Computable Error Bounds for Approximations of Transformed Chi-Squared Variables and Its Statistical Applications

VLADIMIR V. ULYANOV*, GERD CHRISTOPH**

and YASUNORI FUJIKOSHI***

**Faculty of Computational Mathematics and Cybernetics,*

Moscow State University, Moscow, Russia,

***Department of Mathematics, University of Magdeburg,*

Magdeburg, Germany

and

****Department of Mathematics,*

Graduate School of Science and Engineering,

Chuo University, Bunkyo-ku, Tokyo, Japan

Abstract

We get computable error bound of order $O(n^{-1})$ for chi-squared with 1 degree of freedom approximation of transformed chi-squared random variable with n degrees of freedom. The result is applied for likelihood ratio statistics in multivariate case.

AMS 1991 subject classification: primary 62H10; secondary 62E20

Key Words and Phrases: Computable error bound, chi-squared approximation, likelihood ratio statistics, transformation of random variables

Abbreviated title: Computable error bound of chi-squared approximation

1 Introduction

Let $\mathcal{X}_n^2$ be a random variable having chi-squared distribution with n degrees of freedom and density

$$p_{\mathcal{X}_n^2}(x) = \frac{1}{2^{n/2}\,\Gamma(n/2)}\; x^{-1+n/2}\; e^{-x/2}\, I_{(0\,,\infty)}(x)\,,$$

where $I_A(x)$ denotes indicator function of set A. We consider a transformed chi-squared statistic defined by

$$T_1 = \mathcal{X}_n^2 - n\log\frac{\mathcal{X}_n^2}{n} - n.$$

The distribution of T_1 appears as the null distribution of LR (likelihood ratio) statistic for testing a hypothesis that the variance σ^2 is equal to a given value, based on a sample of $n+1$ observations from a normal population $N(\mu,\sigma^2)$. Note that for large n

$$P(T_1 \le x) = G_1(x) + O(n^{-1})\,,$$

where $G_m(x)$ is the distribution function of chi-squared random variable $\mathcal{X}_m^2$ with m degrees of freedom. One of our main purposes is to show that

$$\sup_x |P(T_1 \le x) - G_1(x)| \le B(n),$$

where $B(n)$ is a computable constant and $B(n) = O(n^{-1})$ as $n \to \infty$.
In this paper we also obtain Berry-Esseen type bound for asymptotic approximation of the distribution of

$$T_p = \mathrm{tr}W - n\log|\frac{1}{n}W| - np,$$

which is an extension of T_1, where the random matrix W has a Wishart distribution $W_p(n, I_p)$. The distribution of T_p appears as the null distribution of LR statistic for testing a hypothesis that the covariance matrix Σ is equal to a given covariance matrix. In fact, we prove that

$$\sup_x |\, P(T_p \le x) - G_q(x)\,| \le B(p,n),$$

where $q = \frac{1}{2}p(p+1)$ and $B(p,n)$ is a computable constant, depending only on p and n and $B(p,n) = O(n^{-1})$ as $n \to \infty$.

In Section 2 we formulate the main results. In Section 3 we give an outline for the method of deriving Theorem 1. It is shown that the result will be obtained by estimating the three integrals J_1, J_2 and J_3. In Section 5 we estimate the summands J_1, J_2 and J_3, based on auxiliary results given in Section 4. A proof of Theorem 1 is given in Section 6. In Section 7 we show that Theorem 2 can be obtained from Theorem 1 with help of some basic properties. The possible generalizations of Theorem 1 are discussed in Section 8.

2 Main Results

Let $p \geq 1$ and $n \geq 1$ be integers and $p \leq n$. Put

$$D(n) = \frac{1.9}{n}\left(\frac{n}{n-1}\right)^2 + \frac{15.59}{n}\,0.9906^n + C_1(n)\,0.9894^n$$

with $C_1(n) = 15.21\,/\,(n-4)$ for $n > 32$ or $C_1(n) = 0.5271$ for $4 \leq n \leq 32$.

Theorem 1 *We have*

$$\sup_x |P(T_1 \leq x) - G_1(x)| \leq B(n)\,, \tag{1}$$

where

$$B(n) = 2\,D(n) + \frac{1.877}{n} + \frac{1.1284}{\sqrt{n}}\;0.7788^n\,.$$

Remark 1. It is easy to see that $B(n) = O(n^{-1})$ as $n \to \infty$.

