ASYMPTOTIC BIAS AND VARIANCE OF RATIO ESTIMATES IN GENERALIZED POWER SERIES DISTRIBUTIONS AND CERTAIN APPLICATIONS

By G. P. PATIL*

Indian Statistical Institute

SUMMARY. A discrete probability distribution which forms a generalization of some important discrete distributions like the Binomial, Poisson, Negative Binomial and Logarithmic Series and their truncated forms is introduced. It is called the "generalized power series distribution (gpsd)." In this paper, we suggest what we call the "Ratio Method" for estimation of the parameter of the gpsd and investigate its important properties and study certain applications. The method is applicable not merely for estimating the parameter, but also for its integral powers. The performance of the method is investigated, in particular, in case of truncated Binomial and truncated Poisson distributions and correspondingly certain recommendations are offered.

1. Introduction

Let $g(\theta)$ be a positive function admitting a power series expansion with non-negative coefficients for non-negative values of θ smaller than the radius of covergence of the power series :

$$g(\theta) = \sum_{z=0}^{\infty} a_z \theta^z \qquad \dots \quad (1.1)$$

Noack (1950) defined a random variable Z taking non-negative integral values z with probabilities

$$\text{Prob} \quad \{Z = z\} = \frac{a_z \theta^z}{g(\theta)} \qquad z = 0, 1, 2, \dots. \qquad \dots \quad (1.2)$$

He called the discrete probability distribution given by (1.2) a power series distribution (psd) and derived some of its properties relating to its moments, cumulants, etc.

To be more general, we note that the set of values of an integral-valued random variable Z need not be the entire set of non-negative integers (0, 1, 2, ...). For, let T be an arbitrary non-null subset of non-negative integers[1] and define the generating function

$$f(\theta) = \sum_{z \epsilon T} a_z \theta^z$$

with $a_z \geqslant 0$; $\theta \geqslant 0$ so that $f(\theta)$ is positive, finite and differentiable.

*Now with the University of Michigan.

[1]In fact, one can take T to be a countable subset of real numbers; for purposes of this paper, however, T is chosen to be a subset of non-negative integers.

Then we can define a random variable X taking non-negative integral values in T with probabilities

$$P_x = \text{Prob}\{X = x\} = \frac{a_x \theta^x}{f(\theta)} \quad x \epsilon T \qquad \dots \quad (1.3)$$

and call this distribution analogously a generalized power series distribution (gpsd). It may be noted that gpsd reduces to a psd when T is the entire set of non-negative integers. The properties established by Noack (1950) and Khatri (1959) for psd can be easily deduced for gpsd by following the same lines. Further, it can be easily seen that proper choice of T and $f(\theta)$ reduces the gpsd, in particular, to the Binomial, negative Binomial, Poisson and logarithmic series distributions and their truncated forms. Incidentally, it is obvious that truncated gpsd is itself a gpsd in its own right and hence the properties that hold for a gpsd continue to hold for its truncated forms.

Problems of statistical inference associated with psd's do not seem to have been much investigated. Roy and Mitra (1957) have derived the uniformly minimum variance unbiased estimates in certain particular cases and have provided necessary tables for Poisson distribution truncated at zero. The author (1957) has shown that for gpsd (1.2), the maximum likelihood method and the method of moments give the same estimate of the parameter of the gpsd. The likelihood equation and a method for solving it are derived for the problem of estimation. In this paper, we suggest what we call the, "Ratio Method" for estimation of the parameter of the gpsd and investigate its important properties and study certain applications.

2. Estimation by the ratio method for a g.p.s.d.

[Range T finite and $T = (c, c+1, \dots c+k = d)$ with positive probabilities].

The gpsd that we consider here is of the form :

$$P_x = \text{Prob}\{X = x\} = \frac{a_x \theta^x}{f(\theta)} \qquad \dots \quad (2.1)$$

where $\quad x \epsilon T = (c, c+1, \dots c+k = d), \quad d$ finite

$$f(\theta) = \sum_{x=c}^{d} a_x \theta^x \qquad \dots \quad (2.2)$$

and $\quad a_x > 0$ for $x \epsilon T$.

