INCORPORATING PARTIALLY CLASSIFIED SAMPLE SEGMENTS INTO NASA ACREAGE ESTIMATION PROCEDURES

Robert L. Sielken, Jr., Texas A&M University

ABSTRACT

For the purposes of crop acreage estimation a segment is said to be completely classified if the acreage of each crop of interest can be directly estimated. A segment is said to be partially classified if some individual crops within a family of crops cannot be differentiated and their acreages individually estimated. For example, if it is estimated that 30% of a segment is summer crops but no further differentiation into say corn, soybeans, and other summer crops is made, then the segment is only partially classified. Methods of estimating individual crop acreages using a mixture of completely classified and partially classified segments are discussed.

1. Introduction

A small sample of segments within a large region is selected. Each sample segment is observed via satellite at several different times during the crop growing seasons. The objective is to estimate for each crop of interest the proportion of the region's acreage corresponding to that crop's harvested acreage.

In this paper we shall assume that there are only two crops of interest. Furthermore, we shall assume that only data from the current growing year are to be used in estimating the crop at harvest proportions. The cases where more than two crops are of interest and/or data is available from more than one growing year will be considered in subsequent papers.

The sample segments are all assumed to be of the same size. No assumption is made about the region size or the segment size. The sampled segments are assumed to represent a random sample (without replacement) from the segments in the region.

Each sample segment is assumed to have been observed at least once during the growing year and possibly several times. The two crops of interest will be designated as crop A and crop B. When a sample segment is observed, the observation can have the form (p_A, p_B, p_{other}) where

p_A = the estimated proportion of the segment which will be harvested in crop A,

p_B = the estimated proportion of the segment which will be harvested in crop B, and

$p_{other} = 1 - p_A - p_B$ = the estimated proportion of the segment which will not be harvested in either A or B.

Alternatively, estimates may not be made on A and B separately but only on A and B collectively, so that the observation can have the form (p_{A+B}, p_{other}) where

p_{A+B} = the estimated proportion of the segment that will be harvested in either A or B, and

$p_{other} = 1 - p_{A+B}$ = the estimated proportion of the segment that will not be harvested in A or B.

The most recent segment estimates are assumed to reflect any previous observations made on that sample segment during the current growing year. If a sample segment's current observation is of the form (p_{A+B}, p_{other}), then the sample segment will be called partially classified. If its observation is of the form (p_A, p_B, p_{other}), then it will be called completely classified.

The proportion of the region harvested in crop A will be denoted P_A with P_B similarly defined. The objective is to estimate P_A and P_B using the observations on the sample segments. This estimation may have to be made at more than one time during the growing year. Of course, if there are no completely classified sample segments, only the sum $P_A + P_B$ can be estimated on the basis of the sample segments.

2. Estimation Procedures

There are undoubedly several decision makers within a region who collectively determine the proportional acreages (P_A, P_B, P_{other}) for the region. The exact decision making process is unknown. Hence, certain assumptions will be made about the outcome of this decision making process, and these assumptions will suggest a few reasonable estimation procedures for consideration. The assumptions are at best approximations and are no doubt technically not satisfied. The only role of these assumptions is to suggest estimation procedures. The worth of these procedures will not be determined relative to the underlying assumptions which suggested them, but rather determined by their behavior on real sample data.

The number of acres harvested in a segment, say $(X_A, X_B, N - X_A - X_B)$, will be assumed to follow a multinomial distribution with

$$P(X_A = x_A, X_B = x_B) = \frac{N!}{x_A!\, x_B!\, (N - x_A - x_B)!} P_A^{x_A} P_B^{x_B} (1 - P_A - P_B)^{N-x_A-x_B} . \quad (1)$$

This assumption would be technically correct if every decision maker acted independently and for each unit of acreage chose crop A, crop B, or other with probabilities P_A, P_B, and P_{other} respectively. The satellite based (CAMS) estimated segment proportions (p_A, p_B, p_{other}) are assumed to correspond to X_A/N, X_B/N, $(N-X_A - X_B)/N$ respectively.

The following additional notation will be used throughout this paper. The total number of segments sampled will be denoted by n. The sample segments which have been completely classified will be denoted by $(p_{Ai}, p_{Bi}, 1-p_{Ai}-p_{Bi})$, $i=1,2,\ldots,I$.

