BNWL-1386
UC-15

2-

5-70

STATISTICAL TECHNIQUES
FOR ENHANCING THE ROLE
OF MATERIAL BALANCE ACCOUNTING
IN SAFEGUARDS

K. B. Stewart

May 1970

AEC RESEARCH &
DEVELOPMENT REPORT

DATE TO	P. NO.	LOCATION	FILE NO. / DATE

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or

B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

PACIFIC NORTHWEST LABORATORY

RICHLAND, WASHINGTON

operated by

BATTELLE MEMORIAL INSTITUTE

for the

UNITED STATES ATOMIC ENERGY COMMISSION UNDER CONTRACT AT(45-1)-1830

3 3679 00061 6773

BNWL-1386

UC-15, Standard and
Nuclear Materials
Management

STATISTICAL TECHNIQUES FOR ENHANCING THE ROLE
OF MATERIAL BALANCE ACCOUNTING IN SAFEGUARDS

By

K. B. Stewart

Applied Mathematics Department
Systems and Electronics Division

May 1970

FIRST UNRESTRICTED
DISTRIBUTION DATE

MM 24 70

BATTELLE MEMORIAL INSTITUTE
PACIFIC NORTHWEST LABORATORIES
RICHLAND, WASHINGTON 99352

STATISTICAL TECHNIQUES FOR ENHANCING THE ROLE
OF MATERIAL BALANCE ACCOUNTING IN SAFEGUARDS

K. B. Stewart

ABSTRACT

This report, BNWL-1386, was originally a paper prepared for and delivered at the IAEA Panel on Safeguards Systems Analysis of Nuclear Fuel Cycles in Vienna, Austria, August 25 through 29, 1969. The report discusses the use of a statistical model to assist in making certain basic determinations (such as whether an apparent loss occurred as a single incident or as several incidents over a period of time). From such determinations the proper course of action is determined. The report also stresses the importance of highly accurate measurements and the proper application of as much pertinent data as can be used, to obtain meaningful results.

TABLE OF CONTENTS

INTRODUCTION	1
SUMMARY	1
CONCLUSIONS	3
DISCUSSION	4
1. The Concept of MUF	4
2. A Mathematical Model to Illustrate the Parameters Involved in the MUF Statistics	6
3. Statistical Models and Procedures	7
4. Control on a Unit Period Basis	9
5. Control on the Basis of the Total of n Periods	11
6. Another Approach to Inventory Estimation, MUF Estimation and Hypothesis Testing	15
7. A Comparison of the Three Methods Under a Single Isolated Diversion	20
8. Alternative Hypothesis II - Systematic Diversion	26
9. An Example of Optimization by Variance Minimization	34
10. Variance Minimization by an Optimum Allocation of Effort	43
GLOSSARY OF TERMS AND NOTATIONS	50
Various MUF's	50
BIBLIOGRAPHY	52
APPENDIX	A-1

STATISTICAL TECHNIQUES FOR ENHANCING THE ROLE OF MATERIAL BALANCE ACCOUNTING IN SAFEGUARDS

K. B. Stewart

INTRODUCTION

This report presents the results of a study of statistical techniques designed to develop and maintain tighter controls over the inventory of special nuclear materials. Sample surveys are made at commercial nuclear facilities on a periodic basis by safeguards inspections teams in order to verify the accuracy of the facility's stated inventory amounts of the materials. The results of the decisions are used to formulate decisions about whether certain apparent and unexplained losses may be actually the result of diversion. The following alternatives for employing optimal procedures are discussed:

(1) the sampling effort may be apportioned in such a way that the survey results in an unbiased, minimum variance estimate of the inventory amount for a given level of inspection efforts, and (2) the historical data of the accountability system may be utilized to augment the current survey data to provide a background for more valid decisions about the possibility of diversion.

SUMMARY

The following is a brief introduction of subject matter in the areas discussed in this report.

The Concept of MUF

When the Material Unaccounted For (MUF) indicates that a loss has occurred, this should be an automatic signal to investigate the possibility of diversion.

A Model to Illustrate the Parameters Involved in MUF Statistics

This model is created to indicate the parameters estimated by the use of MUF statistical procedures.

Statistical Models and Procedures

This is a discussion of statistical models of the observed data, and what their parametric counterparts are when using standard MUF procedures.

Short-term Control

To detect diversion it is important to match the correct statistic with the correct alternative hypothesis. The alternative hypotheses are that: (a) diversion is an isolated incident, or (b) diversion is on a systematic basis.

Control on the Basis of n Periods

Under systematic diversion, the long-term MUF and the associated statistics assume an important diversion detection role.

Another Approach to Inventory Verification

How to use specific features of existing data in a system in an optimum manner to create diversion detection procedures which are more sensitive than those used customarily.

A Comparison of the Three Proposed Methods Based on the Alternative Hypothesis That a Single Isolated Diversion has Occurred.

The three methods discussed are the optimum method identified in the preceding paragraph, the usual short-term MUF procedure, and the long-term MUF procedures.

A Comparison of the Three Methods Based on the Alternative Hypothesis That the Diversion Has Been on a Systematic Basis

In this situation the short-term MUF statistics use only a small fraction of the inherent information in the system. The long-term MUF has much greater relative efficiency in detection diversion.

An Example of Variance Minimization

This topic discusses the best way to acquire information from different strata of an inventory in order to attain the estimate of the total amount of material in an inventory which has minimum variance. The material is converted to a scale based on units of strategic value and indicates how relatively inefficient sampling can be when the sampling procedures are not chosen skillfully.

Other Methods of Optimizing the Information Acquired in Inventory Verification

This is a discussion of how to obtain optimum sampling and measurement procedures in assessing the contents of a container.

CONCLUSIONS

- It is useful to develop mathematical and statistical models in order to study the manner in which material balance area accounting systems behave.
- Safeguards and inventory verification are rich areas for formalizing the mathematical and statistical procedures and for studying the consequences of different decision-making procedures.
- Three different kinds of optimization techniques are studied, which under appropriate circumstances, can result in statistical techniques with the most sensitivity.

- The correct type of statistic to use in detecting diversion by material balance accounting procedures depends on whether the diversion is systematic or isolated.

DISCUSSION

1. THE CONCEPT OF MUF

MUF is an acronym for Material Unaccounted For. It is used in connection with material balance accounting in a material balance area to estimate the amount of material which is not accounted for. It is calculated by taking the difference between the book inventory, (the amount of material which is supposed to be present in the inventory) and the physical inventory, (the amount of material which is estimated to be in the inventory by direct inventory estimation techniques). In the absence of measurement, sampling, and bookkeeping errors, a positive MUF indicates a sink (loss) and a negative MUF indicates a source (an unaccounted for gain in material).

In practice, of course, the MUF is composed of many values which are subject to various types of errors of estimation, as for example sampling, measurement, and calibration errors. Without an understanding that these error sources exist and without an estimate of the variation that the MUF statistic can experience under conditions when the material is in a state of control, it is difficult to adjudge the practical or statistical significance of a particular MUF value. Unless the MUF is grossly aberrant it is difficult to say whether a particular MUF represents a random fluctuation from the MUF's expected value (which would be zero under controlled conditions) or whether the MUF indicates a true loss condition.

If the standard deviation of the MUF can be estimated with sufficient accuracy there are several ways to test whether the MUF value indicates, in a statistical sense, that the material is not controlled. The methods are essentially identical. One

way is to determine whether the MUF exceeds a critical value $c_r = t \sigma_{\text{MUF}}$ say, or equivalently to test whether $\text{MUF}/\sigma_{\text{MUF}}$ exceeds the critical value t . MUF and the statistical procedures associated with it are important to safeguards because one of the necessary functions of a material balance accounting system as used in safeguards is to determine when an indicated loss is too large to be accounted for by the error sources which exist in the process. It is important to note that while diversion may precipitate a MUF value which indicates that a loss has occurred, the converse is not true. A statistically significant and positive MUF value does not necessarily indicate diversion. Thus while the consequences of a particular loss mechanism in a process may map onto a MUF with a positive expected value, the converse is not true. It is not possible, for example, to infer diversion from a significant and positive MUF value. That is to say the mapping is onto and not one-to-one. In other words a significant and positive MUF is a necessary but not sufficient basis for inferring diversion based on material balance accounting.

On the other hand, if there is a concern that diversion may exist as a particular loss mechanism, it is of course perfectly proper to evaluate the capability to determine that the material is out-of-control as a function of the amount of diversion without implying that a MUF significantly displaced from zero represents a diversion.

There are certain conditions under which it is useful to consider only an isolated time period when investigating diversion. This is the case when the variance of the MUF is dominated by input-output uncertainties as compared to the uncertainties of the beginning and ending physical inventories. There is a dynamic situation, however, when it is very valuable to consider not only the individual MUF's by themselves but to consider the pattern of the MUF's as they evolve in the process

of controlling the flow of material and detecting indicated losses when diversions are present. This is the case when the variance of the MUF's are dominated by the contributions of the beginning and ending inventory variance components.

2. A MATHEMATICAL MODEL TO ILLUSTRATE THE PARAMETERS INVOLVED IN THE MUF STATISTICS

Let

η_i = the true amount of material in an inventory at the beginning of the i^{th} time period

δ_i = the true amount of material received minus the true amount of material removed from the MBA by legal routes in the i^{th} time period

ϵ_i = the amount of material diverted in the i^{th} period.

Then

$$\epsilon_i = (\eta_i + \delta_i) - \eta_{i+1}$$

denotes the relationship between these parameters. $\eta_i + \delta_i$ is the true book value for a single period, the i^{th} period, and η_{i+1} is the true inventory value at the end of the i^{th} period. After n periods the true relationship is

$$\begin{aligned} \sum_1^n \epsilon_i &= \sum_1^n (\eta_i + \delta_i - \eta_{i+1}) \\ &= \eta_1 - \eta_{n+1} + \sum_1^n \delta_i. \end{aligned}$$

A good tactic for a diverter who wishes to mask diversion is to overstate wastes and/or to overstate the amounts in the inventories. If waste measurements and other input-output measurements are controlled without error, the true amount of material in the inventory after n periods is equal to

$$\eta_{n+1} = \eta_1 + \sum_1^n \delta_i - \sum_1^n \epsilon_i$$

In the absence of measurement errors, the book inventory over the n periods will predict the value

$$\eta_1 + \sum_1^n \delta_i$$

as the amount of material in the inventory so that the physical inventory must be overstated by the amount

$$\sum_1^n \epsilon_i$$

in order to balance the books.

In practice, of course, the true amounts of material are not known and the relationships are obscured by sampling, measurement, and counting errors. The practical problems in detecting a discrepancy which may indicate a diversion are:

- To determine when a discrepancy exists between the physical and book inventories by the use of objective statistical measures
- To detect a discrepancy for as small a discrepancy as is possible
- To detect a discrepancy which may indicate a diversion as soon as possible
- To be efficient in these procedures, and
- To minimize the times when the MUF gives incorrect signals as to whether or not a true discrepancy has occurred.

3. STATISTICAL MODELS AND PROCEDURES

Let

y_i = the measurement of the physical inventory at the beginning of the i^{th} time period.

x_i = the estimate of the input minus output values for the i^{th} time period.

In other words y_i estimates η_i and x_i estimates δ_i . The MUF for the i^{th} period is defined as

$$z_i = (y_i + x_i) - y_{i+1}$$

$$= \left(\begin{array}{l} \text{the estimate of} \\ \text{the amount of} \\ \text{material which} \\ \text{is supposed to} \\ \text{be in the} \\ \text{inventory} \end{array} \right) \text{minus} \left(\begin{array}{l} \text{the measurement} \\ \text{of the amount of} \\ \text{material which} \\ \text{is in the} \\ \text{inventory} \end{array} \right)$$

= the book inventory minus the physical inventory.

The term z_i is then an estimate of ϵ_i , the amount of material which has been diverted within the i^{th} period. The term y_{n+1} estimates the amount of material which is in the inventory after n periods. The book inventory after n periods is equal to

$$y_1 + \sum_1^n x_i$$

so that for the total of n periods

$$u_n = \left(y_1 + \sum_1^n x_i \right) - y_{n+1}$$

= (the book inventory) minus (the physical inventory)

gives a corresponding MUF value for the total of the n periods. u_n , then gives an estimate of $\sum_1^n \epsilon_i$ = the amount of material which has been diverted over the n periods. Note that if the z_i values, the MUF's for the individual periods, are summed

$$\begin{aligned}
z_1 + z_2 + \dots + z_n &= y_1 + x_1 - y_2 \\
&\quad + y_2 + x_2 - y_3 \\
&\quad \dots \\
&\quad + y_n + x_n - y_{n+1} \\
&= \left(y_1 + \sum_{i=1}^n x_i \right) - y_{n+1} \\
&= u_n
\end{aligned}$$

so that u_n is algebraically identical with the sum of the z_i values over the n periods.

This simple treatment indicates that a value of MUF increases over a long period of time (or equivalently for the sum of individual MUF's). Thus, the proportional contribution to the MUF variance induced by the following variance sources increases as the period for the MUF increases,

- Material inputs,
- Material shipped,
- Operating losses, and
- Waste streams.

There are other approaches to this problem. See Section 6 and References 1, 2, and 3.