Let

$$g_r(x) = \frac{a_{x-r}}{a_x} \quad x \epsilon T \qquad \dots \quad (2.3)$$

with r being an integer such that $x - r \epsilon T$. Then

$$\sum_{x=u}^{v} g_r(x) P_x = \theta^r \sum_{x=u-r}^{v-r} P_x \qquad \dots \quad (2.4)$$

where u and v are arbitrary with $c+r \leqslant u \leqslant v \leqslant d$. From (2.4) we get the identity

$$\theta^r = \sum_{x=u}^{v} g_r(x)P_x \Big/ \sum_{x=u-r}^{v-r} P_x \qquad \dots \quad (2.5)$$

which can be made use of in problems of estimation. In a sample of size N, if n_x is the observed frequency for x, then since $E(n_x) = NP_x$, the statistic

$$\sum_{x=u}^{v} g_r(x)n_x \Big/ \sum_{x=u-r}^{v-r} n_x \qquad \dots \quad (2.6)$$

may be taken as an estimate of θ^r for admissible values of $r = 1, 2$, etc. Since u and v are arbitrary, the same method is applicable for estimation in truncated and censored gpsd's also, provided that their range contains a subset of consecutive integers. We call these estimates "ratio estimates."

It is interesting to note that Plackett (1953) and Moore (1952, 1954) applied this ratio method to the special cases of estimating θ in truncated Binomial and Poisson distributions. The method which we call the ratio method is applicable not merely for estimating θ, but also for its integral powers and for any gpsd of this section, truncated or censored.

The ratio estimate is not generally unbiased or efficient, but is always easy to compute. In certain cases (see Section 3), however, unbiased estimates can be obtained by the ratio method. In other cases, such as those in this section, the bias is generally of the order $\frac{1}{N}$. It may be easily verified that no unbiased estimate for θ(and θ^r in general) exists when the range of T is finite as in situations considered here.

Consider the following ratio estimate of θ for gpsd (2.1):

$$\theta' = \frac{t_1}{t_2} \qquad \dots \quad (2.7)$$

where

$$t_1 = \sum_{x=c+1}^{d} \left(\frac{a_{x-1}}{a_x}\right) n_x \qquad \dots \quad (2.8)$$

and

$$t_2 = \sum_{x=c}^{d-1} n_x. \qquad \dots \quad (2.9)$$

Then, writing

$$E(t_2) = N \sum_{x=c}^{d-1} P_x = N(1-P_d) = NP, \text{ say,} \qquad \dots \quad (2.10)$$

where

$$P = 1-P_d, \qquad \dots \quad (2.11)$$

we have

$$E(t_1) = NP\theta. \qquad \dots \quad (2.12)$$

Let

$$t_1-E(t_1) = \delta t_1 \quad \text{and} \quad t_2-E(t_2) = \delta t_2\,. \qquad \dots \quad (2.13)$$

Then

$$\theta' = \frac{t_1}{t_2} = \theta\left(1+\frac{\delta t_1}{NP\theta}\right)\left(1+\frac{\delta t_2}{NP}\right)^{-1}$$

Since the deviations δt_1, δt_2 are stochastically of order $N^{\frac{1}{2}}$, we get on expansion

$$\theta' = \theta \left[1+\frac{\delta t_1}{NP\theta}-\frac{\delta t_2}{NP}-\frac{(\delta t_1)(\delta t_2)}{N^2P^2\theta}+\frac{(\delta t_2)^2}{N^2P^2} \right] \quad \dots \quad (2.14)$$

neglecting terms of order higher than $\frac{1}{N}$. Thus, to this order of approximation,

$$E(\theta') = \theta \left[1+\frac{E(\delta t_2)^2}{N^2P^2}-\frac{E(\delta t_1)(\delta t_2)}{N^2P^2\theta} \right]. \quad \dots \quad (2.15)$$

Now a little computation gives

$$E(\delta t_2)^2 = NP(1-P) \quad \dots \quad (2.16)$$

and

$$E(\delta t_1)(\delta t_2) = N\theta[P(1-P)-P_{d-1}]. \quad \dots \quad (2.17)$$

