Subject: MS uploaded to Annals of Applied Statistics
From: Helene Hill <hzhill@verizon.net>
Date: 3/13/2012 4:26 PM
To: Bradley Efron <brad@stat.Stanford.EDU>

Dear Dr Efron,

I have uploaded the paper along with my first of 4 zipped files of supporting material. I will be glad to send the rest on DVD if you would like it. The remaining supporting materials consist of PDF copies of experiments from the laboratory that were provided in response to subpoena during Discovery of a *qui tam* case. My lawyer assures me that they are in the public domain. We strongly believe that these experimental results should be made public so that your readers can verify for themselves the accuracy of our analyses.

Our paper resembles in a number of respects the paper that you published in 2009 by Baggerly and Coombes (AoAS 3: 1309-1334). In their paper, they analyze results of micro-arrays that were used to predict patient responses to chemotherapeutic agents. Erroneous results could mean that a given patient could be given an ineffective drug or could miss out on receiving the most effective drug. The research that we analyze was designed to determine the levels of radioactive isotopes to be given to patients in diagnostic procedures and/or in therapeutic doses. Miscalculations could result in under and/or over treatment: ineffective or excessive exposures, the latter increasing the risk of carcinogenesis. The research was also designed to set the levels of permissible radiation exposures to workers and the general public exposed during medical procedures. Here, again, miscalculations allowing inappropriately high doses could increase cancer and mutation risks. As in our paper, Baggerly and Coombes analyzed data that formed the background of reported results but were invisible to the readers of the questioned articles.

In your guidelines for Referees, you query
1. Does the paper genuinely concern applied statistics? We believe that it does. It uses basic constructs and probability and it involves distribution functions.
2. Is the paper clearly written? The paper was edited by American Journal Editors where it was reviewed by a professional statistician. The editorial certificate number is DBFD-0979-287A-DD09-FACA.
3. Is the paper correct? For the most part, we have applied standard statistical tests such as chi-squared and Z-scores. The reader should be able to access the supporting material in order to see for him/herself that our calculations are correct.
4. Is the paper interesting? We hold that the paper is very interesting. The "max prob" analysis devised by Dr Pitt allows for the evaluation of anomalous results in triples containing 3 supposedly independent values that are, in fact, distributed in a non-uniform manner. The "max prob" metric is further strengthened by giving the questioned experimenter the benefit of the doubt. The paper is all the more interesting because it demonstrates that a single individual in the laboratory produced anomalous results over almost the entire time period that he was working in there (4 years) and that these anomalies went undetected by referees for as many as 8 publications and 4 study sections that reviewed 4 grant submissions.

The *qui tam* case failed in a court of law because it was judged based on timing and other non-science considerations ("failure to replicate just means failure to replicate") and not on science where replication is the hallmark of the scientific method. We hope you will not allow the judicial decisions to cloud your judgment and that you will review and then accept our paper which is based on sound scientific reasoning and analysis.

Sincerely yours,

Helene Z Hill, PhD
Professor of Radiology
New Jersey Medical School

Subject: Fwd: Decision "Reject" on paper AOAS1203-020
From: Helene Hill <hill@umdnj.edu>
Date: 3/16/2012 9:23 AM
To: Helene Hill <hzhill@verizon.net>

Subject: Decision "Reject" on paper AOAS1203-020
From: Tilmann Gneiting <tilmann@stat.washington.edu>
Date: 3/15/2012 12:16 PM
To: hill@umdnj.edu

Dear Helene Z Hill,

We appreciate your submission of "ANALYSIS OF KEY EXPERIMENTS OBTAINED IN DISCOVERY IN A QUI TAM CASE" to The Annals of Applied Statistics, but regret that it has been rejected.

Please see the Editor's commentary and the test below for the review.

I have read through your paper and have determined it is not suitable for AOAS and thus have decided to return it to you, so you can consider other options.

In a nutshell, while you present an interesting case study in what could be called forensic radiobiology, the statistical contents of your manuscript are not of sufficient interest, nor of sufficient originality, to make it successfully through our review process.

