LETTERS TO THE EDITOR

Letters to the editor will be confined to discussions of papers that have appeared in *The American Statistician* and to important issues facing the statistical community. Letters discussing papers in *The American Statistician* must be received within two months of publication of the paper; the author of the paper will then be given an opportunity to reply, and the letters and reply will be published together. All letters to the editor will be refereed. Corrections of errors that have been noted in papers published in *The American Statistician* will be listed as corrections at the end of this section.

Here is an object lesson for my statistical colleagues on what happens if they turn a blind eye when fraudulent statistical methods are used.

Much of the information here is from a March 3, 1993, response from the Defense Nuclear Agency (DNA) (Irby 1993) to my Freedom of Information Act (FOIA) request. There were 16 numbered FOIA attachments. Attachment 1 was a press release dated June 3, 1985, from the U.S. National Academy of Sciences (NAS)–National Research Council (NRC)–Institute of Medicine (IOM) Medical Follow-up Agency. It starts out as follows:

"WASHINGTON—A National Research Council review of death certificates for a large sample of 'atomic veterans' has found no consistent evidence of increased deaths from cancer or other diseases. . .."

This false negative report (Robinette 1985) was subsequently used by the U.S. Veterans Administration (VA) to deny fair compensation to the atomic veterans for their radiation injuries from exposures to fallout during atmospheric nuclear weapons tests. The fraudulent statistical methods involved both gross falsification of the actual data and the use of at least three fraudulent statistical methods in the analysis that are identified in (Bross 1990).

This is what the DNA itself has to say in the statement of work (FOIA Attachment 11) for a contract for a new study:

Some fifteen thousand military participants were erroneously evaluated or included; and that some twenty-eight thousand need to be included. The latter amounts to a 55% change to the original data base. Combined with the former, there are serious questions on the current validity/usefulness of the 1985 study. Further that study was compared to the general population, not to a control population of soldiers who did not participate in the tests. . .. This study was severely criticized by Congress, veterans organizations and veterans for these reasons.

Except for the data from the last test, PLUMBBOB, the failure to find "consistent evidence" (FOIA Attachment 1) was largely due to this falsification of the basic data. Although some misclassification is likely in a large study, it is extremely unlikely that a 55% change to a database is in any way "accidental."

The failure to find consistent evidence of increased deaths in the more adequate data for PLUMBBOB was due to more subtle kind of scientific fraud. At least three statistical methods were used, and it is hardly coincidental that all three biases are in a direction that would cover up the excess cancer at PLUMBBOB. The bias in comparison of the servicemen with the general population is the notorious "healthy soldier (worker) bias" (Sterling 1985, 1986). Although any competent biostatistician should recognize these biases, other statisticians and epidemiologists might not. Because of a "blind eye" policy, fraudulent statistical methods are barely mentioned—and almost never discussed—in publications of the American Statistical Association.

A simple statistical correction for the healthy soldier bias was given in (Bross and Bross 1987). When applied to the PLUMBBOB data, it shows that contrary to the press release from the NAS–NRC–IOM Medical Follow-up Agency, there is clear and consistent evidence of drastically increased deaths from cancer and other diseases. I subsequently presented this evidence at a panel of the VA on compensation (Bross 1988). The representatives of the Medical Follow-up Agency refused to modify the fraudulent claim in FOIA Attachment 1 and also testified that *there were no serious flaws in the 1985 study*. On the basis of these misrepresentations, the VA continued to deny fair compensation to the atomic veterans.

It might seem inconceivable that any competent funding agency or any person who read the DNA evaluation of the "validity/usefulness of the 1985 study" by the NAS–NRC–IOM Medical Follow-up Agency would award any other contract to *this* agency for a study whose purpose was to provide accurate reassessment of radiation injuries to the atomic veterans.

In fact, on September 8, 1992, DNA awarded contract DNA 001-92-C-0042 to the NAS–NRC–IOM Medical Follow-up Agency for this purpose. Note that both NAS and DNA are in extreme conflict of interest here. Both were involved in the original disastrous decisions that resulted in deliberate exposure of roughly half a million servicemen to fallout. Both gave false assurances that this fallout was harmless. I have recent atomic veteran death certificate data that show that deaths from cancer now outnumber deaths from all other causes combined. This suggests that at least half of all of these deaths are due to that "harmless" fallout.