4. CONTROL ON A UNIT PERIOD BASIS

As indicated before, the MUF for the i^{th} period is defined as

$$z_i = (y_i + x_i) - y_{i+1}.$$

Under the assumption of statistical independence the variance of z_i is equal to

$$\sigma_{z_i}^2 = \sigma_{y_i}^2 + \sigma_{x_i}^2 + \sigma_{y_{i+1}}^2$$

so that the standard deviation of z_i is equal to

$$\sigma_{z_i} = \sqrt{\sigma_{y_i}^2 + \sigma_{x_i}^2 + \sigma_{y_{i+1}}^2}.$$

If the periods are of constant duration with a constant amount of throughput, the assumptions $\sigma_y^2 = \sigma_{y_i}^2$ and $\sigma_x^2 = \sigma_{x_i}^2$, for all i , are reasonable. Then

$$\sigma_{z_i} = \sqrt{2\sigma_y^2 + \sigma_x^2}.$$

If the values σ_y^2 and σ_x^2 are known, then the region in the positive direction from

$$\begin{aligned} c_r &= E(\text{MUF})_{\text{under } H_0} + t_{1-\alpha} \sigma_{z_i} \\ &= 0 + t_{1-\alpha} \sigma_{z_i} \end{aligned}$$

gives the best critical region for rejecting H_0 : the null hypothesis that no loss has occurred, where $E(\text{MUF}) = 0$ under H_0 , and $t_{1-\alpha}$ is the $(1-\alpha) \times 100$ percentile point of the $N(0,1)$ distribution.

The probability by this procedure that ϵ , a particular amount of diverted material, will indicate a loss is determined from

$$P(\epsilon) = \int_{c_r}^{\infty} \frac{1}{\sigma_z \sqrt{2\pi}} \exp - 1/2 \left(\frac{x - \epsilon}{\sigma_z} \right)^2 dx.$$

This assumes that the MUF will be approximately normally distributed. If different values of $P(\epsilon)$ are determined and are plotted as the ordinate versus the abscissa ϵ , the resulting curve, (a power curve), will give the capability to detect by this procedure that a loss has occurred. The power curve depicts this relationship over the whole range of the possible amounts of material diverted.

5. CONTROL ON THE BASIS OF THE TOTAL OF n PERIODS

The MUF for a total of n consecutive periods is given by

$$u_n = y_1 + \sum_{i=1}^n x_i - y_{n+1}$$

which under the assumptions of Section 4 has the standard deviation

$$\sigma_{u_n} = \sqrt{2\sigma_y^2 + n\sigma_x^2}.$$

Suppose that a constant amount of material is being diverted each period. The best critical region for the test of a total of n periods

$$\begin{aligned} c_n &= E(\text{MUF})_{\text{under } H_0} + t_{1-\alpha} \sigma_{u_n} \\ &= 0 + t_{1-\alpha} \sqrt{2\sigma_y^2 + n\sigma_x^2}. \end{aligned}$$

The probability of detecting that a loss has occurred is

$$P\left(\sum_{i=1}^n \epsilon_i\right) = P(n\epsilon) = \int_{c_n}^{\infty} \frac{1}{\sigma_{u_n} \sqrt{2\pi}} \exp - 1/2 \left(\frac{w - n\epsilon}{\sigma_{u_n}} \right)^2 dw.$$

It can be shown by the use of power curves that the n period basis has more power to detect the amount of diversion $n\epsilon$, than the unit period basis has to detect the amount ϵ . This can also be demonstrated in the following way. Consider the relation

$$h = \frac{\text{amount diverted}}{\text{standard deviation of the MUF}}$$

h may be regarded as the number of standard deviations that the expected value of the MUF under diversion is displaced from 0, where 0 is the expected value of the MUF under the null hypothesis that no diversion has occurred. For one period

$$h_1 = \frac{\epsilon}{\sqrt{2\sigma_y^2 + \sigma_x^2}}.$$

For a total of n periods

$$\begin{aligned} h_n &= \frac{n\epsilon}{\sqrt{2\sigma_y^2 + n\sigma_x^2}}. \\ &= \frac{\epsilon}{\sqrt{\frac{2\sigma_y^2}{n^2} + \frac{\sigma_x^2}{n}}}. \end{aligned}$$

If $\sigma_y^2 \gg n\sigma_x^2$, then $h_n \approx n h_1$. If $\sigma_x^2 \gg \sigma_y^2$, then $h_n \approx \sqrt{n} h_1$.

In any situation under systematic diversion the total MUF is more sensitive to the detection of diversion than the unit MUF.

On the other hand, however, assume that only one diversion event has occurred. Then define g as

$$g = \frac{\text{the amount of diversion}}{\text{standard deviation of the MUF}}.$$

$$g_1 = \frac{\epsilon}{\sqrt{2\sigma_y^2 + \sigma_x^2}}, \quad g_n = \frac{\epsilon}{\sqrt{2\sigma_y^2 + n\sigma_x^2}}.$$

It is apparent here that $g_n < g_1$, (unless n=1) so that if only one diversion has occurred there is more discriminatory power associated with shorter MUF's.

There are two important characteristics which any adequate safeguards system should have; (a) sensitivity to detect with high probability that a loss has occurred for the amount of diversion which would constitute a significantly dangerous amount of material and (b) timeliness in detecting such losses. The implications from Sections 3 and 4 are as follows.

- When no diversion is present there is a tendency for the variance of the long-term MUF to be dominated by input/output variance contributions.
- When diversion is on a systematic basis, the long-term MUF has more discriminatory power.
- The shorter term periods of inventory verification are justified on the basis of timeliness and on the fact that when only one diversion event has occurred that the short-term MUF which includes the diversion event has more discriminatory power to detect this loss.
- It becomes increasingly more difficult to mask diversion by overstating inventory when diversion is on a systematic basis, as the magnitude of the falsification must be increased to balance the book inventory.

Extensions

Consider as an example the case where the standard deviation of the MUF is known or can be estimated with accuracy. Diversion will be investigated if the observed MUF exceeds some critical value. If the MUF is composed of a sum of independent components and is not dominated by some individual components, then by the central limit theorem the MUF will tend to be normally distributed. $c_r = t_{1-\alpha} \sigma_{\text{MUF}}$ can then be used as a critical value where $t_{1-\alpha}$ is the $(1-\alpha) \times 100$ percentile point of the $N(0,1)$ distribution.* An observed MUF which is larger than c_r will then be cause to investigate for diversion. The probability that an actual diversion will indicate that a loss has occurred can be determined by evaluating the probability that an observed MUF will exceed the critical value as a function of the amount of diversion. Let D denote the actual amount of diversion. Then if the

* This assumption is made for the sake of exposition. An actual application may require the use of student's t based on the Welch-Satterthwaite estimate of the effective number of degrees of freedom.

probability of detecting a loss is plotted against D , a power-of-the-test curve such as shown in Figure 1 will result. At $D = 0$, the power-of-the-test gives α , the probability of "blowing the whistle" when in fact no loss has occurred.

FIGURE 1. A Typical Power Test

The value of this approach of looking at the power to detect a loss as a function of the amount of diversion is that the clear and graphic representation of the power curve test can be used to determine whether the test is sensitive enough. If the curve indicates that there is too little sensitivity then it may be necessary to increase the accuracy and precision of the measurements or to increase the number of measurements taken in the data points which enter into the evaluation of the MUF. There are essentially three ways that the sensitivity can be increased.

- (a) The α value may be increased. In most instances this is not a satisfactory procedure since it can result in an undue amount of unwarranted investigation for diversion.
- (b) It may be possible to increase the efficiency of the MUF and inventory estimators by a better use of the information which is inherent in the system. A technique here

is to determine the minimum variance, unbiased estimator (MVUE) of the inventory by the correct use of all the information in the system pertinent to the inventory.

- (c) The variance of the MUF and the physical inventory can also be minimized in another manner. This procedure involves the predetermination of the manner in which the data are obtained. It can be called the minimization of the variance of an estimator subject to a practical constraint such as the amount of money (effort) spent on the acquisition of inventory verification information. Its use requires knowledge in the form of variance component estimates.

6. ANOTHER APPROACH TO INVENTORY ESTIMATION, MUF ESTIMATION AND HYPOTHESIS TESTING

The previous section has indicated that a systematic study of the effectiveness of different alternative hypotheses can be very useful in improving the test sensitivity in detecting that a loss has occurred and in enhancing the user's understanding of the underlying relationships. Stewart⁽¹⁾ indicates that there is another approach to the problem of inventory estimation, MUF estimation, and of testing $H_0: E(MUF) = 0$. It is based essentially on the fact that the values

$$y_i + x_i + x_{i+1} + \cdots + x_n \quad (i = 1, 2, \cdots, n)$$

as well as y_n , estimate the parametric value η_{n+1} , the true amount of material in the inventory under H_0 , the null hypothesis that no loss has occurred. The minimum variance unbiased estimator of η_{n+1} is determined by minimizing the variance of

$$I_{n+1} = \sum_{i=1}^n k_i (y_i + x_i + \cdots + x_n) + k_{n+1} y_{n+1}$$

subject to the constraint that $\sum_{i=1}^{n+1} k_i = 1$. The techniques depend on certain assumptions. These assumptions would be very reasonable in a plant with relatively constant throughput and with constant measurement variance in estimating the amount of material in the physical inventory. I_{n+1} is an unbiased estimate of η_{n+1} when the process is in control. A systematic loss (e.g., diversion) causes the I_{n+1} value to be inflated, that is to overestimate η_{n+1} "on the average". For determining whether a loss has occurred this is a desirable feature.

$$\sigma_y^2 = \sigma_1^2, \sigma_x^2 = \sigma_2^2, \text{ and } c = \sigma_2^2 / \sigma_1^2.$$

c then is defined as

$$c = \frac{\text{the variance of the quantity (receipts minus removals)}}{\text{the variance of the physical inventory}}$$

It is assumed that σ_y^2 and σ_x^2 are constant for all inventory periods. This assumption is more necessary for the sake of exposition than for practical reasons. The generalization of a situation where σ_x^2 varies from period to period is relatively easy.* Furthermore, a small percentage error in estimating c has little effect on the results. It is also true that the true value of c can vary significantly from period to period with little effect on results. This is to say that the method is relatively robust when c is not estimated exactly or when c varies from period to period.

Let

$$r_i = y_i + x_i + \cdots + x_n \quad i = 1, \cdots, n.$$

$$r_{n+1} = y_{n+1}.$$

* See Appendix.

The estimate

$$I_{n+1} = p_n r_{n+1} + \sum_{j=1}^n p_{j-1} \left(\prod_{i=j}^{n-1} q_i \right) r_j$$

$$(p_0 = 1)$$

is the minimum variance, unbiased estimate (MVUE) of η_{n+1} , the amount of material in the inventory under H_0 , the hypothesis that no diversion has occurred. $q_i = 1 - p_i$ and p_i is the segment

$$p_i = \underbrace{\frac{1}{1+} \frac{1}{c+1} \frac{1}{1+} \cdots \frac{1}{1+} \frac{1}{c+1}}_{i \text{ c's}}$$

of the continued fraction p . Furthermore the variance of I_{n+1} is $\sigma_1^2 p_n$. I_{n+1} appears to be difficult to calculate. However, the recursion relationship

$$I_{n+1} = p_n y_{n+1} + q_n (I_n + x_n)$$

greatly simplifies the calculation.

A statistic to estimate the amount of material which is unaccounted for using I_j is given by

$$w_n = I_n + x_n - y_{n+1}$$

It is interesting that all such w_j values are statistically independent. As a matter of fact the I_j values could have been generated by finding the unbiased estimates of the inventory which make the covariance of the successive MUF's equal to zero.* Since

$$w_j = (I_j + x_j) - y_{j+1}$$

* See Appendix.

and I_j does not involve x_j and y_{j+1} ,

$$\begin{aligned}\sigma_{w_j}^2 &= \sigma_{I_j}^2 + \sigma_x^2 + \sigma_y^2 \\ &= \sigma_1^2 (p_{j-1} + c + 1).\end{aligned}$$

The MUF as usually defined is

$$z_j = (y_j + x_j) - y_{j+1}$$

and has a variance of

$$\begin{aligned}\sigma_{z_j}^2 &= \sigma_y^2 + \sigma_x^2 + \sigma_y^2 \\ &= \sigma_1^2 (2 + c) .\end{aligned}$$

Let u_n denote the long-term MUF, defined as

$$u_n = (y_1 + \sum_{i=1}^n x_i) - y_{n+1}$$

then

$$\begin{aligned}\sigma_{u_n}^2 &= \sigma_y^2 + n \sigma_x^2 + \sigma_y^2 \\ &= \sigma_1^2 (2 + nc) .\end{aligned}$$

Under a single diversion of the amount ϵ during the j^{th} period, the expected values of the three MUF-type estimators is ϵ .

Under systematic diversion of ϵ units per period it can be shown that

$$E(t_w) = \frac{E(w_j)}{\sigma_{w_j}}$$

$$= \frac{\epsilon (1 + q_{j-1} + q_{j-1}q_{j-2} + \dots + q_{j-1}q_{j-2} \dots q_2q_1)}{\sigma_y \sqrt{1 + c + p_{j-1}}}$$

$$\lim_{j \rightarrow \infty} E(t_w) = \frac{\epsilon / p}{\sigma_y \sqrt{1 + c + p}} = \frac{\epsilon}{\sigma_y p \sqrt{1 + c + p}}$$

where $p = \lim_{j \rightarrow \infty} p_j = \frac{-c + \sqrt{(c+2)^2 - 4}}{2}$.

