Enhancing the Visibility of the Statistics Profession: The IYOS and You

Marie Davidian

Department of Statistics
North Carolina State University

http://www4.stat.ncsu.edu/~davidian
Nate Silver, as well as Drew Linzer, Mark Blumenthal/Simon Jackson, Sam Wang, and others predicted the outcome of the 2012 US presidential election.
Nate Silver, as well as Drew Linzer, Mark Blumenthal/Simon Jackson, Sam Wang, and others predicted the outcome of the 2012 US presidential election using . . .

Statistics

http://fivethirtyeight.blogs.nytimes.com/
"Nate Silver-led statistics men crush pundits in election"
– Bloomberg Businessweek

“Nate Silver has made statistics sexy again”
– Associated Press

“Drew Linzer: The stats man who predicted Obama’s win”
– BBC News Magazine

“The allure of the statistics field grows”
– Boston Globe
“Nate Silver-led statistics men crush pundits in election”
– Bloomberg Businessweek

“Nate Silver has made statistics sexy again”
– Associated Press

“Drew Linzer: The stats man who predicted Obama’s win”
– BBC News Magazine

“The allure of the statistics field grows”
– Boston Globe

But interest in statistics didn’t start with the US elections...
August 6, 2009

For Today’s Graduate, Just One Word: Statistics

By STEVE LOHR

MOUNTAIN VIEW, Calif. — At Harvard, Carrie Grimes majored in anthropology and archaeology and ventured to places like Honduras, where she studied Mayan settlement patterns by mapping where artifacts were found. But she was drawn to what she calls “all the computer and math stuff” that was part of the job.

“People think of field archaeology as Indiana Jones, but much of what you really do is data analysis,” she said.

Now Ms. Grimes does a different kind of digging. She works at Google, where she uses statistical analysis of mounds of data to come up with ways to improve its search engine.

Ms. Grimes is an Internet-age statistician, one of many who are changing the image of the profession as a place for dorkish number nerds. They are finding themselves increasingly in demand — and even cool.

“I keep saying that the sexy job in the next 10 years will be statisticians,” said Hal Varian, chief economist at Google. “And I’m not kidding.”

“I keep saying that the sexy job in the next 10 years will be statisticians” – Hal Varian, Chief Economist, Google
What Are the Odds That Stats Would Be This Popular?

By QUENTIN HARDY

“Most of my life I went to parties and heard a little groan when people heard what I did,” says Robert Tibshirani, a statistics professor at Stanford University. “Now they’re all excited to meet me.”

It’s not because of a new after-shave. Arcane statistical analysis, the business of making sense of our growing data mountains, has become high tech’s hottest calling. There are billions of bytes generated daily, not just from the Internet but also from sciences like genetics and astronomy. Companies like Google and Facebook, as well as product marketers, risk analysts, spies, natural philosophers and gamblers are all scouring the info, desperate to find a new angle on what makes us and the world tick. Computing has become cheap and available enough to process any number of formulas.

What no one has are enough people to figure out the valuable patterns that lie inside the data.
Statistical Habits to Add, or Subtract, in 2013

By CARL BIALIK

In the year ending Monday, we saw some gains in statistical savvy: Data-crunching pollsters accurately forecast the outcome of the presidential election; the Memphis Grizzlies hired a vice president of basketball operations for his statistical expertise; and folks grew comfortable with the phrase "big data," to describe the billions of billions of bytes generated daily by information technology.

The growing importance of statistical analysis is set to be a theme of next year, too, with more than 150 professional organizations worldwide, including the American Statistical Association, designating 2013 as the International Year of Statistics.
Data Crunchers Now the Cool Kids on Campus

By CARL BIALIK

On campuses, at cocktail parties and in American corporations, statisticians are walking a bit taller these days.

The explosive growth in data available to businesses and researchers has brought a surge in demand for people able to interpret and apply the vast new swaths of information, from the analysis of high-resolution medical images to improving the results of Internet search engines.

