PROBLEMS IN INVERSE SAMPLING

OF BULK MATERIALS

by

Richard L. Scheaffer

FSU Statistics Report ML30

December, 1967
Department of Statistics
Florida State University
Tallahassee, Florida

Work supported by the National Institute of General Medical Sciences, Training Grant Number 5 T01 GM913 and by the National Science Foundation grant number GP 6359.
Abstract

It is assumed that a population consists of \( r \) (\( r < \infty \)) types of items, each item being characterized by a real-valued, non-negative random variable called the item's amount. Each type of item may have a different distribution of amounts. It is also assumed that samples are selected from this population so as to contain a fixed total amount rather than a fixed number of items. The measurement of the degree of mixing, the estimation of amount proportions, and the estimation of the average item amount are three problems for which answers are found by the use of methods and theorems from renewal theory. Some results which have already appeared in the literature are derived and extended in the course of this work.
1. Introduction

It is assumed that a population is composed of \( r (r < \infty) \) types of items with each item characterized by a real-valued, non-negative random variable \( X \). The value of this random variable for a particular item will be called the item's amount, where amount may denote weight, volume, monetary value or any similar characteristic. This paper will be concerned with sampling schemes in which the sample is composed of a fixed total amount rather than a fixed number of items.

This concept of sampling to an amount occurs in many bulk sampling problems and, in bulk sampling, three specific problems seem to present themselves frequently. These problems are the measurement of the degree to which items of different types are mixed in the population, the estimation of the proportion of the total population amount due to items of a certain type, and the measurement of the average item amount, or fineness. This paper will show that the questions raised above can be answered in part by the use of renewal theory.

The motivation for this work came primarily from three sources, the papers by Becker, [3], [4], [5] and [6], on probability distributions applicable to the mineral industries, the book by Herdan [16], and the paper by Buslik [8] on the sampling of granular material. Becker finds the first three moments of the sample amount for a particular type of item, under some rather restrictive assumptions. The book by Herdan gives much background material on the problems and methods of sampling small particles.
Buslik derives an asymptotic variance formula which is the same as expression (7.6) in this paper. The derivation in this paper makes use of a different method and includes covariances as well as asymptotic distributions, neither of which is considered by Buslik.

Recently a paper was published by Knott [10] in which he also considers the bulk sampling problems from the point of view of renewal theory. However, his methods seem to be somewhat more complicated than the methods used in this paper and he does not consider asymptotic distributions. Knott's asymptotic variance is the same as the one given in expression (3.12) below, except that he uses \( \{t\} \), denoting the least integer not less than \( t \), in place of \( t \). Knott's asymptotic covariance formula differs somewhat from expression (3.15) below, but this difference may be due to a typographical error in his paper.

2. The Model

Consider a population in which each item is characterized by a random variable \( X \), called the item amount. Assume that the population items are of \( r \) types with the distribution function of the amounts for type \( i \) items being given by \( F_i(\cdot) \), \( i=1,\ldots,r \), \( r<\infty \), where \( F_i(0)=0 \) for each \( i \). The unconditional distribution function for \( X \) is given by

\[
P(x) = \sum_{i=1}^{r} p_i F_i(x),
\]

(2.1)

with \( 0 < p_i < 1 \) and \( \sum_{i=1}^{r} p_i = 1 \), where \( p_i \) is the
probability of randomly selecting an item of type \( i \) from the population.

Suppose that one wishes to select a sample of total amount \( t \). One can form a model for this situation by assuming that the items are selected one at a time with the amounts being accumulated so that after the \( i \)th draw one has a value for \( S_i = x_1 + \ldots + x_i \), with \( S_0 = 0 \). The sampling is stopped the first time \( S_r \) is greater than or equal to \( t \). Assuming that the items are sampled independently, the sequence of sampled item amounts, \( x_i \), forms a renewal process (see [11]).

The random variable \( N(t) \) is defined as the largest integer \( k \) such that \( S_k \) is less than or equal to \( t \). This \( N(t) \) will be referred to as the number of items in the sample, with the understanding that the actual number may be \( N(t) + 1 \). Since this paper will be concerned with asymptotic results, a difference of one item will be of little consequence. Also, \( t \) will be referred to as the sample amount with the understanding that the actual sample amount will usually differ from \( t \), but that difference will be smaller than the largest possible item amount. Since in many bulk sampling problems the sample amount is large compared to the item amounts, the difference between \( t \) and the actual sample amount should be small relative to \( t \).
3. The Sample Amount Proportions

Let the assumptions of section 2 be satisfied for a bulk sampling problem, and define a set of functions as follows:

\[ W^{(j)}(x_i) \equiv W^{(j)}_i = x_i, \text{ if } x_i \text{ is of type } j \]  \quad (3.1)

\[ = 0 \text{ otherwise,} \]

for \( j=1,\ldots, r, r < \infty \), where \( x_i \) is the random variable defined in section 2 with distribution function given by \( (2.1) \).

Let \( b_i \) and \( b^{(2)}_i \) be the first and second moments, respectively, of the distribution function \( F_i(.) \) and assume throughout this paper that these first two moments exist and are finite for \( i=1,\ldots, r \).

