

# Shapiro - Wilk test for Normality

Rationale: Assess whether a sample has been generated by a normal distribution by comparing its distribution to a theoretical normal distribution.

Ex:  $X_1, X_2, \dots, X_n$ , a sample. The test statistic ( $W$ ) is:

$$W := \left( \sum_{i=1}^n a_i X_i \right)^2 / \sum_{i=1}^n (X_i - \bar{X})^2$$

$a_i$ : are coefficients (constants) derived from moments of the order statistics from the appropriate Normal distr.  
 $X_i$ : are the ordered sample values ( $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ )  
 $\bar{X}$  is the sample mean

For large samples  $W$  becomes  $W'$  as:

$$W' := \frac{\left( \sum_{i=1}^n b_i X_i \right)^2}{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n b_i^2}$$

with  $b_i = \Phi^{-1} \left( \frac{i}{n+1} \right)$   $\Phi$  is the  $N(0,1)$  distr. function

see: Shapiro and Francia, 1972