Under systematic diversion of ϵ units per period

$$E(t_z) = \frac{E(z_j)}{\sigma_{z_j}} = \frac{\epsilon}{\sqrt{\sigma_y^2 + \sigma_x^2 \sigma_y^2}} = \frac{\epsilon}{\sigma_y \sqrt{2 + c}}$$

Under systematic diversion of ϵ units per period

$$E(t_u) = \frac{E(u_j)}{\sigma_{u_j}} = \frac{j\epsilon}{\sigma_y \sqrt{1 + jc}}$$

There are other alternative hypotheses which are possible, of course, but they are difficult to study in a systematic manner. The following sections study the three different methods of assessing the material unaccounted for in the context of the two alternative hypotheses--of an isolated diversion and of systematic diversion. The design of this study can be represented as follows.

Alternative Hypothesis	Procedure		
	w	z	u
	isolated		
	systematic		

7. A COMPARISON OF THE THREE METHODS UNDER A SINGLE ISOLATED DIVERSION

It is assumed that within the j^{th} period an amount ϵ has been diverted. Three methods can be used to test for diversion.

- The MVUE approach using the statistic $w_j / \sigma_{w_j} = t_w$
- The standard approach using the statistic $z_j / \sigma_{z_j} = t_z$
- The long term approach using the statistic

$$\left(\sum_{i=1}^j z_i \right) / \left(\sigma \sum_{i=1}^j z_i \right) = t_u \text{ say.}$$

One method of showing their relative efficiencies is to show how $E(t_w)$, $E(t_z)$ and $E(t_u)$ vary with c and j , where $E(t_w)$ for example indicates the expected value of t_w . The following tables give j , c , $\sqrt{c}$, p_n , $\sqrt{p_n}$, E_w , E_z and E_u . E_w , E_z and E_u are defined as follows.

$$E(t_w) = E_w \left(\frac{\epsilon}{\sigma_y} \right)$$

$$E(t_z) = E_z \left(\frac{\epsilon}{\sigma_y} \right)$$

$$E(t_u) = E_u \left(\frac{\epsilon}{\sigma_y} \right).$$

For example if ϵ is twice the standard deviation of the physical inventory measurement then $E(t_w) = 2E_w$, $E(t_z) = 2E_z$, and $E(t_u) = 2E_u$.

Conclusions

The tables indicate that for an isolated diversion, the MVUE approach, under appropriate conditions, leads to techniques with the most power to detect diversion unless $n = 1$.

	j	c	$\sqrt{c}$	p_n	p	E_w	E_z	E_u
n	1	.1000	.3162	1.0000	.2702	.6901	.6901	.6901
	2	.1000	.3162	.5238	.2702	.7848	.6901	.6742
	3	.1000	.3162	.3842	.2702	.8208	.6901	.6594
	4	.1000	.3162	.3262	.2702	.8373	.6901	.6455
	5	.1000	.3162	.2988	.2702	.8455	.6901	.6325
	6	.1000	.3162	.2851	.2702	.8497	.6901	.6202
	7	.1000	.3162	.2780	.2702	.8519	.6901	.6086
	8	.1000	.3162	.2743	.2702	.8530	.6901	.5976
	9	.1000	.3162	.2724	.2702	.8536	.6901	.5872
	10	.1000	.3162	.2713	.2702	.8539	.6901	.5774
	11	.1000	.3162	.2708	.2702	.8541	.6901	.5680
	12	.1000	.3162	.2705	.2702	.8542	.6901	.5590
	13	.1000	.3162	.2703	.2702	.8543	.6901	.5505
	14	.1000	.3162	.2703	.2702	.8543	.6901	.5423
	15	.1000	.3162	.2702	.2702	.8543	.6901	.5345
	16	.1000	.3162	.2702	.2702	.8543	.6901	.5270
	17	.1000	.3162	.2702	.2702	.8543	.6901	.5199
	18	.1000	.3162	.2702	.2702	.8543	.6901	.5130
	19	.1000	.3162	.2702	.2702	.8543	.6901	.5064
	20	.1000	.3162	.2702	.2702	.8543	.6901	.5000

Legend:

$c = \sigma_x^2 / \sigma_y^2 = \frac{\text{the variance of (receipts minus removals)}}{\text{the variance of the physical inventory}}$

$p_j; \sigma_y^2 p_j = \text{the variance of } \hat{I}_j, \text{ the minimum variance, unbiased estimate of } \eta_j.$

$E_w, E_z \text{ and } E_u; E_w, E_z \text{ and } E_u \text{ when multiplied by } \epsilon/\sigma_y,$
are the expected values of the t statistics under the alternative hypothesis that ϵ , an isolated amount of material, has been diverted during the n th period.

$n = \text{the number of periods.}$

	j	c	$\sqrt{c}$	p_n	p	E_w	E_z	E_u
n	1	.2000	.4472	1.0000	.3583	.6742	.6742	.6742
	2	.2000	.4472	.5455	.3583	.7560	.6742	.6455
	3	.2000	.4472	.4271	.3583	.7840	.6742	.6202
	4	.2000	.4472	.3854	.3583	.7942	.6742	.5976
	5	.2000	.4472	.3692	.3583	.7983	.6742	.5774
	6	.2000	.4472	.3628	.3583	.7999	.6742	.5590
	7	.2000	.4472	.3601	.3583	.8006	.6742	.5423
	8	.2000	.4472	.3590	.3583	.8009	.6742	.5270
	9	.2000	.4472	.3586	.3583	.8010	.6742	.5130
	10	.2000	.4472	.3584	.3583	.8011	.6742	.5000
	11	.2000	.4472	.3583	.3583	.8011	.6742	.4880
	12	.2000	.4472	.3583	.3583	.8011	.6742	.4767
	13	.2000	.4472	.3583	.3583	.8011	.6742	.4663
	14	.2000	.4472	.3583	.3583	.8011	.6742	.4564
	15	.2000	.4472	.3583	.3583	.8011	.6742	.4472
	16	.2000	.4472	.3583	.3583	.8011	.6742	.4385
	17	.2000	.4472	.3583	.3583	.8011	.6742	.4303
	18	.2000	.4472	.3583	.3583	.8011	.6742	.4226
	19	.2000	.4472	.3583	.3583	.8011	.6742	.4152
	20	.2000	.4472	.3583	.3583	.8011	.6742	.4082

	j	c	$\sqrt{c}$	p_n	p	E_w	E_z	E_u
n	1	.3000	.5477	1.0000	.4179	.6594	.6594	.6594
	2	.3000	.5477	.5652	.4179	.7322	.6594	.6202
	3	.3000	.5477	.4639	.4179	.7530	.6594	.5872
	4	.3000	.5477	.4331	.4179	.7596	.6594	.5590
	5	.3000	.5477	.4230	.4179	.7618	.6594	.5345
	6	.3000	.5477	.4196	.4179	.7626	.6594	.5130
	7	.3000	.5477	.4185	.4179	.7628	.6594	.4939
	8	.3000	.5477	.4181	.4179	.7629	.6594	.4767
	9	.3000	.5477	.4180	.4179	.7629	.6594	.4613
	10	.3000	.5477	.4179	.4179	.7630	.6594	.4472
	11	.3000	.5477	.4179	.4179	.7630	.6594	.4344
	12	.3000	.5477	.4179	.4179	.7630	.6594	.4226
	13	.3000	.5477	.4179	.4179	.7630	.6594	.4117
	14	.3000	.5477	.4179	.4179	.7630	.6594	.4016
	15	.3000	.5477	.4179	.4179	.7630	.6594	.3922
	16	.3000	.5477	.4179	.4179	.7630	.6594	.3835
	17	.3000	.5477	.4179	.4179	.7630	.6594	.3753
	18	.3000	.5477	.4179	.4179	.7630	.6594	.3676
	19	.3000	.5477	.4179	.4179	.7630	.6594	.3604
	20	.3000	.5477	.4179	.4179	.7630	.6594	.3536

	j	c	$\sqrt{c}$	P_n	p	E_w	E_z	E_u
n	1	.4000	.6325	1.0000	.4633	.6455	.6455	.6455
	2	.4000	.6325	.5833	.4633	.7101	.6455	.5976
	3	.4000	.6325	.4958	.4633	.7263	.6455	.5590
	4	.4000	.6325	.4725	.4633	.7308	.6455	.5270
	5	.4000	.6325	.4660	.4633	.7321	.6455	.5000
	6	.4000	.6325	.4641	.4633	.7324	.6455	.4767
	7	.4000	.6325	.4635	.4633	.7325	.6455	.4564
	8	.4000	.6325	.4634	.4633	.7326	.6455	.4385
	9	.4000	.6325	.4633	.4633	.7326	.6455	.4226
	10	.4000	.6325	.4633	.4633	.7326	.6455	.4082
	11	.4000	.6325	.4633	.4633	.7326	.6455	.3953
	12	.4000	.6325	.4633	.4633	.7326	.6455	.3835
	13	.4000	.6325	.4633	.4633	.7326	.6455	.3727
	14	.4000	.6325	.4633	.4633	.7326	.6455	.3627
	15	.4000	.6325	.4633	.4633	.7326	.6455	.3536
	16	.4000	.6325	.4633	.4633	.7326	.6455	.3450
	17	.4000	.6325	.4633	.4633	.7326	.6455	.3371
	18	.4000	.6325	.4633	.4633	.7326	.6455	.3297
	19	.4000	.6325	.4633	.4633	.7326	.6455	.3227
	20	.4000	.6325	.4633	.4633	.7326	.6455	.3162

	j	c	$\sqrt{c}$	P_n	p	E_w	E_z	E_u
n	1	.5000	.7071	1.0000	.5000	.6325	.6325	.6325
	2	.5000	.7071	.6000	.5000	.6901	.6325	.5774
	3	.5000	.7071	.5238	.5000	.7029	.6325	.5345
	4	.5000	.7071	.5059	.5000	.7061	.6325	.5000
	5	.5000	.7071	.5015	.5000	.7068	.6325	.4714
	6	.5000	.7071	.5004	.5000	.7070	.6325	.4472
	7	.5000	.7071	.5001	.5000	.7071	.6325	.4264
	8	.5000	.7071	.5000	.5000	.7071	.6325	.4082
	9	.5000	.7071	.5000	.5000	.7071	.6325	.3922
	10	.5000	.7071	.5000	.5000	.7071	.6325	.3780
	11	.5000	.7071	.5000	.5000	.7071	.6325	.3651
	12	.5000	.7071	.5000	.5000	.7071	.6325	.3536
	13	.5000	.7071	.5000	.5000	.7071	.6325	.3430
	14	.5000	.7071	.5000	.5000	.7071	.6325	.3333
	15	.5000	.7071	.5000	.5000	.7071	.6325	.3244
	16	.5000	.7071	.5000	.5000	.7071	.6325	.3162
	17	.5000	.7071	.5000	.5000	.7071	.6325	.3086
	18	.5000	.7071	.5000	.5000	.7071	.6325	.3015
	19	.5000	.7071	.5000	.5000	.7071	.6325	.2949
	20	.5000	.7071	.5000	.5000	.7071	.6325	.2887

	j	c	$\sqrt{c}$	P_n	p	E_w	E_z	E_u
n	1	.6000	.7746	1.0000	.5307	.6202	.6202	.6202
	2	.6000	.7746	.6154	.5307	.6710	.6202	.5590
	3	.6000	.7746	.5486	.5307	.6822	.6202	.5130
	4	.6000	.7746	.5346	.5307	.6845	.6202	.4767
	5	.6000	.7746	.5315	.5307	.6840	.6202	.4472
	6	.6000	.7746	.5309	.5307	.6851	.6202	.4226
	7	.6000	.7746	.5307	.5307	.6851	.6202	.4016
	8	.6000	.7746	.5307	.5307	.6851	.6202	.3835
	9	.6000	.7746	.5307	.5307	.6851	.6202	.3676
	10	.6000	.7746	.5307	.5307	.6851	.6202	.3536
	11	.6000	.7746	.5307	.5307	.6851	.6202	.3410
	12	.6000	.7746	.5307	.5307	.6851	.6202	.3297
	13	.6000	.7746	.5307	.5307	.6851	.6202	.3194
	14	.6000	.7746	.5307	.5307	.6851	.6202	.3101
	15	.6000	.7746	.5307	.5307	.6851	.6202	.3015
	16	.6000	.7746	.5307	.5307	.6851	.6202	.2936
	17	.6000	.7746	.5307	.5307	.6851	.6202	.2863
	18	.6000	.7746	.5307	.5307	.6851	.6202	.2795
	19	.6000	.7746	.5307	.5307	.6851	.6202	.2732
	20	.6000	.7746	.5307	.5307	.6851	.6202	.2673

	j	c	$\sqrt{c}$	P_n	p	E_w	E_z	E_u
n	1	.7000	.8367	1.0000	.5569	.6086	.6086	.6086
	2	.7000	.8367	.6296	.5569	.6552	.6086	.5423
	3	.7000	.8367	.5707	.5569	.6636	.6086	.4939
	4	.7000	.8367	.5596	.5569	.6652	.6086	.4564
	5	.7000	.8367	.5574	.5569	.6656	.6086	.4264
	6	.7000	.8367	.5570	.5569	.6656	.6086	.4016
	7	.7000	.8367	.5569	.5569	.6656	.6086	.3807
	8	.7000	.8367	.5569	.5569	.6656	.6086	.3627
	9	.7000	.8367	.5569	.5569	.6656	.6086	.3471
	10	.7000	.8367	.5569	.5569	.6656	.6086	.3333
	11	.7000	.8367	.5569	.5569	.6656	.6086	.3211
	12	.7000	.8367	.5569	.5569	.6656	.6086	.3101
	13	.7000	.8367	.5569	.5569	.6656	.6086	.3002
	14	.7000	.8367	.5569	.5569	.6656	.6086	.2911
	15	.7000	.8367	.5569	.5569	.6656	.6086	.2828
	16	.7000	.8367	.5569	.5569	.6656	.6086	.2752
	17	.7000	.8367	.5569	.5569	.6656	.6086	.2682
	18	.7000	.8367	.5569	.5569	.6656	.6086	.2617
	19	.7000	.8367	.5569	.5569	.6656	.6086	.2557
	20	.7000	.8367	.5569	.5569	.6656	.6086	.2500