Define

\[ z^{(j)}_t = \sum_{i=1}^{N(t)} W^{(j)}_i, \text{ for } N(t)=1,2,\ldots, \]  \quad (3.2)

\[ = 0, \text{ for } N(t) = 0, \]

for \( j=1,2,\ldots, r \). Then it follows that

\[ \hat{B}_j = t^{-1} z^{(j)}_t = \text{(the total amount due to items of type } j \text{ in the sample)}/\text{(the total amount of the sample)}. \]  \quad (3.3)

The quantities \( \hat{B}_j, j=1,\ldots, r, \) will be referred to as the sample amount proportions.
The following quantities are easily obtained:

\[ E(x_i) = \sum_{j=1}^{r} p_j b_j = b, \quad (3.4) \]

\[ E(N_i^{(j)}) = p_j b_j, \quad (3.5) \]

\[ \text{Var}(x_i) = b^{(2)} - b^2, \quad \text{where} \quad b^{(2)} = \sum_{i=1}^{r} b^{(2)}_i p_i, \quad (3.6) \]

\[ \text{Var}(N_i^{(j)}) = p_j b^{(2)}_j - p_j^2 b_j^2, \quad (3.7) \]

\[ \text{Cov}(N_i^{(j)}, x_i) = p_j b^{(2)}_j - p_j b_j b_i, \quad \text{and} \]

\[ \text{Cov}(N_i^{(j)}, N_i^{(k)}) = -p_j b_j b_i b_k, \quad (3.9) \]

**Theorem 3.1**

Let \( B_i = (p_i b_i)/b \). Then

(i) \( \lim_{t \to \infty} \hat{B}_j = B_j \) almost surely, \( (3.10) \)

(ii) \( \lim_{t \to \infty} E(\hat{B}_j) = B_j \) and \( (3.11) \)

(iii) \( \lim_{t \to \infty} \text{Var}(\hat{B}_j) = B_j (1-B_j) \left( b^{(2)}_j + \sum_{i=1}^{r} \frac{b_i^{(2)}}{b_i} \right) \quad (3.12) \)

Proof:

(i) \( \hat{B}_j = (N(t)/t) \sum_{i=1}^{n(t)} W_i^{(j)}/N(t) + B_j \) almost surely, since \( \lim_{n \to \infty} \sum_{i=1}^{n} W_i^{(j)} = p_j b_j \) almost surely by the strong law of large numbers, \( N(t)/t \to 1/b \) almost surely, and \( N(t) \to \infty \) almost surely, all as \( t \to \infty \).
(ii) $E \hat{B}_j = \frac{1}{t} \sum_{i=1}^{\hat{N}(t)} \hat{W}_i = \hat{B}_j [1/b] + o(1)$

either by the theory of cumulative processes, Smith [23], or
by proceeding directly using Wald's equation from sequential
analysis, Johnson [17].

(iii) $t \text{Var}(\hat{B}_j) = t^{-1} \text{Var} \sum_{i=1}^{\hat{N}(t)} \hat{W}_i = (1/b) \text{Var} [\hat{W}_i - B_j \hat{Y}_i] + o(1)$,

again by the theory of cumulative processes or by theorems
on moments of random sums, Chow, Robbins and Teicher [9]. Now

$$\frac{1}{b} \text{Var} [\hat{W}_i - B_j \hat{Y}_i] = B_j (1 - B_j) \frac{b_i(2)}{b_j} + B_j^2 \left[ \sum_{i=1}^{r} \frac{b_i(2)}{b_i} \right]$$

(3.13)

by (3.6), (3.7), and (3.8), and the proof is completed.

Theorem 3.2

The random vector $(\hat{B}_1, ..., \hat{B}_r)$ is asymptotically, as
t$\to \infty$, distributed as an $r$-dimensional multivariate normal
random vector with mean vector $(B_1, ..., B_r)$,

$$\text{Var}^*(t \hat{B}_j) = \text{expression (3.13)}$$

(3.14)

and

$$\text{Cov}^*(t \hat{B}_j, t \hat{B}_k) = B_j B_k \left[ \sum_{i=1}^{r} \frac{b_i(2)}{b_i} \right]$$

(3.15)

where $\text{Var}^*(.)$ and $\text{Cov}^*(.)$ denote the variance and
covariance, respectively, of the asymptotic distribution.
Proof:

From a generalization, given by Smith [23], of a theorem by Anscombe [3], one obtains the result that the quantities

\[ t^{-\frac{1}{2}} \sum_{i=1}^{N(t)} (W_i^{(j)} - B_j X_i), \quad j=1, \ldots, r, \]

are jointly asymptotically distributed as a multivariate normal random vector with covariance terms given by

\[ (1/\nu) \text{Cov} [(W_i^{(j)} - B_j X_i), (W_i^{(k)} - B_k X_i)]. \quad (3.16) \]

This result follows since \( E[W_i^{(j)} - B_j X_i] = 0 \) and \( N(t)/t \rightarrow 1/\nu \) almost surely as \( t \rightarrow \infty \). Equations (3.14) and (3.15) are obtained from (3.16) by use of expressions (3.6) through (3.9).

Now,

\[ t^{-\frac{1}{2}} \sum_{i=1}^{N(t)} (W_i^{(j)} - B_j X_i) = t^{-\frac{1}{2}} \left[ \sum_{i=1}^{N(t)} W_i^{(j)} - B_j t + B_j (t - S_N(t)) \right]. \]

Also,

\[ 0 \leq t^{-\frac{1}{2}} B_j (t - S_N(t)) \leq t^{-\frac{1}{2}} B_j X_N(t) + 1 = \frac{B_j X_{N(t)} + 1}{(N(t) + 1)^{\frac{1}{2}}} \left( \frac{N(t) + 1}{t} \right)^{\frac{1}{2}} \rightarrow 0 \]

almost surely as \( t \rightarrow \infty \). Thus, the proof is completed.

Equations (3.10) and (3.11) show that \( \hat{B}_j \) is a good estimator of \( B_j \) for \( j=1, \ldots, r \) and these quantities \( \hat{B}_j \) are usually available in a bulk sampling problem. Then, the asymptotic variances and covariances of the sample amount proportions can be estimated if the ratio \( b_i^{(2)}/b_i \) can be
obtained, or estimated, for \( i=1, \ldots, r \). The estimation of the parameters of \( F_i(.) \) will be discussed in section 7 for some special cases.

4. The Problem of Mixing

Consider a bulk sampling problem in which the assumptions of section 2 are satisfied. One of the important problems of bulk sampling is the problem of deciding whether or not the different types of items are mixed in the population. A reasonable measure of the degree of mixing can be obtained as follows. Suppose \( n \) independent samples, each of amount \( t \), are taken from the population with each sample being a cluster of items from one segment of the population. The word "cluster" is used to mean that the sample is selected in one composite draw of contiguous items. Then, one can compare each \( \hat{B}_i \), the sample amount proportion for items of type \( i \), with its asymptotic expectation \( B_i \). Each \( \hat{B}_i \) should be close to \( B_i \), for \( i=1, \ldots, r \), if the population is thoroughly mixed and if the sample amount is large compared to the item amounts.

