ESTIMATING THE MEAN OF A RANDOM BINOMIAL PARAMETER

G. MORRIS SOUTHWARD
INTERNATIONAL PACIFIC HALIBUT COMMISSION
and
J. VAN RYZIN
UNIVERSITY OF WISCONSIN

1. Introduction

In studying biological phenomenon, one often observes random variables which are the result of other randomly occurring unobservable events. This is usually the case in the observation of genetic traits. The measurable trait in question has a probability distribution for the population of animals under study. Each individual member of the population of animals carries a value of the measurable trait, but it may or may not (and often is not) directly observable. It is not difficult to envision the probability distribution of the trait in the population as being continuous, while the distribution of the visible expression of the trait is a discrete count depending on the value of the measurable trait.

Such a problem came to the authors' attention during discussion with a poultry scientist who was interested in the probability distribution governing the frequency with which blood spotted eggs occur. Poultrymen wish to determine from examination of a small number of eggs laid early in the life of each hen what the average probability of laying blood spotted eggs is for the flock.

The problem can be conceptualized as follows. The distribution of blood spots in eggs for a given chicken is taken as binomial. That is, if p represents the probability of a given chicken to lay a blood spotted egg and m eggs are laid, then X = number of blood spotted eggs is binomially distributed with parameters m and p assuming the eggs are laid independently. However, the probability p (or propensity) for laying blood spotted eggs (the trait in question), differs from chicken to chicken and can be thought of as having a continuous distribution on the unit interval. The probability distribution of the blood spotting trait p in the population is not directly observable. That is, one might postulate that the binomial parameter p (or trait) has a distribution on the unit interval and that the values of this probability carried by each bird in the flock are independently allocated according to this distribution, denoted $G(p)$. Rarely, if ever, are values of p directly observable.

The investigation of J. Van Ryzin was supported in part by USPHS Research Grant No. GM-10525-08, National Institute of Health, Public Health Service, and in part by NSF Contract No. GP-9324.

Thus, the model is written as

$$f(x; m) = \int_0^1 \binom{m}{x} p^x(1 - p)^{m-x}\, dG(p), \tag{1.1}$$

$x = 0, \cdots, m$, where m is the size of the sample examined and $G(p)$ is the distribution of the probability p.

The problem we consider here is that of estimating the mean μ_p of the probability distribution $G(p)$, that is, $\mu = \mu_p = \int p\, dG(p)$, based on a random sample $X_1, \cdots, X_n$ from $f(x; m)$, where X_i is the number of blood spotted eggs among the m eggs sampled from the ith chicken. The moment estimator of μ_p and its large sample properties along with confidence intervals are developed in Section 2.

It is clear that the above type of sampling, when we wish to estimate μ_p, occurs in many setups similar to that of our chicken example. Also, the allied problem, when the sample size m is allowed to vary from individual to individual, is discussed in Sections 3, 4, and 5. In that case, the distribution of each X_i is given by equation (1.1), where m is now replaced by m_i; that is, X_i is distributed with discrete density $f(x_i; m_i)$ in (1.1), $x_i = 0, \cdots, m_i$, $i = 1, \cdots, n$.

Theorems 5.1 and 5.2 develop the consistency and asymptotic distribution theory for the case of differing sample sizes (m_i different). These theorems concern an estimator μ_0^* for μ which behaves asymptotically like the minimum variance unbiased linear (in the X_i) estimator of μ which is studied in Section 3.

Some aspects of this general problem are covered by Pearson [2] who discusses Bayes theorem in the light of experimental sampling. However, no attack on the above problem is made therein.

2. Estimation of the mean of $G(p)$

Let $X_1, \cdots, X_n$ be a random sample from the model

$$f(x; m) = \int_0^1 \binom{m}{x} p^x(1 - p)^{m-x}\, dG(p). \tag{2.1}$$

We consider the problem of estimating the mean

$$\mu = \int_0^1 p\, dG(p) \tag{2.2}$$

based *only* on the observations $X_1, \cdots, X_n$. Note that in fact there exists a bivariate random sample $\{(X_i, P_i), i = 1, \cdots, n\}$, where we assume that X_i conditional on $P_i = p$ is binomially distributed with m trials and success probability p and the P_i are independent marginally distributed as $G(p)$. We write $X_i|P_i = p \sim b(m, p)$ and $P_i \sim G(p)$, $i = 1, \cdots, n$.

We employ the method of moments to obtain our estimator. Observe that

$$E(X_1) = E\{E[X_1|P_1]\} = mE(P_1) = m\mu. \tag{2.3}$$

From (2.3) and the fact that $X_1, \cdots, X_n$ are independent and identically distributed with distribution (2.1), we have $E\bar{X} = m\mu$. Hence, the method of moments yields the estimator $\hat{\mu}$ of μ given by

$$\hat{\mu} = \frac{\bar{X}}{m} = \frac{1}{mn} \sum_{i=1}^{n} X_i. \tag{2.4}$$

From the strong law of large numbers, we have immediately that $\hat{\mu}$ is a strongly consistent estimator of μ. That is,

$$P(\lim_{n \to \infty} \hat{\mu} = \mu) = 1. \tag{2.5}$$

Next we will obtain the large sample distribution of $\hat{\mu}$ quite directly from the central limit theorem. First, observe that the variance of $\hat{\mu}$ is given by

$$\operatorname{Var}(\hat{\mu}) = \operatorname{Var}\left(\frac{\bar{X}}{m}\right) = \frac{\sigma_1^2}{m^2 n}, \tag{2.6}$$

where $\sigma_1^2 = \operatorname{Var}(X_1)$. Therefore, by the central limit theorem for independent, identically distributed random variables with finite variance, we have as $n \to \infty$

$$\mathcal{L}\left(\frac{m\sqrt{n}(\hat{\mu} - \mu)}{\sigma_1}\right) \to N(0, 1), \tag{2.7}$$

where $\mathcal{L}(Z_n) \to N(\mu, \sigma^2)$ means $\{Z_n\}$ converges in distribution to a random variable Z which is normally distributed with mean μ and variance σ^2.

Besides $\hat{\mu}$ being strongly consistent as in (2.5) and asymptotically normal as in (2.7), we note that $\hat{\mu}$ in (2.4) is the minimum variance unbiased linear (in the X_i) estimator of μ. This is a direct result of Theorem 3.1 in Section 3.