Schools have rushed to keep pace, offering college-level courses to high-school students, while colleges are teaching intro stats in packed lecture halls and expanding statistics departments when the budget allows.

Number of Students (thousands)

Year
David Alexander’s job didn’t exist ten years ago. He works for Pacific Biosciences in Menlo Park, California, writing software that can analyse the data generated by DNA polymerase enzymes, which sequence DNA in real time. A decade ago, it took scientists weeks to sequence DNA, one base at a time, using a seemingly endless series of reactions. Back then, they also thought that they would be able to find the roots of major diseases just by identifying the common genetic variants shared by affected individuals.

Both the technology and the hypotheses have changed greatly since then. In the mid-to-late-2000s, while Alexander was working towards his PhD, scientists were using genome-wide association studies (GWAS) — searching genomes for known genetic variants that are shared by people with a particular disease or trait. By the time he graduated, last June, GWAS had mostly been superseded by techniques that sequence entire genomes. The machines designed to do this sequencing are pouring out huge amounts of data, thereby creating a huge need for mathematics and statistics experts. So Alexander, and many others working on statistical genetics, now have many more opportunities. “Scientifically, there are much richer questions to ask, and there are still a lot of deep discoveries to be made, it’s an interesting time,” he says. His career track reveals just how much opportunities in the field have changed.

CAREER VARIATION

It was not for a lack of trying that GWAS didn’t pan out. The completion of the Human Genome Project in 2003 spurred major funders from around the world to invest millions of dollars to build an international haplotype map, a catalogue of all the common human variants at single bases, called single nucleotide polymorphisms (SNPs), to be used in GWAS. The SNP map should have helped researchers to identify genes that are associated with disease. But instead, it showed that SNPs don’t account for much of the heritability of disease.

Researchers now think that many rare variants play a part in causing disease, but rare variants are much harder to find than the common SNPs. As a result, statistical geneticists are now mining sequence data for directly...
2011 McKinsey Global Institute report:

Big data: The next frontier for innovation, competition, and productivity

“A significant constraint...will be a shortage of...people with deep expertise in statistics and data mining...a talent gap of 140K - 190K positions in 2018 (in the US)”

http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation
Based on the 2003 book by Michael Lewis, *Moneyball: The Art of Winning an Unfair Game*
But it hasn’t all been so rosy…
Data Science: The Numbers of Our Lives

By CLAIRE CAIN MILLER

HARVARD BUSINESS REVIEW calls data science “the sexiest job in the 21st century,” and by most accounts this hot new field promises to revolutionize industries from business to government, health care to academia.

The field has been spawned by the enormous amounts of data that modern technologies create — be it the online behavior of Facebook users, tissue samples of cancer patients, purchasing habits of grocery shoppers or crime statistics of cities. Data scientists are the magicians of the Big Data era. They crunch the data, use mathematical models to analyze it and create narratives or visualizations to explain it, then suggest how to use the information to make decisions.

In the last few years, dozens of programs under a variety of names have sprung up in response to the excitement about Big Data, not to mention the six-figure salaries for some recent graduates.

In the fall, Columbia will offer new master’s and certificate programs heavy on data. The University of San Francisco will soon graduate its charter class of students with a master’s in analytics. Other institutions teaching data science include New York University, Stanford, Northwestern, George Mason, Syracuse, University of California at Irvine and Indiana University.
Big Omission

Articles on Big Data and data science abound

- Computer science, mathematics, engineering, physics, analytics, . . .
- *Statistics* is often MIA
And in science...
“My impression is that scientists view statistics not so much as a science but as a ‘bag of tools.’ You have a visibility problem in Science and AAAS.”

– Alan Leshner, Executive Director of the American Association for the Advancement of Science (AAAS), to representatives of the ASA in 2011
Do we really have a “visibility problem?”

Yes, and no

• We continue to play our traditional and critical roles as respected team members in a host of areas
• We have gotten lots of great press
• Enrollments in undergraduate statistics programs are skyrocketing, as are applications to graduate programs
• But statistics continues to be omitted from much of the discourse on Big Data and data science
• And we do not have the stature we would like among many of our fellow scientists
• Moreover, our discipline and its value continue to be misunderstood and unappreciated
Do we really have a “visibility problem?”

