Записки научных
семинаров ПОМИ
Том 501, 2021 г.

B. Ebner, N. Henze

BAHADUR EFFICIENCIES OF THE EPPS–PULLEY TEST FOR NORMALITY

Abstract. The test for normality suggested by Epps and Pulley [9] is a serious competitor to tests based on the empirical distribution function. In contrast to the latter procedures, it has been generalized to obtain a genuine affine invariant and universally consistent test for normality in any dimension. We obtain approximate Bahadur efficiencies for the test of Epps and Pulley, thus complementing recent results of Milošević et al. (see [15]). For certain values of a tuning parameter that is inherent in the Epps–Pulley test, this test outperforms each of its competitors considered in [15], over the whole range of six close alternatives to normality.

§1. Introduction

The purpose of this article is to derive Bahadur efficiencies for the test of normality proposed by Epps and Pulley [9], thus complementing recent results of Milošević et al. [15], who confined their study to tests of normality based on the empirical distribution function. To be specific, suppose $X_1, X_2, \ldots$ is a sequence of independent and identically distributed (i.i.d.) copies of a random variable X that has an absolutely continuous distribution with respect to Lebesgue measure. To test the hypothesis H_0 that the distribution of X is some unspecified non-degenerate normal distribution, Epps and Pulley [9] proposed to use the test statistic

$$T_{n,\beta} = n \int\limits_{-\infty}^{\infty} \left| \psi_n(t) - \mathrm{e}^{-t^2/2} \right|^2 \varphi_\beta(t)\,\mathrm{d}t.$$

Here, $\psi_n(t) = n^{-1} \sum\limits_{j=1}^{n} \exp\left(\mathrm{i}tY_{n,j}\right)$ is the empirical characteristic function of the so-called *scaled residuals* $Y_{n,1}, \ldots, Y_{n,n}$, where $Y_{n,j} = S_n^{-1}(X_j - \overline{X}_n)$,

---

Key words and phrases: test for normality, empirical characteristic function, Bahadur efficiency, eigenvalues of integral operators.

Primary 62F05; Secondary 62G20.

$j = 1,\dots,n$, and $\overline{X}_n = n^{-1}\sum_{j=1}^{n} X_j$, $S_n^2 = n^{-1}\sum_{j=1}^{n}(X_j - \overline{X}_n)^2$ are the sample mean and the sample variance of $X_1,\dots,X_n$, respectively, and $\beta > 0$ is a so-called *tuning parameter*. Moreover,

$$\varphi_\beta(t) = \frac{1}{\beta\sqrt{2\pi}}\exp\Big(-\frac{t^2}{2\beta^2}\Big), \quad t \in \mathbb{R},$$

is the density of the centred normal distribution with variance β^2. A closed-form expression of $T_{n,\beta}$ that is amenable to computational purposes is

$$T_{n,\beta} = \frac{1}{n}\sum_{j,k=1}^{n}\exp\Big(-\frac{\beta^2}{2}\big(Y_{n,j}-Y_{n,k}\big)^2\Big) - \frac{2}{\sqrt{1+\beta^2}}\sum_{j=1}^{n}\exp\Big(-\frac{\beta^2 Y_{n,j}^2}{2(1+\beta^2)}\Big) + \frac{n}{\sqrt{1+2\beta^2}}. \tag{1.1}$$

Epps and Pulley did not obtain neither the limit null distribution of $T_{n,\beta}$ as $n \to \infty$ nor the consistency of a test for normality that rejects H_0 for large values of $T_{n,\beta}$. Their procedure, however, turned out to be a serious competitor to the classical tests of Shapiro–Wilk, Shapiro–Francia and Anderson–Darling in simulation studies (see [3]). In the special case $\beta = 1$, Baringhaus and Henze [5] generalized the approach of Epps and Pulley to obtain a genuine test of multivariate normality, and they derived the limit null distribution of $T_{n,1}$. Moreover, they proved the consistency of the test of Epps and Pulley against each alternative to normality having a finite second moment. The latter restriction was removed by S. Csörgő [6]. By an approach different from that adopted in [5], Henze and Wagner [12] obtained both the limit null distribution and the limit distribution of $T_{n,\beta}$ under contiguous alternatives to normality. Under fixed alternatives to normality, the limit distribution of $T_{n,\beta}$ is normal, as elaborated by [4] in much greater generality for weighted L^2-statistics. For more information on $T_{n,\beta}$, especially on the role of the tuning parameter β, see [7, Section 2.2].

