

Statistical Thinking, Literacy, and Effective Teaching

Thomas E. Love, Ph. D.
 Case Western Reserve University
 MetroHealth Medical Center
Thomas.Love@case.edu

What Is Statistical Literacy?

- In hospital A, 120 babies are born every day, in hospital B, only 12.
- On average, the ratio of baby boys to baby girls born every day in each hospital is 50/50.
- However, one day, in one of the hospitals twice as many baby girls as boys were born.
- In which hospital (A or B) was this more likely to happen?

Source: <http://statsoft.com/textbook/stathome.html> and Nisbett et al. (1987)

Fire Departments As First Responders

In 2008, fire departments nationwide responded to more than three times as many calls for medical help as they did in 1980.

Number of calls U.S. fire departments responded to 17.5 million calls for...

Year	Medical Help (million calls)	Fires (million calls)
1980	1.5	~1.5
2008	16.8	~1.5

Source: National Fire Protection Association
© 2008 NEW YORK TIMES

Firefighters Become Medics to the Poor

Firefighters from Company 18 responding to an emergency call in Washington in late August. Group Photo: The New York Times

WASHINGTON — Peeling off his latex gloves after treating a 4-year-old boy having a severe asthma attack, J. R. Myleart sighed with a touch of frustration. It was 3 a.m. and in the past 24-hour shift, Mr. Myleart, a firefighter, had responded to at least one emergency call per hour.

But only two of those calls were for fires; most of the others involved heart attacks, diabetic sores, epileptic seizures and people complaining of shortness of breath.

Multimedia

New York Times, September 4, 2009

“I keep saying the sexy job in the next ten years will be statisticians. People think I’m joking, but who would’ve guessed that computer engineers would’ve been the sexy job of the 1990s?”

*Hal Varian, Google’s Chief Economist
 McKinsey Quarterly, January 2009*

Continuation of Varian’s Quote

“The ability to take data—to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it—that’s going to be a hugely important skill in the next decades, not only at the professional level **but even at the educational level** for elementary school kids, for high school kids, for college kids. Because now we really do have essentially free and ubiquitous data. **So the complimentary scarce factor is the ability to understand that data and extract value from it.**”

What do we want to accomplish?

Got data? Got questions?
 Statistics is a necessary evil, so pay up
 Statistics provides rigorous and principled ways to think across your life
 Statistics is fun / exciting / sexy
 Statistics is too important to be left to statisticians

Data: Why Bother? 5 Great Motivators

- Greed
- Exclusivity
- Fear
- Guilt
- Need for Approval

The Netflix Prize

Greed

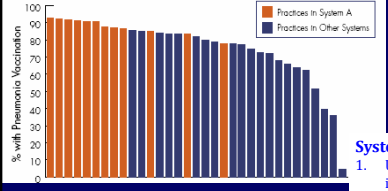
- Data help you draw more knowledge from raw information than anyone else



New York Times: September 21, 2009

Better Health Greater Cleveland

www.betterhealthcleveland.org




Exclusivity

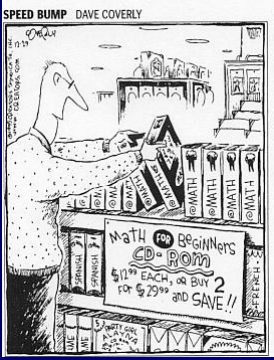
– not everyone knows this stuff.
 “Fixing” large systems
 Meaningful use of EMRs

System A's Strategy

1. Use EMR system each week to identify scheduled patients who need vaccine
2. When a patient on the list checks in, receptionist provides information on vaccine.
3. “Standing orders” from clinic docs authorize nurses to offer and administer vaccine.
4. Nurse documents patient consent and vaccine in EMR.


Guilt

- You and your students need to know this stuff to have a fighting chance –
- Information is ubiquitous, and we as statisticians are selling ways to harness it...



Fear

- People who already understand this stuff can use it to hurt the uninitiated, and do.



How to Lie With Statistics

- Everyone wants to collect bad data that supports them...
- “Would you prefer a delicious flame-broiled Whopper, or a soggy, slimy hunk of chemical-laden ‘meat’ prepared by unattractive workers wearing hairnets?”
- Most businesses use ethics as a competitive tactic, considering ethics to be making maximum returns for shareholders.

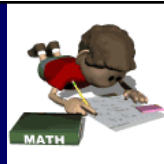
Need for Approval

- Your students will like this stuff – will think it’s exciting and interesting, and see it in their lives daily.

Goals for Today

- Describe what statistics is (and isn’t).
- Help you see the value of statistical thinking in your life and teaching.
- Give you some “news you can use”.
 - Provide some strategies for teaching and learning about statistical ideas
 - Provide strategies for learning from data

Grim Realities



- Teachers do noble work for minimal earthly rewards.
- Our students are not just younger versions of ourselves, and the ways in which they are different are important.
- “It is easy to lie with statistics. But it is easier to lie without them.” – F Mosteller

What Makes Statistical Thinking Hard To Teach?

- Students are uncomfortable with
 - The **messiness** of data
 - The different possible interpretations based on different assumptions
 - The **extensive use of writing and communication** skills
 - The importance of familiar contexts

Barriers to Effective Teaching of Statistical Thinking

- Students equate statistics with mathematics, and expect the focus to be on numbers, computations and one right answer
- Many familiar words have different or more precise meanings statistically

Statistics is ...

- about making sense of **DATA**
 - How do we gather the data?
 - How do we summarize the data, using both pictures and summary numbers?
 - How do we use the data to drive decisions or predictions about the world or the future?

“In God we trust; all others must bring data.”
--W. Edwards Deming, quality improvement pioneer

What Are Data?

- Data are numbers?
- Data are useless without context...
 - **WHAT** is described by these data
 - **HOW** were these data collected
 - **WHO** collected these data?
 - **WHEN** and **WHERE** were these data collected?
 - **WHY** were these data collected?

First Law of Statistics: **DTDP**

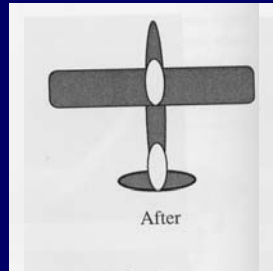
- Draw
- The
- D@\$%
- Picture



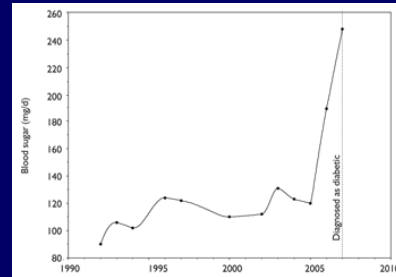
A picture is worth a lot of numbers...