Theorem 2 *We have*

$$\sup_x |P(T_p \leq x) - G_q(x)| \leq B(p,n)\,, \tag{2}$$

where

$$B(p,n) = \sum_{i=1}^{p} B(n_i),.$$

Remark 2. It is easy to see that $B(p,n) = O(n^{-1})$.

n	$B(n)$	$B(2,n)$	$B(3,n)$
100	0.28796	0.58123	0.87994
200	0.07057	0.14192	0.21407
300	0.02932	0.05884	0.08857
400	0.01710	0.03428	0.05153

3 Outline of the Proof of Theorem 1

Put $h(y) = \sqrt{2n}\,y - n\log(1 + \sqrt{\frac{2}{n}}\,y)$. It is easy to see that

$$T_1 = h(V_n) \quad \text{with} \quad V_n = (\mathcal{X}_n^2 - n)/\sqrt{2\,n}\,. \tag{3}$$

Note that V_n is a random variable $\mathcal{X}_n^2$ standardized with $E(\mathcal{X}_n^2) = n$ and $Var(\mathcal{X}_n^2) = 2\,n$. Hence, by the central limit theorem the distribution function

$$F_n(x) = P(V_n \le x) \; = \; P(\mathcal{X}_n^2 - n \le \sqrt{2\,n}\,x)$$

tends to the normal law $\Phi(x) = \frac{1}{\sqrt{2\,\pi}} \int_{-\infty}^{x} e^{-u^2/2}\,du$ as $n \to \infty$.
With $E(V_n^3) = 2\sqrt{2/n}$ and the first-order Chebyshev-Edgeworth expansion

$$\Phi_n(x) = \Phi(x) + \frac{\sqrt{2}\,(1 - x^2)}{3\sqrt{n}} \, \frac{1}{\sqrt{2\,\pi}} e^{-x^2/2}$$

one finds $F_n(x) = \Phi_n(x) + O(1/n)$ as $n \to \infty$.

Dobrić and Ghosh (1996) gave a bound of the remainder term (see Example 3 in the mentioned paper with $a = b = 1/2$), proving

$$\sup_x |F_n(x) - \Phi_n(x)| \le \frac{1.9}{n} \left(\frac{n}{n-1}\right)^2 + \frac{15.59}{n} 0.9906^n + C_1(n)\,0.9894^n \tag{4}$$

with $C_1(n) = 15.21\,/\,(n-4)$ for $n > 32$ or $C_1(n) = 0.5271$ for $4 \le n \le 32$.

Define

$$p_{V_n}(x) = \frac{d}{dx} P(V_n \le x)\,, \quad \varphi(x) = \frac{d}{dx}\Phi(x)\,, \quad \varphi_n(x) = \frac{d}{dx}\Phi_n(x)\,, \tag{5}$$

$$B_x = \{\, y \in \mathbb{R} : h(y) \le x \,\} \quad \text{and} \quad A_x = \{\, y \in \mathbb{R} : |y| \le \sqrt{x} \,\}\,.$$

Then

$$
\begin{aligned}
P(T_1 \le x) - G_1(x) &= P(h(V_n) \le x) - [\Phi(\sqrt{x}) - \Phi(-\sqrt{x})] \\
&= \int_{B_x} p_{V_n}(y)\, dy - \int_{A_x} \varphi(y)\, dy \\
&= J_1 + J_2 + J_3 \,, \qquad (6)
\end{aligned}
$$

where

$$
\begin{aligned}
J_1 &= \int_{B_x} (p_{V_n}(y) - \varphi_n(y))\, dy \,, \\
J_2 &= \int_{B_x} \varphi(y)\, dy - \int_{A_x} \varphi(y)\, dy \,, \\
J_3 &= \int_{B_x} (\varphi_n(y) - \varphi(y))\, dy - \int_{A_x} (\varphi_n(y) - \varphi(y))\, dy \,.
\end{aligned}
$$

Here we used

$$\int_{A_x} (\varphi_n(y) - \varphi(y))\, dy = 0$$

because A_x is a symmetric set and the function $\varphi_n(y) - \varphi(y)$ is odd.

In Section 5 we get bounds for J_1, J_2 and J_3 of order $O(n^{-1})$ using (4) and the following facts:

– Lebesgue measure of the set $\{A_x \triangle B_x\}$ is of order $O(n^{-1/2})$;

– the function $\varphi(y)$ is even;

– the function $\varphi_n(y) - \varphi(y)$ is odd and is of order $O(n^{-1/2})$.

See the detailed proof of Theorem 1 in Section 6. The possible generalizations of Theorem 1 are discussed in Section 8.

4 Auxiliary Results

Let $f(y) = y - \log(1+y)$ for $y > -1$.

Lemma 1 *Let t be a real number such that*

$$0 < t < t_0 \;=\; -5 + \sqrt{40} \;=\; 1.3245 \,. \qquad (7)$$

Assume that y_t and $\overline{y}_t$ satisfy the conditions:

$$f(y_t) = f(\overline{y}_t) = t^2/2 \quad \textit{with} \quad y_t > 0 \quad \textit{and} \quad \overline{y}_t < 0 \,. \qquad (8)$$

Then we have

$$t+\frac{t^2}{3}+\frac{t^3}{45} \,<\, y_t \,<\, t+\frac{t^2}{3}+\frac{t^3}{36} \tag{9}$$

and

$$-t+\frac{t^2}{3}-\frac{t^3}{27} \,<\, \overline{y}_t \,<\, -t+\frac{t^2}{3}-\frac{t^3}{36}\,. \tag{10}$$

Remark 1. Note $f(y)$ is decreasing when $y \in (-1\,,\,0)$ and it is increasing when $y > 0$. Since $f(y)$ is continuous function, $\lim_{y\downarrow -1} f(y) = \lim_{y\uparrow +\infty} f(y) = +\infty$ and $f(0) = 0$, the solutions y_t and $\overline{y}_t$ are defined uniquely by the conditions (8).

