IMPLEMENTATION OF MODERN STATISTICAL METHODS FOR IMPROVING THE ACCURACY OF HIGHWAY LABORATORY AND FIELD DATA

— A Manual —

MICHIGAN DEPARTMENT OF TRANSPORTATION

TESTING AND RESEARCH DIVISION
RESEARCH LABORATORY SECTION

TE192 .K8 1982 c. 2
Implementation of modern statistical methods for improving the accuracy of highway laboratory and field

IMPLEMENTATION OF MODERN STATISTICAL METHODS FOR IMPROVING THE ACCURACY OF HIGHWAY LABORATORY AND FIELD DATA

— A Manual —

W. H. Kuo

A Final Report on a Highway Planning and Research Investigation Conducted by the Michigan Department of Transportation in Cooperation with the U. S. Department of Transportation, Federal Highway Administration

Research Laboratory Section
Testing and Research Division
Research Project 78 G-238
Research Report No. R-1184

Lansing, March 1982

ABSTRACT

Accident rate, traffic count, accident severity index, proportion of wet surface accidents, pavement friction coefficients, highway noise levels, aggregate gradations, etc., are parameters commonly used in the transportation field. Values of these parameters estimated from samples are often the basis of information used in administrative decision making such as the development of effective safety improvement and testing programs. For example, parameter estimates are compared with the designed or desired values to recommend whether or not to reconstruct or install median barriers. Another example is the ranking of accident parameter estimates as a priority basis for resurfacing intersections. Also, highway administrators use parameter estimates to develop effective quality control systems for detecting accidents due to assignable factors and, hope then, to optimally allocate available funding accordingly. For those roadways in the neighborhood of the "critical" conditions required, e.g., for the median barrier installation, it is intuitively clear that the small estimation errors could reverse the "correct" decision. In the course of the following discussion, it will be demonstrated that small estimation errors significantly affect the accuracy of priority lists and, consequently, decisions based on these lists. Also, it will be shown that small estimation errors lead to a more lax quality control system which in the long run fails to detect many accidents associated with assignable causes. Thus, it is of fundamental importance to reduce estimation errors as much as possible.

The usual estimate of each parameter is the "best" information available on a single location. However, this estimate is not the best estimate when the purpose is to develop operational programs involving numerous locations, material sources, etc. Methods which further "improve" usual parameter estimates were first made by Stein and later extended by Effron and Morris. We generalize their theoretical results so that these methods can be used for estimating parameters from various types of transportation data. We show mathematically that the proposed estimation methods are always better than the usual ones in terms of the "global" or overall estimation error. Based on the theoretical development of the proposed estimation methods, we provide four computer programs with examples for estimating parameters which are commonly used in the transportation field.

TABLE OF CONTENTS

Page

Page

INTRODUCTION

Highway administrators must make decisions on how best to improve the roadway system. While their decisions are affected by many factors, usually the main constraint is limited funding. With no funding limitation, new and existing roadways certainly could be designed and updated in conformance with ideal design standards. Since no one has the option of unlimited funding, the interest is focused on developing strategies that would allocate available funding to maximally improve the roadway system.

Information needed for developing optimal strategies must be provided. Thus, highway administrators are faced with the problem of obtaining the best estimates of the commonly used parameters such as accident rates, accident severity indices, highway noise levels, pavement friction coefficients, aggregate gradations, etc. Each of the usual estimates of these parameters is the "best information" for decision making on an isolated location. However, usual estimates generally are not preferable at the program or policy level, such as designing an effective intersection resurfacing program, which must concern the estimation error for the program as a whole (global error). Thus, the shift in concern from individual level to program planning can and should carry with it a corresponding shift in estimation method. Estimation methods for this purpose were first made by Stein (1) and later extended by Effron and Morris (2, 3).

A proposal for a Highway Planning and Research project was submitted to the Federal Highway Administration in 1978 to investigate, revise and extend Stein's estimation methods to handle various types of estimation problems in the transportation field. The objectives of the proposal were to develop a manual of procedure and computer programs to enable highway personnel to take advantage of the procedures now available for improving the accuracy of parameter estimates; and, to familiarize researchers with the issues involved in this type of estimation procedure. This report is submitted in fulfillment of the proposal's two objectives.

First, Stein's results must be generalized. The theoretical development of the proposed estimation procedures are presented in Part III for readers who are interested in the technical development. We show in this part that the proposed estimation methods are always better than the usual ones in terms of the global estimation error. The basic concept of the proposed methods are outlined in Part I for readers who are not familiar with statistical theory. In this part, we explain the logic behind the estimation methods proposed for improving various types of transportation data used in estimating the previously mentioned parameters. The important roles of these parameters in developing highway safety improvement, testing, and maintenance programs are demonstrated in Part II. The potential benefits of using the proposed estimates of these important parameters in decision making are also discussed in this part. This explains why the proposed estimation methods are of considerable value to highway administrators.

Four computer programs are written based on theoretical results presented in Part III to handle general estimation problems in transportation. These computer programs are presented in Parts V through VIII. The general guidelines for using these programs are outlined in Part IV. The detailed input and output systems as well as the use of these programs are explained with examples in Parts V through VIII. These programs are:

Poisson - This program, presented in Part V, is designed for simultaneously estimating many Poisson rates such as accident rates and traffic counts at various locations.

Program Proportion - This program, presented in Part VI, is designed for simultaneously estimating many proportions such as proportions of wet accidents at various locations and severity indices of various fixed object accidents.

Program IMEAN - This program, presented in Part VII, is designed for simultaneously estimating many independent mean values such as friction coefficients at various locations.

Program CMEAN - This program, presented in Part VIII, is designed for simultaneously estimating many correlated mean values such as percentages of aggregate passing various sieve sizes.

Although each of the eight parts in this manual is self-contained, we strongly recommend that users always read Parts I and II before using any of the four computer programs in Parts V through VIII.

I
BASIC CONCEPTS OF PROPOSED ESTIMATION PROCEDURES

In all data gathering such as that required in acceptance testing and accident reduction programs, samples are taken to estimate many constants which either measure product quality or describe important processes. These constants are usually termed 'parameters.' Suppose that we are interested in estimating k parameters denoted as $\theta_1, \ldots, \theta_k$. These parameters could be concrete and coarse aggregate properties such as flexural and compressive concrete strengths, bulk specific gravity, percent deleterious particles, crushed material, loss by washing, etc. These parameters could also be accident rates of various locations treated with different types of safety improvements.

Denote X_i to be the usual estimate of θ_i. In the case of concrete, X_i could be the average compressive strength of 20 samples. For the safety improvement example, X_i could be the number of accidents that occur on a location treated with the i-th type of safety improvement. Note that as usually is the case, each parameter estimation can be considered as an isolated subproblem of the larger estimation problem. If the i-th subproblem is the only one of concern, then X_i is the best parameter estimation for decision making purposes. However, accuracy of the program as a whole rather than that of isolated subproblems characterizes many, if not most cases where data are used as a basis for administrative decision making. Thus, the overall accuracy is more important to the development and administration of programs than isolated location accuracy. To increase the global estimation accuracy, we propose the following estimates:

$$\hat{X}_i = c\,U_i + (1 - c)\,X_i\ ,\ i = 1, \ldots, k \tag{1}$$

where U_i is an initial estimate of θ_i and c takes values between 0 and 1, determined basically by U_j and X_j, $j = 1, \ldots, k$. Methods for determining U_i and c will be discussed later. Thus, $\hat{X}_i$ always lies between U_i and X_i. c is generally called the "shrinking factor" for shrinking X_i toward U_i. We will show in Part III that the proposed estimates are always better than the usual ones in terms of the expected sum of squares of residuals (global estimation error) defined below.

$$E \sum_{i=1}^{k} (\hat{X}_i - \theta_i)^2 \leq E \sum_{i=1}^{k} (X_i - \theta_i)^2 \tag{2}$$

The notation 'E' in Eq. (2) stands for the expectation of a random variable. The left and right hand sides of Eq. (2) are the expected global estimation errors made by the proposed and usual methods, respectively. Thus, the global percentage improvement (reduction of estimation error) of the proposed method over the usual one is

$$\text{GPI} = 100 \; \frac{E \sum_{i=1}^{k} (X_i - \theta_i)^2 - E \sum_{i=1}^{k} (\hat{X}_i - \theta_i)^2}{E \sum_{i=1}^{k} (X_i - \theta_i)^2} \qquad (3)$$

We remark that the global superiority does not guarantee individual location or source superiority. That is, the following inequality may not hold for, say, the i-th component of the problem.

$$E \left| \hat{X}_i - \theta_i \right| \leq E \left| X_i - \theta_i \right| \qquad (4)$$

The left and right hand sides of Eq. (4) are the estimation errors made on the i-th parameter by the proposed and usual method, respectively. The question is, "Under what conditions does Eq. (4) hold for each component?" We are also interested in computing the percentage of time that the proposed estimate of the i-th parameter is closer to θ_i than the usual one. This percentage is defined as:

$$PC(i) = 100 \; \Pr \left(\left| \hat{X}_i - \theta_i \right| \leq \left| X_i - \theta_i \right| \right) \qquad (5)$$

The notation "Pr" in Eq. (5) stands for the probability of an event. There is no way to completely answer the above question due to the complex relationships between initial and usual estimates. However, we provide the following simulated results to more or less answer the above question. These simulated results also serve to demonstrate how to obtain good initial and, consequently, final parameter estimates.

Suppose that X_i is the average of 10 samples randomly obtained from a normal population with mean θ_i and variance 10, i = 1, . . . , 10. Thus, the variance of X_i is 1 for every i. For the usual method, it is known that the expected global and individual estimation errors defined in the right hand sides of Eqs. (2) and (3) are 10 and 0.798, respectively. We now provide in Table 1 the simulated results (based on 5000 simulation points) for the case that $U_i = \theta_i + b$, i = 1, . . . , 10. That is, the initial estimate for each parameter is off by the same amount b. When initial estimates are perfect, i.e., $U_i = \theta_i$ ($b \doteq 0$ in Table 1) for every i, the shrinking factor would be very close to 1. Therefore, every $\hat{X}_i$ would be very close to U_i as it should be. We see from Table 1 (Case 1) that the expected individual

Table 1
Simulated Results for The Case Where
$U_i = \theta_i + b$, i=1,....10

Case	b	%-Global Gain(*1)	Individual Error (*2)	Frequency That The Proposed Estimate Is Better (*3)
1	0.0	87.28	0.192	100.00
2	0.2	83.91	0.298	86.00
3	0.4	74.70	0.420	73.80
4	0.6	62.23	0.527	64.60
5	0.8	49.48	0.607	57.90
6	1.0	38.67	0.661	53.40
7	2.0	13.56	0.751	44.40
8	3.0	6.44	0.776	44.50
9	4.0	3.67	0.786	45.60
10	50.0	0.05	0.797	49.20
11	∞	0.00	0.798	100.00

*1 Based on Eq.(3)
*2 The Left Hand Side of Eq.(4)
*3 Based on Eq.(5)

absolute error is 0.192 which is considerably smaller than the usual error (0.798). Moreover, the proposed estimate is more accurate than the usual one almost 100 percent of the time for every component of the problem. It is known that the larger the sample size, the more reliable the sample average. Therefore, the proposed estimation method, together with reasonably good initial estimates, will substantially improve the estimation accuracy in the small sample size estimation problem.

Now, let b increase. That is, initial estimates are systematically getting worse. This will decrease the value of the shrinking factor and, accordingly, decrease the fraction of U_i used to adjust X_i. As shown in Table 1, this process decreases the global improvement from 87.28 percent to zero. Although Eq. (4) still holds, the frequency that the proposed method is more accurate than the usual one could fall below 50 percent for some values of b. This means that whenever improvement is made, it is made in large magnitude but with less frequency. When b becomes very large (Case 10 in Table 1), the shrinking factor would be very close to zero and consequently, every $\hat{X}_i$ would be very close to X_i. Generally, the difference between $\hat{X}_i$ and X_i becomes too small to be practically significant. If b becomes infinitely large (Case 11, Table 1), the shrinking factor is zero and, thus, $\hat{X}_i = X_i$. This is why PC(i) defined in Eq. (5) is 100 for every component of the problem.

Cases in Table 1 are only hypothetical. In reality, the difference between U_i and θ_i would not be the same for all components. That is, initial estimates are better for some components than others. In this case, those components with poor initial estimates would benefit less than those with good initial estimates. In some cases, components with poor initial estimates might even be sacrificed to achieve the global gain. To clarify these statements, we provide in Table 2 the simulated results for the case that $U_i = \theta_i$ for $i = 1, \ldots, 9$ and $U_{10} = \theta_{10} + b$. That is, we have perfect initial estimates for the first nine components, but are off by a constant b in the tenth one. In this case, the first nine components have the same expected individual absolute error which is smaller than the error of the tenth one. We see from Table 2 that the tenth component benefits much less than the first nine components in Case 1 and is actually sacrificed in Case 2 for a 19.88 percent global gain. Tables 1 and 2 indicate that a particularly bad initial estimate hurts itself and also jeopardizes other components that have good initial estimates. Thus, great effort should be made to isolate components for which bad initial estimates seem likely to occur. This is not as much of a problem as it might first appear.