Furthermore, by defining

$$S^2 = (n - 1)^{-1} \sum_{i=1}^{n} (X_i - \bar{X})^2, \tag{2.8}$$

we have S^2 in an unbiased, consistent estimator of σ_1^2. That is,

$$S^2 \xrightarrow{P} \sigma_1^2 \tag{2.9}$$

as $n \to \infty$. Using (2.9) and (2.7), we obtain (see for example, Rao [3], $(x) - (b)$, p. 102) as $n \to \infty$

$$\mathcal{L}\left(\frac{m\sqrt{n}(\hat{\mu} - \mu)}{S}\right) \to N(0, 1). \tag{2.10}$$

From (2.10), we can immediately give a $100(1 - \alpha)$ per cent, $0 < \alpha < 1$, large sample confidence interval for μ, since

$$\lim_{n \to \infty} P\left\{\frac{1}{m}\left(\bar{X} - \frac{1}{\sqrt{n}} Z_{\alpha/2} S\right) < \mu < \frac{1}{m}\left(\bar{X} + \frac{1}{\sqrt{n}} Z_{\alpha/2} S\right)\right\} = 1 - \alpha, \tag{2.11}$$

where $Z_{\alpha/2}$ is defined by the equation

$$\int_{-Z_{\alpha/2}}^{Z_{\alpha/2}} \left(\frac{1}{2\pi}\right)^{1/2} \exp\left\{-\frac{t^2}{2}\right\} dt = 1 - \alpha. \tag{2.12}$$

For example, if we want a 95 per cent confidence interval for the mean probability of the flock for laying a blood spotted egg, we would choose $\alpha = 0.05$ yielding an approximate large sample confidence interval

$$\left(\frac{1}{m}\left(\overline{X} - \frac{1.96S}{\sqrt{n}}\right), \frac{1}{m}\left(\overline{X} + \frac{1.96S}{\sqrt{n}}\right)\right). \tag{2.13}$$

Some sample intervals are constructed for data randomly generated by a similation process involving varying sample sizes and are here included.

The density given by (2.1) was computed for the case of $G(p)$ being a beta distribution. That is, assume $dG(p) = g(p)\,dp$, where the density $g(p)$ is given by

$$g(p) = \{\beta(r, s)\}^{-1}p^{r-1}(1-p)^{s-1}, \qquad 0 < p < 1, \qquad r > 0, \qquad s > 0, \tag{2.14}$$

and $\beta(r, s) = \int_0^1 p^{r-1}(1-p)^{s-1}\,dp$. A range of values of m, r, and s was chosen and μ computed for each set of values thereof. Random samples were drawn and $\hat{\mu}$ and Var $(\hat{\mu})$ estimated. Table I gives μ, $\hat{\mu}$, and the estimated standard error of $\hat{\mu}$, $S/(m\sqrt{n})$, for a few selected values of m, r, and s. The entry on the first line is for a sample of $n = 50$ and on the second line for a sample of $n = 200$. In most instances $\hat{\mu}$ estimates μ well; $\hat{\mu} \pm 1.96S/m\sqrt{n}$ fails to contain μ only three times out of the 40 cases presented. That is, when $m = 10$, $r = s = 1$, $n = 50$; $m = 15$, $r = 1$, $s = 5$, $n = 200$; and $m = 15$, $r = 2$, $s = 15$, $n = 200$. Many other values of m, r, s, and n were also tried with similar good results.

3. The case of differing sample sizes

Often times in applications the number of trials m connected with each observation may not be the same. That is, consider the case where each X_i is distributed as (2.1) with m replaced by m_i, $i = 1, \cdots, n$. We assume the m_i are all known, fixed, positive integers, but not necessarily equal.

To estimate μ, we again use the method of moments. Similar to (2.3), we have

$$E(X_i) = m_i\mu, \qquad i = 1, \cdots, n. \tag{3.1}$$

Since (3.1) implies both $E\{\sum_{i=1}^n (X_i/m_i)\} = n\mu$ and $E(\sum_{i=1}^n X_i) = (\sum_{i=1}^n m_i)\mu$, we have as possible moment estimators of μ both

$$\hat{\mu}_1 = \frac{1}{n}\sum_{i=1}^n \frac{X_i}{m_i} \tag{3.2}$$

and

$$\hat{\mu}_2 = \frac{\overline{X}}{\overline{m}}, \tag{3.3}$$

where $\overline{m} = (1/n)\sum_{i=1}^n m_i$.

In order to discuss the relative merits of the estimators $\hat{\mu}_1$ and $\hat{\mu}_2$, we compute their variances. With

$$\sigma^2 = \text{Var}(P_1) = \int (p - \mu)^2\,dG(p) \tag{3.4}$$

and

$$\tau = E\{P_1(1 - P_1)\} = \int p(1-p)\,dG(p), \tag{3.5}$$

TABLE I

COMPARISON OF μ AND $\hat{\mu}$ FOR CASE OF BETA DISTRIBUTION, $n = 50$ AND $n = 200$

m	r	s	μ_p	$\hat{\mu}_p$	$S/(m\sqrt{n})$
10	1	1	.500	.594	.0411
				.505	.0217
		2	.333	.314	.0366
				.342	.0198
		5	.167	.170	.0241
				.158	.0125
		10	.091	.100	.0204
				.093	.0087
		15	.063	.060	.0121
				.060	.0068
	2	1	.667	.698	.0366
				.688	.0177
		2	.500	.548	.0388
				.482	.0181
		5	.286	.314	.0287
				.298	.0146
		10	.167	.178	.0225
				.154	.0106
		15	.118	.120	.0232
				.116	.0082
15	1	1	.500	.489	.0470
				.474	.0218
		2	.333	.295	.0397
				.327	.0179
		5	.167	.197	.0268
				.190	.0129
		10	.091	.116	.0209
				.096	.0082
		15	.063	.052	.0107
				.062	.0057
	2	1	.667	.653	.0341
				.657	.0183
		2	.500	.449	.0386
				.485	.0165
		5	.286	.320	.0306
				.285	.0136
		10	.167	.192	.0225
				.167	.0102
		15	.118	.112	.0175
				.118	.0075

we have

(3.6)
$$\operatorname{Var}(\hat{\mu}_2) = (\overline{m}n)^{-2} \sum_{i=1}^{n} \operatorname{Var}(X_i) = n^{-1}(a_n\sigma^2 + b_n\tau),$$

where $a_n = (\overline{m}^2 n)^{-1} \sum_{i=1}^n m_i^2$, $b_n = (\overline{m})^{-1}$, $\overline{m} = n^{-1} \sum_{i=1}^n m_i$.

