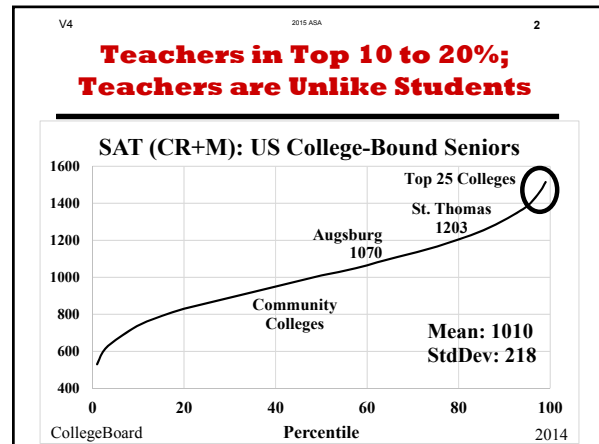


V4 2015 ASA 1

Statistical Inference for Managers

by
Milo Schield, Augsburg College
Member: International Statistical Institute
US Rep: International Statistical Literacy Project
Director, W. M. Keck Statistical Literacy Project
August 11, 2015

Paper: www.StatLit.org/pdf/2015-Schild-ASA.pdf
Slides: www.StatLit.org/pdf/2015-Schild-ASA-6up.pdf



V4 2015 ASA 3

Teachers Mainly Math/Stat; Teachers are Unlike Students

Major: BA/BS	StatEd	
	Teachers	Students
Bus/Econ	7%	41%
Soc/SocWk	2%	21%
Health	7%	15%
Psychology	0%	11%
Biology	5%	10%
Math/Stats	75%	1%
Other	5%	1%
Total	100%	100%

Stat Educators @JSM are a biased sample

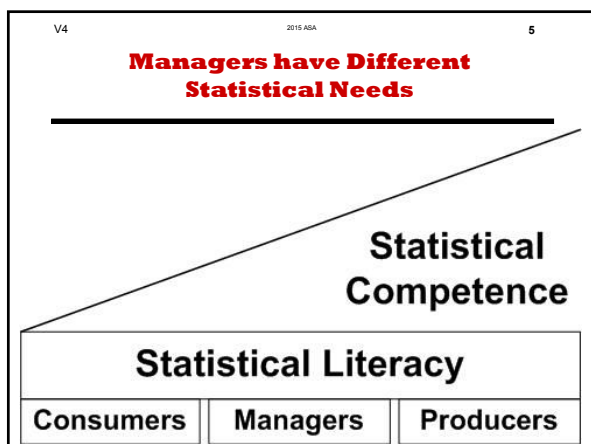
V4 2015 ASA 4

Biz Stat-Teachers at Top End Biz Teachers Unlike Biz Students

Four-Year Colleges	Business Majors	
Student Ability	Quantitative	Qualitative
Calc-based Stats	Highest	Higher
Algebra-based Stats	Lower	Lowest

Quantitative majors (left) focus on problem solving
Qualitative majors (right) focus on critical thinking

Biggest group of Stat-Ed teachers teach upper-left.
Biggest group of business majors is in lower-right.



V4 2015 ASA 6

Managers have unique needs

More breadth than consumers. More on big data, (coincidence & confounding) and on time series.
Less on the “logic of inference” than producers.

Math Colleagues: “Is this STAT LITE???”

Bold reply: “No! It’s not Stat-Lite.”
Yes; Less on formula derivation and test details.
More on *understanding* statistical significance and sampling distributions.

V4 2015 ASA 7

R-sq = 0.49; N = 9.
Is this statistically significant?

Weight vs. Height (N=9)
 R-square = 0.49

Yes! $R > 2/\sqrt{n}$ is sufficient. Schield (2014b)

V4 2015 ASA 8

Correlation = 93.6%.
Is this statistically significant?

Japanese passenger cars sold in the US
 correlates with
 Suicides by crashing of motor vehicle

Correlation: 93.57% ($r=0.935701$)

No! Normal statistical-significance minimums don't apply to time-based correlations.

V4 2015 ASA 9

Chi-sq = 12.5; Six bins.
Is this statistically significant?