Remark 2. One can write the solution $\overline{y}_t < 0$ of $f(\overline{y}_t) = t^2/2$ in terms of the real Lambert's W-function. We recall that the real Lambert's W-function is defined to be the function satisfying $W(u)\, e^{W(u)} = u$, i.e. it is the inverse function of $z\, e^z$ for $z \geq -1$, i.e. if $z\, e^z = u$, then $z = W(u)$. Some properties and applications of the real Lambert's W-function are given in Corless, Gonnet, Hare, Jeffrey and Knuth(1996), e.g. series representation holds

$$W(z) = \sum_{n=1}^{\infty} \frac{(-\,n)^{n-1}}{n\,!}\, z^n = z - z^2 + \frac{3}{2}\, z^3 - \frac{8}{3}\, z^4 + \frac{125}{24}\, z^5 + O(z^6)\,, \quad z \to 0\,,$$

which absolutely converges for $|z| < 1/e$. In order to find a solution of the equation

$$f(y) := y \,-\, \log(1+y) \,=\, t^2/2 \tag{11}$$

provided $-1 < y < 0$, we put $z = 1 + y$. Then (11) can be written in equivalent form as

$$(-z)\, e^{-z} = -\, \exp\{-\,1 - t^2/2)\}.$$

Therefore, using Taylor expansion with $W^{(k)}(0) = (-k)^{k-1}$ for $k = 0, 1, ...$ we get

$$\begin{aligned} \overline{y}_t &= -\,W(-\,\exp\{-1 - t^2/2\}) \,-\, 1 \\ &= -t + (1/3)\, t^2 - (1/36)\, t^3 - (1/270)\, t^4 - (1/4320)\, t^5 + O(t^6)\,, \quad t \to 0\,. \end{aligned}$$

Proof. All four inequalities in (9) and (10) can be proved in the same way. At first we prove the left hand-side inequality of (9). Since $f(y)$ is increasing when $y > 0$, the left-hand side inequality of (9) will be proved if we show that

$$f(t + t^2/3 + t^3/45) < f(y_t) = t^2/2, \qquad (12)$$

provided t satisfies (7). Put

$$\begin{aligned} \lambda_1(t) &= f(t + t^2/3 + t^3/45) - t^2/2 \\ &= t - t^2/6 + t^3/45 - \ln(1 + t + t^2/3 + t^3/45). \end{aligned}$$

We have $\lambda_1(0) = 0$. Therefore, in order to prove (12) it is sufficient to show that the derivative $\lambda_1'(t)$ is negative when t satisfies (7). It is easy to see that

$$\lambda_1'(t) = \frac{t^3\,(t^2 + 10\,t - 15)}{675\,(1 + t + t^2/3 + t^3/45)}.$$

The quadratic equation $t^2 + 10\,t - 15 = 0$ has the solutions $-5 \pm \sqrt{40}$. Hence $\lambda_1'(t) < 0$ and also $\lambda_1(t) < 0$ for t satisfying (7).
Note that $\lambda_1(t) < 0$ for $0 < t < t_0^*$, where t_0^* is the solution of

$$t - t^2/6 + t^3/45 - \ln(1 + t + t^2/3 + t^3/45) = 0,$$

where $1.7 < t_0^* < 1.75$. Hence we may enlarge the interval in (7) until t_0^*.

The other inequality in (9) we find in a similar way. Put

$$\lambda_2(t) = f(t + t^2/3 + t^3/36) - t^2/2.$$

We have $\lambda_2(0) = 0$ and

$$\lambda_2'(t) = \frac{t^4\,(t + 8)}{432\,(1 + t + t^2/3 + t^3/36)}.$$

Hence the derivative $\lambda_2'(t)$ is positive and $\lambda_2(t) > 0$.

Since $f(y)$ is decreasing when $y \in (-1, 0)$, the inequalities of (10) will be proved if we show that

$$f(-t + t^2/3 - t^3/27) < f(\overline{y}_t) = t^2/2 < f(-t + t^2/3 - t^3/36) \qquad (13)$$

provided t satisfies (7). Put

$$\lambda_3(t) = f(-t + t^2/3 - t^3/27) - t^2/2\,, \quad \lambda_4(t) = f(-t + t^2/3 - t^3/36) - t^2/2\,.$$

