V0G 2016 IASE3 1

Teaching Social Statistics: C: Inference & Significance

Milo Schield, Augsburg College

Member: International Statistical Institute

US Rep: International Statistical Literacy Project

VP. National Numeracy Network

IASE Roundtable in Berlin

July 20, 2016

www.StatLit.org/pdf/2016-Schield-IASE-3Slides.pdf

V0G 2016 IASE3 2

More on Confounding Study Design

Study design can eliminate (at least mitigate or 'ward off') different kinds of confounders.

Study design is at least as important in observational studies as it is in experiments.

Study design is seldom identified in news reports or press releases. Students have difficulty remembering the different types.

The following simplified presentation is used.

V0G 2016 IASE3 3

Types and Grades of Studies: Strength in Arguments

Experiment Treatment is assigned	Observational Exposure not assigned
A+ Repeatable	C Longitudinal
A- Randomize	D Snapshot

B Quasi (Queasy)-Experiment
Not repeatable; not randomized
Nature or humans intervene

V0G 2016 IASE3 4

Details on Quasi-Experiments

Quasi (Queasy)-Experiment
Treatment is assigned
Non-repeatable and non-random

Nature Intervenes	Humans intervene
Epidemics Plagues, outbreaks	Politics/Military Change laws & policies
Natural disasters Earthquakes, tornadoes	Business/Education Change pricing/teaching

V0G 2016 IASE3 5

Exploratory vs. Confirmatory: Confirmatory

Classic Statistical Research Cycle

1. Research Question
2. Design Experiment
3. Gather Data
4. Analyze Data
5. Statistical Tests
6. Evaluate Results

V0G 2016 IASE3 6

Exploratory vs. Confirmatory: 'Journalistic' Exploratory

Select research question:→→→→→→→→
Find/select data: →→→→→→→→→→→
Define groups & measures: →→→→→→→
Decide what to control for: →→→→→→→
Generate associations: →→→→→→→→→
Present results: →→→→→→→→→→→→

Choose everything in parallel (simultaneously) for coherence and for "journalistic significance".

V0G 2016 IASE3 7

Status Update

More time for observational data, assembly, confounding and study design must mean that there is less time for statistical inference.

But some new topics involving chance and statistical inference should be added.

V0G 2016 IASE3 8

More Focus on Randomness

1. More precision on statistical significance
2. Frequentists need Bayes for decision rules.
3. Coincidence ↑ as data ↑
4. Show how controlling for a confounder can influence statistical significance.

Decision-makers seldom need any thing more than "statistical significance."

V0G 2016 IASE3 9

1) Statistically Significant

What does 'statistically significant' mean?

The outcome (or anything more extreme) is unlikely

1. *if* due to chance
2. *to be* due to chance
3. due to chance

#1 is OK. For a frequentist, #2 is wrong.
#3 is ambiguous. The differences are subtle!

V0G 2016 IASE3 10

2) Statistical Tests and Frequentist Decision Making

Teaching students to reject the null (accept the alternate) for a statistically-significant outcome is NOT justified by Frequentist theory.

As Frequentists, statistical educators should NEVER allow statistical significance to be used for decision-making. Decision-making should always be left to subject-matter experts.

V0G 2016 IASE3 11

Statistical Tests and Bayesian Decision Making

But focusing on p-values and avoiding decision-making violates the 2016 GAISE guidelines: "Statistics is a decision-making process."

Statistical educators should embrace Bayes.

If the alternate (Ha) is more likely to be true than the null (Ho), then a test statistic with a p-value of P gives at least a (1-P) confidence that Ho is False and Ha is true. Schield (1996)

V0G 2016 IASE3 12

Statistical Tests and Bayesian Decision Making

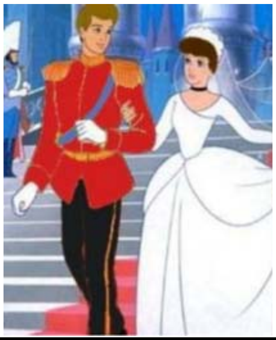
Some may think: "Schield has lost it. We've taught this decision-rule for decades."

My reply: Where is it proven that this decision rule is justified regardless of the alternative?

Remember Utts p-value for ESP: 10^-18.

Should this extremely small p-value justify rejecting the Null (No ESP) and accepting the alternative (H1). Schield (1996)

3) Coincidence in Big Data



Showing Coincidence Law of Very Large Numbers

The unlikely is almost certain given enough tries.

Rare outcome is more likely than not given N tries where P(success) = 1/N.

Selecting unlikely outcomes after-the-fact.