The asymptotic joint distribution of \( (\hat{B}_1, \ldots, \hat{B}_r) \) is given in theorem 3.2, but the covariance matrix of this asymptotic distribution will be singular. Let

\[
\hat{B}^{(i)} = (\hat{B}_1^{(i)}, \ldots, \hat{B}_{r-1}^{(i)}) \quad \text{where} \quad \hat{B}_j^{(i)} \text{ is the sample amount proportion for items of type } j \text{ in the } i\text{th sample},
\]
\[ \hat{\beta} = (\hat{\beta}_1, \ldots, \hat{\beta}_{r-1}) \quad \text{and} \]
\[ \hat{\bar{\beta}} = (\bar{\beta}_1, \ldots, \bar{\beta}_{r-1}) \quad \text{where} \]
\[ \bar{\beta}_j = \frac{1}{n} \sum_{i=1}^{n} \hat{\beta}^{(i)}_j. \]

It follows from theorem 3.2 that \( \hat{\beta}^{(i)} \) is asymptotically distributed as an \((r-1)\)-dimensional multivariate normal random variable with mean vector \( \bar{\beta} \) and a non-singular covariance matrix which will be denoted by \( \frac{1}{\xi} \Sigma \), under the null hypothesis of a thoroughly mixed population.

Define an index of mixing, \( M \), as follows,
\[ M = \frac{1}{n} \sum_{i=1}^{n} (\hat{\beta}^{(i)} - \bar{\beta})' \Sigma^{-1} (\hat{\beta}^{(i)} - \bar{\beta})'. \quad (4.1) \]

Under the null hypothesis that the items in the population are thoroughly mixed, it follows from well known results in multivariate analysis that \( M \) has asymptotically, as \( t \to \infty \), a \( \chi^2 \) distribution with \( n(r-1) \) degrees of freedom, denoted by \( \chi^2_{n(r-1)} \). \( M \) should be close to its mean value for a well-mixed population and much larger than its mean for a poorly mixed population.

The statistic \( M \) assumes that \( \bar{\beta} \) and \( \Sigma \) are known, which often is not the case. If \( \bar{\beta} \) is not known, it can be estimated from the \( n \) samples by \( \hat{\beta} \), which is the maximum likelihood estimate using the limiting distribution of \( \hat{\beta}^{(i)} \) under the null hypothesis. These same \( \hat{\beta}_i \)'s can be used to estimate all of the \( \beta_i \) terms in \( \Sigma \), and such an estimate will be denoted by \( \hat{\Sigma} \).
Let

\[ \hat{M} = \sum_{i=1}^{n} (\hat{B}(i) - \bar{B}) \cdot \left( \hat{V}^{-1} (\hat{B}(i) - \bar{B}) \right)' \]  

(4.2)

**Theorem 4.1**

Under the null hypothesis of a thoroughly mixed population, \( \hat{M} \) has asymptotically, as \( t \to \infty \), a \( \chi^2(n-1)(r-1) \) distribution.

**Proof:**

It follows from well known results in multivariate analysis that the quantity

\[ \bar{M} = \sum_{i=1}^{n} (\hat{B}(i) - \bar{B}) \cdot \left( \hat{V}^{-1} (\hat{B}(i) - \bar{B}) \right)' \]

has asymptotically a \( \chi^2(n-1)(r-1) \) distribution, under the null hypothesis.

Now it will be shown that \( (\hat{M} - \bar{M}) \to 0 \) in probability as \( t \to \infty \), and, hence, that \( \hat{M} \) and \( \bar{M} \) have the same asymptotic distribution. Let \( v_{jk} \) denote the jkth term of \( (\hat{V}^{-1} - \bar{V}^{-1}) \). Since \( B_i + B_i \) almost surely it follows that \( v_{jk} \to 0 \) almost surely and, hence, in probability, as \( t \to \infty \), for every pair \( j, k \). This may be written as \( v_{jk} = o_p(1) \).

Now

\[ P[t^b | \hat{B}_j - E(\hat{B}_j) | > \lambda] \leq \frac{(t/\lambda^2)}{(1/\lambda^2) \sum_j \text{Var} \hat{B}_j = (1/\lambda^2) [B^* + o(1)]}, \]

where \( 0 \leq B^* < \infty \), and
\( P\left[ t^2 | \hat{B}_j - E(\hat{B}_j) | < \lambda \text{ for all } j \right] > 1 - (1/\lambda^2) [B^* + o(1)]. \)

If one chooses \( \lambda \) so that \( \lambda \to \infty \) as \( t \to \infty \), then the probability (4.3) tends to one as \( t \to \infty \). Hence
\[ t^2 [\hat{B}_j - E(\hat{B}_j)] = O_P(1), \text{ for } j = 1, \ldots, r, \]
and it follows that
\[ t^2 (\hat{\beta}(i)_j - \bar{\beta}_j) = O_P(1) \text{ for } j = 1, \ldots, r. \]
Therefore one has that
\[ t (\hat{\beta}(i)_j - \bar{\beta}_j)(\hat{\beta}(i)_k - \bar{\beta}_k) v_{jk} = O_P(1) \text{ o}_P(1) = o_P(1), \]
and
\( (\hat{\lambda} - \bar{\lambda}) + 0 \) in probability, which completes the proof.

5. The Estimation of Population Amount Proportions

In many bulk sampling problems one wants to know what proportion of the total amount of the population is due to a specific type of item. For example, in a population of ore and rock one may want to estimate the fraction of the total weight due to ore. It will be seen that the sample amount proportion is not always a good estimator of the population amount proportion.