	Ideal	Ideal	Actual	Chi-Sq
#1	0.167	10	8	0.4
#2	0.167	10	14	1.6
#3	0.167	10	14	1.6
#4	0.167	10	6	1.6
#5	0.167	10	7	0.9
#6	0.167	10	18	6.4
Total		60	60	12.5

YES! $\chi^2 > 2 * \#bins$ is sufficient. Schield (2014c)

V4 2015 ASA 10

Is Statistical Significance Necessary for Causation?

ZICAM: homeopathic remedy **clinically proven** to reduce symptoms of common cold

Of the millions of users,
 ~ten lost their sense of smell

Zicam defense; Ten is not statistically significant.

US Supreme Court: Lack of statistical significance is not an acceptable defense. See Schield (2011).

V4 2015 ASA 11

Influence of Bias & Confounding on Statistical Significance

Bias:
 Subject bias, measurement bias and sampling bias
 See Schield (2013).

Confounder: A factor related to the predictor and to the outcome in an association that
 (1) has a causal influence on the outcome and
 (2) is not causally influenced by the predictor.
 See Schield (2006 and 2014a)

V4 2015 ASA 12

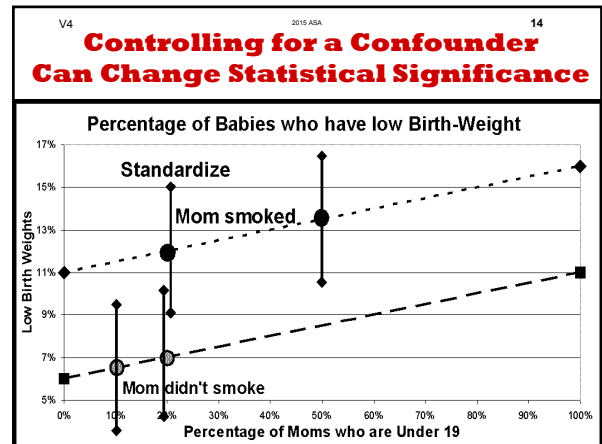
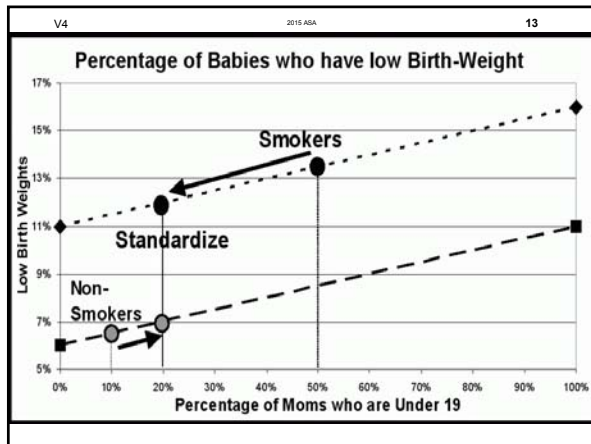
Influence of Bias on Significance

Response bias: Men likely to overstate income

\$5,000 is the 95% margin of error					
Income	Men	Women	Diff	Overlap	Stat. Sig
Stated	\$62,000	\$51,000	\$11,000	No	Yes
Actual	\$53,000	\$51,000	\$2,000	Yes	No

Sample bias: Rich less likely to do surveys

\$3,000 is the 95% margin of error					
Income	Men	Women	Diff	Overlap	Stat. Sig
Responders	\$53,000	\$51,000	\$2,000	Yes	No
Population	\$62,000	\$55,000	\$7,000	No	Yes



Understanding the "Logic of Statistical Inference"

McKenzie (2004) asked statistical educators to pick the top-three core concepts in intro statistics:

- 75% Variation
- 31% Association vs. causation
- 25% Hypothesis tests and **24% Sampling distribution**
- 22% Confidence intervals
- 14% Randomness and statistical significance

#: Percentage of votes by Statistical Educators
Sample size: 56; 95% ME = 12 percentage points

Understanding the "Logic of Statistical Inference"

Teaching randomness and statistical significance is necessary but not sufficient.

Students need to understand and appreciate the sampling distribution.

But deriving the sampling distribution takes time. Randomization takes time and a computer.

What to do with minimal time & no computer?
See the final paper for more on this topic.

