The Annals of Statistics
1981, Vol. 9, No. 6, 1334-1338

A NECESSARY AND SUFFICIENT CONDITION FOR SECOND ORDER ADMISSIBILITY WITH APPLICATIONS TO BERKSON'S BIOASSAY PROBLEM

By J. K. Ghosh and Bimal K. Sinha[1]

Indian Statistical Institute and University of Pittsburgh

A theorem is proved which gives a necessary and sufficient condition for improving, up to $o(n^{-2})$, the mean squared error of the maximum likelihood estimate $\hat{\theta}$ by using an estimate of the form $\hat{\theta} + d(\hat{\theta})/n$. An application is made to a bioassay problem of Berkson.

1. Introduction. Recently in the multiparameter setup, Amemiya (1980) has confirmed Berkson's observation (Berkson, 1955, 1980; Berkson and Hodges, 1961) that the minimum logit χ^2 estimate is often a better point estimate than the maximum likelihood estimate (mle) in the sense of mean squared error. This has been done by computing numerically the Taylor expansion of the mean squared errors up to $o(n^{-2})$ for many parameter values. However, as observed by Amemiya, it is an open question whether this result holds for all values of the parameters.

To attack this problem it is convenient to start with the one-parameter case and work with the Rao-Blackwellised version $\tilde{\theta}$ of Berkson's estimate (Amemiya, 1980, equation (74)). Since $\tilde{\theta}$ is of the form $\hat{\theta} + d(\hat{\theta})/n$, $\hat{\theta}$ being the mle and $d(\theta)$ a continuously differentiable function, the comparison of mean squared errors up to $o(n^{-2})$ is relatively easy. Moreover, as noted in Amemiya (1980) and Ghosh and Subramanyam (1974), $\tilde{\theta}$ has a smaller mean squared error (up to $o(n^{-2})$) than Berkson's estimate. This special case suggests the following more general problem for one-parameter families satisfying the usual Cramér-Rao regularity conditions. Does there exist a function $d(\theta)$ such that

$$(1.1) \qquad E_\theta\left\{\hat{\theta} + \frac{d(\hat{\theta})}{n} - \theta\right\}^2 \le E_\theta(\hat{\theta} - \theta)^2 \quad \text{up to } o(n^{-2}) \quad \text{for all } \theta,$$

with strict inequality for at least one θ? In (1.1), LHS $= a_1(\theta)n^{-1} + a_2(\theta)n^{-2} + o(n^{-2})$, $a_2(\theta) > 0$ and RHS $= a_1(\theta)n^{-1} + a_3(\theta)n^{-2} + o(n^{-2})$, $a_3(\theta) > 0$, and (1.1) requires $a_2(\theta) \le a_3(\theta)$ for all θ.

If the answer to (1.1) is "no", we shall say $\hat{\theta}$ is admissible up to $o(n^{-2})$. In (1.1) as well as in subsequent calculations, we shall use the method of Taylor expansion adopted in Amemiya (1980) and Ghosh and Subramanyam (1974). For a discussion of the nature of approximation involved, see Amemiya's Section 3. Henceforth, admissibility will mean admissibility for mean squared error loss up to $o(n^{-2})$.

In the same vein as (1.1) one may want to study the admissibility of $\tilde{\theta}$ in place of $\hat{\theta}$ and, more generally, that of $\hat{\theta} + c(\hat{\theta})/n$ where c is any given function in the class of estimates of the form $\hat{\theta} + d(\hat{\theta})/n$; c and d will always be assumed to be continuously differentiable. The restriction to estimates of the form $\hat{\theta} + d(\hat{\theta})/n$ can be justified by invoking the second order efficiency of the mle; see Efron (1975), Ghosh and Subramanyam (1974), Ghosh *et al* (1980), Pfanzagl and Wefelmeyer (1978).

