The estimation of mixing proportions by double sampling

by

George H. Brown

BU-418-M April, 1972

Abstract

This paper is concerned with the estimation of mixing proportions for a population consisting of two categories (the generalization to K categories will be considered separately).

Large sample formulae are developed for the combination of information obtained by a double sampling procedure. The first random sample consists of M items for which a bulk measurement only is available (e.g., total weight) and the second (of size m) has individual observations categorized and measured. It is shown that the combined estimate of the mixing proportion is asymptotically unbiased and a simple approximate formula for its variance is derived. Calculations show that the approximations are likely to be accurate for moderate m (greater than thirty) over a wide range of parameter values.

The problem of sequentially observing the categorized sample is dealt with briefly.

The estimation of mixing proportions by double sampling

by

George H. Brown

BU-418-M April, 1972

A. The two category problem

1.1. Defining the problem

Consider a random variable (with mean μ, variance σ^2) in a population that consists of two categories. Category 1 (with mean μ_1 and variance σ_1^2) is a proportion Π_1 of the population and category 2 (with mean μ_2 and variance σ_2^2) accounts for the remaining Π_2 ($\Pi_1+\Pi_2=1$) of the population.

It is assumed that any member of the population may be categorized without error, but that such a determination may be difficult or costly. An example of such a population is one consisting of a mixture of different ages (e.g., immature and mature). In certain natural fish populations it is possible, by examining the scales of the fish in detail, to determine their age. In order to estimate the age distribution we utilize the following two sample procedure:-

Sample 1.

A random sample of size M is drawn from the mixed population and a bulk measurement W (such as total weight) is taken and

$$\bar{W} = W/M \ .$$

Sample 2.

Another random sample of size m is drawn and each item in this sample is categorized and individually measured (giving sample means $\bar{x}_1, \bar{x}_2$ respectively for the two categories).

The problem is to estimate the mixing proportion Π_1 by utilizing information from both samples.

1.2. Preliminary considerations

We have, immediately upon drawing sample 2 an estimator of Π_1, namely

$$p_2 = t/m \quad ;$$

where t is the observed number of items from category 1 in the second sample.

By noting that

$$E[\bar{W}] = \Pi_1\mu_1 + \Pi_2\mu_2$$

we have a second estimator, i.e.,

$$p_1 = (\bar{W}-\bar{x}_2)/(\bar{x}_1-\bar{x}_2) \quad ,$$

providing $t \neq 0,m$.

If there is no within category variation, i.e., $\bar{x}_i = \mu_i$ $(i = 1,2)$ then

$$\mathrm{Var}[p_1|t\neq 0,m] = \sigma^2/M(\mu_1-\mu_2)^2$$

where

$$\sigma^2 = \Pi_1\sigma_1^2 + \Pi_2\sigma_2^2 + \Pi_1\Pi_2(\mu_1-\mu_2)^2$$

i.e., $$\mathrm{Var}[p_1|t\neq 0,m] = \Pi_1\Pi_2/M \quad ,$$

and this variance is that which obtains as if sampling binomially with a sample of size M.

At the other extreme, if σ_1^2, σ_2^2 are not zero but $\mu_1 - \mu_2 = 0$, then $\mathrm{Var}(p_1 | t \neq 0, m)$ is infinite and the bulk measurement cannot be used in the estimation of Π_1. In practice it is likely that an intermediate state between these conditions holds, so we look for some method of combining the two estimators. Consider a linear combination,

$$p = \alpha p_1 + (1-\alpha) p_2 \quad ,$$

where we choose α to satisfy some optimality criterion (e.g., to minimize $\mathrm{Var}(p)$).

If the main cost of sampling is in categorizing the observations, then this criterion corresponds (at least approximately) to the problem of minimizing the variance for a given cost.

By our earlier considerations we have

$$\frac{\Pi_1 \Pi_2}{M+m} \leq \mathrm{Var}(p) \leq \frac{\Pi_1 \Pi_2}{m} \quad .$$

Thus, if the parameters of the population are such that $\mathrm{Var}(p)$ is near the lower limit, then the double sampling scheme represents a considerable saving of cost (or effort) in categorizing individual observations (when $M >> m$).