Conclusion

Managers need a statistics curriculum that is better aligned with their work.

- Less on the derivation of sampling error; More on understanding sampling distributions
- Less on p-value; More on statistical significance

References

Schield, M. (2015). Statistically-Significant Shortcuts. Statchat, Macalester. www.statlit.org/pdf/2015-Schield-StatChat-Slides.pdf

Schield, M. (2014c). Chi-Squared Cutoffs for Statistical Significance. NNN www.statlit.org/pdf/2014-Schield-SS-Shortcut-Chi-Square.pdf

Schield, M. (2014b). Statistically-Significant Correlations. NNN Carleton. www.statlit.org/pdf/2014-Schield-NNN4-Slides.pdf

Schield, M. (2014a). Two Big Ideas for Teaching Big Data: ECOTS. www.statlit.org/pdf/2014-Schield-ECOTS.pdf

Schield, M. (2011). Zicam and the US Supreme Course. www.statlit.org/pdf/2011Schield-ASA-Zicam6up.pdf

Schield, M. (2006). Presenting Confounding & Standardization Graphically. *STATS Magazine*. www.StatLit.org/pdf/2006SchieldSTATS.pdf

McKenzie, John, Jr. (2004). Teaching the Core Concepts. ASA www.statlit.org/pdf/2004McKenzieASA.pdf



