Ann. Inst. Statist. Math.
35 (1983), Part A, 31-39

THE ASYMPTOTIC EXPANSION AS WELL AS THE EXACT MOMENTS OF THE STEIN ESTIMATOR WHEN THE POPULATION MEANS ARE NEARLY EQUAL*

A. J. VAN DER MERWE

(Received Sept. 17, 1981; revised Jan. 26, 1982)

Summary

The purpose of this note is to derive the asymptotic distributions, means and variances of the Stein estimator, as well as that of the quadratic loss function for the vector case when the population means are nearly equal. These results are given in Section 3 and are obtained by using a method similar to the perturbation method, used by Nagao [4]. In Section 4 exact moments of the Stein estimator are also derived.

1. Introduction

Suppose $X_1, \cdots, X_N$ is a sample of N observations from an $N(\theta, I_p)$ population where $\theta(p\times 1)$ is unknown. If it is suspected that the θ_i, $i=1, \cdots, p$ are nearly equal, but not near the origin, then Lindley (in the discussion following Stein's [5] paper) (see also Efron and Morris [3]) suggested the following Stein estimator of θ:

$$\phi(\bar{X})=\bar{\bar{X}}_p e+\left(1-\frac{c}{N(\bar{X}-\bar{\bar{X}}_p e)'(\bar{X}-\bar{\bar{X}}_p e)}\right)(\bar{X}-\bar{\bar{X}}_p e) \tag{1.1}$$

where $\bar{X}=\frac{1}{N}\sum_{j=1}^{N} X_j \sim N(\theta, I_p/N)$ is the maximum likelihood estimator of θ, $c=p-3$,

$$\bar{\bar{X}}_p=\frac{1}{p}\sum_{i=1}^{p}\bar{X}_i \sim N\left(\bar{\theta}_p, \frac{1}{N_p}\right) \quad \text{with} \quad \bar{\theta}_p=\frac{1}{p}\sum_{i=1}^{p}\theta_i ,$$

$$e'=(1, 1, \cdots, 1) \quad \text{and} \quad \bar{X}=(\bar{X}_1, \cdots, \bar{X}_p)' .$$

$\phi(\bar{X})$ is a better estimator than $\bar{X}$, because its risk

---

* Financially supported by the CSIR and the University of the OFS Research Fund.

$$\mathrm{E}\, N(\psi(\bar{X})-\theta)'(\psi(\bar{X})-\theta)=p-(p-3)^2\, \mathrm{E}\, \frac{1}{p-3+2K}<p$$

where K is a Poisson random variable with parameter $\tilde{\lambda}=(1/2)N(\theta-\bar{\theta}_p e)' \cdot(\theta-\bar{\theta}_p e)$ and p is the risk of the maximum likelihood estimator.

Efron and Morris [2] showed that $\bar{X}-\bar{\bar{X}}_p e=Y\sim N\{\Lambda, (1/N)(I_p-(1/p)ee')\}$ and is independent of $\bar{\bar{X}}_p e$ where $\Lambda=\theta-\bar{\theta}_p e$. Let $\tilde{Z}=(\tilde{Z}_1, \cdots, \tilde{Z}_p)=\Gamma Y$ with Γ an orthogonal matrix with last row equal to $(1/\sqrt{p})e'$ and $\zeta=(\zeta_1, \cdots, \zeta_p)'=\Gamma\Lambda$ then $\tilde{Z}\sim N\{\zeta, (1/N)(I_p-\delta\delta')\}$ where $\delta'=(0, 0, \cdots, 0, 1)$. Hence $\tilde{Z}_p=\zeta_p=0$ with probability one and $(\tilde{Z}_1, \cdots, \tilde{Z}_{p-1})'\sim N(\omega, (1/N)I_{p-1})$ where $\omega'=(\zeta_1, \cdots, \zeta_{p-1})$. Therefore (1.1) can be written as

$$\begin{aligned}(1.2) \quad \psi(\bar{X})=&\left(\frac{1}{\sqrt{N}\sqrt{p}}Z_p^* e+\bar{\theta}_p e\right)+\left(1-\frac{c}{N\left(\frac{1}{\sqrt{N}}Z+\zeta\right)'\left(\frac{1}{\sqrt{N}}Z+\zeta\right)}\right)\\ &\times\Gamma'\left(\frac{1}{\sqrt{N}}Z+\zeta\right)\end{aligned}$$

where $\tilde{Z}=\zeta+(1/\sqrt{N})Z$ and $\bar{\bar{X}}_p e=(1/(\sqrt{p}\sqrt{N}))Z_p^* e+\bar{\theta}_p e$, $Z\sim N(0, I_p-\delta\delta')$; $Z_p^*\sim N(0, 1)$ and independent of each other.