	j	c	$\sqrt{c}$	P_n	P	E_w	E_z	E_u
n	1	.8000	.8944	1.0000	.5798	.5976	.5976	.5976
	2	.8000	.8944	.6429	.5798	.6398	.5976	.5270
	3	.8000	.8944	.5906	.5798	.6468	.5976	.4767
	4	.8000	.8944	.5817	.5798	.6480	.5976	.4385
	5	.8000	.8944	.5801	.5798	.6482	.5976	.4082
	6	.8000	.8944	.5799	.5798	.6482	.5976	.3835
	7	.8000	.8944	.5798	.5798	.6482	.5976	.3627
	8	.8000	.8944	.5798	.5798	.6482	.5976	.3450
	9	.8000	.8944	.5798	.5798	.6482	.5976	.3297
	10	.8000	.8944	.5798	.5798	.6482	.5976	.3162
	11	.8000	.8944	.5798	.5798	.6482	.5976	.3043
	12	.8000	.8944	.5798	.5798	.6482	.5976	.2936
	13	.8000	.8944	.5798	.5798	.6482	.5976	.2840
	14	.8000	.8944	.5798	.5798	.6482	.5976	.2752
	15	.8000	.8944	.5798	.5798	.6482	.5976	.2673
	16	.8000	.8944	.5798	.5798	.6482	.5976	.2599
	17	.8000	.8944	.5798	.5798	.6482	.5976	.2532
	18	.8000	.8944	.5798	.5798	.6482	.5976	.2469
	19	.8000	.8944	.5798	.5798	.6482	.5976	.2411
	20	.8000	.8944	.5798	.5798	.6482	.5976	.2357

	j	c	$\sqrt{c}$	P_n	P	E_w	E_z	E_u
n	1	.9000	.9487	1.0000	.6000	.5872	.5872	.5872
	2	.9000	.9487	.6552	.6000	.6256	.5872	.5130
	3	.9000	.9487	.6086	.6000	.6314	.5872	.4613
	4	.9000	.9487	.6014	.6000	.6323	.5872	.4226
	5	.9000	.9487	.6002	.6000	.6324	.5872	.3922
	6	.9000	.9487	.6000	.6000	.6325	.5872	.3676
	7	.9000	.9487	.6000	.6000	.6325	.5872	.3471
	8	.9000	.9487	.6000	.6000	.6325	.5872	.3297
	9	.9000	.9487	.6000	.6000	.6325	.5872	.3147
	10	.9000	.9487	.6000	.6000	.6325	.5872	.3015
	11	.9000	.9487	.6000	.6000	.6325	.5872	.2899
	12	.9000	.9487	.6000	.6000	.6325	.5872	.2795
	13	.9000	.9487	.6000	.6000	.6325	.5872	.2702
	14	.9000	.9487	.6000	.6000	.6325	.5872	.2617
	15	.9000	.9487	.6000	.6000	.6325	.5872	.2540
	16	.9000	.9487	.6000	.6000	.6325	.5872	.2469
	17	.9000	.9487	.6000	.6000	.6325	.5872	.2404
	18	.9000	.9487	.6000	.6000	.6325	.5872	.2344
	19	.9000	.9487	.6000	.6000	.6325	.5872	.2288
	20	.9000	.9487	.6000	.6000	.6325	.5872	.2236

	j	c	$\sqrt{c}$	P_n	P	E_w	E_z	E_u
	1	1.0000	1.0000	1.0000	.6180	.5774	.5774	.5774
	2	1.0000	1.0000	.6667	.6180	.6124	.5774	.5000
	3	1.0000	1.0000	.6250	.6180	.6172	.5774	.4472
	4	1.0000	1.0000	.6190	.6180	.6179	.5774	.4082
	5	1.0000	1.0000	.6182	.6180	.6180	.5774	.3780
	6	1.0000	1.0000	.6181	.6180	.6180	.5774	.3536
	7	1.0000	1.0000	.6180	.6180	.6180	.5774	.3333
	8	1.0000	1.0000	.6180	.6180	.6180	.5774	.3162
	9	1.0000	1.0000	.6180	.6180	.6180	.5774	.3015
n	10	1.0000	1.0000	.6180	.6180	.6180	.5774	.2887
	11	1.0000	1.0000	.6180	.6180	.6180	.5774	.2774
	12	1.0000	1.0000	.6180	.6180	.6180	.5774	.2673
	13	1.0000	1.0000	.6180	.6180	.6180	.5774	.2582
	14	1.0000	1.0000	.6180	.6180	.6180	.5774	.2500
	15	1.0000	1.0000	.6180	.6180	.6180	.5774	.2425
	16	1.0000	1.0000	.6180	.6180	.6180	.5774	.2357
	17	1.0000	1.0000	.6180	.6180	.6180	.5774	.2294
	18	1.0000	1.0000	.6180	.6180	.6180	.5774	.2236
	19	1.0000	1.0000	.6180	.6180	.6180	.5774	.2182
	20	1.0000	1.0000	.6180	.6180	.6180	.5774	.2132

8. ALTERNATIVE HYPOTHESIS II - SYSTEMATIC DIVERSION

It is assumed that the amount ϵ is being diverted every period. The statistics to assess whether a diversion has occurred are t_w , t_z and t_u as before. Values in the tables are given as functions of $c = \sigma_x^2 / \sigma_y^2$ and j , the number of periods. The power of the three methods to detect diversion are given by E_w , E_z and E_u as before. Here

$$E(t_w) = \left(\frac{\epsilon}{\sigma_y} \right) E_w$$

$$E(t_z) = \left(\frac{\epsilon}{\sigma_y} \right) E_z$$

$$E(t_u) = \left(\frac{\epsilon}{\sigma_y} \right) E_u$$

	j	c	$\sqrt{c}$	p_n	p	E_w	E_z	E_u
n	1	.1000	.3162	1.0000	.2702	.6901	.6901	.6901
	2	.1000	.3162	.5238	.2702	1.1584	.6901	1.3484
	3	.1000	.3162	.3842	.2702	1.5671	.6901	1.9781
	4	.1000	.3162	.3262	.2702	1.9144	.6901	2.5820
	5	.1000	.3162	.2988	.2702	2.2009	.6901	3.1623
	6	.1000	.3162	.2851	.2702	2.4308	.6901	3.7210
	7	.1000	.3162	.2780	.2702	2.6113	.6901	4.2601
	8	.1000	.3162	.2743	.2702	2.7505	.6901	4.7809
	9	.1000	.3162	.2724	.2702	2.8564	.6901	5.2850
	10	.1000	.3162	.2713	.2702	2.9361	.6901	5.7735
	11	.1000	.3162	.2708	.2702	2.9956	.6901	6.2476
	12	.1000	.3162	.2705	.2702	3.0397	.6901	6.7082
	13	.1000	.3162	.2703	.2702	3.0724	.6901	7.1563
	14	.1000	.3162	.2703	.2702	3.0964	.6901	7.5926
	15	.1000	.3162	.2702	.2702	3.1141	.6901	8.0178
	16	.1000	.3162	.2702	.2702	3.1270	.6901	8.4327
	17	.1000	.3162	.2702	.2702	3.1365	.6901	8.8379
	18	.1000	.3162	.2702	.2702	3.1434	.6901	9.2338
	19	.1000	.3162	.2702	.2702	3.1485	.6901	9.6210
	20	.1000	.3162	.2702	.2702	3.1522	.6901	10.0000

Legend:

$$c = \sigma_x^2 / \sigma_y^2 = \frac{\text{the variance of (receipts minus removals)}}{\text{the variance of the physical inventory}}$$

$$p_j \sigma_y^2 p_j = \text{the variance of } \hat{I}_j, \text{ the minimum variance unbiased estimate of } \eta_j.$$

E_w , E_z and E_u ; E_w , E_z and E_u , when multiplied by ϵ/σ_y , are the expected values under the alternative hypothesis that ϵ units/period are systematically diverted.

n = the number of periods.

	j	c	$\sqrt{c}$	P_n	p	E_w	E_z	E_u
	1	.2000	.4472	1.0000	.3583	.6742	.6742	.6742
	2	.2000	.4472	.5455	.3583	1.1010	.6742	1.2910
	3	.2000	.4472	.4271	.3583	1.4373	.6742	1.8605
	4	.2000	.4472	.3554	.3583	1.6891	.6742	2.3905
	5	.2000	.4472	.3692	.3583	1.8691	.6742	2.8868
	6	.2000	.4472	.3628	.3583	1.9935	.6742	3.3541
	7	.2000	.4472	.3601	.3583	2.0773	.6742	3.7963
	8	.2000	.4472	.3590	.3583	2.1320	.6742	4.2164
	9	.2000	.4472	.3586	.3583	2.1693	.6742	4.6169
n	10	.2000	.4472	.3584	.3583	2.1930	.6742	5.0000
	11	.2000	.4472	.3583	.3583	2.2083	.6742	5.3675
	12	.2000	.4472	.3583	.3583	2.2182	.6742	5.7208
	13	.2000	.4472	.3583	.3583	2.2246	.6742	6.0613
	14	.2000	.4472	.3583	.3583	2.2287	.6742	6.3901
	15	.2000	.4472	.3583	.3583	2.2313	.6742	6.7082
	16	.2000	.4472	.3583	.3583	2.2330	.6742	7.0165
	17	.2000	.4472	.3583	.3583	2.2341	.6742	7.3156
	18	.2000	.4472	.3583	.3583	2.2348	.6742	7.6064
	19	.2000	.4472	.3583	.3583	2.2353	.6742	7.8893
	20	.2000	.4472	.3583	.3583	2.2356	.6742	8.1650

	j	c	$\sqrt{c}$	P_n	p	E_w	E_z	E_u
	1	.3000	.5477	1.0000	.4179	.6594	.6594	.6594
	2	.3000	.5477	.5652	.4179	1.0506	.6594	1.2403
	3	.3000	.5477	.4639	.4179	1.3321	.6594	1.7617
	4	.3000	.5477	.4331	.4179	1.5215	.6594	2.2361
	5	.3000	.5477	.4230	.4179	1.6423	.6594	2.6726
	6	.3000	.5477	.4196	.4179	1.7167	.6594	3.0779
	7	.3000	.5477	.4185	.4179	1.7615	.6594	3.4571
	8	.3000	.5477	.4181	.4179	1.7881	.6594	3.8139
	9	.3000	.5477	.4180	.4179	1.8037	.6594	4.1514
n	10	.3000	.5477	.4179	.4179	1.8120	.6594	4.4721
	11	.3000	.5477	.4179	.4179	1.8182	.6594	4.7781
	12	.3000	.5477	.4179	.4179	1.8214	.6594	5.0709
	13	.3000	.5477	.4179	.4179	1.8232	.6594	5.3520
	14	.3000	.5477	.4179	.4179	1.8243	.6594	5.6225
	15	.3000	.5477	.4179	.4179	1.8249	.6594	5.8835
	16	.3000	.5477	.4179	.4179	1.8252	.6594	6.1357
	17	.3000	.5477	.4179	.4179	1.8254	.6594	6.3800
	18	.3000	.5477	.4179	.4179	1.8256	.6594	6.6169
	19	.3000	.5477	.4179	.4179	1.8256	.6594	6.8471
	20	.3000	.5477	.4179	.4179	1.8257	.6594	7.0711

	j	c	$\sqrt{c}$	P_n	p	E_w	E_z	E_u
n	1	.4000	.6325	1.0000	.4633	.6455	.6455	.6455
	2	.4000	.6325	.5833	.4633	1.0059	.6455	1.1952
	3	.4000	.6325	.4958	.4633	1.2451	.6455	1.6771
	4	.4000	.6325	.4725	.4633	1.3916	.6455	2.1082
	5	.4000	.6325	.4660	.4633	1.4765	.6455	2.5000
	6	.4000	.6325	.4641	.4633	1.5241	.6455	2.8604
	7	.4000	.6325	.4635	.4633	1.5503	.6455	3.1950
	8	.4000	.6325	.4634	.4633	1.5645	.6455	3.5082
	9	.4000	.6325	.4633	.4633	1.5722	.6455	3.8032
	10	.4000	.6325	.4633	.4633	1.5763	.6455	4.0825
	11	.4000	.6325	.4633	.4633	1.5786	.6455	4.3481
	12	.4000	.6325	.4633	.4633	1.5799	.6455	4.6018
	13	.4000	.6325	.4633	.4633	1.5804	.6455	4.8448
	14	.4000	.6325	.4633	.4633	1.5807	.6455	5.0783
	15	.4000	.6325	.4633	.4633	1.5809	.6455	5.3033
	16	.4000	.6325	.4633	.4633	1.5810	.6455	5.5205
	17	.4000	.6325	.4633	.4633	1.5811	.6455	5.7307
	18	.4000	.6325	.4633	.4633	1.5811	.6455	5.9344
	19	.4000	.6325	.4633	.4633	1.5811	.6455	6.1322
	20	.4000	.6325	.4633	.4633	1.5811	.6455	6.3246