Statistical Inference for Managers

by

Milo Schield, Augsburg College

Member: International Statistical Institute

US Rep: International Statistical Literacy Project

Director, W. M. Keck Statistical Literacy Project

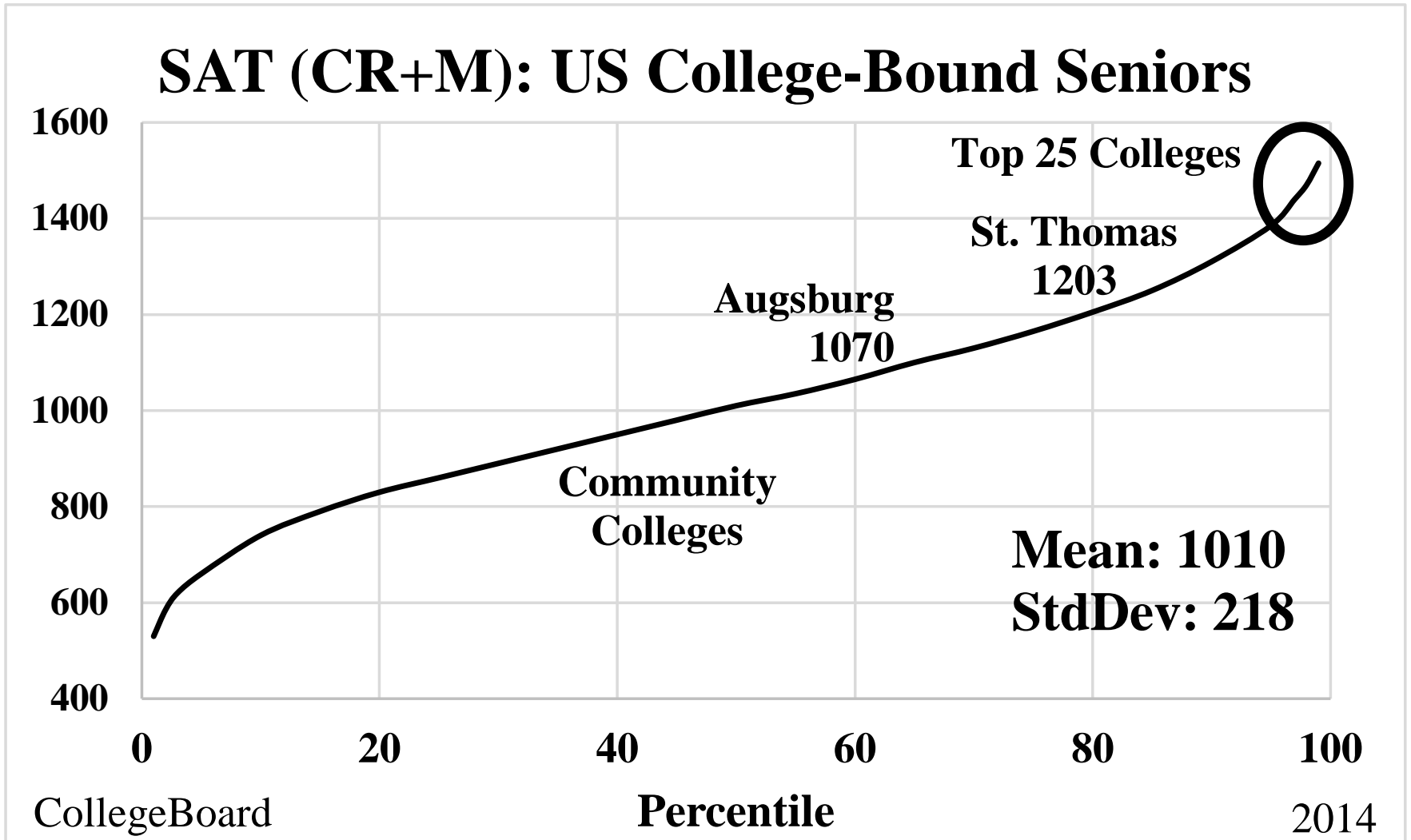
August 11, 2015

Paper: www.StatLit.org/pdf/2015-Schild-ASA.pdf

Slides: www.StatLit.org/pdf/2015-Schild-ASA-6up.pdf



Teachers in Top 10 to 20%; Teachers are Unlike Students





Teachers Mainly Math/Stat; Teachers are Unlike Students

	StatEd	Stat
Major: BA/BS	Teachers	Students
Bus/Econ	7%	41%
Soc/SocWk	2%	21%
Health	7%	15%
Psychology	0%	11%
Biology	5%	10%
Math/Stats	75%	1%
Other	5%	1%
Total	100%	100%

Stat Educators @JSM are a biased sample

Biz Stat-Teachers at Top End Biz Teachers Unlike Biz Students

Four-Year Colleges	Business Majors	
Student Ability	Quantitative	Qualitative
Calc-based Stats	Highest	Higher
Algebra-based Stats	Lower	Lowest

Quantitative majors (left) focus on problem solving
Qualitative majors (right) focus on critical thinking

Biggest group of Stat-Ed teachers teach upper-left.
Biggest group of business majors is in lower-right.



Managers have Different Statistical Needs

**Statistical
Competence**

Statistical Literacy

Consumers

Managers

Producers

Managers have unique needs

More breadth than consumers. More on big data, (coincidence & confounding) and on time series.

Less on the “logic of inference” than producers.

