STATISTICA, anno LXXVII, n. 2, 2017

ASYMPTOTIC PITMAN'S RELATIVE EFFICIENCY

Christopher S. Withers
Industrial Research Limited, Lower Hutt, New Zealand
Saralees Nadarajah [1]
University of Manchester, Manchester M13 9PL, UK

1. Introduction

Suppose T is an estimator of a parameter θ based on an observed sample. The efficiency of T is defined by $1/[I(\theta)\mathrm{Var}(T)]$, where $I(\theta)$ denotes the Fisher information of θ based on the sample. This definition has been modified to suit various application areas, including hypothesis testing, experimental designs, factor analysis, classification analysis, and others. Nikulin (2001) provides a detailed account of an efficiency.

Here, we are interested in efficiencies for a hypothesis testing problem. The classical efficiencies are Pitman efficiency (Pitman, 1948), Chernoff efficiency (Chernoff, 1952), Hodges-Lehmann efficiency (Hodges and Lehmann, 1956) and Bahadur efficiency (Bahadur, 1960b). Some applications involving classical efficiencies are described in Burgio and Nikitin (1998), Burgio and Nikitin (2004). The former describes goodness-of-fit tests for the normal distribution and the latter describes combination of the sign and Maesono tests for symmetry.

The oldest known efficiency is the Pitman efficiency (Pitman, 1948). Its applications have been widespread. Some recent applications have included: test of multivariate linear models using spatial concordances (Choi and Marden, 2005); optimal sign tests for data from ranked set samples (Wang and Zhu, 2005); asymptotic efficiency of the blest-type tests for independence (Stepanova and Wang, 2008); a weighted multivariate signed-rank test for cluster-correlated data (Haataja *et al.*, 2009); assessing the relative power of structural break tests using a framework based on the approximate Bahadur slope (Kim and Perron, 2009); comparison of the Stein and the usual estimators for the regression error variance (Ohtani and Wan, 2009); optimal sign test for quantiles in ranked set samples (Dong and Cui, 2010); testing for increasing mean inactivity time (Zhang and Cheng, 2010); a nonparametric test for a two-sample scale problem based on subsample medians (Mahajan *et al.*, 2011); superior design sensitivity in an observational study of treatments for ovarian cancer (Rosenbaum, 2012).

But there have not been convenient tools to compute the Pitman efficiency. Often only bounds have been obtained for the Pitman efficiency: Capon (1965) derived bounds for the Pitman efficiency of the Kolmogorov-Smirnov test; Ramachandramurty (1966) obtained lower bounds for the Pitman efficiency of the one-sided Kolmogorov test for all normal alternatives and for the Smirnov test for normal shift

[1] Corresponding author. E-mail: mbbsssn2@manchester.ac.uk

alternatives; Yu (1971) computed upper and lower bounds of the Pitman asymptotic efficiency of the Kolmogorov-Smirnov test with respect to the Neyman test and locally most powerful rank test; Rothe (1983) obtained a lower bound for the Pitman efficiency of Friedman type tests; for the group of integral tests of homogeneity, generalizing the omega-square tests, Nikitin (1984) found a lower bound of the Pitman asymptotic relative efficiency relative to the Student test for the shift alternative; Weissfeld and Wieand (1984) presented bounds on the Pitman efficiency for two-sample scale and location statistics along with densities for which these bounds are sharp; Jansen and Ramirez (1993) derived bounds for the Pitman efficiency for efficiency comparisons in linear inference; Tsai (2009) established a lower bound on the Pitman efficiency of the spherical Wilcoxon rank test relative to the spherical T^2-test; Ermakov (2011) derived a lower bound for the Pitman efficiency for a nonparametric signal detection problem; to mention just a few. For other known results on the Pitman efficiency, see the excellent book Nikitin (1995).

The aim of this note is to develop computationally convenient exact expressions for the Pitman efficiency. The given expressions (see Theorem 3.2) are simple and can be computed using software like Maple. Programs written by the authors took only a few seconds to compute the expressions. This development has become possible because of the recent papers Withers and Nadarajah (2009), Withers and Nadarajah (2013b), Withers and Nadarajah (2013a).

Withers and Nadarajah (2009) showed how the asymptotic power (AP) of integral type statistics may be computed and compared to others. Withers and Nadarajah (2013b) introduced two new efficiencies, referred to as the *fixed-α efficiency* and the *fixed-β efficiency*, and developed theoretical tools to evaluate the efficiencies for some of the most usual goodness of fit (gof) tests, including the Kolmogorov-Smirnov tests. Withers and Nadarajah (2013a) introduced a third new efficiency, referred to as the *Bayesian efficiency*, and established its advantages over the fixed-α and fixed-β efficiencies.

The contents of this note are organized as follows. Some theoretical tools to compute the Pitman efficiency are developed in Section 3. Some preliminaries and notations needed for the developments are given in Section 2. The theoretical tools are illustrated numerically in Section 4. Some concluding remarks are given in Section 5. The proofs of all results in Section 3 are provided in Section 6.

