IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 11, Issue 2 Ver. II (Mar - Apr. 2015), PP 18-20
www.iosrjournals.org

A Note on Fisher's Method of Constructing An Efficient Estimator

U.B.AITHAL
The New College, Kolhapur, Maharashtra, India.

Abstract: *Fisher (1925), in his famous paper, gave a method of constructing an efficient estimator by using a successive approximation. If there exists a root-n consistent estimator of the parameter and if the probability density function satisfies certain regularities conditions, it is possible to construct an efficient estimator of the parameter. This can be done by a single approximation (Lehman and Casella, 1998). In order to distinguish between the efficient estimators, Rao((1961, 1963) introduced the concept of second order efficiency (s.o.e.) of an estimators. In this note, we consider the estimator obtained by using two iterations in the approximation and obtain its s.o.e. following the definition of Rao. It follows from this derivation that we can construct a second order efficient estimator, if it exists, by using any root-n consistent estimator of the parameter.*
Keywords: *and phrases: First order efficient estimator, maximum likelihood estimator, root-n consistent estimator, second order efficiency.*

I. Introduction

Consider a sample $x_1,x_2,\ldots x_n$ of size n from a distribution having the probability density function(p.d.f.) $f(x,\theta)$ with respect to a σ-finite measure μ, where θ is real parameter Let $\theta_n^* = \theta^*(x_1,\ldots x_n)$ be any estimator of θ such that, $\sqrt{n}(\theta_n^* - \theta)$ is bounded in probability.

Definition 1:A sequence of estimators θ_n^* is $\sqrt{n}$-consistent for θ, if $\sqrt{n}(\theta_n^* - \theta)$ is bounded in probability, that is, $(\theta_n^* - \theta)=O(1/\sqrt{n})$

Under the following regularity conditions satisfied by $f(x,\theta)$ [Lehman and Casella (1998)] .

i)The parameter space Ω is an open interval(not necessarily finite)

ii)The distributions P_θ of X_i have common support, so that the set $S=\{x:f(x,\theta)>0\}$ is independent of θ.

iii) For every $x \epsilon S$ (sample space), $f(x,\theta)$ is thrice differentiable with respect to θ, and the third derivative is continuous in θ.

iv)The integral can be thrice differentiated under the integral sign.

v) For any given $\theta_0 \epsilon \Omega$ (parameter space), there exists a positive number c and a function M(x) such that,

$| \{\partial^3/ \partial\theta^3\}\{\log f(x,\theta)\}| \le M(x)$ for all $x \epsilon S$, $\theta_0 - c < \theta < \theta_0 + c$ and $E_\theta | M(x) | < \infty$

vi)The Fisher information I is finite.

Theorem 1:If the above regularity conditions are satisfied by $f(x,\theta)$ and θ_n^* is a sequence of consistent asymptotically normal (CAN) estimators of θ,then the estimator sequence

$$\hat{\theta}_n = \theta_n^* -[L'(\theta_n^*)/L''(\theta_n^*)] \quad \ldots(1.1)$$

is asymptotically efficient(BAN),that is, $\sqrt{n}\,(\hat{\theta}_n - \theta)$ asymptotically normal with mean zero and variance 1/I , where L' , L'' are the first and second derivatives of the log-likelihood function L(θ) based on the sample of n observations.(Proof, Lehman and Casella, 1998)

Since $\sqrt{n}\,(\theta_n^* - \theta)$ is bounded in probability, using the first order approximation for $L'(\theta_n^*)$ and $L''(\theta_n^*)$ in (1.1), it follows that

$$|[L'(\theta)/\sqrt{n}] - I \sqrt{n}\,(\hat{\theta}_n - \theta) | \to 0, \text{in probability.} \quad \ldots(1.2)$$

This property of $\hat{\theta}_n$ is termed as the first order efficiency (f.o.e) by Rao (1961, 1963). He defines that second order efficiency (s.o.e) of any first order efficient estimator.

