THE ASYMPTOTIC EXPANSION OF THE STEIN ESTIMATORS
FOR THE VECTOR CASE*

A. J. van der Merwe and D. J. de Waal

(Received Dec. 14, 1977; revised Nov. 14, 1978)

Summary

Charles Stein established the existence of estimators which dominate the maximum likelihood estimators for the problem of simultaneously estimating the means of three or more random variables.

Since the exact distributions of the Stein estimators are not known and because the distributions are of great importance for people studying confidence sets, it was the purpose of this note to derive the asymptotic distributions, means and variances of the Stein estimators, as well as that of the quadratic loss functions for the vector case.

1. Introduction

Suppose \( X_1, \cdots, X_N \) is a sample of \( N \) observations from a \( N(\theta, \Sigma) \) population where \( \Sigma = I_p \) and \( \theta(p \times 1) \) is unknown, then a Stein estimator (James and Stein [2]) of \( \theta \) is defined as

\[
\psi'(\bar{X}) = \left(1 - \frac{c}{N\bar{X}'\bar{X}}\right)\bar{X}
\]

where \( \bar{X} = \frac{1}{N} \sum_{i=1}^{N} X_i \sim N(\theta, I_p/N) \), the maximum likelihood estimator of \( \theta \), and \( c = p - 2 \).

At the Fourth Berkeley Simposium, James and Stein [2] proved that for \( p \geq 3 \), \( \psi(\bar{X}) \) is a better estimator than \( \bar{X} \), because

\[
E N(\psi'(\bar{X}) - \theta)'(\psi'(\bar{X}) - \theta) = p - (p - 2)E \left(\frac{1}{N\bar{X}'\bar{X}}\right) \leq p
\]

where

\[
E N(\bar{X} - \theta)'(\bar{X} - \theta) = p .
\]

* Financially supported by the CSIR and the University of the OFS Research Fund.
Let $\bar{X}=\theta+(1/\sqrt{N})Z$ where $Z \sim N(0, I_0)$ and expanding (1.1) in a series (see Sec. 12) we get for large $N$ that

$$
\Psi(\bar{X}) = h_0(Z) + \frac{1}{\sqrt{N}} h_1(Z) + \frac{1}{N} h_2(Z) + \frac{1}{N\sqrt{N}} h_3(Z) + \frac{1}{N^2} h_4(Z) + \frac{1}{N^2\sqrt{N}} h_5(Z) + \frac{1}{N^3} h_6(Z) + R
$$

where

$$
(1.5) \quad h_0(Z) = \theta
$$

$$
(1.6) \quad h_1(Z) = Z
$$

$$
(1.7) \quad h_2(Z) = -\frac{c\theta}{a_0}
$$

$$
(1.8) \quad h_3(Z) = \frac{c}{a_0} (\theta g_5(Z) - Z)
$$

$$
(1.9) \quad h_4(Z) = \frac{c}{a_0} \{Zg_1(Z) - \theta (g_5(Z) - g_6(Z))\}
$$

$$
(1.10) \quad h_5(Z) = -\frac{c}{a_0} \{\theta (2g_1(Z)g_2(Z) - g_5(Z)) + Z(g_5(Z) - g_6(Z))\}
$$

$$
(1.11) \quad h_6(Z) = -\frac{c}{a_0} \{\theta (g_5(Z) - 3g_5(Z)g_2(Z) + g_1(Z)) + Z(2g_2(Z)g_3(Z) - g_4(Z))\}
$$

and

$$
(1.12) \quad a_0 = \theta \theta
$$

$$
(1.13) \quad g_1(Z) = \frac{2Z'\theta}{a_0}
$$

$$
(1.14) \quad g_2(Z) = \frac{Z'Z}{a_0}
$$

Throughout this paper $R$ will indicate a residual term that consists of order terms in $N$. These order terms will contain the next power terms in $N$ which can be seen from the series under consideration plus additional terms that may arise, because we have ignored the convergence constraints as explained in Sec. 12.

