Uniform priors can be as silly as making the elementary mistake that the probabilities for any binary trial are 0.5. Would anyone expect that in a random sample of two individuals from California that one would be infected with HIV and the other would not? Or expect that the sun would rise one day and not the next?

In addition, uniform priors on the real line, or even normal distributions with mean zero and huge variances, say very silly things, for example, that the parameters are far from zero with high probability. In a logistic regression, such priors induce prior probabilities for events that are concentrated about 50% near zero and about 50% near one (Seaman, Seaman, and Stamey 2012). Pretty stupid. The only advantage to such priors is that they *might* be easily overwhelmed by data. My advice is to check the implications of one's prior, uniform or otherwise.

6. CONCLUSIONS

The state of Bayesian statistics in the twenty-first century is strong. Bayesian statistics provides relatively simple tools to address the increasingly complex models that modern technology and concerns have brought to the fore. But the history of Bayesian methodology was difficult in the twentieth century. It is desirable for us to know where Bayesian statistics has been, not only where Bayesian statistics is going. It is useful for us to know the scientific and philosophical basis for Bayesian statistics, not only the techniques that allow its application.

REFERENCES

Bem, D., Utts, J., and Johnson, W. O. (2011), "Must Psychologists Change the Way They Analyze Their Data? A Response to Wagenmakers, Wetzels, Borsboom and Van der Mass," *Journal of Personality and Social Psychology*, 101, 716–719. [9]

Christensen, R., Johnson, W., Branscum, A., and Hanson, T. (2010), *Bayesian Ideas and Data Analysis: An introduction for Scientists and Statisticians*, Boca Raton, FL: CRC Press. [9,10]

Deely, J., and Lindley, D. (1981), "Bayes Empirical Bayes," *Journal of the American Statistical Association*, 376, 833–841. [10]

Geisser, S. (1993), *Predictive Inference: An Introduction*, Boca Raton, FL: CRC Press. [10]

Gelfand, A., and Smith, A. (1990), "Sampling Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, 410, 398–409. [9]

Samaniego, F. J. (2010), *A Comparison of the Bayesian and Frequentist Approaches to Estimation*, New York: Springer. [9]

Seaman, J. W., Seaman, J. W., Jr, and Stamey, J. D. (2012), "Hidden Dangers of Specifying Noninformative Priors," *The American Statistician*, 66, 77–84. [11]

Utts, J., Norris, M., Suess, E., and Johnson, W. (2010), "The Strength of Evidence Versus the Power of Belief: Are We All Bayesians?" in *Data and Context in Statistics Education: Towards an Evidence-Based Society* (Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia), ed. C. Reading, Voorburg, The Netherlands: International Statistical Institute. Available at *http://www.stat.auckland.ac.nz/ iase/publications.php* [2010 ISI/IASE]. [9]

Discussion: Bayesian Methods: Applied? Yes. Philosophical Defense? In Flux

Deborah G. MAYO

1. INTRODUCTION

I am grateful for the chance to comment on the article by Gelman and Robert. I welcome seeing statisticians raise philosophical issues about statistical methods, and I entirely agree that methods not only should be applicable but also capable of being defended at a foundational level. "It is doubtful that even the most rabid anti-Bayesian of 2010 would claim that Bayesian inference cannot apply" (Gelman and Robert 2013). This is clearly correct; in fact, it is not far off the mark to say that the majority of statistical applications nowadays are placed under the Bayesian umbrella, even though the goals and interpretations found there are extremely varied. There are a plethora of international societies, journals, postdocs, and prizes with "Bayesian" in their name, and a wealth of impressive new Bayesian textbooks and software is available. Even before the latest technical advances and the rise of "objective" Bayesian methods, leading statisticians were calling for eclecticism (e.g., Cox 1978), and most will claim to use a smattering of Bayesian and non-Bayesian methods, as appropriate. George Casella (to whom their article is dedicated) and Roger Berger (2002), in their superb textbook, offer models of a balanced approach.

What about the issue of the foundational defense of Bayesianism? That is the main subject of these comments. Whereas many practitioners see the "rising use of Bayesian methods in applied statistical work" as being in support of a corresponding Bayesian philosophy, Gelman and Shalizi (2012) declared that "most of the standard philosophy of Bayes is wrong" (p. 2). The widespread use of Bayesian methods does not underwrite the classic subjective inductive philosophy that Gelman (2011) associates (correctly) with the description of Bayesianism found on Wikipedia: "Our key departure from the mainstream Bayesian view (as expressed [e.g., in Wikipedia]) is that we do not attempt to assign posterior probabilities to models or to select or average over them using posterior probabilities.

