e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

A Study on Minimum Chi-Square Type Estimators

*1 DHANDAPANI.M, MPhil Scholar, Department of Mathematics Bharath Institute of Higher Education and Research, Chennai,-73, India.

*2 Dr. R. Ishwariya , Associate Professor, Department of Mathematics, Bharath Institute of Higher Education and Research, Chennai,-73, India.

dhandapan2001@gmail.com ishwariyarose@gmail.com

Address for Correspondence

*1 DHANDAPANI.M, MPhil Scholar, Department of Mathematics Bharath Institute of Higher Education and Research, Chennai,-73, India.

*2 Dr. R. Ishwariya , Associate Professor, Department of Mathematics, Bharath Institute of Higher Education and Research, Chennai,-73, India.

dhandapan2001@gmail.com ishwariyarose@gmail.com

Abstract:

The chi-square type estimators for the multinomial set-up. The estimators obtained by minimizing the distance functions $D_n(\text{p}, \pi(\theta))$ are BAN estimators laving the same asymptotic properties as the m.1. estimators. Rao (1961,63) introduced the concept of second order efficiency (s.o.e.) to discriminate different BAN estimators. He has also given a procedure of computing the second order efference of BAN estimators for the multinomial set-up.

Keywords: BAN estimators, Chi-square type estimator, Computation of $\psi(\theta)$ for BAN Estimators.

Introduction:

The second order efficiency of the minimum chi-square type estimators, In section 2 we give some definitions of s.0.e. found in the literature. In section 3 we obtain a general expression for s.0.e. of a BAN estimator. Second order/ efficiencies of some well-known estimators are obtained Sectron 5 deals with the computation of second order efference for the minimum chi-square type estimators. We have also discussed def 1 of the estimators in section 6. In Append1x, proofs of the results obtained in section 3 are given.

e-ISSN 2320 –7876 www.ijfans.org
Research Paper Vol.11, Iss.9, Dec 2022

1.2 Some Definitions of Second Order Efficiency

The term second order efference was introduced by Fao (1961) but the concept as well as the first main result in this area occurs

In Fisher (1925). $F_{1\text{ sher}}$ proposes $E_2' = \lim_{n+\infty}(i, -1_{T_n})$

as a measure of s.0.e. to discriminate between different asymptotically efficient estimators, where 1. is the Fisher information contained1n a single observation and $^1\,T_n$ 15 the information per observation contained in the estimator T_n. Note that, if T_n is a BAN estimator $^1T_n + 1$ as $n + \infty$. The estimator having smaller value of E_2' has more second order efficiency. F1sher proved that the m.1. estimator minimizes E_{2+}' .

Somewhat surprisingly second order efficiency remained neglected t111 it was reconsidered by Rao (1961) who makes major progress in proving Fisher's result that the m.1. estimator possessed highest s.o.e. However, the results actually proved differs in two ways from What Fosher stated. Rao introduces a more eastly computable and a more useful measure E_2 and secondly he restricts attent on to F_1 sher consistent estimators. We def $1ne$ below second order efficiency as given by Rao (1961). A statistic T_n is said to be first order effecting If $\left|n^{1/2}z_n - \alpha - \beta(\theta)n^{1/2}(T_n - \theta)\right| + 0$ in probability under θ; where $\beta(\theta)$ is a function of θ only, and $Z_n = \frac{1}{n} d \log P(x_h, \theta) / d\theta, P(x_n, \theta)$ being the density of the observations.

The condition (2.2.2) Implies that the asymptotic correlation between T_n and z_n is unity. There are several estimators which are first order efficient. To compare among them we consider the random variable $\frac{d\log L}{d\theta} - \alpha n^{1/2} - \beta n(T_n - \theta) - \lambda n(T_n - \theta)^2$ which is the difference between $\frac{d\log L}{\partial\theta}$ and a second order polynomial approximation for $\frac{\partial \log \phi_n}{\partial\theta}$ where ϕ_n15 the probab1lity density function of T_n. The constant λ may be chosen so as to minmize the asymptotic varrance of (2.2.3). Thus, a second measure of second order efference may be defined as E_2, where E_2 is the minimum value of variance of (2.2.3) when minimized with respect to λ.