Table 2
Simulated Results for The Case Where
$U_i = \theta_i$, $i=1,\ldots,9$ And $U_{10} = \theta_{10} + b$

b	%-Global Gain(*1)	Expected Individual Absolute Error (*2)		Frequency That The Proposed Estimate Is Better (*3)	
		Components 1 - 9	Component 10	Components 1 - 9	Component 10
1	79.13	0.225	0.778	100	43.60
5	19.88	0.600	1.380	100	28.10

*1 Based on Eq.(3)
*2 The Left Hand Side of Eq.(4)
*3 Based on Eq.(5)

Many sources, such as past experiments or parallel studies, are available to obtain good initial estimates. When a source is not available or reliable for the problem, other means can always be used to obtain initial parameter estimates. For example, the average of the usual estimates can be used as the initial estimate of each parameter. This method is satisfactory if the total parameter variation, $(\theta_1 - \bar{\theta})^2 + \ldots + (\theta_k - \bar{\theta})^2$, is small. $\bar{\theta}$ is the average of k parameters. On the other hand, if parameters can be arranged into groups such that within-group parameter variations are small, group averages will be good initial estimates for parameters belonging to the same group. Examples for this method are presented in Tables 3 and 4.

For the example in Table 3, we first use the average of $X_1, \ldots,$ and X_{10} (Case 1) as the initial estimate for each parameter. This results in a 52.20 percent global gain. Moreover, Eq. (4) holds in the average

Table 3
Simulated Results for Using Group Averages
As Initial Estimates of Parameters

Component i	True Parameter Value	Case 1			Case 2		
		Group No.	Individual Error (*1)	Freq. (*2)	Group No.	Individual Error (*1)	Freq. (*2)
1	-0.8	1	0.654	55.80	1	0.470	76.30
2	-0.7	1	0.600	60.50	1	0.453	77.30
3	-0.6	1	0.572	64.80	1	0.451	77 80
4	-0.5	1	0.522	69.10	1	0.447	76.20
5	-0.4	1	0.500	73.10	1	0.459	75.70
6	0.4	1	0.500	76.20	2	0.472	75.80
7	0.5	1	0.527	70.90	2	0.458	77.30
8	0.6	1	0.560	66.20	2	0.446	77.90
9	0.7	1	0.595	60.90	2	0.453	76.80
10	0.8	1	0.641	56.40	2	0.472	75.80
%-Global Gain(*3)			52.20 %			65.30 %	

*1 Based on Eq.(3)
*2 The Left Hand Side of Eq.(4)
*3 Based on Eq.(5)

sense and also in terms of frequency for every component. Generally, the individual gain is positively correlated with the closeness of θ_i to $\bar{\theta}$. Since $\bar{\theta}$ is 0 in this example, θ_5 and θ_6 are closer to $\bar{\theta}$ than any other parameters and, therefore, gain the most as shown in Table 3. Observe that there is an apparent gap between θ_5 and θ_6. Thus, the total parameter variation can be substantially reduced by separating components into the following two groups:

Group 1 - Components 1 through 5
Group 2 - Components 6 through 10

In this case, the average of $X_1, \ldots, X_5$ is used as the initial estimate of each of the first five parameters. Similarly, the average of $X_6, \ldots, X_{10}$ is used as the initial estimate of each of the last five parameters. As shown in Table 3 (Case 2), this method substantially increases the individual as well as global gains.

Now, let us widen the gap between θ_5 and θ_6. This will increase the total parameter variation. The simulated results for using the average of usual estimates as the initial estimate of every parameter are presented in Table 4. We see from this table that the global gain has been substantially reduced due to the large total parameter variation. For every component, the individual gain still holds in the average sense, but with less frequency. If these components are rearranged in the same way as in the previous example, we would have the same results shown in Table 3 (Case 2). This indicates that proper grouping which reduces the total parameter variation

Table 4
Simulated Results for Using Group Averages
As Initial Estimates of Parameters

Component i	True Parameter Value	Group No.	Expected Individual Absolute Error (*1)	Frequency That The Proposed Estimate Is Better (*2)
1	-2.0	1	0.782	43.90
2	-1.9	1	0.757	44.20
3	-1.8	1	0.763	44.60
4	-1.7	1	0.732	45.60
5	-1.6	1	0.738	46.90
6	1.6	1	0.738	49.60
7	1.7	1	0.748	46.90
8	1.8	1	0.754	45.20
9	1.9	1	0.754	44.90
10	2.0	1	0.771	44.80
The %-Global Gain (*3) is			12.94 %.	

*1 Based on Eq.(3)
*2 The Left Hand Side of Eq.(4)
*3 Based on Eq.(5)

is the key element determining the performance of the proposed method. The general sources for placing members into proper groups are past experiments, parallel studies, and the known physical properties of components.

We have demonstrated that the overall performance of the proposed method is always better than the usual one. Individual superiority is also achieved if initial estimates are reasonably good. However, components with poor initial estimates gain less and could even be sacrificed to achieve global gain. Thus, obtaining good initial estimates emerges as the crucial element in the practical utilization of the proposed estimation method. This matter is now briefly discussed below:

A) Past experiments and parallel studies are generally good sources of initial parameter estimates. Here are some examples: laboratory test results can be used as initial estimates to adjust field test results; previous ADTs adjusted by general traffic trends can be used as initial estimates to adjust current ADTs.

B) In addition to past experiments and parallel studies, the physical properties of components can also be used to place components into proper groups as shown in the previous examples. For example, highway segments can be separated into two groups; intersected and non-intersected roadways. The accident rate of each group is then used as the initial estimate of every segment belonging to the same group. As previously shown, this method is satisfactory so long as the within-group parameter variations are small.

C) Often, the usual estimators are closely related to other variables. As an example (4), the following equation describes very well the relationship between the number of accidents, X_i, and the total vehicle-miles, t_i, for selected locations.

$$X_i = a\, t_i + b\, t_i^2, \quad i = 1, \ldots, k$$

Here, we first obtain the least squares estimates, $\hat{a}$ and $\hat{b}$, of a and b. We then use $U_i = \hat{a}\, t_i + \hat{b}\, t_i^2$ as the initial estimate of the i-th parameter. Similar situations exist almost everywhere in the transportation field. Other examples are: current and previous ADT figures; in-place aggregate gradations before and after compaction; percent wet surface accidents and percent wet time. This method is satisfactory when the relationship among usual estimates and supplemental variables is reasonably linear.

The above methods will be further explored with examples in Part II (also see Part III). Since information for obtaining good initial estimates exists nearly everywhere in the transportation field, we state that, in addition to the guaranteed global superiority, the proposed method provides superior individual estimates as well. The benefits of using better parameter estimates for decision making will be demonstrated in Part II. Thus, the proposed estimation method is of great value to those charged with program administration and responsibility.

As usual, we require that the usual estimate, X_i, is normally distributed with mean θ_i. If the usual estimate is the sample average, the normality requirement is fulfilled once the sample size is fairly large. In traffic accident analysis, the usual estimate is generally related to Poisson or Binomial processes. For example, the number of accidents, X_i, is approximately distributed (4, 5, 6) according to the Poisson law with parameter λ_i. That is, we have

$$\Pr(X_i = j) = e^{-\lambda_i}\, \lambda_i^j / j!, \quad j = 0, 1, \ldots, \tag{6}$$

For this type of estimator, we use Anscombe's transformation (7) to transform X_i to Z_i defined as

$$Z_i = \sqrt{X_i + 0.375} \tag{7}$$

Anscombe has shown that Z_i is distributed rather more normally than X_i when λ_i is large. Moreover, when $\lambda_i \geqq 5$, the mean and variance of Z_i are approximately $\sqrt{\lambda_i}$ and 0.25, respectively. Thus, the proposed method is applicable to the transformed Poisson data.

Denote Y_i to be the number of subcategory accidents, e.g., injury and fatal accidents. It is known (8) that, given that $X_i = n$, Y_i is a binomial random variable with parameter s_i and n. In this case, s_i is the true proportion of injury and fatal accidents. This proportion is also a standard accident severity index. The usual estimate of s_i is

$$S_i = \frac{Y_i}{n} \tag{8}$$

For this type of estimator, we use Anscombe's transformation (2, 7) to transform S_i to Z_i defined as

$$Z_i = \sqrt{n + 0.5}\ \mathrm{Sin}^{-1}\left[\frac{n}{n + 0.75}(2S_i - 1)\right] \tag{9}$$

Again, the above transformed random variable is distributed rather more normally than S_i when ns_i is large. Moreover, when $ns_i > 4$, the variance of Z_i is approximately equal to one. Therefore, the proposed method is also applicable to the transformed proportion data.

Sample proportions, Poisson rates and sample averages are parameters commonly used in the transportation field. The important role of these parameters in the decision making process will be examined in Part II. Examples are also provided to demonstrate the potential benefits of using the proposed methods to estimate these important parameters.

II

POTENTIAL BENEFITS OF USING THE PROPOSED METHODS TO ESTIMATE ACCIDENT RATES, TRAFFIC VOLUMES, ACCIDENT SEVERITY INDICES AND OTHER RELATED PARAMETERS WHICH PLAY IMPORTANT ROLES IN HIGHWAY SAFETY, TESTING, MAINTENANCE AND CONSTRUCTION

In the transportation field, the most commonly used parameters can be grouped into the following three categories according to the distribution types of their usual estimators. These are:

Poisson - accident rate, traffic count, etc.

Binomial - accident severity index, proportion of wet surface accidents, accident reporting level, etc.

Normal - pavement friction coefficient, highway noise level, aggregate gradation, etc.

In each of the following three sections, the roles of the above parameters in designing highway safety improvements and testing programs will be examined. Since the true parameter values are unknown, the estimated values must be used as substitutes. It is clear that the more accurate the estimated values, the better the program utilizing them will be. Moreover, small estimation errors can have considerable impact on decision consequences as will be shown. Thus, it is of fundamental importance to reduce estimation errors as much as possible. We turn now to the role that the proposed method can play in achieving this end.

Potential Benefits of Using the Proposed Method to Estimate Accident Rates and Related Parameters in Highway Safety Improvements

Some accidents can be prevented either by providing motorists with sufficient and effective road guidance information (9), or by improving roadways through reconstruction in conformance with current standards. Since safety improvement program funding is limited, it is neither practical nor possible for program administrators to fully examine and improve every location or roadway segment. Thus, a monitoring and improvement system is needed to identify roadway segments that become hazardous and, therefore, may require treatment. Such a system can be developed through the use of quality control techniques. The first step is to statistically set upper control limits which call attention to locations operating at suspiciously high accident rates. For discussion purposes, we numerically index every location from 1 through L, in the highway system under consideration. If we denote X_i to be the number of accidents occurring at the i-th location, X_i is assumed to be Poisson distributed with parameter (accident rate) λ_i. The upper control limit for declaring at the 100 $(1 - \alpha)$ percent confidence level that the i-th location will operate at an accident rate higher than λ_i is

$$UCL_i = \lambda_i + Z_{1-\alpha} \sqrt{\lambda_i}, \quad i = 1, \ldots, L \tag{10}$$

where $z_{1-\alpha}$ is the upper 100α percentage point of the standard normal distribution. Note that, due to random variation, X_i could exceed UCL_i even when the i-th location is operating at the usual rate λ_i. The probability of this occurrence is α. Thus, we could examine αL locations even when every location is operating at the usual rate. Naturally, time and money spent for examining these locations are essentially wasted. The only way to trim this waste is to specify a smaller α. Unfortunately, this would also reduce our chances of detecting assignable factors causing an increase in accident rates. Generally, α should be chosen as the best compromise of these conflicting concerns so that one maximizes the total accident reduction for the available funding. This is an optimization problem beyond the scope of this study and will not be discussed here.

Since true accident rates are unknown, the estimated rates must be used as substitutes for λ_i, $i = 1, \ldots, L$, in Eq. (10). The more accurate the estimated accident rates, the closer the system operates at the targeted goal. Thus, it is of fundamental importance to obtain the best possible estimates of λ_i, $i = 1, \ldots, L$. Ordinarily, the past accident rate of the i-th location or a 'similar' location is used to estimate λ_i. As previously mentioned, this estimate is satisfactory if the estimation accuracy of an isolated location is our only concern. This is certainly not the case at the transportation policy level. For the whole system, the proposed estimation method can and should be used to obtain better parameter estimates and hence improve overall system performance.

Two examples are provided below which show how to improve usual estimates of accident rates. The benefits of the estimation improvement are also discussed.

Example 1 - The Michigan Department of Transportation has established a complete inventory of every 0.2-mile segment of roadway in the Michigan trunkline system. For demonstration purposes, we shall only consider those rural, two-way, and two-lane segments located in State Highway

Districts 1 through 4. We group these roadway segments into 24 categories according to the following roadway characteristics:

C1) Intersected or non-intersected roadway segment (I or NI)
C2) Tangent or curve roadway segment (T or C)
C3) No passing or passing roadway segments (NP or P)
C4) Lane width (10, 11, or 12).