In the next section we are going to give some expected values for further use. In Sec. 3 the asymptotic characteristic function of $V^* = \sqrt{N}(\Psi(\bar{X}) - h_0(Z))$ is derived and in Sec. 4 the asymptotic probability density function of $V^*$ is derived. In Sec. 5 the asymptotic values of
E \mathcal{V}(\bar{X}) and Var \mathcal{V}(\bar{X}) are given and in Sec. 6 the quadratic loss function U is defined. In Secs. 7 and 8 the asymptotic characteristic function as well as the asymptotic probability density function of U are derived and in Sec. 9, the asymptotic values of E(U) and Var(U) are given.

The asymptotic distributions derived in Secs. 3–8 are obtained by using a method similar to the perturbation method, used by Nagao [3].

In Sec. 10 the Stein estimator \mathcal{V}(\bar{X}) for the case Σ ≠ I_p is defined and a method for obtaining the asymptotic results is given.

In Sec. 11 the advantages of the Stein estimator over the maximum likelihood estimator are mentioned.

In Sec. 12 the theoretical justification for the convergence of equation (1.4) is given.

2. Expected values

**Lemma 2.1.** Suppose Z ∼ N(0, I_p), M(p×p), θ(p×1), s(p×1), then

\begin{align}
(2.1) & \quad E(\mathcal{V}Z'MZZ') = 2M + I_p\text{ tr } M \\
(2.2) & \quad E(Z'Z)^2 = p(p+2) \\
(2.3) & \quad E(Z'\theta)^2 = 3(\theta'\theta)^2 \\
(2.4) & \quad E(Z'\theta)(Z'Z) = (p+2)(\theta'\theta) \\
(2.5) & \quad E(s'Z)(Z'\theta)(Z'Z) = (p+2)(s'\theta) \\
(2.6) & \quad E(Z'\theta)(Z's) = 3(\theta'\theta)(s'\theta). 
\end{align}

The proofs of equations (2.1)–(2.6) follow easily.

3. On the characteristic function

**Theorem 3.1.** The characteristic function of \( V^* = \sqrt{N}(\mathcal{V}(\bar{X}) - h_0(Z)) \) for large N is given by

\begin{align}
(3.1) & \quad \phi_{\mathcal{V}}(it) = e^{-t^2/2} \left\{ 1 + \frac{1}{\sqrt{N}} \phi_1(t) + \frac{1}{N} \phi_2(t) + \frac{1}{N^2} \phi_3(t) + \frac{1}{N} \phi_4(t) + R \right\} \\
\text{where} & \\
(3.2) & \quad \phi_1(t) = -\frac{i(p-2)}{a_0} (t'\theta) \\
(3.3) & \quad \phi_2(t) = \frac{(p-2)}{a_0} \left\{ (t't) - \frac{1}{2a_0} (p+2)(t'\theta)^2 \right\}
\end{align}
\begin{align}
(3.4) \quad \Phi_s(t) &= \frac{(p-2)i}{a_0^3} \left[ (p-2)\langle t' \theta \rangle - (p+1)\langle t' t \rangle \langle t' \theta \rangle \\
&\quad + \frac{1}{6a_0} (4+8p+p^3)\langle t' \theta \rangle^2 \right] \\
(3.5) \quad \Phi_\ell(t) &= \frac{(p-2)}{a_0^3} \left\{ \frac{p}{2} \langle t' t \rangle^2 - \frac{3(p-2)}{2} \langle t' t \rangle - \frac{p}{2a_0} (p+6)\langle t' t \rangle \langle t' \theta \rangle^2 \\
&\quad + \frac{1}{a_0} (p+2)\langle p-2 \rangle \langle t' \theta \rangle^2 + \frac{1}{24a_0^3} (60p+18p^2+p^3-8)\langle t' \theta \rangle^4 \right\}.
\end{align}

**Proof.**

\begin{align}
(3.6) \quad \phi_v(i \ell) &= E e^{i t h_\ell(x)} \left\{ 1 + \frac{1}{\sqrt{N}} (i t' h_\ell(Z)) + \frac{1}{N} \left( i t' h_\ell(Z) + \frac{1}{2} (i t' h_\ell(Z))^2 \right) \\
&\quad + \frac{1}{N \sqrt{N}} \left( i t' h_\ell(Z) + (i t' h_\ell(Z^2))(i t' h_\ell(Z)) + \frac{1}{6} (i t' h_\ell(Z))^3 \right) \\
&\quad + \frac{1}{N^2} \left( i t' h_\ell(Z) + (i t' h_\ell(Z))(i t' h_\ell(Z)) + \frac{1}{2} (i t' h_\ell(Z))^2 \right) \\
&\quad + \frac{1}{2} (i t' h_\ell(Z))^2 (i t' h_\ell(Z)) + \frac{1}{24} (i t' h_\ell(Z))^4 + R \right\}.
\end{align}