Expanding (1.2) in a series we get for N large enough that

$$(1.3) \quad \psi(X)=h_0(Z)+\frac{1}{\sqrt{N}}h_1(Z)+\frac{1}{N}h_2(Z)+\frac{1}{N\sqrt{N}}h_3(Z)+\frac{1}{N^2}h_4(Z)+R$$

where
$$h_0(Z)=\bar{\theta}_p e+\Gamma'\zeta=\bar{\theta}_p e+\Lambda=\theta ,$$

$$h_1(Z)=\Gamma' Z+\frac{1}{\sqrt{p}}Z_p^* e ,$$

$$h_2(Z)=\frac{-c\Lambda}{a_0} ,$$

$$h_3(Z)=\frac{c}{a_0}(\Lambda g_1(Z)-\Gamma' Z) ,$$

$$h_4(Z)=\frac{c}{a_0}\{\Gamma' Z g_1(Z)-\Lambda(g_1^2(Z)-g_2(Z))\}$$

and
$$a_0=\zeta'\zeta=(\theta-\bar{\theta}_p e)'(\theta-\bar{\theta}_p e) ,$$

$$g_1(Z)=\frac{2}{a_0}Z'\zeta ,$$

$$g_2(Z)=\frac{1}{a_0}Z'Z .$$

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 the convergence constraints have been ignored. This can be done if N and θ are large enough, as explained by Copson [1]. (See also Subsection 4.2 for further clarification).

In the next section some lemmas are given. These lemmas are of importance in the derivation of asymptotic distributions.

2. Lemmas for further use

LEMMA 2.1. *Suppose* $Z(p\times 1)\sim N(0, I_p-\delta\delta')$ *where* $\delta'=(0,0,\cdots,0,1)$, *then* $\mathrm{E}\, h_j(Z)e^{it'h_1(Z)}=e^{-t't/2}\,\mathrm{E}\, h_j\{A(Y^*+iA'\Gamma t)\}$; $j=2,3,4$ *where* $Y^*\sim N(0, I_{p-1})$; $A(p\times p-1)=\begin{bmatrix} I_{p-1} \\ 0\cdots 0\end{bmatrix}$; $t'=(t_1,\cdots,t_p)$ *and* $h_j(Z)$; $j=2,3,4$ *are defined in Section* 1.

LEMMA 2.2.

$$\begin{aligned}&\mathrm{E}\, h_j(Z)'h_k(Z)\exp(i\tilde{t}h_1(Z)'h_1(Z))\\ &\quad=(1-2i\tilde{t})^{-1/2}\,\mathrm{E}\, h_j(Z)'h_k(Z)\exp(i\tilde{t}Z'Z) \qquad j,k=2,3,4\end{aligned}$$

where $\tilde{t}$ *is a scalar and* $Z\sim N(0, I_p-\delta\delta')$.

The proofs of Lemmas 2.1 and 2.2 follow easily. For further details see van der Merwe [7].

3. The distribution of the Stein estimator

3.1. *On the characteristic function*

THEOREM 3.1. *The characteristic function of* $V^*=\sqrt{N}(\psi(\bar{X})-h_0(Z))$ *for N large enough is given by*:

$$\phi_{V^*}(it)=e^{-t't/2}\left\{1+\frac{1}{\sqrt{N}}\Phi_1(t)+\frac{1}{N}\Phi_2(t)+R\right\} \tag{3.1}$$

where $\Phi_1(t)=\dfrac{-ci}{a_0}(t'\Lambda)$,

$$\Phi_2(t)=\frac{c}{a_0}\left\{t'\left(I_p-\frac{ee'}{p}\right)t-\frac{(p+1)}{2a_0}(t'\Lambda)^2\right\} .$$

PROOF.

$$\phi_{V^*}(it)=\mathrm{E}\, e^{it'h_1(Z)}\left\{1+\frac{1}{\sqrt{N}}(it'h_2(Z))+\frac{1}{N}\left\{it'h_3(Z)+\frac{1}{2}(it'h_2(Z))^2\right\}+R\right\} .$$

By using Lemma 2.1 equation (3.1) follows.