- Unlikely sequences in flipping coins
- Unlikely patterns in grains of rice
- Seeing why the Birthday problem works

Why We Should Teach "Practical" Statistics

1. Students value assembly and confounding.
2. Increased focus on critical thinking
3. They need it; they see value in it.

Can we teach statistical significance in less time?

Can we show the criteria for statistical significance?

Can we demonstrate the influence of confounding, assembly and bias on statistical significance?

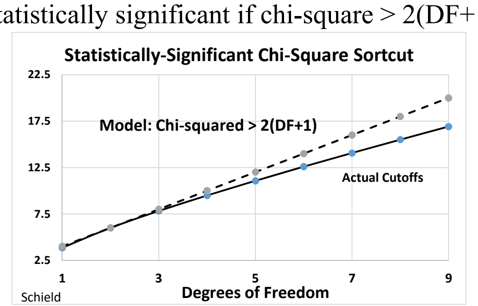
YES!

Teach Statistical Inference Differently

1. Skip derivation of the sampling distribution. See Utts (2014) and de Veaux (2011).
2. Use 1-2-3 rule for Confidence levels.
3. Focus on statistical significance (big idea).
4. Use statistical-significance short-cuts
5. Non-overlap of 95% confidence intervals is sufficient for statistical significance

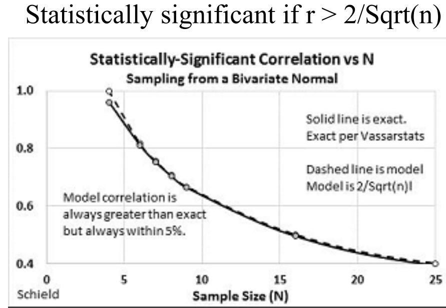
Statistically-Significant Short-Cut: Chi-Square

Statistically significant if chi-square > 2(DF+1)



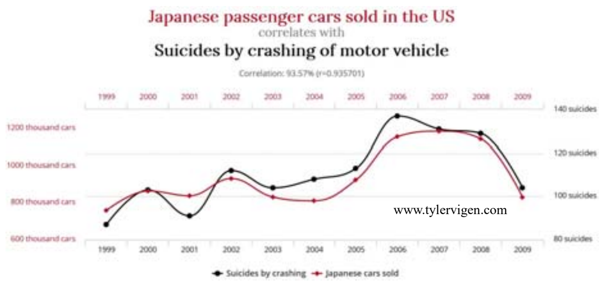
Statistically-Significant Short-Cut: Bivariate Correlation

Statistically significant if r > 2/Sqrt(n)



IASE 2016 - 3 19

3. Correlation = 93.6%. Isn't this statistically significant?



Normal Statistical Significance Cutoffs
Don't Apply **to Time-Based Correlations**

V0G 2016 IASE3 20

Can we show Confounder Influence on Statistical Significance?

Variation – random variation – is at the core of the introductory statistics course.

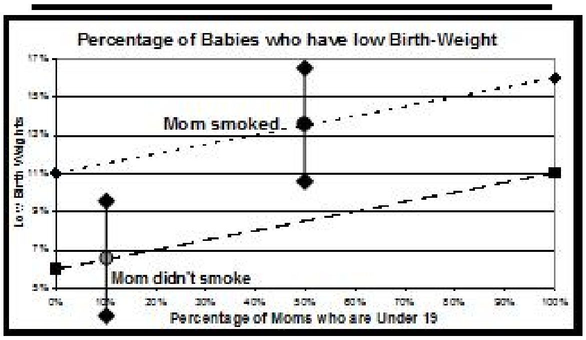
We know that controlling for a confounder can negate or reverse an observed association.

Can we show this with minimal assumptions?

Yes, using Wainer's standardization diagram

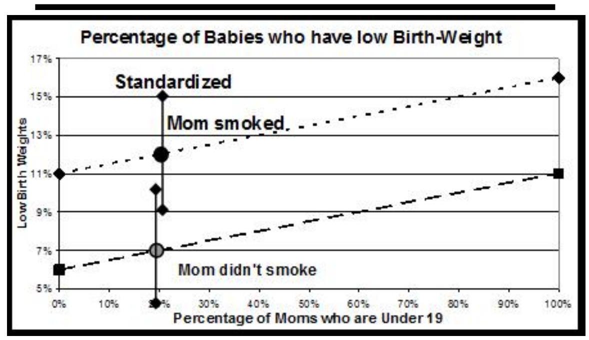
V0G 2016 IASE3 21

Confounder Influence: Non-Overlap = Statistical Significance



V0G 2016 IASE3 22

Confounder Influence on Statistical Significance



V0G 2016 IASE3 23

Showing Influence on Statistical Significance

Variation – random variation – should be at the core of the introductory statistics course.

Wainer's Standardization technique allow us to show students how controlling for a confounder can influence statistical significance.