2. Preliminaries and notation

We need some notation for the theoretical developments in Section 3. Let F_n denote the empirical cumulative distribution function (cdf) of a random sample $X_1,\ldots,X_n$ from a cdf F on $\mathbb{R}=(-\infty,\infty)$. Let F_0 denote some hypothesized cdf for F and assume throughout that F_0 is absolutely continuous. Let ψ denote a non-negative function on $[0,1]$. Define

$$D_{F_0}(F)=\left\| \, |F-F_0|\, \psi(F_0) \right\|_{F_0,\infty}, \tag{1}$$

$$D^+_{F_0}(F)=\sup\,(F-F_0)\,\psi(F_0), \tag{2}$$

$$D^-_{F_0}(F)=\sup\,(F_0-F)\,\psi(F_0), \tag{3}$$

$$V_{F_0}(F)=D^+_{F_0}(F)+D^-_{F_0}(F), \tag{4}$$

$$T_{n,m}(\psi) = \| |F_n - F_0| \psi(F_0) \|_{F_0,m}, \ 0 < m \leq \infty, \tag{5}$$
$$T^+_{n,m}(\psi) = \| (F_n - F_0) \psi(F_0) \|_{F_0,m}, \ m = 1,3,5,\ldots, \tag{6}$$
$$D_n(\psi) = D_{F_0}(F_n), \tag{7}$$
$$D^+_n(\psi) = D^+_{F_0}(F_n), \tag{8}$$
$$D^-_n(\psi) = D^-_{F_0}(F_n), \tag{9}$$
$$V_n(\psi) = V_{F_0}(F_n), \tag{10}$$
$$S_{n,m}(\psi) = \| |F_n - F_0| \psi(F_0) \|_{F_n,m}, \ 0 < m \leq \infty, \tag{11}$$
$$S^+_{n,m}(\psi) = \| (F_n - F_0) \psi(F_0) \|_{F_n,m}, \ m = 1,3,5,\ldots, \tag{12}$$

where $|x|$ denotes the absolute value of x, $\|G\|_{F,m} = \left(\int_{-\infty}^{\infty} G^m dF \right)^{1/m}$ if $0 < m < \infty$, $\|G\|_{F,\infty} = \sup G$ and $\|G\|_m = \|G\|_{U,m}$, where U is the cdf of a uniform$[0,1]$ random variable. The parameter m represents the kind of averaging used: $m = 1$ represents ordinary averaging, $m = 2$ represents quadratic averaging, $m = 3$ represents cubic averaging and so on. The class of statistics given by (5)-(12) includes the integral statistics, Kolmogorov-Smirnov statistics, Kuiper statistics, $T_{n,2}(1)$, $T_{n,\infty}(1)$ and $V_n(1)$. The asymptotic null distributions for these statistics were derived in Anderson and Darling (1952), Kolmogorov (1941) and Stephens (1965).

We test $H_0 : F = F_0$ against $H_1 : F = F_\theta$, where $\{F_\theta\}$ is a set of cdfs on $\mathbb{R}$ disjoint from F_0. We reject H_0 when $T_n(F_n) > r_n$ for some functional $T_n(\cdot)$ taken to be one of (5)-(12). For simplicity of presentation, we exclude randomized tests. Suppose T_n is such that $T_n(F_n) = T(F) + o_p(1)$, a common condition in testing problems (see, for example, Lehmann (1999)). Set

$$I(F,G) = \int_{-\infty}^{\infty} \ln(dF/dG)\,dF \text{ if } F,\ G \text{ are absolutely continuous cdfs,}$$
$$I(A,B) = \inf_{F\in A} \inf_{G\in B} \int_{-\infty}^{\infty} \ln(dF/dG)\,dF \text{ for sets of cdfs } A \text{ and } B,$$
$$I_0(r,F) = \inf_Q I(\{\text{cdfs } Q \text{ on } \mathbb{R} : T(Q) > r\}, \{F\}),$$

where both F and G are assumed to be absolutely continuous cdfs. A weaker condition is to assume that F is absolutely continuous with respect to G and then define $I(F,G) = \infty$ otherwise.

3. Pitman efficiency

The Pitman efficiency is an index for comparing test procedures. Suppose there are two procedures. If the first procedure requires n_1 observations to attain a certain power of a test, or a specified mean squared error, and second procedure requires n_2 observations to achieve the same precision, the Pitman efficiency of the first against second procedure is n_2/n_1. A related efficiency is the Bahadur efficiency. Its definition is of highly theoretical nature, see Bahadur (1960b) for details. Further details including applications of Pitman and Bahadur efficiencies can be found in Chapters 11 to 14 of van der Vaart (1998).

Suppose there are two tests say Test 1 and Test 2 for testing $H_0 : F = F_0$ against $H_1 : F = F_\theta$ based on a sample of size n. Since the limiting Bahadur and Pitman efficiencies are the same, the latter can be expressed as

$$\lim_{n\to\infty} \frac{\text{Bahadur slope for Test 1}}{\text{Bahadur slope for Test 2}} = \lim_{n\to\infty} \frac{\dfrac{1}{n}\ln(p\text{-value for Test 1})}{\dfrac{1}{n}\ln(p\text{-value for Test 2})}.$$

Suppose that the true value of θ is known for Test 1 and Test 2 is the optimal likelihood ratio test. Then, by equation (6) in Otsu (2010),

$$\lim_{n\to\infty} \frac{1}{n}\ln(p\text{-value for Test 2}) = -I(F_\theta, F_0).$$

Suppose Test 1 rejects H_0 if $T_n(F_n) > r_n$, where F_n denotes the empirical cdf of the data. Again, by equation (6) in Otsu (2010),

$$\lim_{n\to\infty} \frac{1}{n}\ln(p\text{-value for Test 1}) = -\lim_{n\to\infty} \inf_Q I(\{\text{cdfs } Q \text{ on } \mathbb{R} : T_n(Q) > r_n\}, F_0).$$

Thus, the Pitman efficiency for the problem formulated in Section 2 is

$$\lim_{n\to\infty} \frac{\inf_Q I(\{\text{cdfs } Q \text{ on } \mathbb{R} : T_n(Q) > r_n\}, F_0)}{I(F_\theta, F_0)} = \lim_{r\downarrow 0} \frac{I_0(r, F_0)}{I(F_\theta, F_0)}.$$

Hence, in order to find the Pitman efficiency, we need the behavior of $I_0(r, F_0)$ near $r = 0$, see Theorem 3.1. For a proof that $I_0(r, F_0)$ is continuous in r, see Groeneboom *et al.* (1979).