Consider a bulk sampling problem which satisfies the conditions of section 2. The probabilities \( p_i \) and \( F_i(.) \) were defined in general terms in previous sections, but now they will be given more specific interpretations based on frequencies.
Define

\[ F_i^*(x) = \text{(the number of items of type } i \text{ which have amounts less than or equal to } x \text{ in the population)}/(\text{the number of items of type } i \text{ in the population)}, \]

assuming the number of items in the population to be finite but very large, and

\[ q_i = \text{(the number of items of type } i \text{ in the population)}/(\text{the number of items in the population)}, \]

with \( q_i > 0, \ i=1, \ldots, r. \)

If the probability of selecting an item from the population were just equal to the item's frequency proportion, then one would have that

\[ P[X_j \in (x, x+dx), X_j \text{ is of type } i] = q_i dF_i^*(x). \]

But in many bulk sampling problems the probability of selecting an item of a certain amount is proportional to that amount. This occurs in situations in which the sampling mechanism tends to select the larger items more often than the smaller ones, or vice-versa. Thus, a more realistic model to investigate is given by

\[ P[X_j \in (x, x+dx), X_j \text{ is of type } i] = cx^\alpha q_i dF_i^*(x), \quad (5.2) \]

where \( c \) and \( \alpha \) are real-valued constants with \( |\alpha| < \infty \) and \( i=1, \ldots, r. \)

The constant \( c \) can be determined by the identity

\[ \sum_{i=1}^{r} \int_{0}^{\infty} c q_i x^\alpha dF_i^*(x) = 1, \quad \text{which gives} \]

\[ c = \frac{1}{\sum_{i=1}^{r} \int_{0}^{\infty} q_i x^\alpha dF_i^*(x)}. \]
\[ c \cdot 1 = \sum_{i=1}^{r} \alpha_i \beta_i(a) \text{ where } \beta_i(a) = \int_0^{\infty} x^a dF_i(x), \]

assuming these moments exist.

Also, one has that

\[ P[X_i \text{ is of type } j] = \alpha_i \beta_j(a) \]

The model given by (5.2) fits into the renewal process model of section 2 with

\[ P_i = \frac{\alpha_i \beta_i(a)}{\sum_{j=1}^{r} \alpha_j \beta_j(a)} \quad \text{and} \quad \alpha F_i(x) = \frac{x^a dF_i(x)}{\beta_i(a)}, \quad (5.3) \]

Now define

\[ N_i(t) = \text{the number of items of type } i \text{ in the sample of amount } t, \quad (5.4) \]

so that

\[ \sum_{i=1}^{r} N_i(t) = N(t). \]

The quantities \( P_i, i=1, \ldots, r, \) can be estimated by

\[ \hat{P}_i = \frac{N_i(t)}{N(t)}, \quad i=1, \ldots, r \text{ and } N(t) > 0, \quad (5.5) \]

and these estimators, in turn, give estimators for the \( \alpha_i \)'s, denoted by \( \hat{\alpha}_i \), as follows:

\[ \hat{\alpha}_i = \frac{\hat{P}_i \beta_i(a)}{\sum_{j=1}^{r} \hat{\alpha}_j \beta_j(a)}. \]
\[ \sum_{j=1}^{r} \beta_j^{(a)} \beta_j = \sum_{i=1}^{r} \beta_i^{(a)} \beta_i \text{ and, under the } \\
\text{restriction } \sum_{i=1}^{r} \alpha_i = 1, \]
\[ \alpha_i = \frac{\alpha_i^{(a)}}{\beta_i^{(a)}} \left[ \frac{r}{\sum_{j=1}^{r} \beta_j^{(a)}} \right]^{-1}, \text{ } i=1,\ldots,r. \quad (5.6) \]

In order to work with the quantities \( \alpha_i^{(a)}, i=1,\ldots,r, \) one must have some information on the distribution of the quantities \( \beta_i^{(a)}, i=1,\ldots,r, \) and this information is contained in the following theorem.

**Theorem 5.1**

Under the assumptions of section 2,

(i) \( \lim_{t \to \infty} t^{-1} \beta_i^{(a)} = (p_i/\rho) \) almost surely,

(ii) \( \lim_{t \to \infty} t^{-1} E[\beta_i^{(a)}] = (p_i/\rho), \)

(iii) \( \lim_{t \to \infty} t^{-1} \text{Var}[\beta_i^{(a)}] = \frac{1}{2} \left[ \left( \frac{p_i}{\rho} \right)^2 \sum_{j=1}^{r} \beta_j^{(2)} \rho_i \rho_j + \rho_i \left( 1 - 2 \frac{p_i b_i}{\rho} \right) \right], \quad (5.7) \)

(iv) The random vector \( t^{-b/2} (\beta_1^{(a)}, \ldots, \beta_r^{(a)}) \) is asymptotically, as \( t \to \infty, \) distributed as an \( r \)-dimensional multivariate normal random vector with mean \( t^{-b/2} (p_1, \ldots, p_r), \)

\[ \text{Var}^* [t^{-b/2} \beta_i^{(a)}] = \text{expression (5.7)}, \]

and
\[
\text{Cov}^*(t^{-k} \bar{N}_1(t), t^{-k} \bar{N}_j(t)) = \frac{D_i D_j}{b^2} \left[ 1 \sum_{k=1}^{r} \sum_{b} \left( \frac{D_k}{b} \right) \left( b_i - b - b_j \right) \right].
\]

**Proof:**

Define

\[\omega^j(x_i) = \omega^j_i = 1 \text{ if } x_i \text{ is of type } j \]

\[= 0 \text{ otherwise.}\]

Then

\[N_j(t) = \sum_{i=1}^{N(t)} \omega^j_i \text{ for } N(t) = 1, 2, \ldots\]

\[= 0 \text{ otherwise.}\]

The proof of parts (i), (ii) and (iii) is analogous to the proof of theorem 3.1 and the proof of part (iv) is analogous to that of theorem 3.2.

