A COMPARISON OF ANALYSES FOR TWO GROUP SMALL SAMPLES WITH A LARGE NUMBER OF MEASURES

Margo Sidell, ScD
Department of Biostatistics Tulane University
School of Public Health and Tropical Medicine

Abstract

In many fields, data are collected that have a greater number of continuous variables (p) than group sample size (n). The classical two sample method for comparing multivariate mean vectors, Hotelling's T2 test, is undefined when $p > 2n - 2$. This study examined four alternative two sample tests that are not restricted by the number of variables: a component wise statistic (Wu, Genton, & Stefanski, 2006), an extension of Hotelling's T2 test (Schott, 2007), an ANOVA type nonparametric test (Bathke, Harrar, & Madden, 2008), and a new two sample U-score test based on a single sample U-score statistic (Wittkowski et al., 2008). Monte Carlo simulations were run for each of the four tests. Multivariate normal data were generated with low, moderate, and high correlation where the number of variables was equal to or exceeded the group sample size. Each test was used for each dataset. This study investigated which of these four methods was most appropriate in terms of type I error and power under various conditions.