Abraham Wald

- WW II
- Trying to determine where to add extra armor to planes
- Data used to drive decisions
- Assumptions?

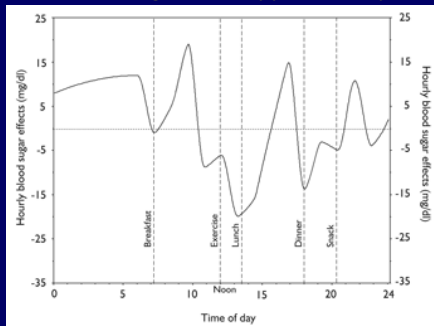


Opportunities to improve your life: 15 years of fasting blood sugar tests



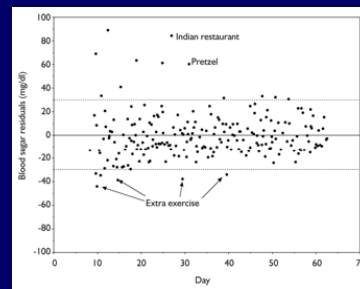
Howard Wainer and Paul Velleman (2008) Looking at blood sugar. *Chance*, 21(4), 56-61.

Average Hourly Blood Sugar Readings in a Typical Day



Howard Wainer and Paul Velleman (2008) Looking at blood sugar. *Chance*, 21(4), 56-61.

Residuals in Blood Sugar Readings after removing long-term trend and daily variation



Some unusual points are annotated to indicate probable cause. These **outliers** are the key to learning how to control blood sugar.

Howard Wainer and Paul Velleman (2008) Looking at blood sugar. *Chance*, 21(4), 56-61.

Statistics is...

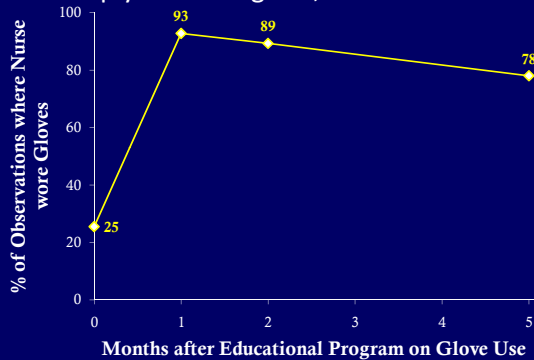
- about **VARIATION**
 - Variation results when two or more things, which we may think are exactly the same, turn out to be different.
 - Understanding and reducing variation in processes are keys to success
 - Variation is everywhere, in everything.

Glove Use Among Nurses

- Experiment about an educational program on importance of using gloves
- Subjects: 23 nurses in an inner-city pediatric hospital emergency dep't.
 - Observed during vascular access procedures before and 1, 2, 5 months after the educational program.
 - Nurse experience ranges from 1-20 years

Source: [Data and Story Library: http://lib.stat.cmu.edu/DASL/Datafiles/Nurses.html](http://lib.stat.cmu.edu/DASL/Datafiles/Nurses.html)

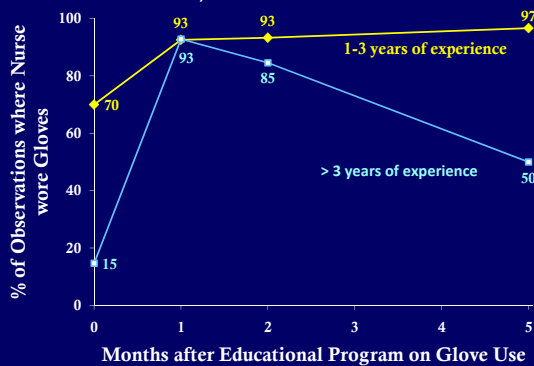
Across All 23 Nurses, Glove Use Jumped Sharply After Program; then Tailed Off



Improvement as Reducing Variation

- To meet this standard more often, we need to reduce individual variation – opposite of creative process...
- We want to do a better job holding the gains...
 - So can we think about a target?

Less Experienced Nurses Wore Gloves More Often, and Held the Gains



Without Understanding Variation, Can Evidence Based Medicine Succeed?

- Management by the last data point
- There's lots of firefighting.
- Using special cause methods to solve common cause problems.
- **Omnipresence of variation** means...
 - Conclusions are uncertain.
 - Avoid inference from coincidence.
 - Avoid inference from short-run irregularity.

Why Do We Need Data To Make Decisions?

- Without data ...
 - Everyone is an expert.
 - Discussions produce more **heat** than **light**.
 - Historical memory is poor.
- Data fuel “truth machines”
- Data help us triage the information flood

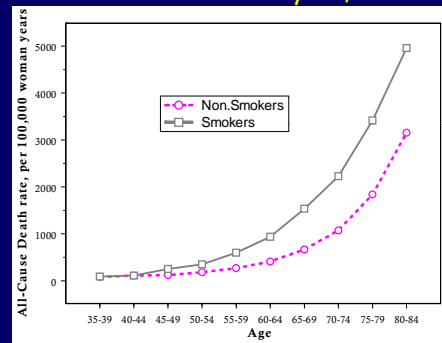
Statistics is ...