Definition 2: The s.o.e of an estimator $T_n = T(x_1,x_2,\ldots x_n)$ is the minimum asymptotic variance

$$S_\lambda^2(T_n) =[L'(\theta) - nI(T_n-\theta) - n\lambda(T_n - \theta)^2] \quad \ldots(1.3)$$

when minimized with respect to λ, where λ is a function of θ only.

II. Second Order Efficient Estimator Of θ

Let us define an estimator δ_n of the parameter θ such that,

$$\delta_n = \hat{\theta}_n - [L'(\hat{\theta}_n) / L''(\hat{\theta}_n)] \quad \ldots(2.1)$$

Theorem 2:The estimator δ_n defined above is a second order efficient estimator of θ.

Proof: By replacing the regularity condition iv) in theorem 1 in to iv)For any given $\theta_0 \varepsilon \Omega$ (parameter space), there exists a positive number c and a function M(x) such that

$$| \partial^4 / \partial\theta^4 \log f(x,\theta)| \leq M(x) \text{ for all } x \in S,$$
$$\theta_0 - c < \theta < \theta_0 + c$$

and

$$E_\theta | M(x) | < \infty$$

Following the definition 2. Using Taylor expansion for $L'(\hat{\theta}_n)$ and $L''(\hat{\theta}_n)$ around the true parameter θ, we can rewrite the equation (2.1) in the form

$$(\delta_n - \theta)[L''(\theta) + (\hat{\theta}_n - \theta) L'''(\theta) + o(\hat{\theta}_n - \theta)^2]$$
$$= (\hat{\theta}_n - \theta)[L''(\theta) + (\hat{\theta}_n - \theta) L'''(\theta) + o(\hat{\theta}_n - \theta)^2]$$
$$-[L'(\theta) + (\hat{\theta}_n - \theta)L''(\theta) + (1/2)(\hat{\theta}_n - \theta)^2 L'''(\theta) + o(\hat{\theta}_n - \theta)^3] \quad ..(2.2)$$

Since $\sqrt{n}(\hat{\theta}_n - \theta)$ is bounded in probability, the above equation reduces to

$$(\delta_n - \theta)\, nI = L'(\theta) + (\delta_n - \theta)\,[L''(\theta)+nI] + (\delta_n - \theta)(\hat{\theta}_n - \theta)L'''(\theta) - (1/2)(\hat{\theta}_n - \theta)^2 L'''(\theta) + o(1/n) \quad ..(2.3)$$

Using the above approximation, we get

$$S_\lambda^2(\delta_n) = - \{ (\delta_n - \theta)[L''(\theta)+nI] + (\delta_n - \theta)(\hat{\theta}_n - \theta)L'''(\theta) - 1/2(\hat{\theta}_n - \theta)^2 L'''(\theta) + n\lambda(\delta_n - \theta)^2\} + o(1/n) \quad ..(2.4)$$

The asymptotic distribution of $S_\lambda^2(\delta_n)$ is unaltered if we replace $(\delta_n - \theta)$ and $(\hat{\theta}_n - \theta)$ by $L'(\theta)/nI$. After some simplification, the asymptotic variance of $S_\lambda^2(\delta_n)$ is

$$V[S_\lambda^2(\delta_n)] = (1/n^2I^2)\ V\{L'(\theta)[L''(\theta)+nI]\} + (1/4n^4I^4)\ V\{L'(\theta)^2L'''(\theta)\} + (\lambda^2/n^2I^4)\ V\{L'(\theta)^2\} + (1/n^3I^3)\ \text{Cov.}\{L'(\theta)[L''(\theta) + nI]\ ,\ L'(\theta)^2L'''(\theta)\} + \{(2\lambda/n^3I^3)\}\ \text{Cov.}\{L'(\theta)[L''(\theta) + nI]\ ,\ L'(\theta)^2\} + (\lambda/n^3I^4)\ \text{Cov.}\{L'(\theta)^2L'''(\theta)\ ,\ L'(\theta)^2\} \quad ..(2.5)$$