Similarly, the sample segments which have only been partially classified will be denoted by $[(p_A + p_B)_j,\ 1 - (p_A - p_B)_j]$, $j=1,2,\ldots,J$. Hence $n = I + J$. Since no separate estimate of P_A and P_B could be made solely on the basis of the sampled segments unless some sampled segments had been completely classified, it will be assumed that $I > 0$.

2.1 Maximum Likelihood Estimates

Under the assumed multinomial distribution (1) the joint likelihood for the CAMS estimates is

$$L = \left\{\prod_{i=1}^{I} \frac{N!}{X_{Ai}!X_{Bi}!(N-X_{Ai}-X_{Bi})!} P_A^{X_{Ai}} P_B^{X_{Ai}} (1 - P_A - P_B)^{N-X_{Ai}-X_{Bi}}\right\}$$

$$\left\{\prod_{j=1}^{J} \frac{N!}{(X_A + X_B)_j!\ [N-(X_A + X_B)_j]!} (P_A + P_B)^{(X_A+X_B)_j} [1-(P_A+P_B)]^{N-(X_A+X_B)_j}\right\} \tag{2}$$

or equivalently,

$$L = \left\{\prod_{i=1}^{I} \frac{N!}{(Np_{Ai})!\ (Np_{Bi})!\ (N-Np_{Ai}-Np_{Bi})!} P_A^{Np_{Ai}} P_B^{Np_{Bi}} (1-P_A-P_B)^{N-Np_{Ai}-Np_{Bi}}\right\} .$$

$$\left\{\prod_{j=1}^{J} \frac{N!}{[N(p_A+p_B)_j]!\ (N[1-(p_A+p_B)_j])!} (P_A+P_B)^{N(p_A+p_B)_j} (1-P_A-P_B)^{N-N(p_A+p_B)_j}\right\} . \tag{3}$$

The corresponding maximum likelihood estimates are

$$\hat{P}_A = \frac{\sum_{i=1}^{I} p_{Ai}}{\sum_{i=1}^{I} p_{Ai} + \sum_{i=1}^{I} p_{Bi}} \left\{\frac{\sum_{i=1}^{I} p_{Ai} + \sum_{i=1}^{I} p_{Bi} + \sum_{j=1}^{J} (p_A+p_B)_j}{I + J}\right\}, \tag{4}$$

$$\hat{P}_B = \frac{\sum_{i=1}^{I} p_{Bi}}{\sum_{i=1}^{I} p_{Ai} + \sum_{i=1}^{I} p_{Bi}} \left\{\frac{\sum_{i=1}^{I} p_{Ai} + \sum_{i=1}^{I} p_{Bi} + \sum_{j=1}^{I} (p+p)_j}{I + J}\right\}, \tag{5}$$

and

$$\widehat{(1 - P_A - P_B)} = 1 - \hat{P}_A - \hat{P}_B \quad . \tag{6}$$

The form of these estimators is really quite intuitive. For example,

$$\hat{P}_A = \left\{\begin{matrix}\text{Estimated Proportion}\\ \text{of Crops A and B}\\ \text{that is Crop A}\end{matrix}\right\} \cdot \left\{\begin{matrix}\text{Estimated Proportion}\\ \text{of the Stratum that}\\ \text{is Crop A or B}\end{matrix}\right\} . \tag{7}$$

Hocking and Oxspring (1971) have considered maximum likelihood estimators for incomplete multinomial data such as the above and for even more general situations.

2.2 Least Squares Estimates

If $(\hat{P}_A, \hat{P}_B, 1 - \hat{P}_A - \hat{P}_B)$ is chosen to minimize

$$\sum_{i=1}^{I} (p_{Ai}-P_A)^2 + \sum_{i=1}^{I} (p_{Bi}-P_B)^2 + \sum_{i=1}^{I} [(1-p_{Ai}-p_{Bi}) - (1-P_A-P_B)]^2 + \sum_{j=1}^{J} [(p_A+p_B)_j - (P_A+P_B)]^2 +$$

$$\sum_{j=1}^{J}\left\{[1-(p_A+p_B)_j] - [1 - (P_A+P_B)]\right\}^2 , \tag{8}$$

with respect to P_A and P_B, then $(\hat{P}_A, \hat{P}_B, 1-\hat{P}_A-\hat{P}_B)$, would be the least squares estimate of $(P_A, P_B, 1 - P_A - P_B)$. The resulting least squares estimates are