Showing confounder influence on statistical significance should be included in every introductory statistic course.

Anything less is **professional negligence**.

V0G 2016 IASE3 24

Statistics for Decision Makers: Recommendations

To uphold statistics as numbers with a context, a new intro statistics course should be offered.

This intro course needs more focus on big ideas:

- **Context** (control), **assembly** (definitions) and **bias** are big ideas for non-statisticians.
- **Randomness** and **statistical significance** are big ideas for statisticians.
- **Seeing how confounding, assembly and bias can influence statistical significance should be central for a "statistics-in-context" course**.

Thesis

Adding context to introductory statistics will

- uphold context as the essence of statistics (e.g., statistics are numbers in context),
- give stronger support for statistics as a liberal art
- separate applied statistics from mathematical statistics,
- improve student retention of key ideas, and
- improve student attitudes on the value of statistics.

Student Evaluations

Claim: Statistical literacy should be required for all students for graduation.

Of the 57 students in my classes this past year,
* 68% chose "Agree" or "Strongly agree " while
* 21% chose "Strongly agree."

This is a strong confirmation that a class focused on observational studies and using statistical associations as evidence for causal connections is of value to future decision makers.

Conclusion

Statistical educators should support three intro classes:

Stat 100: Statistical Literacy (Statistics in the Media)

Stat 101: Statistics for Researchers (Statistical Inference)
Focus on derivations and statistical tests.
Emphasis on random samples & inference

Stat 102: Practical Statistics for Decision Makers.
Focus on all sources of influence on statistics
Emphasis on Assembly and Confounding.
Show confounder influence on significance.

Stat 102 is needed to meet the 2016 GAISE guidelines.

V0G 2016 IASE3 1

Teaching Social Statistics: C: Inference & Significance

Milo Schield, Augsburg College

Member: International Statistical Institute

US Rep: International Statistical Literacy Project

VP. National Numeracy Network

IASE Roundtable in Berlin

July 20, 2016

www.StatLit.org/pdf/2016-Schield-IASE-3Slides.pdf

More on Confounding Study Design

Study design can eliminate (at least mitigate or 'ward off') different kinds of confounders.

Study design is at least as important in observational studies as it is in experiments.

Study design is seldom identified in news reports or press releases. Students have difficulty remembering the different types.

The following simplified presentation is used.

V0G 2016 IASE3 3

Types and Grades of Studies: Strength in Arguments

Experiment
Treatment is assigned

- A+ Repeatable
- A- Randomize

Observational
Exposure not assigned

- C Longitudinal
- D Snapshot

B Quasi (Queasy)-Experiment
Not repeatable; not randomized
Nature or humans intervene

V0G 2016 IASE3 4

Details on Quasi-Experiments

Quasi (Queasy)-Experiment

Treatment is assigned

Non-repeatable and non-random

Nature Intervenes

Epidemics
Plagues, outbreaks

Natural disasters
Earthquakes, tornadoes

Humans intervene

Politics/Military
Change laws & policies

Business/Education
Change pricing/teaching

Exploratory vs. Confirmatory: Confirmatory

Classic Statistical Research Cycle

1. Research Question
2. Design Experiment
3. Gather Data
4. Analyze Data
5. Statistical Tests
6. Evaluate Results

V0G 2016 IASE3 6

Exploratory vs. Confirmatory: 'Journalistic' Exploratory

Select research question:→→→→→→→→

Find/select data: →→→→→→→→→→→

Define groups & measures: →→→→→→→

Decide what to control for: →→→→→→→

Generate associations: →→→→→→→→→

Present results: →→→→→→→→→→→→

Choose everything in parallel (simultaneously) for coherence and for "journalistic significance".

Status Update

More time for observational data, assembly, confounding and study design must mean that there is less time for statistical inference.

But some new topics involving chance and statistical inference should be added.

More Focus on Randomness

1. More precision on statistical significance
2. Frequentists need Bayes for decision rules.
3. Coincidence ↑ as data ↑
4. Show how controlling for a confounder can influence statistical significance.

Decision-makers seldom need any thing more than "statistical significance."

1) Statistically Significant

What does ‘statistically significant’ mean?

The outcome (or anything more extreme) is unlikely

1. *if* due to chance
2. *to be* due to chance
3. due to chance

#1 is OK. For a frequentist, #2 is wrong.
#3 is ambiguous. The differences are subtle!

V0G 2016 IASE3 10

2) Statistical Tests and Frequentist Decision Making

Teaching students to reject the null (accept the alternate) for a statistically-significant outcome is NOT justified by Frequentist theory.

As Frequentists, statistical educators should NEVER allow statistical significance to be used for decision-making. Decision-making should always be left to subject-matter experts.

