International Journal of Applied Mathematics &
Statistical Sciences (IJAMSS)
ISSN(P): 2319-3972; ISSN(E): 2319-3980
Vol. 6, Issue 1, Dec - Jan 2017; 81-88

International Academy of Science, Engineering and Technology
Connecting Researchers; Nurturing Innovations

OBTAINING SUPER-EFFICIENT ESTIMATORS FOR A REAL-VALUED PARAMETER: DELTA METHOD

K. SIVASAKTHI[1], T. M. DURAIRAJAN[2] & MARTIN L. WILLIAM[3]
[1]Department of Statistics, Tagore College of Arts and Science, Chennai, India
[2,3]Department of Statistics, Loyola College, Chennai, India

ABSTRACT

This paper addresses the problem of finding super-efficient estimators for single parameter families. Initially, the original approach of Hodges (1951) is presented. A procedure for obtaining super-efficient estimators with respect to 'Fisher Information', using the 'Delta Method' of asymptotic inference theory is derived and illustrated with examples.

MATHEMATICS SUBJECT CLASSIFICATION: 62F10, 62F12

KEYWORDS: Fisher Information, Maximum Likelihood Estimators, Super-Efficiency

1. INTRODUCTION

Let $X_1, X_2, \cdots, X_n$ be i.i.d. k-dimensional vector-valued random variables, each with p.d.f. $p(x\,|\,\theta),\ x \in S \subset R^k, \theta \in \Omega \subset R$. Let $l_\theta = \dfrac{\partial \log p(x\,|\,\theta)}{\partial \theta}$ be the score function and let $I(\theta)$ =E(l_θ^2) be the Fisher Information (FI).

Denoting $X_{(n)}$ = ($X_1, X_2, \cdots, X_n$), let $T_n(X_{(n)})$ = T (say) be any unbiased estimator of θ, V_θ(T) be the variance of T and C be the class of unbiased estimators of θ. Then, by Cramer-Rao Inequality,

$$V_\theta(T) \geq \frac{1}{n\, I(\theta)}\ ,\ \forall T \in C \tag{1.1}$$

Let $\hat{\theta}_n(X_{(n)}) = \hat{\theta}_n$ (say) be any estimator (not necessarily unbiased) of θ. It could be MLE of θ or any other. In view of the inequality (1.1), $\hat{\theta}_n$ is said to be 'Asymptotically Efficient' (AE) if

$$\sqrt{n}\,(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, I^{-1}(\theta))\,,\ \theta \in \Omega$$

An estimator $\breve{\theta}_n$ of θ is said to be 'Super Efficient at θ_0 ' if

$$\sqrt{n}\,(\breve{\theta}_n - \theta) \xrightarrow{d} N(0, \sigma^2(\theta))\,,\ \theta \in \Omega$$

where $I^{-1}(\theta) \geq \sigma^2(\theta)\ \forall\, \theta \in \Omega$ and $I^{-1}(\theta_0) > \sigma^2(\theta_0)$ for some $\theta_0 \in \Omega$.

Hodges (1951), in an unpublished work, was the first to propose super efficient estimator. For detailed discussion on Hodge's super efficiency, reference is made to Bahadur (1983). Investigations were carried out and interesting results on the principle of super-efficiency were established by Le Cam (1953, 1956, 1960, 1972). Basu (1952) considered the non-regular Uniform family U(θ, 2θ) and established the super-efficiency of the Best Linear Unbiased Estimator over the unbiased estimator based on the MLE of θ. Stein (1956) discussed super efficient estimator of a multivariate normal mean vector with the squared-norm loss function. Sethuraman (2004) carried out interesting investigations on the efficiency of tests based on super-efficient estimators comparing it with tests based on the MLEs. Durairajan (2012) presented the notion of Sub-Score and construction of Super-efficient estimator for estimating a vector parameter of interest in the presence of nuisance parameters.

This paper is organized as follows: Section 2 discusses the idea of Hodges (1951) that led to the notion of 'super-efficiency'. In Section 3, the 'delta method' of asymptotic inference theory is invoked to obtain super-efficient estimators for single parameter families of distributions. Section 4 presents a few applications of the method to obtain super-efficient estimators in different contexts.