Yes, and no

• We continue to play our traditional and critical roles as respected team members in a host of areas
• We have gotten lots of great press
• Enrollments in undergraduate statistics programs are skyrocketing, as are applications to graduate programs
• But statistics continues to be omitted from much of the discourse on Big Data and data science
• And we do not have the stature we would like among many of our fellow scientists
• Moreover, our discipline and its value continue to be misunderstood and unappreciated
Do we really have a “visibility problem?”

Yes, and no

- We continue to play our **traditional and critical roles** as respected team members in a host of areas
- We have gotten **lots** of great press
- Enrollments in undergraduate statistics programs are **skyrocketing**, as are applications to graduate programs
Do we really have a “visibility problem?”

Yes, and no

- We continue to play our **traditional and critical roles** as respected team members in a host of areas
- We have gotten **lots** of great press
- Enrollments in undergraduate statistics programs are **skyrocketing**, as are applications to graduate programs
- But statistics continues to be omitted from much of the discourse on **Big Data** and **data science**
- And we do not have the stature we would like among many of our **fellow scientists**
- Moreover, our discipline and its value continue to be misunderstood and unappreciated
The American Community Survey

- “An ongoing statistical survey that samples a small percentage of the population every year” (about 3 million)
- **Mandatory** participation
- Used to allocate >$400 billion in federal, state funds
- Used by private industry and companies to plan, expand
- [http://www.census.gov/acs/www/](http://www.census.gov/acs/www/)
The American Community Survey

- “An ongoing statistical survey that samples a small percentage of the population every year” (about 3 million)
- **Mandatory** participation
- Used to allocate >$400 billion in federal, state funds
- Used by private industry and companies to plan, expand
- [http://www.census.gov/acs/www/](http://www.census.gov/acs/www/)

House bills to eliminate or make voluntary the ACS

“This is a program that intrudes on peoples lives... We’re spending $70 per person to fill this out. That’s just not cost effective, especially since in the end this is not a scientific survey. It’s a random survey.” – Rep. Daniel Webster (Florida), *New York Times*, May 19, 2012
The tornado that struck Moore, Okla., this week, killing at least 24, is estimated to have left more than $2 billion of damage in its path. When asked if he thought the government should take steps to mandate tornado shelters in the aftermath of the storm, (Sen. Tom) Coburn rejected the idea. “If you’re living in that area of Moore in Oklahoma, the likelihood of being hit by another tornado is about zero in terms of odds,” he said. – Huffington Post, May 23, 2013
What are we doing to address these issues?

- The American Statistical Association (ASA)
- The International Year of Statistics (Statistics2013)
- All of us
The American Statistical Association

- Founded in Boston in 1839, second oldest professional society in the US (175th anniversary in 2014!)
- Largest community of statisticians in the world
- Over 18,000 members from more than 90 countries
- Individual, corporate, institutional members
The American Statistical Association

- Founded in Boston in 1839, second oldest professional society in the US (175th anniversary in 2014!)
- Largest community of statisticians in the world
- Over 18,000 members from more than 90 countries
- Individual, corporate, institutional members

“Promoting the Practice and Profession of Statistics”
ASA strategic plan

Adopted March 2008, revised December 2012
ASA strategic plan

Adopted March 2008, revised December 2012

The ASA as “The Big Tent for Statistics”

• Membership growth
• Meetings
• Publication and information needs
• Financial status
• Organizational effectiveness and efficiency

Increasing the Visibility of the Profession

• Public awareness
• Visibility and impact in policy making
• Education

http://www.amstat.org/about/strategicplan.cfm
Enhancing visibility in policy making

- Historically, the ASA avoided entering public debate on science policy.
- In 2006, ASA president Sallie Keller launched the Science Policy Task Force, continued by president Mary Ellen Bock in 200.
- Investigate the feasibility of the ASA formally entering the science policy arena.
- Extend ASA activities in public policy, particularly science policy.
- Creation of the position of Director of Science Policy.
Steve Pierson, ASA Director of Science Policy