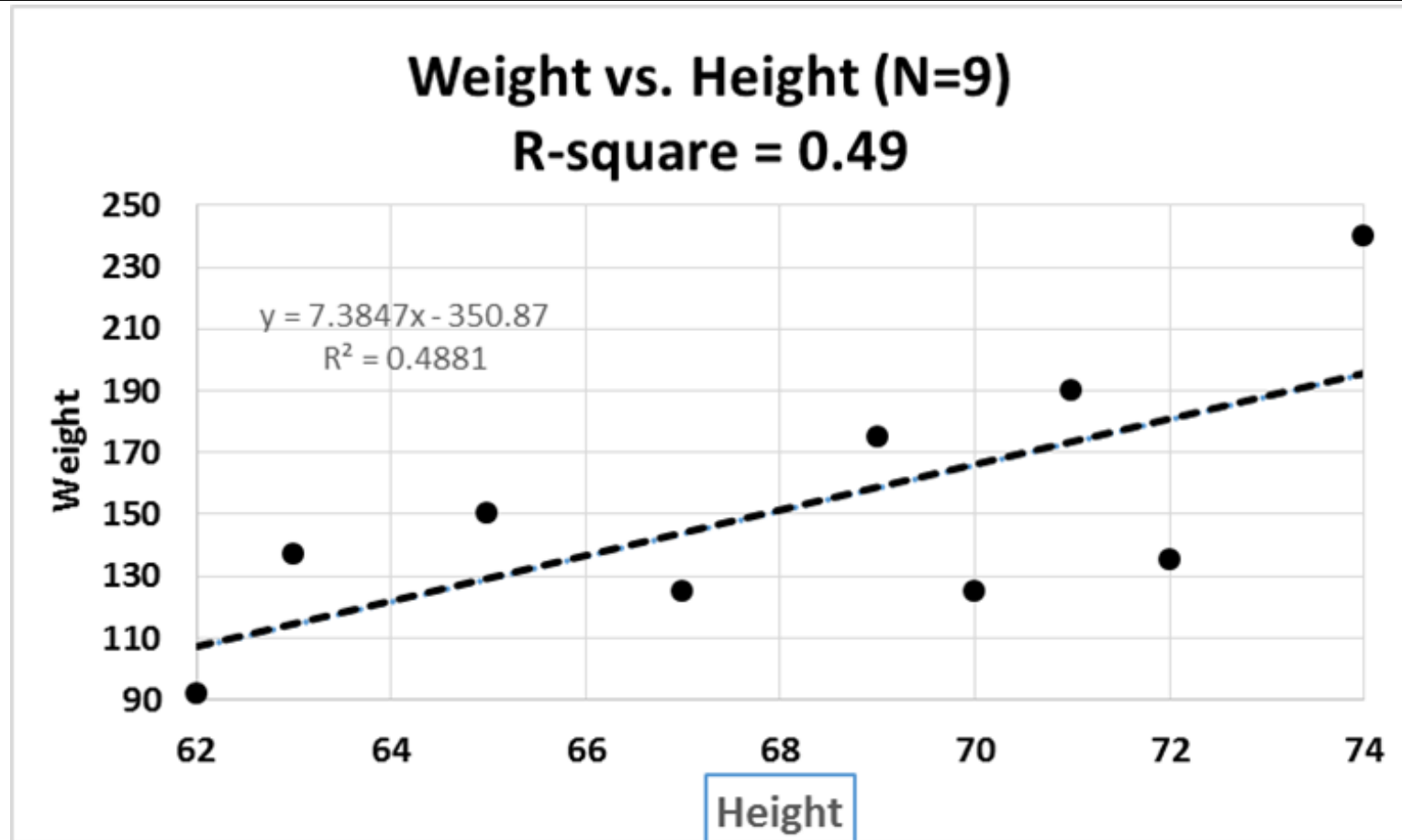
Math Colleagues: “Is this STAT LITE???”

Bold reply: “No! It’s not Stat-Lite.”

Yes; Less on formula derivation and test details.

More on *understanding* statistical significance and sampling distributions.

R-sq = 0.49; N = 9.
Is this statistically significant?

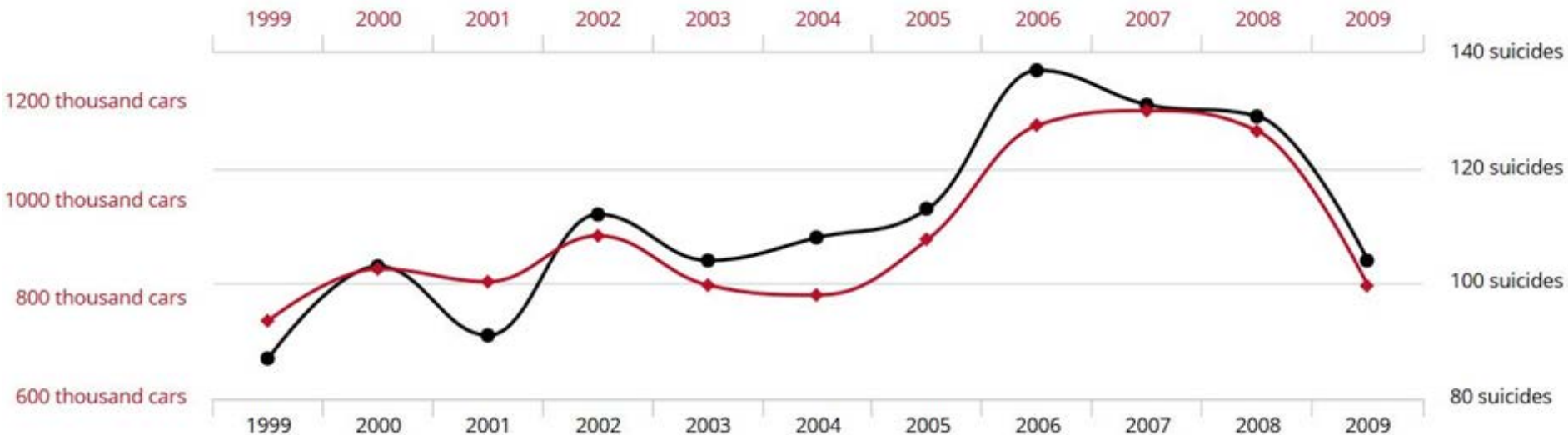


Yes! $R > 2/\text{Sqrt}(n)$ is sufficient. Schield (2014b)

Correlation = 93.6%. Is this statistically significant?

