Journal of the Korean
Statistical Society
Vol. 20, No. 2, 1991

On Estimating the Distributional Parameter and the Complete Sample Size from Incomplete Samples†

Sung-chil Yeo*

ABSTRACT

Given a random sample of size N(unknown) with density $f(x \mid \theta)$, suppose that only n observations which lie outside a region R are recorded. On the basis of n observations, the Bayes estimators of θ and N are considered and their asymptotic expansions are developed to compare their second order asymptotic properties with those of the maximum likelihood estimators and the Bayes modal estimators. Corrections to bias and median bias of these estimators are made. An example is given to illustrate the results obtained.

1. Introduction

In dealing with many lifetime data, obervable data are often restricted to a part of the total(complete) potential data. Among N potential observations, only n remain observable by the result of the restricting process, while (N-n) are eliminated. The set of such incomplete data is usually referred to as censored when N is known, and as truncated otherwise.

In this paper, we are concerned with the estimation of distributional parameters and sample sizes when the sample is truncated. For example, the following life testing situation is considered by Blumenthal and Marcus (1975) : Suppose that M items with failure density $f(x|\theta)$ are put on life test, and that out of these M items there are certain unknown N items with a particular defect identifiable only after the item fails. If the lifetime of an item with this particular defect is the variable of interest, then the sample is truncated in that N, the number of missing observations of lifetime greater than the burn-in or testing period, is unknown. An interesting problem in such situation is to estimate the number of remaining defective items of a particular type after an initial burn-in period.

The problem of estimating θ and N from trucated samples were dealt with in many articles for various cases. Assuming the exponential distribution for lifetime, Blumenthal and Marcus (1975) gave results for the estimation of θ and N and also discussed the second order asymptotic

† This research is supported by Korea Research Foundation, 1989−1990.
* Department of Applied Statistics, Kon-Kuk University, Seoul, 133-701, Korea.

properties of the estimators considered. Dahiya and Gross(1973) discussed the problem of estimating θ and N from a truncated Poisson sample, where all the observations assuming the zero are missing. Blumenthal, Dahiya, and Gross(1974) studied the same problem with emphasis on the second order properties of the estimators of θ and N. For truncated continuous samples, the articles on estimating N include Johnson (1962) and Marcus and Blumenthal (1974) who limited their attention to the completely known distributions. In the discrete case, other articles on estimating θ and N are found in Sanathanan (1972) for the multinomial sample, in Feldman and Fox (1968), Draper and Guttman (1971), and Blumenthal and Dahiya (1981) for the binomial distribution, in Blumenthal and Sanathanan (1980) for the inverse binomial sampling. For the case of general distribution, Sanathanan (1977) gave a result for the asymptotic distribution of the maximum likelihood estimators (m.l.e.'s) of θ and N from truncated samples. The asymptotic properties of the m.l.e.'s and Bayes modal estimators of θ and N were further studied by Blumenthal (1977) for comparing asymptotic biases of estimators, by Watson and Blumenthal (1980) for controlling mean squared error, and by Blumenthal (1982) for bias reduction.

All of the above work were concerned with the m.l.e.'s of θ and N and their modifications, especially Bayes modal estimators. Asymptotic expansions have been obtained and second order asymptotic properties such as bias and mean squared error (m.s.e) have been studied in a restricted class of Bayes modal estimators. However, the Bayes estimators which minimize the expected loss, especially the posterior means which result from the squared error loss function, have not been examined yet. For complete samples, Gusev (1975, 1976) obtained the results about asymptotic expansions for the Bayes estimator.

In this article, we consider the Bayes estimators of θ and N on the basis of truncated samples and examine their asymptotic properties with the m.l.e.'s and the Bayes modal estimators. In Section 2, we briefly review the previous results about asymptotic expansions of the m.l.e., the Bayes modal estimator as well as the Bayes estimator of the distributional parameter for the case of complete samples. In Section 3, extending the results given in Section 2, we develope the asymptotic expansions of these estimators for θ and N from truncated samples. From asymptotic expansions, we give expressions of the bias corrections and the median bias corrections of these estimators. In Section 4, we present an example to illustrate the results given in Section 3. Finally, in Section 5, we give some concluding remarks.

2. Complete Samples

Let $X_1, \cdots, X_N$ be a random sample of size N with density $f(x|\theta)$ (with respect to (w.r.t.) a σ-finite meausre μ), where the values of x are in a sample space $\mathfrak{X}$ and θ is in a real valued parameter space Θ. Consider some available estimators $\hat{\theta}$'s for θ. If as is often the case, $\hat{\theta}$'s are asymptotically equivalent in the sense of usual limitng distribution theory, we may want to know more details about the asymptotic properties such as the asymptotic biases and mean squared errors and to find the corrections to the biases or median biases for $\hat{\theta}$'s. all of these properties can be resolved by the asymptotic expansions such as expansions of random variables, moments, and distributions. For estimators of θ, we consider the m.l.e., the Bayes modal estimator (or the "modified" m.l.e.), and the Bayes estimator. For Bayes modal and Bayes estimators, we assume that θ has the prior density function $\pi(\theta)$. Then the likelihood function is given by

$$l(\mathbf{x}|\theta)=\prod_{i=1}^{N} f(x_i|\theta) \tag{2.1}$$

and the posterior likelihood is expressed as

$$h(\theta|\mathbf{x}) \propto l_m(\mathbf{x}|\theta) \tag{2.2}$$

where $$l_m(\mathbf{x}|\theta)=\pi(\theta)\ l(\mathbf{x}|\theta) \tag{2.3}$$

We assume that $f(x|\theta)$ and $\pi(\theta)$ are smooth functions which allow continuous differentiations w.r.t. θ. Let $\hat{\theta}_0$, $\hat{\theta}_m$, and $\hat{\theta}_B$ denote the m.l.e., the modified m.l.e., and the Bayes estimator of θ respectively. Then $\hat{\theta}_0$, and $\hat{\theta}_m$ are obtained by solving

$$0=\frac{\partial \log l(\mathbf{x}|\theta)}{\partial\theta}\bigg|_{\theta=\hat{\theta}_0}=\sum_{i=1}^{N} S(x_i|\hat{\theta}_0) \tag{2.4}$$

and

$$0=\frac{\partial \log l_m(\mathbf{x}|\theta)}{\partial\theta}\bigg|_{\theta=\hat{\theta}_m}=\sum_{i=1}^{N} S(x_i|\hat{\theta}_m)+\zeta(\hat{\theta}_m) \tag{2.5}$$

respectively, where

$$S(x|\theta)=\frac{\partial \log f(x|\theta)}{\partial\theta}=\frac{f'(x|\theta)}{f(x|\theta)} \tag{2.6}$$

and $$\zeta(\theta)=\frac{\partial \log\pi(\theta)}{\partial\theta}=\frac{\pi'(\theta)}{\pi(\theta)} \tag{2.7}$$

and prime notations indicate differentiation w.r.t. θ.

On the other hand, $\hat{\theta}_B$ is obtained by

$$\hat{\theta}_B=\frac{\int\theta\pi(\theta)\ l(\mathbf{x}|\theta)\ d\theta}{\int\pi(\theta)\ l(\mathbf{x}|\theta)\ d\theta} \tag{2.8}$$

where the integral is taken over Θ.

First, consider a stochastic expansion of the form

$$\hat{\theta}=\theta+\frac{1}{\sqrt{N}}A+\frac{1}{N}B+O(N^{-3/2}) \tag{2.9}$$

where A and B are polynomials in certain sums of i.i.d. random variables, which will be specified on later.