Deborah G. Mayo is Professor, Department of Philosophy, Virginia Tech, Blacksburg, VA 24061 (E-mail: *mayod@vt.edu*).

 DOI: 10.1080/00031305.2012.752410

Instead, we use predictive checks to compare models to data and use the information thus learned about anomalies to motivate model improvements" (p. 71).

From the standpoint of this departure, Gelman and Robert defended their Bayesian approach against Feller's view "that Bayesian methods are absurd—not merely misguided but obviously wrong in principle."

Given that Bayesian methods have widely inundated teaching and applications, a reader might at first be puzzled by the authors' choice to consider Feller's 1950 introduction to probability, the text of which gives a page or two to "Bayes Rule." Noting that "before the ascendance of the modern theory, the notion of equal probabilities was often used as synonymous for 'no advance knowledge,'" Feller questioned the "'law of succession of Laplace' connected with this" (Feller 1950, pp. 124–125, of the 1970 edition). The authors readily concede that "[I]t would be accurate, we believe, to refer to Bayesian inference as being an undeveloped subfield in statistics at that time, with Feller being one of the many academics who were aware of some of the weaker Bayesian ideas but not of the good stuff," (Gelman and Robert 2013).

Yet the authors have a deeper reason to examine Feller. As they reiterate, what strikes them "about Feller's statement was not so much his stance as his apparent certainty." They "doubt that Feller came to his own considered judgment about the relevance of Bayesian inference. . . . Rather, we suspect that it was from discussions with one or more statistician colleagues that he drew his strong opinions about the relative merits of different statistical philosophies," (Gelman and Robert 2013).

Whether or not their suspicion of Feller is correct, they have identified a common tendency in foundational discussions of statistics simply to be swayed by colleagues and often-repeated criticisms, rather than arriving at one's own considered conclusion. Also to their credit, their defense is not "defensive." Indeed, in some ways, they raise stronger criticisms of Bayesian standpoints than Feller himself:

> In the last half of the twentieth century, Bayesians had the reputation (perhaps deserved) of being philosophers who were all too willing to make broad claims about rationality, with optimality theorems that were ultimately built upon questionable assumptions of subjective probability, in a denial of the garbage-in-garbage-out principle, thus defying all common sense. In opposition to this nonsense, Feller (and others of his time) favored a mixture of Fisher's rugged empiricism and the rigorous Neyman–Pearson theory, which "may be not only defended but also applied" (Gelman and Robert 2013).

Perhaps Bayesians have gotten over the reputation cited by Gelman and Robert of "being philosophers who were all too willing to make broad claims about rationality," but, by and large, philosophers have not. I regard the most important message of their article as being a call for a change from all players.

2. PROBABILISM IN CONTRAST TO SAMPLING THEORY STANDPOINTS

Tellingly, the authors begin their article by observing that "[y]ounger readers of this journal may not be fully aware of the passionate battles over Bayesian inference among statisticians in the last half of the twentieth century." They are undoubtedly correct, and that alone attests to the predominance of Bayesian methods and pro-Bayesian arguments in statistics courses. By contrast, few readers are unaware of the litany of criticisms repeatedly raised regarding statistical significance tests, confidence intervals, and the frequentist sampling-theory justifications for these tools. We heartily share their sentiment:

> At the very least, we hope Feller's example will make us wary of relying on the advice of colleagues to criticize ideas we do not fully understand. New ideas by their nature are often expressed awkwardly and with mistakes—but finding such mistakes can be an occasion for modifying and improving these ideas rather than rejecting them. (Gelman and Robert 2013)

The construal of Neyman–Pearson statistics that is so widely lampooned reflects Neyman and Pearson's very early attempt to develop a formalism that would capture the Fisherian and other methods used at the time. As Pearson remarked in his response to Fisher's (1955) criticisms, "Indeed, from the start we shared Professor Fisher's view that in scientific enquiry, a statistical test is 'a means of learning'" (Pearson 1955, p. 206).

Underlying one of the philosophers' examples, Gelman and Robert discuss that (the doomsday argument) "is the ultimate triumph of the idea, beloved among Bayesian educators, that our students and clients don't really understand Neyman-Pearson confidence intervals and inevitably give them the intuitive Bayesian interpretation." The idea "beloved among Bayesian educators" does not merely assert that probability *should* enter to provide posterior probabilities—an assumption we may call probabilism; it assumes that the frequentist error statistician also shares this goal. Thus, whenever error probabilities, be they p-values or confidence levels, disagree with a favored Bayesian posterior, it is alleged to show that frequentist methods are self-contradictory and thus unsound.

