

Discussion Paper
on I.P. David's Critique

by A. Buenaventura¹

Since I am substantially in agreement with the views expressed in Dr. David's paper, my discussion will focus on other instances where the jump from hypothesis to conclusion was made with no thought or care as to whether inferences arrived at were statistically tenable or not.

We at the Statistical Center are no strangers to the phenomenon described by Dr. David, in which researchers approach the statistician with data collected from a completed sample survey. Some of them have absolutely no idea of how to analyze the data. Their only specification is that it must be "statistical." Others specify the type of statistical analysis to be used only to be told that such an analysis is not possible, given the design of the survey, or more often, the absence of any design at all. In some cases, it is possible to perform some salvage operations but when the researcher is a thesis student on the verge of graduation, the only advice that can be given is to describe the characteristics of the sample, making it clear in the report that such descriptions apply only to the individuals selected, and that no inferences can be made about the characteristic of the target population.

In this respect, there are two related problems which usually face the statistician:

(a) The researcher wants to make statistical inference from non-probability samples; and

¹Dean, U.P. Statistical Center

(b) Even if the sample is a probability sample, the statistical analysis specified by the researcher is not warranted by the design of the survey.

These problems are not recent ones. The example I will cite occurred more than thirty years ago. Perhaps, then, the statisticians have to carry the burden of blame for this state of affairs. If the problem existed as far back as thirty years ago, why have they not done anything about it? We are all aware of how, for instance, the doctors zealously guard their monopoly on the practice of medicine. Have the statisticians been closetting themselves in their ivory towers, forcing other social scientists to preempt the roles of survey statistician?

The example I will cite is a 1948 study entitled "Sexual Behavior in the Human Male" by Kinsey, Pomeroy and Martin, hereafter called KPM. I am using this example because when it came out, it was widely publicized (by virtue of its subject matter), quoted, and even used as a benchmark by other researchers in the field.

Three well-known statisticians, William Cochran, Frederick Mosteller, and John Tukey were appointed by the American Statistical Association to review the statistical methods used in the KPM study.

The authors introduced their Report on the sampling aspect of KPM with the following statements:

"Whether by biologists, sociologists, engineers, or chemists, sampling is often taken too lightly. . . . Any excuse for the practice of treating non-random samples as random ones is now entirely tenuous. Wider knowledge of the principles involved is needed if scientific investigations involving samples are to be solidly based."

In the summary of their Report, Cochran, Mosteller and Tukey noted that the KPM study was superior to other studies of sexual behavior so far undertaken, specially in their modes of interview, the amount of data collected and their painstaking attempts to tabulate and organize their data.

It was not clear what their target population was, but all their respondents were white male Americans, with almost one half of the sample coming from the state of Indiana.

Their sampling scheme was to select individuals in clusters (with no attempt at probability sampling) and with equal allocation with respect to age, education, and marital status. Cell means were computed as though the sample was srs. Similarly, standard errors attached to the means for individual cells were presented. Even in KPM has been a probability sample, these standard errors, calculated on the assumption of srs, were under estimates, perhaps by a substantial amount, because sampling was by clusters.

No statistical tests of hypothesis were attempted but KPM arrived at conclusions which did not seem warranted by the data presented. For instance, the respondents were divided into two age groups: those below 33 years (median age: 21.2 years) and those above (median age: 43.1 yrs.) Based on frequency distributions and other descriptive statistics; KPM concluded that differences in sexual activity between generations was slight. This is one of the controversial findings of KPM and was strongly criticized on the grounds that KPM tended to brush aside reported differences in activity as "immaterial", although they did not state what they considered as a "material" difference.

Furthermore, with respect to some of the other conclusions of KPM, the Report states that "We are convinced that unsubstantiated assertions are not, in themselves, inappropriate in a scientific study. In any complex field, where many questions remain unresolved, the accumulated insight of an experienced worker frequently merits recording when no documentation can be given. However, the author who values his reputation for objectivity will take pains to warn the reader, frequently repetitiously, whenever an unsubstantiated conclusion is being presented, and will choose his words with the greatest care.

Many of the most interesting statements in KPM are not based on the tabular material presented and it is not at all clear on what evidence the statements are based. Nevertheless, the statements are presented as if they were well-established conclusion."

