Module 7

Lecture 6: Goodness of fit tests

Goodness of Fit

- The goodness of fit tests measure the compatibility of a random sample with a theoretical probability distribution function. In other words, these tests show how well the distribution one selected fits to the data.

  - ***Kolmogorov-Smirnov Test***

    This test is used to decide if a sample comes from a hypothesized continuous distribution. It is based on the empirical cumulative distribution function (ECDF). Assume that we have a random sample $x_1, \ldots, x_n$ from some distribution with CDF $F(x)$. The empirical CDF is denoted by

$$F_n(x) = \frac{1}{n}\left(Number\ of\ observations \le x\right)$$

The Kolmogorov-Smirnov statistic (K) is based on the largest vertical difference between the theoretical and the empirical cumulative distribution function:

$$K = \max_{1 \le i \le n} \left(F(x_i) - \frac{i-1}{n}, \frac{i}{n} - F(x_i) \right)$$

Hypothesis Testing

The null and the alternative hypotheses are:

H_0: the data follow the specified distribution;

H_A: the data do not follow the specified distribution.

The hypothesis regarding the distributional form is rejected at the chosen significance level (α) if the test statistic, K, is greater than the critical value. The fixed values of (0.01, 0.05 etc.) are generally used to evaluate the null hypothesis (H_0) at various significance levels. A value of 0.05 is typically used for most applications.

- ***Anderson-Darling Test***

The Anderson-Darling procedure is a general test to compare the fit of an observed cumulative distribution function to an expected cumulative distribution function. This test gives more weight to the tails than the Kolmogorov-Smirnov test.

The Anderson-Darling statistic (A^2) is defined as:

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1).\left[ln\left(F(x_i) + ln\left(1 - F(x_{n-i+1})\right)\right)\right]$$

Goodness of Fit **Contd…**

- ***Chi-Squared Test***

The Chi-Squared test is used to determine if a sample comes from a population with a specific distribution. This test is applied to binned data, so the value of the test statistic depends on how the data is binned. Please note that this test is available for continuous sample data only.

The Chi-Squared statistic is defined as

$$\chi^2 = \sum_{i=1}^{n} \frac{\left(O_i - E_i\right)^2}{E_i},$$

where O_i is the observed frequency for bin i, and E_i is the expected frequency for bin i calculated by, $E_i = F(x_2) - F(x_1)$, where F is the CDF of the probability distribution being tested and x_1 and x_2 are the limits for bin i.

Goodness of Fit **Contd…**

❖ **The relative goodness of a model fit may be checked using**

- RMSE (Root Mean Square Error),
- AIC (Akaike Information Criterion) and
- BIC (Bayesian Information Criterion).

RMSE

The root mean square error is defined as the square root of mean sum of square of difference between empirical distribution and theoretical distribution, i.e. $RMSE = \sqrt{MSE}$

Goodness of Fit

AIC

The AIC is used to identify the most appropriate probability distribution.

It includes:

(1) the lack of fit of the model and

(2) the unreliability of the model due to the number of model parameters

It can be expressed as:

AIC = -2(log(maximum likelihood for model)) + 2(no of fitted parameters)

or

AIC = N log(MSE) + 2(no of fitted parameters)

Goodness of Fit **Contd…**

BIC

The BIC (Schwarz, 1978) is another measurement of model selection which can be expressed as :

BIC = N log(MSE) + [(no of fitted parameters)* log(N)

The best model is the one which has the minimum RMSE, AIC and BIC values