Also, we obtain

$$\text{Var}(\hat{\mu}_1) = n^{-2} \sum_{i=1}^{n} \text{Var}(X_i) = n^{-1}(\sigma^2 + c_n\tau), \tag{3.7}$$

where $c_n = n^{-1} \sum_{i=1}^{n} m_i^{-1}$.

From the Cauchy-Schwarz inequality, it is easy to show that

$$a_n \geqq 1 \quad \text{and} \quad b_n \leqq c_n, \tag{3.8}$$

where the inequalities are strict unless $m_i = m$ for all i. Hence, it is clear that neither $\hat{\mu}_1$ nor $\hat{\mu}_2$ is for all G relatively more efficient than the other. In fact, from (3.8) we see that if $\sigma^2 = 0$ and $\tau > 0$, $\hat{\mu}_2$ is more efficient, $\text{Var}(\hat{\mu}_2) \leqq \text{Var}(\hat{\mu}_1)$, than $\hat{\mu}_1$, while the reverse is true if $\sigma^2 > 0$ and $\tau = 0$.

Note the case $\sigma^2 = 0$ implies that $G(p)$ is degenerate at, say, p_0, and hence that $S_n = \sum_{i=1}^{n} X_i$ is binomially distributed with parameter $\sum_{i=1}^{n} m_i$ and p_0. Thus, $\hat{\mu}_2$ in this case becomes the classical (maximum likelihood, moment and minimum variance unbiased estimator) solution to the problem of estimating $\mu = p_0$. *In the remainder of the paper, we omit this case from consideration and shall assume* $\sigma^2 > 0$.

We consider now the question of the existence of an optimal solution in the minimum variance sense. The following theorem gives a solution to the problem for unbiased linear (in X_i) estimators.

Let M_n be the class of all unbiased linear estimators of μ based on $X_1, \cdots, X_n$. That is,

$$M_n = \left\{\hat{\mu} \middle| \hat{\mu} = \sum_{i=1}^{n} c_{in} \frac{X_i}{m_i}, \sum_{i=1}^{n} c_{in} = 1\right\}. \tag{3.9}$$

Observe that the condition $\sum_{i=1}^{n} c_{in} = 1$ implies that $\hat{\mu}$ is unbiased by (3.1). Also, $\hat{\mu} \in M_n$ is clearly linear in the X_i as well as in the X_i/m_i by taking $c'_{in} = c_{in}/m_i$ and defining $\hat{\mu} = \sum_{i=1}^{n} c'_{in} X_i$ in M_n.

THEOREM 3.1. *The minimum variance unbiased linear estimate of* μ *(that is, the* $\hat{\mu} \in M_n$ *of minimum variance) is given by*

$$\hat{\mu}_0 = \sum_{i=1}^{n} \mathring{c}_{in} \frac{X_i}{m_i}, \tag{3.10}$$

where

$$\mathring{c}_{in} = \mathring{c}_{in}(\sigma^2, \tau) = \left\{\sigma^2 + \frac{\tau}{m_i}\right\}^{-1} \Big/ \sum_{i=1}^{n} \left\{\sigma^2 + \frac{\tau}{m_i}\right\}^{-1}, \tag{3.11}$$

with σ^2 *and* τ *as in* (3.4) *and* (3.5).

REMARK. In particular, $\hat{\mu}_0 = \hat{\mu}_1 = \hat{\mu}_2$ if all $m_i = m$ (see Section 1), and $\hat{\mu}_0 = \hat{\mu}_1$ if $\sigma^2 > 0$, $\tau = 0$, and $\hat{\mu}_0 = \hat{\mu}_2$ if $\sigma^2 = 0$, $\tau > 0$.

PROOF. Let $\sigma_i^2 = \text{Var}(X_i/m_i) = \sigma^2 + \tau/m_i$. Then, for $\hat{\mu} \in M_n$, we have $\text{Var}(\hat{\mu}) = \sum_{i=1}^{n} c_{in}^2 \sigma_i^2$, which is minimized by taking $c_{in} = \mathring{c}_{in}$ as in (3.11). (See for example, Rao [3], 2.2, p. 249.)

THEOREM 3.2. *If* $\sigma^2 > 0$, *then* $P\{\lim_{n\to\infty} \hat{\mu}_0 = \mu\} = 1$ *for any sequence* $\{m_n\}$ *of positive integers.*

PROOF. Let $Y_i = \sigma_i^{-2}(X_i/m_i)$, where $\sigma_i^2 = \text{Var}(X_i/m_i) = \sigma^2 + \tau/m_i$. Le $b_n = \sum_{i=1}^n \sigma_i^{-2}$ and observe that (see Loève [1], 16.3, II, A, p. 238) $\hat{\mu}_0 - \mu = b_n^{-1} \sum_{i=1}^n (Y_i - EY_i) \to 0$ with probability 1 as $n \to \infty$ provided

$$\sum_{n=1}^{\infty} b_n^{-2} \text{Var}(Y_n) < \infty. \tag{3.12}$$

But since $b_n \geq n(\sigma^2 + \tau)^{-1}$ and $\text{Var}(Y_i) = \sigma_i^{-2} \leq \sigma^{-2}$, we see that (3.12) holds since $\sum_{n=1}^{\infty} n^{-2} < \infty$.