- Joined ASA in March 2008 as Director of Science Policy
Steve has elevated awareness of statistics in policy making

- Regular visits to Capitol Hill, NSF, the White House OSTP
- Statistical Significance series – one pagers on the role of statistics in areas of science
- *Amstat News* Science Policy column
- Policy Director’s blog, Twitter feed
- ASA letters/statements to policy makers

http://www.amstat.org/policy/index.cfm
ASA science policy and advocacy examples

- 2009, ASA Advisory Committee on Climate Change Policy
- Since 2011, Climate Science Day
- 2012, ASA Ad Hoc Committee on Forensic Science
- 2012, “Stats for Staffers” in conjunction with the Senate Office of Education and Training
- 2012, Amstat News series on Statistics for Policymakers
Why Statistics?

POPULAR MEDIA AND SCIENCE PUBLICATIONS SOUND THE DRUM: “BIG DATA” WILL DRIVE OUR FUTURE, FROM TRANSLATING GENOMIC INFORMATION INTO NEW THERAPIES, TO HARMONIZING THE WEB TO UNRAVEL COMPLEX SOCIAL INTERACTIONS, TO DETECTING INFECTIOUS DISEASE OUTBREAKS. STATISTICS IS THE SCIENCE OF LEARNING FROM DATA, AND OF MEASURING, CONTROLLING, AND COMMUNICATING UNCERTAINTY; AND IT THEREFORE PROVIDES THE NAVIGATION ESSENTIAL FOR CONTROLLING THE COURSE OF SCIENTIFIC AND SOCIETAL ADVANCES. THIS FIELD WILL BECOME EVEN MORE CRUCIAL AS ACADEMIA, BUSINESS, AND GOVERNMENTS RELY INCREASINGLY ON DATA-DRIVEN DECISIONS, EXPANDING THE DEMAND FOR STATISTICS EXPERTISE.

The melding of science and statistics has often propelled major breakthroughs. Last year’s Nobel Peace in Physics was awarded for the discovery of the accelerating expansion of the universe. That discovery was facilitated by sophisticated statistical methods, establishing that the finding was not an artifact of imprecise measurement or misjudgments. Statistical methods also allowed the trial demonstrating that aspirin reduces the risk of HIV transmission from infected pregnant women to their infants to be stopped early, benefiting countless children. Statistical principles have been the foundation for field trials that have improved agricultural quality and for the randomized clinical trial, the gold standard for comparing treatments and the backbone of the drug regulatory system.

Statistics often informs policy development. For example, in the United States, billions of dollars are allocated to school districts based on county-specific estimates of income and poverty, derived by combining data using statistical methods. In evaluating pollutants, statistical modeling isolates true associations with illnesses and deaths. Big Data payoffs can be enormous, but there are many pitfalls. Take the promise of personalized medicine: Achieving this goal will require the integration of vast landscapes of genomic, clinical, and related data from legions of patients. The potential for false discovery looms large.

New statistical methods will be needed to address some of these issues. Similar challenges arise from the haystacks of information on social network, time-use, economic, and other activities that can be mined to benefit science, business, and society. Close collaboration with statisticians is the best way to ensure that critical issues are identified and solutions found.

A dramatic increase in the number of statisticians is required to fill the nation’s needs for expertise in data science. A 2011 report by a private consulting firm projected a necessary increase of nearly 260,000 professionals (a 50% increase) by 2018.* Graduate specialization in statistics are equipped with skills that allow them to pursue diverse careers, and there has been a surge in applications for graduate education in these fields. But available places are limited; for example, the ratio of qualified applicants to slots in the Johns Hopkins Biostatistics program exceeds 10:1. Resources must be found to expand the number and size of graduate programs.

