

BEHAVIORAL STATISTICS TEXTBOOKS: SOURCE OF MYTHS AND MISCONCEPTIONS?

James K. Brewer
The Florida State University

Introduction

Numerous writers, among them Bakan (1966), Brewer (1978), Carver (1978), Coats (1910), Dooling and Danks (1975), Lykken (1968), Morrison and Henkel (1970), Tversky and Kahneman (1971), Wilson (1973) and Winch and Campbell (1969) have argued quite forcefully that users of statistics in the behavioral sciences have been guilty of misunderstanding and misapplying even the most rudimentary concepts and procedures of applied statistics. Why is this the case when almost every university and college in America has several departments teaching applied statistics courses in the behavioral sciences? We are quick to hold researchers responsible for statistical abuses, but it may well be that researchers are only parroting what they have read or been taught. Since the most common element in almost all teaching of behavioral statistics is the textbook, could it not be that the textbook is a source of statistical "myths and misconceptions" so often denounced as misleading and inappropriate?

At the American Educational Research Association annual meeting of 1977, Roger E. Kirk said, in effect, that the essential elements of hypothesis testing will be utilized by researchers only when these elements are adequately presented in textbooks in the behavioral sciences. Although these paraphrased remarks were made relative to hypothesis testing concepts and sample size concerns, they are equally appropriate to all concepts of behavioral statistics. If one believes Kirk's statement to be true, then a form of contrapositive is also true, namely, that a conceivable source of statistical misconceptions and errors occurring in the published literature, theses, and dissertations is the behavioral statistics textbook. To illustrate the nature and extent of myths and misconceptions found in some of the best-selling introductory behavioral statistics textbooks is the purpose of this paper.

Procedure

Requests were sent to 18 statistics textbooks publishers asking for the text and author of their best selling introductory behavioral statistics textbook published during 1982. The publishers were selected because they (1) had previously published statistics texts, (2) had an external review process or textbook manuscripts, and (3) were relatively well-known in the academic world. Introductory texts were selected for consideration because this author believes that the first exposure to statistics is where a form of "misconception imprinting" takes place, crucially affecting the researchers' statistical beliefs and behaviors for years to come.

The authors, text names and responding publishers are coded alphabetically by publisher, and the letters "A" through "F" will be used throughout this report when referring to a particular author and text. The remaining 12 publishers either published no such text during 1982 or chose not to respond to this writer's request. The three major headings under which the example quotes will be grouped are (1) Hypothesis testing concepts, (2) Confidence interval concepts, and (3) Sampling distribution and the Central Limit Theorem. Although there were 43 quotes found in the 6 textbooks, only a subset of these will be used to give the general "flavor" of the type M and M reflected therein. (The complete list of quotes along with texts and authors' names will be provided upon request.)

Hypothesis Testing Concepts

This area of inference is chosen first because it is probably the most misunderstood, confused, and abused of all possible behavioral statistics topics. Each textbook author has his or her own way of addressing the concepts of hypothesis testing using real (and not so real) situations to clarify the concepts and give meaning to behavioral science problems. It may very well be that in this desire to provide simplification for the mathematically naive, authors have fostered rather than minimized the misconceptions.

An Example Hypothesis Testing Misconceptions

When we reject a hypothesis at the 1 percent level, we are saying that the chances are 99 in 100 that it is false (C, page 150).

Confidence Interval Concepts

The misconceptions which are associated most often with confidence intervals appear to be related to their interpretations and their relation to hypothesis testing. The general tendency in published textbooks seems to be to combine hypotheses testing with confidence intervals, giving the impression that they are used interchangeably rather than that we are separate inferential techniques. The misconceptions which may result from this impression is that researchers would fail to understand that hypotheses testing and confidence intervals do not necessarily have anything to do with each other and thereby often involve different prior considerations and minimum sample sizes.

An Example Confidence Interval Misconceptions

In our example, the 99% confidence interval is thus 73 ± 3.87 or 69.13 to 76.87. When we assert that the unknown falls within this range of values, 99% of such assertions will be correct (F page 154).

Sampling Distributions and the Central Limit Theorem

Misconceptions relative to sampling distributions seem to center around misconceptions or misstatements of the Central Limit Theorem (CLT). The major abuses appear to stem from a penchant on the part of textbook authors to argue that almost anything will be approximately normal as long as the sample size is large.