In order to express the coefficients A and B for $\hat{\theta}_0$, $\hat{\theta}_m$, and $\hat{\theta}_B$, we introduce the following notations. Let

$$L_{ijk}=E_\theta[S^i(S')^j(S'')^k]$$
$$=\int S^i(S')^j(S'')^k\ f \tag{2.10}$$

where the integral is taken over $\mathfrak{X}$, and the arguments x and $d\mu(x)$ are omitted from the notation unless needed to avoid ambiguity. Trailng zero's are omitted so that $L_{ij0}=L_{ij}$, etc. Similarly, let

$$Z_{ijk}=\frac{1}{\sqrt{N}}\left[\sum_{p=1}^{N} S_p^i(S_p^{'})^j(S_p^{''})^k-NL^{ijk}\right] \tag{2.11}$$

where $S_p=S(X_p|\theta)$, $S_p^{'}=S^{'}(X_p|\theta)$, etc.

Usually, we encounter only Z_1, Z_{01}, and Z_{001}. From elementary calculation, we see that

$$E(Z_{ijk})=0 \text{ and } Var(Z_{ijk})=V_{ijk},$$

where

$$V_{ijk}=L_{2i\ 2j\ 2k}-(L_{ijk})^2 \tag{2.12}$$

As Pfanzagl (1973) indicated, under suitatble regularity conditins such as interchange of integrals and derivatives, we see that

$$L_1=0,\ L_{01}+L_2=0,\ L_{001}+3L_{11}+L_3=0,$$

$$L_2^{'}=2L_{11}+L_3,\ L_{11}^{'}=L_{21}+L_{101}+L_{02},\ \text{etc.} \tag{2.13}$$

Now, the coefficeients A and B in (2.9) for $\hat{\theta}_0$, $\hat{\theta}_m$, and $\hat{\theta}_B$ are

$$A_0=Z_1/L_2$$

$$A_B=Z_m=A_0$$

$$B_0=L_2^{-2}\{Z_1Z_{01}+Z_1^2L_2^{-1}L_{001}/2\} \tag{2.14}$$

$$B_m=B_0+\nu_m \qquad B_B=B_0+\nu_m+\nu_B$$

where $\zeta=\zeta(\theta)$

$$\nu_m=\nu_m(\theta)=\zeta/L_2 \tag{2.15}$$

and $\nu_B=\nu_B(\theta)=L_{001}/2L_2^2$.

We note that it requires a variety of regularity conditions and a great deal of care in bounding remainder terms to justify that $\hat{\theta}$ allows the form (2.9) with (2.14) for $\hat{\theta}$ and to justify the formal manipulation to go from (2.9) to moments expansion or Edgeworth expansion. Our purpose is to explore what can be learned in a formal way about asymptotic properties of an estimator from asymptotic expansions rather than to examine the mathematical justification of these manipulations. References will be given for the latter.

Note that from (2.9) with (2.14), we see that $\hat{\theta}-\theta=Z_1/L_2\sqrt{N}$ which converges stochastically to zero, hence consistency follows immediately, and $\sqrt{N}(\hat{\theta}-\theta)=Z_1/L_2+B/\sqrt{N}$. Since B has a legitimate limiting distribution, $B/\sqrt{N}\xrightarrow{p}0$ and the central limit theorem applied to Z_1 show that $\sqrt{N}(\hat{\theta}-\theta)$ is asymptotically normal with mean 0 and variance L_2^{-1}. Hence $\hat{\theta}_0$, $\hat{\theta}_m$, and $\hat{\theta}_B$ are all asymptotically equivalent in the first order sense. We note that the stochastic expansion yields a limiting distribution and stochastic covergence results as an easy by-product.

From (2.14), we find that

$$\hat{\theta}_m-\hat{\theta}_0=\nu_m/N+O(N^{-3/2}) \tag{2.16}$$

$$\hat{\theta}_B-\hat{\theta}_m=\nu_B/N+O(N^{-3/2}) \tag{2.17}$$

$$\hat{\theta}_B-\hat{\theta}_0=(\nu_m+\nu_B)/N+O(N^{-3/2}) \tag{2.18}$$

Thus, for asymptotic approximations, we may regard $\hat{\theta}_m$ and $\hat{\theta}_B$ as adjusted m.l.e.'s given by

$$\hat{\theta}_m \approx \hat{\theta}_o+\hat{\nu}_m/N$$

$$\hat{\theta}_B \approx \hat{\theta}_o+(\hat{\nu}_m+\hat{\nu}_B)/N \tag{2.19}$$

where $\hat{\nu}=\nu_m(\hat{\theta}_o)$ and $\nu_B(\hat{\theta}_o)$. (2.20)

Next, from (2.9) the form of the moments expansion for $\hat{\theta}$ is expressed as

$$E(\hat{\theta})=\theta+b/N+O(N^{-3/2}) \tag{2.21}$$

where b=E(B) can be regarded as the asymptotic bias of $\hat{\theta}$. From (2.14) the bias terms b's in (2.21) for $\hat{\theta}_o$, $\hat{\theta}_m$, and $\hat{\theta}_B$ are given by

$$\begin{aligned} b_o&=E(B_o)=L_2^{-2}(L_{11}+L_{001}/2)\\ b_m&=E(B_m)=b_o+\nu_m\\ b_B&=E(B_B)=b_0+\nu_m+\nu_B. \end{aligned} \tag{2.22}$$

We notice that $E(A_o)=0$. From (2.22), the bias-corrected versions of $\hat{\theta}_o$, $\hat{\theta}_m$, and $\hat{\theta}_B$ are given by

$$\begin{aligned} \tilde{\theta}_o&=\hat{\theta}_o-\hat{b}_0/N\\ \tilde{\theta}_m&=\hat{\theta}_m-(\hat{b}_o+\hat{\nu}_m)/N\\ \tilde{\theta}_B&=\hat{\theta}_B-(\hat{b}_o+\hat{\nu}_m+\hat{\nu}_B)/N, \end{aligned} \tag{2.23}$$

where $\hat{b}_o=b_o(\hat{\theta}_o)$. We note that these bias-corrected estimators will be unbiased up to order $O(N^{-1})$.

On the other hand, we consider the Edgeworth expansion for the distribution of $\hat{\theta}$ such that

$$P(\sqrt{NL_2}(\hat{\theta}-\theta)\leq x)=\Phi(x)+\phi(x)D(x)/\sqrt{N}+O(N^{-1}), \tag{2.24}$$

where $\Phi(x)$ is the c.d.f. of the standard normal distribution,
$\phi(x)$ is the p.d.f. of the standard normal distribution.
$D(x)$ is the polynomial of degree 2 such that $D(x)=d+ex^2$.
The coefficients d and e in $D(x)$ for $\hat{\theta}_o$, $\hat{\theta}_m$, and $\hat{\theta}_B$ are given by

$$\begin{aligned} d_o &= \nu_o\sqrt{L_2} \qquad e_o=(L_3-L_{001})/6L_2^{3/2} \\ d_m &= d_o-\nu_m\sqrt{L_2} \\ d_B &= d_o-(\nu_m+\nu_B)\sqrt{L_2} \\ e_B &= e_m=e_o, \end{aligned} \tag{2.25}$$

where

$$\nu_o=\nu_o(\theta)=L_3/6L_2^2. \tag{2.26}$$

Pfanzagl (1973) defined an estimator $\hat{\theta}$ as being median unbiased of order $O(N^{-r})$ if $|F(0)-1/2|$ is $O(N^{-r})$.

Thus, the median unbiased versions of $\hat{\theta}_o$, $\hat{\theta}_m$, and $\hat{\theta}_n$ are given by

$$\begin{aligned} \bar{\theta}_o &= \hat{\theta}_o+\hat{\nu}_o/N \\ \bar{\theta}_m &= \hat{\theta}_m+(\hat{\nu}_o-\hat{\nu}_m)/N \\ \bar{\theta}_B &= \hat{\theta}_B+(\hat{\nu}_o-\hat{\nu}_m-\hat{\nu}_B)/N, \end{aligned} \tag{2.27}$$

where $\hat{\nu}_o=\nu_o(\hat{\theta}_o)$. We note that these median unbiased estimators will be unbiased up to order $O(N^{-1})$.