Then $\lambda_3(0) = 0 = \lambda_4(0)$ and we obtain (13) by

$$\begin{aligned} \lambda_3'(t) &= \frac{t^3\,(t^2 - 6\,t + 9)}{243\,(1 - t + t^2/3 - t^3/27)} > 0 \quad \text{and} \\ \lambda_4'(t) &= \frac{t^4\,(t - 8)}{432\,(1 - t + t^2/3 - t^3/36)} < 0. \end{aligned}$$

Lemma 1 is proved. Now we consider a function

$$h(y) = \sqrt{2n}\; y - n \log\left(1 + \sqrt{2/n}\; y\right)\,.$$

Obviously,

$$h(y) = n\, f\left(\sqrt{2/n}\; y\right)\,. \tag{14}$$

Lemma 2 *Let t_0 be the same as in Lemma 1 and x satisfy*

$$0 < \sqrt{2x/n} < t_0 \;=\; -5 + \sqrt{40} \;=\; 1.3245. \tag{15}$$

Let $y_1(x)$, $y_2(x)$ be such that

$$h(y_1(x)) = h(y_2(x)) = x \quad \textit{with} \quad y_1(x) > 0 \quad \textit{and} \quad y_2(x) < 0\,. \tag{16}$$

Then we have

$$\sqrt{x} + \sqrt{\frac{2}{n}}\,\frac{x}{3} + \frac{2}{n}\,\frac{x^{3/2}}{45} \quad < y_1(x) < \quad \sqrt{x} + \sqrt{\frac{2}{n}}\,\frac{x}{3} + \frac{2}{n}\,\frac{x^{3/2}}{36} \tag{17}$$

$$-\sqrt{x} + \sqrt{\frac{2}{n}}\,\frac{x}{3} - \frac{2}{n}\,\frac{x^{3/2}}{27} \quad < y_2(x) < \quad -\sqrt{x} + \sqrt{\frac{2}{n}}\,\frac{x}{3} - \frac{2}{n}\,\frac{x^{3/2}}{36}\,. \tag{18}$$

Proof: Inequalities (17) and (18) follow from (14) and (9), (10) when we take $t = \sqrt{2x/n}$ in Lemma 1 considering

$$y_t = \sqrt{\frac{2}{n}}\; y_1(x) \quad \text{and} \quad \overline{y}_t = \sqrt{\frac{2}{n}}\; y_2(x)\,.$$

Lemma 2 is proved.

5 Bounds for J_1, J_2 and J_3 from (6) under (15)

Recall that $A_x = (-\sqrt{x}, \sqrt{x})$ and $B_x = \{\, y : h(y) \le x \,\}$. The function $h(y)$ is decreasing for $y \in (-\sqrt{n/2}\,,\, 0)$, $h(0) = 0$ and $h(y)$ is increasing for $y > 0$ (cp. Remark 2 after formulation of Lemma 1). Therefore, the set B_x is in fact an interval $(\, y_2(x)\,,\, y_1(x)\,)$ according to definitions of $y_1(x)$ and $y_2(x)$ in (16). Now we show how to get bounds for

$$
\begin{aligned}
J_1 &= \int_{B_x} (p_{V_n}(y) - \varphi_n(y))\, dy\,, \\
J_2 &= \int_{B_x} \varphi(y)\, dx - \int_{A_x} \varphi(y)\, dy\,, && (19)\\
J_3 &= \int_{B_x} \Big(\varphi_n(y) - \varphi(y)\Big)\, dy - \int_{A_x} \Big(\varphi_n(y) - \varphi(y)\Big)\, dy\,. && (20)
\end{aligned}
$$

At first we consider J_1. We get (see (5))

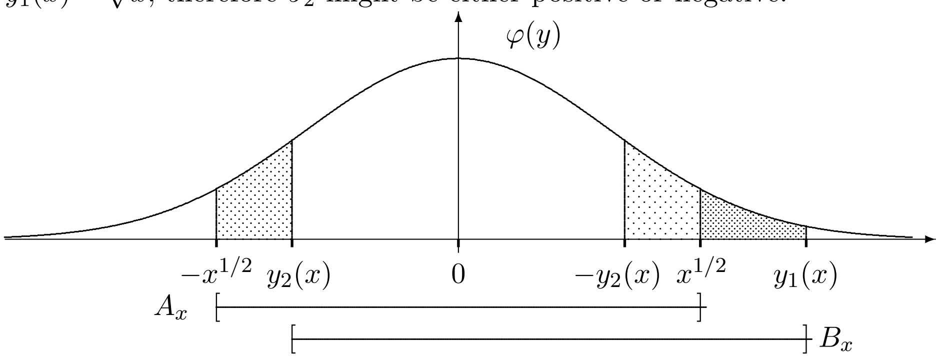
$$
J_1 = \Big(\, F_n(y_1(x)) \;-\; \Phi_n(y_1(x))\Big) \;-\; \Big(\, F_n(y_2(x)) \;-\; \Phi_n(y_2(x))\Big)\,,
$$

which leads to

$$
|J_1| \le 2\, \sup_x\, |F_n(x) - \Phi_n(x)|\,. \tag{21}
$$

With (4) we obtain the bound for J_1.

