Ann. Inst. Statist. Math.
Vol. 54, No. 2, 382–390 (2002)

ASYMPTOTIC EFFICIENCY OF MAESONO STATISTICS FOR TESTING OF SYMMETRY

YA. YU. NIKITIN* AND E. V. PONIKAROV

Department of Mathematics and Mechanics, St. Petersburg State University, 198504 Petrodvorets, Russian Federation

(Received April 4, 2000; revised December 25, 2000)

Abstract. In an interesting paper Maesono introduced a new class of distribution-free statistics for testing of symmetry against shift alternative. The simplest of them coincides with the Wilcoxon statistic while the next is different but has the same Pitman efficiency. Maesono raised the problem of comparison between these two statistics on the basis of *exact* Bahadur efficiency. In this paper we calculate exact local Bahadur indices for all Maesono statistics and show when his statistics are better than the Wilcoxon statistic for sufficiently close alternatives.

Key words and phrases: Bahadur efficiency, large deviations, U-statistics, Wilcoxon test, exact slope, local index, Maesono statistics.

1. Introduction

Let $X_1, X_2, \ldots$ be independent identically distributed observations with absolutely continuous distribution function $F(x-\theta)$, $\theta \in R^1$ and density $f(x-\theta)$, which satisfies at $\theta = 0$ the condition of symmetry $f(x) = f(-x)$ for all x. We are interested in testing the hypothesis of symmetry $H_0 : \theta = 0$ against the alternative $H : \theta > 0$.

Maesono (1987) proposed for that purpose a new class of distribution-free statistics, which differs from the well-known class of linear signed-rank statistics. It consists of U-statistics S_r, defined for any integer $r \geq 2$ by the equality

$$S_r = \binom{n}{r}^{-1} \sum_{1 \leq i_1 < \cdots < i_r \leq n} \Phi_r(X_{i_1}, \ldots, X_{i_r})$$

with the kernels

$$\Phi_r(s_1, \ldots, s_r) = r^{-1} \left(\sum_{i=1}^{r} \prod_{j \neq i}^{r} \mathbf{1}_{\{s_i + s_j > 0\}} - 1 \right). \tag{1.1}$$

For a clear idea of a structure of kernels (1.1) we write out, for example, such kernel for $r = 3$:

$$\begin{aligned}\Phi_3(s_1, s_2, s_3) \\ &= \frac{1}{3}(\mathbf{1}_{\{s_1+s_2>0\}}\mathbf{1}_{\{s_1+s_3>0\}} + \mathbf{1}_{\{s_1+s_2>0\}}\mathbf{1}_{\{s_2+s_3>0\}} + \mathbf{1}_{\{s_2+s_3>0\}}\mathbf{1}_{\{s_1+s_3>0\}} - 1).\end{aligned}$$

*Research supported by the grants of RFBR No. 01-01-0245 and 99-01-0724.

The hypothesis H_0 is rejected for large absolute values of S_r.

Clearly, S_2 is equivalent to the famous signed-rank Wilcoxon statistic W. It is an interesting question how efficient the tests of symmetry based on S_r for $r > 2$ are. Maesono (1987) has found the expression for the Pitman efficiency $e_P(S_r \mid W)$ of the statistic S_r with respect to W, namely

$$e_P(S_r \mid W) = \frac{(r-1)^2 \left(\int_{-\infty}^{\infty} F^{r-2}(x) f^2(x) dx \right)^2}{6 \left(\frac{1}{2r-1} - \frac{((r-1)!)^2}{(2r-1)!} \right) \left(\int_{-\infty}^{\infty} f^2(x) dx \right)^2}. \tag{1.2}$$

In particular, it follows from (1.2) that $e_P(S_3 \mid W) = 1$, so that the statistics S_3 and W are indistinguishable in terms of Pitman efficiency. Therefore Maesono has found the expression for the so-called approximate Bahadur efficiency $e_B^{(a)}(S_3 \mid W; \theta)$ of the statistic S_3 with respect to W and has shown that for all bounded and almost everywhere continuous densities f and for any $\theta > 0$ it is true that

$$e_B^{(a)}(S_3 \mid W; \theta) > 1. \tag{1.3}$$

Certainly, the limiting value of this efficiency as $\theta \to 0$ equals to 1, as it should be because this limiting value coincides with the Pitman efficiency (see Bahadur (1960)).