If *The American Statistician* will publish this letter, there might be some chance of stopping the scientific fraud that has blocked justice for the atomic veterans—a population that is rapidly becoming extinct.

Irwin D. Bross
Biomedical Metatechnology Inc.
Eggertsville, NY 14226

REFERENCES

Bross, I., and Bross, N. (1987), "Do Atomic Veterans Have Excess Cancer? New Results Correcting for the Healthy Soldier Bias," *American Journal of Epidemiology*, 126, 1042–1051.

Bross, I. (1988), *Crimes of Official Science*, Buffalo, NY: Biomedical Metatechnology Press, 40–50.

——— (1990) "How to Eradicate Fraudulent Statistical Methods: Statisticians Must Do Science," *Biometrics*, 46, 1213–1225.

——— (1992), *Scientific Truth vs. Scientific Fraud: The Establishment is the Enemy of the Enterprise*, Buffalo, New York. Biomedical Metatechnology Press.

Irby, D. (1993), "Response to FOIA Request to the Defense Nuclear Agency on March 3, 1993" (with 16 attachments).

Robinette, D., Jablon, S., and Preston, T. (1985), *Mortality of Nuclear Weapons Test Participants*, Washington, DC: National Research Council.

Sterling, T., and Weinkam, J. (1985) "The 'Healthy Worker Effect' on Morbidity Rates," *Journal of Occupational Medicine*, 27, 477–482.

———(1986), "Extent, Persistence, and Constancy of the Healthy Worker or Healthy Person Effect by All and Selected Causes of Death," *Journal of Occupational Medicine*, 28, 348–353.

With regard to Ron Snee's comments on the need for doing actual projects as part of statistical education. In Box, Hunter, and Hunter's (1978 , p. 368) widely used experimental design book appears the following quote:

"One must learn by doing the thing; for though you think you know it, you have no certainty until you try."

—Sophocles (ca. 450 B.C.E.)

So one hardly needs to refer to behavioral psychology to find these ideas. Why aren't we using them, then? I suspect that the answer has more to do with incentive systems that emphasize covering extensive syllabi within a limited number of hours, satisfying prerequisites for other courses, and rewarding research rather than good teaching than it does with not knowing what should be done. Following Deming, we should look to the system, not to the individual worker. Comments anyone?

Bert Gunter
B. H. Gunter and Associates
P.O. Box 9
Hopewell, NJ 08525

REFERENCE

Box, G. P., Hunter, W. G., and Hunter, J. S. (1978), *Statistics for Experimenters*, New York: John Wiley.

There is no question that we need to look at our system of statistical education. It has many problems that need to be fixed. The incentive system is high on the list. Deming emphasizes that we need a theory to use as a basis for improvement: "Experience teaches nothing unless studied with the aid of theory" (see Neave 1990).

Student projects are one of many ways the delivery of statistical education can be improved. Doing actual projects is not a new idea, but many have not been sure why projects are effective. Work in behavioral sciences is providing a theory—a theory we can test and build on as we work to improve our system of education. It seems reasonable that we should use available theory whenever possible, even when that theory comes from outside our own field.

I also note that there is much work underway on using the concepts, methods, and tools of total quality to improve our system of education at all levels: how we run our educational institutions and how we deliver our courses (see Chaffee and Sherr 1992; Coate 1990; Hogg and A. Hogg 1993; Hogg and M. Hogg 1993; *Quality Digest* 1993; Seymour 1992). People are recognizing that 'total quality management' provides the processes to make the necessary changes. The availability of these total quality processes should enable and speed up the improvement of our educational systems.

Ronald D. Snee
Joiner Associates Inc.
3800 Regent Street
Madison, WI 53705

REFERENCES

Chaffee, E. E., and Sherr, L. A. (1992), *Quality: Transforming Postsecondary Education*, Washington, DC: George Washington University.