	j	c	$\sqrt{c}$	P_n	p	E_w	E_z	E_u
n	1	.5000	.7071	1.0000	.5000	.6325	.6325	.6325
	2	.5000	.7071	.6000	.5000	.9661	.6325	1.1547
	3	.5000	.7071	.5238	.5000	1.1716	.6325	1.6036
	4	.5000	.7071	.5059	.5000	1.2875	.6325	2.0000
	5	.5000	.7071	.5015	.5000	1.3494	.6325	2.3570
	6	.5000	.7071	.5004	.5000	1.3815	.6325	2.6833
	7	.5000	.7071	.5001	.5000	1.3977	.6325	2.9848
	8	.5000	.7071	.5000	.5000	1.4060	.6325	3.2660
	9	.5000	.7071	.5000	.5000	1.4101	.6325	3.5301
	10	.5000	.7071	.5000	.5000	1.4121	.6325	3.7796
	11	.5000	.7071	.5000	.5000	1.4132	.6325	4.0166
	12	.5000	.7071	.5000	.5000	1.4137	.6325	4.2426
	13	.5000	.7071	.5000	.5000	1.4140	.6325	4.4590
	14	.5000	.7071	.5000	.5000	1.4141	.6325	4.6667
	15	.5000	.7071	.5000	.5000	1.4141	.6325	4.8666
	16	.5000	.7071	.5000	.5000	1.4142	.6325	5.0596
	17	.5000	.7071	.5000	.5000	1.4142	.6325	5.2463
	18	.5000	.7071	.5000	.5000	1.4142	.6325	5.4272
	19	.5000	.7071	.5000	.5000	1.4142	.6325	5.6028
	20	.5000	.7071	.5000	.5000	1.4142	.6325	5.7735

	j	c	$\sqrt{c}$	P_n	p	E_w	E_z	E_u
n	1	.6000	.7746	1.0000	.5307	.6202	.6202	.6202
	2	.6000	.7746	.6154	.5307	.9303	.6202	1.1180
	3	.6000	.7746	.5486	.5307	1.1086	.6202	1.5390
	4	.6000	.7746	.5346	.5307	1.2021	.6202	1.9069
	5	.6000	.7746	.5315	.5307	1.2485	.6202	2.2361
	6	.6000	.7746	.5309	.5307	1.2709	.6202	2.5355
	7	.6000	.7746	.5307	.5307	1.2815	.6202	2.8113
	8	.6000	.7746	.5307	.5307	1.2865	.6202	3.0679
	9	.6000	.7746	.5307	.5307	1.2889	.6202	3.3085
	10	.6000	.7746	.5307	.5307	1.2900	.6202	3.5355
	11	.6000	.7746	.5307	.5307	1.2905	.6202	3.7510
	12	.6000	.7746	.5307	.5307	1.2908	.6202	3.9563
	13	.6000	.7746	.5307	.5307	1.2909	.6202	4.1527
	14	.6000	.7746	.5307	.5307	1.2909	.6202	4.3412
	15	.6000	.7746	.5307	.5307	1.2910	.6202	4.5227
	16	.6000	.7746	.5307	.5307	1.2910	.6202	4.6978
	17	.6000	.7746	.5307	.5307	1.2910	.6202	4.8671
	18	.6000	.7746	.5307	.5307	1.2910	.6202	5.0312
	19	.6000	.7746	.5307	.5307	1.2910	.6202	5.1904
	20	.6000	.7746	.5307	.5307	1.2910	.6202	5.3452

	j	c	$\sqrt{c}$	P_n	p	E_w	E_z	E_u
n	1	.7000	.8367	1.0000	.5569	.6086	.6086	.6086
	2	.7000	.8367	.6296	.5569	.8978	.6086	1.0847
	3	.7000	.8367	.5707	.5569	1.0540	.6086	1.4816
	4	.7000	.8367	.5596	.5569	1.1305	.6086	1.8257
	5	.7000	.8367	.5574	.5569	1.1661	.6086	2.1320
	6	.7000	.8367	.5570	.5569	1.1822	.6086	2.4097
	7	.7000	.8367	.5569	.5569	1.1895	.6086	2.6649
	8	.7000	.8367	.5569	.5569	1.1927	.6086	2.9019
	9	.7000	.8367	.5569	.5569	1.1941	.6086	3.1239
	10	.7000	.8367	.5569	.5569	1.1947	.6086	3.3333
	11	.7000	.8367	.5569	.5569	1.1950	.6086	3.5319
	12	.7000	.8367	.5569	.5569	1.1951	.6086	3.7210
	13	.7000	.8367	.5569	.5569	1.1952	.6086	3.9020
	14	.7000	.8367	.5569	.5569	1.1952	.6086	4.0756
	15	.7000	.8367	.5569	.5569	1.1952	.6086	4.2426
	16	.7000	.8367	.5569	.5569	1.1952	.6086	4.4039
	17	.7000	.8367	.5569	.5569	1.1952	.6086	4.5598
	18	.7000	.8367	.5569	.5569	1.1952	.6086	4.7108
	19	.7000	.8367	.5569	.5569	1.1952	.6086	4.8574
	20	.7000	.8367	.5569	.5569	1.1952	.6086	5.0000

	j	c	$\sqrt{c}$	P_n	p	E_w	E_z	E_u
	1	.8000	.8944	1.0000	.5798	.5976	.5976	.5976
	2	.8000	.8944	.6429	.5798	.8683	.5976	1.0541
	3	.8000	.8944	.5906	.5798	1.0061	.5976	1.4302
	4	.8000	.8944	.5817	.5798	1.0696	.5976	1.7541
	5	.8000	.8944	.5801	.5798	1.0974	.5976	2.0412
	6	.8000	.8944	.5799	.5798	1.1093	.5976	2.3009
	7	.8000	.8944	.5798	.5798	1.1144	.5976	2.5392
	8	.8000	.8944	.5798	.5798	1.1165	.5976	2.7603
	9	.8000	.8944	.5798	.5798	1.1174	.5976	2.9672
n	10	.8000	.8944	.5798	.5798	1.1178	.5976	3.1623
	11	.8000	.8944	.5798	.5798	1.1179	.5976	3.3472
	12	.8000	.8944	.5798	.5798	1.1180	.5976	3.5233
	13	.8000	.8944	.5798	.5798	1.1180	.5976	3.6918
	14	.8000	.8944	.5798	.5798	1.1180	.5976	3.8534
	15	.8000	.8944	.5798	.5798	1.1180	.5976	4.0089
	16	.8000	.8944	.5798	.5798	1.1180	.5976	4.1590
	17	.8000	.8944	.5798	.5798	1.1180	.5976	4.3041
	18	.8000	.8944	.5798	.5798	1.1180	.5976	4.4448
	19	.8000	.8944	.5798	.5798	1.1180	.5976	4.5813
	20	.8000	.8944	.5798	.5798	1.1180	.5976	4.7140

	j	c	$\sqrt{c}$	P_n	p	E_w	E_z	E_u
	1	.9000	.9487	1.0000	.6000	.5872	.5872	.5872
	2	.9000	.9487	.6552	.6000	.8413	.5872	1.0260
	3	.9000	.9487	.6086	.6000	.9637	.5872	1.3838
	4	.9000	.9487	.6014	.6000	1.0170	.5872	1.6903
	5	.9000	.9487	.6002	.6000	1.0391	.5872	1.9612
	6	.9000	.9487	.6000	.6000	1.0481	.5872	2.2056
	7	.9000	.9487	.6000	.6000	1.0517	.5872	2.4297
	8	.9000	.9487	.6000	.6000	1.0531	.5872	2.6375
	9	.9000	.9487	.6000	.6000	1.0537	.5872	2.8319
n	10	.9000	.9487	.6000	.6000	1.0539	.5872	3.0151
	11	.9000	.9487	.6000	.6000	1.0540	.5872	3.1887
	12	.9000	.9487	.6000	.6000	1.0541	.5872	3.3541
	13	.9000	.9487	.6000	.6000	1.0541	.5872	3.5122
	14	.9000	.9487	.6000	.6000	1.0541	.5872	3.6640
	15	.9000	.9487	.6000	.6000	1.0541	.5872	3.8100
	16	.9000	.9487	.6000	.6000	1.0541	.5872	3.9509
	17	.9000	.9487	.6000	.6000	1.0541	.5872	4.0872
	18	.9000	.9487	.6000	.6000	1.0541	.5872	4.2193
	19	.9000	.9487	.6000	.6000	1.0541	.5872	4.3475
	20	.9000	.9487	.6000	.6000	1.0541	.5872	4.4721

n	j	c	$\sqrt{c}$	P_n	p	E_w	E_z	E_u
	1	1.0000	1.0000	1.0000	.6180	.5774	.5774	.5774
	2	1.0000	1.0000	.8667	.6180	.8165	.5774	1.0000
	3	1.0000	1.0000	.6250	.6180	.9258	.5774	1.3416
	4	1.0000	1.0000	.6190	.6180	.9710	.5774	1.6330
	5	1.0000	1.0000	.6182	.6180	.9888	.5774	1.8898
	6	1.0000	1.0000	.6181	.6180	.9957	.5774	2.1213
	7	1.0000	1.0000	.6180	.6180	.9984	.5774	2.3333
	8	1.0000	1.0000	.6180	.6180	.9994	.5774	2.5298
	9	1.0000	1.0000	.6180	.6180	.9998	.5774	2.7136
	10	1.0000	1.0000	.6180	.6180	.9999	.5774	2.8868
	11	1.0000	1.0000	.6180	.6180	1.0000	.5774	3.0509
	12	1.0000	1.0000	.6180	.6180	1.0000	.5774	3.2071
	13	1.0000	1.0000	.6180	.6180	1.0000	.5774	3.3566
	14	1.0000	1.0000	.6180	.6180	1.0000	.5774	3.5000
	15	1.0000	1.0000	.6180	.6180	1.0000	.5774	3.6380
	16	1.0000	1.0000	.6180	.6180	1.0000	.5774	3.7712
	17	1.0000	1.0000	.6180	.6180	1.0000	.5774	3.9001
	18	1.0000	1.0000	.6180	.6180	1.0000	.5774	4.0249
	19	1.0000	1.0000	.6180	.6180	1.0000	.5774	4.1461
	20	1.0000	1.0000	.6180	.6180	1.0000	.5774	4.2640

Suppose $n = 5$. Then the situation can be summed as follows.

		E_w	E_z	E_u
$c = 0.2$	isolated	0.80	0.67	0.58
	systematic	1.87	0.67	2.89
$c = 0.4$	isolated	0.73	0.65	0.50
	systematic	1.48	0.65	2.50
$c = 0.6$	isolated	0.68	0.62	0.45
	systematic	1.25	0.62	2.24
$c = 0.8$	isolated	0.65	0.60	0.41
	systematic	1.10	0.60	2.04
$c = 1.0$	isolated	0.62	0.58	0.38
	systematic	0.99	0.58	1.89

Conclusions from this Table

- For isolated diversion $E_w \geq E_z \geq E_u$.
- For systematic diversion $E_u \geq E_w \geq E_z$.
- There is no generally best procedure against both types of alternative procedures.
- Letting the rank = 1, 2, 3, the average rank over the two alternative hypotheses is $E_w = 1 \frac{1}{2}$, $E_u = 2$, $E_z = 2 \frac{1}{2}$.
- If the conditions in real life approximate the assumptions pertinent to the development of the theory it would appear that the statistic t_w is the most robust, that it would be a good idea in practice to retain a good long-term and a good short-term t statistic to detect diversion.

Suppose that the material in a material balance area has been in control up to and through the first $n-k$ periods. In period $n-k+1$ and in all the remaining k periods thereafter the amount ϵ units is diverted. It is a question posing interest and insight to ask just what value k must be as a function of c in order that the long-term MUF, $y_{n-k} + x_{n-k+1} + \dots + x_{n+1} - y_{n+1}$, has the same power to detect a loss as the MUF based

on the MVUE of η_n . It is assumed that n and k and $n-k$ are of sufficient magnitude to use p , the limiting value of p_j as $j \rightarrow \infty$. Then the two MUF's will have equivalent power to detect when

$$\frac{1}{p\sqrt{p+c+1}} = \frac{k}{\sqrt{1+kc}}.$$

Utilizing the relationship that $p^2 + pc - c = 0 \rightarrow p^2 = c(1-p)$ and solving for k gives

$$k = 1/2 + \sqrt{1/4 + 1/c}$$

or

$$c = \frac{1}{k(k-1)}.$$

$k = 2$, when $c = 1/2$, $k = 3$ when $c = 1/6$, $k = 4$ when $c = 1/12$. Thus considering the MUF values as the indicators of a process in which the material is to be tangibly controlled, it would be useful to retain two MUF's, one based on the short-term MUF and the other a longer-term MUF whose length can be estimated by the k value in the above formula.

9. AN EXAMPLE OF OPTIMIZATION BY VARIANCE MINIMIZATION

The objective of this section is to indicate procedures by which it is possible to obtain an estimate, $\hat{T}$ say, of the amount of material in an inventory in such a manner that $\sigma_{\hat{T}}^2$, the variance of $\hat{T}$, is a minimum for a given amount of effort. The conditions are as follows:

- There are s strata of material
- There are N_i items in the i^{th} stratum
- The measurement x_{ij} of the j^{th} item in the i^{th} stratum is an unbiased estimate of ξ_{ij} , the actual amount of material in the item.

