L. Daniel Maxim
Everest Consulting Associates
Princeton Jct., NJ 08550
Leigh Harrington
Mathtech
Rosslyn, VA 22209
Mary Kennedy
Great Falls, VA 22066

Alternative "Scale Up" Estimators for Aerial Surveys where both Detection and Classification Errors Exist

The best choice among these estimators may vary from application to application, depending upon the knowns and the actual values of the detection and identification parameters.

Abstract: *This paper develops, characterizes, and illustrates four statistical estimators to "scale-up" observed counts of objects in aerial surveys where both detection and classification errors are present. Tables and charts are provided to illustrate the properties of these estimators.*

Knowledge of partial ground truth enables efficient estimators to be constructed or can be used in lieu of either detection or identification probabilities. Three of the four estimators make use of partial ground truth.

The Problem Setting

In some remote sensing applications the detection, classification, and identification of objects of interest is essentially a deterministic process. For many, if not most, applications, however, one or more of these steps is probabilistic in nature and the estimation process is statistical in character.

For some applications (see Kraus (1977), Green (1977), or Vizy (1974) for examples) *detection* itself is highly uncertain, and detection probabilities need be explicitly factored into the estimation logic. Maxim *et al.* (1981a) have summarized the appropriate estimation logic in this situation.

In yet other applications, object detection may be assured, but the *identification or classification* steps may be probabilistic. This is often the case in agricultural examples (Ulaby *et al.*, 1980), environmental pollutants (Vizy, 1974; Maxim and Cullen, 1977), and/or where automated classification rules are employed (Hsu, 1978; Piech, 1977; Shimabukuro *et al.*, 1980; or Steiner, 1970). Statistical theory relevant to this problem is extant, but (to our knowledge) has not been extensively developed or employed (see, however, Bauer *et al.* (1978) for one clear exposition).

This paper proposes and characterizes statistical estimators when both detection and identification errors are present. The context of discussion is for an individual image or frame. Extensions to related aspects of sampling theory follow directly or are self-evident. Together with other work (see Maxim *et al.* (1980b) for a discussion of detection versus sampling errors, or the well developed theory of mensuration errors), it can be extended to the simultaneous consideration of detection, classification, mensuration, and sampling errors associated with any survey.

Problem Description

The model presented here considers detection and classification as a two-stage process. For simplicity, only two types of objects are considered, termed Type 1 and Type 2. Matrix methods make

0099-1112/81/4708-1227$02.25/0

extension to more types of objects straightforward if tedious. Categories are defined so as to be mutually exclusive and collectively exhaustive. Detection probabilities are denoted p_1 and p_2, and where detection probabilities are equal are represented by the common symbol, p. (False detections are not explicitly considered here but can easily be imbedded in this analytical framework.) Correct classification/identification probabilities are denoted by $1 - \alpha$ and $1 - \beta$ for Type 1 and Type 2 objects, respectively. For some applications, detection and identification cannot be entirely separated into distinct events. For these cases, the detection probability is more properly a detection and preliminary classification probability, while the quantities $1 - \alpha$ and $1 - \beta$ are conditional identification probabilities. In other cases, the detection of an object is not directly related to its classification, for example, when objects are not detected simply because they are too small. For these situations, the classification can be considered a second and distinct step.

Figure 1 presents an illustrative example where radar imagery is used to detect and identify agricultural fields, as studied and reported by Ulaby *et al.* (1980). At the bottom of Figure 1 is reproduced a portion of an experimentally determined classification matrix after Ulaby *et al.* (1980). In the notation of this paper, Type 1 objects are corn fields and Type 2 objects are soybean fields. (For illustrative purposes, it is assumed that the two categories are exhaustive. In the actual example, there were two additional categories: forests and pasture. This would therefore require a four crop model unless some categories were combined.) As shown on Figure 1, 25 of 37 corn fields were correctly classified, so the value for α in this example is $1 - 25/37 = 0.32$. The β value is 4/39 or about 0.10. In this reference, Ulaby cites no value for the detection probability, p, but from the context it appears that p was close to unity. If so, almost all fields actually in the image were detected.

Y_1 objects of Type 1 (e.g., Y_1 corn fields) are assumed to be located in the area covered by the sample frame, while Y_2 objects of Type 2 are likewise situated. The *observed* counts of Type 1 and Type 2 fields are denoted x_1 and x_2, respectively. Note that x_1 is distributed as the sum of two independent binomials. The first binomial has a probability of success of $p_1(1 - \alpha)$ and a sample size equal to Y_1. The second binomial has a probability of success equal to $p_1\beta$ and a sample size equal to Y_2. A similar structure holds for the distribution of x_2. To simplify the exposition, the detection and classification of objects are assumed to be statistically independent. (An obvious and polar alternative is to assume "common mode" identification failures, i.e., if one object of a given type is misclassified in a frame, then so too are the others. For this extreme case of dependent object-to-object classification, a two-point probability mass function results. Actual applications are bounded by these extremes.)

Given this notation and problem description, the questions addressed in this work are

- how can y_1 and y_2 be estimated from observed data (x_1, x_2) and various assumptions regarding knowledge of p_1, p_2, α, β?
- what are the statistical properties of these estimators and which are to be preferred?

In what follows it is assumed that estimation of one of these quantities, say y_2, is of primary interest. Y_1 may be either assumed known (as from another source) or, alternatively, a nuisance parameter. These questions are next explored.

Estimator A: The Method of Moments

For specified values of p_1, p_2, α, β, y_1, and y_2, the expected values of x_1 and x_2, written $E[x_1]$ and $E[x_2]$ are,

$$E\left[x_1\right] = y_1p_1(1 - \alpha) + y_2p_2\beta, \tag{1}$$

and

$$E\left[x_2\right] = y_1p_1\alpha + y_2p_2(1 - \beta). \tag{2}$$

Equation 1, for example, follows from the logic that x_1 consists of two types of objects: those that are Type 1, detected and correctly identified (the expected number of these equals $y_1p_1(1 - \alpha)$), and those that are Type 2, detected but misclassified as Type 1 (the expected number of these equals $y_2p_2\beta$). A convenient estimator of y_1 and y_2 is found by the method of moments (which for this case is

MATHEMATICAL FORM:

TRUE OBJECT CLASSIFICATION	CLASSIFICATION OUTCOME: TYPE 1	CLASSIFICATION OUTCOME: TYPE 2
TYPE 1	CORRECT / 1-α	ERROR / α
TYPE 2	ERROR / β	CORRECT / 1-β

ENTRIES SHOW DECISION/PROBABILITY NOTATION

A PORTION OF A REAL EXAMPLE (AFTER ULABY ET AL.[15]):

TRUE FIELD TYPE	CLASSIFICATION OUTCOME: CORN	CLASSIFICATION OUTCOME: SOYBEANS	TOTAL
CORN	25 / .68	12 / .32	37
SOYBEANS	4 / .10	35 / .90	39
TOTAL	29	47	76