THEOREM 3.3. *If $\sigma^2 > \sigma$, then for any sequence of positive integers $\{m_n\}$,*

$$\mathcal{L}\left(\frac{\hat{\mu}_0 - \mu}{(\text{Var}(\hat{\mu}_0))^{1/2}}\right) \to N(0, 1) \tag{3.13}$$

as $n \to \infty$, where

$$\text{Var}(\hat{\mu}_0) = \left\{\sum_{i=1}^n \sigma_i^{-2}\right\}^{-1} = \left\{\sum_{i=1}^n (\sigma^2 + \tau/m_i)^{-1}\right\}^{-1}. \tag{3.14}$$

PROOF. Let $Y_{in} = c_{in}^{\circ}(m_i^{-1}X_i - \mu)$, $\sigma_{in}^2 = \text{Var}(Y_{in})$, and $s_n^2 = \sum_{i=1}^n \sigma_{in}^2$. Then, by an extended version of the Liapounov theorem (Loève [1], 20.1, a, p. 277), we have

$$\mathcal{L}\left(\frac{\hat{\mu}_0 - \mu_0}{(\text{Var}(\hat{\mu}_0))^{1/2}}\right) = \mathcal{L}\left(\frac{\sum_{i=1}^n Y_{in}}{s_n}\right) \to N(0, 1) \tag{3.15}$$

as $n \to \infty$ provided

$$s_n^{-3} \sum_{i=1}^n E|Y_{in}|^3 \to 0 \tag{3.16}$$

as $n \to \infty$. But since $s_n^2 = \{\sum_{i=1}^n \sigma_i^{-2}\}^{-1}$ and $c_{in}^{\circ} = s_n^2 \sigma_i^{-2}$ with $\sigma_i^2 = \sigma^2 + \tau/m_i$, we have

$$\begin{aligned} s_n^{-3} \sum_{i=1}^n E|Y_{in}|^3 &= s_n^3 \sum_{i=1}^n \sigma_i^{-6} E\left|\frac{X_i}{m_i} - \mu\right|^3 \\ &\leq s_n^3 \sum_{i=1}^n \sigma_i^{-4} \\ &\leq n^{-1/2}(\sigma^2 + \tau)^{3/2}\sigma^{-4}, \end{aligned} \tag{3.17}$$

where the last inequality follows by using $(\sigma^2 + \tau)^{-1} \leq \sigma_i^{-2} \leq \sigma^{-2}$. Hence, (3.16) holds and the theorem is proved.

We note that Theorem 3.3 immediately yields large sample confidence intervals on μ provided σ^2 and τ are known. Under the condition of Theorem 3.3 a $100(1 - \alpha)$ per cent large sample approximate confidence interval for μ is given by

$$(\hat{\mu}_0 - \varepsilon_n, \hat{\mu}_0 + \varepsilon_n), \tag{3.18}$$

where $\varepsilon_n = Z_{\alpha/2}\{\sum_{i=1}^n (\sigma^2 + \tau/m_i)^{-1}\}^{-1/2}$ and $\hat{\mu}_0 = \sum_{i=1}^n c_{in}^{\circ}(X_i/m_i)$. However, in most applications σ^2 and τ remain unknown and we must therefore concern ourselves with this case. Section 4 discusses the question of estimating σ^2 and τ, while Section 5 develops the necessary large sample results for estimating μ when σ^2 and τ are unknown and the sample sizes m_i differ.

4. Estimation of σ^2 and τ

Define the random variables Y_i, Z_i and the indicator variables δ_i, $i = 1, \cdots, n$, as follows:

$$Y_i = \frac{X_i}{m_i}, \tag{4.1}$$

$$Z_i = \begin{cases} \dfrac{X_i(m_i - X_i)}{m_i(m_i - 1)} & \text{if } m_i > 1, \\ 0 & \text{if } m_i = 1, \end{cases} \tag{4.2}$$

and

$$\delta_i = \begin{cases} 1 & \text{if } m_i > 1, \\ 0 & \text{if } m_i = 1. \end{cases} \tag{4.3}$$

Observing that $EX_i^2 = EX_i(X_i - 1) + EX_i = m_i^2 E(P_1^2) + m_i\tau$, one obtains

$$EZ_i = \delta_i\tau. \tag{4.4}$$

Furthermore, using the relationship

$$\sum_{i=1}^n (Y_i - \bar{Y})^2 = \sum_{i=1}^n (Y_i - \mu)^2 - n(\bar{Y} - \mu)^2, \tag{4.5}$$

it can easily be shown that

$$E\left\{\sum_{i=1}^n (Y_i - \bar{Y})^2\right\} = (n - 1)\left\{\sigma^2 + \left(\frac{1}{n}\sum_{i=1}^n \frac{1}{m_i}\right)\tau\right\}. \tag{4.6}$$

From equations (4.4) and (4.6) and defining

$$S_1^2 = (n - 1)^{-1} \sum_{i=1}^n (Y_i - \bar{Y})^2 \tag{4.7}$$

and

$$a_n = \max\left\{\sum_{i=1}^n \delta_i, 1\right\}, \tag{4.8}$$

we obtain as moment estimators of τ and σ^2, when $\sum_{i=1}^n \delta_i > 0$,

$$\hat{\tau} = a_n^{-1} \sum_{i=1}^n Z_i \tag{4.9}$$

and

$$(\sigma^*)^2 = S_1^2 - \hat{\tau}\left(\frac{1}{n}\sum_{i=1}^n \frac{1}{m_i}\right). \tag{4.10}$$

Note that from (4.4) and (4.6) the unbiasedness of τ and $(\sigma^*)^2$ follows. That is, when $\sum_{i=1}^n \delta_i > 0$,

$$E(\hat{\tau}) = \tau, \qquad E\{(\sigma^*)^2\} = \sigma^2. \tag{4.11}$$

The estimator $(\sigma^*)^2$ in (4.9) may be negative as an estimator of $\sigma^2 > 0$. We shall find it convenient to modify $(\sigma^*)^2$ for later purposes and we define $\hat{\sigma}^2$ as the following positive truncation of $(\sigma^*)^2$,

$$\hat{\sigma}^2 = \max\{(\sigma^*)^2, n^{-1}\}. \tag{4.12}$$

The following theorem gives the consistency properties of $\hat{\tau}$, $\hat{\sigma}^2$ (and $(\sigma^*)^2$).