Coate, E. L. (1990), "Implementing Total Quality in a University Setting," *NACUBO Business Officer*, 26–35.

Hogg, R. V., and Hogg, A. C. (1993), "A Quality Journey: Observations and Suggestions Regarding Continuous Process Improvement," *Total Quality Management*, 195–214.

Hogg, R. V., and Hogg, M. C. (1993), "Total Quality Management in Higher Education: A Paradigm Shift," in *The Ellis R. Ott Workshop on Academic Programs in Quality*, Rochester: NY: Rochester Institute of Technology.

Neave, H. R. (1990), *The Deming Dimension*, Knoxville, TN: SPC Press.

Quality Digest (1993), "Special Report: Quality in Education," September 1993, 25–63.

Seymour, D. (1992), *On Q: Causing Quality in Higher Education*, New York: Macmillan.

MEHTA, C. R. AND HILTON J. F. (1993), "EXACT POWER OF CONDITIONAL AND UNCONDITIONAL TESTS: GOING BEYOND THE 2 × 2 TABLE," *THE AMERICAN STATISTICIAN*, 47, 91–98: COMMENT BY BERGER AND REPLY

In this article, Mehta and Hilton argue that as the dimension of a contingency table increases from 2 × 2 to 2 × 3, the power advantage of the unconditional test over the conditional test decreases. Providing that the sample sizes are reasonably large, as they are in the Mehta and Hilton examples, I agree with this point. I think, however, that two other points in the Mehta and Hilton article deserve further comment. These points are Comment 3 (p. 96) and the discussion of randomized tests (p. 92). (I use Mehta and Hilton's notation throughout.)

In Comment 3 Mehta and Hilton, state, "It is possible that there exists a complicated test statistic for which the unconditional test always has greater power than its conditional counterpart. However, we are unaware of any procedure for finding such a test and leave this question open for further research" (p. 96).

For any conditional test, there is always an unconditional test that has the same level and greater than or equal power. The method of constructing this test was described by Boschloo (1970) for the 2 × 2 problem and was discussed by Shuster (1990) for McNemar's test. But the method is quite general, and the test statistic is not complicated. Indeed, the test statistic is just the conditional p value, the first expression in Equation (2.1) of Mehta and Hilton! Call this value $p(x)$. The level-α conditional test rejects H_0 if $p(x) \leq \alpha$. And as Equation (2.3) shows, the conditional test is an unconditional, level-α test, that is, $\sup_{H_0} \Pr(p(X) \leq \alpha) \leq \alpha$. Define

$$p_\alpha = \max\{p' \in \mathcal{P}\colon \sup_{H_0} \Pr(p(X) \leq p') \leq \alpha\},$$

where $\mathcal{P}$ is the set of all possible values of $p(x)$ as x ranges over the entire sample space. (Note that p_α could be defined in two stages like Mehta and Hilton define t_α.) Because $\sup_{H_0} \Pr(p(X) \leq \alpha) \leq \alpha$, either $p_\alpha \geq \alpha$ or $p_\alpha < \alpha$, and there are no values of x with $p_\alpha < p(x) \leq \alpha$. In either case, $\mathcal{R}_c \equiv \{x\colon p(x) \leq \alpha\} \subset \{x\colon p(x) \leq p_\alpha\} \equiv \mathcal{R}_u$. $\mathcal{R}_c$ is the rejection region of the conditional test, and $\mathcal{R}_u$ is the rejection region of an unconditional, level-α test that is at least as powerful as the conditional test. It can happen that $\mathcal{R}_c = \mathcal{R}_u$, and the tests are identical. But typically $\mathcal{R}_c$ is proper subset of $\mathcal{R}_u$, and the unconditional test is uniformly more powerful.

The important point in this construction is to consider $p(x)$ and not $T(x)$ as the test statistic. T defines which sample points x are more extreme in Γ_s. A different function T_s could be used to define this order for each s. And as Mehta and Hilton point out, it may not be appropriate to compare values of $T(x)$ for x's that are in different sets Γ_s. But it is appropriate to compare values of $p(x)$, and by doing so a more powerful test can be constructed.