**Theorem 5.2**

Under the assumptions of section 2,

\[\lim_{t \to \infty} \hat{q}_i = q_i \text{ almost surely for } i=1, \ldots, r.\]

**Proof:**

From equations (5.6) and (5.3) and theorem 5.1 it follows that

\[\hat{q}_i = \frac{N_i(t)}{t \beta_i} \left[ \frac{r}{\Sigma_{j=1}^{r} \frac{N_j(t)}{t \beta_j}} \right]^{-1} \text{ a.s. } \frac{D_i}{\beta_i} \left[ \frac{r}{\Sigma_{j=1}^{r} \frac{D_j}{\beta_j}} \right]^{-1} = q_i,\]

and the proof is complete.
From the definitions given above for \( \Gamma_i^*(x) \) and \( q_i \) one can see that \( \beta_i^{(1)} = \beta_i \) can be thought of as

(the total amount due to type \( i \) items in the population)/(the number of type \( i \) items in the population)

so that

\[
\pi_i \equiv \frac{\beta_i q_i}{\sum_{j=1}^{r} \beta_j q_j} = \text{(the total amount due to type \( i \) items in the population)/(the total amount of the population)}.
\]

The quantities \( \pi_i, i=1,\ldots, r \), will be referred to as population amount proportions.

One can use the quantities

\[
\hat{\pi}_i = \frac{\beta_i \hat{q}_i}{\sum_{j=1}^{r} \beta_j \hat{q}_j}, \quad i=1,\ldots, r,
\]

as reasonable estimators of the \( \pi_i \)'s. From theorem 5.2 it is obvious that \( \hat{\pi}_i \) converges to \( \pi_i \) almost surely as \( t \to \infty \). These quantities \( \hat{\pi}_i, i=1,\ldots, r \), are ratios of random variables, and the following theorem is useful for obtaining their asymptotic distribution. The statement of the theorem 5.3 is taken from Anderson [1], but the theorem is essentially that of Cramer [12].
Theorem 5.3

Let \( Y(n) \) be an \( m \)-component random vector and \( \zeta \) a fixed vector. Assume \( n^{1/2} [Y(n) - \zeta] \) is asymptotically distributed according to \( N(0, \Sigma) \). Let \( f(Y) \) be a function of a vector \( Y \) with first and second derivatives existing in a neighborhood of \( Y = \zeta \). Let the partial derivative of \( f(Y) \) with respect to \( Y_i \), evaluated at the point \( Y = \zeta \), be the \( i \)th component of \( \Phi_\zeta \). Then the limiting distribution of

\[ n^{1/2} [f(Y(n)) - f(\zeta)] \]

is \( N(0, \Phi_\zeta \Sigma \Phi_\zeta^T) \).

Let \( K_i = \beta_i / \beta (\alpha) \) for \( i = 1, ..., r \), so that

\[ f_i = K_i Y_i(t) \left[ \sum_{j=1}^{r} K_j N_j(t) \right]^{-1} \]

The asymptotic distribution of \( f_i \) will now be found.

Let \( C_i = \beta_i / b \) for \( i = 1, ..., r \). It then follows from theorem 5.1 that

\[
t^{-1} \text{Var}^*[K_i Y_i(t)] = b^{-1}\left[ \sum_{k=1}^{r} [b(2)_{k p_k}] \sum_{j=1}^{r} C_k K_i^2 \right] + \sum_{j=1}^{r} K_i^2 \beta_j^2 - 2 \sum_{j=1}^{r} C_k K_j \sum_{j=1}^{r} K_i b_j \beta_j \] (5.11)

and

\[
t^{-1} \text{Cov}^*[K_i Y_i(t), \sum_{j=1}^{r} K_j Y_j(t)] = \]

\[
b^{-1}\left[ \sum_{k=1}^{r} [b(2)_{k p_k}] \sum_{j=1}^{r} C_k K_j K_i C_j + K_i^2 \beta_i^2 \right] - C_i K_i \sum_{j=1}^{r} [K_j^2 b_j \beta_j] - K_i^2 b_i \beta_i \sum_{j=1}^{r} C_j K_j \] (5.12)
Using theorem 5.3, one then sees that

\[ t \text{Var}^*(\hat{\alpha}_i) = b^{-1}_i \left( D^{-2} \sum_{k_i}^2 \nu_i + D^{-3} (C_i K_i)^2 \left( \sum_{j=1}^r K_j \rho_j \right) \right) \tag{5.13} \]

\[ -2 b^{-3} (C_i K_i) K_i^2 \nu_i \}, \text{ where} \]

\[ D = \sum_{j=1}^r C_i K_i \]

and

\[ E^*(\hat{\alpha}_i) = \pi_i. \]

These results are summarized in the following theorem.

**Theorem 5.4**

Under the assumptions of section 2, \( t \frac{1}{b} \hat{\alpha}_i \) is asymptotically, as \( t \to \infty \), normally distributed with mean \( t \frac{1}{b} \pi_i \) and variance given by expression (5.13), for \( i=1, \ldots, r \).

Consider the quantity \( \pi_i \) for the case \( \alpha=0 \), which means that the probability of selecting an item of a certain amount from the population is equal to the item's population frequency proportion. From equation (5.3) one sees that \( \alpha = 0 \) implies that \( \rho_i = q_i \) and \( F_i(.) = F_i^* (.) \). It then follows that

\[ \pi_i = B_i, \text{ for } i=1, \ldots, r. \]

From theorem 3.1 one has that \( \hat{B_i} \) is a good estimator of \( B_i \). Therefore, for \( \alpha=0 \), one could use the sample amount proportion \( \hat{B_i} \) as a consistent, asymptotically unbiased estimator of the population amount proportion \( \pi_i \).
Now consider the estimator \( \hat{n}_i \) for the case \( \alpha = 1 \).

In this case

\[
\hat{n}_i = \frac{n_i(t)}{N(t)}, \text{ for } i = 1, \ldots, r.
\]

That is, a good estimator of the population amount proportion, \( \pi_i \), for items of type \( i \) is the sample frequency proportion for type \( i \) items. Also, for \( \alpha = 1 \),

\[
\hat{p}_i = \beta_i q_i \left( \sum_{j=1}^{r} \beta_j q_j \right)^{-1} = \pi_i
\]

so that the probability of selecting an item of type \( i \) is equal to the population amount proportion for type \( i \) items.

