BELIEF-SUSTAINING INFERENCE*

BY ALEX REINHART AND JERZY WIECZOREK

Department of Statistics, Carnegie Mellon University

Two major paradigms dominate modern statistics: frequentist inference, which uses a likelihood function to objectively draw inferences about the data; and Bayesian methods, which combine the likelihood function with a prior distribution representing the user's personal beliefs. Besides myriad philosophical disputes, neither method accurately describes how ordinary humans make inferences about data. Personal beliefs clearly color decision-making, contrary to the prescription of frequentism, but many closely-held beliefs do not meet the strict coherence requirements of Bayesian inference. To remedy this problem, we propose belief-sustaining (BS) inference, which makes no use of the data whatsoever, in order to satisfy what we call "the principle of least embarrassment." This is a much more accurate description of human behavior. We believe this method should replace Bayesian and frequentist inference for economic and public health reasons.

1. Introduction. Modern statistics is at a crossroads. Frequentist inference, the original foundation of statistical inference, is under attack from many angles due to the low quality of work making use of it (e.g. Ioannidis, 2005, 2008) and its perceived philosophical paradoxes (Meehl, 1967). Using the likelihood function, frequentist inference attempts to infer parameters from the data objectively and with no reference to personal beliefs or subjectivity, making it very scientifically appealing but practically error-prone.

On the other hand, Bayesian inference is presented as a comprehensive system for the updating of personal beliefs on the basis of data. A rational Bayesian holds a coherent system of beliefs and systematically updates them as new data arrives. Computational difficulties rendered this method impractical until computers became sufficiently powerful, and it is now a hot area of research in statistics despite controversy about the appropriateness of subjectivity in science.

However, we believe that neither paradigm accurately represents how humans reason about their beliefs. As demonstrated by Kahneman, Slovic and Tversky (1982), most people do not hold coherent beliefs or update them as a Bayesian should, and the mere existence of subjective prior beliefs rules out frequentism as an accurate model.

*The authors acknowledge the funding and support of the Department of Statistics. The views expressed are those of the authors and will probably upset the Department when they find out they paid for this research.

MSC 2010 subject classifications: 62A01

Keywords and phrases: uninformative likelihood, embarrassment, cognitive dissonance, minimax

We believe both approaches are misguided. Rather than attempting to minimize error, as in frequentism, or attempting to maintain a coherent set of subjective beliefs, most people attempt to *minimize embarrassment*.

This insight leads to a new class of estimators we refer to as *belief-sustaining inference*, or BS inference, which we shall discuss in the following sections. An important discovery is that embarrassment is minimized by ignoring the data altogether. We also show that this approach has important public health benefits.

2. Quantifying embarrassment. The principle of minimum embarrassment may be explained in terms of the theory of cognitive dissonance (Festinger, Riecken and Schachter, 1956). The arrival of new data causes a conflict in the mind of the scientist: he would like to believe he is a rational, intelligent person who holds correct prior beliefs, but the data suggests he is wrong. This dissonance causes psychological distress and can only be resolved by jettisoning one of the contradictory beliefs, such as one's self-esteem. Overwhelming or unimpeachable evidence can thus cause severe embarrassment and psychological breakdown.

To minimize embarrassment, it is first necessary to define the mathematical concept of embarrassment in terms of the change in a personal prior after data is collected.

DEFINITION 2.1. *Let X be a random variable distributed according to some distribution function $f(x;\theta)$, where θ is an unknown parameter, $\pi(\theta)$ a personal belief about that parameter, and $\pi(\theta|x)$ the estimate based on the data. The* ***embarrassment*** *E is the distance between $\pi(\theta)$ and $\pi(\theta|x)$, as measured by the Kullback-Leibler divergence:*

$$E = \int \log\left(\frac{\pi(\theta)}{\pi(\theta|x)}\right) \pi(\theta)\, d\theta. \tag{1}$$

In Bayesian inference, the data (in the form of the likelihood $p(x|\theta)$) is combined with the prior $\pi(\theta)$ to produce a new best estimate of the parameter, using Bayes' famous theorem:

$$\pi(\theta|x) = \frac{p(x|\theta)\,\pi(\theta)}{\int p(x|\theta)\,\pi(\theta)\,d\theta}. \tag{2}$$

Many approaches are taken to choose the appropriate prior distribution, and a great deal of literature deals with the elicitation of priors from subject-matter experts. Other work attempts to eliminate subjectivity by choosing an "uninformative" or "flat" prior (*e.g.* a constant function) which places no special importance on any specific value of the parameter, letting the data decide instead.

In frequentist inference, the posterior estimate does not depend on $\pi(\theta)$, and can be a point or interval estimate solely based on $p(x|\theta)$ (*e.g.* the value of θ which maximizes $p(x|\theta)$). The estimate usually has an asymptotic normal distribution. But if we wish to minimize embarrassment, both approaches are wrong-headed, as demonstrated by the following theorem.

