Comments on the Three Invited Papers at IPM58

P. K. Ito
68/2 Rakuen-cho, Showa-ku,
Nagoya, Japan 466-0822

1. Introduction

First of all I would like to thank the three speakers for presenting very interesting and thought-provoking papers to us. I am sure these contributions are important and valuable additions to the list of references on statistical education and the significance tests controversy.

Before I give my comments on each of the three papers, let me review very briefly the current states of the significance tests controversy in theoretical statistics, applied statistics, and statistical education. Of course the issues are inter-related in these fields; i.e., one can be a theoretical statistician one day and an applied statistician or a teacher of statistics next day, and vice versa.

In theoretical statistics, I note carefully the viewpoints of the purists of each school: Fisherian, Neyman-Pearson, Subjective probability Bayesian, Objective probability Bayesian, and Likelihoodist. There are a kind of eclectics who recognize that each school has its own paradigm in which to build its own theory of statistical inference and/or decision-making. Another kind of eclectics are those who try to unify different schools and create a new unified, say, frequentist and Bayesian theory of hypothesis testing.

In applied statistics I refer to two works which are not quoted in the papers of today. One is "Sense and Nonsense of Statistical Inference" by C.Wang(1993) and the other is "What If There Were No Significance Tests?" edited by L.L.Harlow and others(1997).

In statistical education I refer to the two recent conferences, one is in Chicago, 1996, and the other ICOTS5 in 1998. In view of the shortcomings of the frequentist approach in statistical inference including significance testing, many papers propose Bayesian approach in the university statistics courses, even for beginning students. Noteworthy is one exception, D.S.Moore(1997), who hesitates to do so.

In discussing the topic of today, I note that in practical works the significance test is an informal blending of Fisher's original pure significance test and Neyman-Pearson theory with concepts and interpretations that are not a part of the latter.

2. Comments

In the first paper, Teaching hypothesis testing. Can it still be useful?, Professor Dahl explains that statistics including hypothesis testing and interval estimation was recently started at secondary schools in Norway, which was not without controversy. There were problems in texts, and he stresses key points in teaching statistical tests: the philosophy behind the methods and the basic concepts, with interesting and meaningful examples. It seems, however, the tests he talked about are Neyman-Pearson decision-theoretic tests, quite different from Fisherian significance tests. Neyman-Pearson theory could be legitimately applied in quality control and acceptance sampling, and should be taught as such at schools. Another comment is: Teaching statistical inference at secondary schools is premature, and we had better wait until the university.

In the second paper, Some empirical evidences on learning difficulties about testing hypotheses, Dr. Vallecillos found, by means of going through published results and her own experiments, that pre-university and university students have difficulties in learning basic concepts such as the logic of hypothesis testing, etc. and recommends more researches be made to obtain a more precise knowledge about what our students really learn. However, I wonder if she could consider, at this stage, how to overcome these difficulties and propose sample syllabi of statistics courses for these students.

In the third paper, Beyond the significance test controversy: Prime time for Bayes?, Professor Lecoutre found problems and short-comings of NHST and difficulties of interpretation of confidence intervals in experimental data analysis, especially in psychology. He asserts that a widely accepted objective Bayes theory is of immense theoretical and practical importance, which is not speculative, but desirable and feasible. He and his colleagues are working in a fiducial(for motivation) and Bayesian(for technique) perspective to develop standard("noninformative") Bayesian methods, so that Bayesian interpretations of significance tests and confidence intervals may be made, which are considered quite natural to students. It seems that Professor Lecoutre has made Fisher almost a Bayesian, with which being an eclectic I am not surprised.

However, I would not take the viewpoint of eclectics who want to unify different schools of thought. I would rather join Cox(1987) who states that the comparison of different approaches to the same or similar problems is important, sometimes soothing and occasionally constructively alarming.

3. Concluding remarks

In conclusion, I would like to propose sample syllabi of statistics courses for (1) elementary and secondary school students, (2) non-statistics majors who will later use statistical methods, (2a) liberal arts students, and (3) graduate students interested in experimental data analysis.

BIBLIOGRAPHY

Cox,D.R.(1987). Comment in Berger,J.O.and Delampady,M. Testing precise hypotheses. Statist. Sci.,3, 317-352.

Harlow,L.L.,Mulaik,S.A.,and Steiger,J.H.(eds.)(1997). What If There Were No Significance Tests? Lawrence Erlbaum Associates,Inc.

Moore, D.S.(1997). Bayes for beginners? Some reasons to hesitate. The American Statistician, 51(3), 254-261.

Wang,C.(1993). Sense and Nonsense of Statistical Inference. Marcel Dekker,Inc.