MINIMUM CHI-SQUARED ESTIMATION USING INDEPENDENT STATISTICS

By A. D. Joffe

The University of New South Wales

0. Summary. In a multinomial situation with observed proportions Q_i ($i = 1, \cdots, r + 1$) and corresponding expected proportions p_i, it has been shown by Chapman [1] that for large samples

$$Y_i = \ln Q_i - \ln Q_{i+1}, \qquad i = 1, \cdots, r,$$

and Y_j are independent for $j \neq i - 1, i, i + 1$. In this note the efficiency obtained when estimating the parameters of a distribution from these mutually independent odd (or even) Y_i's is examined in the case of the geometric and Poisson distributions and it is shown that the resulting estimators are inefficient.

1. Introduction. Katti and Gurland [3] discuss applications of a minimum χ^2 approach to the problem of estimating the parameters of a distribution by way of a vector of statistics $t = (t_1, t_2, \cdots, t_k)$. If t is a consistent estimator of $\tau = (\tau_1, \tau_2, \cdots, \tau_k)$ then the required estimates are the ones that minimize

$$S = (t - \tau)M^{-1}(t - \tau)'$$

where M is the variance covariance matrix of the t_i.

In the case of a single parameter (θ) the asymptotic variance of such an estimator $\tilde{\theta}$ will be

$$\text{Var } \tilde{\theta} = [(\partial\tau/\partial\theta)M^{-1}(\partial\tau/\partial\theta)']^{-1}$$

where $\partial\tau/\partial\theta = (\partial\tau_1/\partial\theta, \partial\tau_2/\partial\theta, \cdots)$.

A numerical difficulty associated with this method would be that of inverting M as in general this need not be simple, although with the advent of the high speed computer this is no longer such a serious problem. If, however, all the t_i are independent then M will be a diagonal matrix and M^{-1} will be readily available.

In his paper on the truncated gamma distribution Chapman [1] has obtained such a set of statistics. He in fact suggests applying the above method to these independent statistics in the case of the truncated gamma distribution, where the frequencies in the class intervals are not all the same.

Briefly his results are as follows; if Q_i ($i = 1, \cdots, r + 1$) are observed proportions in a multinomial classification, then for large samples the variance co-

Received 1 December 1965; revised 11 July 1966.

variance matrix of $Y_i = \ln Q_i - \ln Q_{i+1}$ $(i = 1 \cdots r)$ is

$$M = n^{-1}\begin{bmatrix} (p_1^{-1}+p_2^{-1}) & -p_2^{-1} & 0 & \cdots & 0 \\ -p_2^{-1} & (p_2^{-1}+p_3^{-1}) & -p_3^{-1} & \cdots & 0 \\ 0 & -p_3^{-1} & (p_3^{-1}+p_4^{-1}) & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & (p_r^{-1}+p_{r+1}^{-1}) \end{bmatrix}$$

Thus Y_i and Y_j will be mutually independent as long as $j \neq i - 1, i, i + 1$ and consequently if one uses only the odd (or even) Y_i then the resulting variance covariance matrix will be a diagonal matrix with entries n^{-1} $(p_{2i-1}^{-1} + p_{2i}^{-1})$ and its inverse can be readily evaluated. Chapman mentions that using half the Y_i will lead to less efficient estimators than when using all the Y_i but does not pursue the matter any further. In this note the loss in efficiency due to using the even (or odd) Y_i only is investigated for the geometric and Poisson distributions and is shown in most instances to be considerably more than 50%—the figure one would obtain if half the observations were rejected by some random procedure. The efficiency obtained when using the odd Q_i of the geometric distribution is also evaluated and in certain cases this too is less than 50%.

The investigation is confined to distributions with a single parameter.

2. Asymptotic efficiency using all the Y_i. In order to assess the efficiency of estimators based on half the Y_i it is firstly necessary to know the efficiency of any estimators based on all the Y_i. When discussing transformed minimum χ^2 estimators Ferguson [2] indicates that the estimators obtained by that method will be BAN estimators and will therefore be asymptotically efficient relative to the standard maximum likelihood estimators.

Thus in the limiting case, using all the Y_i's results in the estimated parameter having the same variance as when using all the Q_i's and when investigating the variances obtained by using only half of the Y_i's it is adequate to compare the variance of the estimator to the variance of the corresponding maximum likelihood estimator.

Although the results above are all derived in terms of multinomial distributions with a finite number of cells the extension to an infinite number of intervals is simple enough as long as the probability in every cell is non zero and the extension will be used without further comment.