By using Lemma 2.1 and (3.6) and the fact that

\begin{align}
(3.7) \quad E e^{i t h_\ell(Z)} h_\ell(Z) &= e^{-\ell t/2} E h_\ell(Y + i \ell), \quad i = 1, 2, \ldots
\end{align}

where \( Y \sim N(0, I) \) we find that

\begin{align}
(3.8) \quad E i t' h_\ell(Z) e^{i t h_\ell(Z)} &= -\frac{c i}{a_0} \langle t' \theta \rangle e^{-\ell t/2} \\
(3.9) \quad E i t' h_\ell(Z) e^{i t h_\ell(Z)} &= \frac{c}{a_0} \left( \langle t' t \rangle - \frac{2}{a_0} \langle t' \theta \rangle \right) e^{-\ell t/2} \\
(3.10) \quad E i t' h_\ell(Z) e^{i t h_\ell(Z)} &= \frac{ci}{a_0^3} \left( (p-2)\langle t' \theta \rangle - 3\langle t' t \rangle \langle t' \theta \rangle + \frac{4}{a_0} \langle t' \theta \rangle^3 \right) e^{-\ell t/2} \\
(3.11) \quad E i t' h_\ell(Z) e^{i t h_\ell(Z)} &= -\frac{c}{a_0^3} \left( 8 \langle t' t \rangle \langle t' \theta \rangle^2 - \frac{4}{a_0} (p-2)\langle t' \theta \rangle^2 \\
&\quad - \frac{8}{a_0^3} \langle t' \theta \rangle^4 + (p-2)\langle t' t \rangle - \langle t' \rangle \right) e^{-\ell t/2} \\
(3.12) \quad E (i t' h_\ell(Z))^2 e^{i t h_\ell(Z)} &= \frac{c^2}{a_0^5} \left( \frac{4}{a_0} \langle t' \theta \rangle^4 - \langle t' \rangle + \langle t' \rangle^3 - \frac{4}{a_0} \langle t' t \rangle \langle t' \theta \rangle \right) e^{-\ell t/2}.
\end{align}

Substituting (3.8)–(3.12) in (3.6) and (3.1) follows.
4. The distribution of $V^*$

**Theorem 4.1.** The probability density function of $V^*$ for large $N$ is given by

$$f_{V^*}(v^*) = \left( \frac{1}{2\pi} \right)^{p/2} e^{-v^* v / 2} \left[ 1 + \frac{1}{N} \tilde{g}_i(v^*) + \frac{1}{N^2} \tilde{g}_j(v^*) + \mathcal{R} \right]_{-\infty < v^* < \infty}$$

where

$$\tilde{g}_i(v^*) = -\frac{1}{a_0} (p-2)(v^* \theta)$$

$$\tilde{g}_j(v^*) = \frac{(p-2)}{a_0} \left( \frac{1}{2} (p-2) - v^* v^* + \frac{1}{2a_0} (p+2)(v^* \theta)^2 \right)$$

$$\tilde{g}_k(v^*) = \frac{(p-2)}{a_0} \left( \frac{1}{2} (4p-4-p^2)(v^* \theta) + (p+1)(v^* v^*) (v^* \theta) \right) - \frac{1}{6a_0} (4+8p+p^2)(v^* \theta)^3$$

$$\tilde{g}_l(v^*) = \frac{(p-2)}{a_0} \left( \frac{1}{8} (p^2 - 8p + 20) + \frac{p}{2} (v^* v^*) \right) + \frac{1}{2} (p-2)(3-p)(v^* v^*) + \frac{1}{24a_0} (p^2 + 18p + 60p - 8)(v^* \theta)^4$$

$$- \frac{p}{2a_0} (6+p)(v^* v^*)(v^* \theta)^3 + \frac{(p-2)}{4a_0} (p^2 - 12)(v^* \theta)^4.$$  

