THEORY PROBAB. APPL.
Vol. 51, No. 4, pp. 683–695

ASYMPTOTIC EXPANSION OF THE COVERAGE PROBABILITY OF JAMES–STEIN ESTIMATORS*

S. E. AHMED†, A. K. MD. E. SALEH‡, A. I. VOLODIN§, AND I. N. VOLODIN¶

(*Translated by A. I. Volodin*)

Abstract. This paper provides a new approach to the asymptotic expansion construction of the coverage probability of the confidence sets recentered in [W. James and C. Stein, *Estimation with quadratic loss*, in Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, Univ. California Press, Berkeley, CA, 1961, pp. 361–379] and its positive-part Stein estimators [C. Stein, *J. Roy. Statist. Soc. Ser. B*, 24 (1962), pp. 265–296]. The coverage probability of these confidence sets depends on the noncentrality parameter τ^2 as in the case of risks of these estimators. The new approach (which is different than Berger's [J. O. Berger, *Ann. Statist.*, 8 (1980), pp. 716–761] and Hwang and Casella's [J. T. Hwang and G. Casella, *Statist. Decisions*, suppl. 1 (1984), pp. 3–16]) allows us to obtain the asymptotics analysis of the coverage probabilities for the two cases, namely, when $\tau^2 \to 0$ and $\tau^2 \to \infty$. For both cases we provide a simple approximation of the coverage probabilities. Some graphical and tabular results are provided to assess the accuracy of our approximations.

Key words. confidence sets, James–Stein estimators, Stein estimation, multivariate normal distribution, coverage probability, asymptotic expansion

DOI. 10.1137/S0040585X97982712

1. Introduction. In this paper we address the confidence estimation problem of the mean vector $\theta = (\theta_1, \dots, \theta_p)$ of the p-variate normal distribution with independent components and equal unit variances. Let $\overline{X} = (\overline{X}_1, \dots, \overline{X}_p)$ be the sample mean vector that is calculated from samples of equal size n of marginal distributions. The confidence set

$$D_{\overline{X}} = \left\{\theta\colon\ n\sum_1^p (\theta_i - \overline{X}_i)^2 \le c^2\right\}$$

has the given confidence coefficient $1-\alpha$ if c^2 is defined as the quantile of the central chi-square distribution with p degrees of freedom according to the formula $K_p(c^2) = 1-\alpha$, where $K_p(\,\cdot\,)$ is the chi-square distribution function.

*Received by the editors November 17, 2004.
http://www.siam.org/journals/tvp/51-4/98271.html
†Department of Mathematics and Statistics, University of Windsor, Windsor, Ontario N9B 3P4, Canada (seahmed@uwindsor.ca).
‡School of Mathematics and Statistics, Carleton University, Ottawa, Ontario K1S 5B6, Canada (esaleh@math.carleton.ca).
§Department of Mathematics and Statistics, University of Regina, Regina, Saskatchewan S4S 0A2, Canada (andrei@math.uregina.ca).
¶Department of Mathematical Statistics, Kazan State University, Kazan, Kremlevskaya St., 18, 420008, Russia (Igor.Volodin@ksu.ru).

This confidence set $D_{\overline{X}}$ enjoys many optimal properties, namely, it is (i) best unbiased, (ii) best invariant, (iii) minimax, but it is *not* unique; Stein [11] showed that there exists other sets which may provide improved coverage probability retaining the minimax property. As a result, Stein [11] suggested considering the recentered confidence sets D_δ and $D_{\delta+}$ with fixed volume as $D_{\overline{X}}$, defined by

$$D_\delta = \Bigl\{\theta\colon\ n\bigl\|\theta - \delta(\overline{X})\bigr\|^2 \le c^2\Bigr\}, \qquad \text{where} \quad \delta(\overline{X}) = \biggl(1 - \frac{p-2}{n\|\overline{X}\|^2}\biggr)\overline{X},$$

is the James–Stein estimator (cf. [5]) and

$$D_{\delta+} = \Bigl\{\theta\colon\ n\bigl\|\theta - \delta^+(\overline{X})\bigr\|^2 \le c^2\Bigr\},$$

where

$$\delta^+(\overline{X}) = \biggl(1 - \frac{p-2}{n\|\overline{X}\|^2}\biggr)\overline{X} I\bigl(n\|\overline{X}\|^2 > p-2\bigr).$$

In both cases we restrict p to be greater than 2. Subsequently, many studies followed by various authors, notably, Brown [2], Berger [1], Hwang and Casella [7], [8], Robert and Saleh [9], and Saleh [10], among others. A real breakthrough came when Hwang and Casella [7], [8] showed that $D_{\delta+}$ dominates $D_{\overline{X}}$ uniformly in $\tau^2 = n\|\theta\|^2$ for $p \ge 4$ and for approximately $0.8(p-2)$ instead of $p-2$. Also, Hwang and Casella [8] provided an approximation of the coverage probability when $\tau^2 \to \infty$. Our approach to the approximation of the coverage probability is different, combined with a geometrical and analytical methodology, and shows the following order of the coverage probabilities

$$\mathbf{P}(D_{\overline{X}}) \le \mathbf{P}(D_\delta) \le \mathbf{P}(D_{\delta+})$$

for all τ^2 as in the case of risk analysis for point estimation.