ENTRIES SHOW NUMBER OF FIELDS/PROBABILITY

Fig. 1. An abbreviated classification matrix and numerical example.

identical to maximum likelihood) in which the observed counts, x_1 and x_2, are set equal to their expectations, $E[x_1]$ and $E[x_2]$, and the equations solved for the unknown quantities, y_1 and y_2. This leads to the estimators $\hat{y}_1$ of y_1 and $\hat{y}_2$ of y_2 shown below,

$$\hat{y}_1 = \frac{(1-\beta)\,x_1 - \beta x_2}{p_1\,(1-\alpha-\beta)}\,,\ \hat{y}_2 = \frac{(1-\alpha)\,x_2 - \alpha x_1}{p_2\,(1-\alpha-\beta)} \tag{3}$$

(When the p_is are equal to unity, the above results correspond, aside from notational differences, to those given in Bauer *et al.* (1980)). By taking the expectation of Equations 3, it can be shown that the estimator, designated A, is unbiased (i.e., $E[\hat{y}_i] = y_i$). From Equation 3 it can be seen that there are circumstances where the estimators are negative, a logical impossibility. This will occur if (assuming $\alpha + \beta < 1$)

$$\text{either } \frac{x_1}{x_2} < \frac{\beta}{1-\beta} \text{ or } \frac{x_2}{x_1} < \frac{\alpha}{1-\alpha}. \tag{4}$$

If either of these conditions occur, the usual practice (e.g., restricted maximum likelihood) would be to set the appropriate estimator, $\hat{y}_i$, equal to zero. This scheme introduces a small bias into the estimate.

$$\mathrm{Var}\left[\hat{y}_2\right] = \frac{1}{(1-\alpha-\beta)^2 p}\left|y_1\alpha\,(1-\alpha) + y_2\left[\beta\,(1-\beta) + (1-p)\,(1-\alpha-\beta)^2\right]\right| \tag{9}$$

The variance of the estimator is an important quantity because it is necessary for the computation of statistical confidence intervals. The variance of the estimator, $\hat{y}_2$, can be determined from the observation that $\hat{y}_2$ is a linear combination of two binomially distributed random variables, x_1 and x_2. The computation is somewhat complicated by the fact that x_1 and x_2 are statistically correlated. The derivation is summarized below.

- The variance of $\hat{y}_2$, a linear combination of random variables is given by (see Mood *et al.* (1974) for a general discussion)

$$\mathrm{Var}\left[\hat{y}_2\right] = \left|\frac{(1-\alpha)}{p_2\,(1-\alpha-\beta)}\right|^2 \mathrm{Var}\left[x_2\right] + \left|\frac{\alpha}{p_1(1-\alpha-\beta}\right|^2 \mathrm{Var}\left[x_1\right] - \frac{2\alpha\,(1-\alpha)}{p_1p_2\,(1-\alpha-\beta)^2}\,\mathrm{Cov}\left[x_1,x_2\right] \tag{5}$$

- The quantities x_1 and x_2 are sums of independent binomially distributed random variables, hence,

$$\mathrm{Var}\left[x_2\right] = y_1p_1\alpha\,(1-p_1\alpha) + y_2p_2(1-\beta)\,(1-p_2(1-\beta)), \tag{6}$$

and

$$\mathrm{Var}\left[x_1\right] = y_1p_1\,(1-\alpha)\,(1-p_1(1-\alpha)) + y_2p_2\beta\,(1-p_2\beta). \tag{7}$$

- The derivation of the covariance of x_1 and x_2 is more complicated and Appendix A is used to show that,

$$\mathrm{Cov}\left[x_1,x_2\right] = -p_1p_2\alpha\,(1-\alpha)\,y_1 - p_1p_2\beta\,(1-\beta)y_2, \tag{8}$$

and so x_1 and x_2 are negatively correlated. The strength of the association is a function of α, β, p_1, p_2 and the relative magnitudes of y_1 and y_2. Table 1 furnishes some idea of this association for the case $y_1 = y_2$, in terms of the correlation coefficient, ρ, for selected values of α, β, and a common detection probability, p. As can be be seen, ρ values are close to unity when p, α, and β are large. ρ values are zero when α and β are zero and small when the detection probability is small. The detection probability acts to control the maximum value of ρ that is attained when α and β approach unity. The assumption $y_1 = y_2$ makes the problem symmetric in α and β. When this is not the case, the symmetry disappears and either α or β becomes more dominant.

Assuming a common detection probability, Equations 5, 6, 7, and 8 above can be combined and after much simplification, the following formula for the variance of the estimate is obtained:

The equation for the variance of $\hat{y}_2$ is more complicated if the detection probabilities are not equal (but a closed form expression results). A convenient summary of the model, resulting observations and their variance-covariance matrix, and estimators and their variance is shown in Figure 2.

A Numerical Example

To illustrate the above, assume that Type 1 and Type 2 objects are corn and soybean fields, respectively, and that the misclassification matrix is as shown in Figure 1, i.e., $\alpha = 0.32$ and $\beta = 0.10$ in accord with the results of Ulaby *et al.* (1980) using L-Band radar. Assume further that the overall field detection probability, p, is 0.8 (not provided in Ulaby) and that imagery readout of a given quadrat produced $x_1 = 78$ apparent corn fields and $x_2 = 242$

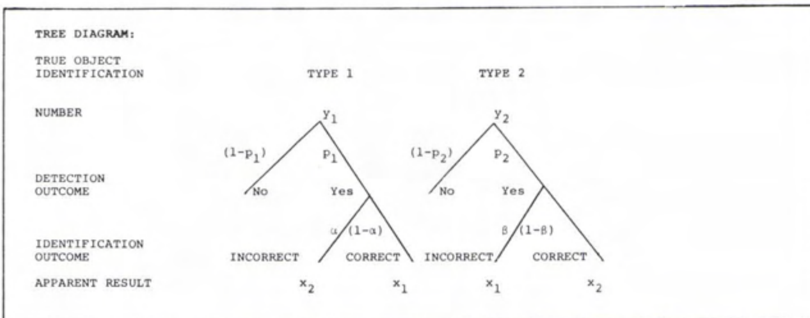
TABLE 1. CORRELATION COEFFICIENTS BETWEEN x_1 AND x_2 AS A FUNCTION OF α, β, AND p WHEN $y_1 = y_2$ (NOTE ALL SIGNS NEGATIVE. VALUES TRUNCATED AT THREE DECIMAL PLACES)

