Conference on

Machine Processing of

Remotely Sensed Data

October 16 - 18, 1973

The Laboratory for Applications of
Remote Sensing

Purdue University
West Lafayette
Indiana

ESTIMATION OF PROPORTIONS OF OBJECTS AND DETERMINATION OF TRAINING SAMPLE-SIZE IN A REMOTE SENSING APPLICATION*

R. S. Chhikara and P. L. Odell

The University of Texas at Dallas
Dallas, Texas

I. ABSTRACT

One of the problems in remote sensing is estimating the expected proportions of certain categories of objects which cannot be observed directly or distinctly. For example, a multi-channel scanning device may fail to observe objects because of obstructions blocking the view, or different categories of objects may make up a resolution element giving rise to a single observation. This will require ground truth on any such categories of objects for estimating their expected proportions associated with various classes represented in the remote sensing data. Considering the classes to be distributed as multivariate normal with different mean vectors and common covariance, we give the maximum likelihood estimates for the expected proportions of objects associated with different classes, using the Bayes procedure for classification of individuals obtained from these classes. An approximate solution for simultaneous confidence intervals on these proportions is given, and thereby a sample-size needed to achieve a desired amount of accuracy for the estimates has been determined.

II. INTRODUCTION

We consider here the application of the remote sensing technique for ascertaining those types of earth resources which are not observable directly or distinctly using remote sensors. For example, the multi-channel spectral measuring devices used as remote sensors may fail to observe any vegetation, flora, etc. grown underneath trees over a large track of land covered by forest. But if one knows how the amount of vegetation and flora is associated with different types of trees, the remote sensing technique can then be easily employed to estimate the expected amounts or proportions of different categories of the former in the region. The other situation may arise when different types of objects fall within the instantaneous field of view of a multispectral scanning device, making up a resolution element that gives rise to a single observation, and it is desired to evaluate the contribution of each type of objects associated with various classes in a region.

---

*This research was supported by NASA Johnson Space Center, Houston, Texas under Contract No. NAS9-12775.

When a region is scanned from above using airborne data scanning devices, identification of observations becomes subject to uncertainty and involves some probability of misclassification which is often an unknown quantity. Also, the true proportions of different categories of objects associated with the underlying classes are generally unknown. Hence the estimation problem for any such types of objects may involve a two-stage sampling process where at the first stage sample observations are from the underlying classes, and at the second stage samples consist of elements from the types of objects under consideration. A general formulation of the problem is stated as follows:

Let $\pi_i, i=1,2,\ldots,m$ be m different classes and every member of these classes be characterized by p common observable features. Denote the measurements associated with an individual I(X) on these features by a vector $X = (X_1, X_2, \ldots, X_p)^T$. Let $0_1, 0_2, \ldots, 0_k$ be k mutually exclusive categories of objects associated with each of the m classes, and assume that there is at least one object associated with each individual from every class. Further, let p_{ij} denote the actual proportion of jth category objects in π_i, $j=1,2,\ldots,k$ and $i=1,2,\ldots,m$. Assuming that the observation X on a selected individual I(X) is the basis of classifying it into one of the given classes and that its identification is subject to uncertainty, let $P(s|i)$ denote the probability of its misclassifying into π_s when it belongs to π_i $(i \neq s)$ and let $P(i|i)$ denote the probability of its correctly classifying into π_i. Then the expected proportion of jth category objects in π_i is given by

$$e_{ij} = \sum_{s=1}^{m} p_{sj} P(s|i), \qquad j=1,2,\ldots,k \tag{2.1}$$

and $i=1,2,\ldots,m$.

It may be noted that for any fixed class π_i, $\sum_{j=1}^{k} e_{ij}=1$, and $e_{ij}=p_{ij}$ for all j if and only if $P(i|i) = 1$. However, the later condition is often not achievable and is highly dependent upon the classification procedure used for discrimination between the classes. In our effort to discriminate well between the classes, we will use the Bayes classification procedure which minimizes the expected probability of misclassification when equal costs of misclassifying individuals from different classes and equal a <u>priori</u> probabilities for the classes are assumed (Anderson, 1958).