THEOREM 2.2. *In Bayesian inference, the embarrassment E is minimized by choosing an* uninformative likelihood*, also known as a* flat *likelihood.*

PROOF. Let $p(x|\theta) = 1$. ($p(x|\theta)$ may differ on zero-measure sets in θ, with respect to the Lebesgue measure.) Using eq. (2) and the fact that the probability density $\pi(\theta)$ integrates to 1, we find that $\pi(\theta|x) = \pi(\theta)$.

Substituting into eq. (1), we determine that

$$E = \int \log\left(\frac{\pi(\theta)}{\pi(\theta)}\right)\pi(\theta)\,d\theta = 0. \tag{3}$$

The Kullback-Leibler divergence is always nonnegative, so this embarrassment is minimal. We leave the proof that this solution is unique as an exercise for the reader. □

That is, we should not allow the data to place special importance on any specific value of the parameter, as the data will not feel obligated to support the value most beneficial to the scientist. It may even prove embarrassing, and this must not be allowed. Belief-sustaining inference requires that we ignore the data instead of taking this risk.

(Some readers may prefer to think of this in the minimax framework. Instead of aiming for the minimax risk, we aim for the minimax embarrassment. The maximum embarrassment would be a final parameter value entirely contrary to the prior, and this embarrassment is minimized by never allowing the parameter estimates to deviate from the prior.)

3. Belief-sustaining inference. Ordinarily, we would use this section to discuss the benefits of belief-sustaining inference and properties of the data-free estimator. However, any attempt to derive these properties could prove embarrassing.

Nonetheless, we will point out several important features of belief-sustaining inference. Because the belief-sustaining estimator has zero variance, it gives narrower confidence intervals than any other technique, and similarly is maximally efficient. Statistical power calculations, usually so complex as to merit entire textbooks on the subject, are made simple: the most cost-effective number of samples is always zero. Computer clusters previously wasted obtaining Bayesian Markov Chain Monte Carlo estimates can be put to more productive uses, for BS inference is computationally efficient.

We must also recognize an important public health benefit of the method of least embarrassment. Bayesian and frequentist inference would have us constantly change our beliefs, subjecting us to cognitive dissonance and causing a great deal of stress. This stress can lead to heart attacks, strokes, and other unpleasant outcomes: according to the American Institute of Stress, stress costs Americans over $300 billion annually in medical, legal and productivity costs. Hence frequentist and Bayesian inference impose a $1,000 per person tax on Americans who are already economically struggling, while the adoption of embarrassment-free inference would give an immediate 2% boost to the recovering American economy.

Considerable evidence thus suggests that evidence should be ignored altogether.

4. Conclusions. We have demonstrated that rational actors following the principle of least embarrassment will rightly ignore new data, for fear it might cause an embarrassing change of position. Ample sociological evidence demonstrates the accuracy of this model; for a particularly high-profile series of experiments, watch any Presidential debate or political talk show. The foundations of Bayesian and frequentist inference are hence falsified, and public health concerns suggest they should be banned entirely.

This work has impact on many questions of current interest. For example, scientific publishing is currently under attack by open-access advocates who would have us make publicly-funded research freely available to anyone who wishes to read it, regardless of the risk of embarrassment and cognitive dissonance to the unsophisticated reader. While professional scientists have spent years developing defensive mechanisms to protect themselves from embarrassing results, the unsuspecting reader might be unintentionally exposed to an idea contradictory to their naïve and unscientific beliefs. Our developments in embarrassment theory clearly demonstrate the foolishness of such proposals.

Additionally, BS inference has important applications in many fields of study. Economists, for instance, will be heartened by the requirement to never test theories against empirical data. Previously computationally-intractable problems in other fields are rendered trivial.

Further research is merited on several questions. For example, is it possible for the weight of evidence to be so strong that *not* changing one's opinion is *more* embarrassing? Researchers are encouraged to send us their results, although we will of course ignore them.

ACKNOWLEDGMENTS

We thank the anonymous reviewers for their undoubtedly insightful comments, which of course we did not read.

REFERENCES

FESTINGER, L., RIECKEN, H. and SCHACHTER, S. (1956). *When Prophecy Fails: A Social and Psychological Study of a Modern Group That Predicted the Destruction of the World*. Harper-Torchbooks.

IOANNIDIS, J. P. A. (2005). Why Most Published Research Findings Are False. *PLOS Medicine* **2** e124.

IOANNIDIS, J. P. A. (2008). Why Most Discovered True Associations Are Inflated. *Epidemiology* **19** 640–648.

KAHNEMAN, D., SLOVIC, P. and TVERSKY, A. (1982). *Judgment under Uncertainty: Heuristics and Biases*. Cambridge University Press.

MEEHL, P. E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science* **34** 103–115.

E-MAIL: areinhar@stat.cmu.edu