If we keep the type 1 and type 2 errors fixed then the Pitman efficiency will equal the limit of the fixed-β efficiency introduced in Withers and Nadarajah (2013b), for a proof see Appendix 2 of Bahadur (1960a).

THEOREM 3.1. *(i) Let ψ be bounded, continuous and non-negative, $0 < m < \infty$. For $T_{n,m}(\psi)$, $T^+_{n,m}(\psi)$, $S_{n,m}(\psi)$, $S^+_{n,m}(\psi)$ of (5), (6), (11), (12),*

$$I_0(r, F_0) = \lambda_1 \frac{r^2}{2} + o\left(r^2\right) \tag{13}$$

as $r \downarrow 0$, where $\lambda_1 = \lambda_1(m, \psi)$ is the minimum positive λ such that $\ddot{H} + \lambda\psi^m H^{m-1} = 0$ has a non-negative solution H_1 in $[0, 1]$ such that $H_1(0) = H_1(1) = 0$, and $\int_0^1 \psi^m H_1^m = 1$, provided that such a solution is unique, where $\ddot{\epsilon}$ denotes the second derivative of $\epsilon(\cdot)$. Furthermore,

$$\lambda_1^{-1} = \sup_{\int_0^1 f^m = 1, f \geq 0} \int_0^1 \int_0^1 [\min(s, t) - st]\, \psi(s)\psi(t) f(s)^{m-1} f(t)^{m-1} ds\, dt, \tag{14}$$

where inf *replaces* sup *for $m > 1$, and*

$$\lambda_1(m, 1) = \frac{2}{m}\left(1 + \frac{m}{2}\right)^{2/m-1} \cdot B\left(\frac{1}{2}, \frac{1}{m}\right)^2, \tag{15}$$

where

$$B(a,b)=\int_0^1 t^{a-1}(1-t)^{b-1}dt$$

denotes the beta function.

(ii) Let ψ be bounded, continuous and non-negative. Then for $D_n(\psi)$, $D_n^+(\psi)$, $D_n^-(\psi)$, (13) holds with $\lambda_1=\lambda_1(\infty,\psi)$ given by

$$\lambda_1^{-1}=\sup_{0<x<1}\left(x-x^2\right)\psi(x)^2.$$

(iii) Let ψ be bounded, continuous and non-negative. Then for $V_n(\psi)$, (13) holds with

$$\begin{aligned}\lambda_1^{-1} &= \sup_{0<x<y<1}\left[\left(x-x^2\right)\psi(x)^2+\left(y-y^2\right)\psi(y)^2-2x(1-y)\psi(x)\psi(y)\right]\\ &= \lambda_V(\psi)^{-1}\end{aligned}$$

say.

THEOREM 3.2. *Let ψ be bounded, $F=F_\theta$ and $p(x)=\left[\frac{\partial}{\partial\theta}F_\theta\left(F_0^{-1}(x)\right)\right]_{\theta=0}$, assuming $p(\cdot)$ is well defined. Then the Pitman efficiency of*

$$\begin{array}{lll} T_{n,m}(\psi) \text{ and } S_{n,m}(\psi) & \text{is} & \lambda_1(m,\psi)\|\,|p|\,\psi\|_m^2/\int_0^1 \dot{p}^2,\ 0<m\le\infty,\\ T_{n,m}^+(\psi) \text{ and } S_{n,m}^+(\psi) & \text{is} & \lambda_1(m,\psi)\|p\psi\|_m^2/\int_0^1 \dot{p}^2,\ m=1,3,5,\ldots,\\ D_n^+(\psi) & \text{is} & \lambda_1(\infty,\psi)(\sup(p\psi))^2/\int_0^1 \dot{p}^2,\\ D_n^-(\psi) & \text{is} & \lambda_1(\infty,\psi)(\sup(-p\psi))^2/\int_0^1 \dot{p}^2,\\ V_n(\psi) & \text{is} & \lambda_V(\psi)(\sup(p\psi)+\sup(-p\psi))^2/\int_0^1 \dot{p}^2, \end{array}$$

where $\lambda_1(m,\psi)$, $\lambda_V(\psi)$ are defined as in Theorem 3.1 and $\dot{\epsilon}$ denotes the first derivative of $\epsilon(\cdot)$.

Suppose the parametric model used in Theorem 3.2 is regular at 0, i.e, F_θ has a density f_θ with respect to the Lebesgue measure μ and

$$\int\left(f_\theta^{1/2}-f_0^{1/2}-\frac{1}{2}\theta\chi f_0^{1/2}\right)^2 d\mu=o\left(\theta^2\right) \tag{16}$$

for some χ in $L_2(F_0)$. Then we have

$$\sup_{t\in\mathbb{R}}\left|F_\theta(t)-F_0(t)-\int_{-\infty}^t \chi\, dF_0\right|=o(\theta).$$

This shows that

$$p(x)=\int_{-\infty}^{F_0^{-1}(x)}\chi\,dF_0$$

and $\dot{p}=\chi\circ F_0^{-1}$ almost everywhere μ. Consequently, $\int_0^1 \dot{p}^2$ is the Fisher information $J=\int \chi^2 f_0 d\mu$ of the model at 0.

The proof of Theorem 3.1 is given in Section 6. Theorem 3.2 follows easily from Theorems 4.1 to 4.4 in Withers and Nadarajah (2013b) and Theorem 3.1.

We have written programs in Maple to compute the Pitman efficiencies given by Theorem 3.2. These programs took only a fraction of a second to compute the efficiencies for a wide range of examples, including those in Section 4. The programs could be of use to the many applications of Pitman efficiency.