Considering both cases of $P(s|i)$ in (2.1) known and unknown, we give an estimate of e_{ij}, $j=1,2,\ldots,k$ and $i=1,2,\ldots,m$, in each case. We restrict our discussion to the case of normally distributed classes. Below in Section III we outline the sampling procedure and introduce the notations being used in this paper. Our main results are given in Sections IV and V where we discuss simultaneous confidence intervals for e_{ij}'s and a determination of sample size that leads to estimates which with a given probability allow only a specified deviation about each e_{ij}.

III. NOTATIONS AND THE SAMPLING PROCEDURE

For the sake of greater illumination, we consider here the case of two classes π_1 and π_2, and let X be distributed as multivariate normal with mean vector μ if the associated individual I(X) is in π_1 and mean vector ν if I(X) is in π_2. Assume the covariance matrix Σ is the same for both classes. Then given an observation X for a randomly selected individual I(X) from the totality of two classes, it follows under the Bayes classification procedure, assuming equal costs of misclassification for individuals and equal a <u>priori</u> probabilities for the classes, that the two types of probability of misclassification (Anderson, 1958),

$$P(1|2) = P(2|1) = \Phi(-\frac{\Delta}{2}), \tag{3.1}$$

where

$$\Delta^2 = (\mu-\nu)^T \Sigma^{-1} (\mu-\nu) \tag{3.2}$$

and

$$\Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-y^2/2}\, dy.$$

Next, the probabilities of correct classification,

$$P(1|1) = P(2|2) = 1 - \Phi(-\tfrac{\Delta}{2}) . \tag{3.3}$$

Regarding these probabilities of classification, two cases arise. If parameters μ, ν and Σ are all known, $P(1|1)$, $P(2|1)$, $P(1|2)$ and $P(2|2)$ will be known quantities. Then, in order to estimate e_{ij}'s, one only needs to estimate p_{ij}'s $j=1,2,\ldots,k$ and $i=1,2$. This can be achieved by sampling N_1 individuals from π_1 and N_2 from π_2 and then determining separately the observed proportions of k categories of objects associated with these N_1 and N_2 sampled individuals. In case the population of objects associated with the sampled individuals from each class is large, a fixed number of objects, say n_1 and n_2 for π_1 and π_2, respectively, should be selected at the second stage for the purpose of finding the observed proportions estimating p_{ij}'s. The sampling scheme for obtaining samples from π_1 and π_2 is simple random or stratified depending upon whether or not each of these classes consists of one or more than one cluster.

In the context of remote sensing, the sampling process should be executed first by clustering the data allowing the possibility of several clusters per class and then sampling at random the pixels in each cluster. The ground truth at these sampled points is determined by on-site inspection.

When μ, ν and Σ are partially or completely unknown, Δ is unknown, and hence $P(i|j)$, i and $j=1,2$ are unknown quantities. Now for estimating e_{ij}'s one needs to obtain estimates for p_{ij}'s as well as for $P(i|j)$'s. We use the same sampling procedure as in the previous situation except now the observations for the sampled N_1 and N_2 individuals will be utilized to estimate Δ and thereby to estimate $\Phi(-\frac{\Delta}{2})$, since estimation of any $P(i|j)$ amounts to estimating this quantity.

IV. SIMULTANEOUS CONFIDENCE INTERVALS FOR e_{ij}'s

4.1. μ, ν AND Σ ALL KNOWN

Let n_{ijr} denote the number of jth category objects associated with rth sampled individual from π_i. Further, let

$$n_{ij} = \sum_{r=1}^{N_i} n_{ijr}, \quad j=1,2,\ldots,k$$

and

$$n_i = \sum_{j=1}^{k} n_{ij}, \quad i=1,2.$$

Then $\hat{p}_{ij} = n_{ij}/n_i$ provides an unbiased estimate of p_{ij}, $j=1,2,\ldots,k$ and $i=1,2$. It also is the maximum likelihood estimate (MLE). Accordingly, both an unbiased estimate and the MLE of e_{ij} is given by

$$\hat{e}_{ij} = \sum_{s=1}^{2} \hat{p}_{sj} P(s|i), \quad j=1,2,\ldots,k \text{ and } i=1,2. \tag{4.1}$$

We note that the sample size n_i of categorized objects is itself a random variable. In our discussion we would regard it fixed and so our results would hold conditionally. However, if $n_i (i=1,2)$ is large, as is the case being considered at present, such restriction should not affect the unconditional results in any significant way. On the other hand, if a second stage sampling is involved for the selection of these objects, the results will not be subject to any condition.

