



## What Makes A Researcher Cringe

Not that anyone on our mailing list would ever be guilty, but here are a few lines we would rather not hear again:

- "We've never done any research before, but . . ."
- "Is this survey statistically significant?"
- "Could you write up a questionnaire, and then we'll try to get your proposal approved?"
- "We want you to run Spearman correlations on everything." [This actually happened. Spearman's *rho* is a seldom-used tool for rank-order correlations and was totally inappropriate to the data, but this ex-client just had to have them.]
- "We've decided we can't do this concept test after all because somebody might find out what we are doing."
- "Can't you just ask consumers what it would take to get them to buy our product?"
- "Should qualitative be done before or after quantitative research?" (Uh, what's our objective?)
- "Sometimes there are two spaces after a period and sometimes only one in these verbatims you sent us. Can't you be consistent?" [Another true story].
- "This is how our CEO likes to have questions asked."

Most professions have similar complaints. Among musicians the dreaded words are "We don't have any money to pay you, but you could get some exposure if you played for us!" To which a veteran will reply "You can die from exposure."



## Away from the Office

Paul Riedesel's annual fall holiday from business will occur the week of September 27. We are already scheduling work around this week. If you are going to need us then, please get in touch sooner rather than later.

---

## Statistical Wisdom (Things We Wish We Had Thought Of)

- All models are wrong, but some are useful. (George E. P. Box)
  - Statistics are like bikinis. What they reveal is suggestive, but what they conceal is vital. (Aaron Levenstein)
  - Conducting data analysis is like drinking a fine wine. It is important to swirl and sniff the wine, to unpack the complex bouquet and to appreciate the experience. Gulping the wine doesn't work. (Daniel B. Wright)
  - Statistical thinking will one day be as necessary a qualification for efficient citizenship as the ability to read and write. (H.G. Wells)
- 

## The Price of Green

Without mentioning names, we recently observed a curious pattern for an eco-friendly concept. These particular products cost substantially more than conventional ones, but many consumers actually preferred the highest-price option. The interpretation that made sense to us was that they did not want to compromise. The lower-priced green products made some performance trade-offs vs. conventional ones. Those opting for green products know they will be paying more, but did not want to pay an additional price in diminished performance. Your mileage could vary.

## Mr. Bayes' Neighborhood



We commonly need to compare two percentages, say from a pre/post test or between two subsamples. Normal significance testing will declare that the difference either does or does not meet some standard. Not that many end clients actually understand the nuances of alpha levels and null hypotheses, but the language is at least familiar. Significant is good; not significant is at least disappointing.

Unfortunately, standard significance testing often leaves us and our clients a bit hungry. Unlike grenades or horseshoes, "close" doesn't count in significance testing. Yet we are not always blessed with large samples that would permit the detection of small but real differences. (Many do resort to describing a difference as "directional", the translation of which is "A difference that Marketing wishes were real").

But there is an alternative perspective. So-called Bayesian methods are increasingly finding their way into marketing research (they are already essential to modern machine learning and email virus filtering). For our purposes, a key and very intuitive feature is that the actual probability of an hypothesis being true is provided.

With some help from our friend and colleague, Dr. Peter Lenk of the University of Michigan, we have developed software that will enable us to report precisely on the probability of an hypothesis (specifically that  $p_2 > p_1$ ) being true.

As a hypothetical but realistic example, assume we have done a pre/post tracking survey with samples of 300 in each wave. In the base wave, 19% volunteered our brand on an unaided basis. In the second wave, 22% did so. This difference falls well short of being "significant." But gee, awareness was higher in the second wave. It may be just noise, but aren't the odds still pretty good that awareness really is higher now?

Plugging these numbers into the Bayesian algorithm, we quickly determine that the probability that true awareness is now higher is 0.82. How hard can that be to explain?

We have also programmed a parallel procedure for comparing two means (typically done with a "t-test"). Say two brands earned means of 1.78 and 1.90 on our valence scale (range of -5 to +5) with sample sizes under 200 each and normal variances. The difference is not significant by "t" but the Bayes method shows a probability of .68 that the latter score is higher. That is still a helpful finding and less ambiguous than "directional."

Your mother taught you that if something seems too good to be true, it probably is. Is there a catch? Are we somehow cheating by doing a Bayesian test where we are likelier to get a desirable answer?

Yes, one can "cheat" with Bayes by only running tests on large observed differences and ignoring small ones. But cherry picking is also an offense with conventional significance tests. No less offensive is wholesale testing where you run dozens of tests "to see what pops." By chance alone, something almost certainly will. We are still obliged to have pre-formed hypotheses in mind, even before a study is fielded.

We are very pleased to have added these Bayesian procedures to our toolbox.