Let $b(\theta)/n$ be the bias of $\hat{\theta} + c(\hat{\theta})/n$ up to $o(n^{-1})$. Then $b(\theta) = b_0(\theta) + c(\theta)$ where $b_0(\theta)/n$ is the bias of the mle up to $o(n^{-1})$. Let $I(\theta)$ be the Fisher information per observation. Assume $b_0(\theta)$ and $I(\theta)$ are continuous and $I(\theta) > 0$ for all θ.

The main result on admissibility can now be stated. The proof is given in Section 2.

Received February, 1980; revised March, 1981.

[1] On leave from the Indian Statistical Institute.

Key words and phrases. Maximum likelihood estimate, second order admissibility.

AMS 1970 subject classification. Primary 62F12; secondary 62C15.

THEOREM. (i) *$\hat\theta + c(\hat\theta)/n$ is admissible iff for some* $-\infty < \theta_0 < \infty$

$$\int_{\theta_0}^{\infty} I(\theta)\exp\left\{-\int_{\theta_0}^{\theta} b(u)I(u)\,du\right\} d\theta = \infty$$

and

$$\int_{-\infty}^{\theta_0} I(\theta)\exp\left\{\int_{\theta}^{\theta_0} b(u)I(u)\,du\right\} d\theta = \infty.$$

(ii) *If $\hat\theta + c(\hat\theta)/n$ is inadmissible then one can choose $d(\cdot)$ so that $\hat\theta + d(\hat\theta)/n$ is admissible and better than $\hat\theta + c(\hat\theta)/n$ up to $o(n^{-2})$.*

The condition (1.2) suggests that for admissibility the bias term $b(\theta)$ should be negative as $\theta \to \infty$ and positive as $\theta \to -\infty$, i.e., the estimate should behave like a shrinker at least as far as bias is concerned. See in this connection the better estimates constructed via (2.7) and (2.8) in the proof of the theorem.

To give an application of the theorem we shall consider Berkson's problem of estimating θ on the basis of observations on independent random variables R_i having the binomial distribution $B(n, \pi_i)$, $i = 1, \cdots, k$, with

$$\pi_i = \{1 + \exp(-\theta - \beta d_i)\}^{-1}, \tag{1.3}$$

where β is known; R_i represents the number of responses among n trials at dose level d_i. It follows from the theorem that the mle $\hat\theta$ is always inadmissible and that the Rao-Blackwellised version $\bar\theta$ of Berkson's estimate is admissible for $k \geq 4$ and inadmissible for $k < 4$. These results are discussed in detail in Section 3.

Returning to the special problem of comparing $\hat\theta$ with $\tilde\theta$, we prove directly that in general neither dominates the other. This theoretical result along with the Berkson-Amemiya calculations suggest that the mle can be improved by Berkson's estimate within a bounded interval (θ_1, θ_2) but generally not beyond. That this phenomenon holds quite generally is the content of the corollary in Section 2.

2. Main Results. It is easy to check by straightforward Taylor expansion that

$$E_\theta\left\{\hat\theta + \frac{d(\hat\theta)}{n} - \theta\right\}^2 - E_\theta\left\{\hat\theta + \frac{c(\hat\theta)}{n} - \theta\right\}^2 = \{g^2(\theta) + 2g(\theta)b(\theta) + 2g'(\theta)/I(\theta)\}/n^2 + o(n^{-2}), \tag{2.1}$$

where $g(\theta) = d(\theta) - c(\theta)$. Hence, the condition that $\hat\theta + d(\hat\theta)/n$ dominates $\hat\theta + c(\hat\theta)/n$ is given by

$$g^2(\theta) + 2g(\theta)b(\theta) + 2g'(\theta)/I(\theta) \leq 0 \qquad \text{for all } \theta, \tag{2.2}$$

with strict inequality for at least one θ.