For example, the fact that a frequentist p-value can differ from a Bayesian posterior (in two-sided testing, assuming one or another prior) has been regarded as showing that p-values overestimate the evidence against a (point) null (e.g., Berger 2003). That a sufficiently large sample size can result in rejecting a null deemed plausible by a Bayesian is thought to show the logical unsoundness of significance tests (Howson 1997a, 1997b; Relevant references are far too numerous; please see, for example, Mayo 1996; Mayo and Spanos 2011, and references therein). Assuming that confidence levels are to give posterior probabilities to the resulting interval estimate, Jose Bernardo declared that non-Bayesians "should be subject to some re-education using well known, standard counter-examples such as the fact that conventional 0.95-confidence regions may actually consist of the whole real line" (2008, p. 453). The situation with all of these alleged "counterexamples" looks very different when error probabilities associated with methods are employed to indicate the parameter values that are or are not well indicated by the data (e.g., Mayo 2003, 2005, 2010, 2011; Mayo and Cox 2010; Mayo and Spanos 2006). Error probabilities are not posteriors, but refer to the distribution of a statistic $d(X)$—the so-called sampling distribution (hence the term *sampling* *theory*). Admittedly, this alone is often claimed to be at odds with mainstream (at least subjective) Bayesian methods where consideration of outcomes other than the one observed is disallowed (i.e., the likelihood principle), at least once the data are available. In Jay Kadane's recent text, "Neyman-Pearson hypothesis testing violates the likelihood principle, because the event either happens or does not; and hence has probability one or zero" (Kadane 2011, p. 439).

It often goes unrecognized that criticisms of frequentist statistical methods assume a certain philosophy about statistical inference (probabilism) and often allege that error probabilities serve only to control long-run error rates (behavioristic goals). Feller, in declaring that "the modern method of statistical tests and estimation is less intuitive but more realistic," also revealed the common tendency to assume a philosophy of probabilism (Feller 1950, pp. 124–125, of the 1970 edition). Our own intuitions go in a different direction: what is intuitively required are ways to quantify how well-tested claims are, and how precisely and accurately they are indicated. Still, we admit that good error probabilities, while necessary, do not automatically satisfy the goal of capturing the well testedness of inferences.

However, when we try to block the unintuitive inferences, for example, by conditioning on error properties that are relevant for assessing well testedness, we hear that "there is a catch" (Ghosh, Delampady, and Samanta 2006, p. 38): we supposedly are led to violate other familiar frequentist principles (sufficiency, weak conditionality), at least according to a famous argument (by Allan Birnbaum in 1962). Once again, critics place us in a self-contradictory position, but we argue that the frequentist is simply presented with a false dilemma and that "the 'dilemma' argument is therefore an illusion" (Cox and Mayo 2010; Mayo 2010).

While the text by Gelman et al. (2003) is a noteworthy exception, it is standard for texts to list, in addition to the above "counterexamples," an assortment of classic fallacies (conflating statistical and substantive significance, fallacies of insignificant results, fallacies of rejection), which, to echo the authors' point about Feller, stem from often-heard strong opinions of frequentist methods, overlooking how frequentists have responded. The current situation in statistical foundations may present an opportunity to reconsider them, free of the traditional frameworks both of Bayesian and frequentist statistics. The appeal to a testing notion may also be relevant to justify the Bayesian account that Gelman and Robert advance.

3. A TESTING DEFENSE FOR BAYESIANISM?

The authors correctly suspect that what has bothered mathematicians such as Feller comes from assuming

> "that Bayesians actually seem to believe their assumptions rather than merely treating them as counters in a mathematical game ... [T]his interpretation may be common among probabilists, whereas we see applied statisticians as considering both prior and data models as assumptions to be valued for their use in the construction of effective statistical inferences," (Gelman and Robert 2013).

Rather than believing their assumptions, the authors suggest that they test them:

> [W]e make strong assumptions and use subjective knowledge in order to make inferences and predictions that can be tested by comparing to observed and new data (see Gelman and Shalizi, 2012, or Mayo, 1996 for a similar attitude coming from a non-Bayesian direction). (Gelman and Robert 2013)

So perhaps some kind of a "non-Bayesian checking of Bayesian models" (Gelman and Shalizi 2012, p. 11) would offer more promise than attempts at a reconciliation of Bayesian and frequentist ideas by way of long-run performance properties.