From equation (5.13) it follows that, for \( \alpha = 1 \),

\[
\text{Var}^*(\hat{n}_i) = \frac{b}{c} p_i (1-p_i), \text{ for } i = 1, \ldots, r.
\]

5. The Problem of Fineness

In the problem of measuring the average item amount, or fineness, of the population one could work with many different probability distributions which have previously been discussed in this paper. Two such distributions are the \( \beta_i \)'s defined in section 3 and the \( p_i \)'s introduced in section 2. However, a reasonable measure of fineness is the unweighted average amount of all items in the population. This average will be denoted by \( \mu_x \) and defined by

\[
\mu_x = \sum_{i=1}^{r} \beta_i q_i.
\]
The parameter $\mu_x$ can be estimated by

$$\hat{\mu}_x = \sum_{i=1}^{r} \beta_i \tilde{c}_i.$$  \hfill (5.2)

It is convenient to define $\mu_x^{-1}$ as the index of fineness (Herdan [16] p. 39) and to use $\hat{\mu}_x^{-1}$ as the estimator of $\mu_x^{-1}$.

Let the assumptions of section 2 be fulfilled for a bulk sampling problem. With $\tilde{c}_i$ given by (5.6) one has that

$$\hat{\mu}_x^{-1} = \left( \sum_{i=1}^{r} L_i N_i(t) \right) \left( \sum_{j=1}^{r} K_j N_j(t) \right)^{-1}$$  \hfill (6.3)

where

$$L_i = [\beta_i^{(1)}]^{-1}.$$

The asymptotic variance of the random variable in the numerator and the one in the denominator of (6.3) can be obtained from (5.11). It remains to find the covariance between these two random variables and, using (5.7) and (5.8), it follows that

$$t^{-1}\text{Cov}^*[\sum_{i=1}^{r} L_i N_i(t), \sum_{j=1}^{r} K_j N_j(t)] =$$

$$b^{-1}(2) \left[ \sum_{i=1}^{r} C_i K_i \left[ \sum_{j=1}^{r} C_j L_j \right] + \sum_{i=1}^{r} K_i L_i \tilde{c}_i \right]$$  \hfill (6.4)

$$- \left[ \sum_{i=1}^{r} K_i b_i \tilde{c}_i \right] \left[ \sum_{j=1}^{r} C_j L_j \right] - \left[ \sum_{j=1}^{r} L_j b_j \tilde{c}_j \right] \left[ \sum_{i=1}^{r} C_i K_i \right].$$

Theorem 6.1 then follows from theorem 5.3.
Theorem 6.1

Under the assumptions of section 2, \( t^{\frac{k}{2}} \hat{\mu}_x^{-1} \) is asymptotically, as \( t \to \infty \), normally distributed with mean \( t^{\frac{k}{2}} \mu_x^{-1} \) and variance given by

\[
\text{Var}^*\left(t^{\frac{k}{2}} \hat{\mu}_x^{-1}\right) = \sigma^{-1} \left\{ \sum_{i=1}^{r} \left[ \sum_{k=1}^{K_i} \left[ \sum_{j=1}^{L_i} \mathbb{D}_j \right] \right] \right\}
\]

An interesting special case to consider is the case \( \alpha = 0 \). In that case, \( \mu_x = \sigma \). One can easily show that \( t^{-1} \hat{N}(t) \) converges almost surely to \( \sigma^{-1} \) and so, for large \( t \), \( t^{-1} \hat{N}(t) \) and \( \hat{\mu}_x^{-1} \) both form reasonable estimates of the fineness. Also, for \( \alpha = 0 \),

\[
\text{Var}^*\left(\hat{\mu}_x^{-1}\right) = \frac{1}{t \sigma} \left[ (\sigma^{(2)/\beta^2}) - 1 \right],
\]

which is the asymptotic variance of \( t^{-1} \hat{N}(t) \).

7. Estimation under some Specific Distributions for \( F^*_1 \)

Three sample quantities of interest in this paper are \( \hat{B}_i \), which is used in the index of mixing, \( \hat{f}_i \), which is used to estimate population amount proportions, and \( \hat{\mu}_x^{-1} \), which is used as a measure of fineness. Asymptotic variances have been found for all of these quantities and these

-22-
variances can be written in terms of the $B_i$'s, some moments of the $F_i^*$ distributions, and $\alpha$. If $\hat{B}_i$ is used to estimate $B_i$ and if $\alpha$ is known, then it remains to estimate the moments of the $F_i^*$ distributions in order to estimate the variances under consideration. Sometimes it may be necessary to estimate $\alpha$ also. In this section, estimation of the parameters of the $F_i^*$ distributions and of $\alpha$ will be discussed for some specific cases.

Assume that one is sampling from a population containing items of $r$ possible amounts, $A_1, \ldots, A_r$, with $0 < A_i < \infty$ for $i=1, \ldots, r$ and $r < \infty$. Also assume that the items are drawn independently from an infinite population. This sampling scheme fits into the renewal process model discussed in section 2 if the "types" are associated with "amounts" by the relation

$$F_i^*(x) = 0 \quad \text{for } x < A_i$$
$$= 1 \quad \text{for } x \geq A_i.$$  \hfill (7.1)

Hence, $P[X_j = A_i] = p_i$, with $0 < p_i < 1$ and

$$\sum_{i=1}^{r} p_i = 1,$$

where $X_j$ is the amount of the $j$th item selected from the population.

This is a reasonable case to consider for practical problems since data is often obtained in discrete form. For example, if an experimenter is interested in particle weight in a coal experiment, he may put a sample of coal through a
series of sieves and assign a weight to a particle that pass through sieve I, but not through sieve J, equal to an average weight of particles that pass through sieve I. Thus the data extracted from the sample is discrete. Other methods of measurement, such as sedimentation, result in similar sets of discrete data.

In the discrete case under consideration \( \beta_i^{(\alpha)} = \lambda_i^\alpha \) for any real, finite \( \alpha \) so no estimation of the moments of the \( F_i(.) \) distribution is necessary if \( A_i \) is known.