4. Numerical computations

Here, we compute the Pitman efficiency given by Theorem 3.2 for seven parametric families given by

(a) the normal with shift alternative specified by

$$F_\theta(x)=\Phi(x-\theta),$$

where $\Phi(\cdot)$ denotes the cdf of a standard normal random variable. For this family,

$$p(x)=-\phi\left(\Phi^{-1}(x)\right)$$

with

$$\int_0^1 \dot{p}^2=1$$

and

$$\sup(p)=0,\quad \sup(-p)=\phi(0),$$

where $\phi(\cdot)$ denotes the probability density function (pdf) of a standard normal random variable;

(b) the logistic with shift alternative specified by

$$F_\theta(x)=\left[1+\exp(-x+\theta)\right]^{-1}.$$

For this family,

$$p(x)=-x+x^2$$

with

$$\int_0^1 \dot{p}^2 = 1/3$$

and

$$\sup(p) = 0, \quad \sup(-p) = 0.25;$$

(c) the double-exponential with shift alternative specified by

$$F_\theta(x) = \begin{cases} \frac{1}{2}\exp(x-\theta), & x \le \theta, \\ 1 - \frac{1}{2}\exp(-x+\theta), & x \ge \theta. \end{cases}$$

For this family,

$$p(x) = \begin{cases} -x, & x \le 0.5, \\ x-1, & x \ge 0.5 \end{cases}$$

with

$$\int_0^1 \dot{p}^2 = 1$$

and

$$\sup(p) = 0, \quad \sup(-p) = 0.5;$$

(d) the Lehmann alternative (Lehmann, 1953) specified by

$$F_\theta(x) = F_0^{\theta+1}(x).$$

For this family,

$$p(x) = x \ln x$$

with

$$\int_0^1 \dot{p}^2 = 1$$

and

$$\sup(p) = 0, \quad \sup(-p) = \exp(-1);$$

(e) the family specified by

$$F_\theta(x) = [\exp(\theta F_0(x)) - 1] / [\exp(\theta) - 1].$$

For this family,

$$p(x) = -\frac{1}{2}\left(x - x^2\right)$$

with

$$\int_0^1 \dot{p}^2 = 1/12$$

and

$$\sup(p) = 0, \quad \sup(-p) = \frac{1}{8};$$

(f) Cauchy with shift alternative specified by

$$F_\theta(x) = 1/2 + 1/\pi \tan^{-1}(x - \theta).$$

For this family,

$$p(x) = -\frac{1}{\pi}\cos^2\left[\pi\left(x - \frac{1}{2}\right)\right]$$

with

$$\int_0^1 \dot{p}^2 = 1/2$$

and

$$\sup(p) = 0, \quad \sup(-p) = \frac{1}{\pi};$$

(g) the normal with scale alternative specified by

$$F_\theta(x) = \Phi\left(x \exp(-\theta)\right).$$

For this family,

$$p(x) = -\Phi^{-1}(x)\phi\left(\Phi^{-1}(x)\right)$$

with

$$\int_0^1 \dot{p}^2 = 2$$

and

$$\sup(p) \approx 0.2419707, \quad \sup(-p) \approx 0.2419707.$$

Throughout, the sup is computed over $x \in [0,1]$.

Examples (a), (b), (c) and (f) are special cases of the location model generated by F_0 with F_0 having finite Fisher information for location. f_0 is absolutely continuous and

$$J_0 = \int \left[\frac{f_0'(x)}{f_0(x)}\right]^2 f_0(x)dx < \infty.$$

This is equivalent to (16) with $f_\theta(x) = f_0(x-\theta)$ and $\chi(x) = -f'(x)/f_0(x)$. In this case, $p = -f_0 \circ F_0^{-1}$ and $\int_0^1 \dot{p}^2 = J_0$. Also $\sup p = 0$ and $\sup -p = \sup f_0$. In examples (a), (b), (c) and (f), the density f_0 is symmetric and unimodal, so $\sup -p$ equals to $f_0(0)$.

Example (g) is a special case of a scale model generated by F_0. Suppose that the function h defined by $h(x) = xf_0(x)$ is absolutely continuous and

$$J_1 = \int \left[\frac{h'(x)}{f_0(x)}\right]^2 f_0(x)dx = \int \left[1 + x\frac{f_0'(x)}{f_0(x)}\right]^2 f_0(x)dx < \infty.$$

Then (16) holds with $f_\theta(x) = f_0(x/(1+\theta))/(1+\theta)$ and $\chi = -h'/f_0$ implying $p = -h \circ F_0^{-1}$ and $\int_0^1 \dot{p}^2 = J_1$.

Example (d) treats the case $f_\theta(x) = (1+\theta)f_0(x)F_0^\theta(x)$ for which (16) holds with $\chi = 1 + \ln F_0$ and $\int \chi^2 dF_0 = 1$.

The distributions specified by example (e) are known as the *truncated-exponential skew-symmetric distributions* (Nadarajah *et al.*, 2014). They are the most flexible skew-symmetric distributions known to date, even more flexible than Azzalini's skew-symmetric distributions (Azzalini, 1985). Some of the advantages of truncated-exponential skew-symmetric distributions established by Nadarajah *et al.* (2014) include: 1) they belong to the exponential family; 2) have closed form expressions for pdf, cdf and quantiles; 3) exhibit the same tail behaviors as those of F_0; 4) the maximum likelihood estimator for θ always exists and is unique; 5) admit a uniformly most powerful test for hypotheses about θ. For details on the distributions, see Nadarajah *et al.* (2014).