Thus

$$E(\theta') = \theta+\frac{1}{N}\left(\frac{\theta P_{d-1}}{P^2} \right) \quad \dots \quad (2.18)$$

from which we get the magnitude of the bias θ', to order $\frac{1}{N}$,

$$b(\theta') = \frac{1}{N}\left(\frac{\theta P_{d-1}}{P^2} \right) = (\theta P_{d-1})/N(1-P_d)^2 = \frac{B(\theta')}{N}, \text{ say.} \quad \dots \quad (2.19)$$

The variance of θ' correct to terms of order $\frac{1}{N}$ is

$$\text{Var}(\theta') = \frac{1}{N^2P^2}\,[E(\delta t_1)^2+\theta^2E(\delta t_2)^2-2\theta E(\delta t_1)(\delta t_2)]. \quad \dots \quad (2.20)$$

Now

$$E(\delta t_1)^2 = N(D-P^2\theta^2) \quad \dots \quad (2.21)$$

where

$$D = \sum_{x=c+1}^{d} \left(\frac{a_{x-1}}{a_x}\right)^2 P_x. \quad \dots \quad (2.22)$$

Thus, to order $\frac{1}{N}$

$$\text{Var}(\theta') = \frac{1}{NP^2}\,[D-P\theta^2+2\theta^2P_{d-1}]. \quad \dots \quad (2.23)$$

3. Unbiased estimation by the ratio method for a g.p.s.d.

[Range T infinite and $T = (c, c+1, \dots)$ with positive probabilities].

Uniformly minimum variance unbiased estimation for psd's has been considered by Roy and Mitra (1957). Tate and Goen (1958) have considered the same for truncated Poisson.

It is easy to demonstrate that the ratio method discussed in Section 2 gives the unique unbiased estimate of θ, linear in frequencies, for a gpsd with range T infinite and $T = (c, c+1, \dots)$ with positive probabilities. For, consider the gpsd

$$P_x = \text{Prob}\,\{X = x\} = \frac{a_x\theta^x}{f(\theta)} \quad x = c, c+1, \dots \qquad \dots \quad (3.1)$$

where

$$f(\theta) = \sum_{x=c}^{\infty} a_x\theta^x \qquad \dots \quad (3.2)$$

and

$$a_x > 0 \quad \text{for all} \quad x = c, c+1, \dots .$$

Now, if in a sample of size N from gpsd (3.1), the frequency of x is n_x and we want an unbiased estimate for θ of the type linear in n_x, we should be able to demonstrate the existence of a function of x, $t(x)$, such that, denoting the corresponding estimate

$$\tilde{\theta} = \sum_{x=c}^{\infty} t(x)n_x \qquad \dots \quad (3.3)$$

we must have $E(\tilde{\theta}) = \theta$ for all θ in the parameter space of (3.1). That is

$$N \sum_{x=c}^{\infty} t(x)a_x\theta^x = \sum_{x=c}^{\infty} a_x\theta^{x+1}.$$

Since this is an identity in θ, equating coefficients of corresponding powers of θ, we get

$$t(x) = \begin{cases} 0 & \text{for} \quad x = c \\ \dfrac{1}{N}\left(\dfrac{a_{x-1}}{a_x}\right) & \text{for} \quad x = c+1, c+2, \dots . \end{cases}$$

Thus, the unique unbiased estimate of θ linear in the frequencies comes out to be the ratio estimate θ'. The exact variance of this estimate is

$$\sigma^2(\theta') = \frac{1}{N}\left[\sum_{x=c+1}^{\infty} \left(\frac{a_{x-1}}{a_x}\right)^2 P_x - \theta^2\right]. \qquad \dots \quad (3.4)$$

An unbiased estimate of $\sigma^2(\theta')$ is

$$\left[\sum_{x=c+1}^{\infty} \left(\frac{a_{x-1}}{a_x}\right)^2 n_x - N\theta'^2\right] \Big/ N(N-1) \qquad \dots \quad (3.5)$$

the proof of which is almost immediate once one recognizes that θ' is the mean of N independent identically distributed random variables Y_i with probability distribution given by (for $i = 1, 2, \ldots, N$)

$$\text{Prob } \{Y_i = 0\} = P_c$$

and
$$\text{Prob } \left[Y_i = \frac{a_{x-1}}{a_x} \right] = P_x \quad \text{for } x = c+1,\ c+2, \ldots .$$

One can compare $\sigma^2(\theta')$ with the asymptotic variance Var $(\hat{\theta})$ of the maximum likelihood estimate of θ and the efficiency of the ratio estimate θ' can be computed.