2. THE IDEA OF HODGES (1951)

Let $\hat{\theta}_n$ be the maximum likelihood estimator of θ such that $\sqrt{n}\,(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, I^{-1}(\theta))$, $\theta \in \Omega$. Let θ_o be an arbitrary value of θ. Consider the estimator

$$\tilde{\theta}_n = \begin{cases} \theta_0 & if \;\; n^{1/4}|\hat{\theta}_n - \theta_0| \le 1 \\ \hat{\theta}_n & if \;\; n^{1/4}|\hat{\theta}_n - \theta_0| > 1 \end{cases} \tag{2.1}$$

Clearly, $[\hat{\theta}_n \ne \tilde{\theta}_n] \Rightarrow [\tilde{\theta}_n = \theta_0]$. Note that this is a one-way implication only.

$$\text{So, } P_\theta[\hat{\theta}_n \ne \tilde{\theta}_n] \le P_\theta[\tilde{\theta}_n = \theta_0] = P_\theta[\, n^{1/4}|\hat{\theta}_n - \theta_0| \le 1] \tag{2.2}$$

Consider $|\theta - \theta_0| \le |\hat{\theta}_n - \theta| + |\hat{\theta}_n - \theta_0|$ from which we get

$$|\hat{\theta}_n - \theta| \ge |\theta - \theta_0| - |\hat{\theta}_n - \theta_0|$$

Thus, $[n^{1/4}|\hat{\theta}_n - \theta_0| \le 1] \Rightarrow [\,|\hat{\theta}_n - \theta| \ge |\theta - \theta_0| - n^{-1/4}]$ so that

$$P_\theta[n^{1/4}|\hat{\theta}_n - \theta_0| \le 1] \le P_\theta[\,|\hat{\theta}_n - \theta| \ge |\theta - \theta_0| - n^{-1/4}] \tag{2.3}$$

Using (2.3) in (2.2), we get

$$P_\theta[\hat{\theta}_n \ne \tilde{\theta}_n] \le P_\theta[\sqrt{n}|\hat{\theta}_n - \theta| \ge \sqrt{n}\,|\theta - \theta_0| - n^{1/4}] \tag{2.4}$$

We now investigate the asymptotic behaviour of $\tilde{\theta}_n$.

For $\theta \ne \theta_0$, $\sqrt{n}|\theta - \theta_0| - n^{1/4} \to \infty$ so, $P_\theta[\sqrt{n}|\hat{\theta}_n - \theta| \ge \sqrt{n}|\theta - \theta_0| - n^{1/4}] \to 0$ as $n \to \infty$, which in turn gives $P_\theta[\hat{\theta}_n \neq \tilde{\theta}_n] \to 0$. That is, $\hat{\theta}_n$ and $\tilde{\theta}_n$ are asymptotically equivalent and have the same asymptotic behaviour. Hence, we infer that

$$\sqrt{n}\,(\tilde{\theta}_n - \theta) \xrightarrow{d} N(0, I^{-1}(\theta)) \text{ for } \theta \neq \theta_0 \tag{2.5}$$

And, for $\theta = \theta_0$, by the definition of $\tilde{\theta}_n$, $P_{\theta_0}[\tilde{\theta}_n = \theta_0] = P_{\theta_0}[n^{1/4}|\hat{\theta}_n - \theta_0| \le 1]$

That is, $P_{\theta_0}[\sqrt{n}\,(\tilde{\theta}_n - \theta_0) = 0] = P_{\theta_0}[\sqrt{n}|\hat{\theta}_n - \theta_0| \le n^{1/4}]$

$$= P_{\theta_0}[-n^{1/4} \le \sqrt{n}|\hat{\theta}_n - \theta_0| \le n^{1/4}]$$

$$\to 1 \text{ as } n \to \infty$$

Hence, we infer that, $\sqrt{n}\,(\tilde{\theta}_n - \theta_0) \xrightarrow{d} N(0,0)$ for $\theta = \theta_0$ (2.6)

From (2.5) and (2.6), it is found that asymptotic variance of $\tilde{\theta}_n$ equals that of $\hat{\theta}_n$ for $\theta \neq \theta_0$ and is less (equal to 0) for $\theta = \theta_0$. Hence, $\tilde{\theta}_n$ defined in (2.1) is 'super efficient' at θ_0.