We compute the Pitman efficiencies in Theorem 3.2 for the seven examples for $\psi \equiv 1$. This requires $\lambda_1(m,1)$, $\lambda_1(\infty,1)$ and $\lambda_V(1)$. A plot of $\lambda_1(m,1)$ versus m is shown in Figure 1. We see that $\lambda_1(m,1)$ is a monotonically decreasing function of m with $\lambda_1(\infty,1) = 4$. Also

$$\lambda_V(1)^{-1} = \sup_{0<x<y<1} \left[x - x^2 + y - y^2 - 2x(1-y)\right] = \frac{1}{4}.$$

Figure 2 shows the Pitman efficiency, $\lambda_1(m,1)\| |p|\|_m^2 / \int_0^1 \dot{p}^2$, versus m for the seven examples. The Pitman efficiency for each of the examples initially increases before decreasing with m. An exception is the double-exponential example with shift alternative. In this case, the Pitman efficiency monotonically increases with m.

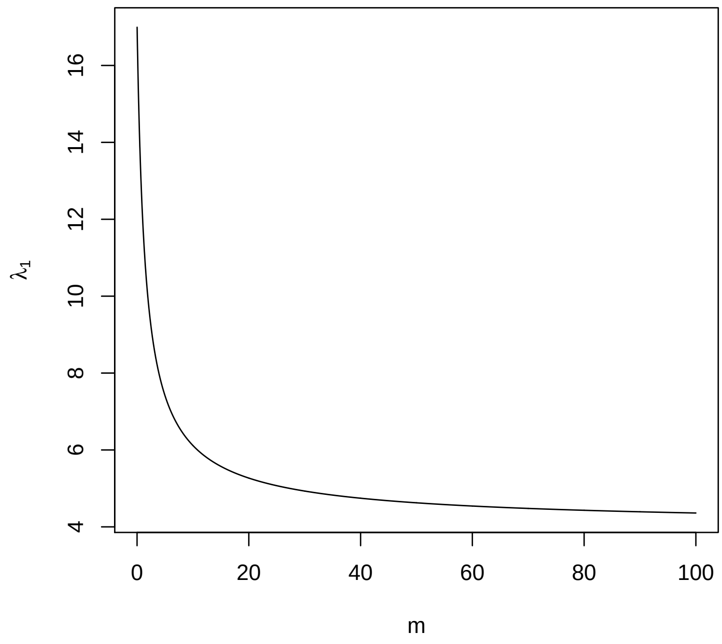

Figure 1 – Plot of $\lambda_1(m,1)$ versus m.

The Pitman efficiency, $\lambda_1(m,1) \| \, |p| \|_m^2 / \int_0^1 \dot{p}^2$, is smallest for all m for the Cauchy example with shift alternative. It is second smallest for all m for the normal example with scale alternative. It is third smallest for all m for the Lehmann alternative example. It is fourth smallest for small m for the double-exponential example with shift alternative. It is fourth smallest for large m for the normal example with shift alternative. It is fifth smallest for small m for the normal example with shift alternative. It is fifth smallest for large m for example (e). It is largest for small m for example (e). It is largest for large m for the double-exponential example with shift alternative.

The fact that $\lambda_1(m,1) \| \, |p| \|_m^2 / \int_0^1 \dot{p}^2$ is smallest for the Cauchy example may be due to it having the heaviest tails among the seven examples. The fact that $\lambda_1(m,1) \| \, |p| \|_m^2 / \int_0^1 \dot{p}^2$ is largest for the double-exponential and normal examples may be due to they having the lightest tails among the seven examples. The different behaviours of $\lambda_1(m,1) \| \, |p| \|_m^2 / \int_0^1 \dot{p}^2$ may be due to how heavy or light the tails are.

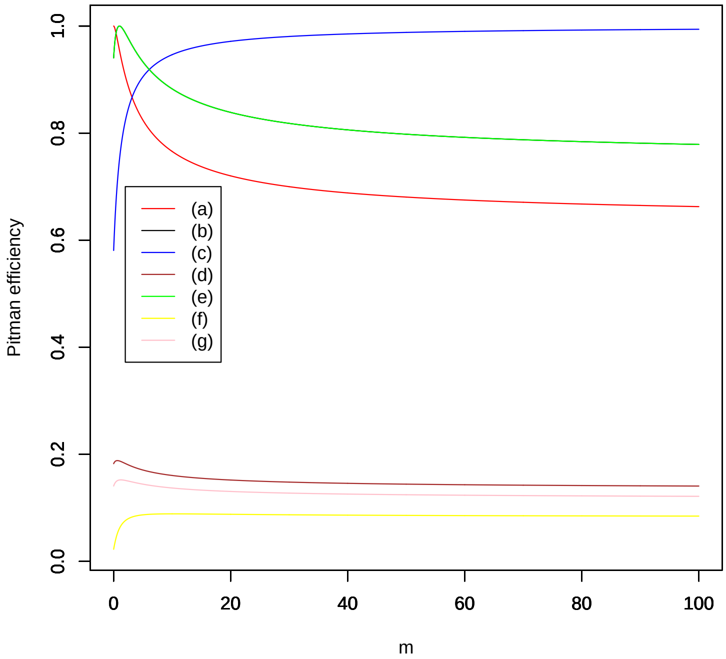

Figure 2 – Plot of $\lambda_1(m,1)\| |p| \|_m^2 / \int_0^1 \dot{p}^2$ versus m for the seven examples.

The Pitman efficiency, $\lambda_1(\infty,1)(\sup(p))^2 / \int_0^1 \dot{p}^2$, is zero for the normal example with shift alternative, zero for the logistic example with shift alternative, zero for the double-exponential example with shift alternative, zero for the Lehmann alternative example, zero for example (e), zero for the Cauchy example with shift alternative and $2 \cdot (0.2419707)^2$ for the normal example with scale alternative.