The second point that deserves comment is the Mehta and Hilton discussion (p. 92) of post hoc randomization to inflate the sizes of both tests so they are equal. They then claim that a generalization of Tocher's result would show that the conditional test is more powerful than the unconditional test; this is not true. Suissa and Shuster (1984) considered exactly this kind of comparison. In their figure 1, for example, the unconditional test has higher power than the conditional test on more than 72% of the alternative parameter space. Unbiased tests, like those considered by Tocher, must have Neyman structure, that is, must satisfy $\Pr(\text{reject } H_0 \mid s) = \alpha$ for all values of s. The unconditional test typically has $\Pr(\text{reject } H_0 \mid s) > \alpha$ for some values of s. One would have to remove points from $\mathcal{R}_u$ to make it unbiased so that Tocher's result would apply.

The whole discussion of unbiased tests should be viewed with some skepticism for these discrete problems because in these problems *every* unbiased test is a randomized test. Any unbiased test must satisfy $\Pr(\text{reject } H_0 \mid S = 0) = \alpha$. But there is only one sample point with $S = 0$, the sample point for which no successes are observed in any population. So any unbiased test must randomize at this point. Typically, unbiased tests must randomize for most or all values of S. As Mehta and Hilton state, "post hoc randomization is unacceptable in practice" (p. 92). So, in these problems, the UMP unbiased test is the best test in a class of tests that no practicing statistician would use. And, finding the UMP unbiased test, then dropping the randomization to obtain the exact conditional test, does not lead to a test with superior power properties.

In conclusion, for any conditional test there is an unconditional test of the same level that is uniformly more powerful. Work should continue to make these unconditional tests computationally available to the practicing statistician. This work is most important for small dimensional or small sample size tables. For, as noted in the beginning, I agree with Mehta and Hilton that for large dimensional and large sample size tables, the power advantage of the unconditional test is likely to be slight. Research should also continue on which ordering statistics, $T(x)$, yield tests with better power properties.

Roger. L. Berger
Department of Statistics
North Carolina State University
Raleigh, NC 27695-8203

REFERENCES

Boschloo, R. D. (1970), "Raised Conditional Level of Significance for the 2 × 2 Table, when Testing for the Equality of Two Probabilities," *Statistica Neerlandica*, 21, 1–35.

Shuster, J. J. (1990), "Conditional vs. Unconditional Tests in 2 × 2 Trials," in *Proceedings of the Biopharmaceutical Section*, Alexandria, VA: American Statistical Association.

Suissa, S., and Shuster, J. J. (1984), "Are Uniformly Most Powerful Unbiased Tests Really Best?" *The American Statistician*, 38, 204–206.

We are grateful to Berger for his thoughtful comments on our article. They helped us probe even deeper into the fundamental principles underlying conditional and unconditional tests.

Berger's claim is technically correct. His test will indeed have unconditional power equal to or greater than that of the conditional test. We wish to clarify, however, that Berger's test adopts a nonstandard approach to determining statistical significance from both the conditional and the unconditional perspectives. Conditional tests as proposed by Fisher are based on a single cutoff value, $t_\alpha(s)$, derived from the conditional distribution of T, given s. Unconditional tests, as proposed by Barnard, are also based on a single cutoff value, t_α, derived from the unconditional distribution of T (nuisance parameters being eliminated through a $\sup_{H_0}$ argument). Thus Fisher and Barnard share a common philosophy on testing statistical hypotheses. They agree that the criterion for rejection is a single cutoff value for the test statistic such that the chance of exeeding that cut-off value under H_0 is small ($\leq \alpha$) in hypothetical repetitions of the original experiment. They merely differ in their definition of what constitutes a hypothetical repetition. Fisher considers only hypothetical repetitions in which s is fixed, whereas Barnard allows s to vary across the hypothetical repetitions.