We have been concerned here only with the expansions of order $O(N^{-1})$ to correct the asymtotic biases or median biases of the estimators considered in this paper. For general expansions of order $O(N^{-r/2})$ and for regularity conditions which allow these expansions with rigorous arguments, the relevant articles, for example, are found in Chibisov(1972a, 1972b, 1973) Gusev(1975, 1976), and Pfanzagl (1973). For less mathematically sophisticated arguments about the asymptotic expansions and approximations of the Bayes estimators, the references are found in Ghosh and Subramanyam(1974) and Lindley (1961, 1980), etc..

3. Truncated Samples

Our view of truncated samples was described in Section 1. Formally, given a random sample $X_1, X_2, \cdots, X_N$, of size N with density $f(x|\theta)$, where $x\in\mathfrak{X}$ and $\theta\in\Theta$, suppose that only the n observations in a region $\bar{R}(=\mathfrak{X}-R)$ are recorded, and the remaining $(N-n)$ are lost. However, we do not even record the fact that these (N-n) observations were in R, hence N is an unknown parameter. On the basis of this truncated observations, we wish to estimate both θ and N and to study the asymptotic properties of the estimators for θ and N.

3.1 Estimation of θ and N

If we assume for notational convenience that $x_1, x_2, \cdots, x_n$ are in $\bar{R}$, and $x_{n+1}, \cdots, x_N$ are in R, we see that the likelihood function is expressed as

$$l(\mathbf{x}, n \mid \theta, N) = l_1(n \mid \theta, N) l_2(\mathbf{x} \mid \theta, n), \tag{3.1}$$

where

$$l_1(n \mid \theta, N) = \binom{N}{n} q^n p^{N-n} \tag{3.2}$$

and

$$l_2(\mathbf{x} \mid \theta, n) = \prod_{i=1}^{n} g(x_i \mid \theta), \tag{3.3}$$

with

$$p = p(\theta) = P(X \in R), \; q = q(\theta) = 1 - p, \text{ and}$$

$$g(x \mid \theta) = \begin{cases} f(x \mid \theta)/q, & x \in \bar{R} \\ 0 \quad , & x \in R \end{cases} \tag{3.4}$$

We note that (3.1) expresses the likelihood as the product of the likelihood for n and the conditional likelihood of the observed x's given n, and (3.2) expresses the binomial distribution for n representing the number of observations falling in $\bar{R}$. To obtain the Bayes modal and the Bayes estimators of θ and N, we assume that θ has the prior density $\pi(\theta)$, $\theta \in \Theta$ and N has the uniform prior given by $\psi(N) = 1/M$, $1 \le N \le M$, if there is an upper bound on N (otherwise, we use the improper prior $\psi(N) = 1$ for all N). As Blumenthal(1977, 1982), Draper and Guttman (1971), and others considered, we assume that θ and N are independent apriori. The posterior likelihood is then given by

$$h(\theta, N \mid \mathbf{x}, n) \propto l_m(\mathbf{x}, n \mid \theta, N), \tag{3.5}$$

where

$$l_m(\mathbf{x}, n \mid \theta, N) = l_1(n \mid \theta, N) \; l_{2.m}(\mathbf{x} \mid \theta, N) \tag{3.6}$$

and

$$l_{2.m}(\mathbf{x} \mid \theta, n) = \pi(\theta) l_2(\mathbf{x} \mid \theta, n) \tag{3.7}$$

One method of estimating θ and N is given by maximizing first either the conditional likelihood $l_2(\mathbf{x} \mid \theta, n)$ or the modified conditional likelihood $l_{2,m}(\mathbf{x} \mid \theta, n)$ w.r.t. θ and then maximizing $l_1(n \mid \theta, N)$ w.r.t. N. The resulting estimators are referred to either the conditinal m.l.e.'s $\hat{\theta}_c$ and $\hat{N}_c$ or the modified conditional m.l.e.'s $\theta_{c,m}$ and $\hat{N}_{c,m}$. Alternatively, we can maximize either the likelihood $l(\mathbf{x}, n \mid \theta, N)$ or the modified likelihood $l_m(\mathbf{x}, n \mid \theta, N)$ w.r.t. θ and N, simultaneously. This gives rise to either the unconditional m.l.e.'s θ_u and $\hat{N}_u$ or the modified unconditional m.l.e.'s $\hat{\theta}_{u,m}$ and $\hat{N}_{u,m}$. Thus, $\hat{\theta}_c$, $\hat{\theta}_{c,m}$, $\hat{\theta}_u$, and $\hat{\theta}_{u,m}$ are obtained by solving

$$O = \sum_{i=1}^{n} \tilde{S}(x_i \mid \hat{\theta}_c) \tag{3.8}$$

$$O=\sum_{i=1}^{n}\tilde{S}(x_i\mid\hat{\theta}_{c,m})+\zeta(\hat{\theta}_{c,m}) \tag{3.9}$$

$$O=\sum_{i=1}^{n}\tilde{S}(x_i\mid\hat{\theta}_u)+U(\hat{\theta}_u)(\hat{N}-\frac{n}{q(\hat{\theta}_u)}) \tag{3.10}$$

$$O=\sum_{i=1}^{n}\tilde{S}(x_i\mid\hat{\theta}_{u,m})+U(\hat{\theta}_{u,m})(\hat{N}-\frac{n}{q(\hat{\theta}_{u,m})})+\zeta(\hat{\theta}_{u,m}) \tag{3.11}$$

where

$$U(\theta)=\frac{p'}{p} \tag{3.12}$$

$$\tilde{S}(x\mid\theta)=\frac{g'(x\mid\theta)}{g(x\mid\theta)}=\frac{f'(x\mid\theta)}{f(x\mid\theta)}+\frac{p'}{q} \tag{3.13}$$

and

$$\hat{N}=N(\hat{\theta})=[n/\hat{q}],\ \hat{q}=q(\hat{\theta}), \tag{3.14}$$

where [·] is the greatest integer function.

On the other hand, the Bayes estimators of θ and N are given by

$$\hat{\theta}_B=\frac{\int\theta\pi(\theta)\ l_2(\mathbf{x}\mid\theta,n)\ (\Sigma l_1(n\mid\theta,N))\ d\theta}{\int\pi(\theta)\ l_2(\mathbf{x}\mid\theta,n)\ (\Sigma l_1(n\mid\theta,N))\ d\theta} \tag{3.15}$$

and

$$\hat{N}_B=\frac{\int\pi(\theta)\ l_2(\mathbf{x}\mid\theta,n)\ (\Sigma\ Nl_1(n\mid\theta,N))\ d\theta}{\int\pi(\theta)\ l_2(\mathbf{x}\mid\theta,n)\ (\Sigma\ l_1(n\mid\theta,N))\ d\theta} \tag{3.16}$$

where the integral is taken over Θ and the summation is taken over the range of N, and l_1 and l_2 are given by (3.2) and (3.3), respectively. We note that $\hat{\theta}_B$ and $\hat{N}_B$ will not have simple closed forms and some approximations will be used on later.

