

Statistics in the Media: How accurate are they?

N. Saidikowski

Day 3, NCJ – National Convention of Journalists

Statistics in the Media: How accurate are they?

Sunday, 2:30-4:00pm

Speaker: Stacy Statis, MA Professional Writing, BS Statistics

Humans love to make predictions about the future, but to do so they need accurate information. In the Information Age, it is imperative that the people disseminating the information do so as correctly and accurately as possible. However, journalists, those often displaying information to consumers, do not always provide as accurate a description of that information as people require, resulting in the spread of misinformation. This talk discusses some basic statistical principles to help journalists offer the public more accurate renderings of statistical studies.

Sunday, 2:25pm

She had studiously prevented herself from looking out at the audience gathering for her talk. The weekend, filled to bursting with conference seminars and activities, had felt long, it was late on the last day, people were tired, she was tired, and everyone would much rather be napping, including her. But she had agreed to come to this conference and talk about statistics in the media and she was determined to do so, even in light of her fear of public speaking. Her topic – the near-constant butchering of statistical principles and statistical findings in media consumed by millions of people – was that important.

Stacy Statis finally managed to connect her laptop to the projector and allowed herself a cursory glance around the auditorium.

Her already shadowy mood dimmed.

Five people were sitting in the 200-seat auditorium.

She fleetingly wondered if perhaps the attendance, like the scheduling of her talk in the latest time slot on the last day of the conference, was indicative of a larger issue. Nonetheless, Stacy swallowed her concern and hoped that perhaps there would be a sudden influx of attendees at the last minute.

By 2:35, she had admitted defeat.

“Welcome,” she chirped, probably a little too brightly as she tried to sound cheerful and excited around her disappointment. The audience shifted listlessly.

“My name is Stacy Statis. I came here today to talk about statistics in the media and some of the ways journalists can better inform the public about topics involving statistical analyses.”

Resounding silence met her opening remarks. Then, someone moved. Stacy glanced hopefully towards the noise, but it turned out to be an older man, grumbling intelligibly. He glared at her, stood up, and stalked out of the room.

Mortified, Stacy tried to play it off.

“I guess he was in the wrong room,” she joked. Trying not to cry, she launched into her talk and tried to ignore the awkwardness.

“You might be wondering why someone who specializes in communication is qualified to talk about statistics,” she began, “or is even interested in talking about statistics. Well, my Bachelors degree is in stats, and I moved into writing and communications because my favorite part of the statistical analysis process was discussing the results. I’m fascinated by the communication of statistics and, more importantly, some of the barriers to doing so. The reason I wanted to speak at this conference is that I hope to impress upon you, the individuals helping to educate and inform

the public, why knowing about and understanding some of the underlying principles of statistical analysis are so important.”

The audience continued to look nonplussed. Only a particularly alert young woman in the first row even bothered to look up at the podium, but that may have been more a side-effect of drinking her coffee than an interest in the talk. Deciding to focus on her anyway, Stacy continued.

“The first question you might ask is, why should we care? What is so important about gathering information and proceeding to analyze it? Well, as I’m sure many of you know, humans like to make predictions about the future, and we like to use those predictions to establish future directions, like policy decisions or disaster preparedness.”

Stacy took a deep breath and then reached for her glass of water. The audience somehow looked less attentive.

“But, as Niels Bohr said, ‘Prediction is very difficult, especially about the future.’” Although she paused for a moment to allow the audience the chance to laugh at the small joke, no one stirred. She barreled on. “So, if we hope to make accurate predictions about ourselves and the world around us in the near and distant future, we need to work from existing data. That is, we need to gather pertinent information from now to as far back as resources allow, and use that information to find patterns and make inferences which we hope can be extrapolated to the future. Otherwise, our predictions are poorly founded and potentially inaccurate, which lead to poor policy decisions, lack of preparedness, and, more generally, the spread of misinformation.”

Her audience perked up a little as she spoke of the consequences. Stacy had a sudden inspiration.

“For example, I sometimes troll mediamatters.org,” Stacy confessed, deviating from her carefully written speech and launching into unknown territory, “which might sound terrible-” The young woman in the front row giggled a little and Stacy felt a slight rush of energy. “-but I think some of the mistakes pointed out on that site are ones perpetuated throughout the media and they don’t need to be.”