Lastly, one may establish that

$$\frac{1}{N} \sum_{x=c+r}^{\infty} \left(\frac{a_{x-r}}{a_x} \right) n_x \qquad \ldots \quad (3.6)$$

is the only unbiased estimate of θ^r (r an integer) which is a linear function of the frequencies.

4. Estimation by the ratio method for singly truncated binomial distribution

Fisher (1936) and Haldane (1932, 1938) discussed uses of the truncated binomial distribution. For instance, in problems of human genetics, in estimating the proportion of albino children produced by couples capable of producing albinos, sampling has necessarily to be restricted to families having at least one albino child. Finney (1949) has cited some more applications. Fisher and Haldane derived the maximum likelihood procedure to estimate the parameter π. Patil (1959) gave a direct method to obtain the maximum likelihood estimate. Moore (1954) suggested a simple "ratio-estimate" based on an identity between binomial probabilities. For a slightly different problem, where, in a sample from a complete binomial distribution, the frequencies in some lowest classes are missing, Rider (1955) suggested a method of estimation, which uses first two moments of the complete binomial and leads to a linear equation.

The probability law of the binomial distribution truncated at c on the left can be written as

$$b^*(x, \pi, n) = \left((B^*(c, \pi, n) \right)^{-1} \binom{n}{x} \pi^x (1-\pi)^{n-x},\ x = c, c+1, \ldots n. \quad \ldots \quad (4.1)$$

where
$$B^*(c, \pi, n) = \sum_{x=c}^{n} \binom{n}{x} \pi^x (1-\pi)^{n-x} \qquad \ldots \quad (4.2)$$

The first two moments about the origin of (4.1), then, are

$$\mu^* = \mu^*(c, \pi, n) = n\pi. \quad B^*(c-1, \pi, N-1)/B^*(c, \pi, n) \qquad \ldots \quad (4.3)$$

and
$$m_2^* = m_2^*(c, \pi, n) = \mu^*(c, \pi, n) \{1+\mu^*(c-1, \pi, n-1\}. \qquad \ldots \quad (4.4)$$

The case of truncation to the right can be dealt with in a similar way by replacing π by $1-\pi$ and the truncation point c by $n-c$.

In this case $a_{x-1} | a_x = x|(n-x+1)$ and since $\theta = \pi/(1-\pi)$, (see Patil, 1959), we have the following "ratio-estimate" for π :

$$\pi' = \frac{t_1}{t_1+t_2} \qquad \dots \quad (4.5)$$

where $$t_1 = \sum_{x=c+1}^{n} \left(\frac{x n_x}{n-x+1} \right) \text{ and } t_2 = \sum_{x=c}^{n-1} n_x.$$

To investigate the efficiency of π' given by (4.5) its asymptotic variance can be written down as:

$$\text{Var } (\pi') = \frac{(1-\pi)^2}{NP^2} \left[(1-\pi)^2 D - P\pi^2 + 2\pi^2 \, . \, P_{n-1}\right] \qquad \dots (4.6)$$

where $$P = \sum_{n=c}^{n-1} b^*(x, \pi, n)$$

$$D = \sum_{x=c+1}^{n} \left(\frac{x}{n-x+1} \right) b^*(x, \pi, n)$$

and $$P_{n-1} = b^*(n-1 \, . \, \pi, n) \; .$$

Also the asymptotic variance of the maximum likelihood estimate $\hat{\pi}$, (Patil, 1959), is given by

$$\text{Var } (\hat{\pi}) = \frac{\pi(1-\pi)^2}{N\mu_2^*}$$

where μ_2^* is the variance of (4.1).

Therefore the asymptotic efficiency of π' takes the form :

$$\text{Eff } (\pi') = \frac{P^2}{\mu_2^*} \left(\left(\frac{1-\pi}{\pi}\right)^2 D - P + 2P_{n-1} \right)^{-1} . \qquad \dots \quad (4.7)$$

The special cases of some importance in genetics are $c = 1$ and $\pi = 1/4$, $1/2$, or $3/4$. The efficiency of the Ratio-Estimate (R) relative to the Maximum Likelihood Estimate (ML) in these cases is tabulated and shown in Table 1.

