therapeutics.

**Statistical Thinking Comment Policy**

Comments are welcomed. Be informed, informative, respectful, and criticize ideas, not people. Please read our Comment Policy before commenting.

37 Comments   Statistical Thinking   🔒 Disqus' Privacy Policy                     ⬤ Frank Harrell ▾

♡ Favorite          🐦 Tweet          f Share                                    Sort by Best ▾

**Join the discussion…**

**Frank Harrell** • 6 years ago
See Section 10.7.4 in Biostatistics for Biomedical Research available from
http://biostat.mc.vanderbil...
⌃ ⌄ • Reply • Share ›

**tibu** • 6 years ago
can you explain the concept of "degree of freedom"?
⌃ ⌄ • Reply • Share ›

**Frank Harrell** • 6 years ago
I'm of two minds about this. (1) CIs are almost impossible for non-statisticians to perfectly understand because of their indirectness. They a lot of the same problems as p-values. (2) they are better than p-values and can be interpreted not matter how large the p-value; so they are a bridge to better methods with likelihood and Bayes.
⌃ ⌄ • Reply • Share ›

**tgs** • 6 years ago
Suggested topic: If I give up p-values, do I also have to give up confidence intervals? Is an "estimation and uncertainty" approach also inherently flawed, even if there are no p-values?

(We've had a brief discussion about this before, but I'm still fuzzy on the issue.)
⌃ ⌄ • Reply • Share ›

**Frank Harrell** • 6 years ago
Good question. I view the bootstrap as an amazingly versatile frequentist and exploratory tool, especially for demonstrating variation (e.g., in feature selection) and getting confidence intervals of strange things. It is not the ultimate answer, e.g., if you sample from a log-normal distribution the accuracy of all bootstrap confidence limits for the population mean is terrible. I seek exact answers much of the time, leading me to Bayes.
⌃ ⌄ • Reply • Share ›

**Frank Harrell** • 6 years ago
I am learning Stan, thanks to Chris Fonnesbeck and John Kruschke. In the past I've used JAGS a good deal, and liked it.
⌃ ⌄ • Reply • Share ›

**Moritz Körber** • 6 years ago

Thank you very much. Last question: What scripts do you use for Bayesian parameter estimation? I am currently relying on John Kruschke's scripts (+ JAGS).

∧ | ∨ • Reply • Share ›

**Leon** • 6 years ago

I'd be interested in how you think about the bootstrap, given your distaste for "backwards" probabilities. The bootstrap seems to me entirely backwards, since it's all about variation in data-space rather than unknown-space.

∧ | ∨ • Reply • Share ›

**Leon** • 6 years ago

This comment has been removed by the author.

∧ | ∨ • Reply • Share ›

**Frank Harrell** • 6 years ago

Yes that is one good approach. The beauty of Bayesian sample size estimation is that in incorporates uncertainty yet doesn't depend on unknown parameter values. Some good papers are here: http://www.citeulike.org/se...

∧ | ∨ • Reply • Share ›

**Moritz Körber** • 6 years ago

Thank you for clarifying your position and I definitely agree. Could you also elaborate how you do a kind of Bayesian power analysis? How can you estimate how many participants/data you need until your posterior is precise enough?

∧ | ∨ • Reply • Share ›

**Frank Harrell** • 6 years ago

It's best to spell out the steps you envision, and best to avoid hypothesis tests altogether in favor of gathering evidence for a positive effect. Bayesian inference can be misleading at early looks only if your prior specifies a high chance of obtaining a large treatment effect. The prior perfectly calibrates early looks and shrinks the posterior mean by a perfect amount depending on the information content of data collected so far.

∧ | ∨ • Reply • Share ›

**Moritz Körber** • 6 years ago

"There are no errors to control in Bayesian posterior inference"

Do you mean in parameter estimation? What if I test hypothesis, then I can falsely reject/accept a hypothesis. The BF can mislead in the beginning of data collection as well. As you see, even if there are no errors, it seems (at least to me) worth a discussion of that fact! I would appreciate.

∧ | ∨ • Reply • Share ›

**Frank Harrell** • 6 years ago

P.S. Bayesian modeling seems very natural to me and frequentist inference seems like voodoo. This is especially true when you get into sequential testing, multiplicity, closed testing procedures, group sequential methods, sample size re-estimation, ...

∧ | ∨ • Reply • Share ›

**Frank Harrell** • 6 years ago

I'm interested in 1, 2, 5, 7, 8. I'm almost finished covering what I wanted to cover for 7 and 8 after two upcoming posts.