The random vector $\hat{e} = (\hat{e}_{i1}, \hat{e}_{i2}, \ldots, \hat{e}_{ik})^T$, $(i=1,2)$, is a linear combination of two independent multinomial distributed random vectors $\hat{p}_1 = (\hat{p}_{11}, \hat{p}_{12}, \ldots, \hat{p}_{1k})^T$ and $\hat{p}_2 = (\hat{p}_{21}, \hat{p}_{22}, \ldots, \hat{p}_{2k})^T$, and component-wise

$$E[\hat{e}_{ij}] = e_{ij} \ ,$$

$$\mathrm{Var}(\hat{e}_{ij}) = \sum_{s=1}^{2} P^2(s|i) \frac{p_{sj}(1-p_{sj})}{n_s} \tag{4.2}$$

and

$$\mathrm{Cov}(\hat{e}_{ij}, \hat{e}_{ij'}) = - \sum_{s=1}^{2} P^2(s|i) \frac{p_{sj}p_{sj'}}{n_s} \ , \qquad j \neq j' \tag{4.3}$$

$$j=1,2,\ldots,k \quad \text{and} \quad i=1,2.$$

Moreover, both $\hat{e}_1$ and $\hat{e}_2$ have multinomial distributions, and their respective components are perfectly correlated.

Considering the large sample case, it follows (Miller, 1966) that the random vector $\hat{e}_i$ has the multivariate normal distribution with mean $e_i = (e_{i1}, e_{i2}, \ldots, e_{ik})^T$ and covariance matrix E consisting of elements in (4.2) and (4.3). So a $100(1-\alpha)\%$ confidence region for e_i is approximately given by the ellipsoid of points e_i's satisfying

$$(\hat{e}_i - e_i)^T \hat{E}^{-1} (\hat{e}_i - e_i) \leq \chi^2_{\alpha,p-1} \tag{4.4}$$

where $\hat{E}$ is an estimate of E obtained by replacing p_{ij}'s by their estimates $\hat{p}_{ij}$'s and $\chi^2_{\alpha,p-1}$ is the $100(1-\alpha)\%$ quantile for χ^2 variate with $(p-1)$ degrees of freedom. Now considering the coordinate-wise projection, this yields simultaneous confidence intervals

$$\hat{e}_{ij} \pm \left[\chi^2_{\alpha,p-1} \sum_{s=1}^{2} P^2(s|i) \frac{\hat{p}_{sj}(1-\hat{p}_{sj})}{n_s} \right]^{1/2} \tag{4.5}$$

for e_{ij}, $j=1,2,\ldots,k$ and $i=1,2$.

4.2. NOT ALL OF μ, ν AND Σ KNOWN

Let $X_1, X_2, \ldots, X_{N_1}$ be the observations for N_1 individuals sampled from π_1 and $Y_1, Y_2, \ldots, Y_{N_2}$ be the observations for N_2 individuals sampled from π_2. The MLE's of the unknown parameters or Δ^2 can now be obtained on the basis of these observations. Two particular cases of interest are (i) μ, ν unknown and Σ known and (ii) μ, ν and Σ all unknown. The MLE of Δ^2 is

$$\hat{\Delta}^2 = (\bar{X} - \bar{Y})^T \Sigma^{-1} (\bar{X} - \bar{Y})$$

in case (i) and

$$\hat{\Delta}^2 = (\bar{X} - \bar{Y})^T S^{-1} (\bar{X} - \bar{Y})$$

in case (ii), where

$$\bar{X} = \frac{1}{N_1} \sum_{1}^{N_1} X_i, \qquad \bar{Y} = \frac{1}{N_2} \sum_{1}^{N_2} Y_i$$

and

$$S = \frac{1}{N_1+N_2-2} \left[\sum_{1}^{N_1} (X_i-\bar{X})(X_i-\bar{X})^T + \sum_{1}^{N_2} (Y_i-\bar{Y})(Y_i-\bar{Y})^T \right] .$$