Notice that $T_{n,\beta}$ is invariant with respect to affine transformations $X_j \mapsto aX_j + b$, where $a, b \in \mathbb{R}$ and $a \neq 0$. Hence, under H_0, both the finite-sample and the asymptotic distribution of $T_{n,\beta}$ do not depend on the parameters μ and σ^2 of the underlying normal distribution $\mathrm{N}(\mu,\sigma^2)$. Under H_0, we will thus assume $\mu = 0$ and $\sigma^2 = 1$. The rest of the paper unfolds as follows: In Section 2, we revisit the notion of approximate Bahadur efficiency. Sections 3 and 4 deal with stochastic limits and local Bahadur slopes, and Section 5 tackles an eigenvalue problem connected with the limit null distribution of the test statistic. The final section 6 contains results regarding local approximate Bahadur efficiencies of the Epps–Pulley test for the six close alternatives considered in [15] and a wide spectrum of values of the tuning parameter β.

§2. Approximate Bahadur efficiency

There are several options to compare different tests for the same testing problem as the sample size n tends to infinity, see [16]. One of these options is asymptotic efficiency due to Bahadur (see [1]). This notion of asymptotic efficiency requires knowledge of the large deviation function of the test statistic. Apart from the notable exception given in [18], such knowledge, however, is hitherto not available for statistics that contain estimated parameters, like $T_{n,\beta}$ given in (1.1). To circumvent this drawback, one usually employs the so-called *approximate* Bahadur efficiency, which only requires results on the tail behavior of the limit distribution of the test statistic under the null hypothesis. To be more specific with respect to the title of this paper, let $X, X_1, X_2, \ldots$ be a sequence of i.i.d. random variables, where the distribution of X depends on a real-valued parameter $\vartheta \in \Theta$, where Θ denotes the parameter space, and only the case $\vartheta = 0$ corresponds to the case that the distribution of X is standard normal. Suppose, $(S_n)_{n\geqslant 1}$, where $S_n = S_n(X_1, \ldots, X_n)$, is a sequence of test statistics of the hypothesis $H_0 : \vartheta = 0$ against the alternative H_1: $\vartheta \in \Theta \setminus \{0\}$. Furthermore, suppose that rejection of H_0 is for large values of S_n. The sequence (S_n) is called a *standard sequence*, if the following conditions hold (see, e.g., [16, p. 10], or [8, p. 3427]):

- There is a continuous distribution function G such that, for $\vartheta = 0$,
$$\lim_{n\to\infty} \mathbb{P}_0(S_n \leqslant x) = G(x), \quad x \in \mathbb{R}. \tag{2.1}$$
- There is a constant a_S, $0 < a_S < \infty$, such that
$$\log(1 - G(x)) = -\frac{a_S x^2}{2}(1 + o(1)) \text{ as } x \to \infty. \tag{2.2}$$
- There is a real-valued function $b_S(\vartheta)$ on $\Theta \setminus \{0\}$, with $0 < b_S(\vartheta) < \infty$, such that, for each $\vartheta \in \Theta \setminus \{0\}$,
$$\frac{S_n}{\sqrt{n}} \xrightarrow{\mathbb{P}_\vartheta} b_S(\vartheta). \tag{2.3}$$

Then the so-called *approximate Bahadur slope*

$$c_S^*(\vartheta) = a_S \cdot b_S^2(\vartheta), \quad \vartheta \in \Theta \setminus \{0\},$$

is a measure of approximate Bahadur efficiency. Usually, it is true that $c_S^*(\vartheta) \sim \ell(S) \cdot \vartheta^2$ as $\vartheta \to 0$. In this case $\ell(S)$ is called the *local* (*approximate*) *index* of the sequence (S_n). We will see that the sequence (S_n), where $S_n := \sqrt{T_{n,\beta}}$, is a standard sequence. To this end, we will derive the stochastic limit of $T_{n,\beta}/n$ for a general alternative in Section 3. In Section 4, we will specialize this stochastic limit for local alternatives, and we will derive the local index for the Epps–Pulley test statistic.

§3. Stochastic limit of $T_{n,\beta}/n$

To calculate the asymptotic Bahadur efficiency of the test of Epps and Pulley, we need the following result.

Theorem 3.1. *Suppose that* $\mathbb{E}(X^2) < \infty$. *Then*

$$\frac{T_{n\beta}}{n} \xrightarrow{\mathbb{P}} \mathbb{E}\left[\exp\left(-\frac{\beta^2(Y_1 - Y_2)^2}{2}\right)\right] - \frac{2}{\sqrt{1+\beta^2}}\mathbb{E}\left[\exp\left(-\frac{\beta^2 Y_1^2}{2(1+\beta^2)}\right)\right] + \frac{1}{\sqrt{1+2\beta^2}}.$$

Here, $\xrightarrow{\mathbb{P}}$ *denotes convergence in probability, and* $Y_j = (X_j - \mu)/\sigma$, $j \geqslant 1$, *where* $\mu = \mathbb{E}(X)$ *and* $\sigma^2 = \mathbb{V}(X)$.

