Journal of the Korean
Statistical Society
Vol. 25, No. 1, 1996

# Kolmogorov-Smirnov Type Test for Change with Sample Fourier Coefficients

**Jaehee Kim** [1]

**Abstract**

The problem of testing for a constant mean is considered. A Kolmogorov-Smirnov type test using the sample Fourier coefficients is suggested and its asymptotic distribution is derived. A simulation study shows that the proposed test is more powerful than the cusum type test when there is more than one change-point or there is a cyclic change.

**Key Words :** Change-point model; Kolmogorov-Smirnov type test; Sample Fourier coefficients

## 1. INTRODUCTION

Detecting a change in the mean of a stochastic process is of interest in a number of areas. One of the nonparametric approach in detecting mean change is using the Fourier series coefficients. While parametric methods

[1] Department of Statistics, Duksung Women's University, Seoul, 132-714, Korea

require very specific quantitative information from the experiment, nonparametric techniques rely more on the data itself. In many function spaces, the Fourier series coefficients contain information about the underlying function. Sample Fourier coefficients, computed from regression data can be regarded as a transformation that contain the information about the data.

Page (1954, 1955) considered testing the null hypothesis that there is no change with the assumption that the initial mean is known. Wolfe and Schechtman (1984) discussed several nonparametric methods using ranks or score function for testing for no change. Csörgó and Horáth (1988) suggested a Kolmogorov-Smirnov type test based on cumulative sum. Buckley (1991) proposed a test based on the CUSUM for testing the hypothesis that the design points have no effect on the response variable. Durbin and Knott (1972) introduced Fourier analysis of a sample distribution function in the context of testing lack-of-fit.

In this paper we are concerned with a development of test statistics for detecting changes especially using Fourier series coefficients.

## 2. TESTING FOR A CONSTANT MEAN

Consider the model

$$y_i = f(x_i) + \epsilon_i, \qquad i = 1, \ldots, n \tag{2.1}$$

where $x_i = i/n$ and $\epsilon_i$'s are iid with mean zero, variance $\sigma^2$ and finite fourth moments. Of interest is testing the null hypothesis that $x$ has no effect on $y$, i.e.,

$$\text{H}_0 : f(x) = C \quad \text{for all } x \in [0,1], \text{ where } C \text{ is a constant,} \tag{2.2}$$

against

$$\text{H}_1 : f(x) \text{ is nonconstant.}$$

Under the alternative hypothesis, $f$ is arbitrary. If $f(x)$ is any smooth function, then $f(x)$ has a Fourier series representation

$$f(x) = a_0 + 2\sum_{j=1}^{\infty} a_j \cos(\pi j x), \quad 0 \le x \le 1, \tag{2.3}$$

where $a_j = \int_0^1 \cos(\pi j x) f(x) dx$, $j = 0, 1, \ldots$. The null hypothesis is equivalent to $a_j = 0$, $j = 1, 2, \ldots$.

The sample Fourier coefficients using a cosine system are defined by

$$\hat{\phi}_j = \frac{1}{n}\sum_{i=1}^{n} y_i \cos(\pi j x_i), \quad j = 0, 1, \ldots, n-1, \tag{2.4}$$

where $x_i = i/n$.

In what follows, we consider the Kolmogorov-Smirnov type test with the sample Fourier coefficients, since the sample Fourier coefficients contain the information about the underlying function. If $|\hat{\phi}_j|$ are significantly large, it would suggest that the test rejects the null hypothesis. Hence, we suggest the following test statistic with the normalized sample Fourier coefficients

$$T_n = \max_{1 \le k \le n-1} \frac{1}{k}\sum_{j=1}^{k} \frac{\sqrt{2n}|\hat{\phi}_j|}{\hat{\sigma}_n} \tag{2.5}$$

where $\hat{\sigma}_n$ is a consistent estimator of $\sigma$.