Statistical Tests and Bayesian Decision Making

But focusing on p-values and avoiding decision-making violates the 2016 GAISE guidelines: “Statistics is a decision-making process.”

Statistical educators should embrace Bayes.

If the alternate (Ha) is more likely to be true than the null (Ho), then a test statistic with a p-value of P gives at least a (1-P) confidence that Ho is False and Ha is true. Schield (1996)

V0G 2016 IASE3 12

Statistical Tests and Bayesian Decision Making

Some may think: “Schield has lost it. We’ve taught this decision-rule for decades.”

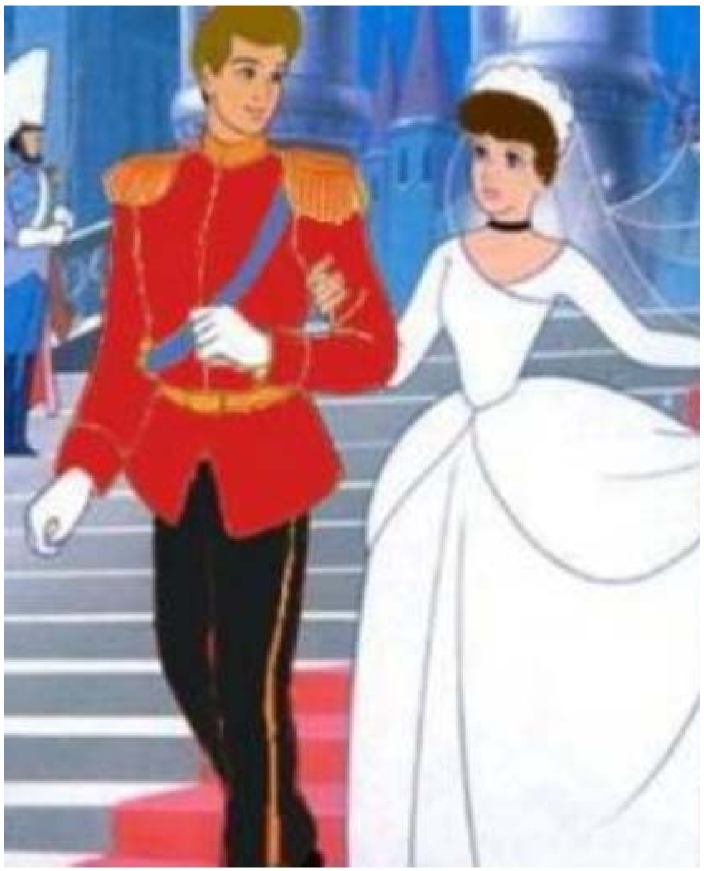
My reply: Where is it proven that this decision rule is justified regardless of the alternative?

Remember Utts p-value for ESP: 10^-18.

Should this extremely small p-value justify rejecting the Null (No ESP) and accepting the alternative (H1). Schield (1996)

V0G 2016 IASE3 13

3) Coincidence in Big Data



Showing Coincidence
Law of Very Large Numbers

The unlikely is almost certain given enough tries.

Rare outcome is more likely than not given N tries where P(success) = 1/N.

Selecting unlikely outcomes after-the-fact.

- Unlikely sequences in flipping coins
- Unlikely patterns in grains of rice
- Seeing why the Birthday problem works

Why We Should Teach "Practical" Statistics

1. Students value assembly and confounding.
2. Increased focus on critical thinking
3. They need it; they see value in it.

Can we teach statistical significance in less time?

Can we show the criteria for statistical significance?

Can we demonstrate the influence of confounding, assembly and bias on statistical significance?

YES!