Proof. From (1.1), we have

$$\frac{T_{n,\beta}}{n} = \frac{1}{n^2}\sum_{j,k=1}^{n} \exp\left(-\frac{\beta^2}{2}\left(\frac{X_j - X_k}{S_n}\right)^2\right)$$

$$- \frac{2}{\sqrt{1+\beta^2}} \cdot \frac{1}{n}\sum_{j=1}^{n} \exp\left(-\frac{\beta^2}{2(1+\beta^2)}\left(\frac{X_j - \overline{X}_n}{S_n}\right)^2\right) + \frac{1}{\sqrt{1+2\beta^2}}$$

$$=: A_{n,1} - \frac{2}{\sqrt{1+\beta^2}} \cdot A_{n,2} + \frac{1}{\sqrt{1+2\beta^2}}$$

(say). By symmetry, it follows that

$$\mathbb{E}\big(A_{n,1}\big) = \frac{1}{n} + \frac{n-1}{n} \cdot \mathbb{E}\Big[\exp\Big(-\frac{\beta^2}{2}\Big(\frac{X_1 - X_2}{S_n}\Big)^2\Big],$$

$$\mathbb{E}\big(A_{n,2}\big) = \mathbb{E}\Big[\exp\Big(-\frac{\beta^2}{2(1+\beta^2)}\Big(\frac{X_1 - \overline{X}_n}{S_n}\Big)^2\Big)\Big].$$

Since $\overline{X}_n \to \mu$ and $S_n \to \sigma$ almost surely as $n \to \infty$ by the strong law of large numbers, it follows from Lebesgue's dominated convergence theorem that

$$\lim_{n\to\infty} \mathbb{E}\big(A_{n,1}\big) = \mathbb{E}\Big[\exp\Big(-\frac{\beta^2(Y_1 - Y_2)^2}{2}\Big)\Big],$$

$$\lim_{n\to\infty} \mathbb{E}\big(A_{n,2}\big) = \mathbb{E}\Big[\exp\Big(-\frac{\beta^2 Y_1^2}{2(1+\beta^2)}\Big)\Big],$$

and thus the expectation of $T_{n,\beta}/n$ converges to the stochastic limit figuring in Theorem 3.1. Likewise, the variance of $T_{n,\beta}/n$ converges to zero. □

§4. Local Bahadur slopes

As was done in Milošević et al. [15], we now assume that $\mathcal{G} = \{G(x;\vartheta)\}$ is a family of distribution functions (DF's) with densities $g(x;\vartheta)$, such that $\vartheta = 0$ corresponds to the standard normal DF Φ and density φ, and each of the distributions for $\vartheta \neq 0$ is non-normal. Moreover, we assume that the regularity assumptions WD in [17] are satisfied. If $X, X_1, X_2, \ldots,$ are i.i.d. random variables with DF $G(\cdot;\vartheta)$, we have to consider the stochastic limit figuring in Theorem 3.1 as a function of ϑ and expand this function at $\vartheta = 0$. To this end, let

$$\gamma = \frac{\beta^2}{2}, \;\; \delta = \frac{\beta^2}{2(1+\beta^2)}. \tag{4.1}$$

Then, putting

$$\mu(\vartheta) = \int x g(x;\vartheta)\,\mathrm{d}x,$$

$$\sigma^2(\vartheta) = \int x^2 g(x;\vartheta)\,\mathrm{d}x - \mu^2(\vartheta),$$

Theorem 3.1 yields

$$\frac{T_{n,\beta}}{n} \xrightarrow{\mathbb{P}_\vartheta} b_{T_\beta}(\vartheta),$$

where $\xrightarrow{\mathbb{P}_\vartheta}$ denotes convergence in probability under the true parameter ϑ, and

$$b_{T_\beta}(\vartheta) = \iint \exp\left(-\frac{\gamma(x-y)^2}{\sigma^2(\vartheta)}\right) g(x;\vartheta)g(y;\vartheta)\,\mathrm{d}x\,\mathrm{d}y \tag{4.2}$$

$$- \frac{2}{\sqrt{1+\beta^2}} \int \exp\left(-\frac{\delta(x-\mu(\vartheta))^2}{\sigma^2(\vartheta)}\right) g(x;\vartheta)\,\mathrm{d}x + \frac{1}{\sqrt{1+2\beta^2}}. \tag{4.3}$$

Here and in what follows, each unspecified integral is over $\mathbb{R}$.