We investigate an asymptotic expansion of the coverage probability of the confidence sets recentered at the James–Stein and positive-part estimators. It is established that the probabilities $Q_p(\tau) = \mathbf{P}(D_\delta)$ and $Q_p^+(\tau) = \mathbf{P}(D_{\delta+})$ depend on $\tau^2 = n\|\theta\|^2$ and are decreasing functions of τ^2.

If $\tau \to 0$, then

$$Q_p(\tau) = K_p\bigl(w_2(c,\tau)\bigr) - K_p\bigl(w_1(c,\tau)\bigr) + O(\tau^2) = K_p\bigl(w_2(c,0)\bigr) - K_p\bigl(w_1(c,0)\bigr) + O(\tau^2), \tag{1.1}$$

where

$$w_{1,2}(c,\tau) = p - 2 + \frac{c^2-\tau^2}{2} \mp \sqrt{\frac{(c^2-\tau^2)^2}{4} + c^2\tau^2 - (p-2)(\tau^2 - c^2)}. \tag{1.2}$$

If we center at the positive-part James–Stein estimator, the coverage probability can be written

$$Q_p^+(\tau) = K_p\bigl(w_2(c,\tau)\bigr) + O(\tau^2) = K_p\bigl(w_2(c,0)\bigr) + O(\tau^2). \tag{1.3}$$

If $\tau \to \infty$ (cf. with asymptotics of Berger [1] and Hwang and Casella [8]), then

$$Q_p(\tau) \sim Q_p^+(\tau) = K_p\bigl(w_2(c,\tau)\bigr) + O(\tau^{-2}) = K_p(c^2) + O(\tau^{-2}). \tag{1.4}$$

We mention that the leading terms of the asymptotic expansion with the function $w_{1,2}(c,\,\tau)$ in place of $w_{1,2}(c,\,0)$ strongly improve the accuracy of the approximation. It is worthwhile to mention that for the confidence sets recentered at the positive-part James–Stein estimator, we obtain the same asymptotic formula for both cases (i.e., $\tau^2 \to 0$ and $\tau^2 \to \infty$). The graphical and numerical illustrations, given in the final section of the paper, show that this formula gives exceptionally accurate values for the coverage probability in the range $\tau \le C$, which are useful for practical application, for any $\tau \ge 0$.

It is well known that the range of James–Stein estimators applications is notably peculiar and requires great care in an interpretation of traditional estimator shrinkage. First of all, this relates to the existence of the real center of shrinkage μ (without loss of generality, in our paper it is assumed that $\mu = 0$). There are some other challenges in these estimator applications, which are discussed in detail in the book of Lehmann and Casella [12]. At the same time there are many publications in which James–Stein estimators are applied to real data and numerical results on quadratic risk gains are presented. In these situations, when the application of shrinkage estimates is justified, the main results of our paper show the region of the parameter values of the normal distributions under consideration. In this region the James–Stein estimators have a substantial advantage over the sample means from the point of view of improved accuracy and/or reliability. This region is defined by the inequality $\tau^2/\sigma^2 \le K_p^{-1}(1-\alpha)$, and, as Table 1 and Figure 5 show, the coverage probability rapidly approaches the nominal confidence coefficient $1-\alpha$ when the parameters of the distribution move away from the region.

2. Integral representation of the coverage probability. Our asymptotic investigation of the coverage probability of the true value of θ by a confidence set centered by James–Stein estimators is based on the following representation of the supporting functions $\sum_1^p(\theta_i - \delta_i)^2$ and $\sum_1^p(\theta_i - \delta_i^+)^2$ of the corresponding confidence sets D_δ and $D_{\delta+}$.

Introduce the standard normal random variables $Z_i = (\overline{X}_i - \theta_i)\sqrt{n}$, $i = 1, \dots, p$, and put

$$a = p-2, \quad \tau = \left[n\sum_1^p \theta_i^2\right]^{1/2}, \quad X = \sum_1^p Z_i^2, \quad Y = \left[\sum_1^p \theta_i^2\right]^{-1/2} \sum_1^p \theta_i Z_i.$$

Let A and B be the events defined by

$$A = \left\{n\sum_1^p \overline{X}_i^2 > a\right\} \quad \text{and} \quad B = \left\{n\sum_1^p \big(\delta_i(\overline{X}) - \theta_i\big)^2 < c^2\right\},$$

where $c^2 = K_p^{-1}(1-\alpha)$ in terms of the introduced variables. Then

$$A = \{X + 2Y\tau + \tau^2 > a\},$$
$$B = \Big\{X\big(X + 2Y\tau + \tau^2\big) - 2aY\tau - 2aX < c^2\big(X + 2Y\tau + \tau^2\big)\Big\},$$

since $X + 2Y\tau + \tau^2 > 0$. Using these events we can represent the coverage probability by the confidence set D_δ and $D_{\delta+}$ as

$$Q_p(\tau) = \mathbf{P}(B) \quad \text{and} \quad Q_p^+(\tau) = \mathbf{P}(A^c)\, I(\tau < C) + \mathbf{P}(A \cap B),$$

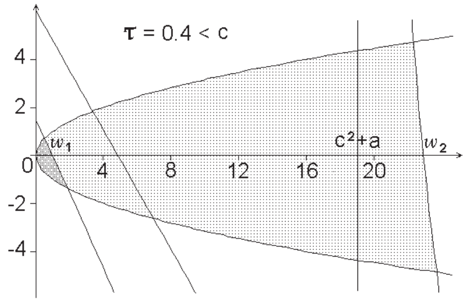

Fig. 1.