3.2. *The distribution of* V^*

THEOREM 3.2. *The probability density function of* V^* *for* N *large enough is given by*

$$(3.2)\qquad f_{V^*}(v^*)=\left(\frac{1}{2\pi}\right)^{p/2}e^{-v^{*\prime}v^*/2}\left\{1+\frac{1}{\sqrt{N}}\tilde{g}_1(v^*)+\frac{1}{N}\tilde{g}_2(v^*)+R\right\}$$
$$-\infty<v^*<\infty$$

where $\tilde{g}_1(v^*)=-\dfrac{c}{a_0}v^{*\prime}\Lambda$,

$$g_2(v^*)=\frac{c}{a_0}\left\{\frac{1}{2}(p-3)-v^{*\prime}\left(I_p-\frac{ee'}{p}\right)v^*+\frac{(p+1)}{2a_0}(v^{*\prime}\Lambda)^2\right\} .$$

PROOF.

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

The proof follows by making use of the fact that

$$\mathrm{E}\,k^*(t)e^{-itv^*}=e^{-v^{*\prime}v^*/2}\,\mathrm{E}\,k^*(\tilde{Y}-iv^*) ,\qquad \text{where } \tilde{Y}\sim N(0, I_p)$$

($k^*(t)$ is a function in t).

Using (1.3) and taking expected values the mean and variance of $\psi(\bar{X})$ are given by

$$\mathrm{E}\,(\psi(\bar{X}))=\theta-\frac{c\Lambda}{Na_0}+\frac{c(p-3)}{N^2a_0^2}\Lambda+R$$

and

$$\mathrm{Var}\,(\psi(\bar{X}))=\frac{1}{N}I_p+\frac{2c}{N^2a_0^2}\left\{\frac{2}{a_0}\Lambda\Lambda'-\left(I_p-\frac{ee'}{p}\right)\right\}+R .$$

3.3. *The quadratic loss function*

The quadratic loss function is defined as $U=N(\psi(\bar{X})-\theta)'(\psi(\bar{X})-\theta)$.

THEOREM 3.3. *The probability density function of* U *for* N *large enough is given by*

$$(3.3)\qquad f_U(u)=\frac{e^{-u/2}u^{p/2-1}}{\Gamma(p/2)2^{p/2}}\left\{1+\frac{1}{N}\frac{c^2}{2a_0}\left(1-\frac{1}{p}u\right)+R\right\}\qquad u>0 .$$

PROOF.

$$f_U(u)=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\phi_U(i\tilde{t})\exp(-i\tilde{t}u)di\tilde{t}$$

where $$\phi_U(i\tilde{t})=(1-2i\tilde{t})^{-p/2}\left\{1+\frac{1}{N}\left(\frac{c^2}{a_0}(i\tilde{t})-\frac{2c(p-3)}{a_0}(i\tilde{t})(1-2i\tilde{t})^{-1}+\frac{2c^2}{a_0}(i\tilde{t})^2(1-2i\tilde{t})^{-1}\right)+R\right\}$$

is the characteristic function of U which is obtained by using Lemma 2.2. For further algebraic details see van der Merwe [7].

