COMPARATIVE EMPIRICAL POWER OF UNIVARIATE NORMALITY TESTS UNDER SYMMETRIC, ASSYMETRIC AND SCALED DISTRIBUTIONS (A CASE STUDY OF 2016 UGANDA PRESIDENTIAL ELECTION RESULTS)

BY

MUYOMBYA SOLOMON MATOVU

MSTAT/36383/151/DU

BSC. STATISTICS

A THESIS SUBMITED TO THE COLLEGE OF ECONOMICS AND MANAGEMENT IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF MASTER OF SCIENCE IN STATISTICS OF KAMPALA INTERNATIONALUNIVERSITY

May, 2017

ABSTRACT

The study aims at conductingan empirical comparison of powers of the univariate normality tests under different distributions and sample sizes to obtain their rankingusing a Monte-Carlo simulation. The objectives are to investigate the effect of symmetry/assymetry and scale contamination on the empirical power of the measures of Normality,explore the effect of sample size on the empirical power of the statistical measures of normality and compare the results of simulation to real-life data for consistency.A total of six normality tests are selected fromthe Empirical Distribution Function (EDF), the Correlation and Regression family of normality tests and moment-based normality tests were considered in this study. From the EDF family of normality tests, the Kolmogorov-Smirnov test (Lilliefors correction) and Anderson-Darling normality tests were chosen. From the regression and correlation family of distributions, Shapiro-Wilk and Shapiro-Francia normality tests were chosen. Jaque-Bera and D'Agostino Pearson normality tests were chosen from the moment family of normality tests. The empirical powers of these normality tests were studied using distributions that are symmetric, asymmetric and scale contaminated normal distributions. From the symmetric distributions,$\beta(1,1)$, $N(0,1)$ and $t(10)$ were chosen to study the empirical power. From the asymmetric family of distributions,$\Gamma(3,2)$,$\beta(2,5)$ and $chi(10)$ were chosen. From the scale contaminated distributions,$N(3,1)$, $N(6,9)$ and a Normal distribution with outliers were chosen in this study. These distributions were studied at different sample sizes. The results of the Uganda general elections were used after transformation for the third objective. The findings indicate that for the symmetric distributions, Kolmogorov-Smirnov normality test is the most powerful test, followed by Anderson-Darling, Shapiro-Wilk, Shapiro-Francia, D'Agostino-Pearson and lastly Jaque-Bera. The analysis thus reveals that for symmetric distributions, theEmpirical Distribution Function based tests have the best statistical power, followed by Regression and Correlation results and lastly those based on the measures of moments for symmetric distributions. For the asymmetric distributions,Anderson-Darling,Shapiro-Wilk, Kolmogorov-Smirnov, Jaque-Bera and lastly D'Agostino-Pearson. The analysis thus reveals that for asymmetric, theEmpirical Distribution Function based tests have the best statistical power, followed by Regression and Correlation results and lastly those based on the measures of moments. For the scale contaminated distributions, Kolmogorov-Smirnov normality test is the most powerful test, Anderson-Darling, Shapiro-Francia, Shapiro-Wilk, D'Agostino-Pearson and lastly Jaque-Bera. Thus, for scaled normal distributions, the Empirical Distribution Function based tests have the best statistical power, followed by Regression and Correlation results and lastly the moment-based tests. Finally, the analysis reveals that the empirical power increases as the sample size increases with the regression-based tests producing better results for smaller sample sizes. The results were consistent with the election results data as it also indicates that S-W produced the most powerful results followed by Shapiro-Francia, D'Agostino-Pearson, Kolmogorov-Smirnov, Jaque-Bera and lastly Anderson-Darling.It was concluded that the empirical-based normality tests provide the best tests of normality though the moment-based tests provide better results if the goal is to only investigate kurtosis and skewness. Due to the robustness and sensitivity of these normality tests, the study recommends that it is imperative to deal with outliers prior to carrying out any further normality tests.