**Remark.** If there is any change in the mean process, some $|\hat{\phi}_j|$ have large values. Let

$$\Delta_n = \max_j |\hat{\phi}_j - \phi_{j0}|. \tag{2.6}$$

Including more terms of the sample Fourier coefficients and incorporating the fact that under $H_0$, $\phi_{j0} = 0$, $j = 1, 2, \ldots$, we can consider the above test statistic $T_n$.

**Theorem 2.1** (Asymptotic distribution of $T_n$)
Assume $\epsilon_i$'s are iid with mean 0 and variance $\sigma^2$ in (2.1). If $f$ is a constant function, $T_n$ converges in distribution to

$$\max_{k \ge 1} \frac{1}{k}\sum_{j=1}^{k} |Z_j|, \tag{2.7}$$

as $n \to \infty$, where $Z_j$'s are independently and identically distributed as $N(0,1)$.

**Proof.** Let $|Z_{jn}| = \sqrt{2n}|\hat{\phi}_j|/\sigma$, $j = 1, \ldots, n-1$. Firstly, we will show that there is a subsequence $\{m_n\}$ such that $m_n \to \infty$, $m_n << n$ and

$$P(T_n > \epsilon) - P(T_{m_n} > \epsilon) \to 0 \quad \text{for } \epsilon > 0. \tag{2.8}$$

Write

$$\begin{aligned} P(T_n > \epsilon) &= P(\{T_{m_n} > \epsilon\} \cup A_n) \\ &= P(T_{m_n} > \epsilon) + P(A_n) - P(\{T_{m_n} > \epsilon\} \cap A_n) \end{aligned}$$

where

$$A_n = \left\{ \max_{m_n+1 \le k \le n-1} \frac{1}{k} \sum_{j=1}^{k} |Z_{jn}| > \epsilon \right\}.$$

Let $j_{n1}$ be the largest integer such that $j^2 < m_n$ and $j_{n2}$ is the largest integer such that $j^2 \le n-1$. Letting $j_{n3} = j_{n2}$ if $j_{n2}^2 < n-1$, otherwise $j_{n3} = j_{n2} - 1$. $A_n^c$ can be written as

$$\begin{aligned} A_n^c &= \left\{ \max_{m_n < k \le n-1} \frac{1}{k} \sum_{j=1}^{k} |Z_{jn}| \le \epsilon \right\} \\ &= \bigcap_{k=m_n+1}^{n-1} \left\{ \frac{1}{k} \sum_{j=1}^{k} |Z_{jn}| \le \epsilon \right\} \\ &\supset \bigcap_{j=j_{n1}}^{j_{n3}} \left[ \left\{ \frac{\sum_{l=1}^{j^2} |Z_{ln}|}{j^2} \le \frac{\epsilon}{2} \right\} \cap \left\{ \frac{\xi_{jn}}{j^2} \le \frac{\epsilon}{2} \right\} \right] \end{aligned}$$

where

$$\xi_{jn} = \max_{1 \le i \le (j+1)^2 - j^2} \sum_{l=j^2+1}^{j^2+i} |Z_{ln}|.$$

By Markov's inequality,

$$P\left( \bigcup_{j=j_{n1}}^{j_{n3}} \left\{ \sum_{l=1}^{j^2} \frac{|Z_{ln}|}{j^2} > \epsilon \right\} \right) \le \frac{4}{\epsilon^2} \sum_{j=j_{n1}}^{j_{n3}} \frac{1}{j^4} \tag{2.9}$$

since $E|Z_{1n}|^2 = E(Z_{1n}^2) = 2nE(\hat{\phi}_1^2)/\sigma^2 = 1$. Since the right side is a part of a convergent series, the above probability approaches 0 as $m_n \to \infty$.