**Proof.**

$$f_{V^*}(v^*) = \left( \frac{1}{2\pi} \right)^p \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \phi_{V^*}(it)e^{-i\theta v^*} \, dt_1 \cdots dt_p$$

where $\phi_{V^*}(it)$ is the characteristic function of $V^*$. By using Lemma 2.1 and the fact that $E k^*(t)e^{-i\theta v^*} = E k^*(Y - v^*)$, the proof follows easily. ($k^*(t)$ is a function in $t$. The first expectation is taken over $t$ and the second over $Y$ where $Y \sim N(0, I_p).$)

It is clear that if $N \to \infty$ then $V^* \sim N(0, I_p)$.

5. E $\bar{X}$ and Var $\bar{X}$

**Theorem 5.1.**
\[
\begin{align*}
\mathbb{E} \Psi(\bar{X}) &= \mathbb{E} \left\{ h_0(Z) + \frac{1}{\sqrt{N}} h_1(Z) + \frac{1}{N} h_2(Z) + \frac{1}{N\sqrt{N}} h_3(Z) \\
&\quad + \frac{1}{N^2} h_4(Z) + \frac{1}{N^2\sqrt{N}} h_5(Z) + \frac{1}{N^3} h_6(Z) + R \right\} \\
\text{and} \\
\text{Var} \, \Psi(\bar{X}) &= \frac{1}{N} \text{Var} \, h_1(Z) + \frac{2}{N^2} \text{Cov} \, (h_1(Z), h_3(Z)) \\
&\quad + \frac{1}{N^3} \{\text{Var} \, h_3(Z) + 2 \text{Cov} \, (h_1(Z), h_3(Z))\} + R
\end{align*}
\]

where
\[
\begin{align*}
\mathbb{E} h_0(Z) &= \theta \\
\mathbb{E} h_1(Z) &= 0 \\
\mathbb{E} h_2(Z) &= -\frac{c}{a_0} \theta \\
\mathbb{E} h_3(Z) &= 0 \\
\mathbb{E} h_4(Z) &= \frac{c(p-2)}{a_0^2} \theta \\
\mathbb{E} h_5(Z) &= 0 \\
\mathbb{E} h_6(Z) &= \frac{c(p-2)(p-4)}{a_0^3} \theta \\
\text{Var} \, h_1(Z) &= I_p \\
\text{Cov} \, (h_1(Z), h_3(Z)) &= \frac{c}{a_0} \left\{ \frac{2\theta' \theta}{a_0} - I_p \right\} \\
\text{Var} \, h_3(Z) &= \frac{c^2}{a_0^3} I_p \\
\text{Cov} \, (h_1(Z), h_3(Z)) &= -\frac{c(p-2)}{a_0^3} \left\{ \frac{1}{a_0} 4\theta' \theta - I_p \right\}.
\end{align*}
\]

**Proof.** By using Lemma 2.1 the proof follows easily \(c=p-2\).

6. The quadratic loss function

The quadratic loss function for \(N\) large is defined as
\[
U = N(\Psi(\bar{X}) - \theta)'(\Psi(\bar{X}) - \theta)
\]
\[
(6.2) \quad = V^* V^* \\
(6.3) \quad = g_0^*(Z) + \frac{1}{\sqrt{N}} g_1^*(Z) + \frac{1}{N} g_2^*(Z) + \frac{1}{N \sqrt{N}} g_3^*(Z) + \frac{1}{N^2} g_4^*(Z) + R \\
\]

where

\[
(6.4) \quad g_0^*(Z) = h_0(Z) h_1(Z) \\
(6.5) \quad g_1^*(Z) = 2 h_1(Z) h_2(Z) \\
(6.6) \quad g_2^*(Z) = h_3(Z) h_4(Z) + 2 h_1(Z) h_3(Z) \\
(6.7) \quad g_3^*(Z) = 2 \{ h_3(Z) h_4(Z) + h_2(Z) h_3(Z) \} \\
(6.8) \quad g_4^*(Z) = h_3(Z) h_4(Z) + 2 h_2(Z) h_4(Z) + 2 h_1(Z) h_3(Z) . \\
\]