1.3. A large sample approach

Let

t = Number of observations (from sample 2) that are found to be in category 1; $t = 0, 1, \ldots, m$.

Now consider

$$p_1 | t \neq 0, m = (\bar{W} - \bar{x}_2)/(\bar{x}_1 - \bar{x}_2)$$

$$= \frac{\mu - \mu_2 + \Delta u}{\mu_1 - \mu_2 + \Delta v}$$

where $\Delta u, \Delta v$ are random variables with

$$E(\Delta u) = E(\Delta v) = 0$$

$$E(\Delta u^2) = \frac{\sigma^2}{M} + \frac{\sigma_2^2}{m-t}$$

$$E(\Delta v^2) = \frac{\sigma_1^2}{t} + \frac{\sigma_2^2}{m-t}$$

$$E(\Delta u, \Delta v) = \frac{\sigma_2^2}{m-t} \quad .$$

Rewriting, we have

$$p_1 | t \neq 0, m = (\pi_1 + \frac{\Delta u}{\lambda})(1 + \frac{\Delta v}{\lambda})^{-1} \quad ,$$

where $\lambda = \mu_1 - \mu_2$ (taken to be non-zero). We now assume $\Delta u/\lambda$ and $\wedge v/\lambda$ to be small relative to unity (or, since they are random variables, we are assuming that

$$P\{|\Delta u| < \lambda\} \text{ and } P\{|\wedge v| < \lambda\}$$

are close to unity) and expand, to give

$$p_1|t\neq 0,m = (\pi_1 + \frac{\Delta u}{\lambda})\ (1 - \frac{\wedge v}{\lambda} + \frac{\wedge v^2}{\lambda^2} - \ldots)\ .$$

A further assumption is now needed to proceed, namely that t and m-t are relatively large, so that we may write as a reasonable approximation:-

$$E[p_1|t\neq 0,m] \doteq E[(\pi_1 + \frac{\Delta u}{\lambda})\ (1 - \frac{\Delta v}{\lambda} + \frac{\Delta v^2}{\lambda^2})]$$

$$\doteq \pi_1 + \frac{\pi_1 E(\Delta v^2)}{\lambda^2} - \frac{E(\Delta u, \wedge v)}{\lambda^2}$$

$$= \pi_1 + \frac{1}{\lambda^2}\left(\frac{\pi_1\sigma_1^2}{t} + \frac{\pi_1\sigma_2^2}{m-t} - \frac{\sigma_2^2}{m-t}\right)$$

$$= \pi_1 + \frac{\pi_1\sigma_1^2}{\lambda^2 t} - \frac{\pi_2\sigma_2^2}{\lambda^2(m-t)}\ .$$

Further,

$$E[p_1^2|t\neq 0,m] \doteq E[(\pi_1 + \frac{\Delta u}{\lambda})^2\ (1 - \frac{2\Delta v}{\lambda} + \frac{3\Delta v^2}{\lambda^2})]$$

$$\doteq \pi_1^2\ [1 + \frac{3E(\Delta v^2)}{\lambda^2}] - \frac{4\pi_1}{\lambda^2} E(\Delta u, \wedge v) + \frac{1}{\lambda^2} 3E(\Delta u^2)$$

$$= \Pi_1^2\left[1 + \frac{3}{\lambda^2}\left(\frac{\sigma_1^2}{t} + \frac{\sigma_2^2}{m-t}\right)\right] - \frac{4\Pi_1}{\lambda^2}\frac{\sigma_2^2}{m-t} + \frac{1}{\lambda^2}\left(\frac{\sigma^2}{M} + \frac{\sigma_2^2}{m-t}\right) .$$