Assume now that the number of items in the sample is fixed at \( n \) and let \( n_i \) denote the number of items of type \( i \) in the sample of size \( n \). Using the symbol \( \sim \) to mean "proportional to", one has that

\[
L \sim p_1^{n_1} p_2^{n_2} \ldots p_r^{n_r}, \quad \text{with} \quad \sum_{i=1}^{r} n_i = n, \quad \text{or} \quad (7.2)
\]

\[
L \sim \left( \sum_{j=1}^{r} q_j \lambda_j^\alpha \right)^{-n} \prod_{i=1}^{r} (q_i \lambda_i^\alpha)^{n_i},
\]

using equation (5.3). Then

\[
\log L \sim -n \log \left( \sum_{j=1}^{r} q_j \lambda_j^\alpha \right) + \sum_{i=1}^{r} n_i \log q_i + \alpha \sum_{i=1}^{r} n_i \log \lambda_i \quad \text{and} \quad (7.3)
\]

\[
\frac{\partial \log L}{\partial \alpha} = -n \frac{\sum_{i=1}^{r} q_i \lambda_i^\alpha \log \lambda_i}{\sum_{j=1}^{r} \lambda_j^\alpha} + \sum_{i=1}^{r} n_i \log \lambda_i. \quad (7.4)
\]
Setting (7.4) equal to zero does not result in a closed solution for the estimator of $\alpha$ unless $r=2$. In the case $r=2$ one obtains, denoting the solution by $\hat{\alpha}^*$,

$$\hat{\alpha}^* = \frac{\log[(q_2n_1)/(q_1n_2)]}{\log (A_1/A_2)}.$$ 

Now, $n_1$ and $n_2$ will be replaced by the random quantities $N_1(t)$ and $N_2(t)$ to form the estimator

$$\hat{\alpha} = \frac{\log[(q_2N_1(t))/(q_1N_2(t))]}{\log (A_1/A_2)},$$

(7.5)

The quantity $\hat{\alpha}$ is a consistent estimator of $\alpha$ which can be seen as follows:

$$\frac{q_2(N_1(t)/t)}{q_1(N_2(t)/t)} \overset{a.s.}{\sim} \frac{q_2\rho_1}{q_1\rho_2} = \frac{\rho_1(p_2/A_2^\alpha)}{\rho_2(p_1/A_1^\alpha)} = \frac{A_1^\alpha}{A_2^\alpha},$$

and hence

$$\log \frac{q_2N_1(t)}{q_1N_2(t)} \overset{a.s.}{\sim} \alpha \log \frac{A_1}{A_2}, \text{ as } t \to \infty.$$ 

Replacing $\alpha$ by $\hat{\alpha}$ in (7.5) results in no solution for $\hat{\alpha}$.

In this discrete case, using equations (3.12) and (3.15), one has that

$$\text{Var}^*(\hat{\alpha}_i) = t^{-1}[A_iB_i(1-B_i) + B_i(\mu_a - A_i)]$$

(7.6)

where

$$\mu_a = \sum_{i=1}^{r} A_iB_i,$$
and

\[
\text{Cov}^{*}(B_i, B_j) = t^{-1} B_i B_j [\mu_a - A_i - A_j]. \quad (7.7)
\]

Consider now the distribution given by

\[
dF_i^*(x) = \frac{1}{\theta_i} \exp[-x/\theta_i] \, dx \text{ for } x > 0,
\]

\[
= 0 \text{ for } x \leq 0,
\]

with \( \theta_i > 0 \). Then

\[
\beta_i^{(a)} = \Gamma(a+1) \theta_i^a, \text{ for } a > 0.
\]

Estimators of the parameters of interest will be

found for the fixed sample size situation, and then the

fixed sample sizes will be replaced by their random

counterparts. Since the distribution function given by

(2.1) is difficult to work with by the method of maximum

likelihood, the parameters for each \( F(.) \) will be estimated

separately.

Assume that a sample of \( n \) independent observations,

\( x_1, \ldots, x_n \), is obtained from the distribution with

probability density given by

\[
\frac{x^\alpha dF_i^*(x)}{\beta_i^{(a)}}, \quad (7.9)
\]

where \( \alpha > 0 \) and \( dF_i^*(x) \) is given by (7.3). The

maximum likelihood estimator of \( \theta_i \) is given by

\[
\hat{\theta}_i^* = [n(a+1)]^{-1} \sum_{i=1}^{n} x_i. \quad (7.10)
\]
Now assume that one is sampling to an amount \( t \) from a distribution of the form (2.1) and let \( Y_{ij} \) denote the jth observation of the ith type. Assume that each of the r distributions \( F_i(.) \) is of the form given in (5.3) with \( dF_i(.) \) as given in (7.8). Then, an estimator of the parameter \( \theta_i \) from the ith distribution is given by

\[
\hat{\theta}_i = \left[ H_i(t)(\alpha+1) \right]^{-1} H_i(t) \sum_{j=1}^{n_i(t)} X_{ij}, \quad \text{for } n_i(t) > 0. \quad (7.11)
\]

**Theorem 7.1**

Under the assumptions of section 2, \( \hat{\theta}_i \) converges almost surely to \( \theta_i \) as \( t \to \infty \).

**Proof:**

Let \( X_j \) denote the jth observation in the sample and let

\[
\omega_j^{(i)} = 1 \text{ if } X_j \text{ is of type } i \quad = 0 \text{ otherwise.}
\]

Then,

\[
\sum_{j=1}^{n_i(t)} X_{ij} = \sum_{j=1}^{n} \omega_j^{(i)} X_j.
\]

Also, \((1/n) \sum_{j=1}^{n} \omega_j^{(i)} X_j + \theta_i(\alpha+1)\) almost surely as \( n \to \infty \) by the strong law of large numbers. It then follows that

\[
\frac{1}{n_i(t)} \sum_{j=1}^{n_i(t)} X_{ij} = \left( \frac{N(t)}{n_i(t)} \right) \frac{1}{N(t)} \sum_{j=1}^{n} \omega_j^{(i)} X_j + \theta_i(\alpha+1)
\]
almost surely, as $t \to \infty$, since $t^{-1} \ln(t) \to b^{-1}$,
$t^{-1} N_i(t) \to b^{-1} \beta_i$ and $\ln(t) \to \infty$ almost surely as $t \to \infty$,
and the proof is completed.