Rao (1963) proposed to compare different estimators in a more direct way by assuming a quadratic loss function. Let

$$b(\theta) = \mathrm{l}\ \ n[E(T_n) - \theta]$$

and

$$T_n^* = T_n - \frac{b(T_n)}{n}.$$

Consider ang a quadratic loss function, if

$$E(T_n^* - \theta)^2 = \frac{1}{n1} + \frac{\psi(\theta)}{n^2} + 0\left(\frac{1}{n^2}\right),$$

then $\psi(\theta)$ can be taken as a third measure (E_3) of second order efficiency. In the next section we discuss the computation of $\psi(\theta)$ for BAN estimators for the multinomial set-up. In this case the latter two definitions become equivalent.

1. 3 Computation of $\psi(\theta)$ for BAN Estimators

In this section we discuss the computation of $\psi(\theta)$ for a BAN estimator in the multinomial set-up. Let $(n_1, \dots, n_k) \sim M(\mathbb{E}; \pi_1(\theta), \dots, \pi_k(\theta))$ where θ is an unknown real parameter. To estimate θ we rave several procedures lake maximum likelihood, mint mum ch1-square and minimum chi-square type etcetera In general the estimating equation is of the form

f (0,p1………p_k)= 0

where $p_1 = \frac{n_1}{n}, 1 = 1, \dots, k$. The following assumptions are made.

(A l) $\pi_1(\theta),\ 1 = 1, \dots, k$ admit derivatives upto the third order, which are continuous in the neighbourhood of the true value of θ.

$\langle A2)$ The estimating equatron is consistent, that 15 ,

e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

$$f(\theta, \pi_1(\theta), \dots, \pi_k(\theta)) \equiv 0$$

The assumption maples that the estimator obtained from the equation

is Fisher consistent.

(A 3) f as a functron of $\theta, p_1, \dots, p_k$ admits third order partial derivatives which are bounded 1 n a closed region P of the cube

$$0 \leq p_1 \leq 1,\ 1 = 1, \dots, k$$

and for values of θ satisfyng (2,3.1) with $(p_1, \dots, p_k) \in P$.

(A 4) The true point $\pi_1(\theta), \dots, \pi_k(\theta)$ is an interior point of p. Ghosh and Subramanyam (1974) call an estimator T_n locally stable of order two (1.s.(11)) if it 15 Fisher consistent and possesses second order derivatives w.r.t. p,'s. If T_n is Fisher consistent and possesses third order derivatives, tt 1 s referted as 1.5.(111). Since $\frac{\partial T_n}{\partial p_1}, \frac{\partial^2 T_n}{\partial p_1 \partial p_J}$ etc. can be written in

terms of $\frac{\partial^2 f}{\partial p_1 \partial p_J}, \frac{\partial^3 f}{\partial \theta^3}$

etc., under the assumptions $(A1)$ to $(A4)$, the estimator is $1.5.(11I)$.

Let $\forall_n$ be a solution of the equatron (2.3.1) such that $\tilde{\theta} \neq \theta$ as $p_1 + \pi_7(\theta)$. Denote by

$f', f_r, f_{rs}, f_{rsk}, f_r', f_r''$ and

$\mathrm{f}_{\mathrm{rs}}'$ the partıal derıvatives $\frac{\partial \mathrm{f}}{\partial \hat{\theta}_n}, \frac{\partial f}{\partial p_{\mathrm{r}}}, \frac{\partial^2 \mathrm{f}}{\partial p_{\mathrm{r}} \partial p_{\mathrm{s}}}, \frac{\partial^3 \mathrm{f}}{\partial p_r \partial p_{\mathrm{s}} \partial p_{\mathrm{K}}}$,

$\frac{\partial^2 f}{\partial \vec{\theta}_n \partial p_r}, \frac{\partial^3 f}{a_\theta^2 \partial p_r}$ and $\frac{\partial^3 f}{\partial \hat{\theta}_n \partial p_r \partial p_s}$ evaluated at $\tilde{\theta}_n = \theta$ and