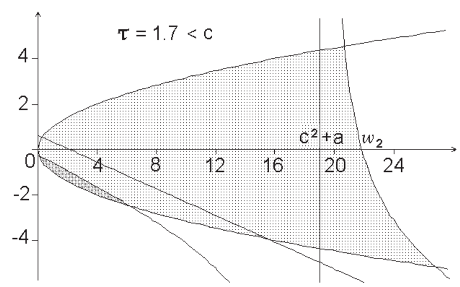

Fig. 2.

where $I(\,\cdot\,)$ is the indicator of the corresponding set. Robert and Saleh [9] obtained the exact expression for $Q_p^+(\tau)$.

The respective orthogonal transformation of the vector Z shows that we can provide the following stochastic representation of the random variables: $Y = \xi$, $X = \chi^2 + \xi^2$, where the normal standard random variable ξ and the chi-square random variable χ^2 with $p-1$ degrees of freedom are independent. Such a representation can be useful for the simulation of confidence sets by the Monte-Carlo method.

The obtained stochastic representation immediately implies the following result.

Proposition 1. *The joint density function of the random variables X and Y is different from zero only in the region $y^2 < x$ and it is represented in this region as*

$$f(x,y) = \frac{1}{\sqrt{2\pi}}\,\frac{1}{2^{(p-1)/2}\Gamma((p-1)/2)}\,(x-y^2)^{(p-1)/2-1}\mathrm{e}^{-x/2}.$$

Now we investigate in detail the geometry of the regions in a plane that correspond to the events A and B. We start with an observation that the confidence region $D_{\delta+}$ is equal to the union of the following disjoint events:

$$\begin{aligned}
&\{Y^2 < X,\, X + 2Y\tau + \tau^2 < a\}\\
&+\left\{Y^2 < X,\ c^2 + a - X > 0,\ X + 2Y\tau + \tau^2 > a,\ Y > \frac{(X-a)^2 - c^2\tau^2 + X(\tau^2 - c^2)}{2\tau(c^2 + a - X)}\right\}\\
&+\left\{Y^2 < X,\ c^2 + a - X < 0,\ X + 2Y\tau + \tau^2 > a,\ Y < \frac{(X-a)^2 - c^2\tau^2 + X(\tau^2 - c^2)}{2\tau(c^2 + a - X)}\right\}.
\end{aligned}$$

If we represent these events as regions in the plane (x, y), then the boundaries of these regions are defined by the following curves in the semiplane $x > 0$:

$$x = y^2, \qquad y = g(x) = \frac{a - x - \tau^2}{2\tau},$$
$$y = h(x) = \frac{(x-a)^2 - c^2\tau^2 + x(\tau^2 - c^2)}{2\tau(c^2 + a - x)} = \frac{1}{2\tau}\left(a - x - \tau^2 + \frac{a(\tau^2 - c^2)}{c^2 + a - x}\right).$$

The view of linear-hyperbolic function $h(x)$ points out the significant difference in the form of confidence regions for the different values of τ, c, and a.

We mention that for small (what we are practically using) values of α, the quantile $c^2 = K_p^{-1}(1-\alpha)$ of the chi-square distribution is bigger than the mean value p of this distribution, and hence in what follows we assume that $a - c^2 < 0$.

If $\tau < c$, then the derivative of the function $y = h(x)$, $x > 0$, equals

$$h'(x) = \frac{1}{2\tau}\left(-1 + \frac{a(\tau^2 - c^2)}{(c^2 + a - x)^2}\right).$$

It is always negative, and hence $h(x)$ it is a *decreasing function* in the region $x > 0$ (cf. Figures 1 and 2). The curve $y = h(x)$ has two monotonically decreasing branches that are separated by the vertical asymptote $x = a + c^2$ and slant asymptote $y = g(x)$ (the line that defines the event A!). One of the branches that is situated to the left of the vertical asymptote and below the slant asymptote $y = g(x)$ has the first OX-intercept at the point $y = (a - c^2\tau^2)/2\tau(c^2 + a) > 0$ and the second OX intercept at the point w_1 (cf. (1.2)) when $\tau < a/c$. In the opposite case, that is, when $\tau \geq a/c$, it does not have the second OX-intercept and always stays below this axis for $x > 0$. Further, this branch tends to $-\infty$, clasping to asymptote $x = a + c^2$ from the left. The other branch of $h(x)$ that is situated to the right of the vertical asymptote and above the slant asymptote $x = g(y)$ begins from $+\infty$ and has OX-intercept at the point w_2. Next, it tends to$-\infty$ for $x \to +\infty$.

The situation described shows us that for $\tau < c$ the confidence coefficient for the region, centered by the positive-part James–Stein estimation δ^+, is defined by the probability of the event $D_2 = \{X > Y^2,\ Y > c^2 + a - X,\ Y < h(X)\}$ (the shaded area in the figures). The gain from using estimation δ^+ instead of δ is defined by the probability of the event $D_1 = \{X > Y^2,\ Y < c^2 + a - X,\ Y > h(X)\}$ (the darker shaded area).