Japanese passenger cars sold in the US
correlates with
Suicides by crashing of motor vehicle

Correlation: 93.57% (r=0.935701)



**No! Normal statistical-significance minimums
don't apply to time-based correlations.**

**Chi-sq = 12.5; Six bins.
Is this statistically significant?**

	Ideal	Ideal	Actual	Chi-Sq
#1	0.167	10	8	0.4
#2	0.167	10	14	1.6
#3	0.167	10	14	1.6
#4	0.167	10	6	1.6
#5	0.167	10	7	0.9
#6	0.167	10	18	6.4
Total		60	60	12.5

YES! $\chi^2 > 2*\#bins$ is sufficient. Schield (2014c)

Is Statistical Significance Necessary for Causation?

ZICAM: homeopathic remedy **clinically proven** to reduce symptoms of common cold

Of the millions of users,
~ten lost their sense of smell

Zicam defense; Ten is not statistically significant.

US Supreme Court: Lack of statistical significance is not an acceptable defense. See Schield (2011).



Influence of Bias & Confounding on Statistical Significance

Bias:

Subject bias, measurement bias and sampling bias

See Schield (2013).

Confounder: A factor related to the predictor and to the outcome in an association that

(1) has a causal influence on the outcome and

(2) is not causally influenced by the predictor.

See Schield (2006 and 2014a)

Influence of Bias on Significance

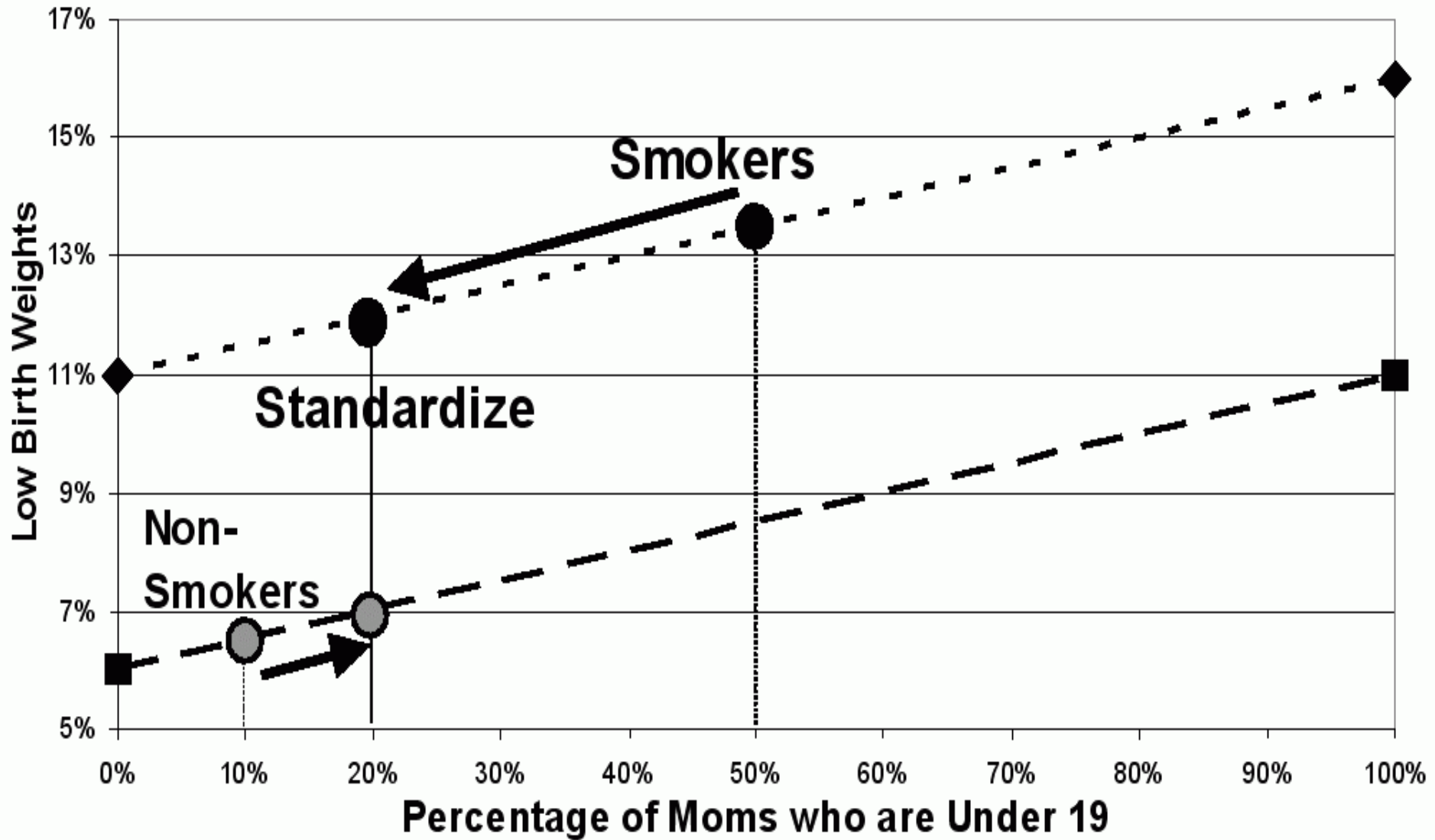
Response bias: Men likely to overstate income