PROOF OF THE THEOREM. (i) We first assume (1.2) and show that this implies that (2.2) has the trivial solution $g(\theta) \equiv 0$. Toward this end, let

$$q(\theta) = I(\theta)\exp\left\{\int_{\theta_0}^{\theta} b(u)I(u)\,du\right\}. \tag{2.3}$$

Note that (2.2) implies, for $-\infty < \alpha < \beta < \infty$,

$$\int_\alpha^\beta g^2(\theta)q(\theta)\,d\theta \leq 2\left\{\int_\alpha^\beta -g(\theta)b(\theta)q(\theta)\,d\theta - \int_\alpha^\beta \frac{g'(\theta)}{I(\theta)}\,q(\theta)\,d\theta\right\}. \tag{2.4}$$

which, on integrating the second integral by parts, is

$$\le 2\sqrt{g^2(\beta)q(\beta)}\cdot\sqrt{q(\beta)}/I(\beta) + 2\sqrt{g^2(\alpha)q(\alpha)}\cdot\sqrt{q(\alpha)}/I(\alpha).$$

Proceeding as in Karlin (1958), it can be shown that (2.4) implies $g(\theta) \equiv 0$ whenever

$$\int_{\theta_0}^{\infty} I^2(\theta)/q(\theta)\,d\theta = \infty = \int_{-\infty}^{\theta_0} I^2(\theta)/q(\theta)\,d\theta.$$

Since (1.2) is equivalent to the above condition, our proof that $g \equiv 0$ is complete.

Conversely we shall show that if (1.2) is violated, then (2.2) admits a non-trivial solution $g(\theta)$. Assume first that the second condition in (1.2) is violated. Let $\psi(\theta)$ be a negative function which is continuous and integrable over $(-\infty, x)\ \forall\ -\infty < x < \infty$. Clearly a solution of

$$1 + \frac{2b(\theta)}{g(\theta)} + \frac{2g'(\theta)}{g^2(\theta)I(\theta)} = \frac{\psi(\theta)q(\theta)}{I^2(\theta)} \tag{2.5}$$

will solve (2.2), provided $g(\theta) \ne 0$ on $(-\infty, \infty)$. Writing $h(\theta) = 1/g(\theta)$, we get

$$h(\theta)b(\theta)I(\theta) - h'(\theta) = \frac{1}{2}\left\{\frac{\psi(\theta)q(\theta)}{I(\theta)} - I(\theta)\right\}. \tag{2.6}$$

It is well-known and easily checked that a solution of (2.6) is

$$h(\theta) = \frac{q(\theta)}{I(\theta)}\left[K + \frac{1}{2}\int_{-\infty}^{\theta}\left\{\frac{I^2(u)}{q(u)} - \psi(u)\right\}du\right], \qquad K \ge 0. \tag{2.7}$$

Thus we have found a non-trivial solution of (2.2).

In case the first condition in (1.2) is violated, define

$$h(\theta) = -\frac{q(\theta)}{I(\theta)}\left[K + \frac{1}{2}\int_{\theta}^{\infty}\left\{\frac{I^2(u)}{q(u)} - \psi(u)\right\}du\right], \qquad K \ge 0. \tag{2.8}$$

Then it follows that $g(\theta) = 1/h(\theta)$ satisfies (2.2), which proves part (i) of the theorem.

(ii) We now show that in case (1.2) is violated ψ can be chosen so that the new improved estimate satisfies (1.2) and hence is admissible.

Consider first the case where only one of the two conditions in (1.2), say the second, is violated. For some $-\infty < \theta_1 < \theta_0 < \theta_2 < \infty$, let

$$\begin{aligned} \psi(\theta) &= -\gamma_1 I^2(\theta)/q(\theta), \qquad -\infty < \theta < \theta_1 \\ &= -\gamma_2 I^2(\theta)/q(\theta), \qquad \theta_2 < \theta < \infty \end{aligned} \tag{2.9}$$

where $0 < \gamma_1 < 1 < \gamma_2$. Take $K = 0$ in (2.7) which defines h; note that $d = h^{-1} + c$. Let

$$q^* = I(\theta)\exp\left\{\int_{\theta_0}^{\theta}(d + b_0)I\,du\right\}. \tag{2.10}$$

Then we have to check (1.2) with q^* in place of q. Clearly, by (2.7) and (2.9),

$$\begin{aligned} q^* &= \text{const.}\ q\{\eta(\theta)\}^{2/(1+\gamma_1)}, \qquad -\infty < \theta < \theta_1, \\ &< \text{const.}\ q\{\text{const.} + \eta(\theta)\}^{2/(1+\gamma_2)}, \qquad \theta_2 < \theta < \infty, \end{aligned} \tag{2.11}$$

where $\eta(\theta) = \int_{-\infty}^{\theta} I^2(u)/q(u)\,du$.