Berger's test is also an unconditional test, for he allows s to change in hypothetical repetitions of the original experiment. But unlike Barnard, he uses a variable cutoff value, $d_\alpha(s)$ say, for determining whether the test statistic T is large enough to reject H_0. This dependence of the cutoff value on s in not evident because the test statistic is expressed in terms of the conditional p value $p(x)$. It is, however precisely the added flexibility of varying the cutoff value with s that enables Berger to ensure that his test is at least as good as the conditional test. In effect Berger augments the rejection region of the conditional test at selected values of s by changing the cutoff from $t_\alpha(s)$ to $d_\alpha(s) < t_\alpha(s)$. He can do this because conditionally $\Pr(T \geq d_\alpha(s) \mid s) > \alpha$ can hold for selected values of s, even though unconditionally the supremum of the weighted sum over all possible values of s satisfies

$$\sup_{H_0} \sum_s \Pr(T \geq d_\alpha(s) \mid s) \Pr(S = s \mid H_0) \leq \alpha.$$

Berger exploits this mathematical artifact by deliberately elevating the Type-1 errors of some of the conditional tests above α but nevertheless ensuring that the overall size of his unconditional test lies below α. There is nothing mathematically wrong with this procedure. There are several reasons, however, why it might be difficult for the test to be adopted in practice.

1. There is no statistical principle or philosophy supporting the way Berger's rejection region is defined. One is accustomed to thinking of a rejection region as all values of the test statistic that are larger than some predetermined cutoff value. In this case, because the rejection region is defined by multiple cutoff values, one for each s, this traditional interpretation of the rejection region does not apply. Instead one has to think of the rejection region as the union of a set of conditional rejection regions artificially inflated to improve the overall unconditional power.

2. If Berger's procedure were adopted, one could actually report a p value greater than α and then claim that the result was statistically significant at the α level.

3. The test is easy to state but very complicated to implement. It requires that one generate every possible conditional p value over the unconditional sample space, sort them in ascending order, and search out the largest one satisfying a criterion that is based on the supremum of a weighted sum taken over all values of the nuisance parameter. It is difficult to see how all this could be achieved for contingency tables of dimension greater than 2×2.

The second issue brought up by Berger concerning Tocher's result is valid. We should have made it clear that randomization leads to the UMP test among unbiased tests only. Our article does not however argue for conditional tests based on Tocher's work, which was merely mentioned in passing.

Cyrus R. Mehta
Harvard School of Public Health
Cambridge, MA

Joan F. Hilton
University of California
San Francisco, CA

FISCH, R. D., AND STREHLAU, G. A. (1993), "A SIMPLIFIED APPROACH TO CALIBRATION CONFIDENCE SETS," *THE AMERICAN STATISTICIAN*, 47, 168–171: COMMENT BY LAMOTTE AND REPLY

Fisch and Strehlau (1993) set out to describe "classical calibration confidence sets for straight line regression...in a new and easy-to-understand way" (p. 168). I believe they have presented an easy-to-understand way, but only in two sentences buried in Section 3: "The limits $C(x)$ are intersection points of the horizontal line at height y_{n+1} with hyperbolas. These hyperbolas result from connecting the upper and lower limits of the $100^*(1-\alpha)\%$ confidence intervals for individual future observations at given x values" (p. 169). This confidence set is clearly justified by Theorem 4 in Lehmann (1959, p. 79), about which Lehmann says, "The [result] shows the structure of the confidence sets $S(x)$ as the totality of parameter values θ for which the hypothesis $H(\theta)$ is accepted when x is observed. A confidence set can therefore be viewed as a combined statement regarding the tests of the various hypotheses $H(\theta)$, which exhibits the values for which the hypothesis is accepted ($\theta \in S(x)$) and those for which it is rejected ($\theta \notin S(x)$)" (p. 79). With this approach, given y and y_{n+1}, we ask whether $y - y_{n+1}$ is significantly different from 0 at x. If not, then x goes in the confidence set; otherwise it does not. Interpretation is straightforward and clear. For each x in the confidence set, a value of Y as far out as y_{n+1} would not be extremely unlikely; for x not in the confidence set, values of Y as far out as y_{n+1} would be extremely unlikely. Simple in concept and interpretation, this approach has the further advantage that it generalizes readily to multiple predictor variables as well as to multivariate Y.