\$5,000 is the 95% margin of error					
Income	Men	Women	Diff	Overlap	Stat. Sig
Stated	\$62,000	\$51,000	\$11,000	No	Yes
Actual	\$53,000	\$51,000	\$2,000	Yes	No

Sample bias: Rich less likely to do surveys

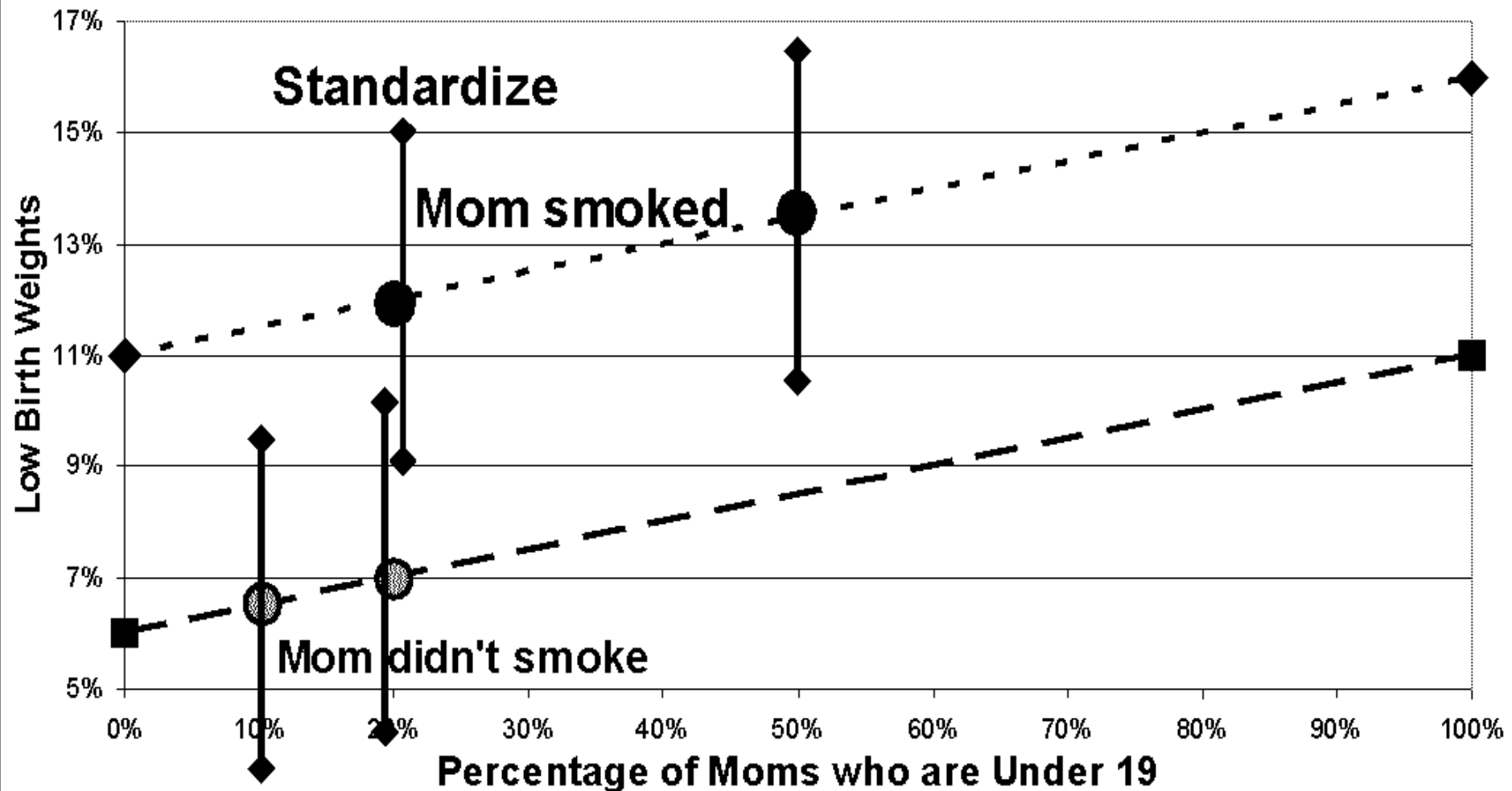
\$3,000 is the 95% margin of error					
Income	Men	Women	Diff	Overlap	Stat. Sig
Responders	\$53,000	\$51,000	\$2,000	Yes	No
Population	\$62,000	\$55,000	\$7,000	No	Yes