Table 5
Accident Statistics And ADTs of Twenty-four Categories

Category i	Description	No. of Roadway Segments (N_i)	Average ADT(T_i)	No. of Accidents X_i
1	NI-T-NP-10	471	1420.97	170
2	NI-T-NP-11	539	1605.43	177
3	NI-T-NP-12	370	2562.92	177
4	NI-T- P-10	691	1439.33	193
5	NI-T- P-11	2477	1611.50	739
6	NI-T- P-12	2180	2468.06	895
7	NI-C-NP-10	595	1182.83	213
8	NI-C-NP-11	669	1595.71	288
9	NI-C-NP-12	348	2527.18	237
10	NI-C- P-10	412	1462.57	130
11	NI-C- P-11	1054	1891.70	428
12	NI-C- P-12	1214	2654.28	634
13	I-T-NP-10	160	1511.23	51
14	I-T-NP-11	183	2167.14	92
15	I-T-NP-12	167	2830.36	103
16	I-T- P-10	224	1601.36	80
17	I-T- P-11	740	1798.27	266
18	I-T- P-12	772	2531.98	354
19	I-C-NP-10	249	1408.66	133
20	I-C-NP-11	338	2028.42	207
21	I-C-NP-12	166	2859.70	107
22	I-C- P-10	122	1601.42	62
23	I-C- P-11	410	1930.16	180
24	I-C- P-12	464	2995.37	258

The number of roadway segments in each category are presented in Table 5. Also included in this table are the number of accidents and the average ADT of roadway segments in each category. We are interested in estimating accident rate for each category. For this purpose, we denote N_i to be the number of roadway segments in the i-th category. Also, denote T_i to be the average ADT of these N_i roadway segments. Define X_i to be the total number of accidents occurring on roadway segments of the i-th category during a period of five years (1971 through 1975). Since accidents occur approximately according to the Poisson law, X_i is Poisson distributed with parameter λ_i, where

$$\lambda_i = \theta_i \times 365 \times 5 \times T_i \times 0.2 \times N_i \times 10^{-8} \quad (11)$$

The parameter θ_i in Eq. (11) is the i-th category accident rate which is the number of accidents per 100 million vehicle-miles. This parameter characterizes the accident behavior of every segment in the i-th category. Thus, the usual estimate of θ_i is

$$A_i = \frac{X_i}{365 \times 5 \times T_i \times 0.2 \times N_i} \times 10^8 \tag{12}$$

The usual estimates of accident rates are presented in Tables 6 and 7.

We observe from Table 5 that the only difference between the first and thirteenth categories is that every roadway segment in the thirteenth category contains an intersection. Actually, the above statement speaks for

Table 6
Estimated Results of Categories 1 through 12
Using Four Group Averages

Category i	Poisson Rate: Usual(X_i)	Poisson Rate: Proposed(X_i)	Accident Rate: Usual(A_i)	Accident Rate: Proposed(A_i)	Difference
1	170	161.47	69.59	66.10	3.49
2	177	179.24	56.04	56.75	-0.71
3	177	184.43	51.14	53.29	-2.15
4	193	188.40	53.17	51.90	1.27
5	739	730.88	50.72	50.17	0.56
6	895	914.59	45.58	46.57	-0.99
7	213	207.92	82.92	80.94	1.98
8	288	291.09	73.91	74.71	-0.79
9	237	239.60	73.83	74.64	-0.81
10	130	128.52	59.11	58.44	0.67
11	428	423.18	58.81	58.15	0.66
12	634	643.76	53.91	54.74	-0.83

Table 7
Estimated Results of Categories 13 through 24
(Using Four Group Averages)

Category i	Poisson Rate: Usual(X_i)	Poisson Rate: Proposed(X_i)	Accident Rate: Usual(A_i)	Accident Rate: Proposed(A_i)	Difference
13	51	51.62	57.79	58.50	-0.71
14	92	91.52	63.56	63.23	0.33
15	103	103.51	59.70	60.00	-0.29
16	80	78.54	61.11	59.99	1.11
17	266	265.08	54.77	54.58	0.19
18	354	358.68	49.62	50.28	-0.65
19	133	128.15	103.89	100.10	3.79
20	207	206.31	82.72	82.44	0.28
21	107	112.46	61.76	64.91	-3.15
22	62	58.87	86.94	82.56	4.38
23	180	179.47	62.32	62.13	0.19
24	258	265.97	50.86	52.43	-1.57

every pair of categories indexed by i and i + 12 for every i = 1, . . . , 12. Thus, the rank order of the accident rates of the last 12 categories should be the same as that of the first 12 categories. Close comparison of Tables 6 and 7 reveals that this is not necessarily the case when ranking is based on usual estimates of category accident rates. For example, A_{13} and A_{15} are, respectively, too small and too large relative to A_{14}. We shall see later that this inconsistency is corrected by the proposed estimation method. We now denote $\hat{X}_i$ to be the proposed estimate of λ_i. Then, the corresponding proposed estimate of θ_i is

$$\hat{A}_i = \frac{\hat{X}_i}{365 \times 5 \times T_i \times 0.2 \times N_i} \times 10^8 \qquad (13)$$

The reliability of the usual estimate X_i is positively correlated with accident exposure which is N_iT_i in this case. We see from Table 5 that X_i is much more reliable than X_{i+12} for every i = 1, . . . , 12. Based on the previous discussions (Part I), we decide to first estimate accident rates of Categories 1 through 12 to avoid the contamination of poor initial estimates. We observe from Table 6 that lane width has less influence on accident rate than other roadway characteristics. For this reason we arrange Categories 1 through 12 into the following four groups by pooling across lanes:

Group 1 - Categories 1 through 3 (Tangent, No Passing)
Group 2 - Categories 4 through 6 (Tangent, Passing)
Group 3 - Categories 7 through 9 (Curve, No Passing)
Group 4 - Categories 10 through 12 (Curve, Passing).

The group accident rate is then used as the initial accident rate estimate for each category in the group. Estimated results obtained from using the computer program of Part V are presented in Table 6. The percentage improvement of the proposed method over the usual one is 18.29 percent. We remark that it is generally not easy to substantially improve reliable usual estimates (large number of accidents in this case). Generally speaking, when the usual estimates are reliable, the global estimation gain is low and, consequently, the differences between the usual and proposed estimates are also small as shown in Table 6. However, two points should be made here:

a) Reliable usual estimates are not always available (because of sampling costs and data validity).

b) A small estimation improvement could result in substantial accident reduction in the long run.

The latter point will now be examined in detail.

We see from Table 6 that the usual and proposed accident rates of the first category are 34 (= 170/5) and 32.29 (= 161.47/5) accidents per year, respectively. The true yearly accident rate is unknown. We shall use these rates as reference for setting up the following hypothetical problem to demonstrate the potential benefits of small estimation improvement in detecting accidents due to assignable factors (preventable accidents).

As discussed, the proposed estimate is generally closer to the true value than the usual one. Thus, for exposition purposes, we assume that true accident rate of the i-th location is 30 accidents per year. Under this assumption, the distribution of the yearly accidents is approximately symmetrical with respect to the true accident rate. Therefore, the probability that the usual estimate (X_i) is, say, 32 is almost the same as the probability that it is 28. For this reason, we first consider the following pair of cases:

Case 1.1: $X_i = 32$ and $\hat{X}_i = 31$

and

Case 1.2: $X_i = 28$ and $\hat{X}_i = 29$

That is, the proposed estimate ($\hat{X}_i$) is closer to the true accident rate than the usual one by one accident per year. To examine the potential benefits of this small estimation improvement in terms of detecting accidents due to assignable factors, we set α in Eq. (10) to be 0.05. Thus, $z_{1-\alpha} = 1.645$. For Case 1.1, the usual 95 percent upper control limit is obtained by substituting 32 for λ_i in Eq. (10):

$$UCL_i = 32 + 1.645\sqrt{32} = 41.3$$

This means that the i-th location will be inspected for possible assignable factors if the yearly accidents of this location are greater than 41. Suppose that there are assignable factors causing this location to operate at the higher accident rate λ_i'. Each time the yearly accidents of this location do not go over the upper control limit, this location will not be inspected. Consequently, $\lambda_i' - 30$ extra accidents are expected to occur. It is known that the reciprocal of the probability of the above occurrence is the expected waiting time before inspecting the i-th location. Thus, the expected number of accidents due to assignable factors which would not be detected by the usual upper control limit is

$$Y_i = (\lambda_i' - 30) / \text{Pr (the yearly accidents are over 41 when the new accident rate is } \lambda_i')$$

Similarly, the proposed upper control limit and the expected number of accidents due to assignable factors which would not be detected by the proposed upper control limit are, respectively

$$\hat{UCL}_i = 31 + 1.645 \sqrt{31} = 40.2$$

and

$$\hat{Y}_i = (\lambda_i' - 30) / Pr \text{ (the yearly accidents are over 40 when the new accident rate is } \lambda_i' \text{)}$$

Since $UCL_i > \hat{UCL}_i$, we have $Y_i > \hat{Y}_i$. Therefore, the difference, $Y_i - \hat{Y}_i$, is the benefit of using the proposed upper control limit to detect accidents due to assignable factors. The benefits for various values of λ_i' in this case are presented in Figure 1.

We similarly obtain the following results for Case 1.2:

$$UCL_i = 28 + 1.645 \sqrt{28} = 36.7$$

$$Y_i = (\lambda_i' - 30) / Pr \text{ (the yearly accidents are over 36 when the new accident rate is } \lambda_i' \text{)}$$

$$\hat{UCL}_i = 29 + 1.645 \sqrt{29} = 37.8$$

and

$$\hat{Y}_i = (\lambda_i' - 30) / Pr \text{ (the yearly accidents are over 37 when the new accident rate is } \lambda_i' \text{)}$$

Since $UCL_i < \hat{UCL}_i$, we have $Y_i < \hat{Y}_i$. Therefore, the difference, $\hat{Y}_i - Y_i$, is the benefit of using the usual upper control limit to detect accidents due to assignable factors. The benefits for various values of λ_i' are also presented in Figure 1.

The above pair of cases shows that, whereas the proposed estimate is better than the usual one, using the proposed upper control limit to detect assignable factors is only beneficial in Case 1.1. However, this benefit is much greater than that of Case 1.2. Since both cases have an equal chance of occurrence, the net benefit is one-half of the difference between the two curves in Figure 1. The net benefits for this pair of cases are presented in Figure 2.

When the usual estimate is smaller than the proposed and true accident rate such as in Case 1.2, use of the usual rather than proposed upper control limit to detect assignable factors will cause the unnecessary inspection

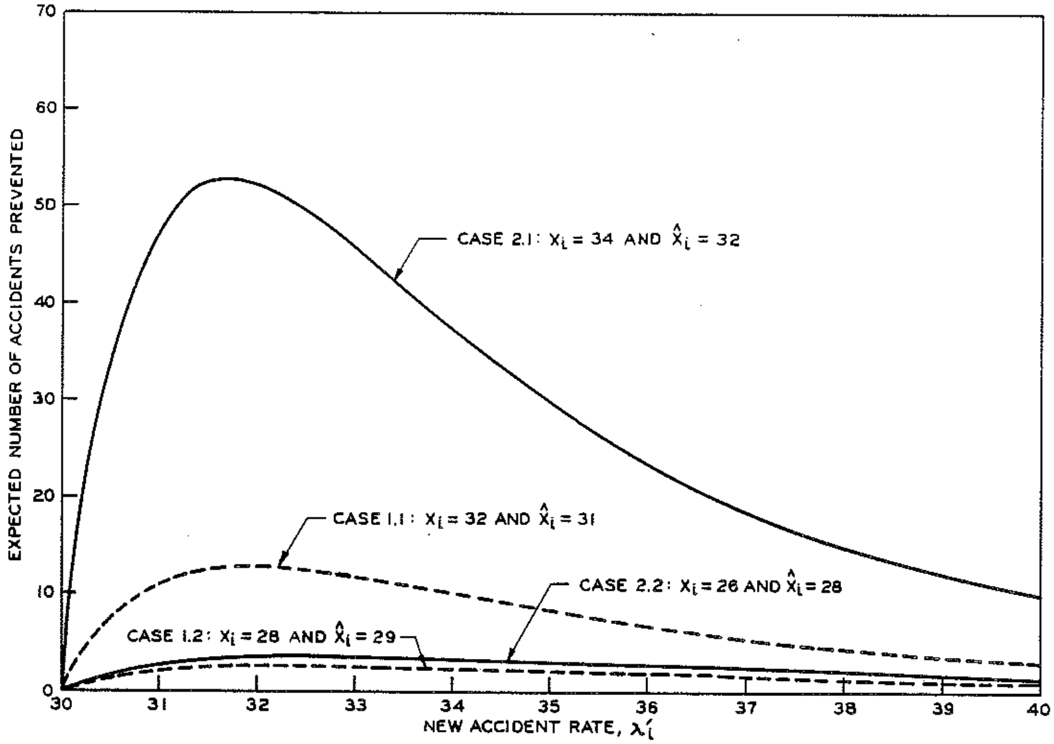

Figure 1. The benefits of using the proposed (solid lines) and usual (dashed lines) upper control limits to detect assignable accident factors.