Hence,

$$\mathrm{Var}[p_1|t\neq 0,m] = E[p_1^2|t\neq 0,m] - E^2[p_1|t\neq 0,m]$$

$$\doteq \Pi_1^2 + \frac{1}{\lambda^2}\left[\frac{\sigma^2}{M} + \frac{3\Pi_1^2\sigma_1^2}{t} + (3\Pi_1^2 - 4\Pi_1 + 1)\frac{\sigma_2^2}{m-t}\right]$$

$$- \left[\Pi_1^2 + \frac{2\Pi_1^2\sigma_1^2}{\lambda^2 t} - \frac{2\Pi_1\Pi_2\sigma_2^2}{\lambda^2(m-t)}\right]$$

$$= \frac{1}{\lambda^2}\left[\frac{\sigma^2}{M} + \frac{\Pi_1^2\sigma_1^2}{t} + \frac{\Pi_2^2\sigma_2^2}{m-t}\right]$$

1.4. Unconditioning $E(p_1|t)$ and $\mathrm{Var}(p_1|t)$

We note that the conditional expectation and variance of p_1 depends upon $1/t$ and $1/m-t$ and in order to avoid the difficulty of infinite expectations (since t may be 0 or m with non-zero probability) we restrict

$$0 < t_1 \leq t \leq t_2 < m .$$

Further, if $\sigma_1 \neq \sigma_2$ we require at least two observations within each category in order to estimate the parameters $(\mu_1, \mu_2, \sigma_1, \sigma_2)$ and, hence, it is reasonable to unconditionalize $p_1|t$ over

$$1 \leq t_1 = t_2 \leq m-1$$

or

$$2 \leq t_1 = t_2 \leq m-2 .$$

Some simplification occurs if $\sigma_1^2 = \sigma_2^2 = \sigma_P^2$ (say).

Let $t_1 = t_2 = t_0$ and consider,

$$E[p_1 | t_0 \le t \le m-t_0] \doteq \pi_1 + \frac{\sigma_P^2}{\lambda^2} E\left(\frac{\Pi_1}{t} - \frac{\Pi_2}{m-t}\right)$$

where

$$E\left(\frac{\Pi_1}{t} - \frac{\Pi_2}{m-t}\right) = \sum_{t=t_0}^{m-t_0}\left(\frac{\Pi_1}{t} - \frac{\Pi_2}{m-t}\right)p(t) \Big/ \sum_{t=t_0}^{m-t_0} p(t)$$

and

$$p(t) = \binom{m}{t}\Pi_1^t \Pi_2^{m-t} .$$

Further,

$$\mathrm{Var}(p_1 | t_0 \le t \le m-t_0) = \mathrm{Var}[E[p_1|t]] + E[\mathrm{Var}(p_1|t]]$$

$$\doteq \frac{\sigma^2}{M\lambda^2} + \frac{\sigma_P^4}{\lambda^4}\mathrm{Var}\left(\frac{\Pi_1}{t} - \frac{\Pi_2}{m-t}\right) + \frac{\sigma_P^2}{\lambda^2} E\left(\frac{\Pi_1^2}{t} + \frac{\Pi_2^2}{m-t}\right),$$

with all expectations over $t_0 \le t \le m-t_0$.

Table 1 gives some values of these expressions for m = 30(10)50, $\Pi_1 = 0.1(0.1)0.5$ and $t_0 = 1,2$.

The maximum bias in $p_1 | t_0 \le t \le m-t_0$, in the range of parameters examined, is $9\sigma_P^2/\lambda^2$ percent (when $\pi_1 = 0.1$, m = 30, $t_0 = 1$) but generally much less. If σ_P^2/λ^2 is not large we may conclude the bias is neglible unless Π_1 is small.

Table 1

Numerical quantities involved in the evaluation of the expectation and variance of p_1 (unconditional)

