

**An open letter to Editors of journals, Chairs of departments, Directors of funding programs, Directors of graduate training, Reviewers of grants and manuscripts, Researchers, Teachers, and Students:**

*Statistical methods have been evolving rapidly, and many people think it's time to adopt modern Bayesian data analysis as standard procedure in our scientific practice and in our educational curriculum. Three reasons:*

- 1. Scientific disciplines from astronomy to zoology are moving to Bayesian data analysis. We should be leaders of the move, not followers.*
- 2. Modern Bayesian methods provide richer information, with greater flexibility and broader applicability than 20th century methods. Bayesian methods are intellectually coherent and intuitive. Bayesian analyses are readily computed with modern software and hardware.*
- 3. Null-hypothesis significance testing (NHST), with its reliance on  $p$  values, has many problems. There is little reason to persist with NHST now that Bayesian methods are accessible to everyone.*

*My conclusion from those points is that we should do whatever we can to encourage the move to Bayesian data analysis. Journal editors could accept Bayesian data analyses, and encourage submissions with Bayesian data analyses. Department chairpersons could encourage their faculty to be leaders of the move to modern Bayesian methods. Funding agency directors could encourage applications using Bayesian data analysis. Reviewers could recommend Bayesian data analyses. Directors of training or curriculum could get courses in Bayesian data analysis incorporated into the standard curriculum. Teachers can teach Bayesian. Researchers can use Bayesian methods to analyze data and submit the analyses for publication. Students can get an advantage by learning and using Bayesian data analysis.*

The goal is encouragement of Bayesian methods, not prohibition of NHST or other methods. Researchers will embrace Bayesian analysis once they learn about it and see its many practical and intellectual advantages. Nevertheless, change requires vision, courage, incentive, effort, and encouragement!

Now to expand on the three reasons stated above.

- 1. Scientific disciplines from astronomy to zoology are moving to Bayesian data analysis. We should be leaders of the move, not followers.*

Bayesian methods are revolutionizing science. Notice the titles of these articles:

**Bayesian computation: a statistical revolution.** Brooks, S.P. Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 361(1813), 2681, 2003.

**The Bayesian revolution in genetics.** Beaumont, M.A. and Rannala, B. Nature Reviews Genetics, 5(4), 251-261, 2004.

**A Bayesian revolution in spectral analysis.** Gregory, P.C. AIP Conference Proceedings, 557-568, 2001.

**The hierarchical Bayesian revolution: how Bayesian methods have changed the face of marketing research.** Allenby, G.M. and Bakken, D.G. and Rossi, P.E. Marketing Research, 16, 20-25, 2004

**The future of statistics: A Bayesian 21st century.** Lindley, DV. *Advances in Applied Probability*, 7, 106-115, 1975.

There are many other articles that make analogous points in other fields, but with less pithy titles. If nothing else, the titles above suggest that the phrase “Bayesian revolution” is not an overstatement.

The Bayesian revolution spans many fields of science. Notice the titles of these articles:

**Bayesian analysis of hierarchical models and its application in AGRICULTURE.** Nazir, N., Khan, A.A., Shafi, S., Rashid, A. *InterStat*, 1, 2009.

**The Bayesian approach to the interpretation of ARCHAEOLOGICAL DATA.** Litton, CD & Buck, CE. *Archaeometry*, 37(1), 1-24, 1995.

**The promise of Bayesian inference for ASTROPHYSICS.** Loredo TJ. In: Feigelson ED, Babu GJ, eds. *Statistical Challenges in Modern Astronomy*. New York: Springer-Verlag; 1992, 275–297.

**Bayesian methods in the ATMOSPHERIC SCIENCES.** Berliner LM, Royle JA, Wikle CK, Milliff RF. In: Bernardo JM, Berger JO, Dawid AP, Smith AFM, eds. *Bayesian Statistics 6: Proceedings of the sixth Valencia international meeting, June 6–10, 1998*. Oxford, UK: Oxford University Press; 1999, 83–100.

**An introduction to Bayesian methods for analyzing CHEMISTRY data:: Part II: A review of applications of Bayesian methods in CHEMISTRY.** Hibbert, DB and Armstrong, N. *Chemometrics and Intelligent Laboratory Systems*, 97(2), 211-220, 2009.