Use of these estimates for Δ^2 leads to the MLE $\Phi(-\frac{\hat{\Delta}}{2})$ of $\Phi(-\frac{\Delta}{2})$ or of $P(i|j)$, given by

$$\hat{P}(i|j) = \begin{cases} 1 - \Phi(-\frac{\hat{\Delta}}{2}), & i=j \\ \Phi(-\frac{\hat{\Delta}}{2}), & i \neq j \end{cases} \tag{4.6}$$

with i and j = 1,2.

Next, as in Section 4.1, let $\hat{p}_{ij}$ be the observed proportion of jth category objects associated with N_i randomly selected individuals from π_i. Assuming stochastic independence between two types of observations in each class, it follows that the MLE of e_{ij} is given by

$$\hat{e}_{ij} = \sum_{s=1}^{2} \hat{p}_{sj} \hat{P}(s|i), \tag{4.7}$$

$$j=1,2,\ldots,k$$

$$\text{and} \quad i=1,2$$

where $\hat{P}(s|i)$ is given by (4.6). The estimates $\hat{e}_{ij}$'s are consistent, and if N_1 and N_2 individuals from π_1 and π_2 are sampled in large numbers, also leading to large values for associated n_1 and n_2 objects in π_1 and π_2 respectively, the asymptotic distribution of $\hat{e}_i=(\hat{e}_{i1},\hat{e}_{i2},\ldots,\hat{e}_{ik})^T$ can be shown to be multivariate normal. In the present case, e_{ij} is a biased estimate and so for the purpose of finding asymptotic simultaneous confidence intervals for e_{ij}'s, we will use their mean square errors (MSE) instead of their variances.

Recognizing the two estimates $\hat{p}_{sj}$ and $\hat{P}(s|i)$ are independent, it follows that

$$\text{Var}\,(\hat{p}_{sj}\,\hat{P}(s|i)) = [E(\hat{p}_{sj})]^2\,[\text{Var}(\hat{P}(s|i)] + [E(\hat{P}(s|i)]^2\,\text{Var}(\hat{p}_{sj})$$

$$+ \text{Var}(\hat{p}_{sj})\,[\text{Var}(\hat{P}(s|i))].$$

Since $\text{Var}\,(\hat{P}(1|i)) = \text{Var}(\hat{P}(2|i)) = \text{Var}(\Phi(-\frac{\hat{\Delta}}{2}))$, i=1,2, we obtain

$$\text{Var}(\hat{e}_{1j}) = \text{Var}(\hat{p}_{1j}\hat{P}(1|1)) + \text{Var}(\hat{P}_{2j}\hat{P}(2|1)) + 2\,\text{Cov}(\hat{p}_{1j}\hat{P}(1|1),\hat{p}_{2j}\hat{P}(2|1))$$

$$= \left[\frac{3p_{1j}(1-p_{1j})}{n_1} + \frac{3p_{2j}(1-p_{2j})}{n_2} + p_{1j}^2+p_{2j}^2-2p_{1j}p_{2j} \right] \text{Var}(\Phi(-\frac{\hat{\Delta}}{2}))$$

$$+ \left[1-E(\Phi(-\frac{\hat{\Delta}}{2}))\right]^2 \frac{p_{1j}(1-p_{1j})}{n_1} + \left[E(\Phi(-\frac{\hat{\Delta}}{2}))\right]^2 \frac{p_{2j}(1-p_{2j})}{n_2} . \tag{4.8}$$