$$\hat{P}_A = \left\{2\sum_{i=1}^{I} p_{Ai} + \sum_{i=1}^{I} p_{Bi} + 2\sum_{j=1}^{J} (p_A+p_B)_j - \hat{P}_B(I+2J)\right\} / 2(I+J) \tag{9}$$

and

$$\hat{P}_B = \left\{(3I + 2J) \sum_{i=1}^{I} p_{Bi} - 2J \sum_{i=1}^{I} p_{Ai} + 2I \sum_{j=1}^{J}(p_A+p_B)_j\right\} / I(3I+4J) . \tag{10}$$

2.3 Weighted Least Squares Estimates

Since one may wish to give more weight to more precise (less variable) estimates than to less precise (more variable) estimates, a weighted least squares estimate of $(P_A, P_B, 1 - P_A - P_B)$ may be preferable. The weighted least squares estimate minimizes

$$\sum_{i=1}^{I} W_{Ai}\,(p_{Ai} - P_A)^2 + \sum_{i=1}^{I} W_{Bi}(p_{Bi} - P_B)^2 +$$

$$\sum_{i=1}^{I} W_{A+B,i}\,[(1 - p_{Ai} - p_{Bi}) - (1 - P_A - P_B)]^2 +$$

$$\sum_{j=1}^{I} W_{A+B,j}\,[(p_A+p_B)_j - (P_A+P_B)]^2 + \sum_{j=1}^{J} W_{A+B,j}\Big\{[1-(p_A+p_B)_j] - [1-(P_A+P_B)]\Big\}^2$$

with respect to P_A and P_B. The corresponding weighted least squares estimates would be

$$\hat{P}_A = \Big\{\sum_{i=1}^{I} W_{Ai}p + \sum_{i=1}^{I} W_{A+B,i}(p_{Ai}+p_{Bi}) + 2\sum_{j=1}^{J} W_{A+B,j}\cdot (p_A+p_B)_j - \hat{P}_B\Big(\sum_{i=1}^{I} W_{A+B,i} + 2\sum_{j=1}^{J} W_{A+B,j}\Big)\Big\} \Big/ \Big\{\sum_{i=1}^{I} W_{Ai} + \sum_{i=1}^{I} W_{A+B,i} + 2\sum_{j=1}^{J} W_{A+B,j}\Big\}$$

and

$$\hat{P}_B = \Big\{[\sum_{i=1}^{I} W_{Bi}p_{Bi} + \sum_{i=1}^{I} W_{A+B,i}(p_{Ai}+p_{Bi}) + 2\sum_{j=1}^{J} W_{A+B,j}(p_A+p_B)_j]\;\cdot$$

$$[\sum_{i=1}^{I} W_{Ai} + \sum_{i=1}^{I} W_{A+B,i} + 2\sum_{j=1}^{J} W_{A+B,j}] -$$

$$[\sum_{i=1}^{I} W_{Ai}p_{Ai} + \sum_{i=1}^{I} W_{A+B,i}(p_{Ai}+p_{Bi}) + 2\sum_{j=1}^{J} W_{A+B,j}(p_A+p_B)_j]\;\cdot$$

$$[\sum_{i=1}^{I} W_{A+B,i} + 2\sum_{j=1}^{J} W_{A+B,j}]\Big\} \Big/ \Big\{[\sum_{i=1}^{I} W_{Ai} + \sum_{i=1}^{I} W_{A+B,i} + 2\sum_{j=1}^{J} W_{A+B,j}]\;\cdot$$

$$[\sum_{i=1}^{I} W_{Bi} + \sum_{i=1}^{I} W_{A+B,i} + 2\sum_{j=1}^{J} W_{A+B,j}] -$$

$$[\sum_{i=1}^{I} W_{A+B,i} + 2\sum_{j=1}^{J} W_{A+B,j}]^2\Big\}\;. \qquad (12)$$