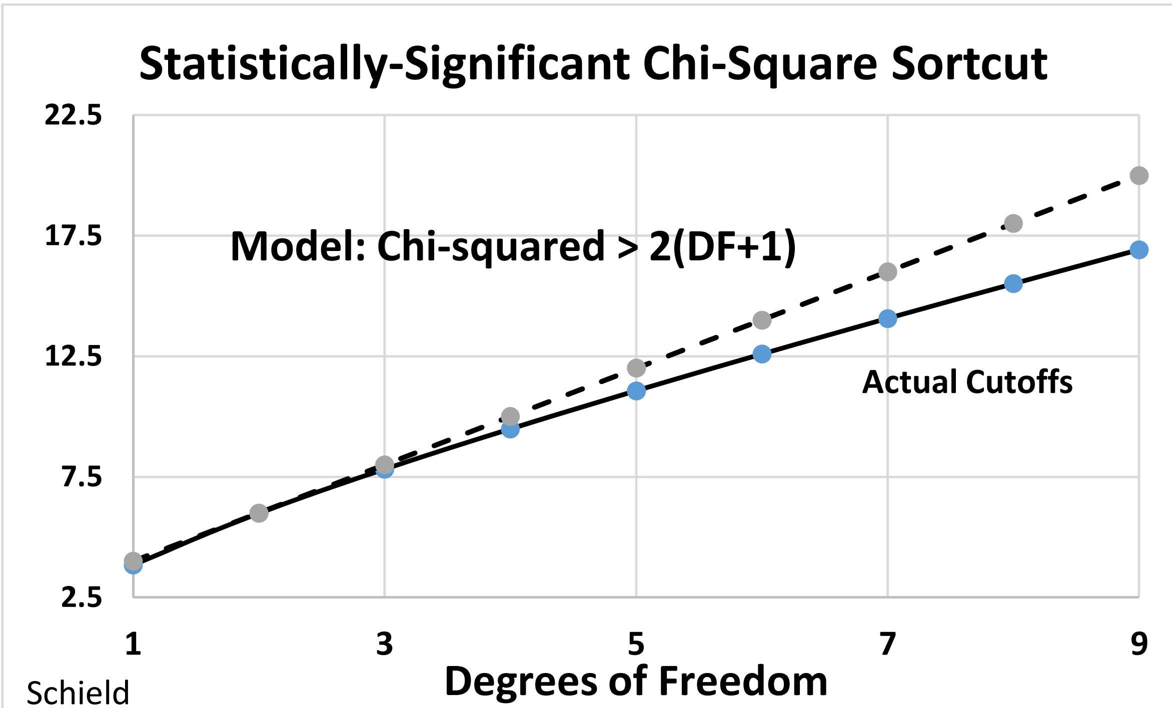
Teach Statistical Inference Differently

1. Skip derivation of the sampling distribution. See Utts (2014) and de Veaux (2011).
2. Use 1-2-3 rule for Confidence levels.
3. Focus on statistical significance (big idea).
4. Use statistical-significance short-cuts
5. Non-overlap of 95% confidence intervals is sufficient for statistical significance

V0G 2016 IASE3 17

Statistically-Significant Short-Cut: Chi-Square

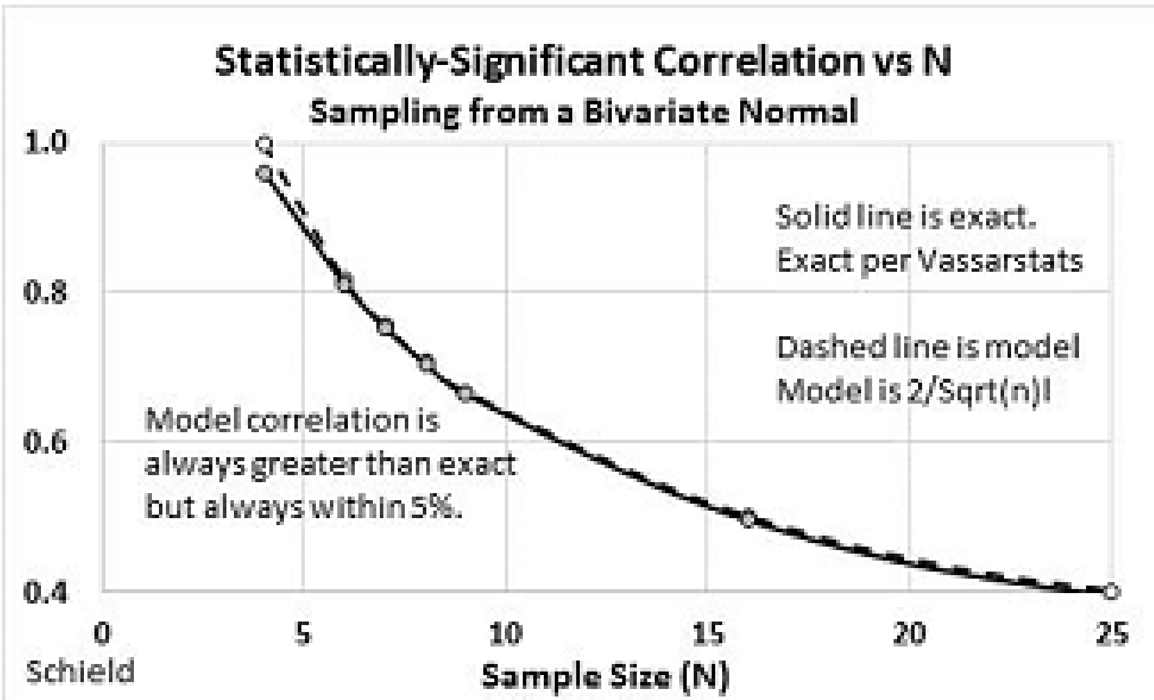
Statistically significant if chi-square > 2(DF+1)



