

Abstract

Tests of Normality and Symmetry of Distributions

Hydar Ali

This thesis presents tests for multivariate normality and proposes some measures and tests of symmetry. Some graphical techniques are also discussed.

Chapter 1 contains a brief literature review of tests of univariate and multivariate normality and tests of symmetry. Some symmetry plots and box plots are also reviewed.

Let $\mathbf{X}_1, \dots, \mathbf{X}_n$ be a random sample of size n from a p -variate normal distribution. Denote the sample mean and sample variance-covariance matrix by $\bar{\mathbf{X}}$ and $\mathbf{S}$, respectively. Then $\mathbf{Y}_i = \mathbf{S}^{-1/2}(\mathbf{X}_i - \bar{\mathbf{X}})$, $i = 1, \dots, n$ are asymptotically, independent, identically distributed as a p -variate normal with zero mean vector and identity covariance matrix. In Chapter 2, use is made of the fact that the np components of all the $\mathbf{Y}_i$'s together can be considered as an asymptotically, independent, identically distributed sample of size $N = np$ from a standard normal. Any test of univariate normality can then be used to test this hypothesis. The ones used are univariate skewness and kurtosis tests, which are sensitive to deviations from normality.

Let $Y_1, \dots, Y_n$ be a random sample from some population and let $\tilde{Y}$ be the sample median. Define the statistics L_n and R_n by $L_n = \sum |Y_i - \tilde{Y}| I_{(Y_i \leq \tilde{Y})} / n$ and

$R_n = \sum |Y_i - \tilde{Y}| I_{(Y_i > \tilde{Y})} / n$, respectively. In Chapter 3, it is shown that L_n and R_n are individually and jointly asymptotically normally distributed. It is further shown that the statistic $H_1 = (L_n - R_n) / (L_n + R_n)$ is also asymptotically normally distributed. For symmetric random variables these results simplify and the asymptotic variance of H_1 is evaluated for the standard normal. The bootstrap technique is used to obtain the critical values at the 5 % level of significance for the statistic H_1 .

In Chapter 4, exact expressions for the expectations and variances of L_n and R_n and their covariance are found in terms of the order statistics from a sample of size n . These expressions are again simplified for symmetric random variables.

In Chapter 5, deviations on either side of the sample median, taken in pairs, are plotted. With an appropriate choice of scale and origin, a symmetry plot is constructed. The plot also displays outliers, which are determined by modifying the usual criterion. The plot is applied to several data sets and for the bootstrap sample considered on Chapter 3.

Keywords: Hydar Ali; tests of univariate normality; tests of multivariate normality; kurtosis and skewness tests; order statistics; tests of symmetry; measures of symmetry; bootstrap estimate; symmetry plot; outliers.