THEOREM 4.1. *Let $\sigma^2 > 0$. If $\{m_n\}$ is any sequence of positive integers for which $a_n \to \infty$ as $n \to \infty$ (see (4.8)), then as $n \to \infty$, $\hat{\tau} \to \tau$ and $\hat{\sigma}^2 \to \sigma^2$ (or $(\sigma^*)^2 \to \sigma^2$) in probability. Furthermore, if $\{m_n\}$ is such that $\sum_{n=1}^{\infty} a_n^{-2} < \infty$, the convergences hold with probability one.*

PROOF. Observe that since $0 \leqq Z_i \leqq \frac{1}{2}$, we have

$$\operatorname{Var}(\hat{\tau}) = a_n^{-2} \sum_{i=1}^{n} \delta_i \operatorname{Var}(Z_i) \leqq (4a_n)^{-1}. \tag{4.13}$$

Hence, $\hat{\tau} \to \tau$ in probability as $n \to \infty$ by Chebyshev's inequality. To prove convergence with probability one for $\hat{\tau}$, it suffices (by Loève [1], 16.3, II, A, p. 238) to verify that $\sum_{n=1}^{\infty} a_n^{-2} \operatorname{Var}(Z_n) < \infty$, which clearly holds in $\sum_{n=1}^{\infty} a_n^{-2} < \infty$, since $\operatorname{Var}(Z_n) \leqq \frac{1}{4}$.

Observe that $(\sigma^*)^2$ is linear in $\hat{\tau}$ in (4.10). Thus, convergence of $(\sigma^*)^2$ to σ^2 in probability ($\xrightarrow{\text{P}}$) or with probability one ($\xrightarrow{\text{a.s.}}$) as $n \to \infty$ follows from Theorem 3.2 provided

$$S_1^2 - \left(\sigma^2 + \frac{\tau}{n} \sum_{i=1}^{n} \frac{1}{m_i}\right) \xrightarrow{\text{a.s.}} 0 \tag{4.14}$$

as $n \to \infty$. But (4.14) follows immediately from (4.5), (4.6), and (4.7) provided as $n \to \infty$,

$$\frac{1}{n} \sum_{i=1}^{n} (Y_i - \mu)^2 - \left(\sigma^2 + \frac{\tau}{n} \sum_{i=1}^{n} \frac{1}{m_i}\right) \xrightarrow{\text{a.s.}} 0. \tag{4.15}$$

Let $X_i' = (Y_i - \mu)^2$ in (4.15) and write the left side of (4.15) as $n^{-1} \sum_{i=1}^{n} (X_i' - EX_i')$. Now, applying a version of the Kolmogorov strong law of large numbers (see Loève [1], 16.3, II, A, p. 238), we see (4.15) holds provided

$$\sum_{n=1}^{\infty} n^{-2} \operatorname{Var}(X_n') < \infty. \tag{4.16}$$

But the convergence of the series in (4.16) is an immediate consequence of the boundedness of X_n' by one. Thus, the theorem is proved.

REMARK. We observe that the condition $a_n \to \infty$ is necessary in Theorem 4.1. To see this consider the case where all the m_i are 1 or 2. Then $a_n \not\to \infty$ implies there exist n_0 such that $\delta_n = 0$ ($m_n = 1$) for $n \geqq n_0$. Thus $\hat{\tau} = a_{n_0}^{-1} \sum_{i=1}^{n_0} Z_i$ for all $n \geqq n_0$ and clearly $\hat{\tau} \overset{\text{P}}{\not\to} \tau$ as $n \to \infty$.

5. Estimation of μ when σ^2 and τ are unknown in the differing sample size case

The minimum variance unbiased linear estimate of μ in Theorem 3.1 depends on *knowing* σ^2 and τ for the optimal choice of the constants c_{in}° in (3.11). To overcome this problem when σ^2 and τ are unknown, we propose and study an estimator of μ, denoted μ_0^*, which chooses the c_{in}° in (3.11) based on the estimators $\hat{\sigma}^2$, $\hat{\tau}$ of σ^2, τ given in the previous section. Theorems 5.1 and 5.2 give the large sample properties of the proposed estimator.

Specifically, let

$$\hat{c}_{in} = c_{in}^\circ(\hat{\sigma}^2, \hat{\tau}) = \begin{cases} \dfrac{1}{n} & \text{if } \hat{\tau} = 0, \\[2ex] \dfrac{\left(\hat{\sigma}^2 + \dfrac{\hat{\tau}}{m_i}\right)^{-1}}{\sum_{i=1}^n \left(\hat{\sigma}^2 + \dfrac{\hat{\tau}}{m_i}\right)^{-1}} & \text{if } \hat{\tau} > 0, \end{cases} \tag{5.1}$$

where $\hat{\tau}$ and $\hat{\sigma}^2$ are defined by (4.9), (4.10), and (4.12). Now define

$$\mu_0^* = \sum_{i=1}^n \hat{c}_{in} \frac{X_{i\cdot}}{m_i} \tag{5.2}$$

It will be shown that under appropriate conditions $\mu_0^* \to \mu$ in probability as $n \to \infty$ (see Theorem 5.1). Before proving this theorem, however, we develop the following lemma.

LEMMA 5.1. *If $\sigma^2 > 0$ and $a_n \to \infty$ as $n \to \infty$ (see (4.3) and (4.8)), then*

$$\begin{aligned} \mu_0^* = \hat{\mu}_0 &+ (\hat{\sigma}^2 - \sigma^2) \sum_{i=1}^n \alpha_{in} \left(\frac{X_i}{m_i} - \mu\right) \\ &+ (\hat{\tau} - \tau) \sum_{i=1}^n \beta_{in} \left(\frac{X_i}{m_i} - \mu\right) \\ &+ [|\hat{\sigma}^2 - \sigma^2| + |\hat{\tau} - \tau|]^2 O_p(1) \hat{\mu}_1, \end{aligned} \tag{5.3}$$

where α_{in} and β_{in} are nonrandom coefficients such that $\sum_{i=1}^n \alpha_{in} = \sum_{i=1}^n \beta_{in} = 0$ and $O_p(1)$ indicates a random factor which is bounded in probability.