V0G 2016 IASE3 18

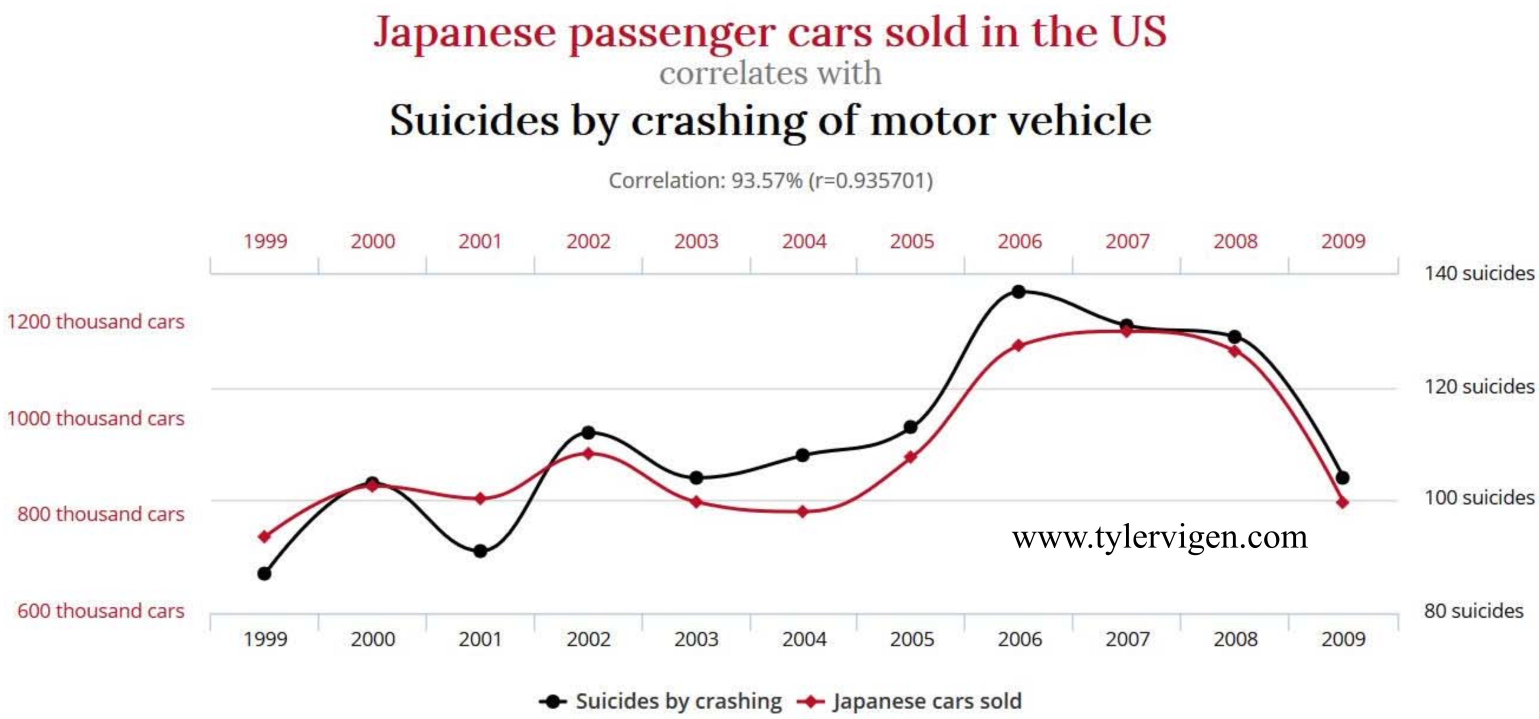
Statistically-Significant Short-Cut: Bivariate Correlation

Statistically significant if r > 2/Sqrt(n)



IASE 2016 - 3 19

3. Correlation = 93.6%. Isn't this statistically significant?



Normal Statistical Significance Cutoffs *Don't Apply* to Time-Based Correlations

Can we show Confounder Influence on Statistical Significance?

Variation – random variation – is at the core of the introductory statistics course.