Next, interchanging subscripts 1 and 2 in (4.8) one can obtain $Var(\hat{e}_{2j})$. Since we are interested in the mean square errors of $\hat{e}_{ij}$'s, we now give similar expressions for these as following:

$$MSE(\hat{e}_{1j}) = \left[\frac{3p_{1j}(1-p_{1j})}{n_1} + \frac{3p_{2j}(1-p_{2j})}{n_2} + (p_{1j}-p_{2j})^2\right] MSE(\Phi(-\frac{\hat{\Delta}}{2}))$$

$$+ (1-\Phi(-\frac{\Delta}{2}))^2 \frac{p_{1j}(1-p_{1j})}{n_1} + (\Phi(-\frac{\Delta}{2}))^2 \frac{p_{2j}(1-p_{2j})}{n_2} \tag{4.9}$$

and

$$MSE(\hat{e}_{2j}) = \left[\frac{3p_{1j}(1-p_{1j})}{n_1} + \frac{3p_{2j}(1-p_{2j})}{n_2} + (p_{1j}-p_{2j})^2\right] MSE(\Phi(-\frac{\hat{\Delta}}{2})$$

$$+ (\Phi(-\frac{\Delta}{2}))^2 \frac{p_{1j}(1-p_{1j})}{n_1} + (1-\Phi(-\frac{\Delta}{2}))^2 \frac{p_{2j}(1-p_{2j})}{n_2} . \tag{4.10}$$

We denote $MSE(\hat{e}_{ij}) = S_{ij}$ and let its estimate s_{ij} be obtained by replacing unknown quantities in (4.9) and (4.10) by their estimates.

Now using the same argument as in the previous case, it follows that approximate $100(1-\alpha)\%$ simultaneous confidence intervals for e_{ij}, $i=1,2$ and $j=1,2,\ldots,k$ are given by

$$\hat{e}_{ij} \pm \left[s_{ij}\chi^2_{\alpha,p-1}\right]^{1/2} \tag{4.11}$$

$i=1,2$ and $j=1,2,\ldots,k$, respectively.

V. DETERMINATION OF SAMPLE SIZE

In this section we consider the problem of determining sample size and, in particular, want to obtain the sample size so that only a specified amount of error for e_{ij}'s is allowed by their estimates with a given probability. In specific terms the problem is to find the size of N_1, N_2 and thereby that of n_1, n_2 so that e_{ij}'s fall simultaneously in intervals given by $e_{ij} \pm \gamma_{ij}$, (γ_{ij} fixed), $i=1,2$ and $j=1,2,\ldots,k$, with probability $(1-\alpha)$. However, this is equivalent to obtaining the sample size when the coordinate-wise length of a $100(1-\alpha)\%$ confidence region for e_i is specified as $2\gamma_i = 2(\gamma_{i1}, \gamma_{i2}, \ldots, \gamma_{ik})^T$, $i=1,2$. Hence, based upon the discussion in Section IV, an asymptotic solution for the sample size is available from equations (4.5) and (4.11) for the two cases considered earlier.

(i) Suppose μ, ν and Σ are known. Then using the result in (4.5), an approximate solution for the sample size is given by the equation

$$\sum_{s=1}^{2} p^2(s|i) \frac{\hat{p}_{sj}(1-\hat{p}_{sj})}{n_s} = \frac{\gamma^2_{ij}}{\chi^2_{\alpha,p-1}}, \qquad i=1,2 \tag{5.1}$$

for any fixed j. Simplifying (5.1) for n_1, n_2, we get

$$n_{1(j)} = \frac{C\,\hat{p}_{1j}(1-\hat{p}_{1j})}{\gamma_{1j}^2[1-\Phi(-\frac{\hat{\Delta}}{2})]^2-\gamma_{2j}^2\Phi^2(-\frac{\hat{\Delta}}{2})} \tag{5.2}$$

and

$$n_{2(j)} = \frac{C\,\hat{p}_{2j}(1-\hat{p}_{2j})}{\gamma_{2j}^2[1-\Phi(-\frac{\hat{\Delta}}{2})]^2-\gamma_{1j}^2\Phi^2(-\frac{\hat{\Delta}}{2})}, \tag{5.3}$$

$$j=1,2,\ldots,k$$

where

$$C = [(1-\Phi(-\frac{\hat{\Delta}}{2}))^2 - \Phi^2(-\frac{\hat{\Delta}}{2})]\,\chi^2_{\alpha,p-1}\,. \tag{5.4}$$

Thus for every j we have a determination for n_1 and n_2 from (5.2) and (5.3), respectively. One way to find a unique solution is to consider n_i given by

$$n_i = \max\ \{n_{i(j)},\quad j=1,2,\ldots,k\},$$

i=1,2. Secondly, by a judicious choice of γ_{ij}, j=1,2,...,k and i=1,2, (5.2) can provide the same value for all $n_{1(j)}$'s and so also (5.3) for all $n_{2(j)}$'s. Then any such common solution, say n_1 and n_2, will be the desired sample size.