PROOF. Let

$$\begin{aligned} c_{in}(u, v) &= \left\{\sum_{i=1}^n \left(u + \frac{v}{m_i}\right)^{-1}\right\}^{-1} \left(u + \frac{v}{m_i}\right)^{-1}, \\ \alpha_{in} &= \left.\frac{\partial c_{in}}{\partial u}\right|_{u=\sigma^2, v=\tau}, \\ \beta_{in} &= \left.\frac{\partial c_{in}}{\partial v}\right|_{u=\sigma^2, v=\tau}. \end{aligned} \tag{5.4}$$

Observe that $c_{in}(u, v)$ has continuous second order partial (and mixed partial) derivatives on the set $\{(u, v): 0 < u < \infty, 0 \leq v < \infty\}$, where the partial derivative is defined from the right at $v = 0$. Hence, we have the following second order Taylor expansion for $c_{in}(\hat{\sigma}^2, \hat{\tau})$,

$$\begin{aligned} c_{in}(\hat{\sigma}^2, \hat{\tau}) = c_{in}(\sigma^2, \tau) &+ (\hat{\sigma}^2 - \sigma^2)\alpha_{in} + (\hat{\tau} - \tau)\beta_{in} \\ &+ \frac{(\hat{\sigma}^2 - \sigma^2)^2}{2} \left.\frac{\partial^2 c_{in}}{\partial u^2}\right|_{u=u_i^*, v=v_i^*} \end{aligned} \tag{5.5}$$

$$+ (\hat{\sigma}^2 - \sigma^2)(\hat{\tau} - \tau) \left.\frac{\partial^2 c_{in}}{\partial u\, \partial v}\right|_{u=u_i^*, v=v_i^*}$$

$$+ \frac{(\hat{\tau} - \tau)^2}{2} \left.\frac{\partial^2 c_{in}}{\partial v^2}\right|_{u=u_i^*, v=v_i^*},$$

where $\min(\sigma^2, \hat{\sigma}^2) \leqq u_i^* \leqq \max(\sigma^2, \hat{\sigma}^2)$ and $\min(\tau, \hat{\tau}) \leqq v_i^* \leqq \max(\tau, \hat{\tau})$. Observe that

$$(5.6)\quad \frac{\partial c_{in}}{\partial u} = \frac{\left\{\left(u + \frac{v}{m_i}\right)^{-1} \sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-2}\right\} - \left\{\left(u + \frac{v}{m_i}\right)^{-2} \sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-1}\right\}}{\left\{\sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-1}\right\}^2},$$

(5.7)

$$\frac{\partial c_{in}}{\partial v} = \frac{\left\{\left(u + \frac{v}{m_i}\right)^{-1} \sum_{j=1}^{n} \frac{1}{m_j}\left(u + \frac{v}{m_j}\right)^{-2}\right\} - \left\{\frac{1}{m_i}\left(u + \frac{v}{m_i}\right)^{-2} \sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-1}\right\}}{\left\{\sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-1}\right\}^2},$$

$$(5.8)\quad \frac{\partial^2 c_{in}}{\partial u^2} = \frac{2\left(u + \frac{v}{m_i}\right)^{-3}}{\sum_{j=1}^{n} \left(u + \frac{v}{m_i}\right)^{-1}} - \frac{2\left(u + \frac{v}{m_i}\right)^{-1} \sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-3}}{\left\{\sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-1}\right\}^2}$$

$$+ \frac{2\left(u + \frac{v}{m_i}\right)^{-1} \left\{\sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-2}\right\}^2}{\left\{\sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-1}\right\}^3} - \frac{2\left(u + \frac{v}{m_i}\right)^{-2} \sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-2}}{\left\{\sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-1}\right\}^2},$$

$$(5.9)\quad \frac{\partial^2 c_{in}}{\partial v^2} = \frac{\frac{2}{m_i^2}\left(u + \frac{v}{m_i}\right)^{-3}}{\sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-1}} - \frac{2\left(u + \frac{v}{m_i}\right)^{-1} \sum_{j=1}^{n} \frac{1}{m_j^2}\left(u + \frac{v}{m_j}\right)^{-3}}{\left\{\sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-1}\right\}^2}$$

$$+ \frac{2\left(u + \frac{v}{m_i}\right)^{-1} \left\{\sum_{j=1}^{n} \frac{1}{m_j}\left(u + \frac{v}{m_j}\right)^{-2}\right\}^2}{\left\{\sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-1}\right\}^3} - \frac{\frac{2}{m_i}\left(u + \frac{v}{m_i}\right)^{-2} \sum_{j=1}^{n} \frac{1}{m_j}\left(u + \frac{v}{m_j}\right)^{-2}}{\left\{\sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-1}\right\}^2},$$

and

$$(5.10)\quad \frac{\partial^2 c_{in}}{\partial u\, \partial v} = \frac{\frac{2}{m_i}\left(u + \frac{v}{m_i}\right)^{-3}}{\sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-1}} - \frac{2\left(u + \frac{v}{m_i}\right)^{-1} \sum_{j=1}^{n} \frac{1}{m_j}\left(u + \frac{v}{m_j}\right)^{3}}{\left\{\sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-1}\right\}^2}$$

$$+ \frac{\left(u + \frac{v}{m_i}\right)^{-2} \sum_{j=1}^{n} \frac{1}{m_j}\left(u + \frac{v}{m_j}\right)^{-2}}{\left\{\sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-1}\right\}^2} - \frac{\frac{1}{m_i}\left(u + \frac{v}{m_i}\right)^{-2} \sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-2}}{\left\{\sum_{j=1}^{n} \left(u + \frac{v}{m_j}\right)^{-1}\right\}^2}$$

$$+\frac{2\left(u+\frac{v}{m_i}\right)^{-1}\left\{\sum_{j=1}^{n}\left(u+\frac{v}{m_j}\right)^{-2}\right\}\left\{\sum_{j=1}^{n}\frac{1}{m_j}\left(u+\frac{v}{m_j}\right)^{-2}\right\}}{\left\{\sum_{j=1}^{n}\left(u+\frac{v}{m_j}\right)^{-1}\right\}^3}$$

$$-\frac{2\left(u+\frac{v}{m_i}\right)^{-2}\sum_{j=1}^{n}\frac{1}{m_j}\left(u+\frac{v}{m_j}\right)^{-2}}{\left\{\sum_{j=1}^{n}\left(u+\frac{v}{m_j}\right)^{-1}\right\}^2}.$$