- about interacting with **TECHNOLOGY**
 - Students should think, machines should calculate.
 - Collecting the data is easier than thinking clearly about what you will do with it.
 - Detailed analyses can be generated with great speed – garbage in, garbage out...
 - Thinking about technology is unavoidable

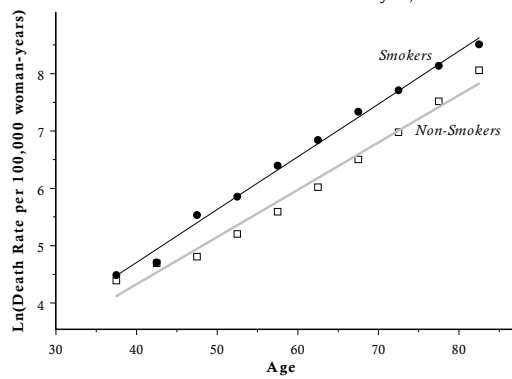
Smoking and Graphing Story from Howard Wainer



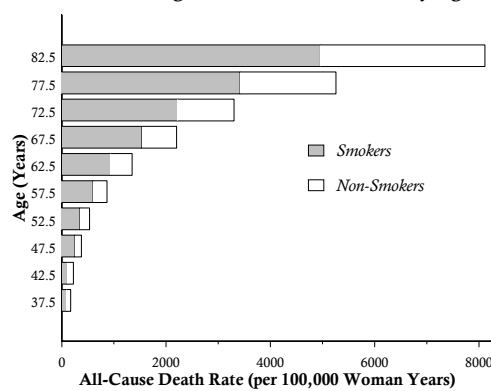
All-Cause Death Rates for Women: Current Smokers and Lifelong Non-Smokers Cancer Prevention Study II, 1982-88



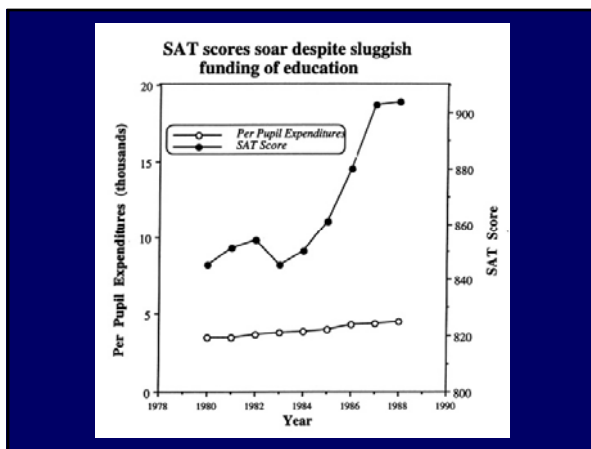
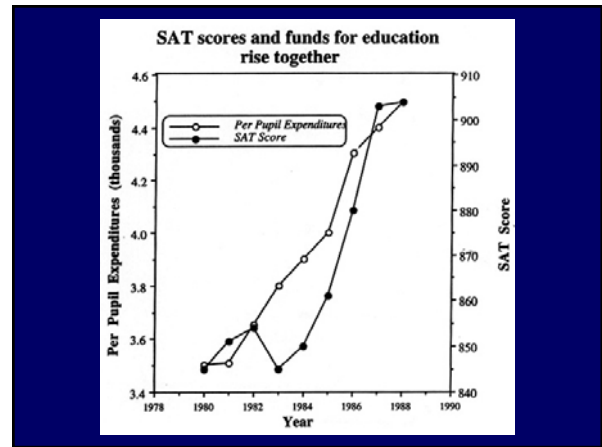
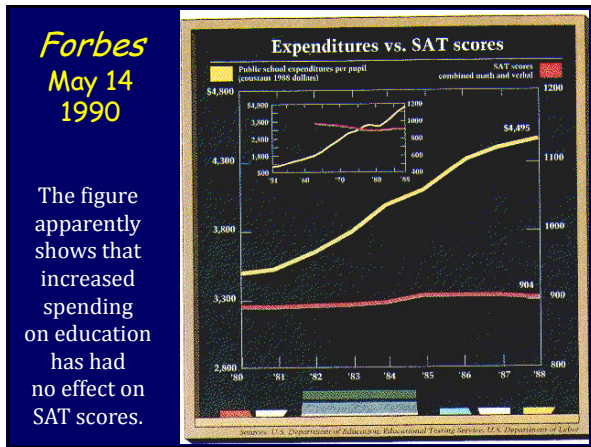
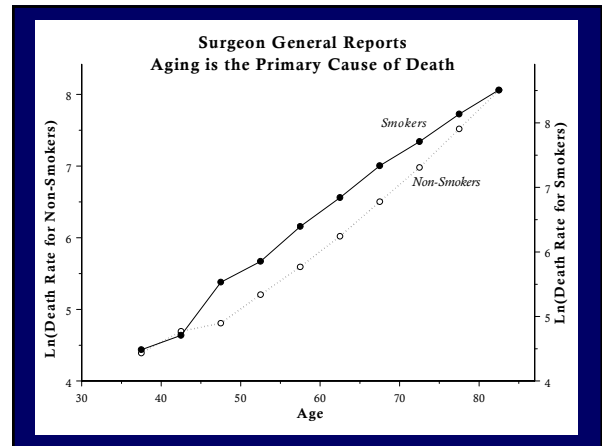
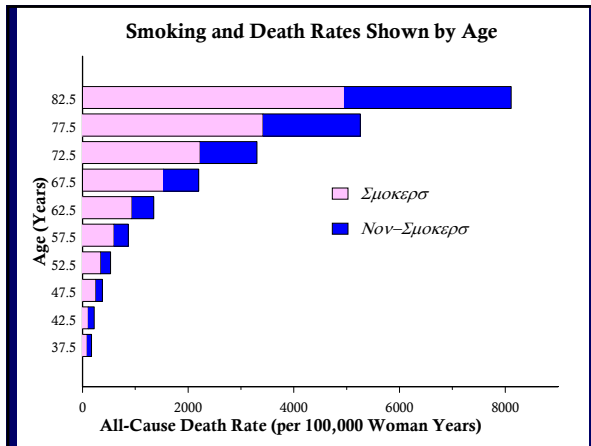
All-Cause Death Rate (Log Scale) Plotted Against Age Women in Cancer Prevention Study II, 1982-1988