Hence, writing $\lambda_i = 1 - 2/(1 + \gamma_i)$ for $i = 1, 2$,

$$\int_{\alpha}^{\beta}\frac{I^2(\theta)}{q^*(\theta)}\,d\theta \ge \text{const.}\frac{\{\eta(\theta)\}^{\lambda_1}}{\lambda_1}\bigg|_{\alpha}^{\theta_1} + \text{const.}\frac{\{\text{const.} + \eta(\theta)\}^{\lambda_2}}{\lambda_2}\bigg|_{\theta_2}^{\beta} \to \infty \tag{2.12}$$

as $\beta \to \infty$ or $\alpha \to -\infty$. This disposes of the case where only the second condition in (1.2) is violated.

Analogously, if only the first condition in (1.2) is violated, defining $\psi(\theta)$ as in (2.9) with $0 < \gamma_2 < 1 < \gamma_1$ and h as in (2.8) with $K = 0$, admissibility of the better estimate is guaranteed.

Finally, suppose that both conditions in (1.2) fail. Let $\gamma > 0$, $\psi = -\gamma I^2/q$, $K = 0$, and define g_1^{-1} and g_2^{-1} by the RHS's of (2.7) and (2.8). Then g_1 and g_2 are, respectively, positive and negative solutions of (2.2). It follows that $g = g_1 + g_2$ is a solution of (2.2) also. That $\hat{\theta} + \{c(\hat{\theta}) + g(\hat{\theta})\}/n$ is admissible follows from straightforward calculations similar to (2.11) and (2.12). This completes the proof.

REMARK. If both conditions are violated and we use the g defined in the previous paragraph then the improvement achieved, i.e. the LHS of (2.2), can be written down explicitly in the form

$$A(\theta) = \frac{-4\gamma}{(1+\gamma)^2}\left(\left[(I/q)\left\{\int_{-\infty}^{\theta}\frac{I^2(u)}{q(u)}\,du\right\}^{-1}\right]^2 + \left[(I/q)\left\{\int_{\theta}^{\infty}\frac{I^2(u)}{q(u)}\,du\right\}^{-1}\right]^2 + \frac{2}{\gamma}(I/q)^2\int_{-\infty}^{\theta}\frac{I^2(u)}{q(u)}\,du\int_{\theta}^{\infty}\frac{I^2(u)}{q(u)}\,du\right).$$

To justify this, use the fact that g_1, g_2 satisfy (2.5).

COROLLARY. *Given any two numbers $a < b$, one can improve $\hat{\theta} + c(\hat{\theta})/n$ by an estimate of the form $\hat{\theta} + d(\hat{\theta})/n$ for $a \le \theta \le b$.*

To prove this, one may take ψ to be a negative constant K', $K = 0$ and use (2.7) replacing $-\infty$ by a or (2.8) replacing ∞ by b. In either case we get a better estimate than $\hat{\theta} + c(\hat{\theta})/n$. The improvement can be shown to be substantial if $|K'|$ is large. Many other choices of ψ are possible.

3. Berkson's example. Consider the bioassay model (1.3). As before let $\tilde{\theta}$ denote the Rao-Blackwellised minimum logit χ^2 estimate and denote its bias and that of $\hat{\theta}$ by $b(\theta)/n$ and $b_0(\theta)/n$ up to $o(n^{-1})$. Then, by (1.3), see (Berkson, 1955, Section 3),