Value of β	Value of α						Value of p
	0	0.1	0.2	0.3	0.4	0.5	
0	0						
0.1	0.631	0.773			SYMMETRIC		
0.2	0.753	0.826	0.858				
0.3	0.801	0.851	0.875	0.888			0.95
0.4	0.822	0.863	0.884	0.895	0.901		
0.5	0.829	0.867	0.886	0.897	0.903	0.904	
0	0						
0.1	0.448	0.618			SYMMETRIC		
0.2	0.592	0.692	0.742				
0.3	0.657	0.730	0.769	0.790			0.90
0.4	0.689	0.750	0.783	0.802	0.812		
0.5	0.700	0.757	0.788	0.806	0.815	0.818	
0	0						
0.1	0.213	0.350			SYMMETRIC		
0.2	0.327	0.429	0.489				
0.3	0.393	0.476	0.526	0.557			0.75
0.4	0.430	0.503	0.547	0.575	0.590		
0.5	0.447	0.515	0.556	0.581	0.595	0.600	
0	0						
0.1	0.082	0.152			SYMMETRIC		
0.2	0.14	0.200	0.242				
0.3	0.179	0.233	0.270	0.295			0.50
0.4	0.204	0.254	0.288	0.311	0.324		
0.5	0.218	0.265	0.297	0.318	0.329	0.333	
0	0						
0.1	0.029	0.056			SYMMETRIC		
0.2	0.051	0.077	0.096				
0.3	0.068	0.092	0.110	0.122			0.25
0.4	0.079	0.102	0.119	0.130	0.137		
0.5	0.086	0.109	0.124	0.135	0.140	0.143	
0	0						
0.1	0.009	0.0196			SYMMETRIC		
0.2	0.017	0.0271	0.034				
0.3	0.023	0.0328	0.039	0.4459			0.10
0.4	0.028	0.0369	0.043	0.0478	0.0506		
0.5	0.030	0.0394	0.045	0.0495	0.0518	0.053	

apparent soybean fields. Substituting these values into Equation 3 yields estimates $\hat{y}_1 \cong 100$ and $\hat{y}_2 \cong 301$ for the actual number of corn and soybean fields, respectively. The variance of the estimated number of soybean fields can be calculated from Equation 9. To do so requires substituting the estimated values, $\hat{y}_1$ and $\hat{y}_2$, for their actual values. Calculated in this manner, the variance is 256 and thus the standard error is about 16. A 95 percent confidence interval on the estimate $\hat{y}_2$ is from 269 to 333. Calculation of the hectareage in corn or soybean requires additional steps and is not treated here.

SOME PROPERTIES OF THE VARIANCE ASSUMING $p_1 = p_2$

The variance of estimator A as shown in Equation 9 is a function of the detection probability, misclassification probabilities, and the true but unknown actual number of objects of each type, y_1 and y_2. As the above example shows, given the observed number of objects x_1 and x_2 and assuming α, β, and p are known, y_1 and y_2 can be estimated and these estimates used in place of the actual values to estimate the variance. In actual application this is an appropriate computational proce-



OBSERVATIONS:

EXPECTATION

$$E[x_1] = y_1 p_1 (1-\alpha) + y_2 p_2 \beta$$

$$E[x_2] = y_1 p_1 \alpha + y_2 p_2 (1-\beta)$$

VARIANCE/COVARIANCE MATRIX

$$\mathrm{Cov}(x_1 x_2) = y_1 \begin{bmatrix} p_1(1-\alpha)(1-p_1(1-\alpha)), & -p_1 p_2 \alpha(1-\alpha) \\ -p_1 p_2 \alpha(1-\alpha), & p_2 \alpha(1-p_2\alpha) \end{bmatrix} + y_2 \begin{bmatrix} p_2\beta(1-p_2\beta), & -p_1 p_2 \beta(1-\beta) \\ -p_1 p_2 \beta(1-\beta), & p_1(1-\beta)[1-p_1(1-\beta)] \end{bmatrix}$$

ESTIMATORS:

$$\hat{y}_1 = \frac{(1-\beta)x_1 - \beta x_2}{p_1(1-\alpha-\beta)} \qquad \hat{y}_2 = \frac{(1-\alpha)x_2 - \alpha x_1}{p_2(1-\alpha-\beta)}$$

$$\mathrm{VAR}\left[\hat{y}_2\right] = \frac{1}{(1-\alpha-\beta)^2 p}\left[y_1 \alpha(1-\alpha) + y_2[\beta(1-\beta) + (1-p)(1-\alpha-\beta)^2]\right], \text{ assuming } p_1 = p_2 = p$$

FIG. 2. Summary of results, Estimator A.

dure. More generally, however, it is interesting to note the sensitivity of the variance to the parameters α, β, and p as these values are often at least partially controllable. For example, the balance between α and β (technology held fixed) can often be varied by altering the multiple classification decision rules or weights (see Maxim and Cullen (1977) for a discussion and Appendix B for a numerical example). The choice of platform, film type, access geometry, or season, for example, can alter both detection and identification probabilities (see Green *et al* (1977) for one illustration).

From Equation 1 the variance is not only a function of α, β, and p but also of y_1 and y_2. The dependence upon y_1 and y_2 must somehow be 'factored out' to see clearly the significance of α, β, and p. This can be done as follows. Let $y = y_1 + y_2$ be the total number of objects. Now when $y_1 = y - y_2$ is substituted into Equation 9, it can be observed that, for fixed y, the numerator is linear in y_2. Thus, the variance is maximized with respect to y_2 when y_2 is set at one of two extreme points, i.e., $y_2 = y$ or $y_2 = 0$, depending upon whether the sign of the coefficient of y_2 is + or −, respectively. After some simplification, it can be shown that the value of this maximum variance is given by

$$\mathrm{var}\left[\hat{y}_2\right] = \begin{cases} \dfrac{y}{p(1-\alpha-\beta)^2}\left[\alpha(1-\alpha) + (1-\alpha-\beta)(1-2\alpha-p(1-\alpha-\beta))\right], \\ \qquad \text{if } \dfrac{1-2\alpha}{(1-\alpha-\beta)} \geqslant p \\ \qquad \text{or,} \\ \dfrac{y}{p(1-\alpha-\beta)^2}\left[\alpha(1-\alpha)\right], \text{ if } \dfrac{1-2\alpha}{(1-\alpha-\beta)} \leqslant p\ . \end{cases} \tag{10}$$

With this substitution, the quantity

$$\underset{y_2}{\mathrm{Max}}\ \mathrm{Var}\left(\frac{[\hat{y}_2],}{y}\right) \tag{11}$$

is an upper bound to the actual variance that is independent of y and solely a function of α, β, and p. Figure 3 shows an illustrative contour (equal to 6.0) for the Function 11 as calculated by Equation

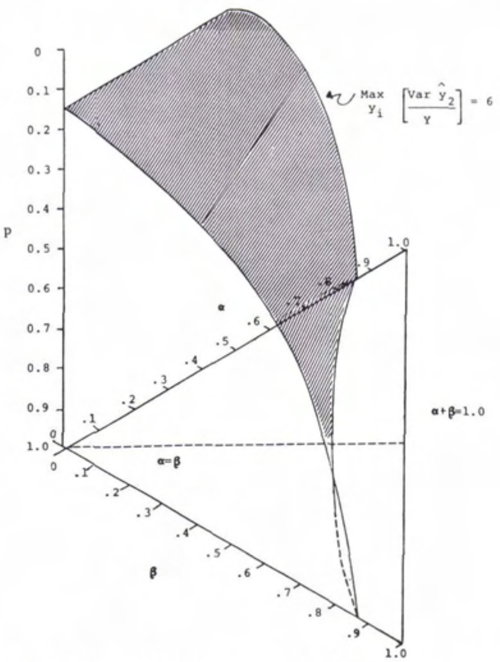


FIG. 3. Surface of maximum value of variance of estimate of Y_2 as a Function of α, β, and p.