- Let e_i denote the amount of effort to obtain a measurement of an item from the i^{th} stratum. e_i can be written in terms of dollars, manhours, or other variables which are functions of effort.
- σ_{2j}^2 is the process variance of the values of ξ_{ij} over the N_i items in the i^{th} stratum. σ_{1j}^2 is the measurement variance in measuring ξ_{ij} . σ_{1j}^2 can be composed of two variability sources, one occurring in sampling and the other resulting from the imprecision of analytical determinations.

It is desired to determine the number of samples n_i to take from the i^{th} stratum in order to minimize the variance of $\hat{T}$, the estimate of the total amount of material in the inventory. (Note: this may involve variables such as $x_{ij} = w_{ij}c_{ij}f_{ij}$, where w_{ij} , c_{ij} and f_{ij} denote respectively weight, concentration, and isotopic ratio values.) Then

- $\bar{x}_i = \sum_{j=1}^{n_i} x_{ij}/n_i$ is the average of the sample in the i^{th} stratum.
- $N_i \bar{x}_i$ is the estimate of the total amount of material in the i^{th} stratum.
- $\hat{T} = \sum_{i=1}^S N_i \bar{x}_i$ is the estimate of the total amount of material in the inventory.
- The variance of $\hat{T}$, as an estimate of the total amount of material in the inventory, is given by

$$\begin{aligned}
\sigma_{\hat{T}}^2 &= \sum_{i=1}^S N_i^2 \sigma_{\bar{x}_i}^2 = \sum_{i=1}^S \left(\frac{N_i}{n_i} \right)^2 \sigma_{n_i}^2 \\
&\quad \left(\sum x_{ij} \right) \\
&= \sum_{i=1}^S N_i^2 \frac{\sigma_{x_i}^2}{n_i} \\
&= \sum_{i=1}^S N_i^2 \left[\frac{\sigma_{1i}^2}{n_i} + \frac{\sigma_{2i}^2}{n_i} \left(1 - \frac{n_i - 1}{N_i - 1} \right) \right]
\end{aligned}$$

e. The total effort $E = \sum_{i=1}^S n_i e_i$ is constant. This is the constraint.

Solution is by the method of Lagrangian multipliers.

Let $H = \sigma_{\hat{T}}^2 + 2\lambda \left(\sum_{i=1}^S n_i e_i - E \right)$. Then H is minimized. Set $\partial H / \partial n_i = 0$, $i = 1, \dots, s$. Set $\partial H / \partial \lambda = 0$, and solve for n_i in terms of N_i , σ_{1i}^2 and σ_{2i}^2 . One difficulty here is that the resulting equations are difficult to solve exactly unless $(n_i - 1)/(N_i - 1)$ is replaced by n_i/N_i . The solution then is

$$\begin{aligned}
n_i &= \frac{N_i \sigma_{3i} E}{\sqrt{e_i} \sum_{i=1}^S \sqrt{e_i} N_i \sigma_{3i}} \\
&= \frac{k N_i \sigma_{3i}}{\sqrt{e_i}}
\end{aligned}$$

where $\sigma_{3i}^2 = \sigma_{1i}^2 + \sigma_{2i}^2$

and $k = \frac{E}{\sum_{i=1}^S e_i N_i \sigma_{3i}}$.

The variance is approximately equal to

$$\begin{aligned}\sigma_{\hat{T}}^2 &= \sum_{i=1}^s \frac{N_i^2 \sigma_{3i}^2}{n_i} \\ &= \frac{\sum_{i=1}^s N_i \sigma_{3i} \sqrt{e_i}}{k} \\ &= E/k^2.\end{aligned}$$

Example 1

E = 300 units.

<u>i</u>	<u>N_i</u>	<u>σ_{3i}</u>	<u>e_i</u>	<u>N_i σ_{3i} √e_i</u>
1	50	1.1	10	173.925
2	80	2.0	5	357.771
3	100	0.4	5	89.771
4	200	2.8	3	969.948
5	500	0.8	8	1131.371
				<u>2722.458</u>

$$k = 0.300/2722.458 = 0.11019453.$$

$$\text{Then in general } n_i = \frac{N_i \sigma_{3i} k}{\sqrt{e_i}}$$

$$= (N_i \sigma_{3i} \sqrt{e_i}) k / e_i$$

to simplify the computations since $N_i \sigma_{3i} \sqrt{e_i}$ are already computed

<u>i</u>	<u>n_i</u>
1	1.9166
2	7.8849
3	1.9712
4	35.6277
5	15.5839

$$\text{As a check on the computations } \sum_{i=1}^5 n_i e_i = 300.00.$$

$$\begin{aligned}\text{Var } \hat{T} &= \sigma_{\hat{T}}^2 = E/k^2 = 300/0.11019453^2 \\ &= 24705.93 \text{ units}^2.\end{aligned}$$

The variance formula can be checked by

$$\begin{aligned}\sigma_{\hat{T}}^2 &= \sum_{i=1}^5 (N_i \sigma_{3i})^2 / n_i \\ &= 24705.86 \text{ units}^2.\end{aligned}$$

In practice, of course, integral-valued sample numbers must be used. Then let $n_i = 2, 9, 2, 36, 16$ ($i = 1, \dots, 5$). For these values of n_i , $E = 306$ and $\sigma_{\hat{T}}^2 = 24,224$.

The exact variance formula is

$$\sigma_{\hat{T}}^2 = \sum \frac{N_i^2}{n_i} \left(\sigma_{1i}^2 + \sigma_{2i}^2 \right) \left(1 - \frac{n_i - 1}{N_i - 1} \right)$$

Let σ_{3i}^2 be composed as follows.

<u>i</u>	<u>σ_{3i}^2</u>	<u>σ_{1i}^2</u>	<u>σ_{2i}^2</u>
1	1.21	0.21	1.00
2	4.00	0.50	3.50
3	0.16	0.10	0.06
4	7.84	0.84	7.00
5	0.64	0.20	0.44

Then $\sigma_{\hat{T}}^2 = 22373 \text{ units}^2$ exactly. $\sigma_{\hat{T}}^2$ is less than $\sigma_{\hat{T}}^2$ since larger n_i values were chosen. The conclusions from this example are under the following assumed conditions:

- The number in a random sample in a stratum should be taken directly proportional to the size of the stratum,
- The number in a random sample in a stratum should be taken directly proportional to the value σ_{3i} which is equal to the square root of the measurement plus process variances, and

- The number in a random sample in a stratum should be taken inversely proportional to the square root of the amount of effort required to obtain a measurement of a unit in the stratum. (Note here that effort can be made synonymous with manhours or cost.)

In other words an analyst observes the manner of relationship between the sample size in a stratum and the stratum size, the process and measurement variation, and the effort required to obtain an observation; and then adjusts the number in the different strata to accrue to the specified amount of effort.

Example 2

In the previous example no explicit provision was made for the worth or strategic importance of the materials. One method of accomplishing this is as follows. In this example the amount of material is considered on a units-of-strategic-value scale. Thus if a_i units of the type of material in stratum i are of strategic value then each unit of material in the stratum is equal to a_i^{-1} units of strategic value. Then one simply makes the transformation $x'_i = x_i/a_i$ and $\sigma_{3i}^2 = \sigma_{3i}^2/a_i^2$. Let E be fixed as before. Let $\hat{B}$ denote the estimate of the total amount of Strategic Value Units (SVU) in the inventory. Then it follows that the number of samples taken in the i^{th} stratum in order to minimize the variance of $\hat{B}$, subject to the constraint of a fixed amount of effort, is

$$n_i = \frac{N_i \sigma_{3i}^2 E}{\sqrt{e_i} a_i \sum_{i=1}^S \frac{N_i \sigma_{3i}^2 \sqrt{e_i}}{a_i}}$$

with variance $\sigma_{\hat{B}}^2 = E/k'^2$

$$\text{where } k' = \frac{E}{\sum_{i=1}^S \frac{N_i \sigma_{3i} \sqrt{e_i}}{a_i}}$$

Here in addition to the considerations in Example 1, the number of items to take in a random sample in the i^{th} stratum is to be inversely proportional to the number of units of material it takes to make a weapon. This procedure then minimizes the variance of the estimate of the amount of strategic material in the inventory for a given fixed amount of effort. As a numerical example assume that N_i , σ_{3i} , and e_i are the same as in Example 1 and that it takes $a_i = 5, 10, 50, 800$ and 1000 ($i = 1, \dots, 5$) units of materials in each stratum to be of strategic value. Let

$$B = \sum_{i=1}^S N_i \bar{x}_i / a_i$$

estimate the total number of SVU in the inventory. Then

<u>i</u>	<u>a_i</u>	<u>N_iσ_{3i}√e_i/a_i</u>	<u>e_i</u>
1	5	34.785	10
2	10	35.777	5
3	50	1.789	5
4	800	1.212	3
5	1000	<u>1.131</u>	8
		74.695	

$$\text{and } k = 300/74.695 = 4.016345$$

$$\text{and } \sigma_B^2 = 300/k^2 = 18.598 \text{ (strategic value units)}^2$$

$$n_i = (N_i \sigma_{3i} \sqrt{e_i} / a_i) k / e_i$$

$$n_1 = 13.97$$

$$n_2 = 28.74$$

$$n_3 = 1.44$$

$$n_4 = 1.62$$

$$n_5 = 0.57$$

As a computational check $E = \sum_{i=1}^5 n_i e_i = 300.02$. The sample numbers must be rounded off (up) to integers.

In this case it is more important to take into account the Finite Population Correction (FPC) factors. Let σ_{3i}^2 be composed as before. Then $\sigma_{\hat{B}}^2$ is actually equal to $\sigma_{\hat{B}}^2 = 13.56$ (SVU)². FPC factors will be ignored in the rest of this section, however, for purposes of exposition.

Since it takes 300 units of effort to obtain the estimates $\hat{T}$ or $\hat{B}$, it is interesting to see what the variance would be in obtaining an estimate $\hat{B}'$ say in (SVU)² using the same sample sizes which were used to estimate $\hat{T}$. Then

$$\hat{B}' = \sum_{i=1}^S \frac{N_i \bar{x}_i}{a_i}$$

where the n_i values are determined from Example 1. Then

$$\begin{aligned} \sigma_{\hat{B}'}^2 &= \sum_{i=1}^5 \frac{N_i^2 \sigma_{x_i}^2}{a_i^2 n_i} \\ &= 95.95 \text{ (SVU)}^2. \end{aligned}$$

Since $\sigma_{\hat{B}}^2 = 18.60$, the use of nonoptimum sample sizes will, in this instance, increase the variance of the estimate of the amount of SVU in the inventory by

$$(95.95/18.60 - 1) \times 100\% = 416\%.$$

Conclusion: When the ground rules of the game are changed to estimate the amount of strategic material, the sampling pattern, as regards the number of units to take in the random samples in the different strata, can vary radically.

Example 3

Suppose that a variance of

$$\sigma_{\hat{T}}^2 = 18.60 \text{ (SVU)}^2$$

is too large. It is desired to reduce the variance of the estimate of the total amount of strategic value material to 6.0 (SVU)^2 . Then the optimum way of proceeding is to take

$$n_i' = n_i (18.60/6.0)$$

samples from the i^{th} stratum. Then

$$n_i = 43.3, 89.1, 4.5, 5.0, 1.8 \text{ (} i = 1, \dots, 5 \text{)}$$

and

$$E' = (18.6/6.0) E = 930$$

units of effort. Then too, of course,

$$\sigma_{\hat{T}}^2 = 6.0 \text{ (SVU)}^2.$$

The following general comments are pertinent here.

- It is vital to efficient safeguards procedures that good records be kept in order that the variance components such as σ_{1i}^2 and σ_{2i}^2 can be estimated.
- It would be very desirable to have a statistician available to the surveys and analyze the results with the objectives of estimation and estimation variance minimization.
- There are some extensions which can be made on these techniques by taking into account such considerations as the portability of the material, the accessibility of the material, the point in the fuel cycle where the inventory fits in, and the possibility of antecedent and subsequent measurements of the same materials as part of the overall fuel cycle. These considerations will, however, take far

more study and possibly more complicated procedures in order to be introduced into the subject of optimum sample numbers.

10. VARIANCE MINIMIZATION BY AN OPTIMUM ALLOCATION OF EFFORT

If a particular nuclide form which occurs in an MBA is divertible then the authentication of the presence of all items of this nuclide form should receive high priority in the inventory verification procedures.

The verification of the presence of all such items by sampling within the item and measurement of the sample contents does not ensure that diversion has not occurred although the verification procedure creates upper bounds for the amount of material which may be diverted and greatly increases the difficulties and risks which the would-be diverter encounters.

This section is concerned with the best procedure to use in obtaining and analyzing sample results for a unit container.

If it is feasible to make the material in the unit more homogeneous by some form of mixing, agitation or sparging, this is advisable.

Three approaches to the problem are:

- s samples are taken and m measurements are made for each sample;
- s samples are taken and composited, one measurement is made of the composite, m such composites are taken;
- s samples are taken and composited and m measurements are made of the overall composite.

It is assumed that the sample units all have equal volume and measurements are made of the concentration η_i .