However the approximate Bahadur ARE is known to be an unreliable means of test comparison. Therefore Maesono ((1987), p. 368) has underlined that "it is important and interesting to obtain the exact Bahadur ARE" of S_r with respect to W.

The problem remained unsolved, possibly because large deviations (LD) asymptotics of Chernoff type for appropriate U-statistics, on which such calculation leans, was unknown. This asymptotics was recently found by the authors (see Nikitin and Ponikarov (1998, 1999*a*, 1999*b*)). We want to emphasize here, that a result of Dasgupta (1984) which was considered by many authors as a solution of the LD problem for non-degenerate U-statistics is wrong, as it is shown in details by Nikitin and Ponikarov (1999*a*).

Using the LD asymptotics for U-statistics, we solve Maesono's problem on the calculation of the exact Bahadur ARE of the statistic S_3 with respect to W. It is interesting enough that the answer depends on the initial density f (see inequality (3.1) below). Furthermore, for all r we calculate the exact local Bahadur efficiency $e_B(S_r \mid W)$, which appears the same as that in (1.2). We also specify distributions for which S_r are locally optimal and hence dominate W. This conclusion is made on the basis of the deep analysis of conditions of local Bahadur efficiency for statistics S_r.

2. Exact Bahadur efficiency

We begin by asymptotic analysis of the statistics S_r under H_0. First of all we will pass from the initial observations $X_1, X_2, \ldots$ to independent observations $U_1, U_2, \ldots$ uniformly distributed on $[0,1]$. As F is continuous, it is possible to define the inverse function F^{-1} in the usual way and to replace the observations X_i by the observations $F^{-1}(U_i)$ coinciding with them in distribution. Under such transformation the events $\{X_i + X_j > 0\}$ are equivalent to the events $\{U_i + U_j > 1\}$.

Therefore under the validity of hypothesis H_0 instead of the statistics S_r we can study the new statistics S_r^*, which are equidistributed with S_r and are based on the uniformly distributed observations $U_1, U_2, \ldots$. These new statistics S_r^* are U-statistics with the kernels

$$\Phi_r^*(s_1, \ldots, s_r) = r^{-1}\left(\sum_{i=1}^{r}\prod_{j\neq i}^{r}\mathbf{1}_{\{s_i+s_j>1\}} - 1\right).$$

It is easy to check that under H_0 $E\Phi_r^*(U_1, \ldots, U_r) = 0$.

We will find now a one-dimensional projection of the kernels Φ_r^*, namely the function

$$\psi_r^*(u_1) = E(\Phi_r^*(U_1, \ldots, U_r) \mid U_1 = u_1).$$

For that end we need to calculate two integrals

$$I_1 = E\mathbf{1}_{\{u_1+U_2>1, u_1+U_3>1, \ldots, u_1+U_r>1\}} = u_1^{r-1} \quad \text{and}$$
$$I_2 = E\mathbf{1}_{\{u_1+U_2>1, U_3+U_2>1, \ldots, U_r+U_2>1\}} = \frac{1}{r-1}(1-(1-u_1)^{r-1}).$$

It follows that

$$\psi_r^*(u_1) = \frac{1}{r}(I_1 + (r-1)I_2) - \frac{1}{r} = \frac{1}{r}(u_1^{r-1} - (1-u_1)^{r-1}).$$

In the sequel the variance

$$\sigma_r^2 = E(\psi_r^*(U_1))^2 = 2r^{-2}\left(\frac{1}{2r-1} - \frac{((r-1)!)^2}{(2r-1)!}\right)$$

will play the important role. As far as for any $r \geq 2$ it is true that

$$\frac{1}{2r-1} > \frac{((r-1)!)^2}{(2r-1)!} \Leftrightarrow \binom{2r-2}{r-1} > 1,$$

and hence $\sigma_r^2 > 0$, we conclude that all the U-statistics S_r are non-degenerate. Consequently, by virtue of the Central Limit Theorem for U-statistics (see, for example, Serfling (1980)) under the validity of H_0 we have

$$\sqrt{n}\frac{S_r^*}{r\sigma_r} \longrightarrow N(0,1)$$

in distribution. It is possible to use this asymptotic relation for the approximate evaluation of significance values.