TABLE 1. ASYMPTOTIC EFFICIENCY OF R FOR $c = 1$

n	$\pi = 1/4$	$1/2$	$3/4$
3	.924	.875	.875
4	.909	.769	.772
5	.919	.715	.664
6	.933	.694	.563
7	.947	.693	.523
8	.952	.705	.481
9	.956	.723	.435
10	.959	.776	.388

Examination of the above table shows that the efficiency of R in case of $\pi = 1/4$ and $\pi = 1/2$ decreases in the beginning with n, reaches a minimum and then increases with increasing values of n. For $\pi = 3/4$, however, the efficiency decreases throughout for $n = 3(1)10$.

Following Section 2, one gets, to order $1/N$, the amount of bias of π' (R) as follows:

$$b(\pi') = \frac{(1-\pi)^2}{NP^2}\,[\pi^2 P+(\pi-2\pi^2)P_{n-1}-(1-\pi)^2 D] = \frac{B(\pi')}{N}.$$

The table below gives $B(\pi')$ for $c = 1$ and $\pi = 1/4$, 1/2 and 3/4.

TABLE 2. N (AMOUNT OF BIAS TO ORDER $1/N$) OF R

n	$\pi = 1/4$	1/2	3/4
3	−.1027	.1458	.0820
4	.1227	.1307	.0833
5	.0801	.1184	.0893
6	.0648	.1062	.0902
7	.0510	.0943	.0931
8	.0420	.0833	.0928
9	.0355	.0736	.0915
10	.0301	.0651	.0896

Table 2 shows that R is an underestimate of π. A closer investigation, however, brings out that the bias to order $1/N$ is quite small for R. One may note that the maximum likelihood estimate also happens to be biased in this case (Patil, 1959).

Illustrative example. The detailed computation procedure of the ratio estimate discussed above will be illustrated with reference to K. Pearson's data on albinism in man. The table below gives the number of families (n_x), each of five children having exactly x children in the family, ($x = 1, 2, 3, 4, 5$).

number of albinos in family (x)	1	2	3	4	5
number of families (n_x)	25	23	10	1	1

If π is the probability for a child to be an albino, we may accept the truncated binomial model:

$$\frac{\binom{n}{x}\pi^x(1-\pi)^{n-x}}{1-(1-\pi)^n} \qquad x = 1, 2, \ldots n$$

for the probability of x albinos in a family of n. Here $n = 5$, and the problem is to estimate π on the basis of the data given in the table above.

Here, $t_2 = \sum_{x=1}^{n-1} n_x = 59$ and t_1 can be computed from the following table :

x	$\frac{x}{n-x+1}$	n_x
2	.5	23
3	1	10
4	2	1
5	5	1

Then,
$$t_1 = \sum_{x=2}^{n} \left(\frac{x}{n-x+1}\right) n_x = 28.50.$$

The ratio estimate is obtained as $\pi' = \frac{28.50}{28.50+59} = 0.3257$.

To compute the variance of π', we require

$$P = 1 - \frac{\pi^n}{1-(1-(1-\pi))^n} = 0.99574$$

$$P_{n-1} = \frac{n(1-\pi)}{\pi}(1-P) = 0.04408$$

and

$$D = \left(\frac{\pi}{1-\pi}\right)\left[\frac{n+1}{1-(1-\pi)^n}\left\{(1-\pi^n)E\left(\frac{1}{x}, n, 1-\pi\right) - \frac{(1-\pi)^n}{n}\right\} - P\right] = 0.56156$$

where
$$E\left(\frac{1}{x}, n, \pi\right) = \sum_{x=1}^{n} \frac{1}{x}\binom{n}{x}\pi^x(1-\pi)^{n-x}/\{1-(1-\pi)^n\}$$

and is tabulated by Grab and Savage (1954).