Lynn Roy LaMotte
Department of Experimental Statistics
Louisiana State University
Baton Rouge, LA 70803-5606

REFERENCE

Lehmann, E. L. (1959), *Testing Statistical Hypotheses*, New York: John Wiley.

We do not share LaMotte's opinion that there are only two sentences in our paper that are easy to understand. The derivation of the result is performed using only an elementary toolkit, namely, least-squares regression, parameter estimation, and confidence sets. Anybody who understands this can easily follow the approach.

Several things are unclear in LaMotte's comment: What plays the role of the Θ ? Are x and y parameters or random variables? What is their relation, $y = \beta_0 + \beta_1 x$? What is the meaning of "unlikely" in this context?

Commenting on his last remark, we emphasize that our approach is easily extended to multiple predictor variables and multivariate Y, though the calculations might become more tedious.

Roland D. Fisch
Günther A. Strehlau
Mathematical Applications
Ciba-Geigy AG
4002 Basel
Switzerland

JUOLA, R. C. (1993), "MORE ON SHORTEST CONFIDENCE INTERVALS," *THE AMERICAN STATISTICIAN*, 47, 117–119: COMMENT BY FERENTINOS AND KOUROUKLIS AND REPLY

We would like to point out that the method presented by Juola (1993) for constructing shortest confidence intervals has previously been used by Ferentinos and Kourouklis (1990).

Kosmas FERENTINOS
Department of Mathematics
University of Ioannina
45110 Ioannina, Greece

Stavros KOUROUKLIS
Department of Mathematics
University of Patras
26110 Patras, Greece

REFERENCE

Ferentinos, K. and Kourouklis, S. (1990), "Shortest Confidence Interval Estimation for Families of Distributions Involving Two Truncation Parameters," *Metrika*, 37, 353–363.

I am really sorry this has happened, but I was not aware of their article. This is especially embarrassing to me since it was Professor Ferrentinos's 1990 paper in this journal that was the primary motivation for the submission of my paper.

I sincerely apologize to Professors Ferentinos and Kourouklis.

Robert JUOLA
Boise State University
Boise, ID 83725

SOLOW, A. R. (1993), "ESTIMATING RECORD INCLUSION PROBABILITY," *THE AMERICAN STATISTICIAN*, 47, 206–208: COMMENT BY DARGAHI-NOUBARY

While reading the article by Solow (1993), I became curious about applying his methods for estimating record inclusion probabilities to earthquake data. This was mainly because this article is a continuation of an approach introduced by Lee and Brillinger (1979) and implemented by Guttorp and Thompson (1990) to earthquake data.

The basic assumption for development of estimation technique is that Y_t $(t = 0, 1, 2, \ldots, n)$, the unknown true number of events in period t, is a sequence of independent Poisson random variables with unknown stationary mean $E(Y_t) = m$. As pointed out in Solow, the assumption of constant mean is crucial, since otherwise it is not possible to distinguish between changes in $E(Y_t)$ and changes in inclusion probability P_t. The question is then to see whether this is a reasonable assumption in the case of earthquake data.

In seismology there has been a great deal of interest in the problem of seismic recurrence. The degree of periodicity of a given fault or seismically active region is often described by means of the interval distribution function $w(t)$. A perfectly periodic system has $w(t) = \delta(t - t_0)$ while a fault system with a constant seismicity is a Poisson process whose intervals are distributed according to