(ii) Suppose not all of μ, ν and Σ are known. Again as in the previous case, an approximate determination of sample size is given by a solution of the equation

$$s_{ij}\chi^2_{\alpha,p-1} = \gamma^2_{ij},\quad i=1,2 \tag{5.5}$$

for any fixed j. Denoting

$$v = \text{estimate of MSE}(\Phi(-\frac{\hat{\Delta}}{2})),$$

it follows from (4.9), (4.10) and (5.5) that

$$[3v + (1-\Phi(-\frac{\hat{\Delta}}{2}))^2]\frac{\hat{p}_{1j}(1-\hat{p}_{1j})}{n_1} + [3v + \Phi^2(-\frac{\hat{\Delta}}{2})]\ \frac{\hat{p}_{2j}(1-\hat{p}_{2j})}{n_2}$$

$$= (\hat{p}_{1j} - \hat{p}_{2j})^2\ v + \frac{\gamma^2_{1j}}{\chi^2_{\alpha,p-1}}$$

and

$$[3v + \Phi^2(-\frac{\hat{\Delta}}{2})]\ \frac{\hat{p}_{1j}(1-\hat{p}_{1j})}{n_1} + [3v+(1-\Phi(-\frac{\hat{\Delta}}{2})^2]\ \frac{\hat{p}_{2j}(1-\hat{p}_{2j})}{n_2}$$

$$= (\hat{p}_{1j} - \hat{p}_{2j})^2\ v + \frac{\gamma^2_{2j}}{\chi^2_{\alpha,p-1}}\ .$$

Simplifying these equations for n_1 and n_2 for a fixed j, we get

$$n_{1(j)} = \frac{(a^2-b^2)\ \hat{p}_{1j}\ (1-\hat{p}_{1j})\ \chi^2_{\alpha,p-1}}{(\hat{p}_{1j}-\hat{p}_{2j})^2\ (1-2\,\Phi(-\frac{\hat{\Delta}}{2}))\,v\chi^2_{\alpha,p-1} + a\gamma^2_{1j} - b\gamma^2_{2j}} \tag{5.6}$$

and

$$n_{2(j)} = \frac{(a^2-b^2)\ \hat{p}_{2j}\ (1-\hat{p}_{2j})\ \chi^2_{\alpha,p-1}}{(\hat{p}_{1j}-\hat{p}_{2j})^2\ (1-2\Phi(-\frac{\hat{\Delta}}{2}))v\chi^2_{\alpha,p-1} + a\gamma^2_{2j} - b\gamma^2_{1j}} \tag{5.7}$$

where

$$a = 3v + (1-\Phi(-\frac{\hat{\Delta}}{2}))^2$$

$$b = 3v + \Phi^2(-\frac{\Delta}{2}) \ .$$

Once again, a unique determination of sample size is obtained either by taking

$$n_i = \max\{n_{i(j)},\quad j=1,2,\ldots,k\},$$

$i=1,2$ or by obtaining a common solution for $n_{i(j)}$, $j=1,2,\ldots,k$ with a judicious choice of γ_{ij}, $j=1,2,\ldots,k$ in (5.6) and (5.7)