Notice that $b_{T_\beta}(0) = 0$. We have to find the quadratic (first non-vanishing) term in the Taylor expansion of b_{T_β} around zero, i.e, we look for some (local index) $\Delta_\beta > 0$ such that

$$b_{T_\beta}(\vartheta) = \Delta_\beta\,\vartheta^2 + o(\vartheta^2) \quad \text{as } \vartheta \to 0.$$

Writing $g'_\vartheta(x;\vartheta)$, $g''_\vartheta(x;\vartheta)$ for derivatives of $g(x;\vartheta)$ with respect to ϑ, we have

$$g(x;\vartheta) = \varphi(x) + \vartheta \cdot g'_\vartheta(x;0) + \frac{\vartheta^2}{2} g''_\vartheta(x;0) + O(\vartheta^3)$$

and thus – since $\mu(0) = 0$ and $\sigma^2(0) = 1$ –

$$\mu(\vartheta) = \vartheta \int x g'_\vartheta(x;0)\,\mathrm{d}x + \frac{\vartheta^2}{2} \int x g''_\vartheta(x;0)\,\mathrm{d}x + O(\vartheta^3),$$

$$\sigma^2(\vartheta) = 1 + \vartheta \int x^2 g'_\vartheta(x;0)\,\mathrm{d}x + \frac{\vartheta^2}{2} \int x^2 g''_\vartheta(x;0)\,\mathrm{d}x - \mu(\vartheta)^2 + O(\vartheta^3).$$

Consequently, putting

$$\mu_1 := \mu'(0),\ \mu_2 := \mu''(0),\ \sigma_1 := (\sigma^2)'(0),\ \sigma_2 := (\sigma^2)''(0) \tag{4.4}$$

for the sake of brevity, it follows that

$$\mu_1 = \int x g'_\vartheta(x;0)\,\mathrm{d}x, \quad \mu_2 = \int x g''_\vartheta(x;0)\,\mathrm{d}x,$$

$$\sigma_1 = \int x^2 g'_\vartheta(x;0)\,\mathrm{d}x, \quad \sigma_2 = \int x^2 g''_\vartheta(x;0)\,\mathrm{d}x - 2\mu'(0)^2.$$

To tackle the integral that figures in (4.2), notice that

$$g(x;\vartheta)g(y;\vartheta) = \varphi(x)\varphi(y) + \vartheta\big[g'_\vartheta(x;0)\varphi(y) + g'_\vartheta(y;0)\varphi(x)\big]$$

$$+ \vartheta^2\Big[\frac{1}{2}g''_\vartheta(x;0)\varphi(y) + \frac{1}{2}g''_\vartheta(y;0)\varphi(x) + g'_\vartheta(x;0)g'_\vartheta(y;0)\Big] + O(\vartheta^3).$$

Moreover, it follows from a geometric series expansion that

$$\frac{1}{\sigma^2(\vartheta)} = 1 - \vartheta\sigma_1 + \vartheta^2\Big[\sigma_1^2 - \frac{\sigma_2}{2}\Big] + O(\vartheta^3) \tag{4.5}$$

(say). From an expansion of the exponential function, we thus obtain

$$\exp\Big(-\frac{\gamma(x-y)^2}{\sigma^2(\vartheta)}\Big) = \mathrm{e}^{-\gamma(x-y)^2}\Big[1 + \vartheta\sigma_1\gamma(x-y)^2 + \vartheta^2\Big\{\frac{1}{2}\sigma_1^2\gamma^2(x-y)^4 - \Big(\sigma_1^2 - \frac{\sigma_2}{2}\Big)\gamma(x-y)^2\Big\}\Big] + O(\vartheta^3).$$

Using

$$\iint \mathrm{e}^{-\gamma(x-y)^2}(x-y)^{2k}\varphi(x)\varphi(y)\,\mathrm{d}x\mathrm{d}y = \frac{4^k\Gamma(k+1/2)}{\sqrt{\pi}(4\gamma+1)^{k+1/2}}, \quad k = 0,1,2,$$

$$\int \mathrm{e}^{-\gamma(x-y)^2}\varphi(x)\,\mathrm{d}x = \frac{1}{\sqrt{1+\beta^2}}\cdot \mathrm{e}^{-\delta y^2}$$

$$\int \mathrm{e}^{-\gamma(x-y)^2}(x-y)^2\varphi(y)\,\mathrm{d}y = \mathrm{e}^{-\delta x^2}\cdot\frac{x^2+\beta^2+1}{(1+\beta^2)^{5/2}},$$

and putting

$$D_0 = \iint \mathrm{e}^{-\gamma(x-y)^2} g'_\vartheta(x;0)g'_\vartheta(y;0)\,\mathrm{d}x\mathrm{d}y, \tag{4.6}$$

$$J_{1,k} = \int \mathrm{e}^{-\delta x^2}x^k g'_\vartheta(x;0)\,\mathrm{d}x, \quad k=0,1,2; \qquad J_2 = \int \mathrm{e}^{-\delta x^2} g''_\vartheta(x;0)\,\mathrm{d}x, \tag{4.7}$$