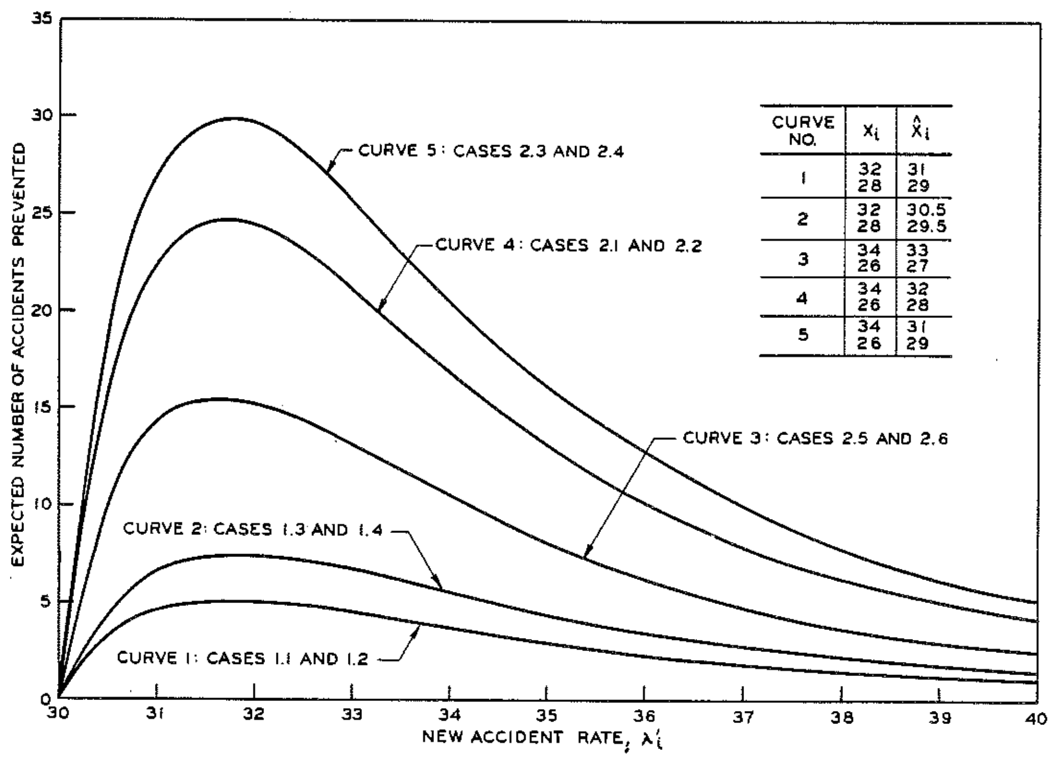

Figure 2. The net benefits of using the proposed upper control limits to detect assignable accident factors.

of additional 'normal' locations (operating at the true accident rate). Therefore, if a second stage monitoring system is implemented to detect the false alarms, the upper control limit can be updated in the near future. Thus, the net benefits of using the proposed upper control limit to detect assignable factors greater than those shown in Figure 2.

To investigate the impact of estimation errors on benefits, we consider a situation which doubles the previous estimation error. That is, we consider the following pair of cases:

Case 2.1: $X_i = 34$ and $\hat{X}_i = 32$

and

Case 2.2: $X_i = 26$ and $\hat{X}_i = 28$

The difference between the usual and proposed estimates is two accidents per year in both cases. The results obtained by repeating the above procedures are also presented in Figures 1 and 2. The results for other similar pairs of cases are also presented in Figure 2 for comparison purposes. We conclude from Figure 2 that:

C1) When major factors are added to the system causing locations to operate at much higher than normal accident rates, the benefits of using the proposed upper control limit to detect assignable factors is rather small (tail of curves in Fig. 2). This is because both upper control limits have a good chance to detect the added-in factors in this case.

C2) When minor factors are added to the system causing locations to operate at slightly higher than normal accident rates, the benefits of using the proposed upper control limit could be very large indeed (the middle section of curves in Fig. 2). This is because when the usual estimate is much higher than the true accident rate, the usual upper control limit has virtually no chance at all to detect the added-in minor assignable factors.

Intuitively, one expects that minor factors frequently affect the traffic system. Thus, in the case where the usual estimate is poor and the proposed estimate is good, the benefits of using the proposed upper control limit to detect assignable factors are considerable as shown by Cases 2.3 and 2.4 in Figure 2. As mentioned in Part I, the proposed estimate could be worse than the usual estimate. In this situation, we suffer from using the proposed upper control limit to detect assignable factors. However, the nature of the proposed estimation method is such that this situation occurs far less frequently than the former case. Moreover, when it does occur, the difference between the usual and proposed estimates is almost always small. Thus, gain is much larger than loss in terms of magnitude and frequency. Therefore, we conclude that using the proposed upper control limit to detect assignable factors results in substantial accident reduction in the long run.

Note that there is no way to mathematically support the above conclusion due to the complex relationships among initial, usual and proposed parameter estimates. However, we did simulate the above problem on 10 locations whose true yearly accident rates are 30, 31, . . . , and 39, respectively. For the cases where initial estimates are all perfect, off by 1, and 5, the simulated results show that at least 99 percent of the preventable accidents undetected by the usual upper control limits will be picked up by the proposed upper control limits. This demonstrates that a relatively small improvement in estimation accuracy results in a substantial improvement in practical benefits.

We now proceed to estimate the accident rates of Categories 13 through 24. This is presented in the following example.

Example 2 - The same grouping method used in Example 1 is used again to estimate accident rates of Categories 13 through 24. Estimated results are presented in Table 7. In this case, the total percentage gain is only 9.20 percent. We also see from Table 7 that the order of accident rates of Categories 13 through 15 remains unchanged. This is because using group average as initial parameter estimates will preserve the rank order of parameters within the group. As mentioned previously, the usual rank order of these three categories is not compatible with that of the first three categories. To correct the rank order, supplemental variables are needed.

The i-th and (i + 12)-th categories have the same geometric configuration except that every roadway segment in the (i + 12)-th category contains an intersection. Thus, θ_i and θ_{i+12}, i = 1, . . . , 12 could be closely related such as $\theta_{i+12} = b\,\theta_i$ or $\theta_{i+12} = a + b\,\theta_i$. The relationship between θ_i and θ_{i+12} can then be used to obtain initial parameter estimates. For the purpose of finding their relationships, we first convert X_i to Z_i so that the i-th and (i + 12)-th categories have the same accident exposure. This is accomplished by the following equation:

$$Z_i = A_i \times 365 \times 5 \times T_{i+12} \times 0.2 \times N_{i+12}, \quad i = 1, \ldots, 12 \qquad (14)$$

Z_i can be interpreted as the number of accidents occurring on roadway segments in the i-th category that has the same accident exposure as the (i + 12)-th category. Graphical examination reveals that X_{i+12} and X_i, i = 1, . . . , 12, are approximately proportionate. Since the proposed method requires the use of Anscombe's transformation of X_{i+12}, i = 1, . . . , 12, we therefore use the following equation to obtain initial estimates

$$\sqrt{X_{i+12} + 0.375} = b \sqrt{X_i + 0.375}\ , \ i = 1, \ldots, 12 \qquad (15)$$

Least squares is used to estimate b in Eq. (15) and, subsequently, to obtain initial accident rates of Categories 13 through 24. The estimated results obtained from using the computer program of Part V are presented in Table 8. The total percentage improvement is 30.18 percent. We see from Table 8 that some of the usual accident rate estimates have been changed substantially. In particular, $\hat{A}_{13}$ is larger than $\hat{A}_{14}$ which is larger than $\hat{A}_{15}$ as expected since wider roadways experience lower accident rates.

Table 8
Estimated Results of Categories 13 through 24
Using The First 12 Category Rates to Obtain Initial Estimates

Category i	Poisson Rate		Accident Rate		Difference
	Usual(X_i)	Proposed(X_i)	Usual(A_i)	Proposed(A_i)	
13	51	56.05	57.79	63.51	-5.72
14	92	90.18	63.56	62.30	1.26
15	103	99.94	59.70	57.93	1.78
16	80	78.09	61.11	59.65	1.46
17	266	264.46	54.77	54.45	0.31
18	354	350.93	49.62	49.19	0.44
19	133	125.96	103.89	98.39	5.50
20	207	203.26	82.72	81.22	1.50
21	107	117.01	61.76	67.53	-5.78
22	62	55.89	86.94	78.37	8.57
23	180	180.39	62.32	62.46	-0.14
24	258	269.16	50.86	53.06	-2.20

Note that the rank order of Categories 13 through 24 based on usual and proposed estimates are different. Since the proposed estimates are closer to the true value than the usual estimates; intuitively, the proposed rank order should be more accurate than the usual one. We postpone discussion on this subject until the next section.

We now suppose that Eq. (10) is operating as part of a high accident location detection program. Each time we do not identify and remove assignable factors causing the i-th location to operate at the higher rate λ_i', extra $\lambda_i' - \lambda_i$ accidents are expected to occur. This occurrence can be reduced by increasing α. The higher α will result in a longer list of locations to be examined. As previously mentioned, some locations in the list will turn out to have no problem at all. They are in the list simply because of random factors in accident occurrence. Moreover, time and funding could prevent examination of the full list. Thus, we should rank locations on a priority basis according to, e.g., accident increase or percent accident increase which, for the location indexed by j are, respectively, defined as

$$I_j = \lambda_j' - \lambda_j \tag{16}$$

and

$$R_j = 100 \frac{\lambda_j' - \lambda_j}{\lambda_j} \tag{17}$$

Again, the estimated rates must substitute for the unknown rates. The accuracy of the priority list then depends on the closeness of the estimated rates to true ones. Conventionally, X_i serves as the estimate of λ_i. This is the point at which the proposed method can be used to improve the usual estimates and hence the priority list as pointed out in Example 2.

Once the priority list is available, we can then start to examine locations according to the order of the list until time or funding runs out. Suppose that this action results in a list of k locations that can be improved. For discussion purposes, we re-index these locations. Suppose that there are n_i improvement alternatives available for treating the i-th location. Denote λ_{ij} to be the accident rate of the i-th location after being treated by the j-th improvement alternative. We also denote b_{ij} and c_{ij} to be the respective accident reduction and improvement cost of implementing the j-th improvement alternative. That is,

$$b_{ij} = \lambda_i - \lambda_{ij}, \; j = 1, \ldots, n_i \text{ and } i = 1, \ldots, k \tag{18}$$

Our goal is to find a strategy for improving these locations which results in maximal accident reduction for the available funding. To this end, we define $A_{ij} = 1$ if the j-th improvement alternative is used to treat the i-th location, otherwise, $A_{ij} = 0$. The best improvement strategy can be found by selecting A_{ij} so as to maximize

$$O(A_{ij}, \; j = 1, \ldots, n_i \text{ and } i = 1, \ldots, k) = \sum_{i=1}^{k} \sum_{j=1}^{n_i} A_{ij} \, b_{ij} \tag{19}$$

subject to the following restrictions

$$\sum_{j=1}^{n_i} A_{ij} \leq 1, \; i = 1, \ldots, k \tag{20}$$

$$\sum_{i=1}^{k} \sum_{j=1}^{n_i} A_{ij} c_{ij} \leq F \tag{21}$$

and

A_{ij} is either 0 or 1 for every i and j. (22)

The right hand side of Eq. (19) is the total accident reduction due to the improvement strategy A_{ij}, $j = 1, \ldots, n_i$ and $i = 1, \ldots, k$. Eqs. (20) and (21) ensure that no more than one improvement action will be performed on a location and the total cost will not be over the available funding F. It is apparent that the improvement strategy is determined by the specified accident reduction and cost of each improvement action and the available funding. Since b_{ij} and c_{ij} are not known, we are faced with the same problem as before of obtaining the best possible estimates of these unknown constants. In this study, we shall not deal with the estimation of improvement costs. However, the proposed estimation method can be used to improve the usual estimates of λ_i' and λ_{ij} in Eq. (18) resulting in a superior improvement strategy.

The above accident rate could be the number of accidents per vehicle-mile. Traffic volume is generally estimated from traffic counts taken in short time intervals. Since traffic count is also distributed approximately according to the Poisson law, the proposed method can also be used to improve usual estimates of traffic volumes. This would benefit many areas in which traffic volume is the main variable used in decision making. For example, the total traffic volumes as well as the ratio of traffic volumes of two intersecting roadways may be used to determine whether signalization is required. In this case, the estimated traffic volumes are compared with the designed values to recommend action or no action. When the true values are not in the neighborhood of the designed values, both usual and proposed estimates probably lead to the same decision. However, if the true values are very close to the designed values, the small estimation improvement could change decisions from no to yes and vice versa. Since the proposed estimates are closer to the true values than the usual estimates, the chance of making a 'wrong' decision would be reduced by using the proposed estimates. The consequences of making wrong decisions could be quite serious. This again demonstrates that small estimation improvement can result in substantial benefit in the long run. Another important example is the use of traffic volumes of various vehicle types to compute highway noise levels. This will be explained in a later section.

Note that the term 'accident' has been used in a very broad sense. That is, if the goal is to reduce fatal accidents, X_i and λ_i then refer to the number of fatal accidents and the fatal accident rate, respectively. However, the above optimization procedures remain unchanged. In passing, we point out that it generally takes a long period of time or a large accident exposure to build a reliable X_i because of the relative rarity of fatal accidents. This could make the above optimization procedures impractical. One way of avoiding this problem is to use the same definitions of X_i and λ_i as before, but replace Eq. (18) with the following equation.

$$b_{ij} = \lambda_i' s_i - \lambda_{ij} s_{ij} \tag{23}$$

s_i is the severity index of an accident occurring at the i-th location. In this case, the severity index is the probability that an accident involves at least one fatality. s_i and s_{ij} can be estimated from the larger data sets. Methods which improve the usual accident severity estimates are discussed in the next section.

We have seen that accident rate and other related parameters play very important roles in highway safety. Since the proposed estimation method is capable of improving the usual estimates of these important parameters thereby increasing the overall system performance, it can be of considerable value to administrators who are in charge of safety improvement program development.