No amount of statistical intervention can circumvent flawed subject-matter models or salvage valid conclusions from poorly designed studies, and even sound statistical analysis may fail to yield straightforward answers. The future demands that scientists, policy-makers, and the public be able to interpret increasingly complex information and recognize both the benefits and pitfalls of statistical analysis. It is a good sign that the new U.S. Common Core K–12 Mathematics Standards† introduce statistics as a key component in precollege education, requiring that students be skilled in describing data, developing statistical models, making inferences, and evaluating the consequences of decisions. Embedding statistics in science and society will prove the route to a data-informed future, and statisticians must lead this charge.

—Marie Davidian and Thomas A. Louis


†www.corestandards.org.
The International Year of . . . Chemistry?

2011 was the International Year of Chemistry

This got a certain 2010 ASA president thinking.
2011 was the International Year of Chemistry
2011 was the International Year of Chemistry

This got a certain 2010 ASA president thinking...
The International Year of … Chemistry?

2011 was the International Year of Chemistry

This got a certain 2010 ASA president thinking…
The International Year of Statistics was born

Founding organizations
• ASA
• Institute of Mathematical Statistics (IMS)
• International Biometric Society (IBS)
• International Statistical Institute and Bernoulli Society (ISI)
• Royal Statistical Society (RSS)

Steering committee of nine representatives from these societies worked intensively and tirelessly

2013 was targeted – the “250th anniversary of Bayes Theorem”
The International Year of . . . Statistics!

The International Year of Statistics was born

Founding organizations

- ASA
- Institute of Mathematical Statistics (IMS)
- International Biometric Society (IBS)
- International Statistical Institute and Bernoulli Society (ISI)
- Royal Statistical Society (RSS)
- Steering committee of nine representatives from these societies worked intensively and tirelessly

2013 was targeted – the “250th anniversary of Bayes Theorem”
The International Year of Statistics

http://statistics2013.org
The International Year of Statistics

http://statistics2013.org

- Comprehensive website featuring continually updated articles, blogs, events, resources, career information
- Over 2000 participating organizations worldwide sponsoring events, awareness campaigns
To designate the year 2013 as the “International Year of Statistics”.

IN THE SENATE OF THE UNITED STATES

MAY 21, 2013

Mrs. HAGAN submitted the following resolution; which was referred to the Committee on the Judiciary

RESOLUTION

To designate the year 2013 as the “International Year of Statistics”.
Enhancing visibility among fellow scientists

The ASA president-elect develops a series of 3-4 *initiatives* based on the themes of the ASA strategic plan.
Enhancing visibility among fellow scientists

The ASA president-elect develops a series of 3-4 initiatives based on the themes of the ASA strategic plan.

“Improving the visibility of statistics within the sciences”
Enhancing visibility among fellow scientists

AAAS – A focal point for science

- **World’s largest general scientific society**, ∼120K members, 261 affiliated scientific societies (including ASA)
- Publisher of *Science*
- 24 AAAS sections, including **Section U on Statistics**
- 2012 was the **50th anniversary** of the founding of Section U
Enhancing visibility among fellow scientists

AAAS – A focal point for science
- World’s largest general scientific society, ~120K members, 261 affiliated scientific societies (including ASA)
- Publisher of *Science*
- 24 AAAS sections, including *Section U on Statistics*
- 2012 was the 50th anniversary of the founding of Section U

Workgroup of Section U officers, ASA reps to AAAS sections
- Develop and carry out initiatives to raise the profile of statistics within AAAS
Enhancing visibility among fellow scientists

- Button campaign to encourage statisticians to join AAAS (2012 JSM and beyond)
- Section U sponsorship of invited session proposals for 2013 AAAS Annual meeting (six were accepted!)
- Campaign to encourage statistics students to enter AAAS Student Poster Competition (10 selected; 2 of 4 prestigious travel awards to statistics grad students!)
- Nominations for AAAS Fellow
Enhancing visibility among fellow scientists

- Section U membership increased 14% from 2012 to 2013
- Sessions proposals for 2014 AAAS meeting
- ASA-AAAS joint symposium, articles on statistical science for AAAS MemberCentral, . . .
Enhancing visibility among fellow scientists

- Section U membership increased 14% from 2012 to 2013
- Sessions proposals for 2014 AAAS meeting
- ASA-AAAS joint symposium, articles on statistical science for AAAS MemberCentral, . . .