Now we estimate J_2. It follows from (17) and (18) that $\sqrt{x} - (-y_2(x)) \le y_1(x) - \sqrt{x}$, therefore J_2 might be either positive or negative.



If $J_2 < 0$ then it follows from Lemma 2 and (19) that

$$\begin{aligned} |J_2| = -J_2 &= [\Phi(\sqrt{x}) - \Phi(-\sqrt{x})] - [\Phi(y_1(x)) - \Phi(y_2(x))] \\ &= 2\,\Phi(\sqrt{x}) - [\Phi(y_1(x)) + \Phi(-y_2(x))] \\ &\le 2\,\Phi\Big(\sqrt{x} + \frac{b_1+b_2}{2}\Big) \\ &\qquad - \big[\,\Phi(\sqrt{x} + a + b_1) + \Phi(\sqrt{x} - a + b_2)\big]\,, \end{aligned}$$

where

$$a = \frac{x}{3}\sqrt{\frac{2}{n}}, \quad b_1 = \frac{x^{3/2}}{45}\,\frac{2}{n} \quad \text{and} \quad b_2 = \frac{x^{3/2}}{36}\,\frac{2}{n}\,.$$

Using the second order Taylor expansion of both functions $\Phi(\sqrt{x} + a + b_1)$ and $\Phi(\sqrt{x} - a + b_2)$ at the point $\sqrt{x} + (b_1 + b_2)/2$ we find

$$|J_2| \le -\frac{1}{2}\Big(\varphi'(y^*) + \varphi'(y^{**})\Big)\Big(a + \frac{b_1 - b_2}{2}\Big)^2 \tag{22}$$

with $\sqrt{x} - a + b_2 \le y{**} \le \sqrt{x} + (b_1 + b_2)/2 \le y{*} \le \sqrt{x} + a + b_1$ and $\varphi'(y) = -\frac{1}{\sqrt{2\pi}} \cdot y \cdot \exp(-y^2/2) < 0$ if $y > 0$.

Note that under condition (15) we have

$$\Big(a + \frac{b_1 - b_2}{2}\Big)^2 = \frac{2\,x^2}{9\,n}\left(1 - \frac{1}{60}\sqrt{\frac{2x}{n}}\right)^2 \le \frac{2\,x^2}{9\,n}.$$

Since $\sqrt{x} \le y^*$ then we find

$$x^2\,(-\,\varphi'(y^*)) \le \frac{1}{\sqrt{2\pi}}\,\max_{y\ge 0}\big\{y^5\,\exp\{-y^2/2\}\big\} = \frac{1}{\sqrt{2\pi}}\,\left(\frac{5}{e}\right)^{5/2}.$$

Let us replace condition (15) by stronger one

$$0 < \sqrt{2x/n} \le 1\,. \tag{23}$$

We have

$$\sqrt{x} - a + b_2 = \sqrt{x}\,(1 - u/3 + u^2/36) = \sqrt{x}\,(1 - u/6)^2 \quad \text{with} \quad u = \sqrt{2x/n} \le 1.$$

Since $\sqrt{x} - a + b_2 \le y^{**}$ and

$$\max_{0<u\le 1} \frac{(1 - u/60)^2}{(1 - u/6)^8} = \left(\frac{59}{60}\right)^2\,\left(\frac{6}{5}\right)^8,$$

we get

$$\begin{aligned}
- \quad \frac{9n}{2}\,\varphi'(y^{**})\Big(a+\frac{b_1-b_2}{2}\Big)^2 &= -\,\varphi'(y^{**})\frac{x^2\,(1-u/6)^8\,(1-u/60)^2}{(1-u/6)^8} \\
&\le \frac{1}{\sqrt{2\pi}}\,\left(\frac{59}{60}\right)^2\,\left(\frac{6}{5}\right)^8\,\max_{y\ge 0}\,\{y^5\,\exp\{-y^2/2\}\} \\
&= \frac{1}{\sqrt{2\pi}}\,\left(\frac{59}{60}\right)^2\,\left(\frac{6}{5}\right)^8\left(\frac{5}{e}\right)^{5/2}.
\end{aligned}$$

By (22) we find

$$|J_2| \le \frac{1}{9\cdot\sqrt{2\pi}\cdot n}\left(\frac{5}{e}\right)^{5/2}\left(1+\left(\frac{59}{60}\right)^2\,\left(\frac{6}{5}\right)^8\right) \le \frac{1.049}{n}. \tag{24}$$