$\mathrm{D} = \pi \sim$ respectively. Expanding $f\left(\tilde{\theta}_n, \mathrm{p}_1, \dots, \mathrm{p}_{\mathrm{k}}\right)$ by Taylor's theorem

at $(\theta, \pi_1(\theta), \dots, \pi_k(\theta))$, we get

The summation extends over $1 = 1, \dots, k$ unless otherwise stated. Since $\tilde{\theta}_n$ is BAN est 1 mator (1.5.3),

$$\frac{f_r}{f^{\top}} = -\frac{1}{1.} \frac{\pi_r}{\pi_r}$$

e-ISSN 2320 –7876 www.ijfans.org
Research Paper Vol.11, Iss.9, Dec 2022

Using the above relation and rearranging the terms in (1.3.2) we can write

$$(\tilde{\theta}_n - \theta) = \phi_r^{(1)} + \phi_{r=}^{(2)} + \phi_{rs\ell}^{(3)} + 0(n^{-2}),$$

where

$$\phi_r^{(1)} = \frac{1}{1} \sum_r \frac{\pi_r}{\|_r}(p_r - \pi_r),$$

$$\phi_{rs}^{(2)} = -\frac{1}{2f} \left[\sum_r \sum_s f_{rs}(p_r - \pi_r)(p_s - \pi_s) + 2(\theta_n - \theta) \right.$$
$$\left. \sum_r f_r'(p_r - \pi_r) + (\tilde{\theta}_n - \theta)^2 f'' \right]$$

and

$$\phi_{rsl}^{(3)} = - \frac{1}{6f'}\left[(\tilde{\theta}_n - \theta)^3 f''' + 3(\sigma_n - \theta)^2 \varepsilon f_r''(p_r - \pi_r) \right.$$
$$+3(\theta_n - \theta) \sum_r \sum_s f_{rs}'(p_r - \pi_r)(p_s - \pi_s)$$
$$\left. + \sum_r \sum_s \varepsilon f_{rs\ell}(p_r - \pi_r)(p_s - \pi_s)(p_\ell - \pi_\ell) \right]$$

From the expression (1.3.4) we compute $E(\tilde{\theta}_n - \theta)$ and $E(\theta_n - \theta)^2$. Since $\tilde{\theta}_n$ is a function of $p_1, \ldots, p_k$, to compute mean and var $_1$ ance of $(\tilde{\theta}_n - \theta)$ erther we can use the exact moments of the multinomial distribution and collect the terms to the relevant order or alternatively we can use the formal multivariate Edgeworth expansion for the joint density of $\sqrt{n}(p_1 - \pi_1), \ldots, \sqrt{n}(p_{k-1} - \pi_{k-1})$. The validity of using such formal Edgeworth expansion follows from Gotze and Hipp (1978) and Ghosh, Sinha and Subramanyam (1979). Harris (1985) also compares the exact moments and moments obtained by formal Edgeworth expansion. The advantage of using this type of approach would be clear when we link the variance of $(\theta_n - \theta)$ obtained to the vartance of the asymptotic distrobution of $\sqrt{n1.}(\tilde{\theta}_n - \theta)$.

e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

The bras and variance of $\tilde{\theta}_n$ are computed using the expression (2.3.4) and the Edgeworth density of length 1. To the order of approximation considered, we need Edgeworth expansion of length one. The bras of $\ddot{\theta}_n$ to $0\left(\frac{1}{n}\right)$ is

$$\frac{b(\theta)}{n} = E_1'\left(\phi_r^{(1)} + \phi_{rs}^{(2)}\right)$$
$$= \frac{1}{2n}\left[-\frac{f''}{1+f} - \frac{2}{1+f'}\Sigma_r f_r^1 \pi_r' - \frac{1}{f'}\Sigma_r f_{rr}\pi_r + \frac{1}{f}\Sigma_r \Sigma_s f_{rs}\pi_r\pi_s\right]$$

and the variance $V(\theta_n)$ to the order $1/n^2$ is $V(\tilde{\theta}_n) = \frac{1}{n1.} + \frac{1}{n^2}\left[2E_1'\left(\phi_r^{(1)}\phi_{rs}^{(2)}\right) + 2E_1'\left(\phi_r^{(1)}\phi_{rsl}^{(3)}\right)\right.$

$$\left. + E_1'\left(\phi_{rs}^{(2)}\phi_{rs}^{(2)}\right)\right] - \frac{b^2(\theta)}{n^2} + 0(n^{-3})$$