some algebra gives

$$\begin{aligned}&\iint \exp\Big(-\frac{\gamma(x-y)^2}{\sigma^2(\vartheta)}\Big) g(x;\vartheta)g(y;\vartheta)\,\mathrm{d}x\,\mathrm{d}y\\ &= \frac{1}{\sqrt{1+2\beta^2}} + \vartheta\Big\{\frac{2J_{1,0}}{\sqrt{1+\beta^2}} + \frac{2\sigma_1\gamma}{(1+2\beta^2)^{3/2}}\Big\} \qquad (4.8)\\ &+\vartheta^2\Big\{\frac{J_2}{\sqrt{1+\beta^2}} + D_0 + \frac{2\sigma_1\gamma\big(J_{1,2}+(\beta^2+1)J_{1,0}\big)}{(1+\beta^2)^{5/2}} + \frac{\beta^2\big((2\beta^2+1)\sigma_2-(\beta^2+2)\sigma_1^2\big)}{2(1+2\beta^2)^{5/2}}\Big\}.\end{aligned}$$

As for the integral figuring in (4.3), we use (4.5). Neglecting terms that are of order $O(\vartheta^3)$, straightforward but tedious calculations give

$$\exp\Big(-\frac{\delta(x-\mu(\vartheta))^2}{\sigma^2(\vartheta)}\Big) = \mathrm{e}^{-\delta x^2}\cdot\Big\{1 + \delta\vartheta U(x) + \frac{\delta\vartheta^2}{2}V(x)\Big\} + O(\vartheta^3),$$

where – recalling (4.4) –

$$U(x) = \sigma_1 x^2 + 2\mu_1 x,$$
$$V(x) = \delta\sigma_1^2 x^4 + 4\delta\mu_1\sigma_1 x^3 + \big(4\delta\mu_1^2 - 2\sigma_1^2 + \sigma_2\big)x^2 - (4\mu_1\sigma_1 - 2\mu_2)x - 2\mu_1^2.$$

Thus,

$$\begin{aligned}
&\int \exp\left(-\frac{\delta(x-\mu(\vartheta))^2}{\sigma^2(\vartheta)}\right) g(x;\vartheta)\,\mathrm{d}x \\
&= \int \mathrm{e}^{-\delta x^2}\Big(1 + \delta\vartheta U(x) + \frac{\delta\vartheta^2}{2}V(x)\Big)\varphi(x)\,\mathrm{d}x \\
&+ \vartheta \int \mathrm{e}^{-\delta x^2}\Big(1 + \delta\vartheta U(x) + \frac{\delta\vartheta^2}{2}V(x)\Big)g'_\vartheta(x;0)\,\mathrm{d}x \\
&+ \frac{\vartheta^2}{2} \int \mathrm{e}^{-\delta x^2}\Big(1 + \delta\vartheta U(x) + \frac{\delta\vartheta^2}{2}V(x)\Big)g''_\vartheta(x;0)\mathrm{d}x \\
&+ O(\vartheta^3) = I_1(\vartheta) + \vartheta I_2(\vartheta) + \frac{\vartheta^2}{2} I_3(\vartheta) + O(\vartheta^3)
\end{aligned}$$

(say). We have

$$\begin{aligned}
I_1(\vartheta) &= \frac{1}{(1+2\delta)^{1/2}} + \vartheta \cdot \frac{\delta\sigma_1}{(1+2\delta)^{3/2}} \\
&+ \vartheta^2\left[\frac{3\delta^2\sigma_1^2}{2(1+2\delta)^{5/2}} + \frac{\delta(4\delta\mu_1^2 - 2\sigma_1^2 + \sigma_2)}{2(1+2\delta)^{3/2}} - \frac{\delta\mu_1^2}{(1+2\delta)^{1/2}}\right].
\end{aligned}$$

Furthermore,

$$I_2(\vartheta) = \int \mathrm{e}^{-\delta x^2} g'_\vartheta(x;0)\mathrm{d}x + \delta\vartheta \int \mathrm{e}^{-\delta x^2}\big(\sigma_1^2 x^2 + 2\mu_1 x\big)g'_\vartheta(x;0)\mathrm{d}x + O(\vartheta^2),$$
$$I_3(\vartheta) = \int \mathrm{e}^{-\delta x^2} g''_\vartheta(x;0)\,\mathrm{d}x + O(\vartheta).$$

Recalling (4.7), we thus obtain – apart from a term which is $O(\vartheta^3)$ –

$$\begin{aligned}
&\int \exp\left(-\frac{\delta(x-\mu(\vartheta))^2}{\sigma^2(\vartheta)}\right) g(x;\vartheta)\,\mathrm{d}x \\
&= \frac{1}{(1+2\delta)^{1/2}} + \vartheta\left[\frac{\delta\sigma_1}{(1+2\delta)^{3/2}} + J_{1,0}\right] \qquad (4.9) \\
&+ \vartheta^2\left[\frac{J_2}{2} + \delta\sigma_1^2 J_{1,2} + 2\delta\mu_1 J_{1,1} - \frac{\delta\left(\left(\delta + \frac{1}{2}\right)(\sigma_2 - 2\mu_1^2) - \left(\frac{\delta}{2} + 1\right)\sigma_1^2\right)}{(2\delta+1)^{5/2}}\right].
\end{aligned}$$