For the purpose of developing asymptotic expansions, Blumenthal (1977, 1982) treated N as a continuous parameter, and hence $\hat{N}$ is obtained by solving

$$O=\partial\log l_1(\mathbf{x},n\mid\theta,N)/\partial N$$
$$=\log p(\hat{\theta})+\sum_{i=\hat{N}-n+1}^{N}1/i \tag{3.17}$$

Using the Euler-McLaurin formula for $\Sigma(1/i)$ and expanding $\log(1+x)$ in a Taylor series for $0<x<1$, Blumenthal (1977) obtained $\hat{N}=n/\hat{q}-1/2+O(n^{-1})$. For asymptotic purpose, we take the definition of $\hat{N}$ as

$$\hat{N}=n/\hat{q}-1/2 \tag{3.18}$$

for any $\hat{\theta}$. We note that (3.18) is simply a linear approximation to $n/\hat{q}$, whch is the discrete solution to N. Using (3.18), we see that (3.10) and (3.11) can be expressed as

$$O=\sum_{i=1}^{n}\tilde{S}(x_i\,|\,\hat{\theta}_u)-U(\hat{\theta}_u)/2 \qquad (3.19)$$

and

$$O=\sum_{i=1}^{n}\tilde{S}(x_i\,|\,\hat{\theta}_{u,m})-U(\hat{\theta}_{u,m})/2+\zeta(\hat{\theta}_{u,m}) \qquad (3.20)$$

3.2 Asymptotic Expansions for $\hat{\theta}$

First, we consider the conditional stochastic expansions of $\hat{\theta}$, given n, of the form

$$\hat{\theta}=\theta+\frac{1}{\sqrt{n}}A+\frac{1}{n}B+O(n^{-3/2}). \qquad (3.21)$$

In order to develop asymptotic expansions from truncated samples, we define $\tilde{L}_{ijk}$ and $\tilde{Z}_{ijk}$ as (2.10) and (2.11) with $\tilde{S}$, n, and g in place of S, N and f, respectively, where the integral is taken over $\bar{R}$, e.g.

$$\tilde{L}_2=\int\tilde{S}^2g\;;\;\tilde{Z}_1=\sum_{p=1}^{n}\tilde{S}_p/\sqrt{n},\ \text{etc..}$$

We denote the conditional expectation, given n, by E^n, and reserve the symbol E for the unconditional expectation.

We note that the regularity conditions imposed on $f(x\,|\,\theta)$ such as interchange of integrals and derivatives will continue to hold for $g(x\,|\,\theta)$ which is defined on a smaller sample space $\bar{R}$. Thus, all relations given in Section 2 will be applied with the appropriate substitutions. For $\tilde{L}_{ijk}$, we still have the same relations as (2.13), e.g., $\tilde{L}_1=0$, $\tilde{L}_{01}+\tilde{L}_{02}=0$, etc..

Thus, from (2.14), the coefficients A and B in (3.21) for $\hat{\theta}_c$ and $\hat{\theta}_{c,m}$ are

$$A_{c,m}=A_c=\tilde{Z}_1/\tilde{L}_2$$

$$B_c=\tilde{L}_2^{-2}(\tilde{Z}_1\tilde{Z}_{01}+\tilde{Z}_1^2\tilde{L}_2^{-1}\tilde{L}_{001}/2) \qquad (3.22)$$

$$B_{c,m}=\tilde{B}_c+\nu_m,$$

where

$$\nu_m=\nu_m(\tilde{\theta})=\zeta/\tilde{L}_2. \qquad (3.23)$$

For convenience, we take the same notations ν's used in Section 2.

Using (3.19) and (3.20) in place of (2.5), from the expressions of A_m and B_m for $\hat{\theta}_m$ in (2.14), we see that the coefficients A and B in (3.21) for $\hat{\theta}_u$ and $\hat{\theta}_{u,m}$ are

$$A_{u,m}=A_u=\tilde{Z}_1/\tilde{L}_2$$

$$B_u=B_c+\nu_u \qquad (3.24)$$

$$B_{u,m}=B_c+\nu_u+\nu_m,$$

where

$$\nu_u = \nu_u(\theta) = -U(\theta)/2\tilde{L}_2. \tag{3.25}$$

For stochastic expansion of $\hat{\theta}_B$, we note that $\hat{\theta}_{u,m}$ is obtained by solving (3.20) which results from the modified likelihood (3.6) and (3.18). Thus, from (2.17), we obtain

$$\hat{\theta}_B - \hat{\theta}_{u,m} = \nu_B/n + O(n^{-3/2}), \tag{3.26}$$

where

$$\nu_B = \nu_B(\theta) = \tilde{L}_{001}/2\tilde{L}_2^2. \tag{3.27}$$

Hence, the coefficients A_B and B_B in (3.21) for θ_B are

$$A_B = \tilde{Z}_1/\tilde{L}_2, \qquad B_B = B_{u,m} + \nu_B. \tag{3.28}$$

From (3.22), (3.24), and (3.28), we find that

$$\begin{aligned} \hat{\theta}_{c,m} - \hat{\theta}_c &= \nu_m/n + O(n^{-3/2}) \\ \hat{\theta}_u - \hat{\theta}_c &= \nu_u/n + O(n^{-3/2}) \\ \hat{\theta}_{u,m} - \hat{\theta}_c &= (\nu_u + \nu_m)/n + O(n^{-3/2}) \\ \hat{\theta}_B - \hat{\theta}_c &= (\nu_u + \nu_m + \nu_B)/n + O(n^{-3/2}). \end{aligned} \tag{3.29}$$

Thus, for asymptotic approximations, we may regard $\hat{\theta}_{c,m}$, $\hat{\theta}_u$, $\hat{\theta}_{u,m}$, and $\hat{\theta}_B$ as adjusted estimators of $\hat{\theta}_c$ given by

$$\begin{aligned} \hat{\theta}_{c,m} &\approx \hat{\theta}_c + \hat{\nu}_u/n \\ \hat{\theta}_u &\approx \hat{\theta}_c + \hat{\nu}_u/n \\ \hat{\theta}_{u,m} &\approx \hat{\theta}_c + (\hat{\nu}_u + \hat{\nu}_m)/n \\ \hat{\theta}_B &\approx \hat{\theta}_c + (\hat{\nu}_u + \hat{\nu}_m + \hat{\nu}_B)/n, \end{aligned} \tag{3.30}$$

where

$$\hat{\nu}_m = \nu_m(\hat{\theta}_c), \quad \hat{\nu}_u = \nu_u(\hat{\theta}_c), \text{ and } \hat{\nu}_B = \nu_B(\hat{\theta}_c)$$

Next, from (3.21), the form of the moments expansion for $\hat{\theta}$ is expressed as

$$E^n(\hat{\theta}) = \theta + b/n + O(n^{-3/2}), \tag{3.31}$$

where $b = E^n(B)$ is the conditional asymptotic bias of $\hat{\theta}$, given n. From (3.22), (3.24) and (3.28) the conditinal bias terms b's in (3.31) for $\hat{\theta}$'s are

$$
\begin{aligned}
&b_c = E^n(B_c) = \tilde{L}_2^{-2}(\tilde{L}_{11} + \tilde{L}_{001}/2) \\
&b_{c.m} = E^n(B_{c,m}) = b_c + \nu_m \\
&b_u = E^n(B_u) = b_c + \nu_u \qquad (3.32) \\
&b_{u.m} = E^n(B_{u,m}) = b_c + \nu_u + \nu_m \\
&b_B = E^n(B_B) = b_c + \nu_u + \nu_m + \nu_B.
\end{aligned}
$$

