Anderson-Darling type goodness-of-fit statistic based on a multifold integrated empirical distribution function

Satoshi Kuriki*[1] and Hsien-Kuei Hwang[2]

[1]The Institute of Statistical Mathematics, 10-3 Midoricho, Tachikawa, Tokyo 190-8562, Japan

[2]Institute of Statistical Science, Academia Sinica, Taipei 11529, Taiwan, R.O.C.

*Corresponding author: Satoshi Kuriki, e-mail: kuriki@ism.ac.jp

Abstract

An Anderson-Darling type goodness-of-fit statistic constructed from multifold integrated empirical distribution function is proposed. The proposed statistic is of an integral form whose integrand is a standardized square of m-fold integrated empirical distribution function. The empirical distribution functions is adjusted in advance so that it does not contain the components of m-th or lower degree polynomials. The proposed statistic is a natural extension of the goodness-of-fit statistic by Anderson and Darling (1952, *AMS*), which corresponds to the case $m = 0$. When $m = 1$, the proposed statistic has much statistical power to detect the discrepancy of dispersion of distribution. The Karhunen-Loève expansion of the limiting integrand process is obtained with Legendre eigenfunctions, and the limiting distribution of the proposed statistic is proved to be a weighted sum of chi-square random variables with the weights $1/\{k(k+1)\cdots(k+2m+1)\}$, $k = 1, 2, \ldots$. The explicit form of the Laplace transform of the limiting distribution without infinite product is derived. The relationship to a boundary-value problem is pointed out. Finally, it is mentioned that the similar type of extension is possible to Watson's (1961, *Biometrika*) statistic for testing uniformity of directional data.

Keywords: Boundary-value problem, Directional data, Karhunen-Lòeve expansion, Laplace transform, Legendre polynomial.

1. Anderson-Darling statistic and its extension

Let $X_1, \ldots, X_n$ be an i.i.d. sequence from the distribution F (the cumulative distribution function, cdf). We consider a goodness-of-fit test for testing $H_0 : F = G$ against $H_1 : F \neq G$, where G is a given cdf. When the distribution is continuous, we can let $G(x) = x$ (i.e., Unif$(0,1)$) without loss of generality. Let $F_n(x) = n^{-1} \sum_{i=1}^{n} \mathbb{1}(X_i \leq x)$ be the empirical distribution function. Test statistics are defined as measures of discrepancy between $F_n(x)$ and $G(x) = x$. Various statistics have been proposed. Among them, one of the most popular statistics is the one proposed by Anderson and Darling (1952).

The Anderson-Darling statistic is given as

$$A_n = \int_0^1 \frac{B_n(x)^2}{x(1-x)} dx, \quad \text{where} \;\; B_n(x) = \sqrt{n}(F_n(x) - x). \tag{1}$$

Here,

$$B_n(x) = \sqrt{n} \int_0^1 h^{(0)}(t; x) dF_n(t), \quad h^{(0)}(t; x) = \mathbb{1}(t \leq x) - x$$

is the inner product of a "template" $h^{(0)}(\cdot; x)$ (step function) and $dF_n(\cdot)$ "empirical density function". As an extension to the Anderson-Darling statistic, we propose a new class of test statistics by replacing $h^{(0)}(\cdot; x)$ with different types of "templates".

Note first that $\int_0^1 h^{(0)}(t; x) \cdot 1 \; dt = 0$. Let $h^{(1)}(\cdot; x)$ be a continuous and piecewise linear function with a break point at x such that $\int_0^1 h^{(1)}(t; x) \cdot (at + b) \, dt = 0$, $\forall a, b$. Let $h^{(2)}(\cdot; x)$ be a C^1 and piecewise quadratic function with a break point at x such that $\int_0^1 h^{(2)}(t; x) \cdot (at^2 + bt + c) \, dt = 0$, $\forall a, b, c$. In general, let $h^{(m)}(t; x)$ be of C^{m-1} and a piecewise polynomial of degree m with a break point at x such that

$$\textstyle\int_0^1 h^{(m)}(t; x) \cdot (\forall \, \text{polynomial of degree } m) \, dt = 0.$$

Such a function is concretely constructed as

$$h^{(m)}(t; x) = \frac{1}{m!}(x - t)^m \mathbb{1}(t \leq x) - \sum_{k=0}^{m} \int_0^x \frac{1}{m!}(x - u)^m L_k(u) du \times L_k(t), \tag{2}$$

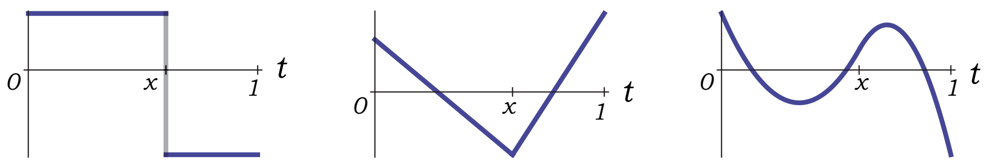

Figure 1: Templates $h^{(0)}(\cdot;x)$, $h^{(1)}(\cdot;x)$ and $h^{(2)}(\cdot;x)$.

where $L_k(\cdot)$ is the normalized Legendre polynomial of degree k on $(0,1)$ such that $\deg L_k(x)=k$, $\int_0^1 L_k(x)L_l(x)dx=\delta_{kl}$ (Figure 1).