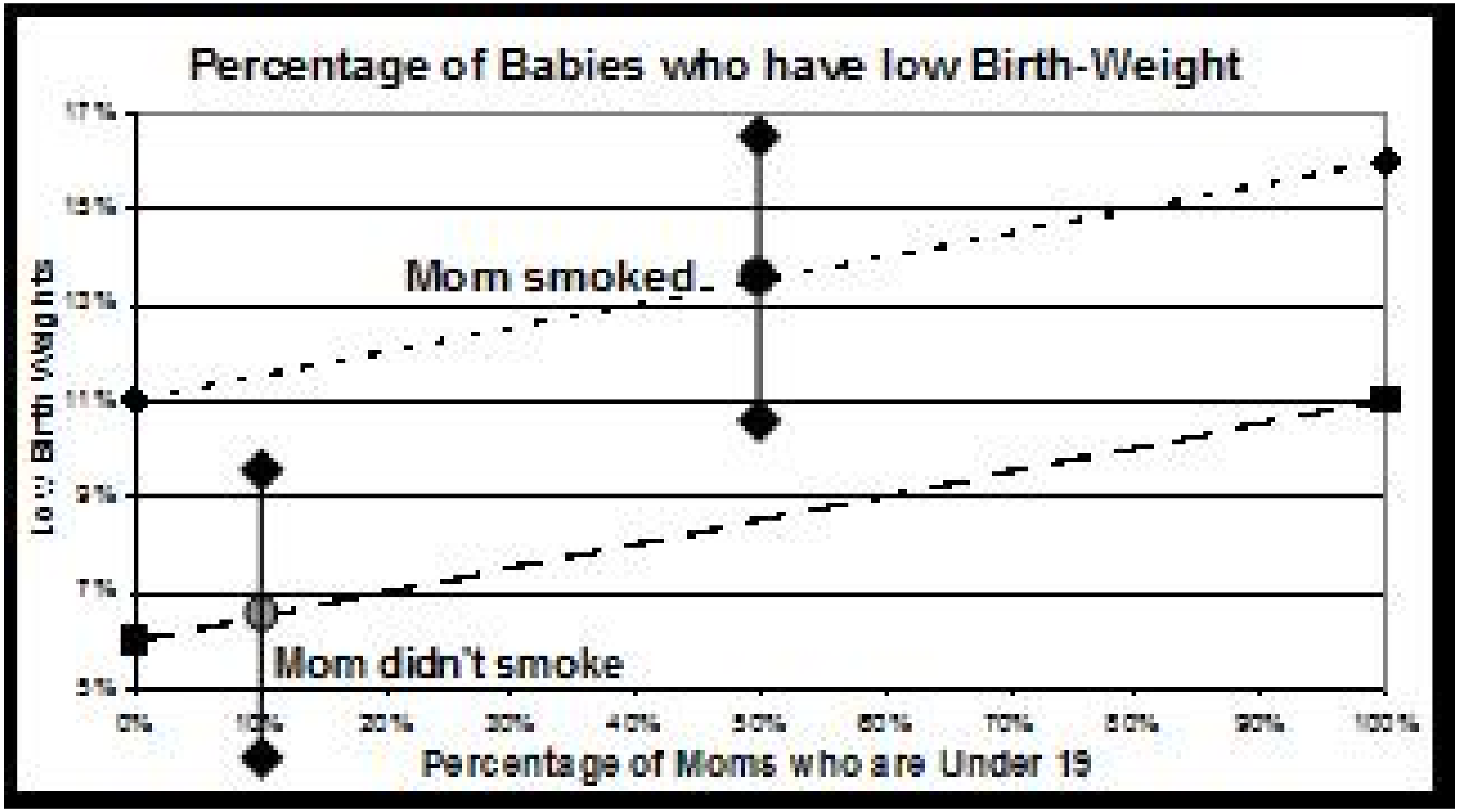
We know that controlling for a confounder can negate or reverse an observed association.

Can we show this with minimal assumptions?