**Bayesian methods in CONSERVATION BIOLOGY.** Wade PR. *Conservation Biology*, 2000, 1308–1316.

**Bayesian inference in ECOLOGY.** Ellison AM. *Ecol Biol* 2004, 7:509–520.

**The Bayesian approach to research in ECONOMIC EDUCATION.** Kennedy, P. *Journal of Economic Education*, 17, 9-24, 1986.

**The growth of Bayesian methods in statistics and ECONOMICS since 1970.** Poirier, D.J. *Bayesian Analysis*, 1(4), 969-980, 2006.

**Commentary: Practical advantages of Bayesian analysis of EPIDEMIOLOGIC DATA.** Dunson DB. *Am J Epidemiol* 2001, 153:1222–1226.

**Bayesian inference of phylogeny and its impact on EVOLUTIONARY BIOLOGY.** Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP. *Science* 2001, 294:2310–2314.

**Geoadditive Bayesian models for FORESTRY defoliation data: a case study.** Musio, M. and Augustin, N.H. and von Wilpert, K. *Environmetrics*. 19(6), 630–642, 2008.

**Bayesian statistics in GENETICS: a guide for the uninitiated.** Shoemaker, J.S. and Painter, I.S. and Weir, B.S. *Trends in Genetics*, 15(9), 354-358, 1999.

**Bayesian statistics in ONCOLOGY.** Adamina, M. and Tomlinson, G. and Guller, U. *Cancer*, 115(23), 5371-5381, 2009.

**Bayesian analysis in PLANT PATHOLOGY.** Mila, AL and Carriquiry, AL. *Phytopathology*, 94(9), 1027-1030, 2004.

**Bayesian analysis for POLITICAL RESEARCH.** Jackman S. *Annual Review of Political Science*, 2004, 7:483–505.

The list above could go on and on. The point is simple: Bayesian methods are being adopted across the disciplines of science. We should not be laggards in utilizing Bayesian methods in our science, or in teaching Bayesian methods in our classrooms.

Why are Bayesian methods being adopted across science? Answer:

2. *Bayesian methods provide richer information, with greater flexibility and broader applicability than 20th century methods. Bayesian methods are intellectually coherent and intuitive. Bayesian analyses are readily computed with modern software and hardware.*

To explain this point adequately would take an entire textbook, but here are a few highlights.

\* In NHST, the data collector must pretend to plan the sample size in advance and pretend not to let preliminary looks at the data influence the final sample size. Bayesian design, on the contrary, has no such pretenses because inference is not based on  $p$  values.

\* In NHST, analysis of variance (ANOVA) has elaborate corrections for multiple comparisons based on the intentions of the analyst. Hierarchical Bayesian ANOVA uses no such corrections, instead rationally mitigating false alarms based on the data.

\* Bayesian computational practice allows easy modification of models to properly accommodate the measurement scales and distributional needs of observed data.

\* In many NHST analyses, missing data or otherwise unbalanced designs can produce computational problems. Bayesian models seamlessly handle unbalanced and small-sample designs.

\* In many NHST analyses, individual differences are challenging to incorporate into the analysis. In hierarchical Bayesian approaches, individual differences can be flexibly and easily modeled, with hierarchical priors that provide rational “shrinkage” of individual estimates.

\* In contingency table analysis, the traditional chi-square test suffers if expected values of cell frequencies are less than 5. There is no such issue in Bayesian analysis, which handles small or large frequencies seamlessly.

\* In multiple regression analysis, traditional analyses break down when the predictors are perfectly (or very strongly) correlated, but Bayesian analysis proceeds as usual and reveals that the estimated regression coefficients are (anti-)correlated.

\* In NHST, the power of an experiment, i.e., the probability of rejecting the null hypothesis, is based on a single alternative hypothesis. And the probability of replicating a significant outcome is “virtually unknowable” according to recent research. But in Bayesian analysis, both power and replication probability can be computed in straight forward manner, with the uncertainty of the hypothesis directly represented.

\* Bayesian computational practice allows easy specification of domain-specific psychometric models in addition to generic models such as ANOVA and regression.

Some people may have the mistaken impression that the advantages of Bayesian methods are negated by the need to specify a prior distribution. In fact, the use of a prior is both appropriate for rational inference and advantageous in practical applications.