The Pitman efficiency, $\lambda_1(\infty,1)(\sup(-p))^2 / \int_0^1 \dot{p}^2$, is $4\phi^2(0)$ for the normal example with shift alternative, $3/4$ for the logistic example with shift alternative, one for the double-exponential example with shift alternative, $4\exp(-2)$ for the Lehmann alternative example, $3/4$ for example (e), $8/\pi^2$ for the Cauchy example with shift alternative and $2 \cdot (0.2419707)^2$ for the normal example with scale alternative.

The Pitman efficiency, $\lambda_V(1)(\sup(p)+\sup(-p))^2 / \int_0^1 \dot{p}^2$, is $4\phi^2(0)$ for the normal example with shift alternative, $3/4$ for the logistic example with shift alternative, one for the double-exponential example with shift alternative, $4\exp(-2)$ for the Lehmann alternative example, $3/4$ for example (e), $8/\pi^2$ for the Cauchy example with shift alternative and $8 \cdot (0.2419707)^2$ for the normal example with scale alternative.

Finally, we compared our exact values for Pitman efficiency with the bounds due to Capon (1965), Ramachandramurty (1966) and Jansen and Ramirez (1993) cited in Section 1. Computations showed that relative errors of the bounds were at best five percent. The details of the computations are not given here for space concerns. This illustrates the importance of having exact means for computing Pitman efficiency.

5. Conclusions

We have given a method for computing the Pitman efficiency exactly for the hypothesis testing problem defined in Section 3. This appears to be the first such method. We have illustrated the method using seven examples. In each example, the Pitman efficiency was computed in a fraction of a second. Programs in Maple for computing the Pitman efficiency are available from the corresponding author. Maple allows for arbitrary precision, so the accuracy of computations was not an issue.

The proposed Pitman efficiency was motivated as the limit of the fixed-β efficiency introduced in Withers and Nadarajah (2013b). A future work is to see if other efficiencies can be introduced by taking the limits of fixed-α and Bayesian efficiencies introduced in Withers and Nadarajah (2013b) and Withers and Nadarajah (2013a).

Other future work are to see how the Pitman efficiency can be computed for hypotheses testing problems involving:

- multivariate distribution functions, matrix variate distribution functions and complex variate distribution functions;
- discrete distributions (Aaberge, 1983).

6. Proofs

We need the following lemma.

Lemma 6.1. *Let ψ be continuous and non-negative. If for some λ_1,*

$$\ddot{H} + \lambda_1 H^{m-1} \psi^m = 0$$

has a unique non-negative solution H_1 on $[0,1]$ which vanishes at 0, 1 such that

$$\int_0^1 H_1^m \, \psi^m = 1,$$

then $f_1 = H_1 \psi$ maximizes if $m \geq 1$ (minimizes if $m < 1$)

$$L(f) = \int_0^1 \int_0^1 K(s,\ t) f(s)^{m-1} f(t)^{m-1} \, ds dt$$

among f such that $\int_0^1 f^m = 1$, where

$$K(s,\ t) = [\min(s,\ t) - st]\, \psi(s) \psi(t)$$

and the maximum (minimum) equals λ_1^{-1}. Furthermore,

$$\lambda_1 = \int_0^1 \dot{H}_1^2.$$

PROOF. Let K be any positive continuous symmetric function on $[0,1]^2$. Set $q = m/(m-1)$. For $0 < m < 1$ it follows from Hölder's inequality that

$$\int |fg| \geq \|f\|_m \|g\|_q.$$

The direction of the inequality is reversed when $m \geq 1$. Applying Hölder's inequality twice, one obtains

$$\int_0^1 \int_0^1 K(s,t)\,|a(s)b(t)|\,ds\,dt \leq (\geq) \|a\|_q \|b\|_q \left(\int_0^1 \int_0^1 K^m \right)^{1/m}$$

for $m \geq (<)1$. Hence, for $m \geq (<)1$, $L(f)$ has a finite positive maximum (minimum) among f such that

$$\int_0^1 f^m = 1. \tag{17}$$

By the calculus of variations, if f_1 is any non-negative function that maximizes (minimizes) $L(f)$ subject to (17), it satisfies

$$f(s) = \lambda \int_0^1 K(s,\ t) f(t)^{m-1}\, dt \tag{18}$$

for some λ.

Now (17) and (18) imply $\lambda^{-1} = L(f)$. Hence, for f_1, $\lambda = \lambda_1$, the minimum (maximum) positive λ such that (17) and (18) have a non-negative solution. For the particular K of the lemma, setting $H = f/\psi$, (18) is equivalent to

$$\ddot{H} + \lambda H^{m-1} \psi^m = 0$$

since $H(0) = H(1) = 0$ (differentiating twice). Hence, if this differential equation has a unique solution (H_1, λ_2) satisfying the boundary conditions, then $\lambda_2 = \lambda_1$ and $f_1 = H_1 \cdot \psi$. Finally,

$$1 = \int_0^1 H_1^m \psi^m = -\int_0^1 H_1 \ddot{H}_1/\lambda_1 = \int_0^1 \dot{H}_1^2/\lambda_1.$$

The proof is complete.

PROOF of Theorem 3.1.
(i) We give the proof for $T_{n,m}(\psi)$, $T_{n,m}^{+}(\psi)$, $m \neq 2$ only. The proof for $m = 2$ follows as the limiting case. The other cases are analogous.

