Ideas on Goodness of Fit and Specification Testing: Evaluating the Koul Stute Test

Samridh Joshi

December 2024

1 Introduction and Brownian Motion

Goodness of fit tests are used to test the hypothesis that data come from a distribution that is either completely specified or specified up to some unknown parameters. A simple example of such a test may be described as the Kolmogorov Smirnov Test (Kolmogorov 1933) which is a nonparametric test of the hypothesis that the population CDF of the data is equal to some hypothesized CDF, testing the goodness of fit of a suggested model distribution of a data, which is achieved through a general Kolmogorov Smirnov Statistic.

$$\mathrm{D}_{\mathrm{KS}} := \sup_{x \in R} |F_n(x) - F(x)| = \|F_n - F\|$$

The Kolmogorov Smirnov Statistic can be shown by elementary transformations to be a distribution free statistic, which reduces to the form (where $\overline{F_n}$ is the empirical CDF of iid uniforms and $\alpha_n(x)$ is $\sqrt{n}\mathrm{D}_{KS}$)

$$\sup_{x \in R} |\alpha_n(x)| = \sup_{u \in [0,1]} \sqrt{n} \left| \overline{F_n(u)} - u \right|$$

which converges to the Kolmogorov Distribution (by the Kolmogorov Theorem) and is given by

$$K = \sup_{t \in [0,1]} |B(t)|$$

here, B(t) is the Brownian Bridge. This is a restricted case of the usual Brownian Motion process, which we shall hereby refer from now on as B(t) instead of the Brownian Bridge.

Calculation of critical values of the process $\sup_{t \in [0,1]} |B(t)|$ is of importance to our effort as we will see later. This task may be accomplished (Bai 2003) by simulation of the distribution of this process. We do it as suggested by approximating sample path of B(t) by normalised partial sums of 1000 iid N(0,1) variables and by simulation of 100000 such paths, ordering the maximum values in a vector and calculating critical value at 5% henceforth. The critical value obtained in our simulation is 2.215 which matches the approximated value of 2.22, suggested in the paper.

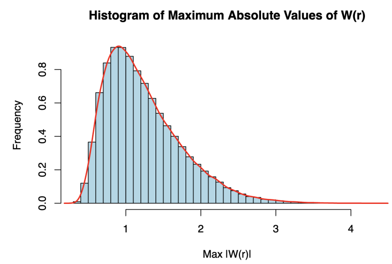

Figure 1: Brownian Motion Simulation

2 The Koul Stute Specification Test

Koul and Stute (1999) discussed some general methods for testing the goodness of fit of a parametric model for a real-valued stationary Markovian time series. The paper discusses a class of tests for testing goodness of fit hypotheses pertaining to a class of implicitly defined autoregressive functions related to first-order autoregressions. The simplest attempt is to test the null hypothesis

$$H_0 : m_\psi = m_0$$

against an alternative (where m_0 is a known function). We now summarise the Koul Stute specification test. The aim is to test for the conditional mean (μ) hypothesis. Since there may be several competing models for a given problem, in order to prevent wrong conclusions, every statistical inference which is based on a model M should be accompanied by a proper model check, that is, by a test for the hypothesis $\mu \in M$. Proposed tests for such model checks in earlier literature include tests based on least square residuals eg. Gauss Newton Regression (MacKinnon 1992). The Koul Stute procedure is described as:

Let ψ be a nondecreasing real-valued function such that $E_\psi\left[|X_1 - r|\right] < \infty$ for each $r \in R$. Define the ψ-autoregressive function m_ψ by the requirement that $E[\psi(X_1 - m_\psi)X_0 \mid X_0] = 0$ a.s.

$$E\left[\psi(X_1 - m_\psi(X_0)) \mid X_0\right] = 0 \text{ a.s.(1)}$$

and the corresponding marked empirical process, based on a sample of size $n+1$, by

$$V_{(n,\psi)}(x) := n^{-1/2} \sum_{i=1}^{n} \left[\psi(X_i - m_\psi(X_{i-1}))I(X_{i-1} \leq x)\right]; x \in R.$$