See my April 2013 column in the Amstat News
Public (non-) relations

- Prior to 2006, the ASA had **no** public relations effort.
- In 2006, then-Executive Director Bill Smith hired a **part-time consultant** to help ASA develop media relations.
- In 2012, ASA Board leadership and staff recognized the need for **full-time public relations assistance**.
- **International Year, 175th anniversary**.
- Board go-ahead to hire **public relations coordinator**.
Jeff Myers, ASA Public Relations Coordinator

- Over 20 years experience with US Army, Insurance Agents and Brokers of America, . . .
- Joined ASA in **July 2012** as Public Relations Coordinator
Public relations, full-time

Immediate impact

- **Press releases** on Statistics2013, the contributions of statisticians
- **Intensive outreach to media**, connecting journalists to expert statisticians
- **Media training** for ASA Board members
2013: The International Year of... Statistics

Posted: 02/12/2013 1:18 pm

It's a scene we statisticians know all too well. The passenger next to us asks, "So, what do you do?" "I'm a statistician," we reply, bracing for impact. "That was the worst course I ever took," our seatmate exclaims. "I had to take it for my major, but didn't understand any of it." It's a recurrent narrative we've experienced for decades.

But times have changed. From Google chief economist Hal Varian's well-circulated 2009 quote in *The New York Times* that "the sexy job in the next 10 years will be statisticians" to the assertions of numerous blog posts and articles, statistics is "hot."

And the dead-on predictions of the 2012 presidential election by not only statistician Nate Silver, but also Huffington Post's Mark Blumenthal and Simon Jackson, political scientist Drew Linzer, and others, have captured the attention of the public and media about the power of statistics like never before. Now, the reactions of our seatmates have transformed from groans into eager curiosity, due in large part to this watershed event.

What is statistics? To many, statistics is the class they took in college or figures on the sports pages. But statistics is so much more. Statistics is the science of learning from data and measuring, controlling, and communicating uncertainty. Statisticians do this by developing models to describe data. These models help us design methods to collect data, draw conclusions from data, and characterize the uncertainty in the findings.
Joint initiative by the ASA presidents (Bob, Marie, Nat) (See my June 2013 *Amstat News* column)

Challenge

- Big Data are a **big topic**
- Statisticians **have not** been very involved
- The statistical community is **not connected** to the data science community
Big Data

Joint initiative by the ASA presidents (Bob, Marie, Nat) (See my June 2013 *Amstat News* column)

Challenge

- Big Data are a **big topic**
- Statisticians **have not** been very involved
- The statistical community is **not connected** to the data science community
- Lack of recognition of what statisticians offer
- Lack of interest among statisticians
- Need for statisticians to develop additional skills
Objective

• Statistical principles are critical to Big Data
• Statisticians must understand new ways of thinking
• The statistical profession should be a valued player in Big Data activities
Big Data

Objective

- Statistical principles are critical to Big Data
- Statisticians must understand new ways of thinking
- The statistical profession should be a valued player in Big Data activities

Projects

- Workgroup on skills and curriculum enhancements, JSM 2014 panel and report
- Meetings between ASA representatives and leaders from business and technology companies, government, academia at the forefront of Big Data
- Training opportunities at 2014 CSP and JSM
How can *you* make a difference?

- Continue to promote the value and sound practice of statistics to collaborators
- Join the AAAS and contribute to ongoing outreach efforts to scientists
- Engage with the data science community (see my July 2013 *Amstat News* column)
- Partner with a local university, school district, or science museum to host a “Statistics Day”
- Host an outstanding speaker for a public lecture
- Engage the local media and pitch stories about statistics
- Offer to speak at elementary and secondary schools about statistics and statistical careers
- Etc. . . Think creatively!
Together, we can all play a role in enhancing the visibility and contributions of the statistics profession!