By the lemma in Withers and Nadarajah (2013b), we seek continuous cdf $H = H(x)$ on $[0,1]$ such that

$$\frac{\ddot{H}(x)}{\dot{H}(x)} + m\lambda[H(x)-x]^{m-1}\ \psi(x)^m = 0,\ H(0)=0,\ H(1)=1,\ \dot{H}(x)\geq 0,$$
$$r^m = \int_0^1 [H(x)-x]^m\ \psi(x)^m,$$

where we suppose for the moment that $H(x) > x$ in $(0,1)$. It is clear that $H \to x$ as $r \to 0$. It is easily shown that

$$H = x + \lambda^\alpha a(x) + o(\lambda^\alpha)$$

as $r \to 0$, where $\alpha = (2-m)^{-1}$, $\ddot{a} + ma^{m-1}\psi^m = 0$ and $a(0) = a(1) = 0$. Hence, $a = \beta H_1$, where λ_1, H_1 are given by Theorem 4.4 in Withers and Nadarajah (2013b) and $\beta = (\lambda_1/m)^{1/(m-2)}$ and $\int_0^1 \dot{a}^2 = \beta^2\lambda_1$ by Lemma 5.1.

For $T_{n,m}(\psi)$, if we did not assume that $H > x$ in $(0,1)$, then we would have instead that a is a non-negative multiple of H_1 such that

$$\ddot{H}_1 + \lambda|H_1|^{m-1}\,\psi^m = 0,\ H_1(0) = H_1(1) = 0,\ \int_0^1 |H_1|^m\,\psi^m = 1;$$

by an analog of Theorem 4.4 in Withers and Nadarajah (2013b), $f_1 = H_1\psi$ maximizes $L(|f|)$ among f such that $\int_0^1 |f|^m = 1$; hence, we may take $f_1 > 0$ in $(0,1)$ which shows that we may assume $H > x$ in $(0,1)$. We have

$$I_0(r,\, F_0) = \frac{\lambda^{2\alpha}}{2}\int_0^1 \dot{a}^2\cdot[1+o(1)],$$
$$r^m = \lambda^{m\alpha}\int_0^1 a^m\psi^m\cdot[1+o(1)],$$

so (13) and (14) of (i) follow. To show (15) of (i), we note that $\dot{f}_1^2 = 2\lambda_1/m\left(c - f_1^m\right)$ for some constant c. Integration over $[0,1]$ yields $\lambda_1 = 2\lambda_1/m(c-1)$. Therefore, $c = 1 + m/2$ and

$$\pm x\cdot\left(\frac{2\lambda_1}{m}\right)^{1/2} = \int_0^{f_1} (c-y^m)^{-1/2}\,dy,$$
$$\pm(2\lambda_1\cdot m)^{1/2}\,c^{1/2-1/m}\ x = B\left(f_1^m/c : \frac{1}{m},\ \frac{1}{2}\right),$$

where

$$B(x:a,b) = \int_0^x t^{a-1}(1-t)^{b-1}dt$$

denotes the incomplete beta function.

Furthermore, $f_1 \geq 0$ and $f_1(0) = f_1(1) = 0$ so that $\dot{f}_1 \geq 0$ in $(0, 1/2]$, $\dot{f}_1 \leq 0$ in $[1/2,\ 1)$. This yields on simplification

$$(m\lambda_1/2)^{1/2}\ c^{1/2-1/m} = B(1/2,\ 1/m).$$

This proves (15) .

Note that f_1 is given by

$$B\left(\frac{f_1^m}{1+\frac{m}{2}} : \frac{1}{m},\ 1/2\right) = 2B\left(\frac{1}{m},\ 1/2\right) \cdot \begin{cases} x, & \text{in} \quad [0, 1/2], \\ 1-x, & \text{in} \quad [1/2,\ 1]. \end{cases}$$

(ii) This follows from Theorem 4.3 in Withers and Nadarajah (2013b) since

$$\begin{aligned} a(x,r) &= \begin{cases} (x+r)\ln\left(1+\frac{r}{x}\right) + (1-x-r)\ln\left(1-\frac{r}{1-x}\right), & 0 < x < 1-r, \\ \infty, & \text{otherwise} \end{cases} \\ &= \frac{r^2}{2}\left(x-x^2\right) + o\left(r^2\right) \end{aligned}$$

as $r \downarrow 0$.

(iii) In a similar way this follows from Theorem 4.4(v) in Withers and Nadarajah (2013b) with $F = F_0$. □

ACKNOWLEDGEMENTS

The authors would like to thank the referee and the Associate Editor for careful reading and comments which greatly improved the paper.

REFERENCES

R. AABERGE (1983). *Asymptotic efficiency of certain tests for comparison of proportions.* Scandinavian Journal of Statistics, 10, pp. 125–133.

T. W. ANDERSON, D. A. DARLING (1952). *Asymptotic theory of certain 'goodness of fit' criteria based on stochastic processes.* Annals of Mathematical Statistics, 23, pp. 193–212.

A. AZZALINI (1985). *A class of distributions include the normal ones.* Scandinavian Journal of Statistics, 12, pp. 171–178.

R. R. BAHADUR (1960a). *On the asymptotic efficiency of tests and estimates.* Sankhya B, 22, pp. 229–252.

R. R. BAHADUR (1960b). *Stochastic comparison of tests.* Annals of Mathematical Statistics, 31, pp. 276–295.

G. BURGIO, Y. NIKITIN (1998). *Goodness-of-fit tests for normal distribution of order p and their asymptotic efficiency.* Statistica, 58, pp. 213–230.

G. BURGIO, Y. NIKITIN (2004). *On the combination of the sign and maesono tests for symmetry and its efficiency.* Statistica, 63, pp. 213–222.

J. CAPON (1965). *On the asymptotic efficiency of the kolmogorov-smirnov test*. Journal of the American Statistical Association, 60, pp. 843–885.

H. CHERNOFF (1952). *A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations*. Annals of Mathematical Statistics, 23, pp. 493–507.