Specifically, the statistical techniques used are routine, with the possible exception of the "max prob" analysis. However, if there is originality in the max prob analysis, then it is probabilistic rather than statistical, and furthermore it is subject to technical concerns, as detailed below. The statistical inference part of the "max prob" analysis again is completely routine. The manuscript is written in the style of a life sciences journal, and parts of it, such as pages 8-11 and 17, are inaccessible and to any non-radiobiological audience.

Let me add a few thoughts you might want to consider in pursuing this work:

1. The "max prob" analysis raises concerns, as you do not state the null hypothesis under which you compute the nominal probabilities, nor do you provide theoretical or computational support for a range of claims about these probabilities. A more formal approach is critical here; let me make an attempt in developing it. As I understand, you consider three independent draws from a Poisson distribution with parameter lambda. Let $M\_1 <= M\_2 <= M\_3$ denote their order statistics, and let mu denote the rounded arithmetic mean of M_1, M_2 and M_3. Then you seem to conjecture that, for all lambda > 0, and all nonnegative integers m_1 and m_3, the conditional probability

Poisson_lambda($M\_2 = mu \mid M\_1 = m\_1, M\_3 = m\_3$)

is bounded above by $1.3/(m\_3 - m\_1 + 1)$ if $m\_3 - m\_1$ is even, and is bounded above by $2.6/(m\_3 - m\_1 + 1)$ if $m\_3 - m\_1$ is odd. I wonder why you cite "simulations" as evidence for these conjectures. Why don't you, instead, compute these conditional probabilities explicitly, given the values of m_1 and m_3 in the data, and using a plausible estimate of lambda? While any such procedure would be computationally intense, it may lead to a much more persuasive test, as it would use actual probabilities, rather than conjectured upper bounds.

2. In the tables, you repeatedly state a p-value of "$< 9.9 \times 10^{\wedge}(-10)$". Why not simply "$< 10^{\wedge}(-9)$"?

While I am sorry to deliver this unwelcome news, I would encourage you to pursue this work along the lines indicated above, and then to submit to a more specialized journal, most likely in radiobiology, or, if the material on pages 8-11 and 17 is cut short, to a quantitative forensic journal.

Sincerely yours,

Tilmann Gneiting
Editor, The Annals of Applied Statistics

Submission URL: http://www.e-publications.org/ims/submission/index.php/AOAS/

Title:
ANALYSIS OF KEY EXPERIMENTS OBTAINED IN DISCOVERY IN A QUI TAM CASE

Authors:
Helene Z Hill, Joel Pitt

Abstract:
Numerical and radiobiological data subpoenaed during Discovery in a qui tam lawsuit were analyzed. Coulter ZM cell counts and survival colony counts that comprised data in 8 publications, a successful R01 grant proposal and its renewal were reviewed. Chi-squared analysis was performed on the terminal digits of the recorded numbers, which were expected to be approximately uniformly distributed. Counts produced by one researcher in the laboratory diverged significantly from this expectation and raised questions about the data. The p-value for the goodness of fit to a uniform distribution applied to the terminal digits of 5,155 Coulter ZM particle counter values from 171 of this individual's experiments was less than 2.2×10^{-16}. The chi-squared p-value for the terminal digits of 3,501 of colony counts in 114 experiments was also less than 2.2×10^{-16}. In data produced by several other members of the same laboratory, the terminal digits in 2,759 cell counts in 99 experiments from the same Coulter ZM particle counter and 1,556 colony counts in 59 experiments were consistent with uniformity: the p-values similarly obtained were 0.12 and 0.57, respectively. In the analyzed data of the questioned researcher, the average or near-average value of triplicate colony counts appears as one of the triples at a frequency that also greatly exceeds expectations. Additionally, results of two key experiments could not be replicated in 22 attempts. Tritiated thymidine survival kinetics in the 22 experiments conform to radiobiological predictions but differ by orders of magnitude from the questioned individual's exponential survivals. This analysis underscores the importance of access to raw data that form the bases of publications, reports and grant applications in order to evaluate the correctness of the conclusions. Methods employed in this study may prove to be useful to others in screening numerical data for anomalous results.

Attachments:

Decision "Reject" on paper AOAS1203-020.eml   5.1 KB