Using the moment relations of the normally distributed random variables (Rao, 1961) (Aithal, and Nagnur 1993), the above expression reduces to

$$V[S_\lambda^2(\delta_n)] = (1/n^2I^2)\ V\{L'(\theta)[L''(\theta)+nI]\} + \{E[L'''(\theta)]^2 / (4n^4I^4)\}\ V\{L'(\theta)^2\} + (\lambda^2/n^2I^4)\ V\{L'(\theta)^2\}$$
$$+ \{E[L'''(\theta)] / (n^3I^4)\}\ \text{Cov.}\{L'(\theta)[L''(\theta) + nI]\ ,\ L'(\theta)^2\}$$
$$+ \{(2\lambda/n^2I^3)\}\ \text{Cov.}\{L'(\theta)[L''(\theta) + nI]\ ,\ L'(\theta)^2\} + \{E[L'''(\theta)] / (n^3I^4)\}\ V\{L'(\theta)^2\} \quad ...(2.6)$$

Using the result,

$$E[L'''(\theta)] = -\{3E[L'(\theta)L''(\theta)] + E[L'(\theta)^3]\} \quad ...(2.7)$$

After simplification, we get

$$V[S_\lambda^2(\delta_n)] = K_{02}/I + [1/I^2]\,[K_{30}{}^2/2 + K_{11}K_{30} - K_{11}{}^2/2] + [2\lambda^2/I^2] - [2\lambda/I^2][K_{11} + K_{30}], \quad ...(2.8)$$

Where,

$$K_{ij} = E[(L'(\theta)/\sqrt{n})^i\ (\{L''(\theta) + nI\}/\sqrt{n})^j]$$

The minimum of $V[S_\lambda^2(\delta_n)]$ occurs when $\lambda = (K_{11}+K_{30})/2$ and the minimum value of the variance is the s.o.e. of δ_n and is denoted by

$$E_2 = (1/I^2)[I\,K_{02} - K_{11}{}^2]. \quad ...(2.9)$$

This result coincides with the s.o.e. of the maximum likelihood estimator of θ which is most efficient among all f.o.e.estimators of θ.(Rao,1961)

III. Relation With Statistical Curvature And Bias Correct Variance Of δ_n

Fisher (1925) has called E_2 as the loss of information in using δ_n as an estimator of θ. The quality $(nI - E_2)$ is defined as the second order information of the maximum likelihood estimator by Hosoya(1985). Efron(1976) gives the relationship $E_2 = I\gamma_\theta{}^2$, where $\gamma_\theta{}^2$ is the statistical curvature of f(x,θ) at θ. We can directly compute the bias and variance of δ_n from (2.3), which reduces to

$$E\ (\delta_n - \theta) = b\ (\theta)/n + o\ (1/n), \quad ...(3.1)$$

and

$$V(\delta_n) = 1/nI + 2b'(\theta)/n^2I + (1/n^2)\ \{[2(I\,K_{02} - K_{11}{}^2) + (K_{11} + K_{30})^2]/2I^4\} + o(1/n^2) \quad ...(3.2)$$

Where, $b(\theta) = -(K_{11} + K_{30})^2/2nI^2$.(Gudi and Nagnur,2004)

The above expression agrees with the expressions of Cox and Hinkley (1975) for the maximum likelihood estimator of θ. The variance of the bias corrected estimator $\delta_n^* = \delta_n - b(\theta)/n$ is

$$V(\delta_n^*) = 1/nI + \psi(\theta)/n^2 + o(1/n^2), \quad ...(3.3)$$

where,

$$\psi(\theta) = [2(I\,K_{02} - K_{11}{}^2) + (\,K_{11} + K_{30})^2]/2I^4$$
$$= E_2/I^2 + (K_{11} + K_{30})^2/2I^4 \quad ...(3.4)$$

IV. Examples

(a) Multinomial Distribution: Rao (1961, 1963) has discussed the computation of E_2 for the mutinomial distribution with the cell probabilities depending on a real parameter θ. Suppose that for a k-cells multinomial distribution $\pi_1(\theta),\ldots,\pi_K(\theta)$ are the cell probability with the observed proportions $p_1,\ldots,p_K$. The log-likelihood function $L(\theta) = \sum_1^k p_r \log \pi_r(\theta)$. Let θ_n^* be an estimator obtain by minimizing the distance function $\Delta(p,\pi) = \Sigma[h(p_r) - h(\pi_r)]^2$ (Taylor 1953). In this case, the estimator θ_n^* is a consistent asymptotically normal (CAN) estimator of θ.By following above procedure (1.1) and (2.1),we get