∧ | ∨ • Reply • Share ›

**philipppa** • 6 years ago

1. Is machine learning statistics or computer science?
2. Methods for sample size calculation (art or science)?
3. Statistical Inference
4. Is logical statistical reasoning reductio ad aburdum? Does it thus always and unavoidably have an agenda?
5. The difference between Bayesian and Frequentist statistics (to be honest, from what I understand Frquentist seems plausible while Bayesian is just vodoo ...)
6. Analysis of contigency tables
7. The p value (and its criticisms)
8. The NHST framework
9. Scientific and statistical misconduct and fraud of the pharmaceutical/nutritional/fitness industry

∧ | ∨ • Reply • Share ›

**Frank Harrell** • 6 years ago

I'll let others address noncentral distributions. There are not errors to control in Bayesian posterior inference. We might concentrate more on robustness, mainly by talking about Bayesian nonparametric and semiparametric models. Thanks for input.

∧ | ∨ • Reply • Share ›

**Moritz Körber** • 6 years ago

1. Error control in Bayesian statistics

2. Robustness of Bayesian statistics: What to do if the assumption of the models are not met?

3. non-central distributions

∧ | ∨ • Reply • Share ›

**Shravan Vasishth** • 6 years ago

Very interesting looking book, thanks for the references. I bought ebook version.

∧ | ∨ • Reply • Share ›

**Brian Byrd** • 6 years ago

Terrific! If you happen to find those papers interesting, I can highly recommend Dr. Glass's book, Experimental Design for Biologists, Cold Spring Harbor Press. It's expansive in scope, spanning philosophy of science to how a Western blot should be designed to considerations relevant to clinical trials. It doesn't dwell on any one topic terribly long, but it's packed with interesting ideas, as you can see from the Table of Contents: http://www.cshlpress.org/de... One other point I would make: his checklist of experimental design is extremely good & is worth the price of the book, even if it's only a page or so.

∧ | ∨ • Reply • Share ›

**Frank Harrell** • 6 years ago

WOW I've been looking for just this type of paper for years. From the title I think I'm going to really like it. I really don't like straw man hypotheses and frequently tell investigators to state a question or sometimes better, state the quantity you want to estimate (often an effect such as a treatment difference). I'll read that paper as soon as I can. Thanks!

∧ | ∨ • Reply • Share ›

**Frank Harrell** • 6 years ago

I would say that a p-value is a probability, and I think of a proportion as something having a finite denominator.

∧ | ∨ • Reply • Share ›

**Brian Byrd** • 6 years ago

Dr. Harrell, I'm delighted that you've started this blog. I would be very interested in your view of David Glass's critique of the hypothesis as a framework for experimentation (as opposed to "the question"). Dr. Glass teaches experimental design at Harvard & his views surprised me & got me thinking. I've included two sources (a Cell paper & a Clinical Chemistry paper) in case you haven't run across his perspective. Thanks in advance for considering this topic for your blog. Go 'Dores. lhahttp://www.sciencedirect... http://clinchem.aaccjnls.or...

∧ | ∨ • Reply • Share ›

**Bill R** • 6 years ago

I was unclear. Exact p-values are proportions (proportion of possible outcomes as large or larger given the current data and some model.) If I understand your meaning, it is a "backward" probability, using the classical definition of probability,

∧ | ∨ • Reply • Share ›

**Frank Harrell** • 6 years ago

p-values are not proportions. But I like your definition of proportions.

∧ | ∨ • Reply • Share ›

**Bill R** • 6 years ago

Forward and Backward probabilities, I like that. My personal view is similar: "Probability is the future tense of a proportion." P-values are proportions and calling them probabilities just confuses everything.

∧ | ∨ • Reply • Share ›

**ramirob** • 6 years ago

Hello, I would be interested in posts about how to "do things right". For example, an analysis of a dataset done properly, with some of the subtleties and nuances explained. In addition, I haven't heard much about design of experiments geared towards a bayesian approach.

∧ | ∨ • Reply • Share ›

**Frank Harrell** • 6 years ago

I don't think the overall approach needs to be changed in that setting, it's just that in Phase II studies we are more open to adaptation/multiple hypotheses, and there is an even greater need for a Bayesian approach because when the effective sample size is not large, it may be necessary to incorporate outside information and this cannot be don't in the traditional frequentist paradigm.

∧ | ∨ • Reply • Share ›

**Roy Tamura** • 6 years ago

Yes, I hope you will be able to give us some insight in this area without releasing proprietary information. In many situations I deal with now, the traditional paradigm of a randomized trial with high power for a modest treatment effect just isn't feasible and I struggle with what alternative approaches are appropriate.