We notice that $E(A_c) = 0$. From (3.32), the bias-corrected versions of $\hat{\theta}$'s are

$$
\begin{aligned}
&\tilde{\theta}_c = \hat{\theta}_c - \hat{b}/n \\
&\tilde{\theta}_{c,m} = \hat{\theta}_{c.m} - (\hat{b}_c + \hat{\nu}_m)/n \\
&\tilde{\theta}_u = \hat{\theta}_u - (\hat{b}_c + \hat{\nu}_u)/n \qquad (3.33) \\
&\tilde{\theta}_{u,m} = \hat{\theta}_{u.m} - (\hat{b}_c + \hat{\nu}_u + \hat{\nu}_m)/n \\
&\tilde{\theta}_B = \hat{\theta}_B - (\hat{b}_c + \hat{\nu}_u + \hat{\nu}_m + \hat{\nu}_B)/n,
\end{aligned}
$$

where

$$\hat{b}_c = b_c(\hat{\theta}_c).$$

On the other hand, for the conditional Edgeworth expansion of the distribution of $\hat{\theta}$, given n, we rewrite (2.24) as

$$P_\theta(\sqrt{n\tilde{L}_2}(\hat{\theta} - \theta) \leq x \mid n) = \Phi(x) + \frac{1}{\sqrt{n}}\phi(x)D(x) + O(n^{-1}), \qquad (3.34)$$

where $\Phi(x)$, $\phi(x)$, and $D(x)$ are defined as (2.24). Using the results of (2.25) with additional term ν_u, we obtain the coefficients d and e in $D(x)$ for $\hat{\theta}$'s as

$$d_c = \nu_c\sqrt{\tilde{L}_2}. \quad e_c = (\tilde{L}_3 - \tilde{L}_{001})/6\tilde{L}_2^{3/2} \qquad (3.35)$$

$$
\begin{aligned}
&d_{c,m} = d_c - \nu_m\sqrt{\tilde{L}_2} \\
&d_u = d_c - \nu_u\sqrt{\tilde{L}_2} \\
&d_{u,m} = d_c - (\nu_U + \nu_m)\sqrt{\tilde{L}_2} \qquad (3.36) \\
&d_B = d_c - (\nu_u + \nu_m + \nu_B)\sqrt{\tilde{L}_2} \\
&e_B = e_{u,m} = e_u = e_{c,m} = e_c,
\end{aligned}
$$

where

$$\nu_c = \nu_c(\theta) = \tilde{L}_3/6\tilde{L}_2^2. \qquad (3.37)$$

Thus, the median unbiased versions of $\hat{\theta}$'s are

$$\begin{aligned}
\bar{\theta}_c &= \hat{\theta}_c + \hat{\nu}_c/n \\
\bar{\theta}_{c,m} &= \hat{\theta}_{c,m} + (\hat{\nu}_c - \hat{\nu}_m)/n \\
\bar{\theta}_u &= \hat{\theta}_u + (\hat{\nu}_c - \hat{\nu}_u)/n \\
\bar{\theta}_{u,m} &= \hat{\theta}_{u,m} + (\hat{\nu}_c - \hat{\nu}_u - \hat{\nu}_m)/n \\
\bar{\theta}_B &= \hat{\theta}_B + (\hat{\nu}_c - \hat{\nu}_u - \hat{\nu}_m + \hat{\nu}_B)/n
\end{aligned} \tag{3.38}$$

where $\hat{\nu}_c = \nu_c(\hat{\theta}_c)$.

Now, we consider the unconditional expansions for $\hat{\theta}$. To express the unconditional stochastic expansion of $\hat{\theta}$ in terms of N, namely

$$\hat{\theta} = \theta + \frac{1}{\sqrt{N}}\tilde{A} + \frac{1}{N}\tilde{B} + O(N^{-3/2}) \tag{3.39}$$

we use the standardized variable Y given by

$$Y = \frac{n - Nq}{\sqrt{Npq}} \tag{3.40}$$

which is asymptotically normal. Substituing (3.40) for n in (3.21), expanding and collecting terms, we see that

$$\tilde{A} = A/\sqrt{q} = \tilde{Z}_1/\tilde{L}_2\sqrt{q}, \quad \tilde{B} = q^{-1}(B - AY\sqrt{p}/2) \tag{3.41}$$

Thus, the coefficients $\tilde{A}$ and $\tilde{B}$ in (3.38) for $\hat{\theta}_c$, $\hat{\theta}_{c,m}$, $\hat{\theta}_u$, $\hat{\theta}_{u,m}$, and $\hat{\theta}_B$ can be obtained by (3.41) with A and B given in (3.22), (3.24) and (3.28), respectively. From (3.39) with (3.41), we see that $\sqrt{N}(\hat{\theta} - \theta)$ is asymptotically normal with mean 0 and variance $(q\tilde{L}_2)^{-1}$.

The unconditional moments expansion for $\hat{\theta}$ in terms of N, has the form

$$E(\hat{\theta}) = \theta + \tilde{b}/N + O(N^{-3/2}) \tag{3.42}$$

where $\tilde{b} = E(\hat{B})$ is the unconditional asymptotic bias of $\hat{\theta}$. Since Y is uncorrelated with A and B, we see that

$$\tilde{b} = b/q, \tag{3.43}$$

where b is given by (3.32) for $\hat{\theta}$'s.

On the other hand, from (3.34), for the unconditional form of the Edgeworth expansion in terms of N, Blumenthal(1982) obtained

$$P(\sqrt{Nq\tilde{L}_2}(\hat{\theta}_c - \theta) \le x) = \Phi(x) + \frac{1}{\sqrt{Nq}}\phi(x)D_c(x) \tag{3.44}$$

where d_c and e_c in $D_c(x)=d_c+e_c x^2$ are given by (3.35). Thus, the unconditional form of the Edgeworth expansions for the distributions of $\hat{\theta}$'s can be expressed as

$$P(\sqrt{Nq\bar{L}_2}(\hat{\theta}-\theta)\leq x)=\Phi(x)+\frac{1}{\sqrt{Nq}}\phi(x)D(x), \tag{3.45}$$

where the coefficients d and e in D(x) for $\hat{\theta}_{c.m}$, $\hat{\theta}_u$, $\hat{\theta}_{u,m}$, and $\hat{\theta}_B$ are given by (3.36)

3.3 Asymptotic expansions for $\hat{N}$

In (3.18), expanding $q(\hat{\theta})^{-1}$ in a Taylor series with the use of (3.21) for $\hat{\theta}$, the conditional stochastic expansion for $\hat{N}$, given n, is written as

$$\hat{N}=nq^{-1}-\sqrt{n}\,\alpha-\beta-O(n^{-1/2}), \tag{3.46}$$

where

$$\alpha=Aq'\,q^{-2} \tag{3.47}$$

$$\beta=\{Bq'+A^2(q''/2-q'^2/2)\}\cdot q^{-2}+1/2.$$

Thus, from (3.22), (3.24) and (3.28), the terms α and β in (3.46) for $\hat{N}$'s are

$$\alpha_c=A_c q' q^{-2}$$

$$\alpha_B=\alpha_{u,m}=\alpha_u=\alpha_{c,m}=\alpha_c$$

$$\beta_c=\{B_c q'+A_c^2(q''/2-q'^2/q)\}\cdot q^{-2}+1/2$$

$$\beta_{c,m}=\beta_c+\rho_m \tag{3.48}$$

$$\beta_u=\beta_c+\rho_u$$

$$\beta_{u,m}=\beta_c+\rho_u+\rho_m$$

$$\beta_B=\beta_c+\rho_u+\rho_m+\rho_B,$$

where

$$\rho_u=\nu_u q' q^{-2},\quad \rho_m=\nu_m q' q^{-2},\quad \rho_B=\nu_B q' q^{-2} \tag{3.49}$$