Now assume that $J_2 > 0$. In this case Lemma 2 implies that

$$\begin{aligned}
0 < J_2 \quad &= \quad \int_{\sqrt{x}}^{y_1(x)} \varphi(y)dy - \int_{-\sqrt{x}}^{y_2(x)} \varphi(y)dy = \int_{\sqrt{x}}^{y_1(x)} \varphi(y)dy - \int_{-y_2(x)}^{\sqrt{x}} \varphi(y)dy \\
&\le \quad \int_{\sqrt{x}}^{\sqrt{x}+a+b_3} \varphi(y)dy - \int_{\sqrt{x}}^{\sqrt{x}+a-b_4} \varphi(y)dy = \int_{\sqrt{x}+a-b_4}^{\sqrt{x}+a+b_3} \varphi(y)dy,
\end{aligned}$$

where

$$b_3 = \frac{2}{n}\;\frac{x^{3/2}}{36} \quad \text{and} \quad b_4 = \frac{2}{n}\;\frac{x^{3/2}}{27} \quad \text{with} \quad b_3+b_4 = \frac{2}{n}\cdot\frac{7\,x^{3/2}}{108}.$$

Therefore,

$$J_2 \le \frac{14}{108\,n}\,x^{3/2}\,\varphi(\sqrt{x}) \le \frac{14}{108\,n\,\sqrt{2\pi}}\,\left(\frac{3}{e}\right)^{3/2} \le \frac{0.06}{n}.$$

Comparing this bound with (24) we get that for J_2 the inequality (24) holds when x satisfies (23).

Now we construct the following bound for J_3:

$$|J_3| \le \frac{0.828}{n}\;. \tag{25}$$

Define $m(y) = (y^3\,-\,3\,y)\,e^{-\,y^2/2}$. Then

$$\varphi_n(y) - \varphi(y) = \frac{1}{\sqrt{2\,\pi}}\;\frac{y^3\,-\,3\,y}{3\,\sqrt{n/2}}\;e^{-\,y^2/2} = \frac{m(y)}{3\,\sqrt{n\,\pi}}\;.$$

Recall that $y_1(x) > 0$ and $y_2(x) < 0$. Since $m(-y) = -\,m(y)$, we have

$$J_3 = \int_{\sqrt{x}}^{y_1(x)} \frac{m(y)}{3\,\sqrt{n\,\pi}}\,dy\;-\;\int_{-\,\sqrt{x}}^{y_2(x)} \frac{m(y)}{3\,\sqrt{n\,\pi}}\,dy\;=\;\int_{-\,y_2(x)}^{y_1(x)} \frac{m(y)}{3\,\sqrt{n\,\pi}}\,dy\,.$$

It follows from (17) and (18) under (23) that

$$y_1(x) \le \frac{49}{36}\,\sqrt{x}, \quad -\,y_2(x) \ge \frac{25}{36}\,\sqrt{x} \;\text{ and }\; y_1(x)\,+\,y_2(x) \le \sqrt{\frac{2}{n}\,\frac{2\,x}{3}}\,. \qquad (26)$$

Constructing bound for J_3 we consider different cases depending on the value of x. Write

$$J_3 = \sum_{i=1}^{5} J_{3i}\,,$$

where $J_{3i} = J_3 \cdot I_{A_i}(x)$ and $I_A(x)$ denotes indicator function of set A. We take

$$\begin{aligned} A_1 &= (11.3001,\, n/2], \quad A_2 = (6.2203,\, 11.3001], \quad A_3 = (5.9877,\, 6.2203], \\ A_4 &= (3.3834,\, 5.9877], \quad A_5 = (0,\, 3.3834]. \end{aligned}$$

It is clear that (25) will be proved when we show for $i = 1, \ldots, 5$

$$|J_{3i}| \le \frac{0.828}{n}\,. \qquad (27)$$

The function $m(y)$ has its extreme points at $y = \pm\sqrt{3 \pm \sqrt{6}}$ and we have to consider only the case $y > 0$. Then $m(y) < 0$ if $0 < y < \sqrt{3}$ and $m(y) > 0$ if $y > \sqrt{3}$.

At first we consider J_{31}. If $x \in A_1$ then $\sqrt{3 + \sqrt{6}} \le (25/36)\sqrt{x}$. Since the function $m(y)$ is decreasing for $y \ge \sqrt{3 + \sqrt{6}}$, we obtain

$$0 < m(y) \le m(-y_2(x)) \le m(25\,\sqrt{x}\,/\,36) \quad \text{for} \quad y \in \Big(-\,y_2(x)\,,\, y_1(x)\Big)$$

and

$$\begin{aligned} \int_{-\,y_2(x)}^{y_1(x)} m(y)\,dy &\le m(25\,\sqrt{x}/36)\,\Big(y_1(x) + y_2(x)\Big) \\ &\le \left(\left(\frac{25}{36}\right)^3 x^{3/2} - 3\left(\frac{25}{36}\right) x^{1/2} \right) e^{-\,(25/36)^2\,x/2} \sqrt{\frac{2}{n}\,\frac{2x}{3}}\,. \end{aligned}$$

The function $v(x) := (a^3\,x^{5/2} \,-\, 3\,a\,x^{3/2})\,e^{-\,a^2\,x/2}$ takes its maximum value at $x^* = (4 + \sqrt{7})\,a^{-2}$ and with $a = 25/36$ we find $v(x) \le v(x^*) = 4.670$. Hence, in the this case we have

$$|J_{31}| \le \frac{1}{\sqrt{2\,\pi}}\;\frac{4\;v(x^*)}{9\;n}\,,$$

which leads to (27) for $i = 1$.