Upon combining (4.8) and (4.9) and recalling (4.1), $b_{T_\beta}(\vartheta)$ figuring in (4.2), (4.3) takes the form

$$b_{T_\beta}(\vartheta) = \Delta_\beta \vartheta^2 + O(\vartheta^3) \text{ as } \vartheta \to 0,$$

where

$$\Delta_\beta = D_0 + \frac{\beta^2}{(\beta^2+1)^{5/2}}\Big(\Big(\Big(J_{1,0} - J_{1,2}\Big)\sigma_1 - 2J_{1,1}\mu_1\Big)\beta^2 + J_{1,0}\sigma_1 - 2J_{1,1}\mu_1\Big) + \frac{\beta^2}{(2\beta^2+1)^{5/2}}\Big(\Big(2\mu_1^2 + \frac{3}{4}\sigma_1^2\Big)\beta^2 + \mu_1^2\Big),$$

and D_0 and $J_{1,0}, J_{1,1}, J_{1,2}$ are defined in (4.6) and (4.7), respectively.

§5. Approximations to solutions of the eigenvalue problem

We now turn to the conditions (2.1) and (2.2). The limit null distribution of $T_{n,\beta}$, as $n \to \infty$, is given by the distribution of

$$T_\beta := \int_{-\infty}^{\infty} Z^2(t)\, \varphi_\beta(t)\,\mathrm{d}t.$$

Here, Z is a centred Gaussian random element of the Fréchet space of continuous real-valued functions having covariance kernel $K(s,t) = \mathbb{E}[Z(s)Z(t)]$, where

$$K(s,t) = \exp\Big(-\frac{(s-t)^2}{2}\Big) - \Big(1 + st + \frac{(st)^2}{2}\Big)\exp\Big(-\frac{s^2+t^2}{2}\Big),\ s,t \in \mathbb{R} \quad (5.1)$$

(see Theorem 2.1 and Theorem 2.2 of [12]). In fact, Z may also be regarded as a Gaussian random element of the separable Hilbert space L^2 (say) of (equivalence classes of) functions that are square integrable with respect to $\varphi_\beta(t)\mathrm{d}t$. The distribution of T_β is that of $\sum_{j=1}^{\infty} \lambda_j(\beta) N_j^2$, where $N_1, N_2, \ldots$ is a sequence of i.i.d. standard normal random variables, and $\lambda_1(\beta), \lambda_2(\beta), \ldots$ is the sequence of positive eigenvalues of the integral operator $\mathcal{K}$ on L^2 defined by

$$\mathcal{K} : L^2 \to L^2, \quad f \mapsto \mathcal{K}f(s) = \int_{-\infty}^{\infty} K(s,t) f(t) \varphi_\beta(t)\,\mathrm{d}t, \quad s \in \mathbb{R}.$$

Since S_n figuring in (2.1) equals $\sqrt{T_{n,\beta}}$, the function G is the distribution function of $\widetilde{Z} := \left(\sum_{j=1}^{\infty} \lambda_j(\beta) N_j^2\right)^{1/2}$. From [21], we thus have

$$\log(1 - G(x)) = \log \mathbb{P}(\widetilde{Z} > x) = \log \mathbb{P}(\widetilde{Z}^2 > x^2) \sim -\frac{x^2}{2\lambda_1(\beta)} \text{ as } x \to \infty,$$

where $\lambda_1(\beta)$ denotes the largest eigenvalue. Hence, the approximate Bahadur slope of the Epps–Pulley test statistic is given by

$$c^*_{T_\beta}(\vartheta) = \frac{b_{T_\beta}(\vartheta)}{\lambda_1(\beta)}. \tag{5.2}$$

Thus, one has to tackle the so-called *eigenvalue problem*, i.e., to find positive values λ and functions f such that $\mathcal{K}f = \lambda f$ or, in other words, to solve the integral equation

$$\int_{-\infty}^{\infty} K(s,t) f(t) \varphi_\beta(t)\, \mathrm{d}t = \lambda f(s), \quad s \in \mathbb{R}. \tag{5.3}$$

Since explicit solutions of such integral equations are only available in exceptional cases (for non-classical goodness-of-fit test statistics, see [10] and [11], we employ a stochastic approximation method. This method is related to the quadrature method in the classical literature on numerical mathematics (see [2, Chapter 3]), and which can also be found in machine learning theory (see [20]).For the approximation of spectra of Hilbert Schmidt operators, see [13]. To be specific, let Y be a random variable having density φ_β. Then (5.3) reads

$$\lambda f(s) = \mathbb{E}\big[K(s,Y) f(Y)\big], \quad s \in \mathbb{R}. \tag{5.4}$$