By application of Theorem A of Serfling (1970),

$$E[\xi_{jn}^2] \le \left( \frac{\log 2(2j+1)}{\log 2} \right)^2 (2j+1). \tag{2.10}$$

Therefore

$$
\begin{aligned}
P\left(\bigcup_{j=j_{n1}}^{j_{n3}} \sum_{l=1}^{j^2} \frac{\xi_{jn}}{j^2} > \frac{\epsilon}{2}\right) &\le \sum_{j=j_{n1}}^{j_{n3}} P\left(\frac{\xi_{jn}}{j^2} > \frac{\epsilon}{2}\right) \\
&\le \frac{4}{\epsilon^2} \sum_{j=j_{n1}}^{j_{n3}} \left(\frac{\log 2(2j+1)}{\log 2}\right)^2 \frac{(2j+1)}{j^4} \\
&\to 0 \quad \text{as} \quad m_n \to \infty.
\end{aligned}
$$

Now we have $P(A_n^c) \to 1$ as $m_n \to \infty$ and $n \to \infty$, hence (2.8) is established.

The second step is to show

$$
\max_{1 \le k \le m_n} \frac{1}{k} \sum_{j=1}^{k} |Z_{jn}| - \max_{1 \le k \le m_n} \frac{1}{k} \sum_{j=1}^{k} |Z_j| \xrightarrow{\mathrm{P}} 0, \tag{2.11}
$$

where $\lim_n Z_{jn} = Z_j$ and $Z_j$ is iid $N(0,1)$ for $i = 1, \dots, m_n$.

We apply Theorem 13.3 of Bhattacharya and Ranga Rao (1976) to the vector $(Z_{1n}, Z_{2n}, \dots, Z_{m_n n})$. Let

$$
P(A_n) = P\left((Z_{1n}, Z_{2n}, \dots, Z_{m_n n}) \in C_n\right)
$$

and

$$
P(A_n^*) = P\left((Z_1, Z_2, \dots, Z_{m_n}) \in C_n\right),
$$

where $C_n$ is a Borel subset of $m_n$ dimensional Euclidean space. The Berry-Esséen type result of Bhattacharya and Ranga Rao (1976) gives that

$$
\sup_{C_n} |P(A_n) - P(A_n^*)| \le a(m_n) m_n^2 E(\epsilon_1/\sigma)^4 / \sqrt{n}, \tag{2.12}
$$

where $a(m_n)$ is a positive constant that depends only on $m_n$. Since $m_n$ can be chosen to grow sufficiently slowly that $m_n^2 a(m_n)/\sqrt{n} \to 0$, we have

$$
|P(A_n) - P(A_n^*)| \to 0. \tag{2.13}
$$

This fact will be used in the following. Note that

$$
\left| \max_{1 \le k \le m_n} \frac{1}{k} \sum_{j=1}^{k} |Z_{jn}| - \max_{1 \le k \le m_n} \frac{1}{k} \sum_{j=1}^{k} |Z_j| \right| \le \max_{1 \le k \le m_n} \left| \frac{1}{k} \sum_{j=1}^{k} (|Z_{jn}| - |Z_j|) \right|
$$

$$
\begin{aligned}
&\leq \max_{1\leq k\leq m_n}\left|\frac{1}{k}\sum_{j=1}^{k}(|Z_{jn}-Z_j|)\right| \\
&\leq \max_{1\leq k\leq m_n}|Z_{kn}-Z_k| \\
&= |Z_{k_0n}-Z_{k_0}|\xrightarrow{\text{P}} 0
\end{aligned}
$$

for some $k_0 \in \{1,\ldots,m_n\}$.

Since $Z_{jn}\xrightarrow{\text{d}}Z_j$ by the central limit approximation and therefore $Z_{jn} - Z_j\xrightarrow{\text{d}}0$ which is equivalent to $Z_{jn}-Z_j\xrightarrow{\text{P}}0$ for $j=1,\ldots,m_n$, and letting $m_n\to\infty$ slowly enough to satisfy (2.13), the proof is complete.