As the kernels of the non-degenerate statistics S_r^* are bounded, they obey the theorems on LD asymptotics of U-statistics, proven in Nikitin and Ponikarov (1998, 1999*a*, 1999*b*). We can claim that under the hypothesis H_0

$$\lim_{n\to\infty} n^{-1}\ln \Pr\{S_r \geq \varepsilon\} = h_r(\varepsilon),$$

where the function $h_r(\varepsilon)$ is analytic for sufficiently small $\varepsilon > 0$ and can be expanded in the series

$$h_r(\varepsilon) = \sum_{j=2}^{\infty} a_j\varepsilon^j, \tag{2.1}$$

Table 1. Expressions for σ_r^2, a_2 and for local indices.

r	σ_r^2	a_2	Local index
2	$\frac{1}{12}$	$-\frac{3}{2}$	$12(\int f^2(x)dx)^2$
3	$\frac{1}{27}$	$-\frac{3}{2}$	$48(\int F(x)f^2(x)dx)^2$
4	$\frac{19}{1120}$	$-\frac{35}{19}$	$\frac{2520}{19}(\int F^2(x)f^2(x)dx)^2$
5	$\frac{23}{2625}$	$-\frac{105}{46}$	$\frac{6720}{23}(\int F^3(x)f^2(x)dx)^2$

with

$$a_2 = -\frac{1}{2r^2\sigma_r^2}, \tag{2.2}$$

and

$$a_3 = \frac{1}{6r^3\sigma_r^6}\int_0^1 \psi_r^3(s_1)ds_1 + \frac{r-1}{2r^3\sigma_r^6}\int_0^1 \cdots \int_0^1 \Phi_r^*(s_1,\ldots,s_r)\psi_r^*(s_1)\psi_r^*(s_2)ds_1\cdots ds_r.$$

The coefficients a_2 for several first values of r, calculated using formula (2.2), are collected in Table 1. They are sufficient to find the main part of local exact Bahadur slopes (the so-called Bahadur local indices). Clearly

$$b_r(\theta) = E_\theta(S_r) = \int_{-\infty}^{\infty} (F(x+2))^{r-1} f(x)dx - \frac{1}{r}.$$

Denote by $\boldsymbol{F}_1^*$ the class of bounded and almost everywhere continuous densities. It follows from Maesono (1987) that for $f \in \boldsymbol{F}_1^*$

$$b_r(\theta) \sim 2(r-1)\left(\int_{-\infty}^{+\infty} F^{r-2}(x)f^2(x)dx\right)\theta, \qquad \theta \to 0. \tag{2.3}$$

Now it is possible to find the local exact slope of the sequence of statistics S_r (see Bahadur (1971) or Nikitin (1995) for corresponding definitions and results)

$$c_r(\theta) = -2h_r(b_r(\theta)). \tag{2.4}$$

Using (2.3) and substituting it in (2.4), we get that the main term of the expansion is of order θ^2 when $\theta \to 0$. Hence it is possible to compare only the coefficients at θ^2. These coefficients (local indices) are also given in Table 1. The integration there and later if not specified is always taken over R^1.

In Table 2 we present for same r the values of local indices, corresponding to the logistic, normal and Cauchy distributions. Note that the potential maximum of these indices coincides with the Fisher information and equals 0.3333... for the logistic density, 0.5 for the Cauchy density and 1 for the normal density (see Section 4 and Nikitin (1995) for explanation).

It turns out that in Table 2 the indices in the first line for $r = 2$ and in the second line for $r = 3$ coincide. It is not by chance as for any symmetric density f it holds true that

$$4\left(\int F(x)f^2(x)dx\right)^2 = \left(\int f^2(x)dx\right)^2. \tag{2.5}$$

Table 2. Local indices for some densities.

r	Logistic density	Cauchy density	Normal density
2	0.3333	0.3040	0.9549
3	0.3333	0.3040	0.9549
4	0.3316	0.2686	0.9764
5	0.3246	0.2242	0.9863

It means that the indices for $r = 2$ and $r = 3$ in Table 1 coincide. Hence, the discrimination between S_3 and W should be done on the basis of a factor at θ^3 in the expression for $c_3(\theta)$.