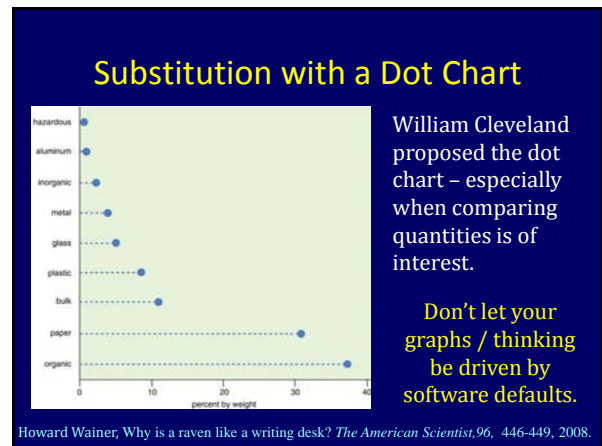
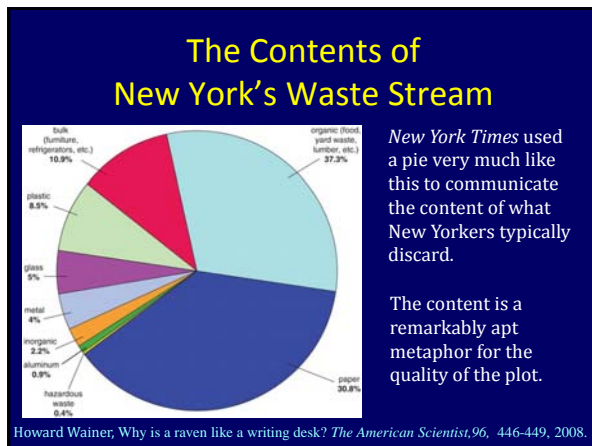
Smoking and Death Rates Shown by Age



Ohio Mathematics and Science Coalition 2009 Conference on Quantitative Literacy
 High-School Breakout Morning Session
 Thomas E. Love, Ph. D.
 Case Western Reserve University / MetroHealth Medical Center



- Graphical Displays Should...**
- Tell a story.
 - Show the data.
 - Induce the viewer to think about the substance of the data.
 - Avoid distorting what the data have to say.
 - Serve a clear purpose.



On Collecting Data Thoughtfully

“Far better an approximate answer to the right question ... than an exact answer to the wrong question.”

- John Tukey

Statistics in Action

The U.S. Federal Aviation Administration (FAA) has a unique device for testing the strength of airplane windshields.

A gun launches a dead chicken into a plane's windshield at the plane's maximum forward airspeed.

If the windshield doesn't crack, it is supposed that the windshield will survive an in-flight collision.

The British wanted to test a windshield of a high-speed locomotive under development. They borrowed the FAA's launcher, loaded the chicken and fired. The ballistic bird shattered the windshield, went through the engineer's chair, shattered an instrument panel and embedded itself in the aft wall of the engine cabin.

The stunned English engineers asked the FAA to review the test thoroughly. Their recommendation?

Thaw the chicken.

Motorcycle Helmets: Necessary?

- In testimony before a committee of the Hawaii State Senate, considering a law requiring all motorcyclists to wear a helmet, one witness declared that despite having been in several accidents during his 20 years of motorcycle riding, a helmet would not have prevented any of the injuries he received.

Who was **unable to testify**? Why?

If you had to do it over again, would you have children?

- Ann Landers asked parents this question.
- 70% of the more than 10,000 people who wrote in said “No”

Who do you suppose would be most likely to participate in such a poll?

- A more carefully designed survey later showed that about 90% of parents actually are happy with their decision.

Unintentional Bias

“What was the most important date in your life?”

- ◆ Most important calendar date?
- ◆ Most important social engagement with a potential partner?
- ◆ Most important shriveled fruit?

Starting and Finishing Fights

- In 90% of all deaths resulting from barroom brawls, the victim was the one who instigated the fight.

When the police asked “Who started this?”
I guess the other 10% did not point at the body
on the floor.

When the Answer is Important

In order to conduct a study properly,

- Get a **representative** sample.
- Get a **large enough** sample.
- Decide whether it should be an **observational study** or an **experiment**.
- Decide precisely what you want to know - what question(s) do you need answered?
- Never spend your entire budget on the 1st run.

Randomized Trials vs. Observational Studies

- We have an outcome measured on two groups of subjects (**treated** and **control**).
- We want to make a **fair comparison** between the treated group and the control group in terms of the outcome.
- **Observational Study**: people come to us as members of a group (treated or control)
- **Randomized Trial**: we assign people to groups (using a random mechanism)

A Randomized Trial: Therapeutic Touch

- In Therapeutic Touch (TT) therapy, the practitioner moves her hands near a patient, attempting to manipulate a “human energy field” (HEF), ostensibly to promote healing.
- We’ll discuss a trial using 15 TT practitioners.
- **Research Question**: Can the TT practitioner detect which of their hands (left or right) the investigator’s hand is hovering over? (chosen at random via a coin flip)

Rosa L. Rose E et al. “A Close Look at Therapeutic Touch” JAMA, 279(13), 1005-1010.

Experimental Procedure



Emily Rosa showing how her apparatus was marked to facilitate placement of her hands.

- Challenge to practitioners:
 - Each of the 15 subjects warmed up first, then attempted 10 trials each.
 - Of 150 trials, there were 70 successes, or 46.7%.
- Can TT practitioners detect a HEF?

Results of the TT Trial

- We want to know if the TT practitioners can detect a HEF at better than chance levels. [Null hypothesis: $pr(\text{detection}) = .50$]
 - Choice of hand **randomized** by coin flip
- Result: 70 out of 170 yields P value = .795.
 - If the true proportion of successes is 50%, then an observed proportion of 46.7% successes or more would occur at random about 8 times out of 10.
 - That's not a rare event, so **there is insufficient evidence to conclude that the practitioners are performing better than if they were just guessing.**

Statistical Thinking in Designing an Experiment

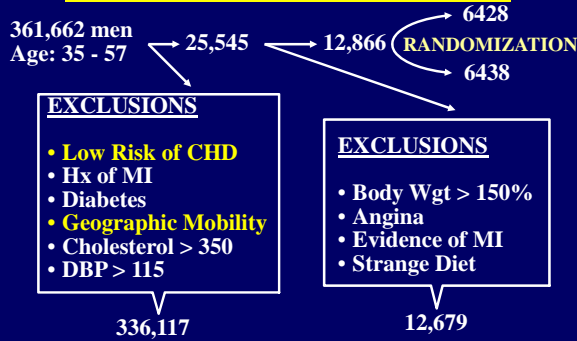
- A credit card bank wanted to test the sensitivity of the market to two factors:
 - the **annual fee** charged for a card and
 - the **annual percentage rate** charged.
- The bank selected people at random from a mailing list and offered ...
 - 50,000 people a low rate card with no fee
 - 50,000 a higher rate card with a \$50 annual fee
- What happened?

Why Not Always Do Randomized Trials?