m	Π_1	$E\left(\frac{\Pi_1}{t} - \frac{\Pi_2}{m-t}\right)$		$Var\left(\frac{\Pi_1}{t} - \frac{\Pi_2}{m-t}\right)$		$\left\{E\left(\frac{\Pi_1^2}{t} + \frac{\Pi_2^2}{m-t}\right)\right\}^{-1}$	
		$t_0=1$	$t_0=2$	$t_0=1$	$t_0=2$	$t_0=1$	$t_0=2$
30	0.1	0.009-	0.001	0.0008	0.0002	28.96	29.46
	0.2	0.006	0.005	0.0006	0.0004	28.68	28.89
	0.3	0.003	0.003	0.0003	0.0003	28.84	28.85
	0.4	0.001	0.001	0.0002	0.0002	28.91	28.91
	0.5	0.000	0.000	0.0002	0.0002	28.92	28.92
40	0.1	0.007	0.002	0.0005	0.0002	38.71	39.24
	0.2	0.003	0.003	0.0002	0.0002	38.75	38.81
	0.3	0.001	0.001	0.0001	0.0001	38.89	38.89
	0.4	0.001	0.001	0.0001-	0.0001-	38.94	38.94
	0.5	0.000	0.000	0.0001-	0.0001-	38.95	38.95
50	0.1	0.005	0.003	0.0003	0.0001	48.60	49.03
	0.2	0.002	0.002	0.0001	0.0001-	48.82	48.83
	0.3	0.001	0.001	0.0001-	0.0000+	48.92	48.92
	0.4	0.000	0.000	0.0000+	0.0000+	48.95	48.95
	0.5	0.000	0.000	0.0000+	0.0000+	48.96	48.96

Note: The expectations above are unconditional over $t_0 \leq t \leq m-t_0$ of the binomial distribution (of t) with the tails removed.

Also, $\mathrm{Var}\left(\frac{\Pi_1}{t} - \frac{\Pi_2}{m-t}\right)$ is less than three percent of $E\left(\frac{\Pi_1^2}{t} + \frac{\Pi_2^2}{m-t}\right)$ and may be ignored unless σ_p^2/λ^2 is large. Finally we note that $\left\{E\left(\frac{\Pi_1^2}{t} + \frac{\Pi_2^2}{m-t}\right)\right\}^{-1}$ is approximated closely by $1/m-1$. Hence, we have two good approximations over a wide range of Π_1;-

$$E[p_1 | t_0 \leq t \leq m-t_0] \doteq \Pi_1$$

and

$$\mathrm{Var}[p_1 | t_0 \leq t \leq m-t_0] \doteq \frac{\sigma^2}{M\lambda^2} + \frac{\sigma_p^2}{\lambda^2(m-2)},$$

the m-2 appearing in the denominator being close to the "rounded down" value of $\left\{E\left(\frac{\Pi_1^2}{t} + \frac{\Pi_2^2}{m-t}\right)\right\}^{-1}$ which will partially compensate for small terms that have been ignored.

It is possible that these approximations are still "good" even when $\sigma_1 \neq \sigma_2$, unless these parameters are widely different. Assuming this not to be so, we proceed.

1.5. Combining the estimators of Π_1

We have

$$\hat{\Pi}_1 = \alpha p_2 + (1-\alpha)p_1$$

where

$$p_2 = \frac{t}{m}$$

and

$$p_1 = \frac{\bar{W}-\bar{x}_2}{\bar{x}_1-\bar{x}_2},$$

both being conditional on $t_0 \le t \le m-t_0$. Under the assumptions made in the foregoing derivation, we also have, ignoring the tails of the distribution,

$$E[p_2] \doteq \Pi_1$$

$$\mathrm{Var}[p_2] \doteq \frac{\Pi_1\Pi_2}{m} \quad .$$

Hence,

$$E[\hat{\Pi}_1] = \Pi_1$$

and

$$\mathrm{Var}[\hat{\Pi}_1] \doteq \alpha^2 \mathrm{Var}(p_2) + (1-\alpha)^2 \mathrm{Var}(p_1) + 2\alpha(1-\alpha)\,\mathrm{cov}(p_1,p_2).$$

However,

$$E[p_1|t] \doteq \Pi_1 \qquad \text{(Under the assumptions made)}$$

so

$$\mathrm{cov}(p_2,p_1) = \mathrm{cov}(p_2,\ E(p_1|t))$$

$$\doteq 0 \quad .$$

We find $\mathrm{Var}(\hat{\Pi}_1)$ is minimized when

$$\alpha = \frac{\mathrm{Var}(p_1)}{\mathrm{Var}(p_1)+\mathrm{Var}(p_2)}$$

in which case

$$\mathrm{Var}(\hat{\Pi}_1) \doteq \frac{\mathrm{Var}(p_1)\mathrm{Var}(p_2)}{\mathrm{Var}(p_1)+\mathrm{Var}(p_2)}$$

where

$$\mathrm{Var}(p_1) \doteq \frac{\Pi_1\Pi_2}{M} + \frac{\sigma_p^2}{\lambda^2}\left(\frac{1}{M} + \frac{1}{m-2}\right)$$

$$\mathrm{Var}(p_2) \doteq \frac{\Pi_1\Pi_2}{m}$$

and σ_p^2, λ^2 may be estimated from the second (categorized) sample - σ_p^2 being a pooled estimate of the within category variation.

