Discussion of the paper: Some Empirical Evidences on Learning Difficulties about Testing Hypotheses by Augustias Vallecillos

Michael F. Capobianco

St. John's University, New York

300 Howard Ave.

Staten Island, NY 10301 USA

capobiam@stjohns.edu

This paper is a summary of observed difficulties encountered by students in the learning of hypothesis testing. It includes some observations of the author himself.

I was happy with the various classifications of these difficulties namely, the four "conceptions" students have of just what a test is, and the three incorrect conceptions of level of significance. I have taken the liberty, for purposes of further discussion, of ranking these three in order of seriousness. I rank the second in the author's list as most serious since this error does not even recognize that significance level is a conditional probability. Next, I would put the third in the list since the error probability is not associated with null hypothesis or with rejection. The first on the list I think is the least serious since it does retain the notion of conditional probabiity. All three are pretty bad, however.

There is no question that the concept of conditional probability, when one first studies it, is a difficult one. I should like to share a little story with you that I've used in class, and which sometimes helps to make the point.

It seems there was a busy executive who necessarily did a a lot of flying around the world, and became very concerned about the many incidents of bombs being found on commercial airliners. He therefore asked his statistician to give him an estimate of the probablity of a bomb being aboard a plane. The statistician collected some data and reported that this probability was very high. The excutive thought for a while and then asked for the probabiity of two bombs being aboard. The statistician researched this question, and this time she reorted that this probability was quite low. From that day one, the executive carried his own bomb!

With regard to section 6 and the understanding of the Null Hypothesis, it seems to me that the very name, NULL hypothesis indicates that this is not what one wants to prove. The confusion of population and sample is very common, and, I suspect, will be with us forever. I believe that Prof. Vallecillos is correct when he states that standard textbooks deal only with simple hypotheses or non-vectorial hypotheses, i.e., those dealing with just one parameter. Composite alternatives are always discussed; these give rise to power functions, but I have never seen, on an elementary level, composite hypotheses mentioned.

In conclusion, I should like to offer still another criticism of hypothesis testing. Essentially, what bothers me is the formation of a question before ever seeing the data. For example, suppose we have some time series data which we wish to test for trend. We have no reason to believe that the trend will be up rather than down, or the reverse. Hence we use the data in some appropriate two-tailed test. I don't like this kind of procedure. If, instead, we plot the data first, we can see the direction of the trend, and use a one-tail test hence gaining power. In any situation, you may have theoretical reasons for considering certain questions, but you may wrong, or even if you are not wrong, there may be other important questions which would not come to your attention without some exploration of data. I believe that more EDA should be taught in our classes, both as an end in itself and in conjuction with the classical procedures of hypotheisi testing and estimation.