Yes, using Wainer's standardization diagram

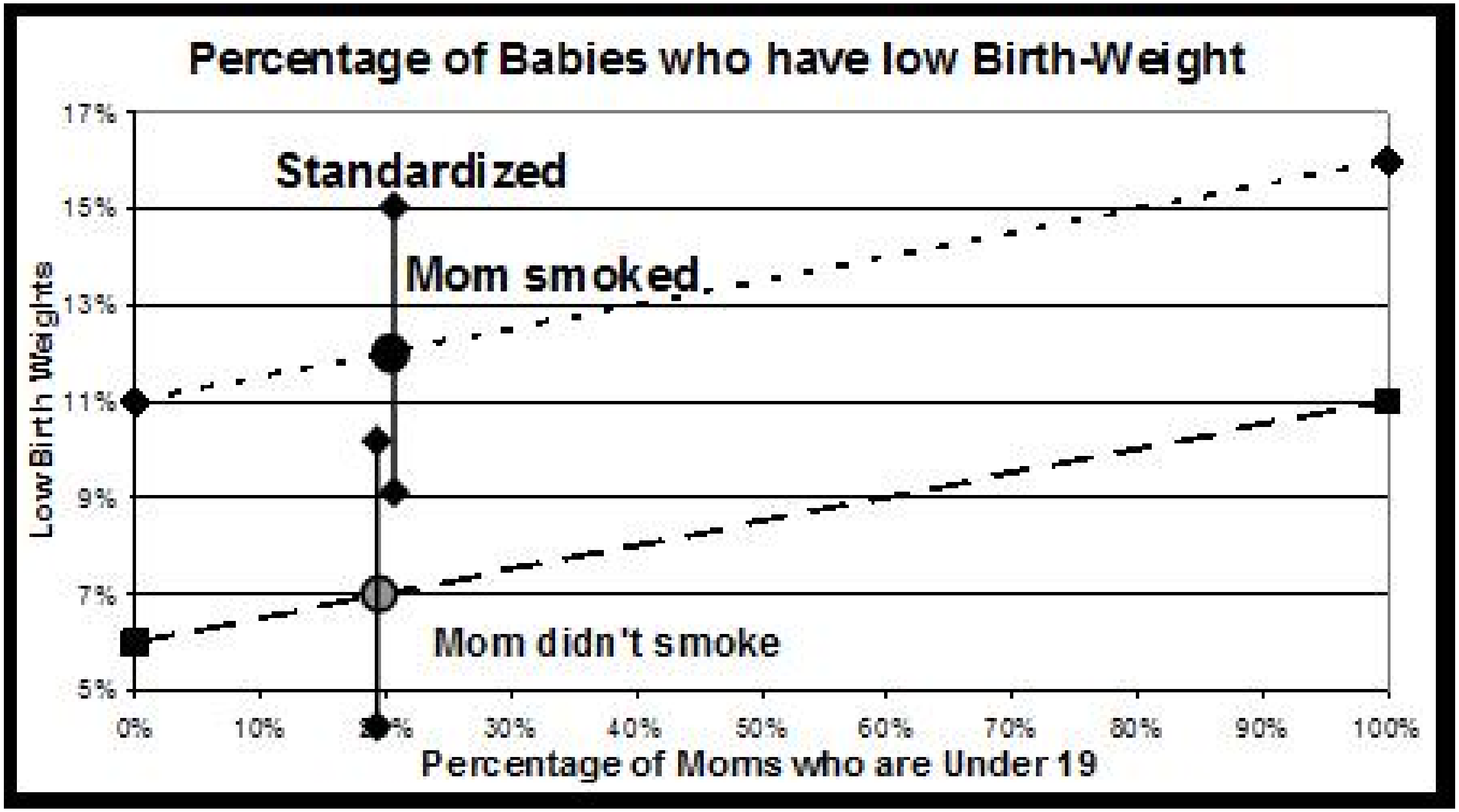
V0G 2016 IASE3 21

Confounder Influence: Non-Overlap = Statistical Significance



V0G 2016 IASE3 22

Confounder Influence on Statistical Significance



Showing Influence on Statistical Significance

Variation – random variation – should be at the core of the introductory statistics course.

Wainer's Standardization technique allow us to show students how controlling for a confounder can influence statistical significance.

Showing confounder influence on statistical significance should be included in every introductory statistic course.

Anything less is **professional negligence**.

V0G 2016 IASE3 24

Statistics for Decision Makers: Recommendations

To uphold statistics as numbers with a context, a new intro statistics course should be offered.

This intro course needs more focus on big ideas:

- **Context** (control), **assembly** (definitions) and **bias** are big ideas for non-statisticians.
- **Randomness** and **statistical significance** are big ideas for statisticians.
- **Seeing how confounding, assembly and bias can influence statistical significance should be central for a "statistics-in-context" course**.

Thesis

Adding context to introductory statistics will

- uphold context as the essence of statistics (e.g., statistics are numbers in context),
- give stronger support for statistics as a liberal art
- separate applied statistics from mathematical statistics,
- improve student retention of key ideas, and
- improve student attitudes on the value of statistics.

Student Evaluations

Claim: Statistical literacy should be required for all students for graduation.

Of the 57 students in my classes this past year,
* 68% chose “Agree” or "Strongly agree " while
* 21% chose "Strongly agree."

This is a strong confirmation that a class focused on observational studies and using statistical associations as evidence for causal connections is of value to future decision makers.

Conclusion

Statistical educators should support three intro classes:

Stat 100: Statistical Literacy (Statistics in the Media)

Stat 101: Statistics for Researchers (Statistical Inference)
Focus on derivations and statistical tests.
Emphasis on random samples & inference

Stat 102: Practical Statistics for Decision Makers.
Focus on all sources of influence on statistics
Emphasis on Assembly and Confounding.
Show confounder influence on significance.

Stat 102 is needed to meet the 2016 GAISE guidelines.