3. DELTA-METHOD TO GET SUPER-EFFICIENT ESTIMATORS

Let T_n be a statistic that is asymptotically normal about parameter θ. That is, suppose we have $\sqrt{n}\,(T_n - \theta) \xrightarrow{d} N(0, \sigma^2(\theta))$. Consider a statistic $g(T_n)$ where 'g' is a function that is at least twice differentiable at θ. Expanding the function g(t) in a neighbourhood of θ, to the first degree of approximation, we have

$$g(T_n) \approx g(\theta) + (T_n - \theta)\, g'(\theta)$$

so that $\sqrt{n}\,[g(T_n) - g(\theta)] \approx \sqrt{n}\,(T_n - \theta)\, g'(\theta) \xrightarrow{d} N(0, [g'(\theta)]^2\sigma^2(\theta))$ (3.1)

In other words, $g(T_n)$ is approximately normal about $g(\theta)$ with variance $[g'(\theta)]^2\sigma^2(\theta)/n$.

We now present a result which helps obtain super-efficient estimator using the delta method.

Theorem 3.1: Let g: $\Omega \to \Omega$ be a function and let θ_0 be a point in the space Ω such that $g(\theta_0) = \theta_0$ and 'g' is differentiable at θ_0 with $|g'(\theta_0)| < 1$. If $\hat{\theta}_n$ be asymptotically efficient estimator of θ, then the estimator

$$\breve{\theta}_n = \begin{cases} g(\hat{\theta}_n) & if \;\; n^{1/4}|\hat{\theta}_n - \theta_0| \le 1 \\ \hat{\theta}_n & if \;\; n^{1/4}|\hat{\theta}_n - \theta_0| > 1 \end{cases} \tag{3.2}$$

is super-efficient at θ_0.

Proof: Given $\hat{\theta}_n$ is asymptotically efficient for θ, we get

$$\sqrt{n}\,(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, I^{-1}(\theta))\,, \theta \in \Omega \tag{3.3}$$

Applying the delta method to the point θ_0, we get

$$\sqrt{n}\,[g(\hat{\theta}_n) - \theta_0)] \xrightarrow{d} N(0, [g'(\theta_0)]^2 I^{-1}(\theta_0)) \tag{3.4}$$

Let A_n = [$n^{1/4} |\hat{\theta}_n - \theta_0| \leq 1$]. Following the discussion in Section 2, we find that $P_{\theta_0}(A_n) \to 1$ and $P_\theta(A_n) \to 0$ for $\theta \neq \theta_0$.

Now, $\breve{\theta}_n = \hat{\theta}_n\, I_{A_n^c} + g(\hat{\theta}_n)\, I_{A_n}$ and hence, $\sqrt{n}\,(\breve{\theta}_n - \theta) = \sqrt{n}\,(\hat{\theta}_n - \theta)\, I_{A_n^c} + \sqrt{n}\,(g(\hat{\theta}_n) - \theta)\, I_{A_n}$

For $\theta \neq \theta_0$, the second term above approaches zero as n $\to \infty$, while for $\theta = \theta_0$, the first term approaches zero as n $\to \infty$ and hence, we have

$$\sqrt{n}\,(\breve{\theta}_n - \theta) \xrightarrow{d} \begin{cases} N(0, I^{-1}(\theta)) & if\ \ \theta \neq \theta_0 \\ N(0, [g'(\theta_0)]^2 I^{-1}(\theta_0)) & if\ \ \theta = \theta_0 \end{cases}$$

Since, $|g'(\theta_0)| < 1$, we conclude that $\breve{\theta}_n$ is super-efficient at θ_0.

The above result, suggests a convenient procedure to obtain super-efficient estimator at a point of the paramter space by suitably choosing a function 'g' on the space. Essentially, the requirements on 'g' are

$$g(\theta_0) = \theta_0 \ \text{ and } \ |g'(\theta_0)| < 1 \tag{3.5}$$

As a generic example, we can have

$$g(\theta) = a_1 (\theta - \theta_0) + a_2 (\theta - \theta_0)^2 + \cdots + a_m (\theta - \theta_0)^m + \theta_0\ ,\ \theta \in \Omega \tag{3.6}$$

where $\theta_0 \in \Omega$ is a chosen point and $|a_1| < 1$. However, to illustrate the possibility of specific choices in various contexts, a few illustrative examples are presented in the next section.