Hence, for $\tau < c$

$$Q_p^+(\tau) = \iint_{D_2} f(x,y)\,dx\,dy, \qquad Q_p(\tau) = \iint_{D_2} f(x,y)\,dx\,dy - \iint_{D_1} f(x,y)\,dx\,dy,$$

and the only thing which is left to do is determine the limits of integration in the double integrals.

There are four points of intersection of the parabola $x = y^2$ and the curve $y = h(x)$. Hence the two equations

$$(x-a)^2 - c^2\tau^2 + x(\tau^2 - c^2) = 2\tau\sqrt{x}(c^2 + a - x), \tag{2.1a}$$
$$(x-a)^2 - c^2\tau^2 + x(\tau^2 - c^2) = -2\tau\sqrt{x}(c^2 + a - x) \tag{2.1b}$$

have two roots each, that is, four roots altogether. Two of them, say u_1 and u_2, for $\tau < a/c$ are situated on the opposite side of the point w_1 and collapse into this point for $\tau \to 0$. Two other roots, say v_1 and v_2, for any $\tau < c$ satisfy the inequalities $v_1 < w_2 < v_2$ and collapse into w_2 for $\tau \to 0$. Hence, for $\tau < a/c$ we have the relations

$$u_1 < w_1 < u_2 < v_1 < w_2 < v_2,$$

where the roots u_1 and v_1 correspond to (2.1a), and u_2 and v_2 correspond to (2.1b). Note that it is sufficient to solve (2.1a) to obtain the roots u_1 and v_1. The roots u_2 and v_2 of (2.1b) can be obtained by the simple substitution of τ by $-\tau$ in the roots obtained.

For all $\tau < c$ represent the event D_2 in the form

$$D_2 = D_{w_2} - D_{v_1} + D_{v_2},$$

where

$$D_{w_2} = \{X > Y^2,\ X < w_2\}, \qquad D_{v_1} = \bigl\{v_1 < Y < w_2,\ h(X) < Y < \sqrt{X}\,\bigr\},$$

and

$$D_{v_2} = \bigl\{w_2 < Y < v_2,\ -\sqrt{X} < Y < h(X)\bigr\}$$

(cf. Figure 1). If $\tau < a/c$, then the analogous representation can be provided for D_1 with the substitution of w_2 by w_1 and v by u. In the region $a/c < \tau < c$ the value $w_1 < 0$, and the equation $h(x) = -\sqrt{x}$ has two roots to the left of the vertical asymptote, say u_1' and u_2' (cf. Figure 2). In this case,

$$D_1 = \bigl\{u_1' < X < u_2',\ -\sqrt{X} < Y < h(X)\bigr\}.$$

Hence for all $\tau < c$, the coverage probability

$$Q_p^+(\tau) = \mathbf{P}(D_2) = \mathbf{P}(D_{w_2}) + \mathbf{P}(D_{v_2}) - \mathbf{P}(D_{v_1})$$

and the probability

$$Q_p(\tau) = \mathbf{P}(D_2) - \mathbf{P}(D_1),$$

where

$$\mathbf{P}(D_1) = \mathbf{P}(D_{w_1}) + \mathbf{P}(D_{u_2}) - \mathbf{P}(D_{u_1})$$

if $\tau < a/c$, and

$$\mathbf{P}(D_1) = \mathbf{P}\bigl\{u_1' < X < u_2',\ -\sqrt{X} < Y < h(X)\bigr\}$$

if $a/c < \tau < c$. Such a representation of the coverage probability is important because the first term in the sum

$$\mathbf{P}(D_{w_2}) = \mathbf{P}\{X < w_2\} = K_p(w_2) = K_p\left(a + \frac{c^2}{2} + \sqrt{\frac{c^4}{4} + ac^2}\right) + O(\tau^2)$$

gives the leading term of $Q_p^+(\tau)$ asymptotics for $\tau \to 0$ (cf. (1.3)). The difference $\mathbf{P}(D_{v_2}) - \mathbf{P}(D_{v_1})$ provides the residual term that, as this will be proved in what follows, has an order $O(\tau^2)$. It is sufficient to investigate the asymptotics of $\mathbf{P}(D_{w_2})$, since the same result is true for the region D_{w_1} with substitution v by u and w_2 by w_1.

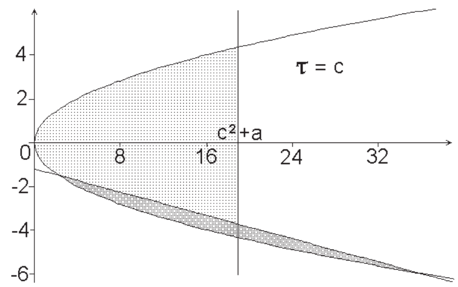

FIG. 3.

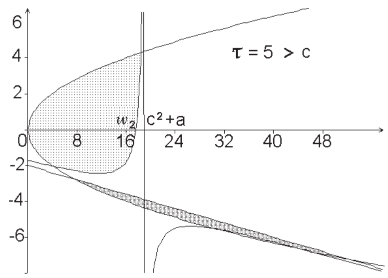

FIG. 4.

Hence we can formulate the result and pass on to the study of the confidence regions geometry for $\tau \geq c$.