Note that from (5.6) and (5.7), we see that $\sum_{i=1}^{n}\alpha_{in}=\sum_{i=1}^{n}\beta_{in}=0$. Thus, we have from (5.2) and the Taylor expansion (5.5),

$$(5.11)\quad \mu_0^* - \left\{\hat{\mu}_0 + (\hat{\sigma}^2-\sigma^2)\sum_{i=1}^{n}\alpha_{in}\left(\frac{X_i}{m_i}-\mu\right) + (\hat{\tau}-\tau)\sum_{i=1}^{n}\beta_{in}\left(\frac{X_i}{m_i}-\mu\right)\right\}$$
$$= \frac{(\hat{\sigma}^2-\sigma^2)^2}{2}\sum_{i=1}^{n}\frac{X_i}{m_i}\left(\frac{\partial^2 c_{in}}{\partial u^2}\bigg|_{u=u_i^*,v=v_i^*}\right)$$
$$+(\hat{\sigma}^2-\sigma^2)(\hat{\tau}-\tau)\sum_{i=1}^{n}\frac{X_i}{m_i}\left(\frac{\partial^2 c_{in}}{\partial u\,\partial v}\bigg|_{u=u_i^*,v=v_i^*}\right)$$
$$+\frac{(\hat{\tau}-\tau)^2}{2}\sum_{i=1}^{n}\frac{X_i}{m_i}\left(\frac{\partial^2 c_{in}}{\partial v^2}\bigg|_{u=u_i^*,v=v_i^*}\right).$$

But the right side of (5.11) is bounded by

$$(5.12)\qquad \{|\hat{\sigma}^2-\sigma^2|+|\hat{\tau}-\tau|\}^2\left\{\frac{5(\underline{u})^{-3}}{(\overline{u}+\overline{v})^{-1}}\right\}\hat{\mu}_1,$$

where $\overline{u}=\max_i u_i^*$, $\underline{u}=\min_i u_i^*$, and $\overline{v}=\max_i v_i^*$, since from (5.8) through (5.10) and repeated use of the inequalities

$$(5.13)\qquad (u_i^*+v_i^*)^{-1}\leqq\left(u_i^*+\frac{v_i^*}{m_i}\right)^{-1}\leqq(u_i^*)^{-1},\qquad \frac{1}{m_i}\leqq 1,$$

it is easy to show that

$$(5.14)\qquad \left|\left(\frac{\partial^2 c_{in}}{\partial u^2}\bigg|_{u=u_i^*,v=v_i^*}\right)\right|\leqq\frac{8(u_i^*)^{-3}}{n(u_i^*+v_i^*)^{-1}},\qquad \left|\left(\frac{\partial^2 c_{in}}{\partial v^2}\bigg|_{u=u_i^*,v=v_i^*}\right)\right|\leqq\frac{8(u_i^*)^{-3}}{n(u_i^*+v_i^*)^{-1}},$$

and

$$(5.15)\qquad \left|\left(\frac{\partial^2 c_{in}}{\partial u\,\partial v}\bigg|_{u=u_i^*,v=v_i^*}\right)\right|\leqq\frac{10(u_i^*)^{-3}}{n(u_i^*+v_i^*)^{-1}}.$$

Bounding the right side of (5.11) by (5.12) and noting that Theorem 4.1 implies that as $n\to\infty$,

$$(5.16)\qquad \frac{(\underline{u})^{-3}}{(\overline{u}+\overline{v})^{-1}}\xrightarrow{P}\frac{\sigma^{-6}}{(\sigma^2+\tau)^{-1}}>0,$$

we have that $5(\underline{u})^{-3}(\overline{u}+\overline{v})$ is $O_p(1)$. Hence, the lemma is proved.

THEOREM 5.1. *If* $\sigma^2 > 0$ *and* $a_n \to \infty$ *as* $n \to \infty$ *(see* (4.3) *and* (4.8)*), then* μ_0^* *defined by* (5.1) *and* (5.2) *is such that* $\mu_0^* \xrightarrow{P} \mu$ *as* $n \to \infty$.

PROOF. Repeated use of $(\sigma^2 + \tau)^{-1} \leqq (\sigma^2 + \tau/m_i)^{-1} \leqq \sigma^{-2}$ and $m_i^{-1} \leqq 1$ in (5.6) and (5.7) imply $\sum_{i=1}^n |\alpha_{in}|$ and $\sum_{i=1}^n |\beta_{in}|$ are bounded by $\sigma^{-4}(\sigma^2 + \tau)^{-1}$. Hence, bounding $|(X_i/m_i) - \mu| \leqq 1$ in (5.3) and invoking the convergences in Theorems 3.2 and 4.1, expansion (5.3) yields the result.

LEMMA 5.2. *If* $n^{1/2}a_n^{-1} \to 0$ *as* $n \to \infty$, *then* $n^{1/4}(\hat\sigma^2 - \sigma^2) \xrightarrow{P} 0$ *and* $n^{1/4}(\hat\tau - \tau) \xrightarrow{P} 0$ *as* $n \to \infty$.

PROOF. From (4.2) we have $0 \leqq Z_i \leqq \frac{1}{2}$ and $\operatorname{Var}(Z_i) \leqq \frac{1}{4}$ and therefore,

$$\text{(5.17)} \qquad \operatorname{Var}(n^{1/4}\hat\tau) = n^{1/2}a_n^{-2} \sum_{i=1}^n \delta_i \operatorname{Var}(Z_i) \leqq \frac{1}{4} n^{1/2}a_n^{-1}.$$

Thus, by Chebyshev's inequality, we have

$$\text{(5.18)} \qquad n^{1/4}(\hat\tau - \tau) \xrightarrow{P} 0$$

as $n \to \infty$.

Next, in (4.5) and (4.7) observe that

$$\text{(5.19)} \qquad (n-1)S_1^2 = \sum_{i=1}^n (Y_i - \mu)^2 - n(\bar Y - \mu)^2 \leqq 2n(\bar Y - \mu)^2.$$

Since S_1^2 is a nonnegative random variable and $|\bar Y - \mu| \leqq 1$, it follows that

$$\text{(5.20)} \qquad \operatorname{Var}(S_1^2) \leqq ES_1^4 \leqq \left(\frac{2n}{n-1}\right)^2 \operatorname{Var}(\bar Y) = \left(\frac{2n}{n-1}\right)^2 \frac{1}{n}\left(\sigma^2 + \frac{1}{n}\sum_{i=1}^n \frac{\tau}{m_i}\right).$$

But this inequality together with $m_i^{-1} \leqq 1$ imply that $\operatorname{Var}(n^{1/4}S_1^2)$ is $O(n^{-1/2}(\sigma^2 + \tau))$. Hence, again by Chebyshev's inequality and (4.6), we have

$$\text{(5.21)} \qquad n^{1/4}\left\{S_1^2 - \left[\sigma^2 + \tau\left(\frac{1}{n}\sum_{i=1}^n \frac{1}{m_i}\right)\right]\right\} \xrightarrow{P} 0$$

as $n \to \infty$. This result together with (5.18) combine in (4.10) to yield $n^{1/4}[(\sigma^*)^2 - \sigma^2] \xrightarrow{P} 0$ as $n \to \infty$ which completes the proof of the lemma by using the definition of $\hat\sigma^2$.