10 in α, β, p space. As can be seen, though the surface is somewhat complex, the maximum value for the variance of Estimator A is an increasing function of α and β and a decreasing function of p. The variance becomes infinite as the sum $\alpha + \beta$ approaches unity or as the detection probability approaches zero.

As the upper bound to the variance of $\hat{y}_2$ was found by maximizing Equation 9 with respect to y_2, given the sum y, so too the lower bound on the variance can be found by minimizing the same expression. It is easily seen that the minimum is given by Equation 10 with the inequality conditions switched, i.e.,

$$\min_{y_2} \mathrm{Var}\left[\hat{y}_2\right] = \begin{cases} \dfrac{y}{p\,(1-\alpha-\beta)^2}\left[\alpha\,(1-\alpha) + (1-\alpha-\beta)\,(1-2\alpha-p)\,(1-\alpha-\beta)\right], \\ \qquad \text{if } \dfrac{1-2\alpha}{(1-\alpha-\beta)} \leqslant p, \\ \qquad \text{or,} \\ \dfrac{y}{p\,(1-\alpha-\beta)^2}\left[\alpha\,(1-\alpha)\right], \text{ if } \dfrac{1-2\alpha}{(1-\alpha-\beta)} \geqslant p. \end{cases} \tag{12}$$

Taken together, equations 11 and 12 bound the actual variance given by Equation 9. These results will be used in a later section to compare alternative estimators.

Knowledge of Y_1

It may happen in practice that other sources of data can be used to estimate y_1 (or, in general, one or more of the quantities of interest). An independent agricultural survey, for example, might have estimated the size of the corn crop. The question posed in this section is, how can this increased knowledge help the estimation process?

The answer to the above question is in two parts. First, knowledge of y_1 can be used to develop estimates of y_2 when one or more parameters (e.g., α, β, p) are not known. Second, knowledge of y_1 enables more efficient estimators for $\hat{y}_2$ to be developed. The efficiency gains are functions of p, α, and β and can be substantial in some cases. These points will be illustrated in what follows.

Estimator B

As an illustration of how knowledge of y_1 can be used as a partial surrogate for knowledge of other parameters, suppose that α and β are known, but that a common detection probability, p, is not. x_1 and x_2 are observed quantities as before. If x_1, x_2, y_1, and y_2 were known, then an obvious estimate of p would be given by

$$\hat{p} = \frac{x_1 + x_2}{y_1 + y_2}, \tag{13}$$

the observed fraction of objects detected. Now y_2 is not known, but an estimate $\hat{y}_2$ is available from Estimator A (Equation 3). Substituting this estimate of y_2 into Equation 13 results in a quadratic equation in p that has a non-zero real root,

$$\hat{p} = \frac{(1-\beta)\,x_1 - \beta x_2}{(1-\alpha-\beta)\,y_1}. \tag{14}$$

Having estimated p in this manner, this can be used in Equation 3 to estimate y_2. Upon substitution and simplification, the estimate, called Estimator B, is

$$\hat{y}_2 = \frac{y_1\left[(1-\alpha)\,x_2 - \alpha x_1\right]}{(1-\beta)\,x_1 - \beta x_2} \qquad (15)$$

A Second Numerical Example

Assume observations and parameters as given in the first numerical example but further that p is unknown and that y_1 is known and equal to 110. The estimate of $\hat{p}$ of p from Equation 14 is computed as 0.721, while the estimate B of $\hat{y}_2$ from Equation 15 is 334, somewhat higher than for Estimator A.

The properties of Estimator B are difficult to determine since it involves a ratio of random variables. Appendix B provides an approximation to the variance of this estimator and some remarks concerning the bias.

Estimator C

As a second illustration of how knowledge of y_1 can be helpful, suppose now that a common detection probability, p, is known, but that α and β are unknown. In this case an obvious choice for estimating y_2 is

$$\hat{y}_2 = \frac{x_1 + x_2}{p} - y_1 \,. \qquad (16)$$

To see this, note that $(x_1 + x_2)/p$ is an estimate of y, the total number of objects. Subtracting the known number of objects of Type 1 from this quantity gives an estimate of y_2. It is easily seen (by taking expectations) that this estimate is unbiased and further that the variance is

$$\mathrm{Var}\left[\hat{y}_2\right] = \frac{1-p}{p}\left[y_1 + y_2\right], \qquad (17)$$

or

$$\frac{\mathrm{Var}\left[\hat{y}_2\right]}{y} = \frac{1-p}{p}\,. \qquad (18)$$

A Third Numerical Example

A numerical example will show that Estimator C can be very efficient. Consider the data given in the first numerical example, $x_1 = 78$, $x_2 = 242$ (α and β values are not needed for this estimate and will be disregarded). Finally, assume that there are 110 known corn fields ($y_1 = 110$). Then Estimator C is calculated as 290 from Equation 16 and the variance of this estimate is computed to be 100 from Equation 17. The standard error, 10, is only 63 percent of that from estimate A (see example one).

A general comparison of the variances of Estimators A and C is complicated by the fact that the variance of Estimator A depends upon both y_1 and y_2 rather than their sum as does Estimator C. A convenient approach is to compare the standard error of estimate C to both the minimum and the maximum values of the standard error of Estimator A. This comparison is shown in Tables 2 and 3, respectively. As can be seen, the standard error of Estimator C is less than or at most equal to the *maximum* standard error of Estimator A (see Table 3). Depending upon the values of α, β, and p, it is sometimes less than the *minimum* value of the standard error of Estimator A (see Table 2). In particular, Estimator C is preferred to Estimator A when p approaches one and α and β grow large. For example, the standard error of Estimator C is only 40 percent that of the minimum possible standard error of Estimator A when $p = 0.95$, $\alpha = 0.15$ and $\beta = 0.20$; excellent performance for an estimator that does not require knowledge of either α or β.