PROPOSITION 2. *If $\tau < c$, then*

$$Q_p^+(\tau) = K_p(w_2) + \int_{w_2}^{v_2} dx \int_{-\sqrt{x}}^{h(x)} f(x,y)\,dy - \int_{v_1}^{w_2} dx \int_{h(x)}^{\sqrt{x}} f(x,y)\,dy.$$

If $\tau < a/c$, then

$$Q_p(\tau) = Q_p^+(\tau) - \left[K_p(w_1) + \int_{w_1}^{u_2} dx \int_{-\sqrt{x}}^{h(x)} f(x,y)\,dy - \int_{u_1}^{w_1} dx \int_{h(x)}^{\sqrt{x}} f(x,y)\,dy\right],$$

and if $a/c \leq \tau < c$, then

$$Q_p(\tau) = Q_p^+(\tau) - \int_{u_1'}^{u_2'} dx \int_{-\sqrt{x}}^{h(x)} f(x,y)\,dy.$$

For $\tau = c$ the branches of the curve $y = h(x)$ merge with its asymptotes and the confidence regions are bounded by the straight lines $x = a + c^2$, $y = (a - x - c^2)/2c$ and the parabola $x = y^2$ (cf. Figure 3). The straight line $y = (a - x - c^2)/2c$ (slant asymptote) has the OX-intercept at the points $x_{1,2} = (c \mp \sqrt{a})^2$ that are situated on the opposite sides of the point $x = c^2 + a$. Hence, we have the following proposition.

PROPOSITION 3. *If $\tau = c$, then the coverage probabilities are*

$$Q_p^+(\tau) = K_p(c^2) + \int_{c^2+a}^{x_2} dx \int_{-\sqrt{x}}^{(a-x-c^2)/(2c)} f(x,y)\,dy,$$

$$Q_p(\tau) = Q_p^+(\tau) - \int_{x_1}^{x_2} dx \int_{-\sqrt{x}}^{(a-x-c^2)/(2c)} f(x,y)\,dy.$$

For the case $\tau > c$, a picture of the region that defines the coverage probabilities is significantly simpler (cf. Figure 4).

The slant asymptote $y = (a - \tau^2 - x)/(2\tau)$ in the region $x > 0$ has negative values and intersects the lower branch of the parabola $y = -\sqrt{x}$ at the points $x_{1,2} = (\tau \mp \sqrt{a})^2$. The region below this asymptote and above the lower branch of the parabola coincides with the region A^c of the zero values of the estimate δ^+. Hence we can simply ignore the right branch of the curve $y = h(x)$, $x > c^2 + a$, situated below the slant asymptote. Next, the left branch of this curve that is situated above the slant asymptote in the region $0 < x < c^2 + a$ begins from the point $y = (a^2 - \tau^2 c^2)/(2\tau(c^2 + a)) < 0$ (recall that $a < c^2$), intersects the lower branch of the parabola at the point $x = v_2$, and has OX-intercept at the point $x = w_2$. It intersects the upper branch of the parabola at the point $x = v_1$. At the same time $v_2 < v_1$. Note that $w_2 \to c^2$ and $v_1 - v_2 \to 0$ when $\tau \to \infty$. Hence we can formulate the following proposition.

PROPOSITION 4. *If $\tau > c$, then*

$$Q_p^+(\tau) = K_p(w_2) + \int_{w_2}^{v_1} dx \int_{h(x)}^{\sqrt{x}} f(x,y)\,dy - \int_{v_2}^{w_2} dx \int_{-\sqrt{x}}^{h(x)} f(x,y)\,dy + \int_{x_1}^{x_2} dx \int_{-\sqrt{x}}^{(a-\tau^2-x)/(2\tau)} f(x,y)\,dy$$

and

$$Q_p(\tau) = K_p(w_2) + \int_{w_2}^{v_1} dx \int_{h(x)}^{\sqrt{x}} f(x,y)\,dy - \int_{v_2}^{w_2} dx \int_{-\sqrt{x}}^{h(x)} f(x,y)\,dy.$$

3. Asymptotic expansions of coverage probability. Integral representations of the coverage probabilities given in Propositions 2 and 4 (compare them with the representations obtained in [7]) reduce the problem of asymptotic analysis to an asymptotic expansion for $\tau \to 0$ and, respectively, as $\tau \to \infty$ of the following integrals:

$$J(\tau) = \int_{w}^{\beta} dx \int_{\varphi(x)}^{\psi(x)} f(x,y)\,dy,$$

where the lower limit w has the form (1.2), the upper limit β is one of the roots of (2.1 a,b), and the functions $\varphi(x)$ and $\psi(x)$ are equal to either $h(x)$ or $\pm\sqrt{x}$.