Also

$$\mathrm{E}\,(U)=p-\frac{(p-3)^2}{Na_0}+R$$

and

$$\mathrm{Var}\,(U)=2p-\frac{4(p-3)^2}{Na_0}-\frac{(p-3)^4}{N^2a_0^2}+R\,.$$

3.4. *The marginal distributions of the Stein estimator*

THEOREM 3.4. *The probability density function of* $V_i^*=\sqrt{N}(\psi(\bar{X}_i)-h_0(Z_i))$ *where* $\psi(\bar{X}_i)$ *is the ith component of the Stein estimator as defined in equation* (1.1), *and* $h_0(Z_i)=\theta_i$ *for large enough* N *is given by*

$$f_{V_i^*}(v_i^*)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}v_i^{*2}\right)\left\{1+\frac{1}{\sqrt{N}}\tilde{g}_1(v_i^*)+\frac{1}{N}\tilde{g}_2(v_i^*)+\frac{1}{N\sqrt{N}}\tilde{g}_3(v_i^*)+R\right\} \tag{3.4}$$

where $\tilde{g}_1(v_i^*)=\frac{-c}{a_0}v_i^*\Lambda_i\,,$

$$\tilde{g}_2(v_i^*)=\frac{c}{a_0}(1-v_i^{*2})\left\{\frac{1}{p}(p-1)-\frac{1}{2a_0}(p+1)\Lambda_i^2\right\}\,,$$

$$\tilde{g}_3(v_i^*)=\frac{c\Lambda_i}{a_0^2}\left\{v_i^*\left(-2p+\frac{\Lambda_i^2}{2a_0}(p+6p-3)\right)-v_i^{*3}\left(\frac{\Lambda_i^2}{6a_0}(p+6p-3)-(p-1)\right)\right\}$$

and $\Lambda_i=\theta_i-\bar{\theta}_p\,.$

PROOF. By putting $t'=(t_1, t_2,\cdots, t_i,\cdots, t_p)$ equal to $(0, 0,\cdots, t_i,\cdots, 0)$ in equation (3.1) the characteristic function of V_i^* is obtained. The density function follows in the same way as in Theorem 3.2.

Also

(3.5) $$\mathrm{E}\,(V_i^*) = \frac{-c}{\sqrt{N}\,a_0}\Lambda_i + \frac{1}{N\sqrt{N}}\,\frac{c\Lambda_i}{a_0^2}(p-3) + R$$

and

(3.6) $$\mathrm{E}\,(V_i^{*2}) = 1 - \frac{2c}{Na_0}\left\{\frac{p-1}{p} - \frac{1}{2a_0}(p+1)\Lambda_i^2\right\} + R\,.$$

The results given by van der Merwe and de Waal [9] can be obtained as special cases if p is taken as $p+1$, $1/p$ as zero and Λ as θ.

3.5. *The case* $\Sigma = \sigma^2 I_p$

If $\Sigma \neq I_p$, but equal to $\sigma^2 I_p$, then the Stein estimator as defined in Section 1 can be written as

(3.7) $$\tilde{\varphi}(\bar{X}) = \bar{\bar{X}}_p e + \left(1 - \frac{\tilde{c}S}{N(\bar{X} - \bar{\bar{X}}_p e)'(\bar{X} - \bar{\bar{X}}_p e)}\right)(\bar{X} - \bar{\bar{X}}_p e)$$

where $\tilde{c} = \dfrac{p-3}{N-p+2} = \dfrac{p-3}{N}\left\{1 + \dfrac{p-2}{N} + \dfrac{(p-2)^2}{N^2} + \cdots\right\}$,

$$S = \sum_{i=1}^{p}\sum_{j=1}^{N}(X_{ji} - \bar{X}_i)^2 \sim \sigma^2\chi_n^2 \qquad \text{with } n = (N-1)p$$

and independent of $\bar{X}$. Also

(3.8) $$\mathrm{E}\,(S) = \sigma^2 n \qquad \text{and} \qquad \mathrm{E}\,(S^2) = \sigma^4 n(n+2)\,.$$

By making use of the same methods as in the previous sections the asymptotic distribution, mean and variance of this estimator as well as that of the quadratic loss function can be obtained.

4. Moments of the Stein estimator

4.1. *Exact moments*

Let $W = \dfrac{N}{\sigma^2}\sum_{i=1}^{p}(\bar{X}_i - \bar{\bar{X}}_p)^2$ where $\bar{X}_i \sim N\left(\theta_i, \dfrac{\sigma^2}{N}\right)$, $i = 1, \cdots, p$; and by using the results from Ullah [6] it follows that

(4.1) $$\mathrm{E}\,(\bar{X}_i - \theta_i)W^{-r} = \frac{\sigma^2}{N}\frac{\partial}{\partial\theta_i}\mathrm{E}\,W^{-r}\,,$$

(4.2) $$\mathrm{E}\,(\bar{X}_i - \theta_i)^2 W^{-r} = \left(\frac{\sigma^4}{N^2}\frac{\partial^2}{\partial\theta_i^2} + \frac{\sigma^2}{N}\right)\mathrm{E}\,W^{-r}\,.$$

To evaluate the expectations involved in (4.1) and (4.2) $\mathrm{E}\,W^{-r}$ $(r = 1, 2)$ and its partial derivatives with respect to θ_i are required.