Potential Benefits of Using the Proposed Method to Estimate Proportions of Wet Surface Accidents, Accident Severity Indices and Other Related Parameters in Highway Safety and Maintenance Programs

In this section, we deal with the estimation of parameters such as accident severity indices and proportions of wet surface accidents. These parameters take values from 0 to 1. Thus, we would not expect to have large differences between the usual and proposed estimates. This is especially true when the usual estimates are reliable (as in the case of large sample sizes). However, the estimated values are often incorporated into, for example, a priority list or values which are either 0 (no) or 1 (yes) used for administrative decision making purposes. As mentioned in the previous section, a relatively small estimation improvement could reverse some decisions resulting in substantial benefit in the long run. This will now be demonstrated through the development of the following two programs.

Pavement Resurfacing Program - Suppose that we are interested in developing a pavement resurfacing program to maximally reduce wet surface accidents at intersections. Naturally, the first step is to obtain a priority list of locations on which friction tests are to be performed. This priority list is typically obtained according to the rank order of proportions of wet surface accidents. Applying this method to two locations which have identical friction coefficients favors testing that location which experiences greater wet time. This is a systematic error. To avoid this error, we should rank only those locations whose proportions of wet surface accidents are suspiciously higher than the value to be maintained. Thus, the first step is to statistically set the upper control limit for each location under the normal condition (average percentage of wet time and the minimum friction coefficient to be maintained). Suppose that the true proportion of wet surface accident of the i-th location is w_i. Then, the upper control limit for declaring at the 100 $(1 - \alpha)$ percent confidence level that the friction coefficient of the i-th location is lower than the minimum value to be maintained is

$$UCL_i = w_i + z_{1-\alpha} \sqrt{\frac{w_i (1 - w)_i}{n_i}} , \quad i = 1, \ldots, L \tag{24}$$

where $z_{1-\alpha}$ is the upper 100α percentage point of the standard normal distribution and n_i is the sample size (number of accidents in this case) for obtaining the proportion of wet surface accidents. Since the true proportion of wet surface accidents is unknown, the estimated value must be used as a substitute for w_i in Eq. (24). The usual estimate of w_i is

$$W_i = \frac{\text{Total Number of Wet Surface Accidents}}{\text{Total Number of Accidents}} \tag{25}$$

Note that this case is parallel to the previous case which detects accidents due to assignable factors. Therefore, the same arguments used there can be used again to conclude that the proposed upper control limits are better than the usual ones in detecting locations with friction coefficients lower than the minimum value to be maintained. We emphasize again that a small difference, such as 0.01, between the usual and proposed estimates can result in a significant wet surface accident reduction in the long run (see discussion in the previous section for obtaining Figures 1 and 2).

Because of limited funding and time, we might not be able to test every location generated by the above quality control procedure. Thus, a selection procedure must be used to test only locations that maximize wet accident reduction for the funding available. For discussion purposes, let us assume that locations are to be tested according to the rank order of proportions of wet surface accidents. The usual and proposed priority lists are obtained by ranking the usual and proposed estimates of proportions of wet surface accidents, respectively. Since the proposed estimates are better than the usual ones; intuitively, the proposed list should be more accurate than the usual list. We provide the following example of simulated results to support this statement.

Simulated Results on Priority Lists - We assume that the true proportion of wet surface accidents of the k-th location is

$$w_k = 0.76 - 0.01\,k, \quad k = 1, \ldots, 50$$

That is, the true proportions of wet surface accidents of these 50 locations range from 0.26 to 0.75. The true rank order is 1, 2, 3, . . . , 50. If funding is available for testing, for example, only 10 percent of the list, our choice is the first five locations. We provide in Table 9 the simulated percentages of correct locations picked up by the usual and proposed lists obtained under the condition that initial parameter estimates randomly deviate from the true values by a constant b. When b = 0, the initial parameter estimate is perfect for every location. The larger the constant b, the worse the initial and, consequently, proposed parameter estimates. We see from Table 9 that the better proposed estimates (corresponding to smaller b) provide the larger percentage of correct locations picked up by the list. Since the proposed estimates (with reasonably good initial estimates) are better than the usual ones, the proposed lists are therefore also better than the usual list as shown in Table 9.

Table 9
Simulated Results on Prority Lists of
Locations with Sample size 20 Each

Number of Top Locations to Be Tested	Percentage of Locations Correctly Detected for Skid Testing Programs				
	Usual Priority List	Proposed Priority Lists with Initial Estimates off by The Constant b			
		b=0	b=0.02	b=0.05	b=0.10
5	39.68	93.28	93.00	76.56	66.72
10	57.76	96.34	89.88	82.08	75.10
20	75.91	98.00	94.55	90.08	84.98
25	79.97	98.44	95.52	89.96	87.04

We note that the differences between usual and proposed estimates decreases as the constant b increases. This again emphasizes the point that a relatively small estimation improvement can significantly increase the accuracy of the priority list (see the case b = 0.1 in Table 9) and, therefore, reduce wet surface accidents in the long run.

The sample size for each usual estimate of proportion of wet surface accidents in Table 9 is 20. Simulated results for the case that the sample size is 100 are presented in Table 10. The increase in sample size substantially increases the reliability of usual estimates of proportions of wet surface accidents. This also increases the accuracy of the usual priority list as can be seen from the comparison of Tables 9 and 10. This shows again that the better the parameter estimates, the more accurate the priority list. The comparison of these two tables also indicates that sample size influences accuracy of the usual list more than the proposed list. This is because it is more difficult to substantially improve reliable usual estimates (large sample sizes). Nevertheless, the proposed lists are still better than the usual lists as shown in Tables 9 and 10.

Table 10
Simulated Results on Prority Lists of Locations with Sample size 100 Each

Number of Top Locations to Be Tested	Percentage of Locations Correctly Detected for Skid Testing Programs				
	Usual Priority List	Proposed Priority Lists with Initial Estimates off by The Constant b			
		b=0	b=0.02	b=0.05	b=0.10
5	65.32	97.92	96.52	82.04	74.36
10	57.76	98.94	91.32	87.48	84.62
20	89.69	99.36	95.43	93.19	91.53
25	91.78	99.34	96.41	93.33	92.46

Roadside Safety Improvement Program - We are interested in developing a roadside safety improvement program that would optimally allocate the available funding to improve roadside safety. Naturally, the first step is to identify roadside obstacles and the corresponding improvement alternatives such as removing curbs, trees, and utility poles; flattening roadside slopes; installing gore attenuation systems; shielding bridge abutments, pier, and wall faces with guardrail; installing median barrier; etc. A more complete list of roadside improvement alternatives can be found in NCHRP 148 (10). For discussion purposes, we numerically index every roadside obstacle (location), say, 1 through k. Suppose that there are n_i improvement alternatives available for treating the i-th obstacle. Denote λ_i and λ_{ij} to be the expected number of accidents involving the i-th obstacle before and after implementing the j-th improvement alternative, respectively. Also, denote c_{ij} to be the cost of treating the i-th obstacle with the j-th improvement alternative. We define $A_{ij} = 1$ if the j-th improvement alternative is used to treat the i-th obstacle, otherwise $A_{ij} = 0$. Then, the best improvement strategy can be obtained by selecting A_{ij} so as to maximize

$$O(A_{ij},\ j = 1, \ldots, n_i \text{ and } i = 1, \ldots, k)$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{n_i} A_{ij}(\lambda_i s_i - \lambda_{ij} s_{ij}) \tag{26}$$

subject to the following restrictions

$$\sum_{j=1}^{n_i} A_{ij} \leq 1, 1, \ldots, k \tag{27}$$

$$\sum_{i=1}^{k} \sum_{j=1}^{n_i} A_{ij} c_{ij} \leq F \tag{28}$$

and

$$A_{ij} \text{ is either 0 or 1 for every i and j.} \tag{29}$$

Eqs. (27) and (28) ensure that no more than one improvement alternative will be used to treat an obstacle and the total cost will not be over the available funding, F. s_i and s_{ij} in Eq. (26) are the severity indices of accidents involving the i-th obstacle before and after being treated by the j-th improvement alternative, respectively. If s_i and s_{ij} are conventional severity indices, the objective function defined in Eq. (26) is the expected total injury and fatal accident reduction due to a roadside safety improvement program specified by A_{ij}, $j = 1, \ldots, n_i$, and $i = 1, \ldots, k$. Accident rates and severity indices are unknown, and estimated values must be used as substitutes for these unknowns. As in the case of the priority list, the better the estimated values, the more accurate the improvement strategy. Since the proposed estimates are better than the usual ones, we should use the proposed improvement strategy to improve roadside safety. To what extent the proposed method can improve the usual improvement strategy depends on the complex relationships among improvement costs and benefits, and initial, usual, and proposed estimates. Two improvement strategies could be the same for one case, and completely different for another case. The point is that it does not hurt to use the proposed method when designing the improvement strategy from an overall system performance point of view. The worst situation is that no improvement is made. On the other hand, the proposed method could substantially improve the system performance. Therefore, the proposed method can be of considerable value in designing highway safety programs.

Let us now examine another area in which accident severity indices also play an important role in the safety decision making process. For a divided highway, we denote p_1 to be the probability that a vehicle encroaching onto the median would collide with the in-place median barrier. Also, denote s_1 to be the severity index appropriate to vehicle-barrier collisions. We define the hazard index as the expected number of accidents in the severity category under consideration. Then, the hazard index (10) of installing a median barrier is

$$H_1 = e \cdot p_1 \cdot s_1 \tag{30}$$

The notation 'e' in the above equation stands for the expected number of vehicles encroaching onto the median. If the median is barrier free, we denote p_2 to be the probability that an encroaching vehicle would cross over the median and collide with vehicle(s) traveling on the opposite roadway. If s_2 is the severity index of this type of accident, the hazard index for this case is

$$H_2 = e \cdot p_2 \cdot s_2 \tag{31}$$

Ignoring installation and maintenance costs, it is beneficial to install a median barrier only if

$$H_1 < H_2 \tag{32}$$

or

$$\frac{H_1}{H_2} = \frac{p_1}{p_2} \cdot \frac{s_1}{s_2} < 1 \tag{33}$$

It is intuitively clear that $p_1 > p_2$. Thus, installing a median barrier is justified only when the reduction of accident severity is large enough to compensate for the increase in accident probability.

Generally, hazard indices are functions of traffic volume and roadway separation for a given type of median barrier. To set up standard guidelines for median barrier installation, we need to know the critical combination of traffic volume and roadway separation satisfying $H_1 = H_2$. This involves computing p_1 and p_2, and estimating s_1 and s_2. It is clear that the accuracy of the estimated values has direct impact on the accuracy of design guidelines. Of course, small estimation errors would not affect the decision made on those roadways that are not in the neighborhood of the critical conditions required for median barrier installation. However, for those roadways at the critical condition threshold, small estimation errors could reverse decisions. Thus, the proposed method has an ideal application in the design of standard guidelines for median barrier installation which maximally prevent accidents and unnecessary construction.

Computations of p_1 and p_2 are beyond the scope of this study and will not be discussed here. The usual estimates of severity indices and related measurements are discussed below.

For discussion purposes, we denote s to be the conventional severity index. That is, the usual estimate of s is

$$S = \frac{\text{Number of Reported Injury and Fatal Accidents}}{\text{Total Number of Reported Accidents}} \quad (34)$$

Assume that accidents behave according to the Poisson law. It is known that, given the total number of reported accidents, the number of injury and fatal accidents is a binomial random variable with s as the parameter of proportionality. That is, S in Eq. (34) is the usual estimate of the conventional severity index which also is the proportion parameter of a binomial process. The usual estimate defined in Eq. (25) is also explained in the same way as above. We now provide an example which demonstrates improvement over the usual estimates of proportional parameters in the binomial processes. This example also serves to demonstrate the following:

1) When sample sizes are large (resulting in reliable usual estimates), the total percentage improvement is generally low. Consequently, the differences between usual and proposed estimates are small. However, if the total percentage improvement is high, the initial and usual estimates must be very close to each other. This indicates that the method used to obtain the initial estimates is a proper one for describing the relationships between supplemental variables and parameters to be estimated. This information could be useful for future research and applications.

2) Although the usual and proposed estimates are very close to each other, small differences become meaningful when the estimated values are transformed into other forms of information used for decision making purposes.