Estimator D

Estimator D assumes knowledge of α, β, p, and y_1 and promises increased efficiency over Estimator C. It follows from the moments Equations 1 and 2 when it is assumed that y_1 is known. In this case there are two estimates of y_2 that are possible. The first, denoted z_1, results from Equation 1, i.e.,

$$z_1 = \frac{x_1 - y_1 p\,(1-\alpha)}{p\beta}, \qquad (19)$$

where x_1 is used as an estimate of $\mathrm{E}\left[x_1\right]$, while the second, denoted z_2, results from Equation 2, i.e.,

$$z_2 = \frac{x_2 - y_1 p\alpha}{p\,(1-\beta)}\,. \qquad (20)$$

In general, z_1 and z_2 as defined above will differ. For instance, using the data from example one and assuming that y_1 is known to be 110 as in the third example, z_1 and z_2 are computed as 227 and 297, respectively. The central idea in Estimator D is to form a convex combination of these estimators as $\hat{y}_2 = w_1 z_1 + (1 - w_1)\,z_2$ where the weights are non-negative and sum to unity. The weights can be chosen so as to minimize the variance of $\hat{y}_2$. By taking expectations of both sides of Equation 19, it is easily shown that $\mathrm{E}\,[z_1] = y_2$ and likewise that $\mathrm{E}\,[z_2] = y_2$, and therefore that these estimators are unbiased. So too will a linear combination $w_1 z_1 + (1 - w_1)\,z_2$, so that $\mathrm{E}\,[\hat{y}_2] = y_2$ and thus this estimator is unbiased. The variance of this linear combination is given by (see Mood *et al.* (1974) or Marpet and Maxim (1979) for a general discussion of the problem)

$$\mathrm{Var}\,\hat{y}_2 = w_1^2\,\mathrm{Var}\,z_1 + (1-w_1)^2\,\mathrm{Var}\,z_2 + 2w_1(1-w_1)\,\mathrm{Cov}\,(z_1, z_2). \qquad (21)$$

Denoting $\mathrm{Var}\,z_1$ by the symbol a, $\mathrm{Var}\,z_2$ by the symbol b, and $\mathrm{Cov}\,(z_1, z_2)$ by the symbol c, it is easy to show that the value of w_1, w_1^*, which minimizes the variance of the estimate of y_2 in Equation 21 is given by

$$w_1^* = \frac{b - c}{(a + b - 2c)} \; . \tag{22}$$

Now

$$a = \text{Var } z_1 = \left[\frac{1}{p\beta} \right]^2 \left[y_1 p \,(1 - \alpha)\,(1 - p\,(1 - \alpha)) + y_2 p\beta\,(1 - p\beta) \right], \tag{23}$$

$$b = \text{Var } z_1 = \left[\frac{1}{p\,(1 - \beta)} \right]^2 \left[y_1 p\alpha\,(1 - p\alpha) + y_2 p\,(1 - \beta)\,(1 - p\,(1 - \beta)) \right], \tag{24}$$

and

$$c = \text{Cov}\,(z_1 z_2) = - \left[\frac{1}{p\beta} \right] \left[\frac{1}{p\,(1 - \beta)} \right] \left[y_1 p^2 \alpha\,(1 - \alpha) + y_2 p^2 \beta\,(1 - \beta) \right]. \tag{25}$$

Thus, the optimal weights can be computed and used for the estimate. This procedure requires an iterative solution, however. It is best illustrated with a numerical example.

A Fourth Numerical Example

Using the above data and parameter estimates, z_1 and z_2 are computed as 227 and 297, respectively. The computation of w_1^* from Equation 22 requires a, b, and c as inputs. These, in turn, require knowledge of x_1, x_2, p, α, β, y_1, assumed known, but also y_2, which is unknown.

An iterative algorithm proceeds as follows:

Step (1)—assume weights w_1, $(1 - w_1)$ then compute $\hat{y}_2$ from $\hat{y}_2 = w_1 z_1 + (1 - w_1)\, z_2$.

Step (2)—from x_1, x_2, y_1, α, β, p, and $\hat{y}_2$ compute a, b, and c from Equations 23, 24, and 25.

Step (3)—from a, b, and c compute w_1^*, from Equation 22 and the optimum variance from Equation 21. If the computed value of w_1^* is sufficiently close to that assumed in Step (1), stop; otherwise go to Step (1) and repeat the procedure using the value of w_1^* computed in this step.

To illustrate, assume in Step (1) that $w_1 = 0.5$. Then from $\hat{y}_2 = w_1 z_1 + (1 - w_1)\, z_2$, $\hat{y}_2 \cong 262$. Proceeding to Step (2), the computed values of a, b, and c from Equations 23, 24, and 25 using $\hat{y}_2 = 262$ are 7276.6, 142.3, and -527.95, respectively. Now, from Equation 22, a new value of w_1^* can be computed as 0.07909. Since this differs appreciably from the value (0.5) assumed in Step (1), the process is repeated, using $w_1 = 0.07909$ for this trial. Table 4 shows the results of subsequent iterations of this algorithm. The procedure converges rapidly; in this instance agreement to four decimal places in the weights occurs after three iterations. The final estimate, $\hat{y}_2$, is 291.4. The computed value of the standard error, 9.839, is slightly smaller than that for Estimator C and substantially beneath that of Estimator A. For this example the efficiency gain of Estimator D over Estimator C does not justify the added computational effort. The ready availability of computers facilitates the use of Estimator D for all but 'on the spot' computations.

The Choice Among Estimators

Four estimators have been proposed, illustrated, and characterized in this paper. The best choice

Table 2. Ratio of STD Error of Estimator C to Minimum STD Error of Estimator A

Assumed Value of p	Assumed Value of β	Value of α: 0	0.05	0.10	0.15	0.20
0.95	0	∞	1.000	1.000	1.000	1.000
	0.05	∞	0.923	0.657	0.634	0.610
	0.10	∞	0.872	0.596	0.488	0.463
	0.15	∞	0.821	0.559	0.438	0.377
	0.20	∞	0.769	0.522	0.407	0.335
0.75	0	∞	2.18	1.50	1.19	1.000
	0.05	∞	2.06	1.42	1.12	0.938
	0.10	∞	1.95	1.33	1.05	0.875
	0.15	∞	1.84	1.25	0.980	0.813
	0.20	∞	1.72	1.17	0.910	0.750
0.50	0	∞	3.08	2.12	1.68	1.41
	0.05	∞	2.92	2.00	1.58	1.33
	0.10	∞	2.76	1.89	1.49	1.25
	0.15	∞	2.60	1.77	1.39	1.15
	0.20	∞	2.43	1.65	1.29	1.06

Table 3. Ratio of STD Error of Estimator C to Maximum STD Error of Estimator A

Assumed Value of p	Assumed Value of β	Value of α: 0	0.05	0.10	0.15	0.20
0.95	0	1.000	0.975	0.671	0.532	0.447
	0.05	0.698	0.678	0.633	0.501	0.419
	0.10	0.557	0.535	0.512	0.470	0.391
	0.15	0.470	0.447	0.425	0.401	0.366
	0.20	0.408	0.387	0.364	0.342	0.318
0.75	0	1.000	1.000	1.000	1.000	1.000
	0.05	0.909	0.900	0.889	0.878	0.865
	0.10	0.832	0.817	0.800	0.781	0.759
	0.15	0.766	0.746	0.724	0.700	0.673
	0.20	0.707	0.684	0.659	0.631	0.600
0.50	0	1.000	1.000	1.000	1.000	1.000
	0.05	0.951	0.946	0.940	0.933	0.925
	0.10	0.905	0.895	0.883	0.870	0.855
	0.15	0.860	0.846	0.830	0.811	0.790
	0.20	0.817	0.798	0.777	0.754	0.727