The above expectation E_1^1 is with respect to the Edgeworth density. The details of obtaining the expectations are given in Append 1x and

we get

where

$$\mu_{1J} = \Sigma\pi_r\left(\frac{\pi_r'}{\pi_r}\right)^1\left(\frac{\pi_r^n}{\pi_r}\right)^J,$$

$$\frac{b(\theta)}{n} = E_1'\left(\phi_r^{(1)} + \phi_{rs}^{(2)}\right)$$
$$= \frac{1}{2n}\left[-\frac{f''}{1+f} - \frac{2}{1+f'}\Sigma_r f_r^1 \pi_r' - \frac{1}{f'}\Sigma_r f_{rr}\pi_r + \frac{1}{f}\Sigma_r \Sigma_s f_{rs}\pi_r\pi_s\right]$$

and the variance $V(\theta_n)$ to the order $1/n^2$ is $V(\tilde{\theta}_n) = \frac{1}{n1.} + \frac{1}{n^2}\left[2E_1'\left(\phi_r^{(1)}\phi_{rs}^{(2)}\right) + 2E_1'\left(\phi_r^{(1)}\phi_{rsl}^{(3)}\right)\right.$

$$\left. + E_1'\left(\phi_{rs}^{(2)}\phi_{rs}^{(2)}\right)\right] - \frac{b^2(\theta)}{n^2} + 0(n^{-3})$$

The above expectation E_1^1 is with respect to the Edgeworth density. The detatis of obtanning the expectatrons are given in Append 1x and

we get

where

$$\mu_{1J} = \Sigma\pi_r \left(\frac{\pi_r'}{\pi_r}\right)^1 \left(\frac{\pi_r^n}{\pi_r}\right)^J,$$

and $4E_1'\left(\phi_{rs}^{(2)}\phi_{rs}^{(2)}\right) = \frac{3f''z}{1.^2f'^2} + \frac{12f''}{1.^2f'^2}\Sigma f_r'\pi_r'$

Substituting (1.3.10), (1.3.11) and (1.3.12) in (1.3.9) we get

The bras corrected estimator 1 s

$$\theta_n^\star = \tilde{\theta}_n - \frac{b(\vec{\theta}_n)}{n}.$$

The variance of the bras corrected estimator θ_n^k[Rao, 1963]is

$$V(\theta_n^\star) = V(\tilde{\theta}_n) - \frac{2b'(\theta)}{n^2 1.} + 0(n^{-3})$$

It follows from (1.3.13) and (1.3.15) that It may be noted that we have used expansion for the estimating equation and thus our approach differs from Ghosh and Subramanyam (1974) and Koorts (1985) where expansion of the estimator is considered. However, unlike Rao (1963) we have considered a direct approach to compute Var $(\theta_n^\star)$ f_r w.r.t. θ (cf. 1.3.3) we get It follows from (1.3.21) that $\psi(\theta)$ involves sum of two terms; a term which does not depend upon the estimating equation and a term that depends on the estimating equation through f′ and f_{rs} only. We wish to rewrite $\psi(\theta) \cap n$ an alternative form.

where $d_r = (p_r - \pi_r)$ and $z_n = \sum \frac{\pi_r}{\pi_r} d_r$

We have and

$$V(q) = \frac{1}{n^2}\left[\frac{1}{1.4}(\Sigma\Sigma f_{rs}\pi_r'\pi_s')^2 + \frac{1}{1.3}(\Sigma\Sigma f_{rs}\pi_r\pi_s')^2\right.$$

Substituting (2.3.24) and (2.3.25) in (2.3.21) we get $u(\theta) = \left[\frac{\mu_{02}-2\mu_{21}+\mu_{40}}{1.3} - \frac{1}{1.} - \frac{(\mu_{11}-\mu_{30})^2}{1.4}\right.$

e-ISSN 2320 –7876 www.ijfans.org
Vol.11, Iss.9, Dec 2022

$$+\frac{\mu_{11}^{2}}{21.4}\bigg] + \frac{n^{2}1\cdot^{2}}{f^{2}}\left[v(Q) - \frac{[\operatorname{cov}(Q, z_{n}^{2})]^{2}}{v(z_{n}^{2})}\right]$$

$$\bigg[\frac{\mu_{0}d - 2\mu_{21} + \mu_{40}}{1.3} - \frac{1}{1\cdot} - \frac{(\mu_{11} - \mu_{30})^{2}}{1^{4}}$$

$$+\frac{\mu_{11}^{2}}{21.^{4}}\bigg] + \frac{n_{1}^{2}1\cdot^{2}\delta^{\star}}{f'^{2}}$$

since $\delta^{*} = v(\theta) - \frac{[\operatorname{cov}(0,z_{n}^{2})]^{2}}{v(z_{n}^{2})} \geq 0$, it 1 s clear that an estimator

we have maximum second order efficiency if $\delta^{\star} = 0$.