The solution for $n_{i(j)}$ in (5.6) and (5.7) depends upon the knowledge of v, an estimate of $MSE(\Phi(-\frac{\hat{\Delta}}{2}))$. The later can be evaluated by considering it a function of the statistic $\hat{\Delta}^2$, which is distributed as non-central χ^2 with p degrees of freedom and non-centrality parameter Δ^2 when Σ is known and asymptotically so if Σ is unknown. Since we are dealing with the large sample case, in both cases we will have the same evaluation given by

$$MSE(\Phi(-\frac{\hat{\Delta}}{2})) = E[\Phi^2(-\frac{\hat{\Delta}}{2})] - 2\Phi(-\frac{\Delta}{2})E[\Phi(-\frac{\hat{\Delta}}{2})] + \Phi^2(-\frac{\Delta}{2}) \tag{5.8}$$

where

$$E[\Phi(-\frac{\hat{\Delta}}{2})] = \frac{1}{\sqrt{2\pi}} \int_0^\infty \left(\int_{-\infty}^{-\sqrt{Q}/2} e^{-y^2/2}\, dy \right) f(Q;\Delta^2)\, dQ$$

and

$$E[\Phi^2(-\frac{\hat{\Delta}}{2})] = \frac{1}{2\pi} \int_0^\infty \left(\int_{-\infty}^{-\sqrt{Q}/2} e^{-y^2/2}\, dy \right)^2 f(Q;\Delta^2)\, dQ$$

with

$$f(Q;\Delta^2) = \sum_{\alpha=0}^{\infty} \frac{e^{-\Delta^2/2}(\Delta^2)^\alpha}{\alpha!} \frac{Q^{P/2+\alpha-1}\ e^{-Q/2}}{2^{P/2+\alpha}\Gamma(\frac{P}{2}+\alpha)} \ .$$

Now for an evaluation of v, it is given by (5.8) with Δ replaced by $\hat{\Delta}$, and for any given $\hat{\Delta}$, the quantities $E[\Phi(-\frac{\hat{\Delta}}{2})]$ and $E[\Phi^2(-\frac{\hat{\Delta}}{2})]$ can easily be computed by a numerical integration technique and $\Phi(-\frac{\Delta}{2})$ is available from the standard normal table.

VI. CONCLUDING REMARKS

The sampling procedure and the other discussion given in this paper are of such a general nature that our results can be easily applicable or adapted in several different situations that may arise in remote sensing. For the specific examples mentioned in the beginning of the paper, one can now see that once samples are obtained from different classes of trees by a suitable sampling technique and subsequently the ground truth regarding vegetation and flora, etc. associated with the sampled trees is ascertained, the above discussion will provide estimates for the amounts of different types of vegetations and flora in a forest covering a large area. Also, a similar approach will lead to estimating contributions of various categories of objects making up a resolution element for remote sensing data in case the instantaneous field of view of a multispectral scanning device covers more than one type of objects. Exclusively addressed to this last situation another approach to the problem (Horwitz, Nalepka, Hyde and Morgenstern, 1971), has been suggested earlier. It consists of estimating the expected mean values and expected covariance matrix, assuming the mean vectors and dispersion matrices for individual categories of objects, and then considering class distributions as multivariate normal with these expected mean vectors and expected covariance matrix. Another situation that falls in the present framework is where an acreage estimate for different types of soils under various crops in a region is desired.

The computations involved are straight-forward except in the evaluation of v which involves numerical integration. In the absence of real data a sampling study can be conducted easily using simulated data obtained for actual proportion vectors associated with different classes and for the classes as well if so needed.

It may as well be pointed out that our determination of sample size was based upon the requirement of achieving certain specified precision by the estimates. If, however, it is desired to have a given fraction of the objects covered with a desired level of confidence, the results will be more interesting. This, of course, would require deriving tolerance limits on e_{ij}'s, and then by a similar procedure as in the fixed precision case, a sample size can be determined so that a desired amount of coverage of objects with a given level of confidence is insured. The authors hope to give results with this approach in a subsequent paper.

VII. REFERENCES

1. Anderson, T. W., An Introduction to Multivariate Statistical Analysis, John Wiley and Sons, 1958.

2. Cochran, W., "Errors of Measurements in Statistics," Technometrics, 1968, Vol. 10, No. 4, pp. 637-666.

3. Hortwitz, H. M., Nalepka, R. F., Hyde, P.D. and Morgenstern, J. P., "Estimating the Proportions of Objects Within a Single Resolution Element of a Multispectral Scanner," Proceedings of the Seventh International Symposium On Remote Sensing of Environment, May, 1971.

4. Miller, R. G., Jr., Simultaneous Statistical Inference, McGraw-Hill, 1966.