We note that $\mathrm{Var}(\hat{\Pi}_1)$ depends upon Π_1, so it is necessary to obtain a solution iteratively, i.e., start with $\Pi_1 = p_2$ (say), obtain α and re-estimate Π_1. This procedure is continued until the estimate of Π_1 is stabilized and then $\mathrm{Var}(\hat{\Pi}_1)$ may be computed.

1.6. Concluding Comments

Clearly there are difficulties in applying this method if the conditions under which the approximations have been made are relaxed. When Π_1 is close to one or zero, then there is appreciable probability that t = 0 or m respectively.

When λ (equal to $\mu_1-\mu_2$) is small there is some chance that p_1 will not be in the parameter range of zero to one.

If Π_1 is known to be close to zero or one, then it may be better (in order to avoid the possibility of one category not being represented) to consider an "inverse" or sequential sampling scheme in order to obtain, say,

at least t of the less frequent category. When λ is small (relative to σ_p^2) there is little to be gained from the double sampling procedure.

Having obtained the weights in order to combine the estimators of Π_1 and the large sample variance of the combined estimate it is desirable to examine its large sample distribution in order to obtain interval estimates. This may be achieved with a simulation study for various parent distributions.

2.1. Sequential sampling

In the concluding remarks made in 1.6 it was noted that for moderate size samples there is an appreciable probability that there will be no representatives of the less frequent category (especially when Π_1 is close to zero or one). Since the estimator utilizing the bulk sample depends upon being able to estimate μ_1 and μ_2 it may be desirable to ensure that at least $t_i (i=1,2)$ representatives from each category are obtained. This may be achieved by sequential or "inverse" sampling and two cases are considered.

Case (a)

If it is known that one category, say category 1, is much less frequent than the other then we sample until t items from category 1 are obtained.

Case (b)

Here we continue sampling until at least $t_i (i=1,2)$ items are obtained from each category (but not more of both).

2.2. Case (a) - Sequential sampling

The sample size m is now a random variable, with the negative binomial distribution

$$P[m=t+r] = \binom{t+r-1}{t-1} \Pi_1^t \Pi_2^r \; ; \quad r = 0,1,2,\ldots$$

It may be shown that

$$p_2 = \frac{t-1}{m-1}$$

is the M.V.U.E. of Π_1 .

Also, by noting that

$$E\left[\frac{(t-1)(t-2)}{(m-1)(m-2)}\right] = \Pi_1^2$$

we have

$$E\left(\frac{t-1}{m-1}\right)^2 - \Pi_1^2 = E\left[\left(\frac{t-1}{m-1}\right)^2 - \frac{(t-1)(t-2)}{(m-1)(m-2)}\right]$$

and, hence, an unbiased estimator of $v(p_2)$ is

$$\widehat{Var(p_2)} = \frac{(t-1)^2(m-2) - (t-1)(t-2)(m-1)}{(m-1)^2(m-2)}$$

$$= \frac{(t-1)(m-t)}{(m-1)^2(m-2)}$$

$$= \frac{p_2(1-p_2)}{m-2} .$$

2.3. Double sampling estimate of Π_1 for Case (a)

Recalling the conditional estimator of Π_1, based on the bulk sampling (except now we condition on r, rather than t) we have:

$$p_1|r = \frac{\bar{W}-\bar{x}_2}{\bar{x}_1-\bar{x}_2} \quad ,$$

which is defined if $r \geq 1$. Since we are assuming that Π_1 is small, the event $r = 0$, which occurs with probability Π_1^t will be negligibly small, even for moderate values of t .