7. The characteristic function of \( U \)

**Theorem 7.1.** The characteristic function of \( U \) is given by

\[
(7.1) \quad \phi_U(it) = (1 - 2it)^{-\nu^2} \left( 1 + \frac{1}{\sqrt{N}} \Phi_0^*(t) + \frac{1}{N} \Phi_1^*(t) + \frac{1}{N \sqrt{N}} \Phi_2^*(t) \right) + \frac{1}{N^2} \Phi_3^*(t) + R \]

where

\[
(7.2) \quad \Phi_0^*(t) = 0 \\
(7.3) \quad \Phi_1^*(t) = \frac{c^2}{a_0} (it) - \frac{2c(p-2)}{a_0} (it) (1-2it)^{-1} + \frac{2c^2}{a_0} (it)^2 (1-2it)^{-1} \\
(7.4) \quad \Phi_2^*(t) = 0 \\
(7.5) \quad \Phi_3^*(t) = \frac{2c}{a_0} (p-4)(p-2)it(1-2it)^{-2} + \frac{c^2}{a_0} (4-p)it(1-2it)^{-1} \\
+ \frac{c^2}{2a_0} (it)^2 - \frac{2c^3}{a_0^3} (p-4)(it)^3 (1-2it)^{-1} + \frac{2c^3}{a_0^3} (p-4)(it)^3 (1-2it)^{-2} \\
+ \frac{4c^3}{a_0^3} (4-p)(1-2it)^{-2} (it) + \frac{2c^4}{a_0^3} (it)^4 (1-2it)^{-1} \\
+ \frac{2c^4}{a_0^3} (it)^4 (1-2it)^{-2} . \\
\]

**Proof.** By using the same method as given in Theorem 3.1, the result follows.
8. The distribution of $U$

**Theorem 8.1.** The probability density function of $U$ defined in (6.1) for large $N$ is given by

\begin{equation}
 f_U(u) = \frac{1}{\Gamma(p/2)2^{p/2}} e^{-u^2/2p^2} \left[ 1 + \frac{1}{\sqrt{N}} \tilde{h}_e(u) + \frac{1}{N} \tilde{h}_o(u) + \frac{1}{N\sqrt{N}} \tilde{h}_2(u) + \frac{1}{N^2} \tilde{h}_o(u) + R \right] \quad u > 0
\end{equation}

where

\begin{align}
(8.2) \quad \tilde{h}_e(u) &= 0 \\
(8.3) \quad \tilde{h}_o(u) &= \frac{(p-2)^2}{2a_0} \left( 1 - \frac{1}{p} u \right) \\
(8.4) \quad \tilde{h}_o(u) &= 0 \\
(8.5) \quad \tilde{h}_o(u) &= \frac{(p-2)^2}{8a_0^2} \left\{ (p^2 - 8p + 20) + \frac{2}{p} u(-p^2 + 6p - 12) + \frac{u(p-2)^2}{p(p+2)} \right\}.
\end{align}

**Proof.**

\begin{equation}
 f_U(u) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \phi(t) e^{-itu} \, dt
\end{equation}

and equation (8.1) follows. For further algebraic details see van der Merwe and de Waal [5].

9. $E(U)$ and $\text{Var}(U)$

The risk is defined as the expected value of the quadratic loss function.

**Theorem 9.1.** For the density function

\begin{equation}
 f_U(u) = \tilde{f}_U(u) \left[ 1 + \frac{1}{\sqrt{N}} \tilde{h}_e(u) + \frac{1}{N} \tilde{h}_o(u) + \frac{1}{N\sqrt{N}} \tilde{h}_2(u) + \frac{1}{N^2} \tilde{h}_o(u) \right] \quad u > 0
\end{equation}

we find that

\begin{equation}
 E(U) = p - \frac{(p-2)^2}{N a_0} + \frac{(p-2)^2}{N^2 a_0^2} (p-4)
\end{equation}
\begin{align}
(9.3) \quad \text{Var}(U) & = 2p - \frac{4(p-2)^4}{Na_0^2} + \frac{4(p-2)^4(p-4)}{N^2a_0^4} - \frac{2(p-2)^4(4-p)}{N^3a_0^6} \\
& \qquad - \frac{(p-2)^4(p-4)^2}{N^4a_0^8} . \\
\end{align}