To pursue such an avenue, one still must reckon with a fundamental issue at the foundations of Bayesian method: the interpretation of and justification for the prior probability distribution, the use of which is arguably what distinguishes it from frequentist error statistics. To their credit, the authors concede "that many Bayesians over the years have muddied the waters by describing parameters as random rather than fixed. Once again, for Bayesians as much as for any other statistician, parameters are (typically) fixed but unknown. It is the knowledge about these unknowns that Bayesians model as random" (Gelman and Robert 2013).

Although many illustrations enable an intuitive grasp of what they seem to have in mind, viewing the knowledge of fixed unknowns as random, if it is to sit at the foundations, calls for explication. The authors are right to observe that most statisticians are comfortable with probability models:

> Bayesians will go the next step and assign a probability distribution to a parameter that one could not possibly imagine to have been generated by a random process, parameters such as the coefficient of party identification in a regression on vote choice, or the overdispersion in a network model, or Hubble's constant in cosmology. There is no inconsistency in this opposition once one realizes that priors are not reflections of a hidden "truth" but rather *evaluations of the modeler's uncertainty about the parameter* (emphasis mine)

But it is precisely the introduction of "the modeler's uncertainty about the parameter" that is so much at the heart of questions involving the understanding and justification of Bayesian methods. It would be illuminating to hear the authors' take on the different conceptions of and debates about this "modeler's uncertainty" about a parameter. Arguably, the predominant uses of Bayesian methods come from those who advocate "objective" or "default" or "reference" priors (we use the neutral term "conventional" Bayesians, but any preferred term will do). Yet contemporary conventional Bayesians have worked assiduously to develop priors that are not supposed to be considered expressions of uncertainty, ignorance, or degree of belief, they are "mathematical concepts" of some sort used to obtain posterior probabilities. While subjective Bayesians urge us to incorporate background information into the analysis of a given set of data by means of a prior probability on alternative hypotheses (perhaps attained through elicitations of degrees of belief), some of the most influential Bayesian methods in practice invite us to employ conventional priors that have the least influence on resulting inferences, letting the data dominate. Conventional priors, unlike what might be expected from measures of initial uncertainty in parameters, are model dependent, "leading to violations of basic principles, such as the likelihood principle and the stopping rule principle" (Berger 2006, p. 394) and thus to Bayesian incoherence. Even within the conventional Bayesian school, there are many from which to choose: priors based on the asymptotic model-averaged information differences (between the prior and the posterior), matching priors that yield optimal frequentist methods, and others besides (Kass and Wasserman 1996; Berger 2006). Cox (2006) summarized some of the concerns he had often articulated:

> [T]he prior distribution for a particular parameter may well depend on which is the parameter of primary interest or even on the order in which a set of nuisance parameters is considered. Further, the simple dependence of the posterior on the data only via the likelihood regardless of the probability model is lost. If the prior is only a formal device and not to be interpreted as a probability, what interpretation is justified for the posterior as an adequate summary of information? (p. 77)

Bayesian testing seems to be in a state of flux. The authors' invitation to test Bayesian models, including priors, is welcome, but the results of testing are clearly going to depend on explicating the intended interpretation of whatever is being tested.

Elsewhere it is suggested that there need not be a uniquely correct conventional nor subjective prior; it may be a "combination of the prior distribution and the likelihood, each of which represents some compromise among scientific knowledge, mathematical convenience, and computational tractability" (Gelman and Shalizi, pp. 12–13). (Without presuming Robert concurs, we assume that the authors endorse some latitude in interpreting priors. See also Gelman 2011.) There is no problem with the prior serving many functions, so long as its particular role is pinned down for the case at hand (Mayo 2012a,b). These authors correctly argue that the assumptions of the likelihood are also just that—assumptions—but we still need to understand what is being represented. If the prior and likelihood is regarded as a holistic model, it is still possible to test for adequacy; but to pinpoint the source of any misfits would seem to require more. (On testing assumptions see Mayo and Spanos 2004.)