- Randomized Trials...
 - Are Expensive and Time-Consuming
 - May be unethical (e.g., smoking)
 - Practice patterns may make them difficult to conduct (e.g., rehab for stroke, right heart cath for seriously ill patients)
 - Often suffer from limited generalizability, owing to the need to exclude subjects

RCT Subject Selection

MRFIT Excluded 96.4% of potential eligibles



Advantages of Observational Studies

- Are widely available, with detailed data
- Enable study of exposures not amenable to randomized trials
- Often work better for exposures with small effect sizes
- Enable study of effects in the "real world"

"Hormone Replacement Therapy Should Be Recommended for Nearly All Women"

Col NF, Eckman MH, *et al.* (1997) Patient-specific decisions about hormone replacement therapy in postmenopausal women. *Journal of the American Medical Association (JAMA)*, 277: 1140-47.

"Do not use estrogen/progestin to prevent chronic disease"

Fletcher SW, Colditz G. (2002) Failure of estrogen plus progestin therapy for prevention. *JAMA*, 288: 366-67.

Manson JE *et al.* for WHI Investigators (2003) Estrogen plus Progestin and the Risk of Coronary Heart Disease. *New England Journal of Medicine*, 349: 523-534.

Observational Studies vs. Randomized Clinical Trials

- In the observational study, patients were either taking HRT or not, on the basis of their desires and their doctor's recommendation.
- In the RCT, 16,608 women ages 50-79 were randomly assigned to either HRT or to a placebo.

Relative Risk of Hormone Replacement Therapy and Coronary Heart Disease

- Observational Studies (1997 and earlier): CHD risk for patients using HRT was **0.60** times as high as those not using HRT
- WHI Randomized Clinical Trial: CHD risk for patients using HRT was **1.29** times as high as those not using HRT

How Did This Happen?

- **Selection Bias!** Women taking HRT may be more likely to...
 - engage in other CHD-related healthy behaviors (exercise, diet, not smoking)
 - have fewer CHD-related co-morbid illnesses (diabetes, high blood pressure)
 - Take other cardio-protective drugs (lipid-lowering agents, etc.)

Statistical Thinking and Your Lives

There are a lot of people in the public who are intimidated by numbers and I think a lot of them become reporters and editors.

– Tom Barrett, political reporter, Vancouver Sun

Veteran journalist Robert Niles...

"When it snowed heavily, we'd call the sheriff, and ask how many fender-benders. Inevitably, we'd have a lead:

A fierce winter storm dumped 8 inches of snow, snarled rush-hour traffic and caused 28 fender-benders on county freeways.

Niles Quote ...

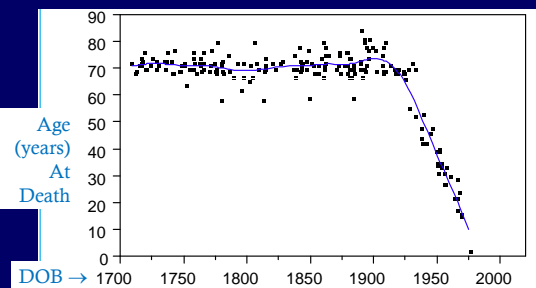
...until one day I dared to ask the sheriff's department how many fender-benders were reported on clear, sunny days...

The answer – 48."

A fierce winter storm dumped 8 inches of snow, snarled rush-hour traffic and prevented 20 fender-benders on county freeways.

The Lombard Surprise

- Data: 204 birth and death dates from NJ cemetery
- Average age at death relatively constant until 1920 birth dates, then declines rapidly ... *Why?*



Statistical Literacy

- **Filters for nonsense:**
 - Triage on the information flood --
 - "Life's too short for me to read this..."
- **Data beat anecdotes.**
 - Think broadly: Is this the right question?
 - Think broadly: Does the answer make sense?
 - Communication: Can you read a graph?

Data Strategies

(details to fit the context)

- **Plot** your data
- **Interpret** what you see
- Remember the **context**
 - HOW, WHAT, WHY
- Consider the best "next step"
 - Numerical summary?
 - Mathematical model?

This is serious stuff.

- Understanding chance variation.
 - Patterns and deviations (fit and residual)
 - Model-data dialogue (diagnostics / checking underlying assumptions)
- One pass through software is not enough.
 - Models as interpretive tools.
 - Strategies, not just methods.
- It means you have to think about things more than you want to think.

Statistics is too important to be left to Statisticians !

- ## Implications for Instruction - Developing Good Habits
- How should we obtain meaningful, relevant data to answer the question?
 - Relation of data to the context
 - Interpretation of conclusions in non-statistical terms
 - Thinking beyond the textbook
 - Help students see relevance of statistics outside of math class
- Source: www.rossmanchance.com/papers/aera.html

- ## Some Additional Examples
- ➔ 1. Is Using a Car Phone Like Driving Drunk?
 - ➔ 2. Should Colleges Institute an SAT-Optional Policy?
 - ➔ 3. How Much Does Extra Time on the SAT Help?
 - ➔ 4. Are Winter Babies at a Disadvantage?
 - ➔ 5. Mapping Disease
- [Final Slide](#)

CHANCE Wiki (#52) at chance.dartmouth.edu/chancewiki

Is Using a Car Phone Like Driving Drunk?

[Back to Menu of Examples](#) [Final Slide](#)

New York Times
 July 19, 2009

CHANCE Wiki (#52)
chance.dartmouth.edu/chancewiki

Other Drivers Are the Problem, Not Me

A national poll found that 48 percent of people say talking on a cellphone is the most dangerous driving distraction. Yet most people who have used a cellphone while driving consider themselves safe drivers.

Which of the following do you feel is the most dangerous distraction for people while driving?

Using technology like a cellphone	48%
Reading	1%
Adjusting the radio	1%
Talking to someone else in the vehicle	1%
Daydreaming	1%
Using a hands-free cellphone	1%
Distractions related to another driver	1%
Eating	1%
Using a GPS device	1%
Don't know	1%

Have you ever been hit or nearly hit by someone using a cellphone?
 YES: 4% NO: 96% DON'T KNOW: 0%

Have you ever talked on your cellphone while driving?
 YES: 81% NO: 19%

Have you ever texted while driving?
 YES: 16% NO: 84%

Do you consider yourself to be a safe driver?
 YES: 81% NO: 19%

Only asked of cellphone owners.

Based on telephone interviews conducted by Nationwide Mutual Insurance from April 15 to 24, 2008, with 1,026 drivers ages 16 to 81, in the 48 contiguous states. Numbers do not add to 100 in all cases because of rounding.