Since $W \sim \chi^2_{p-1}(\lambda)$ where $\lambda = \frac{N}{2\sigma^2} \sum_{i=1}^{p} (\theta_i - \bar{\theta}_p)^2$,

$$\text{(4.3)} \qquad \mathrm{E}\,(W^{-r}) = 2^{-r} \frac{\Gamma((p-1)/2-r)}{\Gamma((p-1)/2)} e^{-\lambda} {}_1F_1\left(\frac{p-1}{2}-r; \frac{p-1}{2}; \lambda\right)$$

where ${}_1F_1(a; c; x) = \sum_{l=0}^{\infty} \frac{(a)_l x^l}{(c)_l}$ and $(a)_l = a(a+1)\cdots(a+l-1)$; and $(a)_0 = 1$. Also

$$\text{(4.4)} \qquad \frac{d^s}{d\lambda^s} \mathrm{E}\, W^{-r} = 2^{-r}(-1)^s \frac{\Gamma(r+s)\Gamma((p-1)/2-r)}{\Gamma(r)\Gamma((p-1)/2+s)} \times e^{-\lambda} {}_1F_1\left(\frac{p-1}{2}-r; \frac{p-1}{2}+s; \lambda\right).$$

The first two derivatives of $\mathrm{E}\, W^{-r}$ with respect to θ_i can easily be obtained and are as follows:

$$\text{(4.5)} \qquad \frac{\partial}{\partial \theta_i} \mathrm{E}\, W^{-r} = \frac{N}{\sigma^2}(\theta_i - \bar{\theta}_p)\left(\frac{d}{d\lambda} \mathrm{E}\, W^{-r}\right),$$

$$\text{(4.6)} \qquad \frac{\partial^2}{\partial \theta_i^2} \mathrm{E}\, W^{-r} = \frac{N^2}{\sigma^4}(\theta_i - \bar{\theta}_p)^2\left(\frac{d^2}{d\lambda^2} \mathrm{E}\, W^{-r}\right) + \frac{N}{\sigma^2}\left(1-\frac{1}{p}\right)\left(\frac{d}{d\lambda} \mathrm{E}\, W^{-r}\right).$$

By making use of the fact that $\frac{\partial \lambda}{\partial \theta_i} = \frac{N}{\sigma^2}(\theta_i - \bar{\theta}_p)$, $\frac{\partial^2 \lambda}{\partial \theta_i^2} = \frac{N}{\sigma^2}\left(1-\frac{1}{p}\right)$ (4.5) and (4.6) follow.

Defining

$$\text{(4.7)} \qquad f_{\mu,v} = \frac{\Gamma((p-1)/2+\mu)}{\Gamma((p-1)/2+v)} e^{-\lambda} {}_1F_1\left(\frac{p-2}{2}+\mu; \frac{p-1}{2}+v; \lambda\right), \qquad v-\mu > 0$$

The following theorem can now be stated:

THEOREM 4.1. *The first two moments of* $\tilde{\phi}(\bar{X}_i) - \theta_i$ *(where* $\tilde{\phi}(\bar{X}_i)$ *is the ith component of the Stein estimator as defined in equation* (3.7)) *are given by*

$$\text{(4.8)} \qquad \mathrm{E}\,(\tilde{\phi}(\bar{X}_i) - \theta_i) = -\frac{1}{2}\tilde{c} n \Lambda_i f_{0,1},$$

$$\text{(4.9)} \qquad \mathrm{E}\,(\tilde{\phi}(\bar{X}_i) - \theta_i)^2 = \frac{\sigma^2}{N} + \tilde{c} n\left\{\Lambda_i^2 f_{0,2} - \frac{\sigma^2}{N}\left(1-\frac{1}{p}\right) f_{0,1}\right\} + \frac{1}{4}\tilde{c}^2 n(n+2) \times \left\{\Lambda_i^2 f_{0,2} + \left(1-\frac{1}{p}\right)\frac{\sigma^2}{N} f_{-1,1}\right\}.$$

PROOF. By using the results given in (3.8)-(4.7) equations (4.8) and (4.9) follow.