In the simple case we will evaluate, $\psi(x) \equiv x$, hence $m_\psi = \mu$. Under the assumption of finiteness of second moment of $\psi(X_1 - m_\psi(X_0))$, it follows from (1) that $V_{(n,\psi)}(x)$ is a mean zero square integrable martingale for each $x \in R$. The martingale central limit theorem shows that finite-dimensional distributions of $V_{(n,\psi)}(x)$ converge weakly to a multivariate normal distribution with mean vector zero and covariance matrix given by the covariance function

$$K_\psi(x,y) = E\left[\psi^2(X_1 - m_\psi(X_0)) I(X_0 \le x \wedge y)\right], \quad x, y \in R$$

Koul and Stute show that all finite-dimensional distributions of $V_{n,\psi}$ converge weakly to those of a centered continuous Gaussian process V_ψ with the covariance function K_ψ. Since $\tau_\psi^2(x) = K_\psi(x,x) = E\left[\psi^2(X_1 - m_\psi(X_0)) I(X_0 \le x)\right]$ is nondecreasing and nonnegative, $V_\psi(x) = B(\tau_\psi^2(x))$ in distribution. Here B is the standard brownian motion as in Section 1. From these and the continous mapping theorem (note that B($\tau_\psi^2(x)$) is continous), we have:

$$\sup_{x \in R} |V_{n,\psi}(x)| \implies \tau_\psi(\infty) \times \sup_{t \in [0,1]} |B(t)| \quad \text{in law}$$

Based on these results, we will devise a KS Test for our simple model on which the KS Test will be evaluated. We consider the following model:

$$y_{t+1} = 0.4 + 0.2 x_t + u_t$$

which is correctly specified if $m_\psi = E[u_t \mid x_t] = 0$. This is a simple model of the kind $y = f(x) + u_t$ which under dynamic effects can be seen as an autoregressive model, although for the sake of simplicity we ignore these effects. The DGP is assumed to be such that the regressor is chosen randomly from uniform distribution and the output variable from an exponential distribution with rate factor $1/0.4 + 0.2x_t$ and the dataset is thus generated.

Under this model, the KS Test is summarised as:

$$H_0 : m_\psi = 0 \quad \text{against} \quad H_1 : m_\psi \neq 0$$

Steps:

1. $$(\tau_{n,\psi})^2 = \frac{1}{T-1} \sum_{t=1}^{T-1} (u_{t+1})^2 I(x_t \le x)$$

2. $$V_{n,\psi} = \sqrt{\frac{1}{T-1} \sum_{t=1}^{T-1} u_{t+1} I(x_t \le x)}$$

3. $$S_{n,\psi}^2 = (\tau_{n,\psi}(\infty))^2 = \frac{1}{T-1} \sum_{t=1}^{T-1} (u_{t+1})^2$$

4. $$KS_n = \sup_x \left[\left(\frac{1}{T-1} \sum_{t=1}^{T-1} (u_{t+1})^2 \right)^{-\frac{1}{2}} \times \left| \sqrt{\frac{1}{T-1} \sum_{t=1}^{T-1} u_{t+1} I(x_t \le x)} \right| \right]$$

The critical value of this Koul Stute Statistic (4) is the same as the one found from the distribution of Brownian Motion in Section 1 at 5% ie. 2.22.

3 Evaluating the Koul Stute Test

The appropriate code to implement the same is written, in the usual manner with generation of a synthetic dataset from the DGP and the test is run for the hypothesis using Monte Carlo Simulation with 100 replications. The false null is tested using draws of errors from a normal with non zero mean. For simplicity, we use the errors as observed and not as residuals. The results are summarised as:

Sample Size	Type I Error	Type II Error	Power
200	0.09	0.72	0.28
500	0.04	0.34	0.66
1000	0.06	0.13	0.87
1500	0.04	0.02	0.98
2000	0.07	0	1

Table 1: Specification Test Results

The Type I error is found to be close to the significance level of 0.05. The Type II error is seen to decrease with increase in sample size and hence the power of the test approaches 1 as n tends to large values (theoretically infinity, practically is 1 for sample sizes greater than 2000). The probability of rejecting the null when the null is false goes to 1 and hence the test is shown to be consistent. Thus the Koul Stute test is expected to be a well formulated, consistent test for testing goodness of fit hypothesis of a class of implicitly defined functions.