If the weights, W, are all taken to be one, then the weighted least squares estimates (12) are equal to the least squares estimates (10). A reasonable alternative to equal weights is to weight each squared difference inversely in proportion to the approximate variance of the CAMS estimate it contains. Specifically,

$$\begin{aligned}
W_{Ai} &= 1/\hat{P}_A(1-\hat{P}_A)\,, & i&=1,2,\ldots,I\,,\\
W_{Bi} &= 1/\hat{P}_B(1-\hat{P}_B)\,, & i&=1,2,\ldots,I\,,\\
W_{A+B,i} &= 1/(1-\hat{P}_A\hat{P}_B)\cdot(\hat{P}_A + \hat{P}_B)\,, & i&=1,2,\ldots I,\\
W_{A+B,j} &= 1/(\hat{P}_A + \hat{P}_B)\cdot(1-\hat{P}_A - \hat{P}_B)\,, & j&=1,2,\ldots J.
\end{aligned} \qquad (13)$$

These weights put more weight on the differences where the segment estimates are more precise (less) variable).

The weights (13) use estimated variances corresponding to the multinomial distribution assumption. The actual calculation of the weighted least squares estimates would be an iterative procedure. Initially, the weights would be set equal to one, then $\hat{P}_A$ and $\hat{P}_B$ determined from (12). These initial estimates $\hat{P}_A$ and $\hat{P}_B$ would be substituted into (13) to determine the weights for the second iteration. The second evaluation of (12) yields the second set of estimates $\hat{P}_A$ and $\hat{P}_B$. After a moderate number of iterations, like 4, the $\hat{P}_A$ and $\hat{P}_B$ should change very little and will be taken as the weighted least squares estimates.

2.4. The Combination of a Least Squares Ratio Estimate and a Maximum Likelihood Proportion Estimate

An alternative approach is to use only the completely classified sample segments to estimate $R_A = P_A/(P_A + P_B)$ where R_A is the relative proportion of crop A among the combined acreage for crops A and B. Similarly, $R_B = P_B/(P_A + P_B) = 1-R$. Then, having estimated R_A and hence R_B, the region proportions P_A and P_B could be estimated by

$$\hat{P}_B = \hat{R}_B \cdot \hat{P}_{A+B} \quad \text{and} \quad \hat{P}_B = \hat{R}_B \cdot \hat{P}_{A+B} \qquad (14)$$

where $\hat{P}_{A+B}$ is an estimate for the combined acreage proportion $P_{A+B} = P_A + P_B$.

The combined acreage proportion P_{A+B} could be estimated by the maximum likelihood estimate

$$\hat{P}_{A+B} = \Big\{\sum_{i=1}^{I} p_{Ai} + \sum_{i=1}^{I} p_{Bi} + \sum_{j=1}^{J} (p_A+p_B)_j\Big\}/(I+J)\,. \quad (16)$$

Other estimates of P_{A+B} could be used.

The relative proportion R_A can be estimated by minimizing

$$\sum_{i=1}^{I} [p_{Ai} - R_A(p_{Ai}+p_{Bi})]^2 + \sum_{i=1}^{I} [p_{Bi} - (1-R_A)(p_{Ai}+p_{Bi})]^2 \qquad (17)$$

with respect to R_A. The resulting least squares estimator of R_A is

$$\hat{R}_A = \sum_{i=1}^{I} p_{Ai}(p_{Ai}+p_{Bi}) / \sum_{i=1}^{I} (p_{Ai}+ p_{Bi})^2 \quad . \tag{18}$$

3. Empirical Behavior of Some Estimation Procedures

Someone once said, "The proof of the pudding is in the eating." Similarly, here the proof of an estimation procedure's value lies in its actual performance on real data. To learn about the empirical behavior of the four alternative estimation procedures

(1) maximum likelihood estimators, (4)-(6),
(2) least squares estimators, (10),
(3) weighted least squares estimators, (12), and
(4) the combination of a least squares ratio estimator, (18), and the maximum likelihood proportion estimator, (16),

a Monte Carlo study was performed.

3.1 The Structure of the Monte Carlo Study

The simulation of the behavior of the estimation procedures was based around two real sets of CAMS estimates. The 1977 North Dakota CAMS estimates of crop A = spring wheat and crop B = spring small grains other than spring wheat formed one set. The second set was the 1978 North Dakota CAMS estimates of crop A = barley and crop B = spring small grains other than barley. There were 83 segments comprising the first set and 76 segments in the second set. All simulations were done on each data set separately.

