

## DISCUSSION OF: STATISTICAL ANALYSIS OF AN ARCHEOLOGICAL FIND

BY HOLGER HÖFLING<sup>1</sup> AND LARRY WASSERMAN<sup>2</sup>

*Stanford University and Carnegie Mellon University*

*There are no small coincidences and big coincidences! There are only coincidences!*

From “The Statue” episode of *Seinfeld*.

**1. Introduction.** Andrey Feuerverger has undertaken a serious challenge. The subject matter is controversial and finding a sensible way to formulate the problem in a rigorous statistical manner is difficult.

The paper is notable for its thoroughness. We have rarely seen a paper on an applied problem that provides so much background material. Most importantly, the author is very careful to document all his assumptions and to remind the reader that the conclusion is sensitive to these assumptions. He resists the temptation to present his results in a sensationalistic way. Rather, he conveys his analysis in a dispassionate, understated tone. Nonetheless, he could still end up on *Oprah*.

We are trying to assess the probability of a hypothesis when the hypothesis is formed after seeing the data. This is a notoriously difficult problem. As Feuerverger notes, coincidences are common. But just how common?

One response—the nihilistic approach—is to say that it is impossible and stop there. We have much sympathy with the nihilists in a problem like this. Perhaps the scientifically honorable path is to say that any answer is misleading so it is better to provide no answer. But ultimately this is unsatisfying and we accept the author’s approach to provide an analysis with many caveats.

The question may be framed formally as follows. We observe an outcome  $x$ —a tomb with interesting names—and we want to know: is this outcome surprising? One way to quantify surprisingness is to perform the following steps:

1. Construct a sample space  $\mathcal{X}$  that contains  $x$ .
2. Identify all the outcomes  $A$  that would have been considered surprising if they had been observed.
3. Construct an appropriate null distribution  $P_0$ .
4. Compute the  $p$ -value  $p = P_0(A)$ .

---

Received September 2007.

<sup>1</sup>Supported by an Albion Walter Hewlett Stanford Graduate Fellowship.

<sup>2</sup>Supported by NSF Grant CCF-06-25879. Thanks to Rob Tibshirani and Isa Verdinelli for helpful comments.