A New Goodness-of-Fit Test for a Distribution by the Empirical Characteristic Function

M. Towhidi* and M. Salmanpour

Extended Abstract. Suppose \( n \) i.i.d. observations, \( X_1, \ldots, X_n \), are available from the unknown distribution \( F(\cdot) \), goodness-of-fit tests refer to tests such as

\[
H_0 : F(x) = F_0(x) \quad \text{against} \quad H_1 : F(x) \neq F_0(x).
\]

Some nonparametric tests such as the Kolmogorov–Smirnov test, the Cramer-Von Mises test, the Anderson-Darling test and the Watson test have been suggested by comparing empirical distribution, \( F_n(x) \), and the known distribution \( F_0(x) \).

The characteristic function is important in characterizing the probability distribution theoretically. Thus it have been expected that the empirical characteristic function, \( c_n(t) \), can be used for suggesting a goodness-of-fit test.

Let \( c(\cdot) \) denote the characteristic function of \( F(\cdot) \), i.e.,

\[
c(t) = E[\exp(itX)] = E[\cos(tX)] + iE[\sin(tX)],
\]

and similarly, \( c_0(\cdot) \) be the characteristic function of \( F_0(\cdot) \). The hypothesis \( H_0 \) is equivalent to \( H'_0 : c(t) = c_0(t) \), for any \( t \in \mathbb{R} \). \( c(t) \) can be estimated consistently by the empirical characteristic function (ECF), \( c_n(t) \), defined as

\[
c_n(t) = \frac{1}{n} \sum_{j=1}^{n} \exp(itX_j) = \frac{1}{n} \sum_{j=1}^{n} \cos(tX_j) + i \left\{ \frac{1}{n} \sum_{j=1}^{n} \sin(tX_j) \right\}.
\]

To obtain a consistent test, Fan (1996) proposed a test for \( H_0 \) on the basis of a quadratic measure between \( c_n(t) \) and \( c_0(t) \) evaluated at \( m \) points, \( t_1, \ldots, t_m \in \mathbb{R} \), where \( m \) is a positive integer.

* Corresponding author.
Let $t_m = (t_1, \ldots, t_m)$ be an arbitrary point in $\mathbb{R}^m$ and the vectors $Z_n(t_m)$, $Z_0(t_m)$ and $Z(t_m)$ are defined as:

$$Z_n(t_m) = \begin{pmatrix} \Re c_n(t_1) \\ \Re c_n(t_2) \\ \vdots \\ \Im c_n(t_1) \\ \Im c_n(t_2) \\ \vdots \\ \Im c_n(t_{m}) \end{pmatrix}, \quad Z_0(t_m) = \begin{pmatrix} \Re c_0(t_1) \\ \Re c_0(t_2) \\ \vdots \\ \Re c_0(t_1) \\ \Im c_0(t_1) \\ \vdots \\ \Im c_0(t_{m}) \end{pmatrix}, \quad Z(t_m) = \begin{pmatrix} \Re c(t_1) \\ \vdots \\ \Re c(t_1) \\ \Im c(t_1) \\ \vdots \\ \Im c(t_{m}) \end{pmatrix}$$

By using the principal components method, we introduce a new goodness-of-fit test which is more powerful than the other previous tests.

Suppose that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$ be the $k$ largest eigen values of $\Omega_0$ and $\beta_1, \ldots, \beta_k$ be the eigen vectors related to $\lambda_1, \ldots, \lambda_k$ respectively. We define a test statistic as

$$T_n^* = (Z_n(t_m) - Z_0(t_m))^T B_k \Lambda_k^{-1} B_k^T (Z_n(t_m) - Z_0(t_m)),$$

where $\Lambda_k$ is a $k \times k$ diagonal matrix with elements $\lambda_1, \ldots, \lambda_k$ and $B_k$ is a $2m \times k$ matrix which its columns are the corresponding eigen vectors, $\beta_1, \ldots, \beta_k$. We refer to $T_n^*$ as "Wald’s principal components" test statistic. In this paper we consider the asymptotic distribution of $T_n^*$ as $n \to +\infty$ and it is proved that $nT_n^* \to \chi^2_k$. Then the new test is compared via simulation to other omnibus tests for normality. It is shown that, the proposed test is more powerful.

**Keywords.** characteristic function; consistent test; eigen values; goodness-of-fit test; multivariate central limit theorem; principal components method.

Mina Towhidimahdi
Department of Statitic, College of Science, Shiraz University, Shiraz 71454, Iran
e-mail: mltowhidi@susc.ac.ir

Mahdi Salmanpour
Department of Statitic, College of Science, Shiraz University, Shiraz 71454, Iran
e-mail: mhr-salman@hotmail.com

---

The full version of the paper, in Persian, appears on pages 1-13.