The other three members of her audience stirred. A smartly-dressed middle-aged woman in the third row sat up straighter. Two men in khakis and blazers sitting towards the back shifted, leaning forward slightly. Stacy took a deep breath and gathered her thoughts.

“In fact, even [mediamatters](http://mediamatters.org) isn’t completely error-free. I found an article about Hillary Clinton and her supposed fading popularity.” Hearing the name of one of the frontrunners in the upcoming 2016 election made the entire audience lean forward as a single entity. Stacy had found her hook.

“Hannah Groch-Begley, the author, made the claim that although many figures in the media argued that Hillary Clinton was losing popularity and that Americans were tired of talking about her, actually quite the opposite was true. Groch-Begley then discussed several articles and polls that showed Clinton was still favored among Americans...” Stacy faded out for dramatic effect, reaching for her glass of water and taking a sip.

Chairs creaked as the audience shifted in anticipation.

“But here’s the thing,” Stacy said, her tone almost conspiratorial, “while the studies Groch-Begley cited technically supported her claim, she buried the actual intent of each one. First, she backed up the claim that ‘a majority of Americans’ still viewed Clinton favorably by citing a Gallup Poll from that summer which attributed to Clinton a 54% favorability rating. However, that poll and the resulting article actually showed that Clinton’s favorability was declining, and had been since February of that year when she left her position as Secretary of State. In fact, previous research had shown that non-political figures are generally considered more favorably than political ones. So, the important story, the nuanced argument supported by the data about Clinton’s declining popularity, was ignored so Groch-Begley could make the point that despite claims by the press, Clinton is still popular. Beautiful irony, if you ask me.” Stacy smiled at the audience as, this time, she was rewarded with chuckles.

“Keep in mind that in the summer of 2014, Clinton hadn’t yet announced whether she was running for president or not. So, in light of the research that suggests Americans view political figures less favorably than non-political ones, we might be led to extrapolate that, far from Clinton staying popular with Americans, her popularity might continue to decline, especially after she made the announcement.”

Stacy paused to let the information sink in, then slipped in a quick statistical lesson. “Extrapolation, though, can be tenuous. The data might suggest a pattern in a certain direction, but that pattern only continues if all exogenous factors remain constant. Any major or seemingly minor changes could affect the pattern, so qualifying predictions with that caveat is actually a pretty solid way to protect yourself from completely losing credibility down the line if your prediction fails.”

By this point, the young woman in the front row had yanked out a small notepad and was scribbling furiously on it with one of the free NCJ pens.

With a small smile, Stacy continued, the rising excitement in her voice directly correlated with the increased attentiveness of her audience. “Gallup is great because they include the methodology and some of the raw numbers from their data collection. Generally, the researchers at Gallup use phone surveys to gather data for their analyses, which has been working relatively well for them, but may face some issues in the future. You may or may not have heard about some of the concerns arising in the statistical community concerning data collection, but survey statisticians are finding it difficult to acquire random and representative samples. This difficulty is traceable back to the propagation of cell phones.”

A small gasp from the audience alerted Stacy to the piqued interest of those assembled.

“The popularity of telephone surveys arose in the seventies, probably because conducting a survey via telephone is more feasible than conducting one face-to-face. Pollsters have historically relied on computer-assisted methods of calling individuals for these surveys because the computer could access a large database of landline telephone numbers. That access allowed statisticians to establish random cross-sectional samples that were relatively representative of the population, even though they did face some problems of coverage since not all families owned a landline and those families that didn’t differed demographically from those that did. That problem of differing demographics between accessible and inaccessible samples makes it imperative for survey statisticians to figure out ways to account for the issue of non-coverage, or failure to collect data on certain subgroups of the population.”

Taking a deep breath, Stacy took a risk. “Would anyone like to hazard a guess as to why non-coverage might be an issue?”

The young woman in the front stared intently ahead for a few seconds. Stacy forced herself to wait, to allow the audience time to think it over. Then, the young woman raised her hand.

When Stacy nodded at her, she said, “Because then the sample won’t reflect the opinions of everyone.”

“Exactly,” Stacy affirmed. “In statistical terms, non-coverage results in a sample that isn’t representative of the population, which leads to results that aren’t necessarily applicable to the whole population. Basically, if we can’t be entirely sure that the sample gathered by Gallup is representative of Americans, we have reason to doubt the claims made. We might doubt that Clinton’s favorability is actually declining, we might wonder if perhaps Groch-Begley is correct when she claimed that Clinton’s popularity was as strong as ever. We just can’t know if we don’t have reliable data.”

