4-15-2003

A Different Approach to the Behren Fisher Problem

Madhur Mathur
Indiana University - Purdue University Fort Wayne

A Different Approach to the Behren Fisher Problem

Madhur Mathur
C. K. Chauhan
Department of Mathematical Sciences
Indiana University-Purdue University Fort Wayne

The Problem of testing the equality of means of two populations plays an important role in many areas such as Biology, Agriculture, and Psychology . It is not surprising that its solution under various situations has been discussed by different authors. In the simplest case of two independent populations with known standard deviations, the hypothesis is tested by the well known Z test statistic. However, in practice standard deviations are unknown and therefore assumed unequal.

Unfortunately, testing the equality of means is not straightforward when the standard deviations are **unequal** and **unknown**. This problem, known as **Behren Fisher Problem**. has been considered by many authors. One approach is to use sample standard deviations as estimates for the population standard deviations. The main problem with this approach is that the statistic thus obtained does not have an exact distribution.

Let x_i (i= 1,2,...,n) be a random sample from a normal population with mean $_x$ and variance $_x^2$ and y_i (i= 1,2,...,m), (m<n) be a random sample from another normal population with mean $_y$ and variance $_y^2$
Some authors have proposed the following approach: Consider the difference, $d_i = x_i - y_i$, i= 1,2,....m..Note for equal population means, the mean value of di is equal to zero. Thus a test statistic based on the di is used with an approximate t-distribution. However this approach has two problems: It results in wastage of data from the larger sample in unequal samples and the numerical value of the statistic is order dependent. Schaffe has suggested a modified approach in which all the data are used, but still order dependent.

In this research project, we propose a different procedure for testing the equality of means from two independent populations with unknown and unequal variances. Our proposed test does not require equal sample sizes from the two populations and is independent of the order in which the data is written. Our proposed test is based on the d_i's, where di are calculated by taking all possible differences, $x_i - y_j$,i=1,...,n, j=1,...,m . Note, however, some of the di and dj are pairwise independent, while others are not. Our proposed test statistic is a function of di, and the covariance between di and dj. Our approach involves the use of The Moore-Penrose Inverse of a singular matrix calculated by MATLAB.