**Corollary 2.1** If $\hat{\sigma}_n$ is a consistent estimator of $\sigma$, then

$$\max_{1\leq k\leq n-1}\frac{1}{k}\sum_{j=1}^{k}\frac{\sqrt{2n}|\hat{\phi}_j|}{\hat{\sigma}_n}\xrightarrow{\text{d}}\max_{k\geq 1}\frac{1}{k}\sum_{j=1}^{k}|Z_j|.$$

**Proof.** Note that

$$
\begin{aligned}
\max_{1\leq k\leq n-1}\frac{1}{k}\sum_{j=1}^{k}\frac{\sqrt{2n}|\hat{\phi}_j|}{\hat{\sigma}_n} &= \max_{1\leq k\leq n-1}\frac{1}{k}\sum_{j=1}^{k}\frac{\sqrt{2n}|\hat{\phi}_j|}{\sigma}\cdot\frac{\sigma}{\hat{\sigma}_n} \\
&= \max_{1\leq k\leq n-1}\left\{\frac{1}{k}\sum_{j=1}^{k}\frac{\sqrt{2n}|\hat{\phi}_j|}{\sigma}+\frac{1}{k}\sum_{j=1}^{k}\frac{\sqrt{2n}|\hat{\phi}_j|}{\sigma}\left(\frac{\sigma}{\hat{\sigma}_n}-1\right)\right\}.
\end{aligned}
$$

Since $\hat{\sigma}_n$ is consistent, it follows that

$$\max_{1\leq k\leq n-1}\frac{1}{k}\sum_{j=1}^{k}\frac{\sqrt{2n}|\hat{\phi}_j|}{\hat{\sigma}_n}\xrightarrow{\text{P}}\max_{k\geq 1}\frac{1}{k}\sum_{j=1}^{k}|Z_j|.$$

**Theorem 2.2** (Consistency of the test)

Let $k_0$ be the smallest $k$ such that $|\phi_k|>0$. If $\hat{\phi}_{k_0}\xrightarrow{\text{P}}\phi_{k_0}$, then

$$P\left(\max_{1\leq k\leq n-1}\frac{1}{k}\sum_{j=1}^{k}\frac{\sqrt{2n}|\hat{\phi}_j|}{\hat{\sigma}_n}>c_\alpha\right)\to 1 \quad \text{as} \quad n\to\infty$$

where $c_\alpha$ is the $\alpha$ critical value from the asymptotic distribution of $T_n$.

**Proof.** For all $n > k_0$,

$$\max_{1\le k\le n-1} \frac{1}{k}\sum_{j=1}^{k} \frac{\sqrt{2n}|\hat{\phi}_j|}{\hat{\sigma}_n} \ge \frac{1}{k_0}\sum_{j=1}^{k_0} \frac{\sqrt{2n}|\hat{\phi}_j|}{\hat{\sigma}_n} \ge \frac{1}{k_0}\frac{\sqrt{2n}|\hat{\phi}_{k_0}|}{\hat{\sigma}_n},$$

and therefore, as $n \to \infty$,

$$\begin{aligned} P\left(\max_{1\le k\le n-1} \frac{1}{k}\sum_{j=1}^{k} \frac{\sqrt{2n}|\hat{\phi}_j|}{\hat{\sigma}_n} > c_\alpha\right) &\ge P\left(\frac{1}{k_0}\frac{\sqrt{2n}|\hat{\phi}_{k_0}|}{\hat{\sigma}_n} > c_\alpha\right) \\ &\ge P\left(|\hat{\phi}_{k_0}| > k_0 c_\alpha \hat{\sigma}_n/\sqrt{2n}\right) \\ &\to 1 \quad \text{by the assumption.} \end{aligned}$$

## 3. SIMULATION STUDY

A numerical study was done to determine the critical values and to do power comparison in various models. The data were generated from the model

$$y_i = f(x_i) + \epsilon_i$$

where $\epsilon_i$'s are iid from $N(0,1)$. Small sample critical values were found by simulation from the empirical null distribution of the test statistic. We used 1000 sample sets of size 50. For a consistent estimator of the variance, the following nonparametric estimator was used in the simulation

$$\hat{\sigma}_n^2 = \frac{1}{2(n-1)}\sum_{i=1}^{n-1}(y_{i+1} - y_i)^2.$$

We compare $T_n$ with Buckley's (1992) test

$$T_B = n^{-2}\sum_{i=1}^{n}\left(\sum_{j=1}^{i}(Y_j - \bar{Y})\right)^2 / \hat{\sigma}_n^2,$$

where $\bar{Y} = \sum_{i=1}^{n} Y_i/n$. The models considered are:
(i) One change-point model

$$f(x) = \begin{cases} 0, & 0 \le x < 0.5 \\ \beta, & 0.5 \le x \le 1 \end{cases}$$