Example: Michigan implemented the 55 mph speed limit in 1974. We shall use the severity data of 1972, 1973, and 1975 published in "Michigan Traffic Accident Facts" (11) to examine the effect of the new speed limit on fixed object accident severity. These data are presented in Table 11. We note that the sample sizes (total objects hit) are very large. Unless there are major changes in safety policies, we would not expect changes, other than small random fluctuation, in severity indices from one year to another. This assumption is supported by the 1972 and 1973 severity data in Table 11. Nevertheless, the 1972 severity indices can be used as initial estimates

Table 11
Usual Severity Indices of Fixed Object Accidents
(Michigan Traffic Accident Facts pp. 22-23, 1972,1973 & 1975)

Type of Object Hit	Proportion of Injury & Fatal Accidents		
	1972	1973	1975
1. Guard Rail Or Post	0.3565	0.3709	0.3422
2. Highway Sign	0.2569	0.2772	0.2476
3. Street Light, Utility Pole	0.4742	0.4714	0.4538
4. Culvert	0.5532	0.5554	0.5659
5. Ditch, Embankment, Stream	0.4140	0.4034	0.3804
6. Bridge Pier Or Abutment	0.5700	0.5692	0.5470
7. Bridge Rail Or Deck	0.4560	0.4604	0.4023
8. Tree	0.5227	0.5342	0.4987
9 Highway Or Railroad Signal	0.2601	0.3133	0.2724
10. Building	0.4106	0.3946	0.3791
11. Mailbox	0.2228	0.2272	0.2080
12. Fence	0.2921	0.3004	0.2393
13. Traffic Isle Or Curb	0.3946	0.4023	0.3908
14. Other On-trafficway Object	0.2832	0.2816	0.2846
15. Other Off-trafficway Object	0.3512	0.3501	0.3329
16. Overhead Fixed Object	0.1136	0.1255	0.0738
17. Not Kknown	0.2198	0.2162	0.1626

to estimate 1973 severity indices. The estimated results obtained from using the computer program of Part VI are presented in Table 12. The total percentage improvement of the proposed method over the usual one is 43.65 percent. Since the sample size (total objects hit) for each type of accident severity is very large, the usual severity indices are very reliable. As mentioned previously, high percentage improvement is obtained only when the initial estimates are excellent. That is, the initial estimates are very close to the usual estimates. Consequently, most differences

Table 12
Estimated Severity Indices of
Fixed Object Accidents in 1973

Type of Object Hit	Severity Index	
	Usual	Proposed
1	0.3709	0.3644
2	0.2772	0.2680
3	0.4714	0.4727
4	0.5554	0.5544
5	0.4034	0.4082
6	0.5692	0.5696
7	0.4604	0.4584
8	0.5342	0.5290
9	0.3133	0.2890
10	0.3946	0.4018
11	0.2272	0.2252
12	0.3004	0.2967
13	0.4023	0.3988
14	0.2816	0.2823
15	0.3501	0.3506
16	0.1255	0.1201
17	0.2162	0.2178

between usual and proposed estimates are small. This is precisely the case in this problem as seen from Table 12. When this occurs, we conclude that the method used to obtain initial estimates is proper for describing the relationship between parameters and supplemental variables. For this case, it means that there are no major changes in safety policies which affect accident severities under consideration.

We now combine 1972 and 1973 severity data to form the estimates of severity indices of various fixed object accidents before implementation of the new speed limit. These are presented in Table 13. We observe that the new speed limit appears to have greater effect on the lower severity indices. This seems reasonable for our particular definition of accident severity. Moreover, lowering the limit should reduce severity indices of all types of fixed object accidents. This means that the fourth and fourteenth usual severity indices in Table 13 run counter to the overall trend.

Table 13
Estimated Severity Indices of Fixed Object Accidents
Before and After Implementing 55-MPH Speed Limit

Type of Fixed Object	Severity Index			%-decrease(increase)		
	Before (Usual)	After (Usual)	After (Proposed)	Usual	Proposed	Difference
1	0.3632	0.3422	0.3429	5.782	5.600	0.182
2	0.2667	0.2476	0.2461	7.162	7.707	-0.545
3	0.4728	0.4538	0.4551	4.019	3.742	0.277
4	0.5543	0.5659	0.5564	(2.093)	(0.371)	1.722
5	0.4092	0.3804	0.3830	7.038	6.396	0.642
6	0.5697	0.5470	0.5428	3.985	4.725	-0.740
7	0.4584	0.4023	0.4036	12.238	11.960	0.278
8	0.5283	0.4987	0.5011	5.603	5.145	0.458
9	0.2915	0.2724	0.2665	6.552	8.562	-2.010
10	0.4025	0.3791	0.3791	5.814	5.817	-0.003
11	0.2250	0.2080	0.2058	7.556	8.543	-0.987
12	0.2961	0.2393	0.2444	19.183	17.471	1.712
13	0.3984	0.3908	0.3872	1.908	2.799	-0.891
14	0.2823	0.2846	0.2804	(0.815)	0.674	-1.489
15	0.3508	0.3329	0.3323	5.103	5.276	-0.173
16	0.1195	0.0738	0.0745	38.243	37.656	0.587
17	0.2191	0.1626	0.1628	25.787	25.696	0.091

We now use the proposed method to estimate severity indices of various fixed object accidents occurring after implementation of the 55 mph speed limit. For the i-th type of fixed object, we denote X_i and Y_i to be Anscombe's transformation of severity indices before and after implementation of the new speed limit, respectively. The data indicate that the following equation holds approximately.

$$Y_i = a + b X_i + c X_i^2 .$$

The above linear relationship is, therefore, used to obtain least squares estimates as initial estimates of after-period severity indices. The estimated results obtained from using the computer program of Part VI are also presented in Table 13. The total percentage improvement of the proposed method over the usual one is only 11.39 percent. Consequently, the differences between usual and proposed estimates are very small as shown in Table 13. However, after converting these estimates to percentages of decline for the purpose of measuring speed limit impact on accident severity, we see from Table 13 that the proposed method substantially reduces the percent increase of the fourth severity index and reverses the fourteenth severity index from a 0.815 percent increase to a 0.673 percent decrease in line with expectation. If better information is available, the proposed method might also reverse the fourth severity index's direction. This supports the previous remark that small estimation improvement can become meaningful depending on the evaluation criterion.

Since injury and fatal accidents tend to have a higher reporting level than property damage accidents, the usual severity index estimate defined in Eq. (34) is likely to be an overestimate of the true value. To correct this bias, we introduce the accident reporting level, r; defined as the probability that an accident will be reported. The usual estimate of the reporting level is

$$R = \frac{\text{Total Number of Reported Accidents}}{\text{Total Number of Accidents}} \tag{35}$$

As suggested, the reporting level is positively correlated with the severity index. Special experiments are required to estimate the reporting level for any accident class such as those with median barriers. Once estimated severity indices and reporting levels are available, the decision on installing a median barrier can be made based on the ratio of the two estimated hazard indices. That is,

$$\frac{H_1}{H_2} = \frac{p_1}{p_2} \cdot \frac{S_1}{S_2} \cdot \frac{R_1}{R_2} \tag{36}$$

Again, the proposed method can be used to improve usual estimates of reporting level for the purposes of designing standard guidelines for median barrier installation.

Based on the above discussions, we conclude that the proposed method is a very useful tool for the development of better highway safety improvement programs.

<u>Potential Benefits of Using the Proposed Method to Estimate Mean Values in Highway Maintenance, Aggregate Testing, and Noise Programs</u>

In highway construction and maintenance, sample averages are statistics commonly used to estimate parameters that measure product quality or characterize processes. Thus, methods for using sample averages to obtain better estimates of these parameters would be of great value to highway administrators. To illustrate this, we present the following three areas in which sample averages play very important roles in decision making or in developing useful programs.

Pavement Resurfacing Program - As mentioned before, to develop a pavement resurfacing program for maximally reducing wet surface accidents at intersections, the first step is to obtain a priority list of locations on which friction tests are to be performed. We have shown previously that the proposed list is better than the usual one. Once a priority list is established, the next step is to test locations according to the order of the list until funding or time runs out. Numerous friction test results are obtained for each intersection. The test result average is the usual estimate of the true coefficient of friction. This estimate, together with the proportion of wet time, determines the net benefit of resurfacing an intersection (12). Net benefits and improvement costs are then used to develop an optimal pavement resurfacing program. We remark that this optimal program can be obtained either by solving a mathematical system such as Eqs. (19) through (22) or by using a priority list based on net benefits (per dollar spent). As mentioned previously, better estimates of parameters (friction coefficients) would lead to a better improvement strategy. Since the proposed estimates are superior to the usual ones, the proposed method is of great value in developing effective pavement resurfacing programs.

Highway Noise Program - The noise level generated by a vehicle traveling on a roadway is a function of vehicle type, speed, distance to receiver as well as geometric configuration and environmental conditions. The noise levels of various combinations of these variables are essential parameters of a computer program (13) which computes highway noise levels such as L_{10}, L_{50} and L_{eq}, etc. Since these parameters are unknown and must be estimated from the data, the reliability of the computed value certainly depends on the accuracy of the estimated parameters. As usual, averages of the observed noise levels serve as estimates of those parameters in the computer program. Thus, the proposed method for improving sample averages would be of great value in designing a more accurate computer program.

The noise levels of roadway segments are generally transformed into a priority list for administrative decision making purposes such as noise barrier installation. Since the true noise levels are unknown, the estimated values must be used as substitutes. As mentioned in the previous section, small estimation errors could significantly change the priority list and, consequently, affect decisions based on the list. Thus, it is of fundamental importance to reduce estimation errors as much as possible. Use of the proposed method for this purpose is discussed below.

One way to obtain the noise level estimates is through the use of a published noise computer program. We shall assume that this program is valid. Then, the reliability of the computed values depends on the accuracy of the input variables such as traffic volumes and average speeds of various types of vehicles, geometric configuration and environmental factors. We note that traffic volumes and average speeds are sample estimated values and are therefore subject to random error. Moreover, each roadway segment is unique in terms of geometric configuration and environmental conditions. It is not possible for the computer program to cover all different types of roadway segments. Therefore, for some roadways, the computed value may not be a good estimate of the true noise level due to imperfection in input variable estimates. Alternatively, we may take random samples to estimate the true noise level of each location. In this case, the sample average is the usual estimate of the true noise level. Of course, this estimate is also subject to error. This error can be reduced by the proposed method which uses the computed values to adjust sample averages as shown in Part III. Thus, the proposed estimates give better information for decision making purposes.

Aggregate Testing and Acceptance - Aggregate testing and acceptance is a major portion of the daily operations in highway construction and maintenance. Therefore, it is important to have methods for obtaining good estimates of parameters used to design aggregate testing and acceptance procedures.

The Michigan Department of Transportation is currently experimenting with the mechanical testing method for aggregate testing. It has been shown by experiment (14) that this method produces different results from the conventional hand testing method. Intuitively, aggregate degrades if it is over-shaken. Thus, the shaking time may be the key element in mechanical testing method design at least if it is to produce the same results as the conventional method. The proper time setting can be determined by using sample averages of a well designed experiment. The proposed estimation method can further improve these averages to design a better testing method.

It has been shown that aggregate significantly degrades when transported and compacted (15). In order to properly adjust specification limits for the purpose of designing an in-place aggregate inspection plan, we need to know aggregate degradation rates. These rates can be estimated from

Table 14
Peak Noise Levels of Commercial Vehicles

No of Axles	Speed	Sample size	Usual Estimate (Sample Average)	Proposed Estimate	Difference
2	25.2	5	72.240	71.558	0.682
2	30.0	19	72.170	72.385	-0.215
2	35.5	60	73.760	74.030	-0.270
2	40.0	34	75.341	75.493	-0.152
2	44.6	22	75.852	76.527	-0.675
2	50.2	13	79.208	78.928	0.280
2	55.3	32	80.293	80.290	0.003
2	59.3	27	82.048	81.737	0.311
2	65.5	4	83.475	83.439	0.036
3	25.5	2	79.600	78.485	1.115
3	31.7	9	76.746	77.964	-1.218
3	37.0	7	79.329	79.617	-0.288
3	41.5	11	78.929	79.937	-1.008
3	46.5	2	82.700	82.054	0.646
3	54.1	14	85.178	83.912	1.266
3	58.0	19	84.208	83.929	0.279
3	65.7	3	83.800	84.592	-0.792
4	27.7	3	77.900	77.155	0.745
4	35.4	5	78.920	78.918	0.002
4	41.8	5	79.320	80.196	-0.876
4	48.0	2	81.950	82.370	-0.420
4	53.5	22	84.783	84.508	0.275
4	57.9	19	85.348	85.508	-0.160
4	62.3	7	86.016	86.551	-0.535
4	66.0	2	89.700	88.732	0.968
5	27.8	4	77.400	77.177	0.223
5	32.3	18	77.367	77.848	-0.481
5	36.7	22	79.533	79.421	0.112
5	42.0	9	80.600	80.673	-0.073
5	45.0	1	81.000	81.296	-0.296
5	52.9	39	84.564	83.985	0.579
5	56.9	83	85.691	85.064	0.627
5	61.4	55	85.925	85.846	0.079
5	66.3	8	85.189	86.283	-1.094
5	71.0	2	88.850	88.527	0.323
6	27.0	2	78.500	77.720	0.780
6	36.3	4	79.325	79.623	-0.298
6	41.0	2	80.950	81.090	-0.140
6	52.8	8	85.774	85.083	0.691
6	56.8	12	86.850	86.203	0.647
6	61.4	16	86.923	87.004	-0.081
7	20.0	1	69.900	73.356	-3.456
7	36.0	1	80.200	80.341	-0.141
7	46.0	2	82.300	82.894	-0.594
7	57.3	6	87.567	86.988	0.579
7	62.0	5	87.520	87.756	-0.236
8	35.5	2	82.750	81.724	1.026
8	41.0	2	82.800	82.667	0.133
8	54.0	5	88.800	87.352	1.448
8	60.3	4	88.300	88.199	0.101
8	65.0	1	86.800	88.360	-1.560
9	33.3	3	82.467	81.639	0.828
9	55.3	4	88.425	87.815	0.610
9	59.8	4	89.425	88.987	0.438
10	55.3	6	88.035	88.054	-0.019
10	59.8	6	89.733	89.518	0.215
11	24.0	5	81.382	80.431	0.951
11	31.4	7	81.871	81.876	-0.005
11	40.3	11	85.046	84.694	0.352
11	45.0	4	86.525	86.099	0.426
11	53.3	8	88.238	88.206	0.032
11	57.9	14	88.691	89.166	-0.475
11	61.7	3	88.767	89.835	-1.068
11	66.0	1	89.800	90.987	-1.187

sample averages of a well-designed experiment. Again, the proposed estimation method can further improve sample averages to obtain better estimates for the purpose of developing a better in-place aggregate inspection plan.