Finally, if we agree with these authors that the key goal is "to make inferences and predictions that can be tested by comparing to observed and new data," we need a notion of adequate/inadequate tests. A basic intuition is that a test has a good capacity, or at least some capability, of detecting inadequacies and flaws in whatever is being tested. The philosophy of statistics we favor employs frequentist error probabilities to appraise and ensure the probative capacity, or severity, of tests, being sensitive to the actual data and claim to be inferred. Admittedly, in developing this statistical philosophy, mistakes and shortcomings in the typical behavioristic construal of frequentist methods were used as "an occasion for modifying and improving these ideas rather than rejecting them," to echo these authors. Possibly this can offer a nontraditional avenue for a philosophical defense of the Bayesian testing these authors advance.

4. CONCLUDING REMARKS

Bayesian methods are widely applied, but when the discussion turns to foundations, there is a question as to whether the success stories are properly credited to mainstream philosophical subjective Bayesianism. Gelman and Robert, as we understand them, say no. Failure to provide an alternative defense for widely used Bayesian methods is at the heart of criticisms that continue. Stephen Senn (2011, p. 58) called attention to "A very standard form of argument . . . frequently encountered in many applied Bayesian articles where the first paragraphs laud the Bayesian approach on various grounds, in particular its ability to synthesize all sources of information," and in the rest of the article, the authors engage in non-Bayesian inexplicit reasoning. The objection loses its force, if some nonstandard or even non-Bayesian defense is involved, but that is something that requires development. We do not deny that there is an epistemological foundation for the authors' Bayesian approach, only that the foundations for Bayesian testing are in some flux and deserve attention.

Our take-home message in a nutshell is this: contemporary Bayesianism is in need of new foundations, whether they are to be found in non-Bayesian testing, or elsewhere. Hopefully philosophers of probability will turn their attention to these tantalizing problems of statistics. In contrast to the heady golden era of philosophy of statistics of 25 or 40 years ago, contemporary philosophers of science are far more focused on probability than statistics. While some of the issues have trickled down to the philosophers, by and large, we see "formal epistemology" assuming the traditional justifications for probabilism that are being questioned by contemporary statisticians, Bayesian and non-Bayesian. Gelman and Robert are among the philosophically minded statisticians who are taking the lead currently. An incomplete list of others includes the late G. Casella, D. R. Cox, J. Berger, R. Berger, J. Bernardo, R. Kass (2011), S. Senn (2011), C. Shalizi, and L. Wasserman (2011). As practitioners, it suffices that their methods are useful and widely applied; we philosophical underlaborers, by contrast, cannot escape the need to make explicit core philosophical defenses.

REFERENCES

Berger, J. O. (2003), "Could Fisher, Jeffreys and Neyman Have Agreed on Testing?" *Statistical Science*, 18, 1–12. [12]

——— (2006), "The Case for Objective Bayesian Analysis" and Rejoinder, *Bayesian Analysis*, 1, 385–402; 457–464. [14]

Bernardo, J. M. (2008), Comment on "Objections to Bayesian Statistics" by A. Gelman, *Bayesian Analysis*, 3, 451–454. [12]

Birnbaum, A. (1962), "On the Foundations of Statistical Inference," in *Breakthroughs in Statistics* (Vol. 1), eds. S. Kotz and N. Johnson, Springer Series in Statistics, New York: Springer-Verlag, pp. 478–518; First published (with discussion) in *Journal of the American Statistical Association*, 57, 269–306. [13]

Casella, G., and Berger, R. L. (2002), *Statistical Inference (*2nd ed.), Pacific Grove, CA: Duxbury Press. [11]

Cox, D. R. (1978), "Foundations of Statistical Inference: The Case for Eclecticism," *Australian Journal of Statistics*, 20, 43–59; Knibbs Lecture, Statistical Society of Australia, 1977. [11]

——— (2006), *Principles of Statistical Inference*, Cambridge: Cambridge University Press. [14]

Cox, D. R., and Mayo, D. G. (2010), "Objectivity and Conditionality in Frequentist Inference," in *Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science*, eds. D. Mayo and A. Spanos, Cambridge: Cambridge University Press, pp. 276–304. [13]

Feller, W. (1950), *An Introduction to Probability Theory and Its Applications*, New York: Wiley. [12,13]

Fisher, R. A. (1955), "Statistical Methods and Scientific Induction," *Journal of the Royal Statistical Society,* Series B, 17, 69–78. [12]

Gelman, A. (2011), "Induction and Deduction in Bayesian Data Analysis" (Special Topic: Statistical Science and Philosophy of Science), *Rationality, Markets and Morals (RMM)*, 2, 67–78. [11,14]