TABLE 4. ITERATIVE SOLUTION TO OPTIMAL WEIGHTED ESTIMATOR SHOWS RAPID CONVERGENCE FOR EXAMPLE PROBLEM (WEIGHTS STABLE TO FOUR DECIMAL PLACES AFTER THREE ITERATIONS)

Iteration	Assumed Weights		Resulting Estimate	Computed Variances			Computed Optimal Weights		Resulting Estimate	Δ	Computed Estimate of σ Given Assumed Weights
	W_1	W_3	Y_2	VAR Z_1	VAR Z_2	COV (Z_1, Z_2)	W_1	W_2	Estimate Y_2	Estimate as %	
1	0.5	0.5	262	7276.6	142.3	−527.95	0.07909	0.92091	291.46	11.24	39.88
2	0.7909	0.92091	291.46	7615.39	153.76	−557.42	0.08005	0.91995	291.36	0.0218	9.841
3	0.08005	0.91995	291.396	7614.65	153.74	−557.35	0.08005	0.91995			
Assumptions:				$Y_1 = 110$	$\alpha = 0.32$		$\beta = 0.10$	$p = 0.80$			
				$x_1 = 78$	$x_2 = 242$		$z_1 = 227$	$z_2 = 297$			

among these estimators may vary from application to application, depending upon the knowns and the actual values of the detection and identification parameters. First, the knowns serve to limit the available choices. Table 5 summarizes the estimators, their variance, and the required input parameters. Note, for example, that only one choice (Estimator A) is available if y_1 is not known. Required knowledge about the other parameters may further limit options. Thus, if neither α nor β is known, Estimator C is the only choice. If data exist to support more than one estimator, the variance of each estimator can be computed from the appropriate equations and the minimum variance estimator selected.

EXTENSIONS

As mentioned, perhaps the most useful and easiest generalization of this work is to consider multiple categories of objects of interest. Estimates analogous to Estimator A generalize readily using matrices. The properties of these estimators also follow directly, though the dimensionality of the problem varies as the square of the number of distinct objects considered. The cases that follow where y_1 is known increase in combinatorial fashion with the number of distinct objects, though, aside from tedious algebraic bookkeeping, this is not a formidable task.

The extension to multiple quadrats is trivial (if α, β, and p are approximately constant, the quadrats can be 'pieced together' and regarded as one quadrat), except when the quadrats are decidedly non-homogeneous. In this situation, various stratification schemes can be employed.

When multiple quadrats are considered, it may be the case that (in the notation of this paper) y_1 is not known on a quadrat by quadrat basis, but rather only in aggregate. In this case the multidimensional proration ideas of Deming *et al.* (1940) and Stephan (1942) are applicable, though computation of the properties of the resulting estimator is likely to be difficult.

ACKNOWLEDGMENTS

The authors wish to thank the referees and the editors for their useful comments and suggestions on an earlier draft of this paper.

APPENDIX A

To show that the covariance of x_1 and x_2 is as given in Equation 8, consider a single Type 1 field. If it has been detected, then it is either classified correctly as a Type 1 field and so $x_1 = 1, x_2 = 0$ or misclassified as a Type 2 field and $x_1 = 0$ and $x_2 = 1$. By definition,

$$\text{Cov}\ (x_1 x_2) = \text{E}\ (x_1 x_2) - \text{E}\ (x_1)\ \text{E}\ (x_2).$$

Since $x_1 \cdot x_2 = 0$ regardless of the classification, the above expression reduces to

TABLE 5. CHARACTERISTICS OF ALTERNATIVE ESTIMATORS ASSUMING A COMMON DETECTION PROBABILITY

Assumed Knowns			Estimator		
Detection Probability (p)	Misclassification Probability (α,β)	Y_1	Designator	Formula	Variance
Yes	Yes	No	A	$\hat{y}_2 = \dfrac{(1-\alpha)x_2 - \alpha x_1}{p(1-\alpha-\beta)}$	$\dfrac{1}{p(1-\alpha-\beta)^2}\left\lvert y_1\alpha(1-\alpha) + y_2[\beta(1-\beta) + (1-p)(1-\alpha-\beta)^2]\right\rvert$
No	Yes	Yes	B	$\hat{p} = \dfrac{(1-\beta)x_1 - \beta x_2}{(1-\alpha-\beta)y_1}$ $\hat{y}_2 = \dfrac{(1-\alpha)x_2 - \alpha x_1}{\hat{p}(1-\alpha-\beta)}$	Complex, approximation available, see Appendix B
Yes	No	Yes	C	$\hat{y}_2 = \dfrac{x_1 + x_2}{p} - y_1$	$\dfrac{(1-p)}{p}[y_1 + y_2]$
Yes	Yes	Yes	D	$\hat{y}_2 = W_1Z_1 + W_2Z_2$ $\hat{Z}_1 = \dfrac{x_1 - y_1p(1-\alpha)}{p\beta}$ $\hat{Z}_2 = \dfrac{x_2 - y_1p\alpha}{p(1-\beta)}$	Iterative algorithm required, see text

$$\begin{aligned}\mathrm{Cov}\,(x_1x_2) &= -\mathrm{E}\,(x_1)\,\mathrm{E}\,(x_2)\\ &= -p_1\,(1-\alpha)\cdot p_2\,(\alpha).\end{aligned}$$

If there are y_1 Type 1 fields, the covariance is simply the sum of y_1 such outcomes, and so

$$\mathrm{Cov}\,(x_1x_2) = -y_1p_1p_2\,(1-\alpha)\,\alpha.$$

A similar argument applies assuming there are y_2 Type 2 fields. Summing these two expressions gives Equation 8.

APPENDIX B

This appendix calculates the approximate expected value and variance of Estimator B:

$$\hat{y}_2 = \frac{y_1\left[(1-\alpha)\,x_2 - \alpha x_1\right]}{(1-\beta)\,x_1 - \beta x_2},\ \text{or} \tag{B1}$$

$$\hat{y}_2 = y_1\left[v_1/v_2\right],\ \text{say}, \tag{B2}$$

where v_1 and v_2 are random variables that are linear functions of the observation x_1 and x_2. Now, the expected value and variance of the ratio of two random variables is given approximately by the formulas (Mood *et al.*, 1974)

$$\mathrm{E}\,(v_1/v_2) \cong \frac{\mathrm{E}\,(v_1)}{\mathrm{E}\,(v_2)} - \frac{\mathrm{Cov}\,(v_1v_2)}{\mathrm{E}^2\,(v_2)} + \frac{\mathrm{E}\,(v_1)}{\mathrm{E}^3\,(v_2)}\,\mathrm{Var}\,(v_2), \tag{B3}$$

and

$$\mathrm{Var}\,(v_1/v_2) \cong \frac{\mathrm{E}^2\,(v_1)}{\mathrm{E}^2\,(v_2)}\left[\frac{\mathrm{Var}\,(v_1)}{\mathrm{E}^2\,(v_1)} + \frac{\mathrm{Var}\,(v_2)}{\mathrm{E}^2\,(v_2)} - \frac{2\,\mathrm{Cov}\,(v_1v_2)}{\mathrm{E}\,(v_1)\,\mathrm{E}\,(v_2)}\right]. \tag{B4}$$