3. Solution of Maesono's problem

To calculate the factor at θ^3, we must take into consideration the second term in the expansion of the function h_3.

To find the coefficient a_3 for arbitrary r we notice, first of all, that by antisymmetry of the function $\psi_r^*(s)$ with respect to $s = 1/2$ the first integral in the expression for a_3 given above is equal to zero. Now denote for brevity for $r > 1$

$$\tau_r(s) = s^r - (1-s)^r, \qquad 0 \le s \le 1.$$

Now we must evaluate the integrals

$$\begin{aligned} J_r^{(1)} &= \int_0^1 \cdots \int_0^1 \mathbf{1}_{\{s_1+s_2>1,\dots,s_1+s_r>1\}} \tau_{r-1}(s_1)\tau_{r-1}(s_2)ds_1 \cdots ds_r \\ &= \frac{1}{r}\int_0^1 s_1^{r-2}\tau_{r-1}(s_1)(1 - s_1^r - (1-s_1)^r)ds_1 \end{aligned}$$

and (for $r > 3$)

$$\begin{aligned} J_r^{(2)} &= \int_0^1 \cdots \int_0^1 \mathbf{1}_{\{s_3+s_1>1,s_3+s_2,\dots,s_3+s_r>1\}} \tau_{r-1}(s_1)\tau_{r-1}(s_2)ds_1 \cdots ds_r \\ &= \frac{1}{r^2}\int_0^1 s_3^{r-3}\left(1 - s_3^r - (1-s_3)^r\right)^2 ds_3. \end{aligned}$$

It is easily found that

$$a_3 = \frac{r-1}{2r^3\sigma_r^6}\left(\frac{2}{r^3}J_r^{(1)} + \frac{r-2}{r^3}J_r^{(2)}\right).$$

For $r = 2$ the integral $J_2^{(2)}$ vanishes and

$$J_2^{(1)} = \frac{1}{2}\int_0^1 (2s-1)(1-s^2-(1-s)^2)ds = 0.$$

Therefore $a_3 = 0$ and in formula (2.1) the member with ε^3 disappears so that

$$h_2(\varepsilon) = -\frac{3}{2}\varepsilon^2 + O(\varepsilon^4).$$

For $r = 3$ we have

$$J_3^{(1)} = \frac{1}{3}\int_0^1 s(s^2 - (1-s)^2)(1 - s^3 - (1-s)^3)ds = \frac{1}{60},$$
$$J_3^{(2)} = \frac{1}{9}\int_0^1 (1 - s^3 - (1-s)^3)^2 ds = \frac{1}{30},$$

whence $a_3 = \frac{9}{5}$.

Consequently the asymptotics (2.1) looks as follows:

$$h_3(\varepsilon) = -\frac{3}{2}\varepsilon^2 + \frac{9}{5}\varepsilon^3 + O(\varepsilon^4).$$

Denote by $\boldsymbol{F}_2^*$ the class of densities with bounded and almost everywhere continuous derivative. For $f \in \boldsymbol{F}_2^*$ we can find the expansion of the function $b_r(\theta)$ with the accuracy up to θ^3 as $\theta \to 0$, namely

$$\begin{aligned} b_2(\theta) &= 2\int f^2(x)dx\theta + O(\theta^3), \\ b_r(\theta) &= 2(r-1)\int F^{r-2}(x)f^2(x)dx\theta \\ &\quad + (r-1)(r-2)\int F^{r-3}(x)f^3(x)dx\theta^2 + O(\theta^3), \qquad r \geq 3. \end{aligned}$$

Substituting the expansion of the function $b_r(\theta)$ in (2.4), we obtain for $r = 2$ with the accuracy up to θ^3 as $\theta \to 0$ the expression

$$c_2(\theta) = 12\left(\int f^2(x)dx\right)^2 \theta^2 + O(\theta^4).$$

Similarly for $r = 3$ we have a formula

$$\begin{aligned} c_3(\theta) &= 12\left(\int f^2(x)dx\right)^2 \theta^2 \\ &\quad + 24\int f^2(x)dx\left(\int f^3(x)dx - \frac{6}{5}\left(\int f^2(x)dx\right)^2\right)\theta^3 + O(\theta^4), \end{aligned}$$

where some simplifications were obtained due to (2.5) and due to the integration by parts.