∧ | ∨ • Reply • Share ›

**Frank Harrell** • 6 years ago

I can't disagree with any of that. I'm spending a lot of time discussing p-values and NHST in an attempt to show the emperor has no cloths and we need to change not just for the sake of change but to have better solutions with clearer interpretations. You're right about the perception of 'error control' driving many choices of statistical approaches, in industry, academia (e.g., NIH-funded research), and regulatory. Few people think about whether the false positive referred to in type I error is really an error. I'll be writing more about advantages of direct forward probabilities because in my opinion what really needs to be known is the probability that the conclusion you are about to make is true, not the probability of getting data more extreme than the current data if the effect happens to be exactly zero. The probability of being wrong about efficacy is quite simply one minus the posterior probability of efficacy given the data. In the future I'd like to expand on what type I error means.

∧ | ∨ • Reply • Share ›

**Daniel Lakens** • 6 years ago

I've long wondered why many statisticians embrace Bayesian statistics (and many have done so for decades), but the FDA does not (yet, as I understand it) fully Bayesian inferences. The FDA seems to value error control. Maybe this is related to point 5? Showing concrete examples of how Bayesian statistics would improve cumulative science would also be a strong argument towards adopting those practices (preferably illustrated with real lines of research). I think most of your readers are not novices at statistics - criticism on NHST is available in dozens of articles. I would personally feel that writing about NHST is a bit of a waste of time (I'm sure you can point out the most important issues as sidenotes in other blogs). Best, Daniel

∧ | ∨ • Reply • Share ›

**Frank Harrell** • 6 years ago

That's not my area, but one I may have to learn about in my work for FDA. I think that analysis that is only exploratory has its own set of problems, and that some inference is needed. Bayesian methods are being used more and more in Phase 2 studies, and allows adaptation and intensive sequential testing to obtain new results. And one could use the same inferential methods as for Phase III studies but just relax the criteria a bit.

∧ | ∨ • Reply • Share ›

**Frank Harrell** • 6 years ago

Excellent points. I think we can teach a lot of useful things without the need for the student knowing calculus (algebra is another matter). Bayesian modeling moves us more to careful model specification, and because simulation methods are often used to get solutions, away from calculus. We should capitalize on that. Yes, the frequentist method gives some simplification hence its heavy use and ease of programming.

∧ | ∨ • Reply • Share ›

**Frank Harrell** • 6 years ago

An excellent question. Don Berry's intro Bayesian textbook has a wonderful introduction to descriptive data analysis. You can spend a lot of good time teaching students how to describe and explore data without getting into inference. But once inference is introduced we have some tough decision to make! More people, such as Tim Hesterberg, are rightfully pushing the bootstrap as a substitute for classical inference at this level. I would also seek a way to introduce Bayesian analysis. Software is starting to help.

∧ | ∨ • Reply • Share ›

**Roy Tamura** • 6 years ago

Frank:

I am glad to read this new blog. Concerning future topics: I would be interested in your views on clinical trial design. Perhaps not so much on large Phase 3 type trials but more on the role of statistics in smaller, more exploratory Phase 2 trials. How can we get away from hypothesis testing of the null in these situations? How should we design and analyze trials for rare diseases?

Roy Tamura
University of South Florida

Reply • Share ›

**therandomtexan** • 6 years ago

In your initial post, you identified a major problem in our scientific culture

"Statistics has been and continues to be taught in a traditional way, leading to statisticians believing that our historical approach to estimation, prediction, and inference was good enough."

It's worse than that. Traditional statistics is what we're teaching our undergraduates in business and the sciences, so we're perpetuating ideas that were already threadbare decades ago (NOW we're worried about p-values? After a hundred years?)

Much of the appeal of the Fisher frequentist methods is that they can be applied by anyone competent in basic algebra, and can be taught as rote formulas, requiring only a simple calculator and a few tables of critical values. And colleges have seized upon this as a way to promote "quantitative literacy," feeding cargo-cult statistics to the math-averse undergraduate masses.

Teachers of statistics need ways to starting changing the direction of statistics education towards modern techniques. But it can't start with requiring everyone to complete a year of calculus first (even though I'd make it a requirement to graduate, or even to vote).

Mike Anderson
University of Texas at San Antonio

Reply • Share ›

**Kevin** • 6 years ago

I have read enough to know the pitfalls of using null hypoth testing. But as a teacher of stats in HS, the texts are focused on this process for inference. So is the AP exam the students take.

My question is....what would you do as a teacher in my position?

Thanks.

Reply • Share ›

## Related

- [Scoring Multiple Variables, Too Many Variables and Too Few Observations: Data Reduction](#)
- [Statistical Criticism is Easy; I Need to Remember That Real People are Involved](#)
- [Continuous Learning from Data: No Multiplicities from Computing and Using Bayesian Posterior Probabilities as Often as Desired](#)
- [Bayesian vs. Frequentist Statements About Treatment Efficacy](#)
- [Integrating Audio, Video, and Discussion Boards with Course Notes](#)