An Example Statement Reflecting a Central Limit Theorem Misconception

In section 10.2 we said that from the basic result of the central limit theorem, we learn that any variable that's sampled repeatedly and randomly tends to be distributed normally, the larger the size of the sample is (D, page 152).

Implications and Recommendations

In their attempts to explain inferential statistics in such a way as to make sense to the readers, authors of some behavioral statistics texts have sacrificed correctness. The general effect of an incorrect (but "sounds good") approach to statistics is that readers and users are led to believe that statistical inference allows the researchers to say more than is permissible and thereby to draw more conclusions than are warranted from the data. To illustrate this effect consider the recent instructions from a researcher to a graduate student for whom the researcher was major professor: "You must reach significance at the .05 level or your dissertation will not be acceptable." A student faced with such an unnecessary dilemma could understandably take random subsamples of a larger random sample until H_0 was rejected at the .05 level. Although this would be pure capitalization on chance, the criterion would have been met and the student would have learned the value of violating ethical standards.

Given that there are statistical myths and misconceptions in well-used behavioral statistics texts, a natural question is, "why has this happened?". A possible answer is the same as the answer to most of the world's ills – ignorance. If an author is ignorant of statistical theory, the editor ignorant of which reviewers are knowledgeable of statistical theory, and textbook purchasers ignorant of statistical theory, then it is no wonder that misconceptions are fostered through published texts. Nothing can stop authors from writing textbooks, but editors can see that manuscripts are reviewed by people with collectively more knowledge of statistical theory and applications than the author of the manuscript. To send manuscripts for review, as some editors do, to authors of previously published behavioral statistics texts often only serves to perpetuate the misconceptions. Editors, however, cannot be held totally responsible since theoretical statisticians quite often do not wish to review behavioral statistics textbook manuscripts regardless of the fee. Thus the books get inferior reviews (less theoretical scrutiny) by default.

It may be of some interest to note that in the 6 textbooks reviewed for this study only 3 of the 14 authors were listed as members of the American Sta-

tistical Association (ASA) or its affiliates in the 1978 Directory of Statisticians. Of the 17 people mentioned by the authors as having reviewed the texts or been instrumental in publication of the texts, five were listed as members of the ASA or its affiliates. Clearly, membership in a learned society does in no way bestow qualifications upon a textbook author, but it does imply that the vast majority of authors and reviewers of these textbooks have primary commitments which are not to statistics, either theoretical or applied. This in itself may reflect an undesirable gulf between statistician and statistics user.

Lest we be too hard on the editors, it must be remembered that they operate on a "cash validity" basis, i.e., "if it sells it's good." and statistics instructors are, after all, the ultimate purchasers and validators of textbooks. The instructors, therefore, must bear the major burden of responsibility to see that only theoretically sound, well-written textbooks reach the market. To do this they must become educated in both the theory of statistics and its applications to behavioral problems. It is this writer's opinion that many instructors of behavioral statistics have had no statistical theory by way of formal coursework in mathematical statistics, and have instead had two or three courses in elementary behavioral statistics taught from textbooks just like the one they are selecting for their courses. It is also a sad but probably true commentary, that it is precisely these instructors who will eventually write the future behavioral statistics texts, still without the proper theoretical exposure. Thus, the statistical myths will continue to be nurtured and fostered.

References

- Baken, D. (1966). The test of significance in psychological research. Psychological Bulletin, 66, 423-437.
- Brewer, J.K. (1978). Everything you always wanted to know about statistics but didn't know how to ask. Dubuque, Iowa: Kendall Hunt Publishing Company.
- Carver, R.P. (1983). The case against statistical significance testing. Harvard Educational Review, 48(3), 378-398.
- Dooling, D.J., & Danks, H.D. (1975). Going beyond tests of significance: Is psychology ready? Bulletin of the Psychonomic Society, 5(1), 15-17.
- Lykken, D.T. (1968). Statistical significance in psychological research. Psychological Bulletin, 70, 151-159.
- Morrison, D. E. & Henkel, R. E. (1970). Significance tests in behavioral research: Skeptical conclusions and beyond. In D.E. Morrison & R.E. Henkel (Eds.), The significance test controversy - A reader. Chicago: Aldine.
- Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. Psychological Bulletin, 76(2), 105-110.

Wilson, J. (1973). Three myths in educational research. Educational Research, 16(1), 17-19.

Winch, R.F., & Campbell, D.T. (1969). Proof? no. evidence? yes. The significance of tests of significance. The American Sociologist, 4, 140-143.