(ii) Two change-points model

$$f(x) = \begin{cases} 0, & 0 \leq x < 0.3 \\ \beta, & 0.3 \leq x < 0.7 \\ 0, & 0.7 \leq x \leq 1 \end{cases}$$

(iii) Smooth change model

$$f(x) = \cos(\pi j x), \quad 0 \leq x \leq 1.$$

Table 3.1 gives the empirical critical values of the test statistics. Table 3.2 shows power comparisons when the alternative model is true. The proposed test is more powerful unless there is one change-point.

**Table 3.1** Empirical Critical Values of $T_n$ and $T_B$ with $n = 50$ based on 10000 Repetitions

| | $T_n$ | $T_B$ |
|---|---|---|
| $\alpha = 0.05$ | 2.9005 | 0.4679 |
| $\alpha = 0.10$ | 2.4557 | 0.3643 |

**Table 3.2** Powers of $T_n$ and $T_B$ with $n = 50$ in 1000 Repetitions

| | $\alpha = 0.05$ | | $\alpha = 0.10$ | |
|---|---|---|---|---|
| model | $T_n$ | $T_B$ | $T_n$ | $T_B$ |
| one change-point ($\beta = 1.0$) | 0.850 | 0.865 | 0.923 | 0.928 |
| two change-points ($\beta = 1.0$) | 0.420 | 0.343 | 0.618 | 0.503 |
| smooth change ($j = 1$) | 0.993 | 0.993 | 0.998 | 0.998 |
| smooth change ($j = 3$) | 0.642 | 0.431 | 0.835 | 0.631 |
| smooth change ($j = 5$) | 0.161 | 0.102 | 0.434 | 0.208 |

## 4. CONCLUSION

The objective of this research is to develop powerful statistical tests to detect change when the data are independent. Fourier series and Kolmogorov-Smirnov test idea was used to derive a test statistic. The proof for the asymptotic distribution of the test can be extended to other orthogonal bases and unevenly spaced design points. The power study shows that when there is only one change-point, the power of the proposed test is as same as that of Buckley's cusum type test. But if the shape of change is cyclic or multiple step function, the proposed test is more powerful than the cusum type test.

## REFERENCES

( 1) Bhattacharya, R. J. and Ranga Rao, R. (1976). *Normal Approximation and Asymptotic Expansions*, Wiley, New York.

( 2) Buckley, M. J. (1991). Detecting a smooth signal: Optimality of cusum based procedures. *Biometrika*, **78**, 253–262.

( 3) Csörgó, M. and Horáth, L. (1988). Nonparametric methods for change-point problems. *Handbook of Statistics*, **7**, 403–425.

( 4) Durbin, J. and Knott, M. (1972). Components of Cramér-von Mises statistics I. *Journal of the Royal Statistical Society, B*, **34**, 290–307.

( 5) Eubank, R. and Hart, J. D. (1992). Testing goodness-of-fit in regression via order selection criteria. *Annals of Statistics*, **20**, 1412–1425.

( 6) Page, E. S. (1954). Continuous inspection schemes. *Biometrika*, **41**, 100–115.

( 7) Page, E. S. (1955). A test for a change in a parameter occurring at an unknown point. *Biometrika*, **44**, 248–252.

( 8) Serfling, R. J. (1970). Moment inequalities for the maximum cumulative sum. *Annals of Mathematical Statistics*, **41**, 1227–1234.

( 9) Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

(10) Wolfe, D. A. and Schechtman, E. (1984). Nonparametric statistical procedures for the shangepoint problem, *Journal of Statistical Planning and Inference*, **9**, 384–396.