From (3.48), we find that

$$\hat{N}_{c,m}-\hat{N}_c=-\rho_m+O(n^{-1/2})$$

$$\hat{N}_u-\hat{N}_c=-\rho_u+O(n^{-1/2}) \tag{3.50}$$

$$\hat{N}_{u,m}-\hat{N}_c=-(\rho_u+\rho_m)+O(n^{-1/2})$$

$$\hat{N}_B-\hat{N}_c=-(\rho_u+\rho_m+\rho_B)+O(n^{-1/2})$$

Thus, for asymptotic approximations we may regard $\hat{N}_{c,m}$ $\hat{N}_u$, $\hat{N}_{u,m}$, and $\hat{N}_B$ as adjusted estimators of $\hat{N}_c$ given by

$$\begin{aligned}\hat{N}_{c,m}&\approx\hat{N}_c-\hat{\rho}_m\\ \hat{N}_u&\approx\hat{N}_c-\hat{\rho}_u\\ \hat{N}_{u,m}&\approx\hat{N}_c-(\hat{\rho}_u+\hat{\rho}_m)\\ \hat{N}_B&\approx\hat{N}_c-(\hat{\rho}_u+\hat{\rho}_m+\hat{\rho}_B),\end{aligned}\tag{3.51}$$

where

$$\hat{\rho}_m=\rho_m(\hat{\theta}_c),\quad \hat{\rho}_u=\rho_u(\hat{\theta}_c),\quad \hat{\rho}_B=\rho_B(\hat{\theta}_c).$$

From (3.46), the conditional moments expansion of $\hat{N}$, given n, has the form

$$E^n(\hat{N})=nq^{-1}-g-O(n^{-1/2}),\tag{3.52}$$

where

$$\begin{aligned}g&=E^n(\beta)\\ &=\{bq'+\tilde{L}_2^{-1}(q''/2-q'^2/q)\}q^{-2}+1/2.\end{aligned}\tag{3.53}$$

We notice that $E^n(\alpha)=0$, $E^n(B)=b$, and $E(A_c^2)=\tilde{L}_2^{-1}$.
From (3.48), the bias terms g's in (3.52) for $\hat{N}$'s are

$$\begin{aligned}g_c&=\{b_cq'+\tilde{L}_2^{-1}(q''/2-q'^2/q)\}q^{-2}+1/2\\ g_{c,m}&=g_c+\rho_m\\ g_u&=g_c+\rho_u\\ g_{u,m}&=g_c+\rho_u+\rho_m\\ g_B&=g_c+\rho_u+\rho_m+\rho_B.\end{aligned}\tag{3.54}$$

For the unconditional stochastic expansion for $\hat{N}$, substituting (3.40) for n in (3.46), expanding and collecting terms, we see that

$$\hat{N}=N-\tilde{\alpha}\sqrt{N}-\tilde{\beta}+O(N^{-1/2})\tag{3.55}$$

where

$$\begin{aligned}\tilde{\alpha}&=\alpha\sqrt{q}-Y\sqrt{p/q}=Aq'q^{-3/2}-Y\sqrt{p/q}\\ \tilde{\beta}&=\beta+\alpha Y\sqrt{p}/2\end{aligned}\tag{3.56}$$

$$=\{Bq'+A^2(q''/2-q'^2/q)\}q^{-2}+Aq'q^{-2}Y\sqrt{p}/2$$

Thus, from (3.22), (3.24), and (3.28), the terms $\tilde{\alpha}$ and $\tilde{\beta}$ in (3.55) for $\hat{N}$'s are

$$\begin{aligned}
&\tilde{\alpha}_c=A_c\ q'\ q^{-3/2}-Y\sqrt{p/q}\\
&\tilde{\alpha}_B=\tilde{\alpha}_{u,m}=\tilde{\alpha}_u=\tilde{\alpha}_{c,m}=\tilde{\alpha}_c\\
&\tilde{\beta}_c=\{B_cq'+A_c^2(q''/2-q'^2/q)\}q^{-2}+A_cq'q^{-2}+A_cq'q^{-2}Y\sqrt{p}/2\\
&\tilde{\beta}_{c,m}=\tilde{\beta}_c+\rho_m\\
&\tilde{\beta}_u=\tilde{\beta}_c+\rho_u\\
&\tilde{\beta}_{u,m}=\tilde{\beta}_c+\rho_u+\rho_m\\
&\tilde{\beta}_B=\tilde{\beta}_c+\rho_u+\rho_m+\rho_B
\end{aligned}\tag{3.57}$$

Since Y is uncorrelated with A and B, we find that $E(\tilde{\alpha})=0$ and $Var(\tilde{\alpha})=q'^2/\tilde{L}_2q^3+p/q$. Thus, from (3.55) with (3.56), we see that $(\hat{N}-N)/\sqrt{N}$ is asymptotically normal with mean 0 and varince $q'^2/\tilde{L}_2q^3+p/q$.

From (3.55) with (3.56), the unconditional moments expansion for N is expressed as

$$E(\hat{N})=N-g+O(N^{-1/2}).\tag{3.58}$$

We notice that the bias term in (3.58) is exactly the same as in (3.52). Thus from (3.54), the bias-corrected versions of $\hat{N}$'s are

$$\begin{aligned}
&\tilde{N}_c=\hat{N}_c+\hat{g}_c\\
&\tilde{N}_{c,m}=\hat{N}_{c,m}+\hat{g}_c+\hat{\rho}_m\\
&\tilde{N}_u=\hat{N}_u+\hat{\rho}_c+\hat{\rho}_u\\
&\tilde{N}_{u,m}=\hat{N}_{u,m}+\hat{\rho}_c+\hat{\rho}_u+\hat{\rho}_m\\
&\tilde{N}_B=\hat{N}_B+\hat{\rho}_c+\hat{\rho}_u+\hat{\rho}_m+\hat{\rho}_B,
\end{aligned}\tag{3.59}$$

where

$$\hat{g}_c=g_c(\hat{\theta}_c),\quad \hat{\rho}_u=\rho_u(\hat{\theta}_c),\quad \hat{\rho}_m=\rho_m(\hat{\theta}_c),$$

and $\hat{\rho}_B=\rho_B(\hat{\theta}_c)$.

Now, we turn our attention to Edgeworth expansions.
Defining

$$V_c=\sqrt{q^3\tilde{L}_2}\,(\hat{N}_c-N)\,/\,\sqrt{N(q'^2+pq^2\tilde{L}_2)}\tag{3.60}$$

Blumenthal (1982) obtained

$$P\{V_c \leq x\} = \Phi(x) + q^3 \tilde{L}_2^{3/2}\phi(x) / \sqrt{Nq(q'^2+q^2p\tilde{L}_2)^3} \times [G_o + qpg_c + x^2(-G_o + q'^2/q\tilde{L}_2)g_c] \tag{3.61}$$

where

$$G_o = p(p-q)/6 + q'^2/(2q^2\tilde{L}_2) - q'^3\tilde{L}_3/6q^3\tilde{L}_2^3 \tag{3.62}$$

For Edgeworth expansions of the distributions of other estimators $\hat{N}$'s than $\hat{N}_c$, let V_a be defined as (3.60) with $\hat{N}_a$ in place of $\hat{N}_c$. From(3.61), we find that

$$P\{V_a \leq x\} = P\{V_c \leq x\} + \frac{\nu_a q'\sqrt{\tilde{L}_2}}{\sqrt{Nq(q'^2+q^2p\tilde{L}_2)}}\phi(x)$$

$$= P\{V_c \leq x\} + \frac{\rho_a q^{3/2}\sqrt{\tilde{L}_2}}{\sqrt{N(q'^2+q^2p\tilde{L}_2)}}\phi(x) \tag{3.63}$$

Setting $x=0$ in (3.60) and (3.63), we see that the median unbiased versions of $\tilde{N}$'s are

$$\begin{aligned} \tilde{N}_c &= \hat{N}_c - \hat{\rho}_c \\ \tilde{N}_{c,m} &= \hat{N}_{c,m} - \hat{\rho}_c + \hat{\rho}_m \\ \tilde{N}_u &= \hat{N}_u - \hat{\rho}_c + \hat{\rho}_u \\ \tilde{N}_{u,m} &= \hat{N}_{u,m} - \hat{\rho}_c + \hat{\rho}_u + \hat{\rho}_m \\ \tilde{N}_B &= \hat{N}_B - \hat{\rho}_c + \hat{\rho}_u + \hat{\rho}_m + \hat{\rho}_B, \end{aligned} \tag{3.64}$$

where $\hat{\rho}_c = \rho_c(\hat{\theta}_c)$.