Using this template function, we propose an extension of (1) as

$$A_n^{(m)}=\int_0^1 \frac{B_n^{(m)}(x)^2}{\{x(1-x)\}^{m+1}}dx, \quad \text{where} \quad B_n^{(m)}(x)=\sqrt{n}\int_0^1 h^{(m)}(t;x)dF_n(t).$$

Note that $B_n^{(m)}(x)$ is rewritten as $B_n^{(m)}(x)=\int_0^1 h^{(m)}(t;x)dB_n(t)$. $B_n^{(m)}(x)$ is constructed from m-fold integrals of empirical distribution function F_n. $A_n^{(0)}$ is the original Anderson-Darling statistic.

Remark 1 *The idea of use of the function $h^{(m)}(\cdot;x)$ $(m=0,1)$ is due to Hirotsu (1986). In the one-way ANOVA model, he proposed two test statistics χ^{*2} and $\chi^{\dagger 2}$ for testing "the equality of means" against "the monotonicity of mean profile", and for testing "the monotonicity of mean profile" against "the convexity of mean profile", respectively. These two are discrete analogs to $A_n^{(0)}$ and $A_n^{(1)}$.*

2. Limiting null distribution

To state results on distributions, we prepare several notations. Let

$$L_k^{(m)}(x)=(-1)^m\sqrt{\frac{(k-m)!}{(k+m)!}}\{x(1-x)\}^{m/2}\left(\frac{d}{dx}\right)^m L_k(x)$$

be the normalized associated Legendre functions on $(0,1)$. For each m, $\left\{L_k^{(m)}(\cdot)\right\}_{k=m,m+1,\ldots}$ forms an ONB. Let $W(\cdot)$ be the Winer process on $[0,1]$. Let $B(x)=W(x)-xW(1)$ be the Brownian bridge.

The following lemma follows from a weak convergence argument in L^2 (Section 1.8 and Example 1.8.6 of van der Vaart and Wellner (1961)). Let μ be any finite measure on $(0,1)$.

Lemma 1 *As* $n \to \infty$, $B_n^{(m)}(\cdot) \to^d B^{(m)}(\cdot)$ *in* $L^2(\mu)$, *where*

$$B^{(m)}(x) = \int_0^1 h^{(m)}(t;x)dB(t), \quad \textit{and}$$

$$A_n^{(m)} = \int_0^1 \frac{B_n^{(m)}(x)^2}{\{x(1-x)\}^{m+1}}dx \ \xrightarrow{d} \ A^{(m)} := \int_0^1 \frac{B^{(m)}(x)^2}{\{x(1-x)\}^{m+1}}dx.$$

Theorem 1 (Karhunen-Loève expansion)

$$\frac{B^{(m)}(x)}{\{x(1-x)\}^{(m+1)/2}} = \sum_{k=m+1}^{\infty} \sqrt{\frac{(k-m-1)!}{(k+m+1)!}} L_k^{(m+1)}(x)\, \xi_k, \tag{3}$$

$$\textit{where} \quad \xi_k = \int_0^1 L_k(t)dB(t), \quad \textit{i.i.d. } N(0,1).$$

The convergence is uniformly in x *with probability one.*

By preparing a differential equation that the Legendre polynomials satisfy, the expansion (3) is derived with formal manipulations. This manipulation is validated by Mercer's theorem and continuity of sample path (Theorem 3.8 of Adler (1990)).

Corollary 1 (Limiting null distribution of $A_n^{(m)}$)

$$A^{(m)} = \sum_{k=m+1}^{\infty} \frac{(k-m-1)!}{(k+m+1)!}\xi_k^2, \quad \xi_k^2 \sim \chi^2(1) \quad \textit{i.i.d.} \tag{4}$$

3. Laplace transform

The Laplace transform (moment generating function) of $A^{(m)} = \sum_{k=1}^{\infty} \lambda_k^{-1}\xi_k^2$, $\lambda_k = k(k+1)\cdots(k+2m+1)$, is easily given by

$$E\big[e^{-sA^{(m)}}\big] = \prod_{k=1}^{\infty}\left(1 - \frac{2s}{\lambda_k}\right)^{-\frac{1}{2}}.$$

However, the infinite product is not convenient for further analyses such as analyzing tail behavior or numerical calculations. Here, we provide a more neat expression.