$E_2 = (\mu_{02} - 2\mu_{21} + \mu_{40})/I - I - (\mu_{11} - \mu_{30})^2/I^2$, …(4.1)
where, $I = \sum_1^k (\pi_r')^2/\pi_r$, $K_{11} = \mu_{11} - \mu_{30}$ $K_{02} = \mu_{02} - 2\mu_{21} + \mu_{40} - I^2$ and $\mu_{ij} = \sum_1^k (\pi_r'/\pi_r)^i (\pi_r''/\pi_r)^j \pi_r$.
(b)Cauhcy Distribution(location parameter): Consider a Cauchy distribution having the p.d.f.

$f(x,\theta)=(1/\pi)[1/\{1+(x-\theta)^2\}]$, $-\infty<x<\infty$, $-\infty<\theta<\infty$.

The log-likelihood function $L(\theta)=(2/\pi)\Sigma(x_i- \theta)/[1 + (x_i +\theta)^2]$.Let θ_n^* be the sample median which is a $\sqrt{n}$-consistent estimator of θ.By following (1.1) and (2.1) , routine computation yield I=1/2, $K_{11}=0, K_{02}=5/8$ .Substituting these values in (2.9) we get,$E_2=1.25$. …(4.2)
(c) Cauchy Distribution (scale parameter): Consider a Cauchy distribution having the p.d.f.

$f(x,\theta) =[\theta/\pi] [1/(\theta^2 + x^2)]$, $-\infty<x<\infty, \theta>0$.

The log-likelihood function $L(\theta)=(1/\pi)\Sigma[1/\theta - 2\theta/(x_i^2 +\theta^2)]$.Let θ_n^* be the sample quartile deviation and is a $\sqrt{n}$-consistent estimator of θ.By following (1.1) and (2.1) , routine computation yield $I = 1/2\theta^2$, $K_{11} = -1/2\theta^3$, $K_{02} =5/8\theta^4$. The log-likelihood function $L(\theta)=(1/\pi)\Sigma[1/\theta - 2\theta/(x_i^2 +\theta^2)]$.Let θ_n^* be the sample quartile deviation which is a $\sqrt{n}$-consistent estimator of θ.By following (1.1) and (2.1) , routine computation yield
Substituting these values in (2.9), we get $E_2 = 1/4\theta^2$. …(4.3)

Reference

[1]. Aithal,B.U and Nagnur,B.N (1990) Second order efficiency of the BAN estimators for the multinomial family.Comm.Statist.Theory.Meth. 19(5)1763-1776.
[2]. Cox , D.R. and Hinkley, D.V. (1974) : Theoretical statistics, Chapman and Hall, London, 279-363.
[3]. Efron, B. (1975): "Defining the Curvature of a Statistical of a Statistical Problem with applications to second order efficiency", Ann,Statist, 3.
[4]. Fisher, R.A. (1925): "Theory of Statistical Estimation", Canb.Phil.Soc. 22, 700-725.
[5]. Gudi,S.V and Nagnur,B.N (2004) Deficiency of Two Estimator in One-Parameter Exponential Family of DistributionComm.Statist.Theory.Meth 33(8) 1779-1800.
[6]. Hosoya, Y. (1988): "The second order Fisher information" ,Biometrika, 75(2), 233-235.
[7]. Lehmann.E.L and Casella.G (1998): Theory of Point Estimation, Springer-Verlag, New York.
[8]. Rao, C.R. (1961): "Asymptotic efficiency and limiting information", Proc. Fourth Berk.Symp.Math.Statist.Prob., 1,531-546.
[9]. ------, (1963): "Criteria of estimation in large samples", Sankhya,Ser A, 34, 133-144.
[10]. Taylor.W.F.(1953).Distance functions and regular best asymptotically normal estimators,Ann.Math.Statist.24,85-92.