So, for small $\theta > 0$ the statistic S_3 wins in comparison with W, if

$$24\int f^2(x)dx\left(\int f^3(x)dx - \frac{6}{5}\left(\int f^2(x)dx\right)^2\right) > 0, \tag{3.1}$$

and loses in case of inverse inequality. We emphasize that the sign of the inequality depends on the distribution of observations, more precisely (and somewhat mysteriously) on the sign of the expression (3.1).

This is the local solution of the "Maesono's problem" on the comparison of W and S_3 stated above.

For example, for the Cauchy distribution the left-hand side of (3.1) equals 0.0290, for the double exponential distribution it equals 0.05, for the normal distribution it equals -0.0244 and for the logistic distribution it is zero.

So, it is preferable to use S_3 for the Cauchy and double exponential distribution, and it is recommended to use the more complicated statistics S_r for the normal distribution. The question about the logistic distribution remains open, because the exact Bahadur slopes S_3 and W coincide for small θ in first two members.

These results are in good accordance with the results of Maesono (1987) who compared S_3 and W using the expansions of the asymptotic power under contiguous alternatives. In the same time they are in contradiction to the conclusion of Maesono (1987), who got the formula (1.3) using not the exact but the approximate Bahadur efficiency. It is well-known that these two types of efficiency may lead to different issues (see Bahadur (1960, 1971), Nikitin (1995)) but the exact variant is certainly a preferable one.

4. Local optimality

We return now to the Bahadur local exact slope of the statistics S_r. It follows from (2.4) that for $r \geq 2$

$$c_r(\theta) \sim \frac{4(r-1)^2}{r^2\sigma_r^2}\left(\int_{-\infty}^{\infty} F^{r-2}(x)f^2(x)dx\right)^2\theta^2, \qquad \theta \to 0. \tag{4.1}$$

In order to simplify (4.1) we denote

$$\Delta_r = \frac{4(r-1)^2}{r^2\sigma_r^2}, \qquad \varphi(x) = f(F^{-1}(x)).$$

Now (4.1) can be written in the form

$$c_r(\theta) \sim \Delta_r\left(\int_0^1 x^{r-2}\varphi(x)dx\right)^2\theta^2, \qquad \theta \to 0.$$

Now let us consider the class $\boldsymbol{F}_3^*$ of densities f from $\boldsymbol{F}_1^*$ for which the following relation (4.2) between the Kullback-Leibler and Fisher informations is true. This condition is well-known in asymptotic statistics. Suppose that

$$K(\theta) = \int \ln\frac{2f(x-\theta)}{f(x-\theta)+f(x+\theta)}f(x-\theta)dx \sim \frac{1}{2}I(f)\theta^2, \qquad \theta \to 0, \tag{4.2}$$

where $0 < I(f) < \infty$ and

$$I(f) = \int \frac{f'^2(x)}{f(x)}dx$$

is the Fisher information. The condition $I(f) < \infty$ implies the boundary conditions

$$\varphi(0) = \varphi(1) = 0 \tag{4.3}$$

(see, e.g., Nikitin (1995), Chapter 6). It is well known that the following inequality of Bahadur-Raghavachari is true under for densities of $\boldsymbol{F}_3^*$ (see Bahadur (1971) or Nikitin (1995))

$$\Delta_r\left(\int_0^1 x^{r-2}\varphi(x)dx\right)^2 \leq \int_0^1 \varphi'^2(x)dx. \tag{4.4}$$

We are interested in the question, whether the equality is possible in (4.4) for some φ corresponding to the density $f \in \boldsymbol{F}_3^*$. This direction of research was initiated by one of the authors in 80-s, see for details Nikitin (1995), Chapter 6.