3. The geometric distribution. Define

$$p_i = (1 - \omega)^{i-1}\omega, \qquad i = 1, 2, \cdots, \infty$$

then the asymptotic variance of the maximum likelihood estimator $(\hat{\omega})$ can be shown to be

$$\text{Var } \hat{\omega} = \omega^2(1 - \omega)/n.$$

For the estimator ($\tilde{\omega}$) based on odd Y_i we have if

$$\eta_i = \ln p_i - \ln p_{i+1}$$

then

$$\eta_{2i-1} = -\ln (1 - \omega), \qquad i = 1, 2, \cdots, \infty,$$

and

$$(p_{2i-1}p_{2i})/(p_{2i-1} + p_{2i}) = (1 - \omega)^{2i-1}\omega/(2 - \omega).$$

Consequently the asymptotic variance of $\tilde{\omega}$ is

$$\begin{aligned} \text{Var } \tilde{\omega} &= [(\partial\eta/\partial\omega)(M^{-1})(\partial\eta/\partial\omega)']^{-1} \\ &= [(1 - \omega)^2/n][\textstyle\sum p_{2i-1}p_{2i}/(p_{2i-1} + p_{2i})]^{-1} \\ &= (1 - \omega)(2 - \omega)^2/n. \end{aligned}$$

The efficiency of $\tilde{\omega}$ relative to $\hat{\omega}$ is RE $= 100\omega^2/(2 - \omega)^2$ and values of this quantity are tabulated for selected values of ω in Table I.

As can be seen from Table I the efficiencies are very low for all but very large values of ω.

4. The Poisson distribution. Here $p_i = e^{-\lambda}\lambda^i/i!$, $i = 0, 1, 2, \cdots$, and the asymptotic variance of the maximum likelihood estimator ($\hat{\lambda}$) is Var $\hat{\lambda} = \lambda/n$. Using the even Y_i to obtain an alternative estimator of $\lambda(\tilde{\lambda})$ it can be shown that

$$\begin{aligned} \text{Var } \tilde{\lambda} &= (\lambda^2 e^{+\lambda}/n)[\textstyle\sum_{i=0}^{\infty} \lambda^{2(i+1)}/(2i + 1)!(\lambda + 2i + 1)]^{-1} \\ &\geqq (\lambda^2 e^{+\lambda}/n)[\textstyle\sum_{i=0}^{4\lambda} \lambda^{2(i+1)}/(2i + 1)!(\lambda + 2i + 1) + 16(\lambda^{4\lambda})/75(4\lambda)!]^{-1}. \end{aligned}$$

The above equation gives a lower bound to the variance of $\tilde{\lambda}$, which for most values of λ will be quite close to the true value.

TABLE I

Asymptotic efficiencies of $\tilde{\omega}$ relative to $\hat{\omega}$

ω	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
RE (%)	0	0	1	3	6	11	18	29	44	67	100

TABLE II

Asymptotic efficiencies of $\tilde{\lambda}$ relative to $\hat{\lambda}$

λ	1	2	3	4	5	6	7	8	9	10
RE (%)	23	11	8	6	5	4	3	3	3	2

TABLE III

Asymptotic efficiencies of $\tilde{\omega}_0$ relative to $\hat{\omega}$

ω	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
RE (%)	50	52	55	57	59	59	58	55	46	30	00

In Table II an upper bound is given for the efficiency of $\tilde{\lambda}$ relative to $\hat{\lambda}$ for selected values of λ and here too the efficiencies are very low.

The method discussed in Table II, although appealing because of the ease of manipulation associated with it, is of disappointingly low efficiency in the cases examined.

5. Asymptotic efficiency using all the odd Q_i. In this section the effect of using the odd Q_i is considered for the geometric distribution. If X is a random variable such that

$$\Pr(X = i) = (1 - \omega)^{i-1}\omega$$

then

$$\Pr(X = 2i - 1 \mid X \text{ is odd}) = (2 - \omega)\omega(1 - \omega)^{2i-2}.$$

The parameter ω can be estimated from the odd X's, which is equivalent to using the odd Q_i, and the asymptotic variance of the conditional maximum likelihood estimator ($\tilde{\omega}_0$) is

$$\text{Var}(\tilde{\omega}_0) = \omega^2(2 - \omega)^3/4n.$$

The efficiency of $\tilde{\omega}_0$ relative to $\hat{\omega}$ (the full maximum likelihood estimator) is RE $= 4(1 - \omega)/(2 - \omega)^3$ and this is tabulated for different values of ω in Table III.

The difference between these results and those of Table I is quite striking. The explanation is that whereas using all the Q_i is asymptotically equivalent to using all the Y_i, using the odd (or even) Q_i is not equivalent to using the odd (or even) Y_i. It is of interest to note that here too the relative efficiency of the estimator is less than 50% for certain values of the parameter.

6. Acknowledgments. I am indebted to Mr. A. G. Elliott for the assistance that he has given me in the preparation of the above manuscript, and to the referee for indicating the result that occurs when using the odd Q_i for estimating the parameter of the geometric distribution.

REFERENCES

[1] CHAPMAN, D. G. (1956). Estimating the parameters of a truncated Gamma distribution. *Ann. Math. Statist.* **27** 498–506.

[2] FERGUSON, T. S. (1958). A method of generating best asymptotically normal estimates with application to the estimation of bacterial densities. *Ann. Math. Statist.* **29** 1046–1062.

[3] KATTI, S. K. and GURLAND, J. (1962). Some methods of estimation for the Poisson binomial distribution. *Biometrics* **18** 42–51.