To obtain a single observation on $(\hat{P}_A, \hat{P}_B, 1-\hat{P}_A-\hat{P}_B)$ for each estimator the following procedure was performed:

i) Simulate a region by sampling without replacement segments from the data set.

ii) Select a sample of segments from the simulated region by sampling without replacement segments from the simulated region.

iii) Calculate all estimators.

This procedure was repeated 1000 times for each data set, stratum size, segment sample size, and fraction of the segment sample that would be only partially classified.

Regions of size 40 segments and 20 segments were considered. The segment sample sizes were 5, 10, and 15 when the region size was 40 segments and were 5 and 10 when the region size was 20 segments. For each region size and segment sample size, 20%, 40%, 60%, and 80% of the segments were partially classified.

3.2 Evaluation Criteria

There are several possible ways to measure the sample behavior of the estimators. For each estimator and each of $\hat{P}_A$, $\hat{P}_B$, and $1-\hat{P}_A-\hat{P}_B$ the following measures were calculated for each data set:

(i) average absolute error = the average over 1000 simulations of $|\hat{P} - P_{region}|$ where P_{region} represents the actual crop proportion in the particular simulated region,

(ii) average squared error = the average over 1000 simulations of $(\hat{P} - P_{region})^2$

(iii) bias of average estimate = the difference between the average $\hat{P}$ in 1000 simulations and P_{set} where P_{set} is the actual crop proportion in the entire set of segments, and

(iv) sample variance of the estimator.

Some information is, of course, lost when some segments are only partially classified. To assess this loss, the maximum likelihood estimators were also calculated using the complete classification for all sampled segments. Since these estimations have complete information for the entire sample of n segments instead of complete information on only some of the n segments and partial information on the remainder, these latter estimators should perform better.

To show that the inclusion of the partially classified segments into the estimation procedure is better than simply ignoring them, the maximum likelihood estimators, least squares estimators, and weighted least squares estimators were also calculated using only the subset of the n sample segments corresponding to the completely classified segments.

3.3 Indications of the Monte Carlo Study

A summary of the observed average absolute errors for the better estimators is given in Table 1 for the 1977 data set when the simulated region size was 20 segments. The average absolute errors when the simulated sample size was 40 are similar to those in Table 1 and have been omitted to save space. The results for the 1978 data set have not been tabled for the same reason. The average over the entire data sets of $(p_A, p_B, 1-p_A-p_B)$ is (.235, .073, .692).

The unweighted least squares estimators given in (10) are substantially inferior to the other estimators, and hence their behavior is not summarized in Table 1. There are much smaller differences among the reamining estimators.

When both the partially and completely classified sample segment estimates are incorporated into the maximum likelihood estimator (MLE) given in (4)-(6), the weighted least squares estimator (WLS) given in (12), and the combination of the least squares ratio estimator (LSR) given in (18) and the maximum likelihood combined acreage proprotion estimator given in (16), all three estimators have similar behavior. Overall, the MLE is slightly better than the WLS, and the WLS is slightly better than the LSR. The differences are only slight since the average absolute error for LSR is never more than 110% of that for MLE. The average squared errors tell the same story. There are no significant differences among the biases or sample variances for the MLE, WLS, and LSR.

The effects of the sample size and amount of partial classification are essentially the same for

MLE, WLS, and LSR. When the underlying data set proportional acreage is nearly .25 as it is for P_A in 1977, then the average absolute error decreases as the segment sample size, n, increases roughly as follows

.04 for n = 5

.025 for n = 10, and

.02 for n = 15.

When the underlying data set proportional acreage is nearly .05 as it is for P_B in 1977, then the average absolute decreases with increasing segment sample size roughly as follows

.02 for n = 5,

.015 for n = 10, and

.01 for n = 15.

There is roughly a quadratic increase in average absolute error as the amount of partial classification increases from 20% to 40% to 60% to 80%. The average absolute error increases around twice as much when the amount of partial classification increases from 60% to 80% as it does when the amount of partial classification increases from 40% to 60%.

All of the estimators seem to slightly (.005) overestimate the combined proportional acreage for A and B when the segment sample size, n, is 5. This bias drops to .0025 when n = 10 and essentially disappears at n = 15.

The WLS does not seem to be as good as either MLE or LSR when there is only one completely identified sample segment.