4. An example

In order to illustrate the results given in Section 3, we consider a truncated Poisson sample studied in Dahiya and Gross(1973). Let a random variable X have the Poission distribution with the density function

$$f(x|\theta) = \frac{e^{-\theta}\theta^x}{x!} \quad x=0,1,2,\cdots; \quad \theta>0 \tag{4.1}$$

Consider a truncated random sample consisting of n observations from (4.1) where all observations for which $x=0$ are missing. Let n_x be the number of sample observations for which $X=x$, and let $n=\Sigma_{x=1}^K n_x$, where K is the largest observed value of X. Dahiya and Gross (1973) considered the case in which the complete sample size $N(=n+n_o)$ is an unknown constant, but in whith n is assumed to be a random variable. They obtained $\hat{N}_c$ as an estimator of N, and hence $\hat{n}_{oc}$. Under the same situation considered in Dahiya and Gross(1973), assuming the conjugate prior on θ, Blumenthal, Dahiya, and Gross (1978) obtained $\hat{N}_u$ and $\hat{N}_{u,m}$ and investigated the second order asymptotic properties of these estimators. Using the general results given in Section 3, we examine their work with our new estimators $\hat{\theta}_B$ and $\hat{N}_B$. From (3.1) (3.2) and (3.3), we find that

$$l(\mathbf{x},\ n\mid\theta,N)=l_1(n\mid\theta,N)l_2(\mathbf{x}\mid\theta,N) \qquad (4.2)$$

where

$$l_1(n\mid\theta,N)=\binom{N}{n}q^n p^{N-n} \qquad (4.3)$$

and

$$l_2(\mathbf{x}\mid\theta,\ N)=(p/q)^n\prod_{x=1}^{K}(\theta^x/x\,!)^{n_x}, \qquad (4.4)$$

with

$$p=p(x=0)=e^{-\theta},\qquad q=1-p$$

As Blumenthal, Dahiya, and Gross (1978) considered, we assume that θ has the gamma prior density given by

$$\pi(\theta)=\frac{r^{s+1}\theta^s e^{-r\theta}}{\Gamma(s+1)},\quad \theta>0\ ;\ r>0,\ s>-1 \qquad (4.5)$$

and N has the uniform prior given by

$$\psi(N)=1/M,\qquad 1\leq N\leq M, \qquad (4.6)$$

where M is an upper bound on N. Then, we find that

$$\begin{aligned}
&p=e^{-\theta},\ q=1-e^{-\theta}\\
&\tilde{S}(x\mid\theta)=x/\theta-1/q\\
&\tilde{S}'(x\mid\theta)=-x/\theta^2+p/q^2\\
&\tilde{S}''(x\mid\theta)=2x/\theta^3-p(1+p)/q^3\\
&U(\theta)=-1,\ \zeta(\theta)=s/\theta-r
\end{aligned} \qquad (4.7)$$

Thus, from (3.8)−(3.11), we see that

$$0=\sum_{x=1}^{K}n_x(x/\hat{\theta}_c-1/q) \qquad (4.8)$$

$$0=\sum_{x=1}^{K}n_x(x/\hat{\theta}_{c,m}-1/q)+(s/\hat{\theta}_{c,m}-r) \qquad (4.9)$$

$$O=\sum_{x=1}^{K} n_x(x/\hat{\theta}_m-1/q)-(\hat{N}-n/q) \tag{4.10}$$

$$O=\sum_{x=1}^{K} n_x(x/\hat{\theta}_{u,m}-1/q)-(\hat{N}-n/q)+(s/\hat{\theta}_{u,m}-r) \tag{4.11}$$

From (4.8)−(4.11), we find that

$$\frac{\hat{\theta}_c}{1-e^{-\hat{\theta}_c}}=w/n$$

$$\hat{\theta}_{c,m}=(w+s)/(r+n/q)$$

$$\hat{\theta}_u=w/\hat{N}_u \tag{4.12}$$

$$\hat{\theta}_{u,m}=(w+s)/(\hat{N}_{u,m}+r).$$

where $$w=\sum_{x=1}^{K} xn_x.$$

And $\hat{\theta}_B$ is expressed as (3.15), where l_1, l_2, and $\pi(\theta)$ are given by (4.3), (4.4), and (4.5), respectively. Since $\hat{\theta}_B$ will not have a simple colsed form, the approximate formula given in (3.30) will be used for computation. For estimators of N, we have

$$\hat{N}=N(\hat{\theta})=[n/q(\hat{\theta})] \tag{4.13}$$

for appropriate estimator $\hat{\theta}$. From the definition of $\tilde{L}_{ijk}$ given in Section 3, we find that

$$\tilde{L}_2=(\theta q)^{-1}(q-\lambda p)$$

$$\tilde{L}_3=(\theta^2 q^3)^{-1}\cdot[q^2+\theta p\{\theta(p+1)-3q\}] \tag{4.14}$$

$$\tilde{L}_{001}=q^{-1}\cdot\{2/\theta^2-p(1+p)/q^3\}$$

Using the definitions given in Section 3, we obtain the following quantities:

$$\nu_u=\frac{\theta q^2}{2(q-\theta p)} \tag{4.15}$$

$$\nu_m=\frac{\theta q^2}{q-\theta p}\cdot(s/\theta-r) \tag{4.16}$$

$$\nu_B=\frac{1}{2(q-\theta p)^2}\cdot\{q^3-\frac{\theta^2}{2}p(1+p)\} \tag{4.17}$$

$$\nu_c=\frac{1}{6(q-\theta p)^2}\cdot[q^3+\theta pq\{\theta(p+1)-3q\}] \tag{4.18}$$

$$b_c=\frac{1}{(q-\theta p)^2}\cdot\{\theta pq^2-\frac{\theta^2}{6}p(1+p)(1+2q)\} \tag{4.19}$$

$$\rho_u = \frac{\theta p}{2(q-\theta p)} \tag{4.20}$$

$$\rho_m = \frac{\theta p}{q-\theta p} \cdot (s/\theta - r) \tag{4.21}$$

$$\rho_B = \frac{p}{2q^2(q-\theta p)^2} \{q^3 - \frac{\theta^2}{2} p(1+p)\} \tag{4.22}$$

$$g_c = q^{-2}\{b_c p - \tilde{L}_2^{-1}(p/2+p^2/q)\} + 1/2 \tag{4.23}$$

$$\rho_c = \frac{q\tilde{L}_2}{\sqrt{p^2+q^2p\tilde{L}_2}} (G_o + qpg_c), \tag{4.24}$$

where

$$G_o = p(p-q)/6 + p^2/2q^2\tilde{L}_2 - p^3\tilde{L}_3/6q^3\tilde{L}_2^3 \tag{4.25}$$

Thus, the approximate solutions of $\hat{\theta}_{c,m}$, $\hat{\theta}_u$, $\hat{\theta}_{u,m}$ and $\hat{\theta}_B$ are expressed as (3.30), where ν_u, ν_m and ν_B are given by (4.15)−(4.17) with $\hat{\theta}_c$, $\hat{p}=p(\hat{\theta}_c)$ and $\hat{q}=q(\hat{\theta}_c)$ in place of θ, p and q, respectively. The bias-corrected and median unbiased versions of $\hat{\theta}$'s are expressed as (3.33) and (3.38), where ν_c and b_c are given by (4.18) and (4.19) with $\hat{\theta}_c$, $\hat{p}$, and $\hat{q}$ in place of θ, p, and q, respectively. On the other hand, the approximate solutions of $\hat{N}_{c,m}$, $\hat{N}_u$, $\hat{N}_{u,m}$ and $\hat{N}_B$ are obtained from (3.51), where ρ_u, ρ_m and ρ_B are given by (4.20)−(4.22) with $\hat{\theta}_c$, $\hat{p}$, and $\hat{q}$ in place of θ, p, and q, respectively. The bias-corrected and median unbiased versions of $\hat{N}$'s are given by (3.59) and (3.64), respectively, where g_c and ρ_c are given by (4.23) and (4.24) with θ, p, and q in place θ, p, and q respectively.