We start with the asymptotics of the roots of (2.1 a,b). As was mentioned before, it is sufficient to consider only (2.1a). By making the substitution $x = w + z$ (recall, $h(w) = 0$) we can represent this equation in the form

$$2z(w-a) + z^2 + z(\tau^2 - c^2) = 2\tau\sqrt{w+z}\,(c^2 + a - w - z). \tag{3.1}$$

LEMMA. *If $\tau \to 0$, then the following asymptotic expansion takes place for the root of* (3.1):

$$z = \sum_{k=1}^{\infty} \lambda_k(w)\,\tau^k. \tag{3.2}$$

If $\eta = 1/\tau \to 0$, then

$$z = \sum_{k=1}^{\infty} \mu_k(w)\,\eta^k. \tag{3.3}$$

Proof. We can rewrite (3.1) for the case $\tau \to 0$ in the form $z = \tau/f(z)$, where

$$f(z) = \frac{2(c^2 + a - w - z)\sqrt{w+z}}{2(w-a) + z + \tau^2 - c^2}$$

is a holomorphic function in a neighborhood of the point $z = 0$ and $f(0) \neq 0$. Hence, the Burman–Lagrange formula is applicable (cf., for example, [3, section 2.2]) and the expansion (3.2) takes place with

$$\lambda_k(w) = \frac{1}{k!}\left[\frac{d^{k-1} f^k(z)}{dz^{k-1}}\right]_{z=0}, \qquad k = 1, 2, \dots .$$

For the case $\tau \to \infty$ we represent (2.1a) in terms of $\varepsilon = 1/\tau$. Hence it can be written as

$$\eta(x-a)^2 - c^2 + x = \eta^2 x c^2 = 2\eta(c^2 + a - x)\sqrt{x}.$$

Next, make the substitution $x = w + z$ and rewrite the equation in the form $z = \eta/f(z)$ with

$$f(z) = \frac{2(c^2 + a - w - z)\sqrt{w+z}}{1 + \eta^2(2(w-a) + z - c^2)}.$$

Applying the Burman–Lagrange formula again yields expansion (3.3).

The fact that for both cases the function $f(z)$ depends on a small parameter (τ^2 or, respectively, η^2), which will move on the expansion coefficients, does not play any significant role. It is possible to provide an additional expansion of the coefficient by this parameter, and the form of the expansion will remain the same.

Hence, the following asymptotic expansions are true for the roots u_1, u_2, v_1, v_2 from Proposition 2 for $\tau \to 0$:

$$u_{1,2} = w_1 + \sum_{k=1}^{\infty} \lambda_k(w_1)(\pm\tau)^k, \qquad v_{1,2} = w_2 + \sum_{k=1}^{\infty} \lambda_k(w_2)(\pm\tau)^k.$$

For the roots v_1, v_2 from Proposition 4 ($\eta = 1/\tau \to 0$) the following expansions take place:

$$v_{1,2} = w_2 + \sum_{k=1}^{\infty} \lambda_k(w_2)(\pm\eta)^k.$$

A significant advantage of the expansions obtained above is the fact that the expansions of v_1 and v_2 (and u_1 and u_2 also) have the same coefficients for even exponents of τ, and for the odd exponents of τ, they are different only by a sign. This fact provides us with the order of the remaining terms for the coverage probability asymptotics for $\tau \to 0$ (cf. formulas (1.1) and (1.3)) as was stated in the introduction. Certainly, the expansions of the roots v_1 and v_2 by the exponents of η, which provide the form of formula (1.4) from the introduction, have the analogous properties. Now we will prove this formula.

THEOREM. *If $\tau \to 0$, then the following asymptotic expansion is true for the coverage probability of the true value by the confidence region D_δ centered by the*

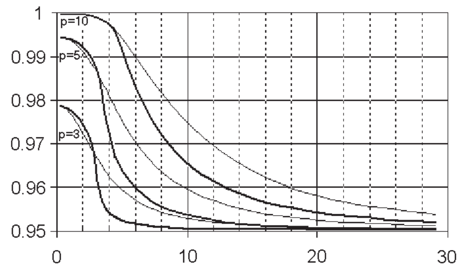

FIG. 5.

James–Stein estimation:

$$
\begin{aligned}
Q_p(\tau) &= K_p\left(p-2+\frac{c^2-\tau^2}{2}+\sqrt{\frac{(c^2-\tau^2)^2}{4}+c^2\tau^2-(p-2)(\tau^2-c^2)}\right) \\
&\quad - K_p\left(p-2+\frac{c^2-\tau^2}{2}-\sqrt{\frac{(c^2-\tau^2)^2}{4}+c^2\tau^2-(p-2)(\tau^2-c^2)}\right)+O(\tau^2) \\
&= K_p\left(p-2+\frac{c^2}{2}+\sqrt{\frac{c^4}{4}+(p-2)\,c^2}\right) \\
&\quad - K_p\left(p-2+\frac{c^2}{2}-\sqrt{\frac{c^4}{4}+(p-2)\,c^2}\right)+O(\tau^2).
\end{aligned}
$$

The following asymptotic expansion is true for the coverage probability of the true value by the confidence region $D_{\delta+}$ centered by the positive-part James–Stein estimation:

$$
\begin{aligned}
Q_p^+(\tau) &= K_p\left(p-2+\frac{c^2-\tau^2}{2}+\sqrt{\frac{(c^2-\tau^2)^2}{4}+c^2\tau^2-(p-2)(\tau^2-c^2)}\right)+O(\tau^2) \\
&= K_p\left(p-2+\frac{c^2}{2}+\sqrt{\frac{c^4}{4}+(p-2)\,c^2}\right)+O(\tau^2).
\end{aligned}
$$