We have then

$$E[p_1|r\geq 1] \doteq \Pi_1 + \frac{\Pi_1\sigma_1^2}{\lambda^2 t} - \frac{\Pi_2\sigma_2^2}{\lambda^2} E\left(\frac{1}{r}\right)$$

and

$$\mathrm{Var}[p_1|r\geq 1] = \mathrm{Var}[E[p_1|r]] + E[\mathrm{Var}(p_1|r)]$$

$$\doteq \frac{1}{\lambda^2}\left[\frac{\sigma^2}{M} + \frac{\Pi_1^2\sigma_1^2}{t}\right] + \frac{\Pi_2^4\sigma_2^4}{\lambda^4}\mathrm{Var}\left(\frac{1}{r}\right) + \frac{\Pi_2^2\sigma_2^2}{\lambda^2}E\left(\frac{1}{r}\right) .$$

Now

$$P[r|r\geq 1] = \frac{\Pi_1^t}{1-\Pi_1^t}\binom{t+r-1}{r}\Pi_2^r \; ; \quad r \geq 1$$

so

$$E\left[\frac{1}{r}|r\geq 1\right] = \frac{\Pi_1^t}{1-\Pi_1^t}\sum_{r=1}^{\infty}\binom{t+r-1}{r}\frac{\Pi_2^r}{r}$$

$$= \frac{\Pi_1^t}{1-\Pi_1^t}\int_0^{\Pi_2}\sum_{r=1}^{\infty}\binom{t+r-1}{r}x^{r-1}dx$$

$$= \frac{\Pi_1^t}{1-\Pi_1^t}\int_0^{\Pi_2}\frac{1-(1-x)^t}{x(1-x)^t}\,dx$$

$$= \frac{\Pi_1^t}{1-\Pi_1^t} \int_{\Pi_1}^{1} \frac{1-y^t}{(1-y)y^t}\, dy$$

$$= \frac{\Pi_1^t}{1-\Pi_1^t} \int_{\Pi_1}^{1} \sum_{j=0}^{t-1} y^{j-t}\, dy$$

$$= \frac{\Pi_1^t}{1-\Pi_1^t} \left[\sum_{j=0}^{t-2} \frac{y^{j-t+1}}{j-t+1} + \ln y \right] \Bigg|_{\Pi_1}^{1}$$

$$= \frac{\Pi_1^t}{1-\Pi_1^t} \left[\sum_{j=0}^{t-2} \frac{\Pi_1^{j-t+1}-1}{t-j+1} - \ln \Pi_1 \right]$$

2.4. Numerical example

We take $t=3$, $\Pi_1=0.1$ (and later, for comparison purposes we suppose $\sigma_1=\sigma_2=\sigma_P$) and get

$$E\left(\frac{1}{r}\middle| r\geq 1\right) = \frac{\Pi_1^3}{1-\Pi_1^3} \left[\sum_{1-\Pi_1^3}^{1} \frac{\Pi_1^{\alpha-2}-1}{2-j} - \ln \Pi_1 \right]$$

$$\doteq \frac{0.001}{0.999} \left[58.5 + 2.30 \right]$$

$$\doteq 0.061 .$$

Now

$$\frac{\Pi_1}{t} - \Pi_2 E\left(\frac{1}{r}\middle| r\geq 1\right) \doteq 0.033 - 0.055$$

$$= - 0.022$$

i.e., this bias factor is more than twice that of the fixed sample size with m=30 (and 3 representatives of category 1 expected).

Similarly

$$\left\{\Pi_2^2 E\left(\frac{1}{r}\middle| r\geq 1\right) + \frac{\Pi_1^2}{t}\right\}^{-1} \doteq \frac{1}{0.0494 + 0.0033}$$

$$\doteq 19$$

compared with 29 in the fixed sample case.