THEOREM 5.2. *If* $a_n \to \infty$ *and* $n^{1/2}a_n^{-1} \to 0$ *as* $n \to \infty$ *(see* (4.3) *and* (4.8)*) and* $\sigma^2 > 0$, *then*

$$\text{(5.22)} \qquad \mathcal{L}\left(\frac{\mu_0^* - \mu}{(\operatorname{Var}(\hat\mu_0))^{1/2}}\right) \to N(0, 1)$$

as $n \to \infty$. *Furthermore, replacing*

$$\operatorname{Var}(\hat\mu_0) = 1 \Big/ \left[\sum_{i=1}^n \left(\sigma^2 + \frac{\tau}{m_i}\right)^{-1}\right]$$

by

$$\text{(5.23)} \qquad U^2 = 1 \Big/ \left[\sum_{i=1}^n \left(\hat\sigma^2 + \frac{\hat\tau}{m_i}\right)^{-1}\right],$$

the result (5.22) *still holds.*

PROOF. Using the Taylor expansion (5.3) of Lemma 5.1, write

$$(5.24)\quad \frac{\mu_0^* - \mu}{(\operatorname{Var}(\hat\mu_0))^{1/2}} = \frac{\hat\mu_0 - \mu}{(\operatorname{Var}(\hat\mu_0))^{1/2}} + \frac{n^{1/4}(\hat\sigma^2 - \sigma^2)\left[n^{1/4}\sum_{i=1}^n \alpha_{in}\left(\frac{X_i}{m_i} - \mu\right)\right]}{(n\operatorname{Var}(\hat\mu_0))^{1/2}}$$

$$+ \frac{n^{1/4}(\hat\tau - \tau)\left[n^{1/4}\sum_{i=1}^n \beta_{in}\left(\frac{X_i}{m_i} - \mu\right)\right]}{(n\operatorname{Var}(\hat\mu_0))^{1/2}}$$

$$+ \frac{\{n^{1/4}|\hat\sigma^2 - \sigma^2| + n^{1/4}|\hat\tau - \tau|\}^2 O_p(1)\hat\mu_1}{(n\operatorname{Var}(\hat\mu_0))^{1/2}}.$$

From (5.24) and Theorem 3.3, the theorem will be completed provided the last three terms on the right side of (5.24) converge to zero in probability as $n \to \infty$. However, such is the case from Lemma 5.2 provided

$$(5.25)\quad \liminf\,\{n\operatorname{Var}(\hat\mu_0)\} \geqq c > 0$$

and that as $n \to \infty$,

$$(5.26)\quad \left\{n^{1/4}\sum_{i=1}^n \alpha_{in}\left(\frac{X_i}{m_i} - \mu\right)\right\} \xrightarrow{P} 0$$

and

$$(5.27)\quad \left\{n^{1/4}\sum_{i=1}^n \beta_{in}\left(\frac{X_i}{m_i} - \mu\right)\right\} \xrightarrow{P} 0.$$

The result (5.25) follows by noting that

$$(5.28)\quad n\operatorname{Var}(\hat\mu_0) = n\left\{\sum_{i=1}^n \left(\sigma^2 + \left(\frac{\tau}{m_i}\right)\right)^{-1}\right\}^{-1} \geqq \sigma^2 > 0 \qquad \text{for all } n.$$

The result (5.26) follows immediately from Chebyshev's inequality upon observing from (5.6) we have $|\alpha_{in}| \leqq 2\sigma^{-4}/n(\sigma^2 + \tau)^{-1}$, which implies

$$(5.29)\quad \operatorname{Var}\left\{\alpha_{in}\left(\frac{X_i}{m_i} - \mu\right)\right\} = \alpha_{in}^2\left(\sigma^2 + \frac{\tau}{m_i}\right)$$

$$\leqq \frac{4\sigma^{-6}}{n^2(\sigma^2 + \tau)^{-2}}.$$

A similar argument implies (5.27) holds. This completes the proof of the theorem.

REMARK. A $100(1 - \alpha)$ per cent, $0 < \alpha < 1$, large sample confidence interval for μ when σ^2 and τ are *unknown* which is close to being optimal in the sense that asymptotically it is the same as that based on $\hat\mu_0$ in (3.10) (the minimum variance unbiased linear estimator of μ) is given from (5.22) and (5.23) by

$$(5.30)\quad \lim_{n\to\infty} P\{\mu_0^* - Z_{\alpha/2}U < \mu < \mu_0^* + Z_{\alpha/2}U\} = 1 - \alpha,$$

where μ_0^*, U, and $Z_{\alpha/2}$ are defined in (5.2), (5.23), and (2.12), respectively.

6. Summary

This paper has examined the question of estimating the mean μ in (2.2) of a random binomial parameter having distribution $G(p)$. Such a problem arose in the context of measuring an unobservable genetic trait in flocks of chickens. The sampling scheme upon which our procedures are based involves observations X_i from a density $f(x_i; m_i)$ given by (1.1), $i = 1, \cdots, n$, where the m_i are known, fixed, positive integers. The case in which $m_i = m$ for $i = 1, \cdots, n$ is treated in Section 2 with the confidence intervals for μ being given by (2.11) based on the estimator $\hat{\mu}$ in (2.4).

The case in which the m_i differ is developed in Sections 3, 4, and 5 with the corresponding confidence interval for μ given by (3.18) based on $\hat{\mu}_0$ in (3.10) of σ^2 and τ are known and by (5.30) based on μ_0^* in (5.2) if σ^2 and τ are unknown.

◇ ◇ ◇ ◇ ◇

The authors would like to thank Mr. Min-Chiang Wang for doing the computer programming of the simulation results in Section 2.

REFERENCES

[1] M. Loève, *Probability Theory*, Princeton, Van Nostrand, 1960 (2nd ed.).
[2] E. S. Pearson, "Bayes theorem in the light of experimental sampling," *Biometrika*, Vol. 17 (1925), pp. 388–442.
[3] C. R. Rao, *Linear Statistical Inference and Its Applications*, New York, Wiley, 1965.