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003), *Bayesian Data Analysis* (2nd ed.), London: Chapman and Hall Press. [13]

Gelman, A., and Robert, C. (2013), "Not Only Defended but Also Applied: The Perceived Absurdity of Bayesian Inference." *The American Statistician*, 67, 1–5. [11,12,13]

Gelman, A., and Shalizi, C. R. (2012), "Philosophy and the Practice of Bayesian Statistics" (with discussion), *British Journal of Mathematical and Statistical Psychology*, 1–31. [11,13]

Ghosh, J. K., Delampady, M., and Samanta, T. (2006), *An Introduction to Bayesian Analysis*, New York: Springer. [13]

Howson, C. (1997a), "A Logic of Induction," *Philosophy of Science*, 64, 268–290. [12]

——— (1997b), "Error Probabilities in Error," *Philosophy of Science*, 64, 194. [12]

Kadane, J. (2011), *Principles of Uncertainty*, Boca Raton: Chapman & Hall. [13]

Kass, R. (2011), "Statistical Inference: The Big Picture," *Statistical Science*, 26, 1–9. [14]

Kass, R., and Wasserman, L. (1996), "The Selection of Prior Distributions by Formal Rules," *Journal of the American Statistical Association*, 91, 1343–1370. [14]

Mayo, D. G. (1996), *Error and the Growth of Experimental Knowledge*, Chicago: University of Chicago Press. [12,13]

——— (2003), "Could Fisher, Jeffreys and Neyman Have Agreed on Testing? Commentary on J. Berger's Fisher Address," *Statistical Science*, 18, 19–24. [12]

——— (2005), "Evidence as Passing Severe Tests: Highly Probed vs. Highly Proved," in *Scientific Evidence*, ed. P. Achinstein, Baltimore: Johns Hopkins University Press, pp. 95–127. [12]

——— (2010), "An Error in the Argument From Conditionality and Sufficiency to the Likelihood Principle," in *Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science*, eds. D. Mayo and A. Spanos, Cambridge: Cambridge University Press, pp. 305–314. [12,13]

——— (2011), "Statistical Science and Philosophy of Science: Where Do/Should They Meet in 2011 (and Beyond)?" (Special Topic: Statistical Science and Philosophy of Science), *Rationality, Markets and Morals (RMM)*, 2, 79–102. [12]

——— (2012a), "Statistical Science Meets Philosophy of Science Part 2: Shallow versus Deep Explorations" (Special Topic: Statistical Science and Philosophy of Science), *Rationality, Markets, and Morals (RMM)*, 3, 71–107. [14]

——— (2012b), Comments on "Philosophy and the Practice of Bayesian Statistics," by A. Gelman and C. Shalizi, *British Journal of Mathematical and Statistical Psychology*. [14]

Mayo, D., and Cox, D. (2010), "Frequentist Statistics as a Theory of Inductive Inference," in *Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability and the Objectivity and Rationality of Science*, eds. D. Mayo and A. Spanos, Cambridge: Cambridge University Press, pp. 247–275; This paper appeared in *The Second Erich L. Lehmann Symposium: Optimality*, 2006, Lecture Notes-Monograph Series, Volume 49, Institute of Mathematical Statistics, 247–275. [12]

Mayo, D. G., and Spanos, A. (2011), "Error Statistics," in *Philosophy of Statistics: Handbook of Philosophy of Science* (Vol. 7), eds. P. Bandyopadhyay and M. Forster (volume eds.); D. M. Gabbay, P. Thagard, and J. Woods (general eds.), The Netherlands: Elsevier, pp. 1–46. [12]

——— (2004), "Methodology in Practice: Statistical Misspecification Testing," *Philosophy of Science*, 71, 1007–1025. [14]

——— (2006), "Severe Testing as a Basic Concept in a Neyman-Pearson Philosophy of Induction," *British Journal of Philosophy of Science*, 57, 323–357. [12]

Pearson, E. S. (1955), "Statistical Concepts in Their Relation to Reality," *Journal of the Royal Statistical Society,* Series B, 17, 204–207. [12]

Senn, S. (2011), "You may Believe You Are a Bayesian but You Are Probably Wrong" (Special Topic: Statistical Science and Philosophy of Science), *Rationality, Markets and Morals (RMM)*, 2, 48–66. [14]

Wasserman, L. (2011), "Low Assumptions, High Dimensions" (Special Topic: Statistical Science and Philosophy of Science), *Rationality, Markets, and Morals (RMM)* 2, 201–209. [14]