If $\tau \to \infty$, then the coverage probabilities for both regions are asymptotically equivalent:

$$
\begin{aligned}
&Q_p(\tau) \sim Q_p^+(\tau) \\
&\quad = K_p\left(p-2+\frac{c^2-\tau^2}{2}+\sqrt{\frac{(c^2-\tau^2)^2}{4}+c^2\tau^2-(p-2)(\tau^2-c^2)}\right)+O(\tau^{-2}) \\
&\quad = K_p(c^2)+O(\tau^{-2}).
\end{aligned}
$$

Proof. We prove that for $\tau \to 0$,

$$
\Delta(\tau)=\int_{w_2}^{v_2} dx \int_{-\sqrt{x}}^{h(x)} f(x,y)\,dy-\int_{v_1}^{w_2} dx \int_{h(x)}^{\sqrt{x}} f(x,y)\,dy=O(\tau^2).
$$

Then the analogous asymptotics takes place for integrals that have limits in terms of u_1 and u_2.

Perform the substitutions $x - w_2 = t(v_2 - w_2)$ in the first integral and $w_2 - x = t(w_2 - v_1)$ in the second integral, and make use of the asymptotic expansion

$$v_2 - w_2 \sim w_2 - v_1 = -\lambda_1 \tau + O(\tau^2),$$

which implies that

$$h\bigl(w_2 + t(v_2 - w_2)\bigr) \sim h\bigl(w_2 - t(w_2 - v_1)\bigr) = O(\tau^2)$$

since $h(w_{1,2}) = 0$, and asymptotic equalities

$$\bigl[w_2 + t(v_2 - w_2)\bigr]^{1/2} \sim \bigl[w_2 - t(w_2 - v_1)\bigr]^{1/2} = -\lambda_1 \tau t \sqrt{w_2} + O(\tau^2)$$

and

$$f\bigl(w_2 + t(v_2 - w_2), y\bigr) \sim f\bigl(w_2 - t(w_2 - v_1), y\bigr) = -\lambda_1 \tau t f(w_2, y) + O(\tau^2).$$

As a result we obtain that

$$\Delta(\tau) = -\lambda_1 \tau \biggl[\int_0^1 dt \int_{-\sqrt{w_2}}^0 f(w_2, y)\, dy - \int_0^1 dx \int_0^{\sqrt{w_2}} f(w_2, y)\, dy \biggr] + O(\tau^2) = O(\tau^2),$$

since $f(x, y)$ is an even function of y.

As $\tau \to \infty$, by the same reason the difference of integrals has the order $O(\varepsilon^2)$ and the last integral in the representation of $Q_p^+(\tau)$ (cf. Proposition 4) has the order

$$\int_{\tau - \sqrt{a}}^{\tau + \sqrt{a}} dx \int_{-\sqrt{x}}^{(a - \tau^2 - x)/(2\tau)} f(x, y)\, dy = O\biggl(\tau^{p-2} \exp\biggl\{-\frac{\tau^2}{2}\biggr\}\biggr).$$

To prove this, it is sufficient to substitute $y = -\sqrt{x}t$ in the dy-integral.

4. Concluding remarks. The method of variable substitution in integration with the subsequent application of Taylor's expansion for the dy-integrals, used in the proof of the theorem, allowed us to produce asymptotic expansions of the coverage probability for any small values of τ. Of course, for the case $\tau \to \infty$ we will obtain the same asymptotic formulas as in the papers of Berger [1] and Hwang and Casella [8] in this case, too. We decided to not provide the subsequent terms of the expansion due to the complexity of the coefficients, which are in the double integrals form as in the formulas for the exact values of coverage probability.

We can judge the accuracy of the $K_p(w_2(c, \tau))$ asymptotics for $Q_p^+(\tau)$ for values $p = 3$, 5, 10 and $\alpha = 0.05$ according to Figure 5 and Table 1.

The thick lines in Figure 5 correspond to the exact values of $Q_p^+(\tau)$, while the thin lines correspond to the asymptotic values.

Notations in the table are $Q = Q_p^+(\tau) - 0.95$ and the approximation error (with the sign) $\Delta = Q_p^+(\tau) - K_p(w_2(c, \tau))$.

Figure 5 and the table show the high accuracy of the asymptotic formula for $\tau \le c$. For $\tau > c$ the exact value of coverage probability decreases significantly faster than the asymptotics obtained in the paper. Starting from $p = 5$ the approximation error begins to decrease, and its maximum value at $p = 5$ is equal to 0.0134. Hence it is possible to conclude with a practical acceptability of the coverage probability calculations according to the asymptotic formulas which are provided in the theorem.