Percentage of Babies who have low Birth-Weight



Controlling for a Confounder Can Change Statistical Significance

Percentage of Babies who have low Birth-Weight



Understanding the “Logic of Statistical Inference”

McKenzie (2004) asked statistical educators to pick the top-three core concepts in intro statistics:

75% Variation

31% Association vs. causation

25% Hypothesis tests and

24% **Sampling distribution**

22% Confidence intervals

14% Randomness and statistical significance

#: Percentage of votes by Statistical Educators

Sample size: 56; 95% ME = 12 percentage points

Understanding the “Logic of Statistical Inference”

Teaching randomness and statistical significance is necessary but not sufficient.

Students need to understand and appreciate the sampling distribution.

But deriving the sampling distribution takes time. Randomization takes time and a computer.

What to do with minimal time & no computer?
See the final paper for more on this topic.

Conclusion

Managers need a statistics curriculum that is better aligned with their work.

- Less on the derivation of sampling error;
More on understanding sampling distributions
- Less on p-value;
More on statistical significance

References

- Schield, M. (2015). Statistically-Significant Shortcuts. Statchat, Macalester.
www.statlit.org/pdf/2015-Schield-StatChat-Slides.pdf
- Schield, M. (2014c). Chi-Squared Cutoffs for Statistical Significance.
NNN www.statlit.org/pdf/2014-Schield-SS-Shortcut-Chi-Square.pdf
- Schield, M. (2014b). Statistically-Significant Correlations. NNN Carleton.
www.statlit.org/pdf/2014-Schield-NNN4-Slides.pdf
- Schield, M. (2014a). Two Big Ideas for Teaching Big Data: ECOTS.
www.statlit.org/pdf/2014-Schield-ECOTS.pdf
- Schield, M. (2011). Zicam and the US Supreme Course.
www.statlit.org/pdf/2011Schield-ASA-Zicam6up.pdf
- Schield, M. (2006). Presenting Confounding & Standardization Graphically.
STATS Magazine. www.StatLit.org/pdf/2006SchieldSTATS.pdf
- McKenzie, John, Jr. (2004). Teaching the Core Concepts. ASA
www.statlit.org/pdf/2004McKenzieASA.pdf