieldBurnhamMAA.pdf
Slides 2008SchieldBurnhamMethodology6up.pdf
www.StatLit.org Schield@Augsburg.edu

7/10/2008

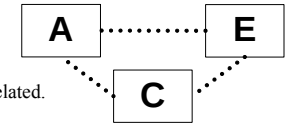
Confounding: Can Cause Variation

Observational studies: associations often *confounded*



Confounding Factor

E is Effect of interest
A is Associated predictor
C is confounder



Confounding is not chance-related.

As sample size increases,
* margin of error (influence of chance) decreases,
* influence of confounder is unchanged.

7/10/2008

Confounder Resistance: Rules of Thumb

Sir Richard Doll: No single study is persuasive unless the lower limit of its 95% confidence level falls above a **threefold** increased risk.

"As a general rule of thumb," says Angell of the New England Journal, "we are looking for a relative risk of **three or more"** before accepting a paper.

Robert Temple, FDA Director of Drug Evaluation, puts it bluntly: "My basic rule is if the relative risk isn't at least **three or four**, forget it."

7/10/2008

Problem of Confounding in Epidemiology

John Bailar, epidemiologist: "*There is no reliable way of identifying the dividing line*."

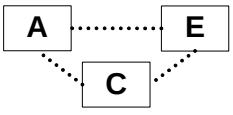
Epidemiologists need an abstract description of confounding that can generate confounder significance and confounder intervals for Relative Risk. RP(E:C) ≡ [P(E|C) / P(E|~C)]

This description must handle binary data, be meaningful, useful and easy to understand.

7/10/2008

Modeling Spuriosity: *A* has "no effect" on *E*

Model relation between 3 binary variables using an ordinary least-squares regression model:



E is Effect of interest
A is Associated predictor
C is confounder

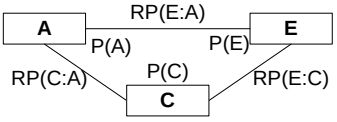
Non-Interactive model: E(A,C) = bo + b1*A + b2*C

Association between A and E is **spurious** if **b1 = 0.**

7/10/2008

Simplifying: Defining an S Confounder

A, C and E are binary. Top three ratios are observed. Confounder is defined by the bottom three ratios.



$$RP(E:A)=\frac{P(E\mid A)}{P(E\mid \sim A)}$$

$$RP(E:C)=\frac{P(E\mid C)}{P(E\mid \sim C)}$$

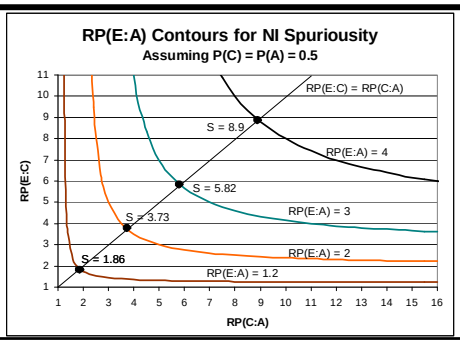
$$RP(A:C)=\frac{P(A\mid C)}{P(A\mid \sim C)}$$

Definition: An "**S confounder**" is a binary confounder with RP(C:A) = RP(E:C) = S. P(C) = P(A).

An association is "confounder resistant" to a size S confounder if it withstands nullification.

7/10/2008

Spuriousity Contours



7/10/2008

S Confounder Needed to Nullify Relative Risk

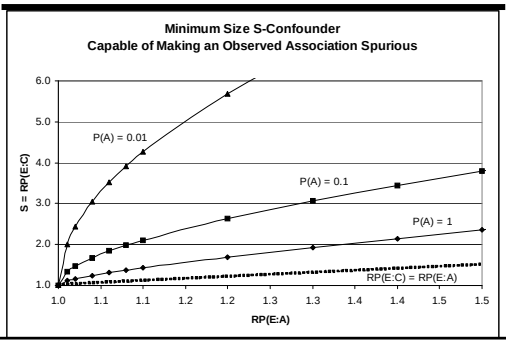
Given an observed association, RP(E:A), and a predictor prevalence, P(A), determine the size of the S confounder needed for nullification.

$$\frac{S-1}{[RP(E:A)-1]} = 1 + \sqrt{1 + \frac{1}{P(A)[RP(E:A)-1]}}$$

If P(A) = 0.5 and S = 5, then RP(E:A) = 2.6.
If P(A) = 0.4 and S = 6, then RP(E:A) = 3.1

7/10/2008

S Confounder Needed to Nullify a Relative Risk



7/10/2008

Confounder Intervals: Influence of S Confounder

Lower limit: determined by removing the influence of a confounder where RP(C:A) = RP(E:C) = S.

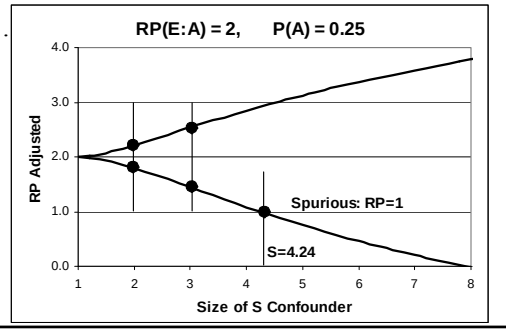
Upper limit: determined by removing the influence of a confounder where RP(C:A) = 1/RP(E:C) = 1/S.

Given the size of an S confounder, the confounder influence can be determined algebraically and illustrated using a standardization diagram.

For a S confounder of size 2, the confounder interval for RP(E:A) = 2 with P(A) = 0.5 is [1.67, 3.0].

7/10/2008

S Confounder Intervals



7/10/2008

Recommendations

Review/critique Schield & Burnham (2006) MAA paper: *Confounders as Mathematical Objects*. Copy at www.StatLit.org/pdf/2006SchieldBurnhamMAA.pdf

This 26 page paper is dense: 150 equations with new concepts and new ratio-comparison notation.

In-depth reviews will be acknowledged in the final paper.

If published, this paper will help make confounding a mathematical object. This could open up an entirely new area for statistics and statistical education.