TABLE 1

τ	$p=3$		$p=4$		$p=5$		$p=7$		$p=10$	
	Q	Δ	Q	Δ	Q	Δ	Q	Δ	Q	Δ
0.0	.0288	.0000	.0396	.0000	.0445	.0000	.0483	.0000	.0496	.0000
0.5	.0280	+.0003	.0386	-.0007	.0443	+.0000	.0482	+.0000	.0496	+.0000
1.0	.0278	+.0008	.0374	-.0009	.0438	+.0001	.0480	+.0000	.0496	+.0000
1.5	.0265	+.0016	.0356	-.0011	.0430	+.0004	.0476	+.0000	.0495	+.0000
2.0	.0246	+.0020	.0332	-.0013	.0417	+.0006	.0470	+.0001	.0493	+.0000
2.5	.0216	+.0020	.0295	-.0031	.0399	+.0008	.0462	+.0002	.0491	+.0000
3.0	.0133	-.0038	.0185	-.0104	.0373	+.0007	.0450	+.0003	.0487	+.0001
3.5	.0068	-.0080	.0126	-.0133	.0303	-.0036	.0433	+.0003	.0482	+.0001
4.0	.0043	-.0084	.0096	-.0134	.0219	-.0091	.0375	-.0034	.0473	+.0001
4.5	.0031	-.0078	.0077	-.0127	.0171	-.0110	.0311	-.0075	.0444	-.0017
5.0	.0025	-.0070	.0064	-.0117	.0140	-.0114	.0266	-.0096	.0402	-.0045
5.5	.0020	-.0062	.0054	-.0106	.0118	-.0111	.0230	-.0105	.0367	-.0064
6.0	.0017	-.0055	.0046	-.0096	.0100	-.0106	.0201	-.0110	.0334	-.0078
6.5	.0014	-.0050	.0040	-.0087	.0086	-.0099	.0176	-.0110	.0302	-.0089
7.0	.0012	-.0044	.0034	-.0078	.0075	-.0092	.0156	-.0108	.0274	-.0096
7.5	.0011	-.0039	.0030	-.0071	.0066	-.0085	.0138	-.0104	.0248	-.0100
8.0	.0009	-.0036	.0027	-.0061	.0058	-.0080	.0123	-.0100	.0225	-.0102
8.5	.0008	-.0032	.0024	-.0058	.0052	-.0073	.0110	-.0094	.0205	-.0102
9.0	.0007	-.0030	.0022	-.0053	.0046	-.0068	.0099	-.0089	.0187	-.0100
9.5	.0007	-.0026	.0020	-.0049	.0042	-.0062	.0090	-.0084	.0171	-.0098
10.0	.0006	-.0024	.0018	-.0044	.0038	-.0057	.0082	-.0079	.0157	-.0094
11.0	.0005	-.0020	.0015	-.0038	.0031	-.0050	.0068	-.0070	.0133	-.0087
12.0	.0004	-.0017	.0012	-.0032	.0026	-.0043	.0058	-.0062	.0114	-.0080
14.0	.0003	-.0013	.0009	-.0024	.0019	-.0033	.0043	-.0048	.0086	-.0066
16.0	.0002	-.0010	.0007	-.0019	.0015	-.0026	.0033	-.0039	.0067	-.0054
18.0	.0002	-.0008	.0006	-.0015	.0012	-.0021	.0026	-.0032	.0054	-.0045
20.0	.0002	-.0006	.0005	-.0012	.0009	-.0016	.0022	-.0026	.0044	-.0038
22.0	.0001	-.0006	.0004	-.0010	.0008	-.0014	.0018	-.0022	.0036	-.0032
25.0	.0001	-.0004	.0001	-.0010	.0006	-.0011	.0014	-.0017	.0028	-.0025
26.0	-	-	-	-	-	-	.0013	-.0016	.0026	-.0024
28.0	-	-	-	-	-	-	.0011	-.0014	.0023	-.0020
30.0	-	-	-	-	-	-	.0010	-.0012	.0020	-.0018

REFERENCES

[1] J. O. BERGER, *A robust generalized Bayes estimator and confidence region for a multivariate normal mean*, Ann. Statist., 8 (1980), pp. 716–761.

[2] L. D. BROWN, *On the admissibility of invariant estimators of one or more location parameters*, Ann. Math. Statist., 37 (1966), pp. 1087–1136.

[3] N. G. DE BRUIJN, *Asymptotic Methods in Analysis*, North-Holland, Amsterdam, 1958.

[4] G. CASELLA AND J. T. HWANG, *Confidence sets and the Stein effect*, Comm. Statist. Theory Methods, 15 (1986), pp. 2043–2063.

[5] W. JAMES AND C. STEIN, *Estimation with quadratic loss*, in Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, Univ. California Press, Berkeley, CA, 1961, pp. 361–379.

[6] V. M. JOSHI, *Inadmissibility of the usual confidence sets for the mean of a multivariate normal population*, Ann. Math. Statist., 38 (1967), pp. 1868–1875.

[7] J. T. HWANG AND G. CASELLA, *Minimax confidence sets for the mean of a multivariate normal distribution*, Ann. Statist., 10 (1982), pp. 868–881.

[8] J. T. HWANG AND G. CASELLA, *Improved set estimators for a multivariate normal mean*, Special issue on recent results in estimation theory and related topics, Statist. Decisions, suppl. 1 (1984), pp. 3–16.

[9] C. ROBERT AND A. K. MD. E. SALEH, *Recentered confidence sets: A review*, in Data Analysis from Statistical Foundations, Nova Sci. Publ., Huntington, NY, 2001, pp. 295–308.

[10] A. K. Md. E. Saleh, *Asymptotic properties of Stein-type confidence sets—a U-statistics approach*, Sankhyā, 65 (2003), pp. 411–421.
[11] C. Stein, *Confidence sets for the mean of a multivariate normal distribution*, J. Roy. Statist. Soc. Ser. B, 24 (1962), pp. 265–296.
[12] E. L. Lehmann and G. Casella, *Theory of Point Estimation*, 2nd ed., Springer-Verlag, New York, 1998.