K. M. CHOI, J. I. MARDEN (2005). *Test of multivariate linear models using spatial concordances*. Journal of Nonparametric Statistics, 17, pp. 167–185.

X. F. DONG, L. R. CUI (2010). *Optimal sign test for quantiles in ranked set samples*. Journal of Statistical Planning and Inference, 140, pp. 2943–2951.

M. ERMAKOV (2011). *Nonparametric signal detection with small type i and type ii error probabilities*. Statistical Inference for Stochastic Processes, 14, pp. 1–19.

P. GROENEBOOM, J. OOSTERHOFF, F. H. RUYMGAART (1979). *Large deviation theorems for empirical probability measures*. Annals of Probability, 7, pp. 553–586.

R. HAATAJA, D. LAROCQUE, J. NEVALAINEN, H. OJA (2009). *A weighted multivariate signed-rank test for cluster-correlated data*. Journal of Multivariate Analysis, 100, pp. 1107–1119.

J. L. HODGES, E. L. LEHMANN (1956). *The efficiency of some nonparametric competitors of the t-test*. Annals of Mathematical Statistics, 27, pp. 324–335.

D. R. JANSEN, D. E. RAMIREZ (1993). *Efficiency comparisons in linear inference*. Journal of Statistical Planning and Inference, 37, pp. 51–68.

D. KIM, P. PERRON (2009). *Assessing the relative power of structural break tests using a framework based on the approximate bahadur slope*. Journal of Econometrics, 149, pp. 26–51.

A. N. KOLMOGOROV (1941). *Confidence limits for an unknown distribution function*. Annals of Mathematical Statistics, 12, pp. 461–463.

E. L. LEHMANN (1953). *The power of rank tests*. Annals of Mathematical Statistics, 24, pp. 23–43.

E. L. LEHMANN (1999). *Elements of Large-Sample Theory*. Springer Verlag, New York.

K. K. MAHAJAN, A. GAUR, S. ARORA (2011). *A nonparametric test for a two-sample scale problem based on subsample medians*. Statistics and Probability Letters, 81, pp. 983–988.

S. NADARAJAH, V. NASSIRI, A. MOHAMMADPOUR (2014). *Truncated-exponential skew-symmetric distributions*. Statistics, 48, pp. 872–895.

Y. Y. NIKITIN (1984). *Asymptotic comparison of a class of nonparametric tests with the student test*. Journal of Soviet Mathematics, 27, pp. 3271–3273.

Y. Y. NIKITIN (1995). *Asymptotic Efficiency of Nonparametric Tests*. Cambridge University Press, New York.

M. S. NIKULIN (2001). *Efficiency of a statistical procedure*. Encyclopedia of Mathematics.

K. OHTANI, A. T. K. WAN (2009). *Comparison of the stein and the usual estimators for the regression error variance under the pitman nearness criterion when variables are omitted*. Statistical Papers, 50, pp. 151–160.

T. OTSU (2010). *On bahadur efficiency of empirical likelihood*. Journal of Econometrics, 157, pp. 248–256.

E. J. G. PITMAN (1948). *Mimeographed Lecture Notes on Nonparametric Statistics*. Columbia University, New York.

P. V. RAMACHANDRAMURTY (1966). *On the pitman efficiency of one-sided kolmogorov and smirnov tests for normal alternatives*. Annals of Mathematical Statistics, 37, pp. 940–944.

P. R. ROSENBAUM (2012). *An exact adaptive test with superior design sensitivity in an observational study of treatments for ovarian cancer*. Annals of Applied Statistics, 6, pp. 83–105.

G. ROTHE (1983). *A lower bound for the pitman efficiency of friedman type tests*. Statistics and Probability Letters, 1, pp. 239–242.

N. STEPANOVA, S. WANG (2008). *Asymptotic efficiency of the blest-type tests for independence*. Australian and New Zealand Journal of Statistics, 50, pp. 217–233.

M. A. STEPHENS (1965). *The goodness-of-fit statistic v_n: Distribution and significance points*. Biometrika, 50, pp. 309–321.

M. T. TSAI (2009). *Asymptotically efficient two-sample rank tests for modal directions on spheres*. Journal of Multivariate Analysis, 100, pp. 445–458.

A. W. VAN DER VAART (1998). *Asymptotic Statistics*. Cambridge University Press, New York.

Y. G. WANG, M. ZHU (2005). *Optimal sign tests for data from ranked set samples*. Statistics and Probability Letters, 72, pp. 13–22.

L. A. WEISSFELD, H. WIEAND (1984). *Bounds on efficiencies for some two-sample nonparametric statistics*. Communications in Statistics—Theory and Methods, 13, pp. 1741–1757.

C. S. WITHERS, S. NADARAJAH (2009). *Power of a class of goodness-of-fit test i*. ESAIM: Probability and Statistics, 13, pp. 283–300.

C. S. WITHERS, S. NADARAJAH (2013a). *Bayesian efficiency*. Statistics and Probability Letters, 83, pp. 1203–1212.

C. S. WITHERS, S. NADARAJAH (2013b). *Fixed-α and fixed-β efficiencies*. ESAIM: Probability and Statistics, 17, pp. 224–235.

C. S. YU (1971). *Pitman efficiencies of kolmogorov-smirnov tests*. Annals of Mathematical Statistics, 42, pp. 1595–1605.

S. H. ZHANG, H. T. CHENG (2010). *Testing for increasing mean inactivity time*. Statistics, 44, pp. 467–476.

SUMMARY

Pitman efficiency is the oldest known efficiency. Most of the known results for computing the Pitman efficiency take the form of bounds. Based on some recent developments due to the authors and some calculus of variations, we develop tools for computing the Pitman efficiency exactly. Their use is illustrated numerically.

Keywords: Goodness-of-fit tests; Pitman efficiency.