By linear interpolation from the table by Grab and Savage, $E\left(\frac{1}{x}, n, 1-\pi\right) = 0.33870$ taking $\pi' = 0.3257$ as the estimate for π throughout. Then the variance of π' is estimated from the formula

$$\text{Var}(\pi') = \frac{(1-\pi)^2}{NP^2}[(1-\pi)^2 D - P\pi^2 + 2\pi^2 . P_{n-1}] = 0.0013410$$

so that the standard error of π' is S.E. $(\pi') = 0.03662$.

Incidentally, the maximum likelihood estimate $\hat{\pi}$ comes out (Patil, 1959) to be in this case, $\hat{\pi} = 0.3088$ with S.E. $(\hat{\pi}) = 0.03210$.

6. Estimation by ratio method for truncated Poisson distribution

Problems of estimation in a truncated Poisson distribution with known truncation points have been discussed by various authors. The case of truncation on the left has been considered by David and Johnson (1948) who gave the maximum likelihood estimate, by Plackett (1953) who gave a simple and highly efficient ratio estimate, and by Rider (1953) who used first two moments. Truncation on the right has been discussed by Tippett (1932), Bliss (1948), and Moore (1952). Tippett derived the maximum likelihood solution; Bliss developed an approximation to it; and Moore suggested a simple ratio estimate. Uniformly minimum variance unbiased estimates have been obtained by Roy and Mitra (1957) and by Tate and Goen (1958). For both types of truncations, the author (1959) has provided neat and compact equations for estimation by the method of maximum likelihood. He has also presented numerical tables and some suitable charts to facilitate the solution of these equations in certain special cases. In this section, we study the Ratio Method as applied to truncated Poisson distributions.

The probability law of the singly truncated Poisson distribution with truncation point on the right at d can be written as:

$$p^*(x, \mu) = [P(d, \mu)]^{-1} e^{-\mu} \frac{\mu^x}{x!} \quad x = 0, 1, 2, \ldots d \qquad \ldots \quad (6.1)$$

where

$$P(d, \mu) = \sum_{x=0}^{d} e^{-\mu} \frac{\mu^x}{x!}. \qquad \ldots \quad (6.2)$$

In this case, $a_{x-1}/a_x = x$, and since $\theta = \mu$, the ratio estimate for μ takes the form

$$\mu' = \sum_{x=0}^{d} x n_x \Big/ \sum_{x=0}^{d-1} n_x \qquad \ldots \quad (6.3)$$

as first suggested by Moore (1954). The following table gives the asymptotic efficiency of μ' relative to $\hat{\mu}$ for values of $d = 5$ with $\mu = .25$, $.5(.5)2.5$, and $d = 10$ with $\mu = .5(.5)4.5$.

TABLE 3. EFFICIENCY OF R

μ	.25	.50	1.00	1.50	2.00	2.50
			Case (i) $d = 5$			
Eff.	.999	.990	.979	.967	.951	.923

μ	.5	1.0	1.5	2.0	2.5	3.0	3.5	4	4.5
				Case (ii) $d = 10$					
Eff.	1.000	1.000	1.000	1.000	.999	.992	.981	.894	.817

Thus, R seems to be highly efficient on the whole and its efficiency always exceeds 82 percent in situations considered in Table 3.

The following table gives $B(\mu')$ for values of $d = 5$ with $\mu = .25$, $.5(.5)2.5$ and $d = 10$ with $\mu = .5(.5)5$.

TABLE 4. N (AMOUNT OF BIAS TO ORDER $1/N$) OF R

μ	.25	.50	1.00	1.50	2.00	2.50				
				Case (i) $d = 5$						
$B(\mu')$	.0003	.0008	.0015	.0719	.1977	.4461				
				Case (ii) $d = 10$						
μ	.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
$B(\mu')$	.0000	.0000	.0003	.0004	.0022	.0081	0231	.0536	.1063	.1876

Table 4 shows that, though over-estimate, R involves almost negligible bias.