An empirical counterpart to (5.4) emerges if we let $y_1, y_2, \ldots, y_N$, $N \in \mathbb{N}$, be independent realizations of Y. An approximation of the expected value in (5.4) is then

$$\mathbb{E}\big[K(s,Y) f(Y)\big] \approx \frac{1}{N} \sum_{j=1}^{N} K(s, y_j) f(y_j), \quad s \in \mathbb{R}. \tag{5.5}$$

If we evaluate (5.5) at the points $y_1, \ldots, y_n$, the result is

$$\lambda f(y_i) = \frac{1}{N} \sum_{j=1}^{N} K(y_i, y_j) f(y_j), \quad i = 1, \ldots, N, \tag{5.6}$$

$\lambda\backslash\beta$	0.25	0.5	0.75	1	2	3	5	10
λ_1	0.00040	0.01065	0.03829	0.07507	0.15207	0.16149	0.13552	0.08791
λ_2	0.00003	0.00304	0.01735	0.04454	0.12921	0.14577	0.12606	0.08178
λ_3	0.00000	0.00021	0.00220	0.00846	0.04894	0.07676	0.08703	0.06879
λ_4	0.00000	0.00004	0.00076	0.00417	0.03966	0.06642	0.07997	0.06459
λ_5	0.00000	0.00000	0.00011	0.00098	0.01692	0.03755	0.05678	0.05518

Table 1. Approximate first five eigenvalues of $\mathcal{K}$ for different weight functions φ_β, each entry is the mean of 10 simulation runs

which is a system of N linear equations. Writing $v = (f(y_1), \dots, f(y_N)) \in \mathbb{R}^N$ and $\widetilde{K} = (K(y_i, y_j)/N)_{i,j=1,\dots,N} \in \mathbb{R}^{N\times N}$, we can rewrite (5.6) according to

$$\widetilde{K}v = \lambda v$$

in matrix form, from which the (approximated) eigenvalues $\lambda_1, \dots, \lambda_N$ can be computed explicitly. Note that for each eigenvalue λ_j we have an eigenvector $v_j \in \mathbb{R}^N$ (say), the components of which are the (approximated) values of the eigenfunctions (say) f_j computed at $y_1, \dots, y_N$.

The simulation of eigenvalues was performed in the statistical computing language R, see [19]. As parameters for the simulation we chose $N = 1000$, and we considered the tuning parameters $\beta \in \{0.25, 0.5, 0.75, 1, 2, 3, 5, 10\}$. Each entry in Table 1 stands for the mean of 10 simulation runs.

§6. Alternatives

As in Milošević et al. ([15]), we consider the following close alternatives:

- a Lehmann alternative with density

$$g_1(x;\vartheta) = (1+\vartheta)\Phi^\vartheta(x)\varphi(x);$$

- a first Ley-Paindaveine alternative with density (see [14]

$$g_2(x;\vartheta) = \varphi(x)\mathrm{e}^{-\vartheta(1-\Phi(x))}(1+\vartheta\Phi(x));$$

- a second Ley-Paindaveine alternative with density (see [14])

$$g_3(x;\vartheta) = \varphi(x)(1-\vartheta\pi\cos(\pi\Phi(x)));$$

- a contamination alternative (with $\mathrm{N}(\mu,\sigma^2)$ for several pairs $(\mu,\sigma^2) \neq (0,1)$) with density

$$g_4^{[\mu,\sigma^2]}(x;\vartheta) = (1-\vartheta)\varphi(x) + \frac{\vartheta}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right).$$

Alt.\\β	0.25	0.5	0.75	1	2	3	5	10
Lehmann	0.996	0.895	0.854	0.743	0.514	0.406	0.328	0.267
1st Ley-Paindaveine	0.947	0.944	0.998	0.937	0.745	0.612	0.507	0.417
2nd Ley-Paindaveine	0.824	0.872	0.986	0.981	0.881	0.754	0.641	0.533
Contamination with N(1,1)	0.760	0.649	0.592	0.499	0.328	0.255	0.205	0.166
Contamination with N(0.5,1)	0.945	0.824	0.766	0.654	0.438	0.343	0.276	0.224
Contamination with N(0,0.5)	0.084	0.267	0.474	0.587	0.675	0.606	0.526	0.442

Table 2. Approximate local Bahadur efficiency of $T_{n,\beta}$ with respect to the LRT