Note that it is necessary to take into account the condition of symmetry $\varphi(x) = \varphi(1-x)$, $x \in [0,1]$. For such φ consider the identity

$$\int_0^1 x^{r-2}\varphi(x)dx = \frac{1}{2}\int_0^1 x^{r-2}\varphi(x)dx + \frac{1}{2}\int_0^1 (1-x)^{r-2}\varphi(x)dx$$
$$= \int_0^1 \frac{x^{r-2}+(1-x)^{r-2}}{2}\varphi(x)dx.$$

Therefore inequality (4.4) takes the equivalent form

$$\frac{1}{4}\Delta_r\left(\int_0^1 (x^{r-2}+(1-x)^{r-2})\varphi(x)dx\right)^2 \le \int_0^1 \varphi'^2(x)dx. \tag{4.5}$$

We minimize now the right-hand side of (4.5) under the normalizing condition

$$\int_0^1 (x^{r-2}+(1-x)^{r-2})\varphi(x)dx = 1$$

and the boundary conditions (4.3). The Euler-Lagrange equation with the undetermined multiplier λ has after the simplification the following form:

$$\varphi''(x) + \lambda(x^{r-2}+(1-x)^{r-2}) = 0, \qquad \varphi(0) = \varphi(1) = 0.$$

It is well-known that the unique solution of such equation is given by means of the Green formula for the two-point boundary problem:

$$\varphi(x) = C\int_0^1 (\min(t,x) - tx)(t^{r-2}+(1-t)^{r-2})dt = \frac{C}{r(r-1)}(1-x^r-(1-x)^r).$$

It is easy to verify that for such function φ inequality (4.4) becomes the equality.

Thus to find the "optimum" distribution for any r we must solve the equation

$$F_r'(x) = \varphi(F_r(x)). \tag{4.6}$$

In particular, for the Wilcoxon statistic W when $r = 2$ and for the statistic S_3 when $r = 3$ we get $\varphi(x) = C_1x(1-x)$, $C_1 > 0$ that corresponds to the logistic distribution with $F(x) = (1+\exp(-x))^{-1}$.

In the general case for arbitrary $r > 3$ the probabilistic solution of equation (4.6) cannot be found explicitly. However this decision F_r always exists, so for any $r > 3$ the appropriate local Bahadur exact slope $c_r(\theta)$ under F_r attains its maximum and will be larger for close alternatives than the slopes $c_2(\theta)$ and $c_3(\theta)$.

Acknowledgements

The authors express their gratitude to the Editor and to the referee for careful reading of manuscript and valuable comments.

References

Bahadur, R. R. (1960). Stochastic comparison of tests, *Ann. Math. Statist.*, **31** (2), 276–295.

Bahadur, R. R. (1971). *Some Limit Theorems in Statistics*, Siam, Philadelphia.

Dasgupta, R. (1984). On large deviation probabilities of U-statistics in non-i.i.d. case, *Sankhyā Ser. A*, **46** (1), 110–116.

Maesono, Y. (1987). Competitors of the Wilcoxon Signed Rank Test, *Ann. Inst. Statist. Math.*, **39**, 363–375.

Nikitin, Ya. Yu. (1995). *Asymptotic Efficiency of Nonparametric Tests*, Cambridge University Press, New York.

Nikitin, Ya. Yu. and Ponikarov, E. V. (1998). On Chernoff type large deviations for U- and V-statistics, *Asymptotic Methods in Probability and Mathematical Statistics, Abstracts of Communications of International Conference, St. Petersburg, June 24–28, 1998*, 226–228.

Nikitin, Ya. Yu. and Ponikarov, E. V. (1999*a*). Rough large deviation asymptotics of Chernoff type for von Mises functionals and U-statistics, *Proceedings of St. Petersburg Mathematical Society*, 1999, Vol. 7, 124–167 (English translation (2001), *Amer. Math. Soc. Trans. Ser. 2*, **203**, 107–146).

Nikitin, Ya. Yu. and Ponikarov, E. V. (1999*b*). Large deviations of Chernoff type for U- and V-statistics, *Doklady Russian Academy of Sciences*, **369** (1), 10–12.

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