Suppose now that $6.2203 < x \le 11.3001$, i.e. $(25/36)\sqrt{x} > \sqrt{3}$, then with $0 < m(y) \le m(\sqrt{3+\sqrt{6}}) = 0.3749$ for $y \in \Big(-y_2(x)\,,\, y_1(x)\Big)$ and $x \le 11.3001$ we obtain

$$|J_{32}| \le \frac{1}{\sqrt{2\,\pi}}\,\frac{4\,x}{9\,n}\,m(\sqrt{3+\sqrt{6}}) \le \frac{0.7512}{n}\,.$$

Note that $\max_{y>0}|m(y)| = -\,m(\sqrt{3\,-\,\sqrt{6}}) \le 1.3802$. Therefore, we get (27) for $i = 5$.

Next we consider J_{34}. Put $x_1 = 3.3834$ and $x_2 = 5.9877$. By (26) we have

$$-y_2(x_1) \ge 1.2736\,, \quad y_1(x_2) \le 3.3307\,, \quad m(3.3307) \le -m(1.2736) \le 0.7799$$

and therefore we get (27) for $i = 4$.

Finally, we construct bound for J_{33}. Put $x_3 = 6.2203$. By (26) we have

$$-y_2(x_2) \ge 1.699\,, \quad y_1(x_3) \le 3.3947$$

and

$$-m(1.699) \le m(3.3947) \le 0.0911\,,$$

which leads to (27) for $i = 3$, since $|J_{33}| \le 0.101/n$.

Thus, for all x satisfying (23) we proved (25).

6 Proof of Theorem 1

Note that J_2 and J_3 are estimated in (24) and (25) provided that x satisfies (15), whereas the bound for J_1 is uniform.

Note now the following fact: let $F(x)$ and $G(x)$ be distribution functions and suppose that for some $x_0 > 0$ we have

$$\sup_{|x|\le x_0} |F(x) - G(x)| \le \delta \tag{28}$$

and

$$\max\{G(-x_0)\,,\ 1-G(x_0)\} \le \varepsilon\,. \tag{29}$$

Then

$$\sup_{x\in\mathbb{R}^1} |F(x)-G(x)| \le \delta+\varepsilon\,. \tag{30}$$

In fact, (30) follows immediately from (28) and (29) because (29) implies that

$$\max\{F(-x_0)\,,\ 1-F(x_0)\} \le \delta+\varepsilon$$

and

$$\sup_{|x|\ge x_0} |F(x)-G(x)| \le \max\{G(-x_0)\,,\ F(-x_0)\,,\ 1-F(x_0)\,,\ 1-G(x_0)\}\,.$$

Since $G_1(x)=P(|Y|^2\le x)$, and

$$\begin{aligned} P(|Y|^{\,2}>x) &= \frac{2}{\sqrt{2\pi}}\int_{\sqrt{x}}^{\infty}\exp\{-y^2/2\}\,dy \le \sqrt{\frac{2}{\pi x}}\int_{\sqrt{x}}^{\infty} y\ \exp\{-y^2/2\}\,dy \\ &= \sqrt{\frac{2}{\pi x}}\int_{x/2}^{\infty}\exp(-z)\,dz = \sqrt{\frac{2}{\pi x}}\ \exp\{-x/2\} \end{aligned}$$

we get for $x\ge n\,/\,2$

$$P(|Y|^2>x) \le \sqrt{\frac{4}{\pi\,n}}\ \exp\Big(-\frac{n}{4}\Big)\,.$$

Therefore, by (4), (6), (21), (24), (25), and (30) we obtain the desired bound (1).

7 Proof of Theorem 2

We show how (2) can be obtained from (1). For $i=1,\ldots,p$, put

$$n_i = n-i+1 \quad\text{and}\quad X_i = \mathcal{X}^2_{n_i} - n_i\log\frac{\mathcal{X}^2_{n_i}}{n_i} - n_i.$$

Here all the $\mathcal{X}^2_{n_i}$-variates are independent. Using a well known Bartlett decomposition theorem, see M. Siotani, T. Hayakawa and Y. Fujikoshi (1985) we can write T_p as

$$T_p = \sum_{i=1}^{p}(X_i+Z_i), \tag{31}$$

where $Z_i \sim \mathcal{X}^2_{i-1}$, $\mathcal{X}^2_0 = 0$ and all Z_i's and X_i's are independent.
Now we show that if D_i is such that

$$\sup_x |\, P(X_i \leq x) - P(U_i \leq x)\, | \leq D_i \tag{32}$$

with U_i distributed as $\mathcal{X}^2_1$ then

$$\sup_x |P(T_p \leq x) - G_q(x)| \leq D_1 + \cdots + D_p \tag{33}$$

with $q = p(p+1)/2$.