From the results given in Figure 2, it can be shown that

$$\begin{aligned}\mathrm{E}\,(v_1) &= py_2,\\ \mathrm{E}\,(v_2) &= py_1,\end{aligned}$$

and further that,

$$\begin{aligned}\mathrm{Var}\,(v_1) &= p\left[y_1\alpha\,(1-\alpha) + y_2\{\beta\,(1-\beta) + (1-p)\,(1-\alpha-\beta)^2\}\right]\\ \mathrm{Var}\,(v_2) &= p\left[y_2\beta\,(1-\beta) + y_1\{\alpha\,(1-\alpha) + (1-p)\,(1-\alpha-\beta)^2\}\right]\\ \mathrm{Cov}\,(v_1v_2) &= -p\,(\alpha\,(1-\alpha)\,y_1 + \beta\,(1-\beta)\,y_2)\end{aligned}$$

Substituting these quantities into Equations B3 and B4, the approximate mean and variance of $\hat{y}_2$ is given by

$$\begin{aligned}\mathrm{E}\,(\hat{y}_2) = {} & y_2 + \frac{\alpha\,(1-\alpha)}{p} + \frac{\beta\,(1-\beta)}{p}\left(\frac{y_2}{y_1}\right)^2\\ & + \left(\frac{y_2}{py_1}\right)\left(\alpha\,(1-\alpha) + \beta\,(1-\beta) + (1-p)\,(1-\alpha-\beta)^2\right)\end{aligned} \tag{B5}$$

$$\begin{aligned}\mathrm{Var}\,(\hat{y}_2) = {} & \frac{1}{p}\,y_1\alpha\,(1-\alpha) + y_2\,(\beta\,(1-\beta) + (1-p)\,(1-\alpha-\beta)^2\\ & + \left(\frac{y_2}{y_1}\right)^2\left(y_2\beta\,(1-\beta) + y_1\left(\alpha\,(1-\alpha) + (1-p)\,(1-\alpha-\beta)^2\right)\right)\\ & + 2\left(\frac{y_2}{y_1}\right)(\alpha\,(1-\alpha)\,y_1 + \beta\,(1-\beta)y_2)\end{aligned} \tag{B6}$$

For the numerical example given earlier, when $p = 0.8$, $\alpha = 0.32$, $\beta = 0.1$ and assuming $y_1 = 110$, Equations B5 and B6 become

$$\mathrm{E}\,(\hat{y}_2) \cong 0.272 + 1.0043\,y_2 + 0.0000093\,y_2^2$$

and

$$\begin{aligned}\mathrm{Var}\,(\hat{y}_2) \cong {} & 29.92 + 0.7406\,y_2 + 0.0042\,y_2^2\\ & + 0.00000744\,y_2^3.\end{aligned}$$

Substituting the estimated value of y_2, 334, into these results, $\mathrm{E}\,(\hat{y}_2)$ is estimated to be 336.75, indicating a positive bias equal to 2.75, while the variance is estimated to be 1023.03. The standard deviation (31.98) is substantially larger than those calculated for three other estimators proposed in this report. This estimate, however, does not require knowledge of the detection probability, p (although estimates of the mean and variance, Equations B5 and B6, do).

APPENDIX C

This appendix discusses the relationship between the misclassification errors, α and β, and the variance of Estimator A given in Table 5, i.e.,

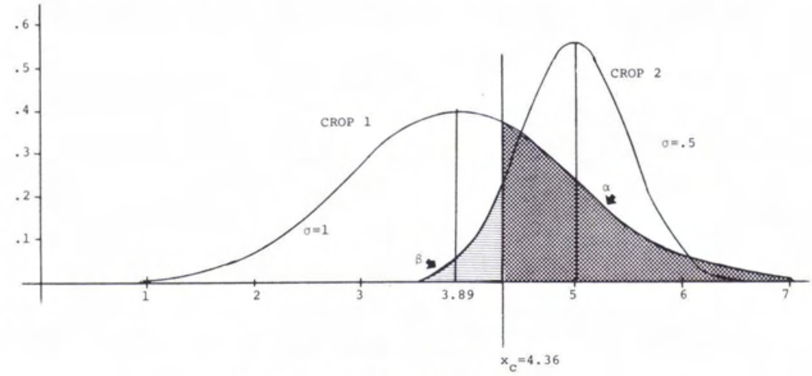


FIG. C1. A two crop classifier.

$$\hat{y}_2 = \left[(1-\alpha)\,x_2 - \alpha x_1\right]/\left[p\,(1-\alpha-\beta)\right].$$

In many applications, especially when the information from the aerial survey has been digitized, statistical discrimination schemes are developed to classify agricultural fields into the various crops. Such classification rules typically have the form that, if a variable or linear combination of variables exceeds some critical value, then the field would be classified as crop 1, else it would be classified as crop 2 (in the two crop problem). Figure C1 illustrates this structure, assuming some numeri-

TABLE C1. NUMERICAL SEARCH SHOWS HOW α, β AND VAR($\hat{y}_2$) DEPEND UPON DECISION THRESHOLD (BASE CASE COMPUTATION ENCLOSED BY LINES)

x_c Assumed	$t_1 = \frac{x_c - 3.89}{1}$	α	$t_2 = \frac{x_c - 5}{0.5}$	β	$\alpha + \beta$	Var ($\hat{y}_2$)	
3.89	0	0.50	−2.22	0.013	0.513	227.1	
3.90	0.010	0.496	−2.20	0.014	0.510	226.7	
3.95	0.060	0.476	−2.10	0.018	0.494	222.7	
4.00	0.110	0.456	−2.00	0.023	0.479	220.3	α, β choice
4.05	0.160	0.436	−1.90	0.029	0.465	219.3	to minimize
4.10	0.210	0.417	−1.80	0.036	0.453	220.1	variance
4.15	0.260	0.397	−1.70	0.045	0.442	222.9	
4.20	0.310	0.378	−1.60	0.055	0.433	227.0	
4.25	0.360	0.359	−1.50	0.067	0.426	233.5	
4.30	0.410	0.341	−1.40	0.081	0.422	242.6	
4.35	0.460	0.323	−1.30	0.097	0.420	253.9	α, β choice
4.36	0.470	0.319	−1.28	0.100	0.419	255.4	to minimize
4.40	0.510	0.305	−1.20	0.115	0.420	267.2	$\alpha + \beta$
4.45	0.560	0.288	−1.10	0.136	0.424	285.1	
4.50	0.610	0.271	−1.00	0.159	0.430	305.3	
4.55	0.660	0.255	−0.90	0.184	0.439	329.4	
4.60	0.710	0.239	−0.80	0.212	0.451	358.3	
4.65	0.760	0.221	−0.70	0.242	0.463	388.2	
4.70	0.810	0.209	−0.60	0.274	0.483	431.4	
4.75	0.860	0.195	−0.50	0.309	0.504	480.2	
4.80	0.910	0.181	−0.40	0.345	0.526	534.6	
4.85	0.960	0.169	−0.30	0.382	0.551	601.2	
4.90	1.010	0.156	−0.20	0.421	0.577	677.9	
4.95	1.060	0.145	−0.10	0.460	0.605	771.3	
5.00	1.110	0.134	0	0.500	0.634	883.1	