This example illustrates that the inverse sampling procedure may give rise to estimators with a larger mean square error than the comparable fixed sample method (which has the same number of expected representatives from category 1). When m=30, Π_1=0.1 the fixed sample example has a probability of about 0.042 that it will contain no representatives from category 1. Thus, it is likely that the fixed sampling procedure is generally superior (by considering further numerical examples) unless σ_p^2/λ^2 is very small, when it becomes imperative that the bulk sample be utilized. If this is so then the superiority of the inverse sampling method depends upon Π_1 - the smaller Π_1, the more there is to be gained by inverse sampling. However, it must be remembered that if σ_p^2/λ^2 is small enough the members of the population may be categorized by the measurement used (with high accuracy) and the assumption that the cost of categorizing is the main cost of sampling may no longer hold. Also, we recall, that the conditional expectation and variance of $p_1|r$ depends upon t and r being fairly large, a condition that is violated when Π_1 is small unless sample sizes with large expectations are allowed.

2.5. Case (b) - Sequential sampling

We now specify that we keep sampling until we obtain at least $t_i (i=1,2)$ representatives from each category (but not more of both).

Then we have,

$$P[m=t_1+t_2] = \binom{t_1+t_2}{t_1} \pi_1^{t_1} \pi_2^{t_2}$$

and now consider $P[m=t_1+t_2+r]$ for $r \geq 1$. After the $(t_1+t_2)^{th}$ trial, let there be t_1+s items from category 1 and t_2-s from category 2. Then $m=t_1+t_2+r$ if, in the next r-1 trials there are s-1 items from category 2 and the r^{th} trial results in an item from category 2 $(s=1,2,\ldots,\min(r,t_2))$. Similarly, if after the $(t_1+t_2)^{th}$ trial there are t_1-u items from category 1 and t_2+u from category 2 then, $m=t_1+t_2+r$ if, in the next r-1 trials there are u-1 items from category 1 and the r^{th} is from category 1 $(u=1,2,\ldots,\min(r,t_1))$.

i.e., for $r \geq 1$

$$\begin{aligned} P[m=t_1+t_2+r] &= P[t_1+r \text{ from category 1 and } t_2 \text{ from category 2}] \\ &\quad + P[t_1 \text{ from category 1 and } t_2+r \text{ from category 2}] \end{aligned}$$

$$= \sum_{s=1}^{\min(r,t_2)} \binom{t_1+t_2}{t_1+s} \pi_1^{t_1+s} \pi_2^{t_2-s} \binom{r-1}{s-1} \pi_1^{r-s} \pi_2^{s-1} \pi_2$$

$$+ \sum_{s=1}^{\min(r,t_1)} \binom{t_1+t_2}{t_2+s} \pi_1^{t_1-s} \pi_2^{t_2+s} \binom{r-1}{s-1} \pi_1^{s-1} \pi_2^{r-s} \pi_1$$

$$= \sum_{s=1}^{t_2} \binom{t_1+t_2}{t_1+s} \binom{r-1}{s-1} \pi_1^{t_1+r} \pi_2^{t_2}$$

$$+ \sum_{s=1}^{t_1} \binom{t_1+t_2}{t_2+s} \binom{r-1}{s-1} \pi_1^{t_1} \pi_2^{t_2+r}$$

$$= \binom{t_1+t_2+r-1}{t_2-1} \Pi_1^{t_1+r} \Pi_2^{t_2} + \binom{t_1+t_2+r-1}{t_1-1} \Pi_1^{t_1} \Pi_2^{t_2+r} .$$

In the event $t_1=t_2=t$ (say) we have

$$P[m=2t+r] = \binom{2t+r-1}{t-1} \Pi_1^t \Pi_2^t \left(\Pi_1^r+\Pi_2^r\right) ; \quad r \geq 0$$

We note that $P[m=2t+r]$ is a function only of $\Pi_1\Pi_2$, since $\Pi_1^r+\Pi_2^r$ may be expressed as a polynomial in $\Pi_1\Pi_2$. Hence, it is not possible to construct a simple to construct a simple estimator of Π_1 from the second sample. Thus, this type of sampling is only feasible when we are prepared to dispense with a combined estimator and we use only the estimator arising from the bulk measurement. As we have already shown, this corresponds to a situation where σ_p^2/λ^2 is small and then we have the same difficulties noted with Case (a) sequential sampling (in the closing sentences of 2.4). For these reasons this type of sampling will not be considered any further.

5/1 aw