Situation 1

The estimate of the average concentration is $\hat{\eta}$ where

$$\hat{\eta} = \sum \hat{\eta}_i / s,$$

where $\hat{\eta}_i$ is the estimate of the concentration of the i^{th} sample. The variance of $\hat{\eta}$ is

$$\sigma_{\hat{\eta}}^2 = \frac{\sigma_s^2}{s} + \frac{\sigma_m^2}{ms}$$

where σ_s^2 and σ_m^2 are the sampling and measurement variance components.

It is desired to determine the values of s and m which minimize the variance subject to a constraint on the total cost of making the s measurements and taking the ms samples.

If a sample costs c_s \$ and a measurement costs c_m \$, then the total cost is $C = s c_s + ms c_m$. Letting

$$H = \sigma_{\hat{\eta}}^2 - \lambda (C - s c_s - ms c_m)$$

and setting $\frac{\partial H}{\partial s}$, $\frac{\partial H}{\partial m}$ and $\frac{\partial H}{\partial \lambda} = 0$, one finds that

$$\sqrt{\lambda} = \frac{\sigma_s \sqrt{c_s} + \sigma_m \sqrt{c_m}}{C},$$

$$s = \frac{\sigma_s}{\sqrt{c_s} \sqrt{\lambda}},$$

$$m = \frac{\sigma_m}{\sigma_s} \sqrt{\frac{c_s}{c_m}}$$

$$\sigma_{\hat{\eta}}^2 = \lambda C = (\sigma_s \sqrt{c_s} + \sigma_m \sqrt{c_m})^2 / C$$

Situation 2

If $\sigma_s \sqrt{c_m} > \sigma_m \sqrt{c_s}$ then m turns out to be a fractional number less than 1. In this case if

$$\sigma_s \sqrt{c_m} \gg \sigma_m \sqrt{c_s}$$

It is optimum to take more than one sample, to composite and then to take a measurement of the composite. This is done m times. The variance would then be

$$\sigma_{\hat{\eta}}^2 = \frac{\sigma_m^2}{m} + \frac{\sigma_s^2}{ms}$$

and the same sort of minimization procedure would obtain. Then

$$m = \frac{\sigma_m}{\sqrt{c_m} \sqrt{\lambda}}$$

$$s = \frac{\sigma_s}{\sigma_m} \sqrt{\frac{c_m}{c_s}}$$

$$\sqrt{\lambda} = (\sigma_s \sqrt{c_s} + \sigma_m \sqrt{c_m}) / C$$

as before and

$$\sigma_{\hat{\eta}}^2 = \lambda C$$

as before.

Situation 3

Another possibility is to composite all samples. Then the variance is

$$\sigma_{\hat{\eta}}^2 = \frac{\sigma_s^2}{s} + \frac{\sigma_m^2}{m}$$

$$\sqrt{\lambda} = \frac{\sigma_s \sqrt{c_s} + \sigma_m \sqrt{c_m}}{C}$$

$$s = \frac{\sigma_m}{\sqrt{\lambda c_s}}$$

$$m = \frac{\sigma_m}{\sqrt{\lambda c_m}}$$

$$\sigma_{\hat{\eta}}^2 = \lambda C$$

Note that under this procedure $\sigma_{\hat{\eta}}^2 = \lambda C$ is the same.

Example

$$c_s = 2, c_m = 5$$

$$\sigma_s = 3, \sigma_m = 8$$

$$\sqrt{\lambda} = (8\sqrt{5} + 3\sqrt{2})/50 = 0.44262$$

$$s = \frac{3}{0.44262\sqrt{2}} = 4.793$$

$$\frac{m}{s} = \frac{8}{3} \left(\frac{2}{5}\right)^{1/2} = 1.687$$

$$\sigma_{\hat{\eta}}^2 = (0.44262)^2 (50)$$

$$\begin{aligned}
 &= \frac{9}{4793} + \frac{64}{8.084} \\
 &= 9.8 \text{ units}^2.
 \end{aligned}$$

In practice, of course, the practitioner must make some decision as to the manner of converting the s and m values to integers. For contrast, consider an example where the effort is badly apportioned

$$\begin{aligned}
 c_s &= 5, \quad c_m = 2 \\
 \sigma_s &= 8, \quad \sigma_m = 3
 \end{aligned}$$

Then the minimum variance procedure is to take

$$s = \frac{8}{3} \sqrt{\frac{2}{5}} = 1.7$$

samples per measurement and this procedure should be repeated

$$m = \frac{3}{\sqrt{\lambda} \sqrt{2}} = 4.8$$

times resulting in a variance of 9.8 units^2 . As a contrast to this see what variance results from the unskillful allocation of sampling and measurement variance. If one sample is taken and all the rest of the effort is expended on making measurements then nine measurements would be taken. The variance of the average of these nine measurements is

$$\sigma_{\hat{\eta}}^2 = \frac{64}{1} + \frac{9}{9} = 65 \text{ units}^2$$

as compared to the minimum variance estimate of 9.8 units^2 .

Conclusions

- (a) If it is feasible to composite all samples then $\sigma_{\hat{\eta}}^2$ is minimized in an optimum manner when the s samples are taken according to

$$s = \frac{\sigma_s}{\sqrt{\lambda} \sqrt{c_s}}.$$

The m measurements of the overall composite are taken according to

$$m = \frac{\sigma_m}{\sqrt{\lambda} \sqrt{c_m}}.$$

- (b) If, for practical reasons, it is a better procedure to measure smaller composites, and if $\sigma_s \sqrt{c_m} \ll \sigma_m \sqrt{c_s}$ then

$$s = \frac{\sigma_s}{\sqrt{\lambda} \sqrt{c_s}}$$

samples are taken as before and

$$m = \frac{\sigma_m}{\sigma_s} \sqrt{\frac{c_s}{c_m}}$$

measurements are taken per sample.

- (c) If $\sigma_s \sqrt{c_m} \gg \sigma_m \sqrt{c_s}$, and smaller sample composites are desired, then s samples are taken and a measurement is made of the composite. This is repeated m times. Then m in this case is

$$m = \frac{\sigma_m}{\sqrt{\lambda} \sqrt{c_m}}, \quad s = \frac{\sigma_s}{\sigma_m} \sqrt{\frac{c_m}{c_s}}.$$

All three methods will result in a variance of λC .

- (d) The unthinking apportionment of effort as shown in an example indicates how poor the estimation results can be when the process is little understood and when procedures to extract the maximum information from the data are not followed.

It was not necessary to develop the three cases separately as the results follow easily, one from another, by analogy. However, the procedure of considering three situations is a useful illustrative tool.

GLOSSARY OF TERMS AND NOTATIONS

- η_i = the true amount of material in an inventory at the beginning of the i^{th} time period.
- δ_i = the true amount of material received minus the true amount of material removed from the material balance area during the i^{th} time period by legal routes.
- ϵ_i = the amount of material diverted in the i^{th} period.
- y_i = the measurement of η_i , the amount of material in the inventory at the beginning of the i^{th} period, i.e., the physical inventory estimate.
- x_i = the estimate of δ_i , the input minus the output for the i^{th} time period.
- I_i = the minimum variance unbiased estimate (MVUE) of η_i , the amount of material in the inventory at the beginning of the i^{th} time period.

VARIOUS MUF'S

- $z_i = (y_i + x_i) - y_{i+1}$
 = the usual definition of MUF, the estimate of $(\eta_i + \delta_{i+1}) - \eta_{i+1} = \epsilon_{i+1}$, the amount of material unaccountably lost during the i^{th} time period.

$$u_n = (y_1 + \sum_1^n x_i) - y_{n+1}$$

= a long-term MUF which estimates $\eta_1 + \sum_1^n \delta_i - \eta_{n+1}$, the amount of material unaccountably lost during the first n time periods. u_n is an estimate of $\eta_1 + \sum \delta_i - \eta_{n+1} = \sum_1^n \epsilon_i$.

$$w_n = I_n + x_n - y_{n+1} = \text{the MUF based on the MVUE of } y_n.$$

$$t_w = w_j / \sigma_{w_j}$$

$$t_z = z_j / \sigma_{z_j}$$

$$t_u = u_j / \sigma_{u_j}$$

$$E_w = E(t_w) \frac{\sigma_y}{\epsilon}$$

$$E_z = E(t_z) \frac{\sigma_y}{\epsilon}$$

$$E_u = E(t_u) \frac{\sigma_y}{\epsilon}$$

where $E(t_w)$, $E(t_z)$ and $E(t_u)$ denote the expected values of the t statistics.

MBA = Material Balance Accounting

MUF = Material Unaccounted For

MVUE = Minimum Variance, Unbiased Estimator (as used here
it is I_n , the estimate of η_n)

BIBLIOGRAPHY

1. K. B. Stewart. B-PID and Inventory Estimates with Minimum Variance, HW-56536. Available from Clearinghouse for Federal Scientific and Technical Information, Springfield, Virginia, 1958.
2. K. B. Stewart. Some Statistical Aspects of B-PID's and Ending Inventories, AEC and Contractor SS Materials Management Meeting, May 25-28, 1959, TID-7581. Available from Clearinghouse for Federal Scientific and Technical Information, Springfield, Virginia.
3. K. B. Stewart. "A New Weighted Average" to be published in Technometrics.
4. C. A. Bennett and N. L. Franklin. Statistical Analysis in Chemistry and the Chemical Industry. Wiley and Sons, 1954, pp. 367-368.

APPENDIX

APPENDIX

APPENDIX

A.1 A MVUE Approach When $\sigma_{y_1}^2 = 0$.

In the development of the MVUE of η_n denoted as I_n in Section 6, it was assumed that the first physical measurement y_1 has a variance of σ_y^2 . In many instances where the inventories are cleaned out at the end of a campaign, $\sigma_{y_1}^2 = 0$. Under this assumption it can be shown that

$$\begin{aligned} I'_n &= u_{n-1} y_n + v_{n-1} (\hat{I}_{n-1} + x_n) \\ &= u_{n-1} y_n + \sum_{j=1}^{n-1} u_{j-1} \prod_{k=j}^{n-1} v_k \theta_j \end{aligned}$$

has minimum variance where

$$\begin{aligned} \theta_j &= y_j + x_j + \dots + x_n \\ u_i &= \underbrace{\frac{1}{1+c} \frac{1}{1+c} \frac{1}{1+c} \frac{1}{1+c} \dots \frac{1}{1+c}}_{i \text{ c's}}, \quad v_i = 1 - u_i. \end{aligned}$$

Furthermore

$$\sigma_{I'_n}^2 = \sigma_y^2 u_{n-1}$$

u_{n-1} converges to p as defined in Section 6, but the convergence is from below p . The rate of convergence depends on c --the smaller c is, the slower is the convergence.

For example with $c = 0.1$

i	u_i	i	u_i
1	0.09091	7	0.26452
2	0.16031	8	0.26714
3	0.20654	9	0.26854
4	0.23462	10	0.26930
5	0.25072	15	0.27012
6	0.25966	∞	0.27016

The following table shows the value of u_n as a function of c and n .

c

	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
1	.09091	.16667	.23077	.28571	.33333	.37500	.41176	.44444	.47368	.50000
2	.16031	.26829	.34673	.40678	.45455	.49367	.52646	.55446	.57871	.60000
3	.20654	.31894	.39274	.44653	.48837	.52237	.55086	.57527	.59656	.61538
4	.23462	.34164	.40924	.45844	.49708	.52883	.55572	.57900	.59945	.61765
5	.25072	.35134	.41494	.46192	.49927	.53026	.55668	.57965	.59991	.61798
6	.25966	.35540	.41689	.46292	.49982	.53057	.55687	.57977	.59999	.61803
7	.26452	.35708	.41755	.46321	.49995	.53064	.55691	.57979	.60000	.61803
8	.26714	.35777	.41778	.46329	.49999	.53066	.55692	.57980	.60000	.61803
9	.26855	.35806	.41785	.46332	.50000	.53066	.55692	.57980	.60000	.61803
10	.26930	.35818	.41788	.46332	.50000	.53066	.55692	.57980	.60000	.61803
11	.26970	.35822	.41789	.46332	.50000	.53066	.55692	.57980	.60000	.61803
12	.26991	.35824	.41789	.46332	.50000	.53066	.55692	.57980	.60000	.61803
13	.27003	.35825	.41789	.46332	.50000	.53066	.55692	.57980	.60000	.61803
14	.27009	.35826	.41789	.46332	.50000	.53066	.55692	.57980	.60000	.61803
15	.27012	.35826	.41789	.46332	.50000	.53066	.55692	.57980	.60000	.61803
16	.27014	.35826	.41789	.46332	.50000	.53066	.55692	.57980	.60000	.61803
17	.27015	.35826	.41789	.46332	.50000	.53066	.55692	.57980	.60000	.61803
18	.27015	.35826	.41789	.46332	.50000	.53066	.55692	.57980	.60000	.61803
19	.27015	.35826	.41789	.46332	.50000	.53066	.55692	.57980	.60000	.61803
20	.27015	.35826	.41789	.46332	.50000	.53066	.55692	.57980	.60000	.61803
21	.27016	.35826	.41789	.46332	.50000	.53066	.55692	.57980	.60000	.61803
22	.27016	.35826	.41789	.46332	.50000	.53066	.55692	.57980	.60000	.61803

This table illustrates the speed of convergence of u_j to p when the variance of the first inventory measurement is zero.

A.2 Development of an Inventory Estimate Based on the Constraint That All MUF's Have Zero Covariance.