Assumptions: $\mu_2 = 5$ $\quad \sigma_2 = 0.5$ $\quad p = 0.8$ $\quad y_1 = 100$ $\quad y_2 = 300$
$\mu_1 = 3.89$ $\quad \sigma_1 = 1.0$

Decision Rule = if $x \geqslant x_c$ call type 2
if $x < x_c$ call type 1

cal attribute (e.g., reflectance) of both crops is normally distributed with regard to the chosen variable (the normality assumption is not required for that follows, however).

As can be observed, lowering the critical value x_c will reduce the number of times crop 2 is misclassified as crop 1, β, but will increase the misclassification rate for crop 1, α*. Everything else being equal, x_c would often be chosen, ultimately, so as to minimize the sum of the misclassification rates. However, if the aerial survey is being performed to estimate crop 2 alone, using Estimator A, above, then α and β might better be selected so as to reduce the variance of this estimator (see Equation 9) given by

$$\begin{aligned}\text{Var}\,(y_2) = {} & \left[1/p\,(1-\alpha-\beta)^2\right]\left[y_1\alpha\,(1-\alpha)\right. \\ & + y_2\left[\beta\,(1-\beta) - (1-p)\right. \\ & \left.(1-\alpha-\beta)^2\right].\end{aligned}$$

To illustrate the possible gains from such an approach, note that when the critical value is 4.36, as in Figure C1, then $\alpha = 0.32$ and $\beta = 0.1$ and, in fact, this x_c minimizes the sum, $\alpha + \beta$. Assuming the true number of type 1 crops, y_1, is 100 and type 2 crops, y_2, is 300 as in previous examples with $p = 0.8$, then Var $(\hat{y}_2) = 255.4$. Table C1 presents the results of this and similar calculations for values of x_c from 3.89 to 5.00. This table shows that a cutoff point, x_c, equal to 4.05 minimizes Var (y_2) and at this point the variance is equal to 219.3. This is a 14 percent reduction over the base case and illustrates the benefits of a judious selection of α and β errors. Similar remarks obtain for other estimators whose variance depends upon α and β.

* For an interesting paper on actual error ratio in discriminant analysis based upon small samples, see Sayre, J. W., The Distribution of the Actual Error Rates in Linear Discriminant Analysis, *Journal of the American Statistical Association,* Theory and Methods Section, Vol. 75, Number 369, March 1980, pp. 201-205.

References

Bauer, M. E., M. M. Hixson, B. J. Davis, and J. B. Etheridge, 1978. Area Estimation of Crops by Digital Analysis of Landsat Data, *Photogrammetric Engineering and Remote Sensing,* Vol. 44, No. 8, pp. 1033-1043.

Deming, W. Edwards, and Frederick F. Stephan, 1940. On a Least Squares Adjustment of a Sample Frequency Table When the Expected Marginal Totals are Known, *Annals of Mathematical Statistics,* Vol. XI, No. 4, pp. 427-444.

Green, L. R., J. K. Olson, W. G. Hart, and M. R. Davis, 1977. Aerial Photographic Detection of Imported Fire Ant Mounds, *Photogrammetric Engineering and Remote Sensing,* Vol. 43, No. 8, pp. 1051-1058.

Hsu, Shin-Yi, 1978. Texture-Tone Analysis for Automated Land-Use Mapping, *Photogrammetric Engineering and Remote Sensing,* Vol. 44, No. 11, pp. 1393-1404.

Kraus, S. P., J. E. Estes, S. G. Atwater, J. R. Jensen, and R. R. Vollmers, 1977. Radar Detection of Surface Oil Slicks, *Photogrammetric Engineering and Remote Sensing,* Vol. 43, No. 12, pp. 1523-1531.

Marpet, M. I., and L. D. Maxim, 1979. Optimal Aggregation of Statistical Estimators, paper presented to an optimization workshop, Montreal, Que., Canada.

Maxim, L. D., and D. E. Cullen, 1977. A Cost Model for Remote Inspection of Ground Sites, *Photogrammetric Engineering* and *Remote Sensing,* Vol. 43, No. 8, pp. 1009-1025.

Maxim, L D., L. Harrington, and M. Kennedy, 1980a. A Capture-Recapture Approach for Estimation of Detection Probabilities in Aerial Surveys, *Photogrammetric Engineering and Remote Sensing,* Vol. 47, No. 6, pp. 779-788.

Maxim, L D., L. Harrington, H. D. Weed, and M. Kennedy, 1980b. Intensity Versus Extent of Coverage, *Photogrammetric Engineering and Remote Sensing,* Vol. 47, No. 6, pp. 789-797.

Mood, A. M., F. A. Graybill, and D. C. Boes, 1974. *Introduction to the Theory of Statistics:* McGraw Hill, New York, Third Edition, pp. 178 et seq.

Piech, K. R., D. W. Gaucher, J. R. Schott, and P. G. Smith, 1977. Terrain Classification Using Color Imagery, *Photogrammetric Engineering and Remote Sensing,* Vol. 43, No. 4, pp. 507-513.

Shimabukuro, Y. E., P. Hernandez F., N. F. Koffler, and S. C. Chen, 1980. Automatic Classification of Reforested Pine and Eucalyptus Using Landsat Data, *Photogrammetric Engineering and Remote Sensing,* Vol. 46, No. 2, pp. 209-216.

Steiner, D., 1970. Time Dimensions of Crop Surveys From Space, *Photogrammetric Engineering and Remote Sensing,* Vol. 36, No. 2, pp. 187-194.

Stephan, Frederick F., 1942. An Iterative Method of Adjusting Sample Frequency Tables When Expected Marginal Totals are Known, *Annals of Mathematical Statistics,* Vol. XIII, No. 2, pp. 166, 178.

Ulaby, F. T., P. P. Batlivala, and J. E. Bare, 1980. Crop Identification with L-Band Radar, *Photogrammetric Engineering and Remote Sensing,* Vol. 46, No. 1, pp. 101-105.

Vizy, K. N., 1974. Detecting and Monitoring Oil Slicks with Aerial Photos, *Photogrammetric Engineering and Remote Sensing,* Vol. 40, No. 6, pp. 697-708.

(Received 10 May 1980; revised and accepted 24 November 1980)