4. APPLICATIONS

4.1 Mean of a Normal Distribution with Known Variance

Let $X_1, X_2, \ldots, X_n$ be i.i.d. $N(\theta, 1)$, $\theta \in R.$ Consider $g(\theta) = \theta^c$, $\theta \in R$, where $c > 1$ is any known positive integer. We have $g(0) = 0$ and $g'(0) = 0$ satisfying (3.5).

$\overline{X}$ is UMVUE of θ with $V_\theta(\overline{X}) = 1/n = I^{-1}(\theta)$

Define

$$\breve{\theta}_n = \begin{cases} \overline{X}^c & if\ \ \sqrt{n}\,|\overline{X}| \leq n^{1/4} \\ \overline{X} & if\ \ \sqrt{n}\,|\overline{X}| > n^{1/4} \end{cases}$$

Then, $\sqrt{n}\left(\breve{\theta}_n - \theta\right) \xrightarrow{d} \begin{cases} N(0,\ 0) & for\ \theta = 0 \\ N\left(0,\ I^{-1}(\theta)\right) & for\ \theta \neq 0 \end{cases}$

Showing that $\breve{\theta}_n$ is super-efficient at 0.

If we wish to have non-degenerate normal at $\theta = \theta_0$, choose g(θ) = c sin$\left(\theta - \theta_0\right) + \theta_0$, $\theta \in R$ where | c | < 1 and θ_0 is any fixed number, or choose g(θ) as in (3.6).

4.2 Variance of a Normal Distribution with Known Mean

Let $X_1, X_2, \ldots, X_n$ be i.i.d. N(0, θ), $\theta > 0$. Consider g(θ) = θ^d , $\theta > 0$, where $0 < d < 1$ is any known number. We have g(1) = 1 and g'(1) = d satisfying (3.5).

$S^2 = \sum_{i=1}^{n} X_i^2 / n$ is the MLE of θ with $V_\theta\left(\sum_{i=1}^{n} X_i^2 / n\right) = 2\theta^2 / n = I^{-1}(\theta)$

Define

$$\breve{\theta}_n = \begin{cases} (S^2)^d & if\ \sqrt{n}\left|S^2 - 1\right| \leq n^{1/4} \\ S^2 & if\ \sqrt{n}\left|S^2 - 1\right| > n^{1/4} \end{cases}$$

Then, $\sqrt{n}\left(\breve{\theta}_n - \theta\right) \xrightarrow{d} \begin{cases} N\left(0,\ d^2 I^{-1}(1)\right) & for\ \theta = 1 \\ N\left(0,\ I^{-1}(\theta)\right) & for\ \theta \neq 1 \end{cases}$

Showing that $\breve{\theta}_n$ is super-efficient at 1.

4.3 Poisson Parameter

Let $X_1, X_2, \ldots, X_n$ be i.i.d. Poisson(θ), $\theta > 0$. Consider the same g(θ) as in Illustration 4.2.

$\overline{X}$ is MLE of θ with $V_\theta\left(\overline{X}\right) = \theta / n = I^{-1}(\theta)$.

Define

$$\breve{\theta}_n = \begin{cases} \overline{X}^d & if\ \sqrt{n}\left|\overline{X} - 1\right| \leq n^{1/4} \\ \overline{X} & if\ \sqrt{n}\left|\overline{X} - 1\right| > n^{1/4} \end{cases}$$

Then, $\sqrt{n}\left(\breve{\theta}_n - \theta\right) \xrightarrow{d} \begin{cases} N\left(0,\ d^2 I^{-1}(1)\right) & for\ \theta = 1 \\ N\left(0,\ I^{-1}(\theta)\right) & for\ \theta \neq 1 \end{cases}$

Showing that $\breve{\theta}_n$ is super-efficient at 1.

4.4 Binomial Parameter

Let $X_1, X_2, \ldots, X_n$ be i.i.d. Bernoulli (1, *p*), 0 < *p* < 1. Consider g(*p*) = 2*p*(1− *p*), 0 < *p* < 1. We have g(1/2) = 1/2 and g'(1/2) = 0 satisfying (3.5).