Table 2 indicates the slight increase in average absolute error caused by only partially classifying a segment's acreages instead of completely classifying them. Only the behavior of MLE is summarized since WLS and LSR should behave similarly. Naturally the increase in the average absolute error increases as the amount of partial classification increases. For less than 80% partial classification there is always an increase of less than .01. For 80% the largest increase is less than .017. At 80% partial classification the average absolute error can double. Of course, if there is 100% partial classification, then there are no individual crop estimates.

Table 3 indicates the average absolute errors both with and without the partially classified sample segments being included in the estimators MLE and WLS. This table is presented to exemplify the value of at least partially classifying a sample segment instead of simply disregarding it if it can't be completely classified. The percentage increase in the average absolute errors when the underlying proportional acreage is nearly .25 and the partially classified segments are ignored is usually at least

5% - 10% for 20% partial classification
10% - 25% for 40% partial classification
25% - 50% for 60% partial classification
50% - 100% for 80% partial classification .

3.5 A Follow-up Investigation

In the Monte Carlo Study the CAMS estimates of the segment's crop acreage proportions were simulated as if they contained no errors. In order to ascertain the impact of any such errors, the Monte Carlo study was repeated with a normal deviate added to each of the segment's crop acreage proportion estimates. The normal deviates had mean zero. Their variances were $P_A(1-P_A)/M$ and $P_B(1-P_B)/M$ for crops A and B respectively. The segment's acreage proportion for "other" was taken to be one minus the proportion estimates for A and B. Tables 4 and 5 indicate the resulting average absolute errors for MLE, WLS, and LSR for M=100 and M=20 respectively.

When M=100, the average absolute errors increased over what they were in the original Monte Carlo study which had no errors or equivalently $M=\infty$. However, the relative behaviors of MLE, WLS, and LSR did not change. On-the-other-hand, when M=20, not only did the average absolute errors increase, but also the relative behaviors changed. For M=20, MLE is still generally preferable to WLS, but here LSR is slightly preferable to MLE.

4. Conclusions

On the basis of the limited Monte Carlo study and the small follow-up investigation the following conclusions were reached:

1) As long as there are some completely classified sample segments, it is reasonable to estimate the individual crop proportions in the region.
2) It is prudent to avoid having a large percentage (say 80%) of only partially classified sample segments.
3) It is much better to incorporate the partially classified sample segments into the estimators than it is to disregard the partially classified sample segments.
4) When there are either no errors or only very small errors in the estimates of the segment's crop acreage proportions, the maximum likelihood estimators seem to be the best estimators, but they are not greatly superior to weighted least squares estimators or the use of a least squares ratio estimator.
5) When there are fairly substantial errors in the estimates of the segment's crop acreage proportions, the combination of the least squares ratio estimator with the maximum likelihood estimator of the combined crop proportion is the superior estimator.

REFERENCE

Hocking, R.H. and Oxspring, H.H. (1971). Maximum likelihood estimation with incomplete multinomial data. JASA 66, 65-70.

Table 1. A summary of the average absolute errors for the MLE, WLS, and LSR on the 1977 data set when the simulated region contained 20 segments.

Segment Sample Size	Percent of Segments Partially Classified	$\hat{P}_A$			$\hat{P}_B$			$1 - \hat{P}_A - \hat{P}_B$		
Estimator:		MLE	WLS	LSR	MLE	WLS	LSR	MLE	WLS	LSR
5	20%	.0387	.0387	.0395	.0149	.0150	.0149	.0473	.0472	.0473
	40%	.0385	.0387	.0394	.0172	.0173	.0176	.0463	.0463	.0463
	60%	.0417	.0425	.0426	.0199	.0199	.0202	.0477	.0476	.0477
	80%	.0464	.0610	.0464	.0287	.0439	.0287	.0472	.0474	.0472
10	20%	.0231	.0231	.0236	.0085	.0086	.0087	.0272	.0272	.0272
	40%	.0232	.0233	.0240	.0105	.0108	.0107	.0271	.0271	.0271
	60%	.0241	.0242	.0247	.0131	.0132	.0136	.0269	.0270	.0269
	80%	.0287	.0290	.0296	.0185	.0192	.0193	.0279	.0278	.0279

Table 2. A Summary of the average absolute errors for MLE both with all sample segments completely classified and with only some completely classified and the remainder partially classified (1977 data set, 20 segments in the simulated region).