The small differences between the usual and proposed estimates probably have no impact on the inspection of very poor and good quality aggregates. This is because any reasonably good inspection plan has a high chance of rejecting poor and accepting good quality material. However, when the aggregate quality is on the margin of the specification limits, small differences become important in the role of rejecting poor quality aggregate. If there is a penalty system imposed upon the inferior material, small estimation improvements could mean a substantial saving in construction cost.

We note that usual estimates (sample averages) are statistically independent in pavement resurfacing and highway noise programs, but correlated in aggregate testing and acceptance programs. In general, the correlated case requires more samples than the independent case to achieve the same percentage estimation improvement. Thus, good supplemental data in the correlated case are almost essential to obtain substantial estimation improvement. We now provide one example for each case to demonstrate how to use the proposed method to improve sample averages.

Example 1: The peak noise level generated by an isolated commercial vehicle traveling on a roadway is measured. The speed and number of axles of this vehicle are also observed. Our initial investigation on 226 data points surveyed by the Michigan Department of Transportation found that the peak noise level is almost a linear function of speed and number of axles. Moreover, the variance of the peak noise level does not depend on traveling speed and vehicle type. Capitalizing on this linear relationship, we combined vehicles that have speeds within 2 mph to increase the reliability and normality of sample averages of peak noise levels and speeds. The data thereby obtained are presented in Table 14. The linear relationship among these variables is used to obtain initial estimates of true noise levels. The estimated results obtained from using the computer program of Part VII are also presented in Table 14. The total percentage improvement of the proposed method over the usual one is 47.94 percent. We observe from Table 14 that some of the differences between sample averages and proposed estimates are substantial and statistically significant. Also, the usual and proposed rank orders are significantly different.

Example 2: To estimate aggregate degradation due to transporting and compaction processes, we take samples from production and construction

Table 15
Aggregate Gradation Differences
Between Production and Construction Sites

Sieve Size	Changes in Gradation	
	Sample Average	New Estimate
3/4-In.	1.9933	2.0167
1/2-In.	3.3967	3.3670
3/8-In.	4.0433	4.0129
No. 4	4.0333	4.0543
No. 8	3.8300	3.8636
No. 16	3.5200	3.5308
No. 30	3.1200	3.0930
No. 50	2.4667	2.4529
No. 100	1.7833	1.8027
NO. 200	1.5200	1.5048
L.B.W.	1.2750	1.2829

Table 16
Aggregate Gradation Differences
Between Production and Construction Sites

Sieve Size	Previous Estimate	Change in Gradation	
		Sample Averge	Proposed Estimate
3/4-In.	2.39	1.9933	2.1501
1/2-in.	3.78	3.3967	3.5481
3/8-in.	4.33	4.0433	4.1566
No. 4	4.62	4.0333	4.2651
No. 8	3.95	3.8300	3.8774
No. 16	3.56	3.5200	3.5358
No. 30	3.01	3.1200	3.0765
No. 50	2.31	2.4667	2.4048
No. 100	1.82	1.7833	1.7978
No. 200	1.47	1.5200	1.5002
L.B.W.	1.28	1.2750	1.2770

sites. The changes in aggregate gradation (percent passing various sieves), based on 30 samples each, are presented in Table 15. We observe from the data that the change in gradation is approximately a third order polynomial function of sieve size. Note that, if this relationship is used to compute initial estimates, the percentage improvement will be low because four unknown coefficients of a polynomial function are too many relative to the number of parameters, 11, to be estimated in this probelm (as explained in Part III and, also, Part VIII). The estimated results obtained from using the computer program of Part VIII for this case are also presented in Table 15. As expected, the percentage improvement is only 8.6 percent. Consequently, sample averages and proposed estimates are almost the same. We now use the experimental results of other locations as initial estimates, the data and estimated results are presented in Table 16. We see from this table that the two experimental test results almost agree on fine aggregates. This is reasonable because the variances of test results on fine aggregates are much less than those on coarse aggregates (see examples in Part VIII). The total percentage improvement of the proposed method over the usual one in this case is 23.77 percent. However, the usual and proposed estimates are practically the same. This is because two experiments provide almost the same usual aggregate degradation estimates. This ensures us that either the averages of two experimental results or the proposed estimates (weighted averages of two experimental results) can be used to adjust specification limits. Of course, the reliability of these estimates is higher than the usual ones (based on one experiment). Thus, even in this situation, the proposed method indirectly provides better estimates than the usual ones.

III
THEORETICAL DEVELOPMENT OF STEIN-LIKE ESTIMATION PROCEDURES

1) Background

Let $X = (X_1, \ldots, X_k)'$ be k-variate normal with mean vector $= (\theta_1, \ldots, \theta_k)'$ and covariance matrix Σ. The superscript $'$ stands for the transpose of a vector or matrix. We are interested in estimating θ under the quadratic loss function

$$L(\theta, \hat{\theta}) = (\hat{\theta} - \theta)' \Sigma^{-1} (\hat{\theta} - \theta) \tag{37}$$

where $\hat{\theta}$ is an estimate of θ.

The usual estimator of θ is X. The risk of the usual estimator is

$$R(\theta, X) = E\,L(\theta, X) = E\,(X - \theta)' \Sigma^{-1} (X - \theta) = k \tag{38}$$

The notation 'E' stands for the expectation of a random variable. For $k \leq 2$, the usual estimator is shown to be admissible by Stein (16). However, for $k \geq 3$, James and Stein (1) have shown that the following estimator is better than the usual one.

$$\delta(X) = (1 - c)\,X \tag{39}$$

where, if the covariance matrix is the identity matrix,

$$c = (k - 2)/X'X \tag{40}$$

if the covariance matrix is σ^2 times the identity matrix, σ^2 is unknown, S^2/σ^2 is a chi-square with n degrees of freedom and is independent of X,

$$c = \frac{k-2}{n+2} \cdot \frac{S^2}{X'X} \tag{41}$$

and, if the covariance matrix is unknown, S is a k by k Wishart matrix with parameter n and Σ and independent of X,

$$c = \frac{k-2}{n-k+3} \cdot \frac{1}{X'S^{-1}X} \tag{42}$$

This estimator shrinks X toward the origin. One simple extension is to shrink X toward a given vector $U = (U_1, \ldots, U_k)'$ which is independent of X. That is, we have

$$\delta(X) = U + (1 - c)(X - U) \tag{43}$$

The column vector U can be interpreted as an initial estimate of θ. The above estimator is substantially better than the usual one when U is near θ. Thus, the key to obtaining a good estimate is the providing of a good initial estimate.

Past experiments and parallel studies are the usual sources for initial estimate of θ. However, for many reasons, these sources may not be available or reliable enough for our problem. Naturally, one is interested in constructing estimators that use the usual estimate X and supplemental data to obtain an initial estimate for the purpose of obtaining a better final estimate of θ. In this case, U is no longer independent of X and, consequently, the above estimator needs to be modified. For example, for the case that $\Sigma = I$,

$$\delta(X) = U + \left[1 - \frac{k-3}{(X-U)'(X-U)}\right](X - U) \tag{44}$$

where

$$U = (\bar{X}, \ldots, \bar{X})' \tag{45}$$

and

$$\bar{X} = \frac{1}{k}\sum_{i=1}^{k} X_i \tag{46}$$

This estimator, suggested by Lindley (in Stein (17), pp 285-297), shrinks all X_i toward the sample average $\bar{X}$. The improvement of this estimator over the usual one is substantial if all θ_i are near $\bar{\theta}$, the average of $\theta_1, \ldots$, and θ_k. In another words, this estimator is an excellent estimator if the total parameter variation, $\sum_{i=1}^{k} (\theta_i - \bar{\theta})^2$, is small. We shall extend this idea to cover more general forms of initial parameter estimates.

It can be verified that the initial estimate U defined in Eqs. (45) and (46) satisfies $X - U = PX$ and P is a k by k matrix of the following form:

$$P = \begin{bmatrix} (k-1)/k & -1/k & \cdot\cdot & -1/k \\ -1/k & (k-1)/k & \cdot\cdot & \cdot \\ \cdot & \cdot & \cdot\cdot & \cdot \\ \cdot & \cdot & \cdot\cdot & \cdot \\ \cdot & \cdot & \cdot\cdot & \cdot \\ -1/k & -1/k & \cdot\cdot & (k-1)/k \end{bmatrix} \qquad (47)$$

Furthermore, P is symmetrical and idempotent of rank $(k - 1)$. That is, $P' = P$ and $PP = P$. In this case, Rank (P) = Trace (P). Thus, there is motivation to investigate estimators of the form defined in Eq. (41) with U satisfying the following three conditions:

C.1) $X - U = Px$

C.2) Elements of P are independent of X

and

C.3) P is an idempotent matrix.

Throughout this part, $U = (U_1, \ldots, U_k)'$ stands for the initial estimate of θ and P is the corresponding matrix satisfying the above three conditions. The major theorem and techniques are presented in the next section. This section deals with the simplest case that the population covariance matrix Σ is known. Common examples are given in Section 3. The applications of the simplest case are discussed in Sections 4 and 5. We then study the case that the population covariance is of the form $\sigma^2\Sigma$ The results for the case that σ^2 is unknown and Σ is known are presented in Section 6. The reverse case is studied in Section 7.

2) Population Covariance Σ is Known

Throughout this section we assume that $X = (X_1, \ldots, X_k)'$ is k-variate normal with mean vector $\theta = (\theta_1, \ldots, \theta_k)'$ and the known covariance matrix Σ. We study an estimator of the form

$$\delta(X) = U + \left[1 - \frac{b}{(X - U)'\Sigma^{-1}(X - U)}\right](X - U) \qquad (48)$$

In order to express the risk function of δ, we first establish the following lemmas.

Lemma 1 - If Y is noncentral chi-square with k degrees of freedom and noncentrality parameter η, we have, for $m < k/2$,

$$E\ Y^{-m} = 2^{-m}\ E\ \frac{\Gamma(k/2 - m + W)}{\Gamma(k/2 + W)} < \infty \tag{49}$$

where Γ is the usual gamma function and W is a Poisson random variable with parameter $\lambda = \eta/2$.

Proof: The result is obtained through straightforward integration.

Lemma 2 - Let Z be k-variate normal with the identity covariance matrix. For any k by k symmetric and non-negative definite matrix A such that Rank (A) $> m > 0$, we have

$$E\ (Z'AZ)^{-m/2} < \infty \tag{50}$$

Proof: There exists an orthogonal matrix Q such that $A = Q'DQ$ where D is a diagonal matrix of non-negative elements. The number of positive diagonal elements is $n > m$. Denote a to be the minimum of those positive elements. Since $(QZ)'QZ$ is a noncentral chi-square with n degrees of freedom, we have from Lemma 1 that

$$E\ (Z'AZ)^{-m/2} = E\left[(QZ)'DQZ\right]^{-m/2} < a^{-1}\ E\left[(QZ)'QZ\right]^{-m/2} < \infty$$

Lemma 3 - For any square matrix Q, Rank $(Q'Q)$ = Rank (Q).

Proof: The proof is quite simple and is omitted.

Lemma 4 - Let Z be k-variate normal with mean vector θ and identity covariance matrix. If Q is a k by k idempotent matrix of rank > 2, we have

$$E\ \frac{(Z - \theta)'QZ}{Z'\theta'\theta Z} = \frac{\text{Trace}\ (Q) - 2}{Z'\theta'\theta Z} < \infty \tag{51}$$

Proof: The result can be directly established by using Hudson's Natural Identity (18) together with Lemmas 1 through 3.

We are now ready to express the risk function of the estimator defined in Eq. (48). We first note that $Z = \Sigma^{-1/2}X$ is k-variate with mean vector $\Sigma^{-1/2}\theta$ and identity covariance matrix. Also, $Q = \Sigma^{-1/2}\ P\Sigma^{1/2}$ is an idempotent matrix with Trace (Q) = Trace (P) and Rank (Q) = Rank (P).

Under the assumption that Rank (P) $>$ 2, we obtain from Lemma 4 that

$$R(\theta, \delta) = k - 2bE\left[\frac{\text{Trace}(P) - 2}{X' P' \Sigma^{-1} PX}\right] + b^2 E\left[\frac{1}{X' P' \Sigma^{-1} PX}\right]$$

When Trace (P) $>$ 2, the above risk function is minimized at b = Trace (P) - 2. This completes the proof of the following main theorem.

Theorem 1 - If Trace (P) $>$ 2 and Rank (P) $>$ 2, the estimator defined in Eq. (48) with b = Trace (P) - 2 is better than the usual one. The risk function of this estimator is

$$R(\theta, \delta) = k - \left[\text{Trace}(P) - 2\right]^2 \cdot E\left[\frac{1}{(X - U)' \Sigma^{-1} (X - U)}\right] \tag{52}$$

We note that when $\Sigma = I$ and U = 0 (consequently, P = I), this estimator is the Stein estimator defined in Eq. (38). We remark that, if P in Theorem 1 is also symmetrical, $(X - U)' \Sigma^{-1} (X - U)$ is a chi-square with n degrees of freedom and noncentrality parameter η, where

$$n = \text{Trace}(P) = \text{Rank}(P) \tag{53}$$

and

$$\eta = \theta' P' \Sigma^{-1} P \theta \tag{54}$$

Consequently, $R(\theta, \delta)$ is computable. The risk is k - Trace (P) + 2, when $\eta = 0$, and increases to k as η increases to infinity. The behavior of the risk function can be better explained by Figure 3. This figure indicates that it is essential to have η low and Trace (P) high in order to obtain good parameter estimates. Unfortunately, we shall see later that Trace (P) and η generally operate in the same direction.