Theorem 2 *Let $x_j(s)$ $(j = 0, 1, \ldots, 2m+1)$ be the solution of $x(x+1)\cdots(x+2m+1) - 2s = 0$. Then*

$$E\big[e^{-sA^{(m)}}\big] = \prod_{j=0}^{2m+1} \left(\frac{\Gamma(1 - x_j(s))}{j!}\right)^{\frac{1}{2}}.$$

When $m = 0$, $\lambda_k = k(k+1)$, and $E\big[e^{-sA^{(0)}}\big] = \sqrt{2\pi s/(-\cos\frac{\pi}{2}\sqrt{1+8s})}$ (Anderson and Darling, 1952). When $m = 1$, $\lambda_k = k(k+1)(k+2)(k+3)$, and

$$E\big[e^{-sA^{(1)}}\big] = \frac{\pi s}{\sqrt{3\cos(\frac{\pi}{2}\sqrt{5-4\sqrt{1+2s}})\cos(\frac{\pi}{2}\sqrt{5+4\sqrt{1+2s}})}}.$$

When $m = 2$, $\lambda_k = k(k+1)(k+2)(k+3)(k+4)(k+5)$, and

$$E\big[e^{-sA^{(2)}}\big] = \frac{(\pi s)^{3/2}}{\sqrt{-4320\cos(\pi\sqrt{\eta_1})\cosh(\pi\sqrt{\eta_2})\cosh(\pi\sqrt{\eta_3})}},$$

where $\eta_1 = \frac{1}{12\eta}\big(4\eta^2 + 35\eta + 112\big)$, $\eta_2 = \frac{1}{12\eta}\big(4e^{-\pi i/3}\eta^2 - 35\eta + 112e^{\pi i/3}\big)$, $\eta_3 = \frac{1}{12\eta}\big(4e^{\pi i/3}\eta^2 - 35\eta + 112e^{-\pi i/3}\big)$ with $\eta = \sqrt[3]{27s + 80 + 3\sqrt{81s^2 + 480s - 1728}}$.

4. Statistical power

By following the proof of Theorem 1, we have the sample counterpart of the KL-expansion (3):

$$\frac{B_n^{(m)}(x)}{\{x(1-x)\}^{(m+1)/2}} = \sum_{k \geq m+1} \sqrt{\frac{(k-m-1)!}{(k+m+1)!}}\, L_k^{(m+1)}(x)\, \widehat{\xi}_k,$$

$$\text{where} \quad \widehat{\xi}_k = \int_0^1 L_k(x)\,dB_n(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} L_k(X_i).$$

The sample counterpart of (4) in Corollary 1 is

$$A_n^{(m)} = \sum_{k=m+1}^{\infty} \frac{(k-m-1)!}{(k+m+1)!}\, \widehat{\xi}_k^{\,2}.$$

First two components of $\widehat{\xi}_k$'s are $\widehat{\xi}_1 = \sqrt{12n}\,m_1$ and $\widehat{\xi}_2 = 6\sqrt{5n} \times (m_2 - 1/12)$ where $m_k = n^{-1}\sum_{i=1}^{n}(X_i - 1/2)^k$. From this observation, we see that $A_n^{(0)} =$

$\widehat{\xi}_1^2/2+\cdots$ has much power for mean-shift alternative, and $A_n^{(1)}=\widehat{\xi}_2^2/24+\cdots$ has much power for dispersion-change alternative.

5. Summary and supplements

We have proposed a class of goodness-of-test statistics $A_n^{(m)}$ $(m=0,1,\ldots)$ based on m-fold integrated empirical distribution function as an extension to the Anderson-Darling statistic. The limiting null distribution $A^{(m)}$ is explicitly derived as weighted infinite sums of chi-square random variables (without solving eigenvalue problems). Using a newly proved theorem, we provide Laplace transforms of $A^{(m)}$ for $m=0,1,2$ without using infinite product. The statistic $A_n^{(0)}$ has much power in direction to mean-shift, and the statistic $A_n^{(1)}$ has much power in direction to dispersion-change.

Similar extensions to Watson's (1961) statistic based on m-fold integrals of F_n is possible (see Henze and Nikitin (2002) when $m=1$). The template function $h^{(m)}(\cdot;x)$ in (2) is regarded as a Green function of a boundary-value problem (see Chang and Ha (2001) for the Watson statistic).

The authors thank Yoichi Nishiyama for his valuable comments.

References

Adler, R. (1990). *Introduction to Continuity Extreme and Related Topics*. IMS Lecture Notes–Monograph Series **12**, IMS, Hayward.

Anderson, T. W. and Darling, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math. Statist.*, **23** (2), 193–212.

Chang, C.-H. and Ha, C.-W. (2001). The Green functions of some boundary value problems via the Bernoulli and Euler polynomials. *Arch. Math.*, **76** (5), 360–365.

Henze, N. and Nikitin, Ya. Yu. (2002). Watson-type goodness-of-fit tests based on the integrated empirical process. *Math. Methods. Stat.*, **11** (2), 183–202.

Hirotsu, C. (1986). Cumulative chi-squared statistic as a tool for testing goodness of fit. *Biometrika*, **73** (1), 165–173.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer, New York.

Watson, G. S. (1961). Goodness-of-fit tests on a circle. *Biometrika*, **48** (1,2), 109–114.