Segment Sample Size	Percent of Segments Partially Classified	$\hat{P}_A$		$\hat{P}_B$	
		All	Some	All	Some
5	20%	.0383	.0387	.0137	.0149
	40%	.0378	.0385	.0139	.0172
	60%	.0389	.0417	.0140	.0199
	80%	.0385	.0464	.0137	.0387
10	20%	.0227	.0231	.0076	.0085
	40%	.0222	.0232	.0081	.0105
	60%	.0221	.0241	.0078	.0131
	80%	.0229	.0286	.0080	.0185

Table 3. A summary of the average absolute errors for MLE and WLS both with and without the inclusion of the partially classified sample segments (1977 data set, 20 segments in the simulated region).

Segment Sample Size	Percent of Segments Partially Classified	$\hat{P}_A$				$\hat{P}_B$			
Estimator:		MLE		WLS		MLE		WLS	
		With	Without	With	Without	With	Without	With	Without
5	20%	.0387	.0443	.0387	.0443	.0149	.0163	.0150	.0163
	40%	.0385	.0526	.0387	.0526	.0172	.0194	.0173	.0194
	60%	.0417	.0682	.0425	.0682	.0199	.0242	.0199	.0242
	80%	.0464	.1005	.0610	.1005	.0287	.0359	.0439	.0359
10	20%	.0231	.0278	.0231	.0278	.0085	.0096	.0086	.0096
	40%	.0232	.0338	.0233	.0338	.0105	.0121	.0108	.0121
	60%	.0241	.0453	.0242	.0453	.0131	.0163	.0132	.0163
	80%	.0286	.0706	.0394	.0706	.0185	.0241	.0192	.0241

Table 4. A summary of the average absolute errors for the MLE, WLS, and LSR when the segment proportions contain small errors (1977 data set, 20 segments in the simulated region, M=100).

Segment Sample Size	Percent of Segments Partially Classified	$\hat{P}_A$			$\hat{P}_B$			$1 - \hat{P}_A - \hat{P}_B$		
		MLE	WLS	LSR	MLE	WLS	LSR	MLE	WLS	LSR
5	20%	.0453	.0452	.0456	.0174	.0173	.0164	.0514	.0514	.0514
	40%	.0417	.0415	.0422	.0226	.0216	.0215	.0516	.0515	.0516
	60%	.0481	.0488	.0486	.0225	.0222	.0214	.0528	.0529	.0528
	80%	.0698	.1081	.0698	.0569	.1008	.0569	.0470	.0520	.0470
10	20%	.0260	.0260	.0261	.0107	.0109	.0107	.0289	.0290	.0289
	40%	.0286	.0285	.0292	.0130	.0132	.0131	.0320	.0321	.0320
	60%	.0276	.0275	.0272	.0159	.0162	.0157	.0301	.0300	.0301
	80%	.0335	.0347	.0342	.0232	.0238	.0228	.0312	.0311	.0312

Table 5. A summary of the average absolute errors for the MLE, WLS, and LSR when the segment proportions contain moderate errors (1977 data set, 20 segments in the simulated region, M=20).

Segment Sample Size	Percent of Segments Partially Classified	$\hat{P}_A$			$\hat{P}_B$			$1 - \hat{P}_A - \hat{P}_B$		
		MLE	WLS	LSR	MLE	WLS	LSR	MLE	WLS	LSR
5	20%	.0546	.0548	.0535	.0269	.0265	.0230	.0595	.0595	.0595
	40%	.0568	.0551	.0538	.0342	.0323	.0295	.0647	.0646	.0647
	60%	.0660	.0664	.0631	.0378	.0397	.0328	.0665	.0671	.0665
	80%	.0882	.0999	.0882	.0729	.0888	.0729	.0579	.0583	.0579
10	20%	.0352	.0351	.0342	.0178	.0180	.0167	.0387	.0389	.0387
	40%	.0378	.0377	.0373	.0212	.0213	.0198	.0410	.0411	.0410
	60%	.0396	.0388	.0368	.0255	.0249	.0226	.0406	.0406	.0406
	80%	.0587	.0846	.0519	.0464	.0756	.0370	.0424	.0448	.0424