$$w(t) = \exp(-\lambda t). \tag{1}$$

The function (1) has found wide application in seismic hazard assessment, but over the years it has been found that Equation (1) does not adequately represent the detailed evolution of seismically active regions. An alternative to the Poisson process is the so-called seismic gap hypothesis, which states that the likelihood of earthquake occurrence is small immediately following the previous earthquake and increases with time since the last event on certain fault or plate boundaries (see Sykes and Nishenko [1984, p. 5911]). That is, an earthquake reduces the potential energy stored in the system to a low value, from which point the process of increase of potential energy begins once again. The seismic gap hypothesis has been used in long-term forecasting of earthquakes around the whole Pacific rim (e.g., Nishenko 1991), in California (Bakun and Lindh, 1985; Sykes and Niskenko 1984; Working Group On California Earthquake Probabilities, 1988), and in other Pacific regions (e.g., Kelleher et al. 1973). In fact, many seismologists (e.g., Oppenheimer et al. 1990, p. 8483; Thatcher 1989, p. 432;) treat the seismic gap hypothesis as one confirmed by observations. Recently Kagan and Jackson (1991) tested the seismic gap hypothesis using a set containing large earthquakes around the Pacific rim. They considered the times of large earthquakes as a point process and used the coefficient of variation, defined as the ratio of standard deviation of interval times to the mean recurrence time. A quasi-periodic process has a coefficient of variation of less than one, while a clustering process has a coefficient of variation of greater than one. The Poisson process has a coefficient of variation equal to one and predicts an earthquake occurrence independent of the time of previous seismic activity. Note that the seismic gap hypothesis assumes that earthquake occurrence is a quasi-periodic process. Earthquake potential is low when the time elapsed since the last large earthquake is less than the recurrence time, and earthquake potential is high afterward. It is interesting to note that the paleoseismic record at Palette Creek, California, is the longest reliable record and suggests a clustering process, although a Poisson process cannot be rejected (see Sieh et al. 1989). This means the occurrence of an earthquake reveals that the region is active and one should expect more earthquakes in the near future.

Thus from this research I learned that, when applying a standard statistical technique, one should remember that data alone often do not provide a complete picture. If I want to perform a meaningful statistical analysis, I should incorporate other available information, especially that coming from the discipline related to the origin of the data.

G. R. DARGAHI-NOUBARY
Department of Mathematics
Bloomsburg University
Bloomsburg, PA 17815

REFERENCES

Bakun, W. H., and Lindh, A. G. (1985), "The Parkfield, California, Earthquake Prediction Experiment," *Science*, 229, 619–624.

Guttorp, P., and Thompson, M. L. (1990), "A Note on Point Process Analysis of Chinese Earthquake Data," *Pure and Applied Geophysics*, 133, 541–546.

Kagan, Y. Y., and Jackson, D. D. (1991), "Seismic Gap Hypothesis: Ten Years After," *Journal of Geophysical Research*, 96, 21419–21431.

Kelleher, J. A., Syskes, L. R., and Oliver, J. (1973), "Possible Criteria for Predicting Earthquake Locations and Their Applications to Major Plate Boundaries of the Pacific and Caribbean," *Journal of Geophysical Research*, 78, 2547–2585.

Lee, W. H. K., and Brillinger, D. R. (1979), "On Chinese Earthquake History—An Attempt to Model an Incomplete Data Set by Point Process Analysis," *Pure and Applied Geophysics*, 117, 1229–1257.

Nishenko, S. P. (1991), "Circum-Pacific Seismic Potential 1989–1999," *Pure and Applied Geophysics*, 135, 169–259.

Oppenheimer, D.H., Bakun, W. H., and Lindh, A. G. (1990), "Slip Partitioning of the Calaveras Fault, California and Prospects for Future Earthquake," *Journal of Geophysical Research*, 8483–8498.

Sieh, K., Stuiver, M., and Brillinger, D. (1989), "A More Precise Chronology of Earthquakes Produced by the San Andreas Fault in Southern California," *Journal of Geophysical Research*, 94, 603–623.

Sykes, L. R., and Nishenko, S. P. (1984), "Probabilities of Occurrence of Large Plate Rupturing Earthquakes for the San Andreas, San Jacinto, and Imperial Faults, California, 1983–2003," *Journal of Geophysical Research*, 89, 5905–5927.

Thatcher, W., (1989), "Earthquake Recurrence and Risk Assessment in Circum-Pacific Seismic Gaps," *Nature*, 341, 432–434.

Working Group on California Earthquake Probabilities, (1989), "Probabilities of Large Earthquakes Occurring in California on the San Andreas Fault," U.S. Goelogical Survey, Open File Rep., 88–398, 62 pp.