The probability law of the singly truncated Poisson distribution with truncation point on the left at c can be written as :

$$p^*(x, \mu) = [P^*(c, \mu)]^{-1} e^{-\mu} \frac{\mu^x}{x!} \quad x = c, c+1, \ldots \infty \qquad \ldots \quad (6.8)$$

where

$$P^*(c, \mu) = \sum_{x=c}^{\infty} e^{-\mu} \frac{\mu^x}{x!} \qquad \ldots \quad (6.9)$$

In this case, $a_{x-1}/a_x = x$ and since $\theta = \mu$, we have the following "ratio estimate" for μ:

$$\mu' = \sum_{x=c+1}^{\infty} x n_x \Big/ N \qquad \ldots \quad (6.10)$$

when $c = 1$, i.e., when only "zero" counts are truncated, the estimate takes the form suggested by Plackett (1953):

$$\mu' = \sum_{x=2}^{\infty} x n_x \Big/ N. \qquad \ldots \quad (6.11)$$

The unique unbiased estimate of μ linear in the frequencies (*ibid.*, Section 3) is provided in (6.10). The exact variance of this estimate is

$$\sigma^2(\mu') = \frac{1}{N}\Big[\sum_{x=c+1}^{\infty} x^2 P_x - \mu^2\Big] \qquad \ldots \quad (6.12)$$

and an unbiased estimate of $\sigma^2(\mu')$ is

$$\Big\{\sum_{x=c+1}^{\infty} x^2 n_x - N\mu'^2\Big\} \Big/ N(N-1) \qquad \ldots \quad (6.13)$$

when $c = 1$, (6.12) reduces to

$$\sigma^2(\mu') = \frac{1}{N}[\mu+\mu^2/(e^{\mu}-1)] \qquad \dots \quad (6.14)$$

first derived by Plackett (1953). Plackett computed also the efficiency of μ' in this special case. The following table gives the efficiencies of μ' relative to $\hat{\mu}$ as computed by Plackett.

TABLE 5. EFFICIENCY OF R FOR $c = 1$.

μ	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
Eff.	.9693	9559	9539	9586	9662	9743	9815	9872

The author is thankful to Professor C. R. Rao and Dr. J. Roy for helpful discussions at the Indian Statistical Institute and to the referee for helpful comments.

REFERENCES

DAVID, F. N. and JOHNSON, N. L. (1952): Truncated Poisson. *Biometrics*, **8**, 275-285.

FISHER, R. A. (1936): The effects of method of ascertainment upon estimation of frequencies. *Ann. Eugen.*, **6**, 13-25.

FINNEY, D. J. (1949): The truncated binomial distribution. *Biometrics*, **9**, 176-200.

FINNEY, D. J. and VARLEY, G. C. (1955): An example of the truncated Poisson distribution. *Biometrics*, **10**, 417-451.

GRAB, E. L. and SAVAGE, I. R. (1954): Tables of the expected value of $1/X$ for positive Bernoulli and Poisson variables. *J. Amer. Stat. Ass.* **49**, 169-177.

HALDANE, J. B. S. (1941): The fitting of binomial distribution. *Ann. Eugen.*, **11**, 179-181.

——— (1945): On a method of estimating frequencies. *Biometrika*, **33**, 222-225.

MOORE, P. G. (1952): The estimation of the Poisson parameter from a truncated distribution. *Biometrika*, **39**, 247-251.

——— (1954): A note on truncated Poisson distributions. *Ann. Math. Stat.*, **14**, 389-400.

PATIL, G. P. (1957): Thesis submitted for the Associateship of the Indian Statistical Institute.

——— (1959): Doctoral Thesis submitted to the University of Michigan.

PEARSON, K. (1913): A monograph on albinism in man. *Drapers Company Research Memoirs.*

PLACKETT, R. L. (1953): The truncated Poisson distribution. *Biometrics*, **9**, 485-488.

RIDER, P. R. (1953): Truncated Poisson distributions. *J. Amer. Stat. Ass.*, **48**, 826-830.

——— (1955): Truncated binomial and negative binomial distributions. *J. Amer. Stat. Ass.*, **50**, 877-883.

ROY, J. and MITRA, S. K. (1957): Unbiased minimum variance estimation in a class of discrete distributions. *Sankhyā*, **18**, 371-378.

TATE and GOEN (1958): Minimum variance unbiased estimation for the truncated Poisson distribution. *Ann. Math. Stat.*, **29**, 755-765.

TIPPETT, L. H. C. (1932): A modified method of counting particles. *Proc. Roy. Soc.*, Series A, **137**, 434-446

Paper received : April, 1960.
Revised : August, 1960.