$\hat{p}_n = \sum_{i=1}^{n} X_i / n$, is the MLE of '*p*' with $V_p(\hat{p}) = p(1-p)/n = I^{-1}(p)$

Define

$$\breve{p}_n = \begin{cases} 2\,\hat{p}_n\,(1-\hat{p}_n) & if \;\; \sqrt{n}\,|\,\hat{p}_n - 1/2\,| \le n^{1/4} \\ \hat{p}_n & if \;\; \sqrt{n}\,|\,\hat{p}_n - 1/2\,| > n^{1/4} \end{cases}$$

Then, $$\sqrt{n}\left(\breve{p}_n - p\right) \xrightarrow{d} \begin{cases} N(0,\,0) & for \;\; p = 1/2 \\ N\left(0,\, I^{-1}(p)\right) & for \;\; p \ne 1/2 \end{cases}$$

Which shows that $\breve{p}_n$ is super-efficient at 1/2.

4.5 Exponential Parameter

Let $X_1, X_2, \ldots, X_n$ be i.i.d. exponential variates with density $f_\theta(x) = (1/\theta)\exp(-x/\theta)$, $\theta > 0$. Let θ_0 be any specified value of θ and consider $g(\theta) = (\theta+\theta_0)/2$. Clearly, $g(\theta_0) = \theta_0$ and $g'(\theta_0) = 1/2$ satisfying (3.5).

$\overline{X}$ is MLE of θ with $V_\theta\left(\overline{X}\right) = \theta^2/n = I^{-1}(\theta)$

Define

$$\breve{\theta}_n = \begin{cases} (\overline{X} + \theta_0)/2 & if \;\; \sqrt{n}\,\left|\,\overline{X} - \theta_0\,\right| \le n^{1/4} \\ \overline{X} & if \;\; \sqrt{n}\,\left|\,\overline{X} - \theta_0\,\right| > n^{1/4} \end{cases}$$

Then, $$\sqrt{n}\left(\breve{\theta}_n - \theta\right) \xrightarrow{d} \begin{cases} N\left(0,\, I^{-1}(\theta_0)/4\right) & for \;\; \theta = \theta_0 \\ N\left(0,\, I^{-1}(\theta)\right) & for \;\; \theta \ne \theta_0 \end{cases}$$

Which shows $\breve{\theta}_n$ is super-efficient at θ_0.

5. CONCLUDING REMARKS

The foregoing discussion presents an approach towards construction of super-efficient estimators using the delta method of asymptotic inference in single parameter situations. The delta method of asymptotic inference has provided a convenient way of constructing super-efficient estimators. Numerical investigations on the estimators proposed in the illustrative examples of Section 4 are at a preliminary stage of investigation. Other applications of this approach are also being addressed separately. The outcomes of these ongoing research work will be reported in future communications.

REFERENCES

1. Bahadur, R. R. Hodges Super-Efficiency. In: *Encyclopedia of Statistical Sciences*. 3, (1983), John-Wiley, 645-646.
2. Durairajan, T. M. Sub-score and Super-efficient Estimator. In: Navin Chandra and Gopal, G., eds. *Applications of Reliability Theory and Survival Analysis*, (2012), Bonfring, 120-129.
3. Hodges, J. L. *Unpublished*. (1951)
4. Le Cam, L. On some asymptotic properties of maximum likelihood estimates and related Baye's estimates. *University of California Publications in Statistics*. 1, (1953), 277-330.
5. Le Cam, L. On the asymptotic theory of estimation and testing hypotheses. In: Neyman, J. ed. *Proceedings of the Third Berkeley Symposium on Mathematics and Probability*. 1, (1956), University of California Press, Berkeley, 129-156.
6. Le Cam, L. Locally asymptotically normal families of distributions. *University of California Publications in Statistics*. 3 (1960), 37-98.
7. Le Cam, L. Limits of experiments. In: *Proceedings of the Sixth Berkeley Symposium on Mathematics and Probability*. 1, (1972), University of California Press, Berkeley. 245-261.
8. Sethuraman, J. Are Super-efficient Estimators Super-powerful?. *Communications in Statistics - Theory and Methods*. 33(9), (2004), 2003-2013.
9. Stein, C. Inadmissibility of the usual estimator for the mean of a normal distribution. In: Neyman, J. ed. *Proceedings of the Third Berkeley Symposium on Mathematics and Probability*. 1, (1956), University of California Press, Berkeley, 197-206.