Let

$J_1 = y_1$ = the first physical inventory measurement

$$J_2 = p'_1 y_2 + (1 - p'_1) (J_1 + x_2)$$

Define the first MUF as

$$\theta_1 = J_1 + x_1 - y_2$$

Let

$$J_3 = p'_2 y_3 + (1 - p'_2) (J_2 + x_2)$$

$$\theta_2 = J_2 + x_2 - y_3 = p'_1 y_2 + (1 - p'_1) (J_1 + x_1) - y_3$$

then

$$\begin{aligned} \text{Cov } \theta_1 \theta_2 &= - p'_1 \sigma_y^2 + (1 - p'_1) \sigma^2 (J_1 + x_1) \\ &= - p'_1 \sigma_y^2 + (1 - p'_1) (\sigma_y^2 + \sigma_x^2). \end{aligned}$$

Set $\text{Cov } \theta_1 \theta_2 = 0$. Then

$$p'_1 = \frac{\sigma_y^2 + \sigma_x^2}{\sigma_y^2 + \sigma_x^2 + \sigma_y^2} = \frac{1 + c}{2 + c} = \frac{1}{1 + \frac{1}{c}}$$

so that $p'_1 = p_1$ as defined before in Section 6.

Assume that $p'_{i-1} = p_{i-1}$. Then

$$J_{i+1} = p'_i y_{i+1} + (1 - p'_i) (J_i + x_i)$$

$$\theta_i = J_i + x_i - y_{i+1}$$

$$\theta_{i+1} = J_{i+1} + x_{i+1} - y_{i+2}$$

$$\begin{aligned}
\text{Cov } \theta_{i+1} \theta_i &= \text{Cov } [p'_i y_{i+1} + (1 - p'_i) (J_i + x_i) + x_{i+2} - y_{i+2}] \\
&\quad [J_i + x_i - y_{i+1}] \\
&= - p'_i \sigma_y^2 + (1 - p'_i) (p_{i-1} \sigma_y^2 + c \sigma_y^2).
\end{aligned}$$

$\text{Cov } \theta_{i+1} \theta_i = 0$, when

$$\begin{aligned}
p'_i &= \frac{c \sigma_y^2 + p_{i-1} \sigma_y^2}{c \sigma_y^2 + p_{i-1} \sigma_y^2 + \sigma_y^2} \\
&= \frac{c + p_{i-1}}{c + p_{i-1} + 1} \\
&= \frac{1}{1 + \frac{1}{c + p_{i-1}}} \\
&= p_i
\end{aligned}$$

as defined in Section 6. Thus the procedure of determining unbiased estimates of η_i say, based on the condition that successive MUF estimates have zero covariance lead to the same inventory estimate.

A.3 The MVUE When $\sigma_{x_i}^2$ Is Variable

If $\sigma_{y_i}^2 = \sigma_y^2$ for all i , but $\sigma_{x_i}^2$ depends on i , then the MVUE of y is given by

$$\begin{aligned}
L_n &= p_{n-1} y_n + (1 - p_{n-1}) L_{n-1} \\
&= p_{n-1} y_n + \sum_{j=1}^{n-1} p_{j-1} \prod_{k=j}^{n-1} (1 - p_k) \theta_j
\end{aligned}$$

where $\theta_j = y_j + x_j + \cdots + x_n$

$$p_0 = 1$$

$$p_i = \frac{1}{1+c_i} \frac{1}{c_i+1} \frac{1}{1+\cdots} \frac{1}{c_1+1} \quad (\text{with } i \text{ } c_i \text{ values})$$

$$c_i = \sigma_{x_i}^2 / \sigma_y^2$$

In addition the variance of $\sigma_{L_n}^2$ is given by

$$\sigma_y^2 p_{n-1}$$

There is another approach to this problem which leads to the same estimator. The element σ_{ij} in the covariance matrix of the θ_j values is given by

$$\sigma_{ij} = \delta_{ij} \sigma_y^2 + \sum_{k=\max(i,j)}^{k=n} \sigma_{x_k}^2$$

where δ_{ij} is Kronecker's δ . Let Z denote a column vector of one's. Then the generalized least squares estimate of η_n is given by

$$\hat{u}_n = (Z' \Sigma_{\theta}^{-1} Z)^{-1} Z' \Sigma_{\theta}^{-1} \theta$$

with variance of $\sigma_{\hat{u}_n}^2 = (Z' \Sigma_{\theta}^{-1} Z)^{-1}$, where θ and Σ_{θ} are respectively a column vector and the covariance matrix of the θ values. It can be shown that

$$\hat{u}_n = L_n$$

and that $\sigma_{\hat{u}_n}^2 = p_{n-1} \sigma_y^2$.

Example

$$\text{If } \sigma_{y_1}^2 = 1, \sigma_{x_1}^2 = 3$$

$$\sigma_{x_2}^2 = 1$$

$$\sigma_{x_3}^2 = 1/2$$

$$\text{then } p_1 = \frac{1 + 3}{2 + 3} = 4/5$$

$$p_2 = \frac{4/5 + 1}{4/5 + 2} = \frac{9}{14}$$

$$p_3 = \frac{9/14 + 1/2}{9/14 + 1/2 + 1} = \frac{8}{15}$$

$$\text{Now } k_1 = (1 - p_1)(1 - p_2)(1 - p_3) = 1/30$$

$$k_2 = p_1(1 - p_2)(1 - p_3) = 4/30$$

$$k_3 = p_2(1 - p_3) = 9/30$$

$$k_4 = p_3 = 16/30$$

$$k_1 + k_2 + k_3 + k_4 = 1, \text{ and } \sigma_{L_4}^2 = p_3 \sigma_y^2 = 8/15.$$

By the generalized least-squares approach

$$\sum_{\theta} = \begin{pmatrix} 5 & 1/2 & 1 & 1/2 & 1/2 & 0 \\ 1 & 1/2 & 2 & 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 1 & 1/2 & 1 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\sum_{\theta}^{-1} = \frac{1}{64} \begin{pmatrix} 14 & -8 & -2 & 0 \\ -8 & 32 & -8 & 0 \\ -2 & -8 & 46 & 0 \\ 0 & 0 & 0 & 64 \end{pmatrix}$$

$$\hat{u} = (Z' \Sigma_{\theta}^{-1} Z)^{-1} Z' \theta = \frac{\theta_1}{30} + \frac{4\theta_2}{30} + \frac{9\theta_3}{30} + \frac{16\theta_4}{30}$$

$$\sigma_{\hat{u}}^2 = (Z' \Sigma_{\theta}^{-1} Z)^{-1} = 16/30 = 8/15$$

as with the other approach.

In practice the easiest approach when confronted with this situation is to calculate the L_i values successively as needed by the iterative formula

$$L_i = p_{i-1} + (1 - p_{i-1}) L_{i-1}$$

$$\begin{aligned} p_i &= \frac{p_{i-1} + \left(\sigma_{x_i}^2 / \sigma_y^2 \right)}{p_{i-1} + \left(\sigma_{x_i}^2 / \sigma_y^2 \right) + 1} \\ &= \frac{1}{1 + \left[1 / p_{i-1} + \left(\sigma_{x_i}^2 / \sigma_y^2 \right) \right]} \end{aligned}$$

A.4 The MVUE When $\sigma_{y_i}^2$ and $\sigma_{x_i}^2$ Are Both Variable

Suppose that $\sigma_{y_i}^2$ and $\sigma_{x_i}^2$ are both variable from one period to another. It can be shown that the value

$$I''_{n+1} = p''_n y_{n+1} + \sum_{j=1}^n p''_{j-1} \theta_j \prod_{k=j}^n (1 - p''_k)$$

is the MVUE of η_{n+1} where

$$p_1'' = \frac{\sigma_{x_1}^2 + \sigma_{y_1}^2}{\sigma_{x_1}^2 + \sigma_{y_1}^2 + \sigma_{x_2}^2}$$

and

$$p_i'' = \frac{p_{i-1}'' \sigma_{y_{i-1}}^2 + \sigma_{x_{i-1}}^2}{p_{i-1}'' \sigma_{y_{i-1}}^2 + \sigma_{x_{i-1}}^2 + \sigma_{y_i}^2}$$

where $\sigma_{I_{n+1}}^2 = \sigma_{y_{n+1}}^2 p_n''$.

As an example consider the following

<u>i</u>	<u>$\sigma_{y_i}^2$</u>	<u>$\sigma_{x_i}^2$</u>
1	1	1
2	2	3
3	2	2
4	3	

Then $p_1'' = \frac{1 + 1}{1 + 1 + 2} = \frac{1}{2}$, $\sigma_{I_2}^2 = 2\left(\frac{1}{2}\right) = 1$

$$p_2'' = \frac{2\left(\frac{1}{2}\right) + 3}{2\left(\frac{1}{2}\right) + 3 + 2} = \frac{2}{3}, \quad \sigma_{I_3}^2 = 2\left(\frac{2}{3}\right) = \frac{4}{3}$$

$$p_3'' = \frac{2\left(\frac{2}{3}\right) + 2}{2\left(\frac{2}{3}\right) + 2 + 3} = \frac{10}{19}, \quad \sigma_{I_4}^2 = 3\left(\frac{10}{19}\right) = \frac{30}{19}.$$

Then $q_1 q_2 q_3 = k_1 = 3/38$

$$p_1 q_2 q_3 = k_2 = 3/38$$

$$p_2 q_3 = k_3 = 12/38$$

$$p_3 = k_4 = 20/38$$

and $\sigma_{\hat{\mu}_4}^2 = \frac{30}{19}$, $\sum k_i = 1$.

As a check the generalized least squares solution can be employed. Here

$$\Sigma_{\theta} = \begin{pmatrix} 7 & 5 & 2 & 0 \\ 5 & 7 & 2 & 0 \\ 2 & 2 & 4 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

$$\Sigma_{\theta}^{-1} = \frac{1}{240} \begin{pmatrix} 72 & -48 & -12 & 0 \\ -48 & 72 & -12 & 0 \\ -12 & -12 & 72 & 0 \\ 0 & 0 & 0 & 80 \end{pmatrix}$$

Then $\hat{\mu}_4 = (Z' \Sigma_{\theta}^{-1} Z)^{-1} Z' \Sigma_{\theta}^{-1} \theta$

where Z is a column vector of one's, and θ is a column vector of θ_i values.

$$(Z' \Sigma_{\theta}^{-1} Z)^{-1} = \frac{240}{152} = \frac{30}{19} = \sigma_{\hat{\mu}_4}^2 \text{ as before}$$

$$\text{and } (Z' \Sigma_{\theta}^{-1} Z)^{-1} Z' \Sigma_{\theta}^{-1} \theta = \frac{30}{19} \left(\frac{12\theta_1 + 12\theta_2 + 48\theta_3 + 80\theta_4}{240} \right).$$

Then

$$\begin{aligned} k_1 &= 3/38 \\ k_2 &= 3/38 \\ k_3 &= 12/38 \\ k_4 &= 20/38 \end{aligned}$$

as before.

DISTRIBUTION

No. of
Copies
OFFSITE

1	<u>AEC Chicago Patent Group</u> G. H. Lee
25	<u>AEC Division of International Affairs</u> W. A. Strauser
5	<u>AEC Division of Nuclear Materials Safeguards</u> R. G. Page
186	<u>AEC Division of Technical Information Extension</u>
1	<u>AEC New Brunswick Laboratory</u> C. J. Rodden
115	<u>AEC Office of Safeguards and Materials Management</u> G. M. Inman (75) S. C. T. McDowell (30) J. M. Williams (10)
1	<u>Allied Chemical Corporation</u> P. O. Box 35 Florham Park, New Jersey 07932 G. A. Huff
20	<u>Argonne National Laboratory</u> M. A. Kanter
2	<u>Arms Control and Disarmament Agency</u> L. Stieff
10	<u>Brookhaven National Laboratory</u> H. Kouts
1	<u>Idaho Nuclear Corporation</u> F. H. Tingey
1	<u>Kerr McGee Corporation</u> Field Box 315 Crescent, Oklahoma 37028 R. E. Hubert
1	<u>Lawrence Radiation Laboratory, Livermore</u> R. E. Batzel
1	<u>Los Alamos Scientific Laboratory</u> Jane H. Hall

No. of
Copies

- 5 National Bureau of Standards
 G. Suzuki
- 1 Nuclear Audit Testing Co.
 Suite 312, Barr Bldg.
 910 17th NW
 Washington, D. C. 20006
 R. P. Wischow
- 2 Nuclear Fuel Services, Inc.
 Wheaton Plaza Office Building
 Wheaton, Maryland 20902
 R. J. Jones
- 2 Stanford Research Institute
 E. M. Kinderman

ONSITE-HANFORD

- 1 AEC Chicago Patent Group
 R. K. Sharp
- 3 AEC Richland Operations Office
 H. R. Freitag (2)
 C. L. Robinson
- 1 Atlantic Richfield Company
 L. L. Zahn
- 3 Atlantic Richfield Hanford Company
 J. W. Jordan
 ARHCO Files (2)
- 3 Battelle Memorial Institute
- 1 Jersey Nuclear Company
 R. Nilson
- 2 RDT Assistant Director for Pacific Northwest Programs
- 74 Battelle-Northwest
 C. A. Bennett
 R. O. Budd
 A. W. DeMerschman
 K. Drumheller (3)
 J. L. Jaech
 R. A. Schneider (20)
 K. B. Stewart (30)
 L. F. Wirfs (10)
 Technical Publications (2)
 Technical Information Files (5)