Source: Nationwide Mutual Insurance Company Driving While Distracted Survey, 2008

NYT July 2009

- A national poll found that 48% of people say talking on a cellphone is the most dangerous driving distraction...

Is Using a Car Phone Like Driving Drunk?

Donald A. Redelmeier and Robert J. Tibshirani

What's wrong with this picture?

<http://www-nrd.nhtsa.dot.gov/departments/Human%20Factors/driver-distraction/PDF/redel2.PDF>

Case-Crossover Design

Each individual is evaluated on two intervals

- A **hazard** interval, in which event occurs
- and one in which the event does not occur (**control** interval).

Potential exposures are then measured during each interval, and possible associations are tested by comparing the exposures in the hazard interval to the exposures in the control interval.

Primary Analysis on 699 drivers who had a collision with significant damage but no injury

- Phone linked to detailed billing data.
- Were drivers more likely to have made a cell phone call during the 10-minute interval immediately before the collision or during a similar interval on the day before?

The Results

- Drivers **WERE** more likely to have made a cell phone call during the 10-minute interval immediately before the collision.
- Much** more likely: 24% vs. 5%.
- Implications of this?

Cell Phone Driving Risks

		Cellular telephone call during Control Interval?	
		yes	no
Cellular telephone call during Hazard Interval?	yes	13	157
	no	24	505

Relative risk: $157/24 = 6.5$
 95% confidence interval: (4.5 - 10.0)

Drivers were at a **6.5-fold increased risk** of a motor vehicle collision when they were using a telephone as compared to when they were not using a telephone.

CHANCE Wiki (#52) chance.dartmouth.edu/chancewiki/index.php/Chance_News_52

How does this compare to driving while intoxicated?

- Driving with a blood alcohol level at the legal limit is associated with a relative risk of 4 in several studies.
- Driving with a blood alcohol level 50% above the legal limit is associated with a 10-fold increase in risk.

<http://www-nrd.nhtsa.dot.gov/departments/Human%20Factors/driver-distraction/PDF/redel2.PDF>

Howard Wainer: What happens if the SAT is made optional?
 Bowdoin College as an example. *The American Scientist* 2009, to appear.
 OR <http://blog.criteriacorp.com/blog/bid/6927>

Should Colleges Institute an SAT-Optional Policy?

[Back to Menu of Examples](#) [Final Slide](#)

National Association for College Admission Counseling Report

NY Times 9/22/2008

- Report critical of SAT and ACT
 - Commission chaired by dean of admissions and financial aid at Harvard.
- Conclusions ...
 - Schools should consider making their admissions "SAT optional," that is allowing their applicants to submit their SAT/ACT scores only if they wish.
 - The commission cites the success that pioneering schools with this policy have had in the past as proof of concept.

Has the admissions process been hampered in schools that have instituted an SAT optional policy?

- Bowdoin College instituted such a policy in 1969.
- How did the students who didn't submit SAT scores do at Bowdoin in comparison to those students that did submit them?
- Would the non-submitters performance at Bowdoin have been predicted by their SAT scores?
- Howard Wainer looked at six colleges with similar observed mean SAT scores (to Bowdoin) for the entering class of 1999.

6 Colleges/Universities with similar observed mean SAT scores for entering class of 1999

Institution	All Students	Submitted SAT Scores		Did not Submit
	N	N	Mean	N
Northwestern University	1654	1505	1347	149
Bowdoin College	379	273	1323	106
Carnegie-Mellon University	1132	1039	1319	93
Barnard College	419	399	1297	20
Georgia Inst. of Technology	1667	1498	1294	169
Colby College	463	403	1286	60
Means and Totals	5714	5117	1316	597

How did the students who didn't submit SAT scores do at Bowdoin in comparison to those students that did submit them?

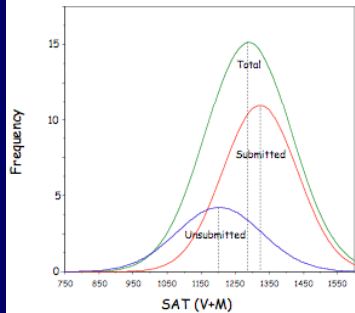
- Non-SAT submitters did about **a standard deviation worse** than students who did submit SAT scores.
- Curves here are approximations based on the normal distribution – they're pretty reasonable...

<http://blog.criteriacorp.com/blog/bid/6927/Don-t-Ask-Don-t-Tell-The-New-Rules-of-the-SAT-and-College-Admissions>

Who Are the Non-SAT Submitters?

- If the Bowdoin admissions office were using other variables to make up for the missing SAT scores, those variables did not contain enough information to prevent them from admitting a class that was academically inferior to the rest.
- But would their SAT scores have provided information missing from other submitted information?
- All of these students actually did take the SAT, and ETS gathered those results...

Distribution of SAT Scores for those who submitted them and those who did not at Bowdoin (Class of 1999)



- Applicants behaved rationally.
- Students who don't submit SAT scores to Bowdoin score about 120 points lower than those who do submit their scores.

Would these SAT scores have led to accurate first year GPA predictions?

- Correlation(grades, SAT scores) is 12% higher for those who didn't submit their SAT scores than for those who did.
- So not having this information doesn't improve the academic performance of Bowdoin's entering class, it diminishes it.
- So why would a school opt for such a policy?

If all students had submitted SAT, Bowdoin's mean SAT would rank below Carnegie-Mellon and Barnard

Institution	All Students		Submitted SAT Scores		Did not Submit	
	N	Mean	N	Mean	N	Mean
Northwestern	1654	1338	1505	1347	149	1250
Bowdoin	379	1288	273	1323	106	1201
Carnegie Mellon	1132	1312	1039	1319	93	1242
Barnard	419	1293	399	1297	20	1213
Georgia Tech	1667	1288	1498	1294	169	1241
Colby	463	1278	403	1286	60	1226
ALL	5714	1307	5117	1316	597	1234

Wainer, Bridgeman, Najarian and Trapani (2004)
 How much does extra time on the SAT help? *Chance*, 17:2, 19-24

How Much Does Extra Time on the SAT Help?