To Stacy’s great surprise, the young woman in the front spoke again. “So what do we do?”

Grinning broadly at the participant, Stacy explained. “Well, that’s The Question. Statisticians are still trying to figure out ways to ensure samples are representative. Fortunately, statisticians have found a few methods that work, and with further development, will work better.”

After a quick pause to allow the audience a chance to breathe, Stacy continued. “The development of the internet offered a new method of data collection that is extremely desirable

because of the feasibility of implementation. Unfortunately, internet surveys come fraught with error potential because they tend to maximize the problem of non-coverage that we discussed earlier. Because everyone does not have equal access to the internet and those that do differ from those that don't, samples gathered this way tend to be limited in representation. This is to say nothing of non-response or response bias, which are both quite high for internet surveys, further exacerbating the problem of failing to represent the population--"

"I'm sorry," the woman in the third row interrupted. Stacy jolted to a halt more out of surprise that someone else was speaking than the interjection itself.

"What are non-response and response bias?"

"Oh!" Stacy gasped. "I'm sorry. Non-response refers to failure of some sampled individuals to respond to the survey. Unfortunately, we can rarely assume that those who don't respond differ in no way from those who do. Response bias, on the other hand, is a problem that arises when only certain subgroups choose to respond to a survey, generally out of extreme dislike or extreme passion for the topic. Thus, both of these issues lead to a lack of representation." She looked questioningly at the woman.

"Ah, okay. Thank you."

"Of course. I apologize, I get caught up in talking about statistics so easily." She smiled. "So, we have several powerful tools for collecting data, but on their own, none of them work as well as we'd like. Fortunately, there is a way to fix that. In 2005, Edith de Leeuw published an article on mixed-modes of data collection, or using multiple methods to collect data."

Stacy paused, partly for dramatic effect, partly to take another drink of water. Her audience still looked engaged and interested, so she happily talked on.

"For example, researchers can combine telephone surveys with internet surveys or face-to-face surveys in an effort to maximize representation. The method Gallup used in the poll I talked about earlier--" Stacy surveyed the room, checking for comprehension. "On Clinton's favorability ratings," she added. The sudden understanding of her audience was obvious, so she continued.

"Well, the pollsters built a sample made of 50% landline respondents and 50% cellphone respondents. Mixing modes like that maximizes the advantages of each mode, or method of collection, and helps to account for some of the weaknesses of each. So, doing so helped Gallup account for those individuals who don't have a landline anymore or only use their cellphone." Her audience nodded along as she spoke.

"Further research is still needed to ascertain how detrimental some of the drawbacks of mixed-modes are." The young woman in the front, still scribbling, looked up in dismay. Stacy gave her a small smile to indicate that all was not lost.

"One such drawback is the introduction of new sources of error into the analysis. For example, sometimes the question arises of whether a significant result is attributable to a factor of interest or whether it actually arose because of the mixed-mode of data collection. Statisticians have found some safeguards to implement which can help reduce the frequency of these errors, but, like I said, more research is still needed before this method of mixing modes of data collection can become commonplace."

Having hit her stride, Stacy confidently continued through the rest of her talk, citing more examples from the article on Hillary Clinton's popularity and discussing some of the more nuanced explanations in the de Leeuw article. Even though she had to make up half of her speech on the fly, her audience remained engaged and interested in the topic.

By the end of the talk, Stacy had a beautiful fantasy in her mind of these four people going out into the world and spreading the information she had given them here, in this room, to every journalist they knew. She entertained the hope that someday widespread understanding of statistics would result in a more informed public, making decisions based on credible findings and questioning studies that sounded dubious.

She was certain crazier fantasies had come true.

Further reading:

1. de Leeuw, Edith D. (2005). To Mix or Not to Mix Data Collection Modes in Surveys. *Journal of Official Statistics* 21(2), 233-255.
2. Groch-Begley, Hannah. Clinton Fatigue Syndrome: What The Press Gets Wrong About Hillary's Popularity. *Media Matters for America*. Retrieved from <http://mediamatters.org/blog/2014/09/19/clinton-fatigue-syndrome-what-the-press-gets-wr/200828>.
3. Krosnick, Jon A. (1995). Survey Research. *Annual Review of Psychology*, 50, 537-567.
4. Tetlock, Philip E. (2006). *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton, NJ: Princeton University Press