Like in Milošević et al. [15], we computed the local (as $\vartheta \to 0$) relative approximate Bahadur efficiencies with respect to the likelihood ratio test (LRT). The LRT is the best test regarding exact Bahadur effiency, and it is often used as a benchmark test. Table 2 displays the local approximate Bahadur efficiencies of $T_{n,\beta}$ with respect to the LRT, for each of the six alternatives considered in [15], and for $\beta \in \{0.25, 0.5, 0.75, 1, 2, 4, 5, 10\}$. A comparison with Table 1 of [15] shows that the Epps–Pulley test with $\beta = 0.5$ dominates the Kolmogorov–Smirnov test for each of the six alternatives, and for $\beta = 1$, $\beta = 2$ and $\beta = 3$, it outperforms the tests of Cramér–von Mises, the Watson variation of this test, and the Watson–Darling variation of the Kolmogorov–Smirnov test, respectively. If $\beta = 0.75$, the Epps–Pulley test dominates the Anderson–Darling test for each of the alternatives with the exception of the final contamination alternative. As a conclusion, the test of Epps and Pulley should receive more attention as a test for normality.

References

1. R. Bahadur, *Rates of convergence of estimates and test statictics.* — **38**, No. 2 (1967), 303–324.
2. C. T. H. Baker, *The numerical treatment of integral equations*, Oxford University Press, 1977.
3. L. Baringhaus, R. Danschke, N. Hense, *Recent and classical tests for normality – a comparative study.* — Comm. Statistics, Simulation and Computation **18** (1989), 363–379.
4. L. Baringhaus, B. Ebner, N. Hense, *The limit distribution of weighted L^2-goodness-of-fit statictics under fixed alternatives, with applications.* — Ann. of the Institute of Statist. Math. **69**, No. 5 (2017), 969–995.
5. L. Baringhaus, N. Hense, *A consistent test for multivariate normality based on the empirical characteristic function.* — Metrika **35**, No. 1 (1988), 39–348.
6. S. Csörgó, *Consistency of some tests for multivariate normality.* — Metrika **36** (1989), 107–116.

7. B. Ebner, N. Henze, *Tests for multivariate normality – a critical review with emphasis on weighted L^2-statictics.* — TEST **29**, No. 4 (2020), 845–892.
8. B. Ebner, N. Henze, Y. Y. Nikitin, *Integral distribution-free statistics of L_p-type and their asymptotic comparison.* — Computational Statistics & Data Analysis **53**, No. 9 (2009), 3426–3438.
9. T. W. Epps, L. B. Pulley, *A test for normality based on the empirical characteristic function.* — Biometrika bf 70, No. 3 (1983), 723–726.
10. N. Henze, Y. Y. Nikitin, *A new approach to goodness-of-fit testing based on the integrated empirical process.* — J Nonparametric Statist. **12**, No. 3 (2000), 391–416.
11. N. Henze, Y. Y. Nikitin, *Watson-type goodness-of-fit test based on the integrated empirical process.* — Mathematical Methods of Statist. **11**, No. 2 (2002), 183–2002.
12. N. Henze, T. Wagner, *A new approach to the BHEP tests for multivariate normality.* — J. Multivariate Anal. **62**, No. 1 (1997), 1–23.
13. V. Koltchinskii, E. Giné, *Random matrix appriximation of spectra of integral operators.* — Bernoulli **6**, No. 1 (2000), 113–167.
14. C. Ley, D. Paindaveine, *Le Cam optimal tests for symmetry against Ferreira and Steel's general skewed distributions.* — J. Nonparametric Statist. **21**, No. 8 (2009), 943–967.
15. B. Milošević, Y. Y. Nikitin, M. Obradović, *Bahadur efficiency of EDF based normality tests when parametres are estimated.* — ArXiv e-prints, ArXiv:2106.07437, 2021.
16. Y. Nikitin, *Asymptotic Efficiency of Nonparametric Test*, Cambridge university Press, 1995.
17. Y. Y. Nikitin, I. Peaucelle, *Efficiency and local optimality of nonparametric tests based on U and V-statistic.* — METRON **62**, No. 2 (2004), 185–2000.
18. Y. Y. Nikitin, A. V. Tchirina, *Lilliefors test for exponentiallity: Large deviations, asymptotic efficiency, and conditions of local optimality.* — Math. Methods of Statistics **16**, No. 1 (2007), 16–24.
19. R. Core Team R, *A Language and Environment for Statist. Computing.* R. Foundation for Statistical Computing, Vienna, Austria, 2021.
20. C. Rasmussen, K. Christopher, *Gaussian Processes for Machine Learning*, MIT Press, 2006.
21. V. M. Zolotarev, *Concerning a certain probability problem.* — Theory of Probability and its Applications **6**, No. 2 (1961), 201–204.

Institute of Stochastics,
Karlsruhe Institute of Technology (KIT),
Englerstr. 2, D-76133 Karlsruhe

E-mail: `Bruno.Ebner@kit.edu`

E-mail: `Norbert.Henze@kit.edu`

Поступило 27 июня 2021 г.