In fact, (33) follows from (31), Lemma 3 (see below) and the fact that the sum of two independent random variables distributed as $\mathcal{X}^2_m$ and $\mathcal{X}^2_n$ resp. has chi-square distribution with $m + n$ degrees of freedom.

Lemma 3 *Let X_1, X_2, U_1, U_2 and Z be independent random variables. Let D_1 and D_2 be such that (32) holds for $i = 1, 2$. Then*

$$\sup_x \big|P(X_1 + X_2 + Z \leq x) - P(U_1 + U_2 + Z \leq x)\big| \leq D_1 + D_2. \tag{34}$$

Remark 1. We do not make in Lemma 3 any assumptions about a form of distributions of X_1, X_2, U_1, U_2 and Z. Its independence is important only.
Proof(cp. the beginning of the proof of Theorem 3.1 in V. V. Ulyanov, H. Wakaki, Y. Fujihoshi (2005)). Write

$$\begin{aligned} & \sup_x \big|P(X_1 + X_2 + Z \leq x) - P(U_1 + U_2 + Z \leq x)\big| \\ \leq \; & \sup_x \big|P(X_1 + X_2 + Z \leq x) - P(U_1 + X_2 + Z \leq x)\big| \\ & + \sup_x \big|P(U_1 + X_2 + Z \leq x) - P(U_1 + U_2 + Z \leq x)\big|. \end{aligned} \tag{35}$$

Since for any independent random variables X, U and Z we have

$$\begin{aligned} & \sup_x \big|P(X + Z \leq x) - P(U + Z \leq x)\big| \\ \leq \; & \sup_x E\big|P(X \leq x - Z|\ Z) - P(U \leq x - Z|\ Z)\big| \\ \leq \; & \sup_x \big|P(X \leq x) - P(U \leq x)\big|, \end{aligned}$$

we get (34) from (35) and Lemma's assumptions.

8 Generalization of Theorem 1

In Theorem 1 we constructed error bound for distribution of T_1 which allowed (see (3)) representation $T_1 = h(V_n)$. Therefore, the possible generalizations can be made in the directions when we replace either function h or random variable V_n or h and V_n simultaneously by similar objects. Here we give generalization connected with replacing of V_n. Concerning V_n we used in the proof of Theorem 1 only the fact that the distribution of V_n can be approximated by the first-order Chebyshev-Edgeworth expansion and error bound of the approximation is known (see (4)). Therefore, the following generalization holds.

Theorem 3 *Let a random variable W_n allow an approximation*

$$\sup_x |P(W_n \leq x) - \Phi_{1n}(x)| \leq B_1(n), \tag{36}$$

where

$$\Phi_{1n}(x) = \Phi(x) + p(EW_n^3, x)\,\varphi(x)/\sqrt{n},$$

$p(EW_n^3, x)$ is a polynomial depending on the third moment of W_n and moreover $p(EW_n^3, x)$ is an even function and $B_1(n) = O(n^{-1})$ as $n \to +\infty$. Let $h(y) = \sqrt{2n}\,y \,-\, n\log(1+\sqrt{\frac{2}{n}}\,y)$ and $T = h(W_n)$. Then

$$\sup_x |P(T \leq x) - G_1(x)| \leq 2B_1(n) + \frac{c}{n} + \frac{1.1284}{\sqrt{n}}\ 0.7788^n\,,$$

where c is a bounded and computable constant depending on the coefficients of the polynomial p.

Remark. Different examples of W_n when W_n is a normalized sum of independent identically distributed random variables and W_n satisfies (36) could be found e.g. in Dobrić and Ghosh (1996). We have noted that Theorem 1 can be applied to LR statistic for testing a hypothesis that the variance σ^2 is equal to a given value in a normal population $N(\mu, \sigma^2)$ in both cases when μ is known and μ is unknown parameter. By using Theorem 3 it is possible to obtain an error bound for the same statistic in a nonnormal population with known μ.

Acknowledgment. The research was supported in part by Russian Foundation of Basic Research, project no. 05-01-02941, and by Grant-in-Aid for Scientific Research (B) in 2005, no. 1530092.

References

R.M. Corless, G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey and D.E. Knuth (1996), On the Lambert W function, *Advances in Computational Mathematics*, **5**, 329-359.

V. Dobrić and B.K. Ghosh (1996), Some analogs of the Berry-Esséen bound for first-order Chebyshev-Edgeworth expansions, *Statistics & Decisions*, **14**, 383 - 404.

M. Siotani, T. Hayakawa, and Y. Fujikoshi, *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook.* The American Sciences Press Series in Mathematical and Management Sciences, Vol. 9. Columbus, Ohio: American Sciences Press, Inc. XIV, 1985.

V. V. Ulyanov, H. Wakaki and Y. Fujihoshi (2005), Berry-Esseen bound for high dimensional asymptotic approximation of Wilks' Lambda distribution, accepted to *Statistics and Probability Letters*.