[Back to Menu of Examples](#)

[Final Slide](#)

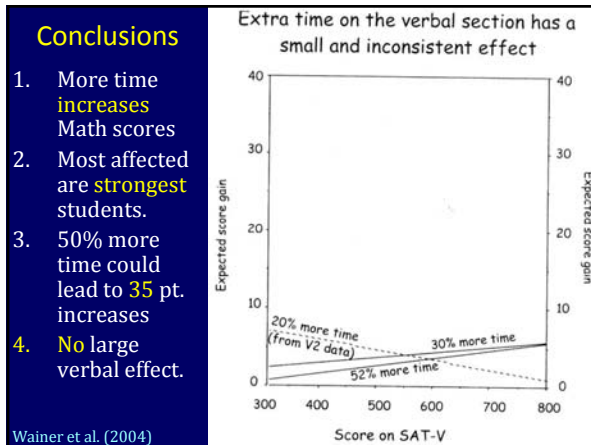
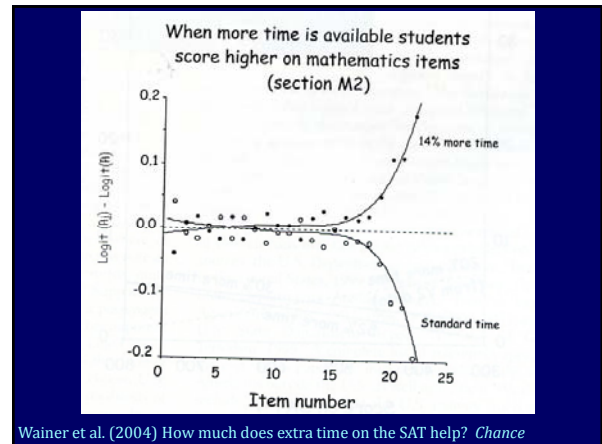
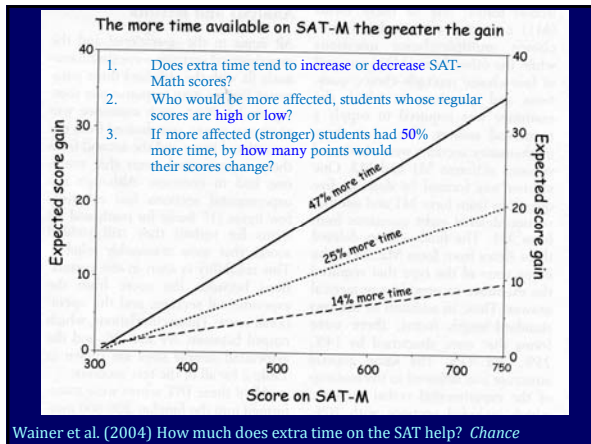
How Much Does Extra Time on the SAT Help?

- SAT I: Reasoning Test
 - Is speed particularly important?
 - # of students requesting extra time is way up
 - Are students gaming the system effectively?
- Do students benefit from additional time?
 - October 2000 administration (3 math sections, 3 verbal and 1 experimental)
 - 100,000 examinees randomly allocated to 10 configurations (sizes) of the experimental section

Does Extra Time on the SAT Help? Four Key Questions

1. Does extra time tend to **increase** or **decrease** student SAT-Math scores?
2. Who would be **more** affected, students whose scores on the regular test are **high** or **low**?
3. If the more affected students had **50%** more time, by **how many** points would you expect their scores to change?
4. Would scores be more affected on the **Math** or the **Verbal** sections of SAT?

P
L
O
T



Conclusions

- More time **increases** Math scores
- Most affected are **strongest** students.
- 50% more time could lead to **35 pt.** increases
- No large** verbal effect.

Are Winter Babies at a Disadvantage?

Back to Menu of Examples Final Slide

When do you want to be born?

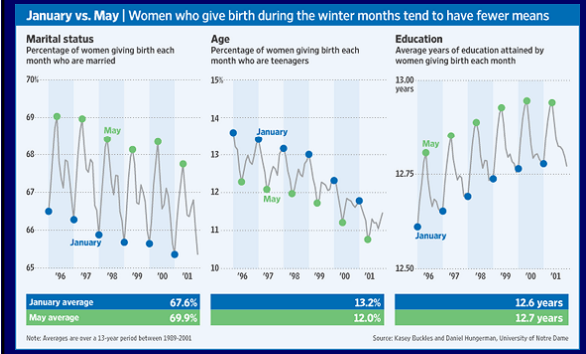
- Multiple studies have found that winter kids...
 - test poorly,
 - don't get as far in school,
 - earn less,
 - are less healthy,
 - and don't live as long

http://chance.dartmouth.edu/chancewiki/index.php/Chance_News_55
<http://online.wsj.com/article/SB125356566517528879.html>

Are Winter Babies at a Disadvantage?

- Researchers have spent years on this. Why might this be?
 - School attendance laws (dropout at 16)
 - Seasonal Sunshine or Pesticide levels
- Key Assumption:
 - Backgrounds of winter kids are the same as those of other kids
- CDC birth certificate data...?

Women giving birth in winter (1989-2001) were less often married, younger & less educated

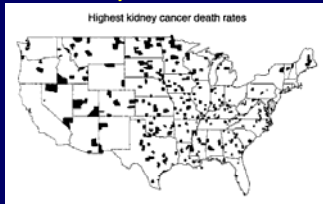


Mapping Disease

[Back to Menu of Examples](#)

[Final Slide](#)

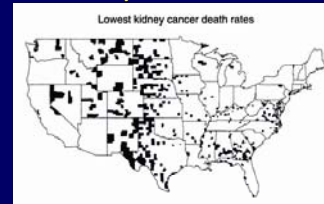
Highest Kidney Cancer Death Rates



- Map shows U.S. counties with **highest** 10% age-standardized death rates for cancer of kidney / ureter for white males in 1980-1989.
- Why are most of the shaded counties in the center-west of the country?

Source: Gelman and Nolan *Teaching Statistics: A Bag of Tricks*

Lowest Kidney Cancer Death Rates



- Map shows U.S. counties with **lowest** 10% age-standardized death rates for cancer of kidney / ureter for white males in 1980-1989.
- Why is this pattern somewhat similar to the map of the highest rates?

Source: Gelman and Nolan *Teaching Statistics: A Bag of Tricks*

Stumped?

- Consider a county with 100 people.
 - If it has even one kidney cancer death in the 1980s – its rate is 1 per thousand per year – among the highest in the nation.
 - If it has no kidney cancer deaths – its rate will be (tied for) the lowest nationwide.