\textbf{Proof.} \quad \mathbb{E}(U^r) = \int_0^\infty u^rf_U(u)du. \quad \text{For } r=2, \text{ we find that}

\begin{align}
(9.4) \quad \mathbb{E}(U^2) & = p(p+2) - \frac{2(p-2)^4(p+2)}{Na_0^2} + \frac{(p-2)^4}{N^2a_0^4} (3p^2 - 8p - 12) \\
\end{align}

\text{Var}(U) = \mathbb{E}(U^2) - \{\mathbb{E}(U)\}^2 \text{ and (9.3) follows (Ullah [4] derived the first four moments as well as their approximations for the marginal distributions of the Stein estimator).}

\section{The case \(\Sigma \neq I_p\)}

The Stein estimator, for the case \(\Sigma\) unknown, is defined as

\begin{align}
(10.1) \quad \psi(\bar{X}) & = \left(1 - \frac{c_1}{NX'X^{-1}X} \right) \bar{X} \\
\end{align}

where

\[ c_1 = \frac{p-2}{N-p+2} \]

and

\[ A = \sum_{i=1}^N (X_i - \bar{X})(X_i - \bar{X})' \sim W(\Sigma, N-1) . \]

By making use of the fact that \(NX'X^{-1}X \sim NX'X/\Sigma\), where \(S \sim \chi^2_{N-p}\) and independent of \(X\), (see Wijsman [6]), (10.1) can be written as:

\begin{align}
(10.2) \quad \psi(\bar{X}) & = \left(1 - \frac{c^*}{NX'X} \right) \bar{X} \\
\end{align}

where

\begin{align}
(10.3) \quad c^* & = \frac{(p-2)}{(N-p+2)} S \\
\end{align}

and

\begin{align}
(10.4) \quad \bar{X} & = D^{-1} \left( D\theta + \frac{1}{\sqrt{N}} Z \right) \\
\end{align}

where \(Z \sim N(0, I_p)\) and \(D\Sigma D' = I_p\)

\begin{align}
(10.5) \quad a_0 & = \theta'\Sigma^{-1}\theta . \\
\end{align}
By making use of the same methods as in Secs. 1–9 the asymptotic
distribution, mean and variance of this estimator as well as that of the
quadratic loss function can be obtained.

11. Advantages of the Stein estimator over the maximum likelihood
estimator

The advantages of the Stein estimator over the maximum likeli-
hood estimator, is that the variance, risk (expected value of the quad-
ritic loss function) and variance of the quadratic loss function in the
case of the Stein estimator, are smaller than in the case of the maxi-
umum likelihood estimator.

The distribution of the Stein estimator as well as that of the quad-
ractic loss function will be useful for studying confidence sets.

12. Justification for the use of the inverse binomial series

If we put \( \bar{X} = \theta + (1/\sqrt{N})Z \) where \( Z \sim N(0, I_p) \), then \( \psi(\bar{X}) \) as defined
in equation (1.1) can be written as

\[
\psi(\bar{X}) = \left(1-\frac{2}{Na_0(1+(2/a_0\sqrt{N})Z'\theta+(1/Na_0)Z'Z)}\right)\left(\theta+\frac{1}{\sqrt{N}}Z\right).
\]

By making use of the inverse binomial theorem (12.1) can be written
as equation (1.4) which will converge, if

\[
1 < \frac{2}{a_0\sqrt{N}}Z'\theta + \frac{1}{Na_0}Z'Z < 1
\]

where

\[a_0 = \theta'\theta .\]

However by taking expected values in the previous sections, we in-
tegrated \( Z \) over the entire real space. This can be done if \( N \) and \( \theta \) are
large enough because the difference in the integration taken over the
region for convergence and the whole space is small, as explained by
de Bruijn [1].

By calculation of exact values of the mean and variance of the
Stein estimator and by comparing them with the asymptotic values this
justification was confirmed. The asymptotic distribution of the quad-
ratic loss function also compared well with the exact distribution ob-
tained from Monte Carlo simulation experiments. See van der Merwe
and de Waal [5].
REFERENCES