\* It is *inappropriate not* to use a prior. Consider the well known example of random disease screening. A person is selected at random to be tested for a rare disease. The test result is positive. What is the probability that the person actually has the disease? It turns out, even if the test is highly accurate, the posterior probability of actually having the disease is surprisingly small. Why? Because the prior probability of the disease was so small. Thus, incorporating the prior is crucial for coming to the right conclusion.

\* Priors are explicitly specified and must be agreeable to a skeptical scientific audience. Priors are not capricious and cannot be covertly manipulated to predetermine a conclusion. If skeptics disagree with the specification of the prior, then the robustness of the conclusion can be explicitly examined by considering other reasonable priors. In most applications, with moderately large data sets and reasonably informed priors, the conclusions are quite robust.

\* Priors are useful for cumulative scientific knowledge and for leveraging inference from small-sample research. As an empirical domain matures, more and more data accumulate regarding particular procedures and outcomes. The accumulated results can inform the priors of subsequent research, yielding greater precision and firmer conclusions.

\* When different groups of scientists have differing priors, stemming from differing theories and empirical emphases, then Bayesian methods provide rational means for comparing the conclusions from the different priors.

To summarize, priors are not a problematic nuisance to be avoided. Instead, priors should be embraced as appropriate in rational inference and advantageous in real research.

If those advantages of Bayesian methods are not enough to attract change, there is also a major reason to be repelled from the dominant method of the 20th century:

3. *20th century null-hypothesis significance testing (NHST), with its reliance on  $p$  values, has many severe problems. There is little reason to persist with NHST now that Bayesian methods are accessible to everyone.*

Although there are many difficulties in using  $p$  values, the fundamental fatal flaw of  $p$  values is that they are ill defined, because any set of data has many different  $p$  values.

Consider the simple case of assessing whether an electorate prefers candidate A over candidate B. A quick random poll reveals that 8 people prefer candidate A out of 23 respondents. What is the  $p$  value of that outcome if the population were equally divided? There is no single answer! If the pollster intended to stop when  $N=23$ , then the  $p$  value is based on repeating an experiment in which  $N$  is fixed at 23. If the pollster intended to stop after the 8th respondent who preferred candidate A, then the  $p$  value is based on repeating an experiment in which  $N$  can be anything from 8 to infinity. If the pollster intended to poll for one hour, then the  $p$  value is based on repeating an experiment in which  $N$  can be anything from zero to infinity. There is a different  $p$  value for every possible intention of the pollster, even though the observed data are fixed, and even though the outcomes of the queries are carefully insulated from the intentions of the pollster.

The problem of ill-defined  $p$  values is magnified for realistic situations. In particular, consider the well-known issue of multiple comparisons in analysis of variance (ANOVA). When there are several groups, we usually are interested in a variety of comparisons among them: Is group A significantly different from group B? Is group C different from group D? Is the average of groups A and B different from the average of groups C and D? Every comparison presents another opportunity for a false alarm, i.e., rejecting the null hypothesis when it is true. Therefore the NHST literature is replete with recommendations for how to mitigate the “experimentwise” false alarm rate, using corrections such as Bonferroni, Tukey, Scheffe, etc. The bizarre part of this practice is

that the  $p$  value for the single comparison of groups A and B depends on what other groups you intend to compare them with. The data in groups A and B are fixed, but merely intending to compare them with other groups enlarges the  $p$  value of the A vs B comparison. The  $p$  value grows because there is a different space of possible experimental outcomes when the intended experiment comprises more groups. Therefore it is trivial to make any comparison have a large  $p$  value and be nonsignificant; all you have to do is intend to compare the data with other groups in the future.

The literature is full of articles pointing out the many conceptual misunderstandings held by practitioners of NHST. For example, many people mistake the  $p$  value for the probability that the null hypothesis is true. Even if those misunderstandings could be eradicated, such that everyone clearly understood what  $p$  values really are, the  $p$  values would still be ill defined. Every fixed set of data would still have many different  $p$  values.

To recapitulate: Science is moving to Bayesian methods because of their many advantages, both practical and intellectual, over 20th century NHST. It is time that we convert our research and educational practices to Bayesian data analysis. I hope you will encourage the change. It's the right thing to do.

John K. Kruschke, Revised 14 November 2010, <http://www.indiana.edu/~kruschke/>