Thank You

Thomas E. Love, Ph. D.
 Case Western Reserve University
 MetroHealth Medical Center
Thomas.Love@case.edu
www.chrp.org/love

An 20-page color PDF handout is at <http://www.chrp.org/love/91009slides.pdf>

Netflix Prize Learn more about the prize at www.netflix.com

Better Health Greater Cleveland One of my day jobs www.betterhealthcleveland.org

Abraham Wald A brief description of Wald's work on airplane armor is provided in Howard Wainer's wonderful book *Visual Revelations: Graphical Tales of Fate and Deception from Napoleon Bonaparte to Ross Perot*. The relevant article in *Chance* magazine is by Wainer, Palmer and Bradlow, as part of a survey of selection anomalies and appeared in the second issue of 1998. And see slide 21 in "Statistics in Aviation: Celebrating 100 Years of Flight" by Fritz Scholz at <http://www.stat.uiowa.edu/SRC2003/papers/f-scholz.pdf>. A more technical description is given by Mangel M and Samaniego FJ (1984) "Abraham Wald's work on aircraft survivability," *Journal of the American Statistical Association*, 79, 259-267. Sadly, Wald's life was tragically cut short at age 48 in an airplane crash while on a lecture tour of India.

Fasting Blood Sugar Example comes from Howard Wainer and Paul Velleman in *Chance* magazine (2008) – volume 21, #4. *Chance* is now published by the American Statistical Association – visit <http://www.amstat.org/publications/chance/2009/22.2Web/index.cfm> and also <http://www.amstat.org/education/onlineresources.cfm> for some great (free) resources. They also publish *STATS* – a student magazine <https://www.amstat.org/publications/stats>

Glove Use Among Nurses is from the Data and Story Library at <http://lib.stat.cmu.edu/DASL/>

Smoking and Death of Women and Forbes SAT Scores vs. Education and Other Graphs

Details on the Cancer Prevention Study II, including source material for my first plot can be found at http://www.cdc.gov/tobacco/sgr/sgr_forwomen/pdfs/chp3.pdf. The age-adjusted death rates for lung and breast cancer among U.S. women come from the CDC as well. The main story I told happened to Howard Wainer, who (using a different data set) recounts it in his book *Visual Revelations*. Howard also tells the story (including the Forbes SAT and Education example) much better than I can in his special one-hour lecture *How to Display Data Badly*, which is both informative and hilarious, and free at <http://www.dartmouth.edu/~chance/> Howard's more recent books are also great, and I mentioned those during the session.

Therapeutic Touch Detailed links to sites both supportive of and skeptical of TT and related phenomena can be found at <http://www.phact.org/e/tt/> The picture I present comes from that site. A very nice CNN piece is at <http://www.cnn.com/HEALTH/9804/01/therapeutic.touch/>

Credit Card Bank Experiment This was described in the best textbook I am aware of for teaching statistics to high-school students – by David Bock, Paul Velleman and Richard DeVeaux, entitled *Stats: Modeling the World*, (2004) Addison-Wesley.

MRFIT and Randomized Trial Subject Selection The Multiple Risk Factor Intervention Trial is discussed in the September 1982 issue of the *Journal of the American Medical Association*, and, later, in many other places. MRFIT was a randomized primary prevention trial to test the effect of a multifactor intervention program on mortality from coronary heart disease (CHD) in 12,866 high-risk men aged 35 to 57 years. Men were randomly assigned either to a special

intervention (SI) program consisting of stepped-care treatment for hypertension, counseling for cigarette smoking, and dietary advice for lowering blood cholesterol levels, or to their usual sources of health care in the community (UC). Over an average follow-up of 7 years, risk factor levels declined in both groups, but to a greater degree for the SI men.

WHI Studies of Hormone Replacement Therapies The Women's Health Initiative (WHI) is a major 15-year research program to address the most common causes of death, disability and poor quality of life in postmenopausal women -- cardiovascular disease, cancer, and osteoporosis. In July 2002, the investigators stopped the estrogen-plus-progestin study after finding that the associated health risks outweighed the benefits. The researchers are continuing to analyze and report on data from this trial while they complete other WHI studies. More is available at <http://www.nhlbi.nih.gov/whi/>

The Lombard Surprise This comes from Wainer, Palmer and Bradlow in *Chance* magazine, as part of a survey of selection anomalies in 1998.

Additional Examples

1. Is using a car phone like driving drunk? Comes from the CHANCE NEWS wiki #52 (see below)
2. Should colleges institute an SAT-Optional policy? Comes from Howard Wainer, and will appear in an article in *The American Scientist* later this year.
3. How much does extra time on the SAT help? Also from Howard Wainer and colleagues, and appeared in 2004 in *Chance* magazine.
4. Are winter babies at a disadvantage? Comes from the CHANCE NEWS wiki (#55) and was also reported on by the Freakonomics blog at the New York Times and by the Wall Street Journal – which has a “Numbers Guy” columnist.
5. The Mapping kidney disease example comes from Andrew Gelman and Deborah Nolan’s book *Teaching Statistics – A Bag of Tricks*. <http://www.stat.columbia.edu/~gelman/blog/>

General Resources

The CHANCE NEWS wiki <http://chance.dartmouth.edu/chancewiki> is an outstanding (again, free) resource of material for incorporating statistical and quantitative literacy ideas from current events into mathematics and statistics courses. I have also found useful materials at the online *Journal of Statistics Education*, <http://www.amstat.org/publications/jse/> and at www.statlit.org Some serious thinking about statistics education has been done by Allan Rossman and Beth Chance – they have a website at www.rossmanchance.com

I do virtually all of my work in R – visit the R project at <http://www.r-project.org/> which is so filled with stuff you can’t believe it, if you’re interested in really learning a lot about statistics and programming. The StatCrunch software site <http://www.statcrunch.com> has some nice stuff, including a free Google plotter to build maps of data.

“I’m talking about simple stuff. I’m talking about percentages, proportions, ratios, and rates. It’s the sort of thing that gets handled, if handled at all, in the first week of the introductory statistics course and then it’s assumed that ‘Hey, we all know this stuff. We can move on.’ Yet those of us who spend our time talking with students know full well that, in fact, this isn’t something they all understand. Their confusion affects the way they read the newspaper, it affects the way that they vote, and it affects the way that they understand the world around them.” - **Joel Best**