We can establish the relationship between the two measures second order efficiency E_2 and E_3. It follows from (2.3.26) that

$$\psi(\theta) = \frac{n^{2}1\cdot 4}{f'^{2}}\left[\operatorname{var}(0) - \frac{[\operatorname{cov}(0, z_{n}^{2})]^{2}}{\operatorname{var}(z_{n}^{2})}\right] + 1.2\psi(m.1+)$$

where e $\psi(\mathrm{m}, 1.)$ is the value of $\psi(\theta)$ for the m.1. estimator. Ne can write $\mathrm{H}(\mathrm{m}.1.)$ as [Ra0,1963]

$$1.^{2}\,\psi(\mathrm{m}.1.) = \mathrm{E}_{2}(\,\mathrm{m}.1.) + \frac{\mu_{11}^{2}}{21.^{2}}$$

where -e $E_{2}(m.1.)$ is the value of E_2 for m.l. estimator. Hence we can

$$1.^{2}\,\psi(\theta) \quad = \frac{n^{2}1.4}{f^{2}}\left[\operatorname{var}(0) - \frac{[\operatorname{cov}(0, z_{n}^{2})]^{2}}{\operatorname{Var}(z_{n}^{2})}\right]$$

$$+E_{2}(m\cdot 1.) + \frac{\mu_{11}^{2}}{21.^{2}}$$

$$= E_{2} + \frac{\mu_{11}^{2}}{21.2}.$$

the above equation we have written

$$E_{2} = E_{2}(m+1.) + \frac{n^{2}1.4}{f^{2}}\left[\operatorname{Var}(Q) - \frac{[\operatorname{cov}(0, z_{n}^{2})]^{2}}{\operatorname{Var}(2_{n}^{2})}\right]$$

which is the expression obtained in Rao (1961).

Conclusion:

e-ISSN 2320 –7876 www.ijfans.org
Vol.11, Iss.9, Dec 2022

We have corrupted the values of deficiencies of estimators obtained by minimizing $0_n^{(n)}$ and is presented in Table 2.6.2. A glance at the table indicates that when $n \in 0$, the magnitude of deforciency is quite substantial. For exarple, when $\theta = 0.5625$ and $n = -2$, the deficiency 15 116.0727, indicati_ng that if m.1. estimator requires 100 observations to have the sane preciston, the estimator obtained through $b_n^{(-2)}$ fequires 216 conservatories. However, when $n = 0.5$ or 1 , the maximum value of disfluency 151.8 when $\theta = 0.5625$; $n = 1$. For these values of n, there is hardly any saving in the sample size through the use of ro. 1. estimator.

A question that naturally arises 15 the equivalence between the variance of $\sqrt{n_1} \cdot (\tilde{\theta}_n - \theta)$ obtained by the usual Taylor's expansion and the varıance of the asymptotic distribution of $\sqrt{n1} \cdot (\theta_n - \theta)$. The asymptatic distribution is the formal Edgeworth expansion of the density of $\sqrt{n1}.(\tilde{\theta}_n - \theta)$. Since $\tilde{\theta}_n$ is a function of $(\mathrm{w}_1, \ldots, \mathrm{m}_k)$. were $w_1 =$ ln $(p_1 - \pi_1)$, and since we have used the multivariate Edgeworth expansion for the joint density of $(n_1, \ldots., w_{k-1})$, it is clear that the asymptotic variance is equal to the variance of the asymptotic distribution. Thas is similar to the observation made in Ghosh and Subramanyan (1974).

References

[1] Aithal, U.B. and Nagnur, B.N. Second order efficiency of the BAN estimators for the multinomial family. To appear in Communications in Statistics - Theory and Methods.

[2] Akahira, M. (1981). On asymptotic deficiency of estimators. Austral. J. Statist., 23, 67 -72

[3] Akahira, M. (1983). Asymptotic deficiency of the jackknife estimator. Austral. J. Statist., 25, 123-129.

[4] Akahira, M. and Takeuchi, K. (1981). Asymptotic Efficiency of Statistical Estimators

e-ISSN 2320 –7876 www.ijfans.org
Vol.11, Iss.9, Dec 2022

: Concepts and Higher Order Asymptotic Efficiency. Lecture Notes in Statistics 7. New York: Springer-Verlag.

[5] Albers, W., Bickel, P.J. and Van Zwet, W. (1976). Asymptotic expansions for the power of distribution free tests in one-sample problem. Ann. Statist., 4, 108-156.