Suppose now that the distribution $F_i^*(\cdot)$
is given by

$$dF_i^*(x) = \frac{1}{x \sigma_i^2 (2\pi)^{1/2}} \exp \left[ (-1/2 \sigma_i^2) \left( \log x - \mu_i \right)^2 \right] dx \text{ for } x > 0, \quad (7.12)$$

$= 0 \text{ otherwise.}$

This log-normal distribution is used extensively in bulk
sampling problems (see Epstein [14]). In this case

$$\beta_i = \exp \left[ \mu_i + \frac{1}{2} \sigma_i^2 \right].$$

Again, under the assumption that a fixed number of
observations are drawn from the distribution given by (7.9)
with $dF_i^*(x)$ given by (7.12), the maximum likelihood
estimators are obtained separately for each distribution and
then the fixed sample sizes are replaced by their random
counterparts yielding

$$\hat{\mu}_i = \frac{1}{N_i(t)} \sum_{j=1}^{N_i(t)} \log x_{ij} - \frac{\sigma_i^2}{2} \quad \text{and} \quad (7.13)$$

$$\hat{\sigma}_i^2 = \frac{1}{N_i(t)} \left[ \sum_{j=1}^{N_i(t)} (\log x_{ij})^2 - \frac{1}{N_i(t)} \left( \sum_{j=1}^{N_i(t)} \log x_{ij} \right)^2 \right]. \quad (7.14)$$

Since $\sigma^2$ remains the same over all $x$ distributions, a
pooled estimator can be formed as follows:
\[ \phi = \left( \sum_{i=1}^{r} N_i(t)i \sigma_i^2 \right)^{-1} \sum_{i=1}^{r} \sum_{j=1}^{r} \log x_{ij} - \sum_{i=1}^{r} N_i(t) \mu_i. \]

If \( \mu_i \) and \( \sigma_i^2 \) are replaced by \( \hat{\mu}_i \) and \( \hat{\sigma}_i^2 \) respectively, no solution for \( \phi \) is obtained. By methods similar to those of theorem 7.1, \( \hat{\mu}_i \), \( \hat{\sigma}_i^2 \) and \( \phi \) can be shown to be consistent estimators of \( \mu_i \), \( \sigma_i^2 \) and \( \phi \) respectively.

8. The Sample Amount Proportions under Markov Dependence

Consider a bulk sampling problem for which the assumptions of section 2 hold, with the exception of the independence assumption. Items are often sampled in a cluster from a population in which the different types are not thoroughly mixed, thus forming a sample in which observations are dependent. One can characterize this dependence by assuming that the items are sampled sequentially and that this sequence forms a Markov chain with respect to item types. That is, the conditional probability of drawing a type i item given the type of the preceding item should not be the same as the unconditional probability of drawing an item of type i. This sampling of items in a sequence can be accomplished, for example, in sampling from conveyor belts or in cores.

By assuming \( r=2 \), for simplicity, and by restricting the transition matrix, \( P \), so that the stationary absolute probability distribution is \((p_1, p_2)\), with \( p_1 \) as
defined in section 2, one can arrive at the following form for $P$,

$$
P = \begin{bmatrix}
  \frac{\phi}{P_1} & \frac{\phi}{P_2} - \frac{\phi}{P_1} \\
  P_1 - \frac{\phi}{P_2} & P_2 + \frac{\phi}{P_2}
\end{bmatrix}
$$

(8.1)

where

$$P_i + \frac{\phi}{P_i} = P[X_j \text{ is type } i \mid X_{j-1} \text{ is type } i].$$

A sufficient condition for $P$ to be irreducible and recurrent is

$$\max(-P_1^2, -P_2^2) < \phi < P_1P_2.$$

(8.2)

Note that this Markov dependence only applies to types, and the item amounts are still assumed to be independent. This sequence of sampled item amounts then forms a Markov renewal process (see [19], [20] and [21]) with the distribution function for the sojourn time in state $i$ given by $F_i(.)$, where $F_i(.)$ is as defined in section 2. Using theorems from Pyke and Schaufele, one obtains the following theorem.

Theorem 8.1

If the transition matrix $P$ is given by (8.1) and if $\phi$ satisfies (8.2), then $\hat{P}_1$, as defined in section 3, is asymptotically, as $t \to \infty$, distributed as a normal random
variable with mean $B_1$ and variance given by

$$\text{Var}^* \left( t_{31}^b \right) = B_1 B_2 \left\{ \left[ B_2 \frac{b^{(2)}}{b_1} + B_1 \frac{b^{(2)}}{b_2} \right] + \frac{2\phi}{B_1 B_2 \phi} \left[ B_2 \frac{b_1}{b_2} + B_1 \frac{b_2}{b_1} \right] \right\}.$$ 

If $\phi$ is positive, this variance is larger than the corresponding variance under independence. This supports the claim that the theoretical variance under independence is usually much smaller than the variance observed in practice (see Vance [25]).

The parameter $\phi$ can be estimated but in order to form a good estimator one must have some knowledge of the number of transitions from type i to type i as well as the number of type i items in the sample, for each i.

9. Conclusion

A population consisting of $r$ types of items, each item being characterized by a non-negative, real-valued random variable called the item's amount, is considered. An index of mixing, an estimator of population amount proportions, and an index of fineness are defined and their asymptotic distributions obtained. Estimation of the variances of these random quantities is discussed for three particular distributions of item amounts, a discrete distribution, the negative exponential and the log-normal.
The author is indebted to Dr. W. A. Thompson, Jr. for suggesting this problem and giving much assistance in the solution, and to the members of his graduate committee at the Florida State University for their valuable assistance.
References