$E(\tilde{\psi}(\bar{X}_i-\theta_i)^3)$ and $E(\tilde{\psi}(\bar{X}_i-\theta_i)^4)$ were also calculated but for lack of space are not given here. The interested reader is referred to van der Merwe [8].

Ullah [6] actually derived the moments for the Stein estimator if θ_i, are all near the origin. If we put $1/p$, p, Λ_i, n and σ^2/N equal to zero, $K+1$, β_i, $T-K$ and σ^2 respectively in equations (4.8) and (4.9), Ullah's results follow.

4.2. *Asymptotic moments*

Copson [1] showed that for large values of λ; $f_{\mu,v}$ can be approximated by

$$f_{\mu,v}\sim\frac{1}{\lambda^{v-\mu}}\sum_{l=0}^{\infty}\frac{(v-\mu)_l(1-(p-1)/2-\mu)_l}{l!\lambda^l}\ .$$

By using this approximation for $f_{0,1}$, $f_{0,2}$, $f_{-1,1}$ and taking $\sigma^2=1$; i.e. $\bar{c}$ as c the first two moments of the Stein estimator about θ_i are now given by

$$E(\psi(\bar{X}_i)-\theta_i)=\frac{-c}{Na_0}\Lambda_i+\frac{c(p-3)}{N^2a_0^2}\Lambda_i+R$$

and

$$E(\psi(X_i)-\theta_i)^2=\frac{1}{N}-\frac{2c}{N^2a_0}\left\{\frac{p-1}{p}-\frac{1}{2a_0}(p+1)\Lambda_i^2\right\}+R\ ,$$

which correspond to the asymptotic moments obtained in Subsection 3.4 (equations (3.5) and (3.6)). Higher order moments will also correspond which is a further indication that the asymptotic distributions will approximate the exact distributions quite well if N and a_0 are large enough and that term by term integration as used in Theorems 3.1–3.4 are permissible.

Acknowledgement

The author wishes to express his gratitude to the referee for his helpful comments in revising the paper.

UNIVERSITY OF THE ORANGE FREE STATE

REFERENCES

[1] Copson, E. T. (1948). *An Introduction to the Theory of Functions of a Complex Variable*, Oxford University Press, London.

[2] Efron, B. and Morris, C. (1972). *Simultaneous Estimation of Parameters*, The Rand Corporation, 4835.

[3] Efron, B. and Morris, C. (1973). Stein's estimation rule and its competitors—an empirical Bayes approach, *J. Amer. Statist. Ass.*, **68**, 117-130.

[4] Nagao, H. (1972). Non-null distributions of the likelihood ratio criteria for independence and equality of mean vectors and covariance matrices, *Ann. Inst. Statist. Math.*, **24**, 67-79.

[5] Stein, C. (1962). Confidence sets for the mean of a multivariate normal distribution, *J. R. Statist. Soc.*, B, **24**, 265-296.

[6] Ullah, A. (1974). On the sampling distribution of improved estimators for coefficients in linear regression, *J. Econometrics*, **2**, 143-150.

[7] van der Merwe, A. J. (1978). The asymptotic expansion of the Stein estimator for the vector case when θ_i, $i=1,\cdots,p$ are nearly equal and the distribution of the quadratic loss function if $\theta_i=0$; $i=1,\cdots,p$, *Technical Report* No. 34, Department of Math. Statistics, University of the OFS.

[8] van der Merwe, A. J. (1980). The asymptotic expansion as well as the exact moments of the Stein estimator when the population means are nearly equal, *Technical Report* No. 55, Department of Math. Statistics, University of the OFS.

[9] van der Merwe, A. J. and de Waal, D. J. (1978). The asymptotic expansion of the Stein estimators for the vector case, *Ann. Inst. Statist. Math.*, **30**, A, 385-395.