$$I = \Sigma\pi_i(1-\pi_i),\quad b_0 = \Sigma\pi_i(1-\pi_i)(2\pi_i - 1)/2I^2,$$

$$b = \Sigma\pi_i(1-\pi_i)(2\pi_i - 1)/I^2 - \Sigma(2\pi_i - 1)/2I.$$

Since $I \sim \text{const.} \exp(-|\theta|)$ as $\theta \to \pm\infty$, $b_0 I \to \pm\frac{1}{2}$ as $\theta \to \pm\infty$, it follows that (1.2) is violated at both ends. Hence from the theorem, $\hat{\theta}$ is inadmissible. Similarly, since $bI \to \mp(k-2)/2$, as $\theta \to \pm\infty$, the theorem implies that $\tilde{\theta}$ is inadmissible (admissible) if $k < 4 (k > 4)$. An analysis of the same sort shows that if $k = 4$, then θ is admissible.

To return to the Berkson-Amemiya problem of comparing $\tilde{\theta}$ and $\hat{\theta}$ directly, note that

$$\{b^2(\theta) + 2b'(\theta)/I(\theta)\}I^4(\theta) = \{\Sigma(\pi_i - 1)\}^2\cdot(6 + k^2/2 - 4k)/2\{1 + o(1)\} \quad \text{as} \quad \theta \to \infty$$

$$= \{\Sigma\pi_i\}^2\cdot(6 + k^2/2 - 4k)/2\{1 + o(1)\} \quad \text{as} \quad \theta \to -\infty$$

and

$$\{b_0^2(\theta) + 2b_0'(\theta)/I(\theta)\}I^4(\theta) = \{\Sigma(\pi_i - 1)\}^2\cdot(5/4)\{1 + o(1)\} \quad \text{as} \quad \theta \to \infty$$

$$= \{\Sigma\pi_i\}^2\cdot(5/4)\{1 + o(1)\} \quad \text{as} \quad \theta \to -\infty.$$

This shows that if $k \ge 8$, $\exists(\theta_1, \theta_2)$ such that for $\theta \notin (\theta_1, \theta_2)$ $\hat{\theta}$ is better than $\tilde{\theta}$; while for $k \le 7$, $\exists(\theta_1', \theta_2')$ such that for $\theta \notin (\theta_1', \theta_2')$ $\tilde{\theta}$ is better than $\hat{\theta}$.

REFERENCES

AMEMIYA, T. (1980). The n^{-2}-order mean squared errors of the maximum likelihood and the minimum chi-square estimator. *Ann. Statist.* **8** 488–503.

BERKSON, J. (1955). Maximum likelihood and minimum χ^2 estimates of the logistic function. *J. Amer. Statist. Assoc.* **50** 130–162.

BERKSON, J. (1980). Minimum chi-square, not maximum likelihood! *Ann. Statist.* **8** 457–487.

BERKSON, J. and HODGES, J. L. JR. (1961). A minimax estimator for the logistic function. *Proceedings of the fourth Berkeley Symposium on Mathematical Statistics and Probability* **4** 77–86.

EFRON, BRADLEY (1975). Defining the curvature of a statistical problem (with applications to second order efficiency). *Ann. Statist.* **3** 1189–1242.

GHOSH, J. K., SINHA, B. K. and WIEAND, H. S. (1980). Second order efficiency of the maximum likelihood estimate with respect to any bounded bowl-shaped loss function. *Ann. Statist.* **8** 506–521.

GHOSH, J. K. AND SUBRAMANYAM, K. (1974). Second order efficiency of maximum likelihood estimators. *Sankhyā*, Ser. A, **36** 325–358.

KARLIN, S. (1958). Admissibility for estimation with quadratic loss. *Ann. Math. Statist.* **29** 406–436.

PFANZAGL, J. and WEFELMEYER, W. (1978). A third order optimum property of the maximum likelihood estimator. *J. Multivariate Anal.* **8** 1–29.

STAT-MATH DIVISION
INDIAN STATISTICAL INSTITUTE
203 B.T. ROAD
CALCUTTA-700035
INDIA

DEPARTMENT OF MATHEMATICS AND STATISTICS
FACULTY OF ARTS AND SCIENCES
UNIVERSITY OF PITTSBURGH
PITTSBURGH, PENNSYLVANIA 15260