Now, for a numerical illustration of these estimators of θ and N, we use the data given in Dahiya and Gross(1973) referring to an epidemic of cholera in a village in India :

x	1	2	3	4	Total
n_x	32	16	6	1	55

where x is the number of cholera cases in a house and n_x denotes the number of houses with x cholera cases. In addition to the 55 households having at least one cholera case, 168 other households had no cases. Let N be the total number of households which were infected. Then $N=n_0+n$, where the observed value of n is 55 and n_o is the number of households which were infected but did not have any active case of cholera. Our problem is to estimate n_o. Dahiya and Gross(1973) showed that

$$\hat{\theta}_c = 0.970, \quad \hat{N}_c = 89, \text{ and hence } \hat{n}_{oc} = \hat{N}_c - n = 34.$$

Blumenthal, Dahiya, and Gross(1978) chose $r=1/3$ and $s=1$ to minimize the maximum asymptotic bias of $\hat{N}_{u,m}$. With this choice of r and s, they obtained $\hat{N}_u=\hat{N}_{u,m}=87$, and hence $\hat{n}_{ou}==\hat{n}_{ou,m}=$ 32. Since $\hat{\theta}_c=0.970$, $\hat{p}=0.3791$, and $\hat{q}=0.6209$, from (4.14)−(4.25), we find the following quantities :

$\tilde{L}_2=0.6770$, $\tilde{L}_3=0.8546$, $\tilde{L}_{001}=-0.0943$

$\hat{\nu}_u=0.7385$ $\hat{\nu}_m=1.0304$, $\hat{\nu}_B=-0.1028$

$\hat{\nu}_c=0.3108$ $\hat{b}_c=-0.6558$, $\hat{\rho}_u=0.7262$

$\hat{\rho}_m=1.0132$, $\hat{\rho}_B=-0.1011$, $\hat{g}_c=-1.7580$

$\hat{G}_o=0.1550$, $\hat{\rho}_c=-0.0424$

Thus, from (3.30), (3.33), (3.38), the estimates of θ are as follows:

$\hat{\theta}_{c,m}=0.9887$, $\hat{\theta}_u=0.9834$, $\hat{\theta}_{u,m}=1.0022$,

$\hat{\theta}_B=1.0003$, $\tilde{\theta}_c=0.9819$, $\bar{\theta}_c=0.9757$,

And from (3.51) (3.59) and (3.64), the estimates of N are as follows:

$\hat{N}_{c,m}=87$, $\hat{N}_u=88$, $\hat{N}_{u,m}=87$,

$\hat{N}_B=87$, $\tilde{N}_c=87$, $\bar{N}_c=89$.

Hence, the Bayes estimate of n_o is $\hat{n}_{oB}=\hat{N}_B-n=32$.

5. Conclusion

In this paper, we have examined the second order asymptotic properties of the Bayes estimators of θ and N. We have given the expressions of the aysmptotic biases and median biases of θ and N. For asymptotic purpose, we have shown that the Bayes estimators of θ and N can also be regarded as adjusted estimators of m.l.e.'s of θ and N as well as the Bayes modal estimators.

For further research, we will need to examine the mean squared error to control the bias terms of the Bayes estimators of θ and N. We will also need Monte Carlo simulation study to investigate the small sample behavior of the Bayes estimators. We hope that these further results will be reported in a future time.

Acknowledgement

I wish to express my gratitude to Professor Saul Blumenthal for suggesting this topic and for valuable discussions during my doctoral study at The Ohio State University.

References

1. Blumental, S. (1977), Estimating Populatin Size with Truncated Sampling, *Communications in Statistics*, A 6, 297-308.
2. ________(1982), Stochastic Expansions for Point Estimation from Truncated Samples, *Sankhya*. A 44, 436-451.

3. Blumenthal, S. and Dahiya, R.C.(1981), Estimating the Binomial Parameter n, *Journal of the American Statistical Association*, 76, 903-909.
4. Blumenthal, S., Dahiya, R.C., and Gross, A.J.(1978), Estimating the Complete Sample Size from an Incomplete Poisson Sample, *Journal of the American Statistical Association*, 73, 182-187.
5. Blumenthal, S. and Marcus, R.(1975), Estimating Population Size with Exponential Failure, *Journal of the American Statistical Association*, 28, 913-922.
6. Blumenthal, S. and Sanathanan, L.P.(1980), Estimation with Truncated Inverse Binomial Sampling, *Communications in Statistics*, A 9, 997-1017.
7. Chibisov, D.M.(1972a), Asymptotic Expansions for Maximum Likelihood Estimates, *Theor. Prob. Applications*, 17, 368-369.
8. _______(1972b), An Asymptotic Expansion for the Distribution of a Statistic Admitting an Asymptotic Expansion, *Theor. Prob. Applications*, 17, 620-630.
9. _______(1973), An Asymptotic Expansion for a Class of Estimators Containing Maximum Likelihood Estimators, *Theor. Prob. Applications*, 18, 295-303.
10. Dahiya, R.c. and Gross, A.J.(1973), Estimating the Zero Class from a Truncated Poisson Sample, *Journal of the American Statistical Association*, 68, 731-733.
11. Draper, N. and Guttman, I.(1971), Bayesian Estimation of the Binomial Parameter, *Technometrics*, 13, 667-673.
12. Feldman, D. and Fox, M.(1968), Estimation of the Parameter n in the Binomial Distribution, *Journal of the American Statistical Assoication*, 63, 150-158.
13. Ghosh, J.K. and Subramanyam, K.(1974), Second Order Efficiency of Maximum Likelihood Estimators, *Sankhya*, A 36, 325-358.
14. Gusev, S.I.(1975), Asymptotic Expansions Associated with Some Statistical Estimators in the Smooth Case, I. Expansions of Random variables, *Theor. Prob. Applications*, 20, 470-498.
15. _______(1976), Asymptotic Expansions Associated with Some Statistical Estimators in the Smooth Case, II. Expansions of Moments and Distributions, *Theor. Prob. Applications*, 21, 14-33.
16. Johnson, N.L.(1962), Estimation of Sample Size, *Technometrics*, 4, 59-67.
17. Lindley, D.V.(1961), The Use of Prior Probability Distributions in Statistical Inference and Decision, *Proceedings of the fourth Berkeley Symposium on Mathematical Statistics*, 1, 453-468.
18. _______(1980), Approximate Bayesian Methods, *Trabajos Estadistica*, 31, 223-237.
19. Marcus, R and Blumenthal, S.(1974), A Sequential Screening Procedure, *Technometrics*, 16, 229-234.
20. Pfanzagl, J.(1973), Asymptotic Expansions Related to Minimum Contrast Estimators, *The Annals of Statistics*, 1, 993-1026.
21. Sanathanan, L.P.(1972), Estimating the Size of a Multinomial Population, *The Annals of Mathematical statistics*, 43, 142-152.
22. _______(1977), Estimating the Size of a Truncated Sample, *Journal of the American statistical Association*, 72, 669-672.
23. Watson, D. and Blumenthal, S.(1980), Estimating the Size of a Truncated Sample, *Communications in Statistics*, A 9, 1535-1550.