In the next section, we provide practical forms of the initial estimate of the parameter. These practical forms will be the options of computer programs written for solving practical problems such as those discussed in Part II.

3) Practical Forms of Initial Estimates

Example 1: Shrinking X Toward an Independent Initial Estimate - When the initial estimate is obtained from past experiments or parallel studies, we shall use the following estimator to estimate the unknown parameter.

$$\delta(X) = U + \left[1 - \frac{k - 2}{(X - U)' \Sigma^{-1} (X - U)}\right](X - U) \quad , \tag{55}$$

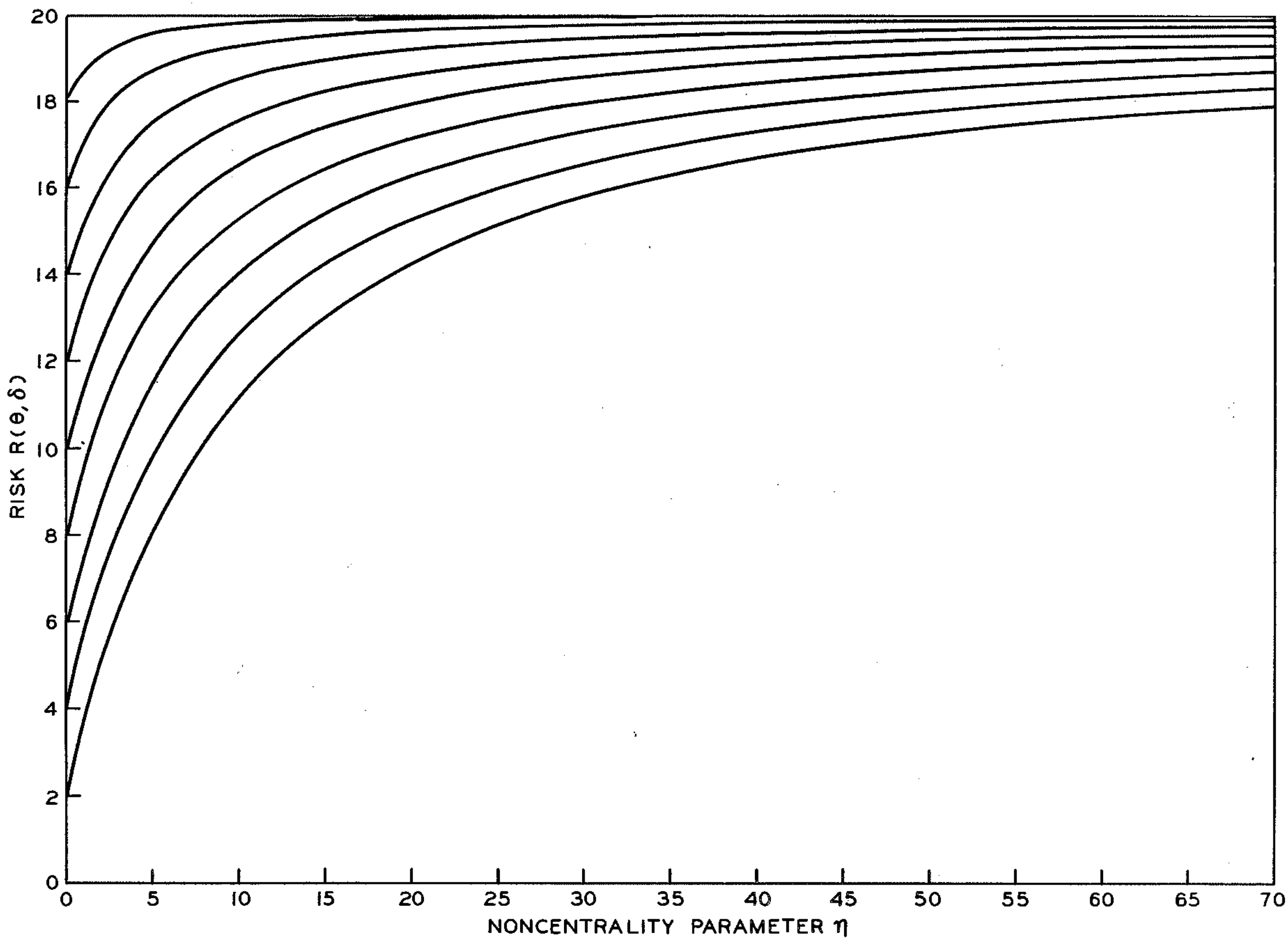

Figure 3. Family of risk functions for the case that k = 20.

The improvement of this estimator over the usual one is substantial if the initial estimate is near the parameter.

Example 2: Shrinking Every X_i Toward the Average of $X_1, \ldots, $ and X_k - In this case, U and P were defined in Eqs. (45) and (46), respectively. The trace of P is k - 1. Thus, by Theorem 1, the following estimator is better than the usual one for $k > 3$.

$$\delta(X) = U + \left[1 - \frac{k-3}{(X-U)'\Sigma^{-1}(X-U)}\right](X-U) \tag{56}$$

The improvement of this estimator over the usual one is substantial if the total parameter variation, $\sum_{i=1}^{k} (\theta_i - \bar{\theta})^2$, is small

Example 3: Shrinking Every X_i Toward the Weighted Average of $X_1, \ldots, X_k$ - For a given set of real numbers, w_i, $i = 1, \ldots, k$, satisfying $w_1 + \ldots + w_k = 1$, we define $U_1 = \ldots = U_k = \sum_{i=1}^{k} w_i X_i$. That is, U_i is the weighted average of usual estimates. Then, $X - U = PX$ with P defined as

$$P = \begin{bmatrix} 1-w_1 & -w_2 & & -w_k \\ -w_1 & 1-w_2 & \cdots & -w_k \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ -w_1 & -w_2 & \cdots & 1-w_k \end{bmatrix} \tag{57}$$

It can be easily verified that the above matrix is idempotent with Rank (P) = Trace (P) = k - 1. We note that P is not a symmetrical matrix unless $w_1 = \ldots = w_k$. Thus, by Theorem 1, the estimator defined in Eq. (56) is better than the usual one. The improvement of this estimator over the usual one is substantial if the total weighted parameter variation is small.

The next two examples demonstrate the techniques of using current data (usual estimate) and supplemental variables to obtain an initial estimate of θ.

Example 4: Shrinking X Toward an Initial Estimate Determined by Supplemental Data (Method I) - Often, θ_i can be expressed as the product of an unknown parameter b_i and a known variable t_i: i.e., $\theta_i = b_i t_i$. Generally speaking, t_i is supplemental data such as sample size, number of years or traffic volume, etc., involved in observing X_i. We shall term t_i as the 'supplemental' variable for θ_i. If every b_i is near the unknown constant b, we have approximately that $\theta_i = b\, t_i$ for every i. In this situation, we may take $U_i = \hat{b}\, t_i$ where $\hat{b}$ is an unbiased estimate of b defined as

$$\hat{b} = \sum_{i=1}^{k} X_i \Big/ \sum_{i=1}^{k} t_i \tag{58}$$

We now define $w_i = t_i \Big/ \sum_{j=1}^{k} t_j$. Then, we have that $U_i = w_i \sum_{j=1}^{k} X_j$ with $w_1 + \ldots + w_k = 1$ and $X - U = P'X$, where P was defined in Eq. (57). Thus, by Theorem 1, the estimator defined in Eq. (56) is better than the usual one if $k > 3$. For this estimator, we estimate θ_i to be

$$\delta_i(X) = \hat{b}\, t_i + \left[1 - \frac{k-3}{(X-U)'\Sigma^{-1}(X-U)}\right](X_i - \hat{b}\, t_i) \tag{59}$$

The improvement of this estimator over the usual one is substantial if $\sum_{i=1}^{k} (b_i - \hat{b})^2$ is small.

Example 5: Shrinking X Toward an Initial Estimate Determined by Supplemental Information (Method II) - Suppose that there are n supplemental variables satisfying the equation $\theta_i = b_{i1} t_{i1} + \ldots + b_{in} t_{in}$. Again, we assume that $B_i = (b_{i1}, \ldots, b_{in})'$ is near $B = (b_1, \ldots, b_n)'$ for every i. Then, we have approximately that $\theta = T\,B$, where T is the supplemental matrix of the following form.

$$T = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ t_{i1} & t_{i2} & \cdots & t_{in} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ t_{k1} & t_{k2} & \cdots & t_{kn} \end{bmatrix} \tag{60}$$

The least square estimate of B is $\hat{B} = (T'T)^{-1} T'X$. If we take $U = T\hat{B}$ as the initial estimate of θ, we have

$$P = I - T\,(T'T)^{-1}\,T' \tag{61}$$

It can be easily verified that P is a symmetrical idempotent matrix with Trace (P) = k - n. Thus, by Theorem 1, the estimator

$$\delta(X) = U + \left[1 - \frac{k - n - 2}{(X - U)'\Sigma^{-1}(X - U)}\right](X - U) \tag{62}$$

is better than the usual one for $k > (n + 2)$. As previously mentioned, keeping Trace (P) high is a good practice for obtaining a good estimate of θ. Thus, the vector size of B should be kept small relative to the vector size of θ.

Example 6: Mixture of Above Examples - Consider a k by k matrix of P of the following form

$$P = \begin{bmatrix} P_1 & 0 & \cdots & 0 \\ 0 & P_2 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & P_m \end{bmatrix} \tag{63}$$

where every P_i is an idempotent matrix and 0 is the null matrix. It can be easily verified that P is also an idempotent matrix such that Rank (P) = $\sum_{i=1}^{k}$ Rank (P_i) and Trace (P) = $\sum_{i=1}^{k}$ Trace (P_i). Thus, if Trace (P) > 2 and Rank (P) > 2, the estimator

$$\delta(X) = U + \left[1 - \frac{\text{Trace }(P) - 2}{(X - U)'\Sigma^{-1}(X - U)}\right](X - U) \tag{64}$$

is better than the usual one. In this case, $U = (I - P)X$. We interprete the above estimator as follows:

a) θ is partitioned into m groups,

b) If θ_i is in the j-th group, X_i is shrunk toward the initial estimate U_j determined by P_j.

For example: supplemental information is available for the first k_1 parameters, but not for the last k_2 parameters, $k_1 + k_2 = k$. Furthermore, we have that $\theta_i = b_i t_i$, $i = 1, \ldots, k_1$. If b_i is near a constant b for $i = 1, \ldots, k$ and θ_i is near a constant $\bar{\theta}$ for $i = k_1 + 1, \ldots, k$, we may take P_1 and P_2 as defined in Example 5 ($k = k_1$) and Example 2 ($k = k_2$), respectively. In this case, Trace (P) = k - 2 and the above estimator shrinks X_i toward U_i which is defined as

$$U_i \begin{cases} \hat{b}\, t_i \,, & i = 1, \ldots\ldots, k_1 \\ \bar{X}_2 \,, & i = k_1 + 1, \ldots, k \end{cases} \tag{65}$$

where $\hat{b} = \sum_{i=1}^{k} t_i X_i \Big/ \sum_{i=1}^{k} t_i^2$ and $\bar{X}_2$ is the average of X_i, $i = k_1 + 1, \ldots, k$.

As previously mentioned, the key to obtaining good estimates is to keep the trace of P high and η defined in Eq. (54) low. Unfortunately, the only way to reduce η is to properly increase the number of groups which decreases Trace (P) and, consequently, increase the total error. In general, if the reduction on η with an additional group does not compensate the loss of extra degrees of freedom (use higher up curves in Figure 3 to determine the squared error), the extra group is not beneficial and, therefore, should not be used.

4) Application of Theorem 1 to the k-Variate Poisson Problem

In this section, we assume that X_i is a Poisson random variable with parameter λ_i. Many measurements in the transportation field, such as traffic count and traffic accidents, possess this property. We also assume that $X_1, \ldots$, and X_k are independent. The usual estimator of $\lambda = (\lambda_1, \ldots, \lambda_k)'$ is $X = (X_1, \ldots, X_k)'$. We are interested in estimating λ by the method presented in the previous sections. For this purpose, we first use Anscombe's transformation (7) to transform X_i to Z_i defined as

$$Z_i = \sqrt{X_i + 0.375} \tag{66}$$

This transformed random variable is distributed rather more normally than X_i when λ_i is large. The mean of Z_i is approximately $\sqrt{\lambda_i}$ when λ_i is fairly large. Based on our computation, when $\lambda_i \geqq 5$, the variance of Z_i is near 0.25. That is, $Z = (Z_1, \ldots, Z_k)'$ is almost a k-variate normal with mean vector $\theta = (\theta_1, \ldots, \theta_k)'$ and covariance matrix 0.25 I if every λ_i is large enough, say ≥ 5. The parameter θ_i is defined as

$$\theta_i = E\, Z_i = \sum_{j=0}^{\infty} \sqrt{j + 0.375}\; \frac{e^{-\lambda_i} \lambda_i^j}{j\,!} \tag{67}$$

Applying Theorem 1 to the transformed random variables, we estimate θ to be

$$\delta_i(Z) = U_i + \left[1 - \frac{\text{Trace }(P) - 2}{4(Z - U)'(Z - U)} \right] (Z_i - U_i) \tag{68}$$

We then estimate λ_i to be the solution of Eq. (67) with θ_i replaced by $\delta_i(Z)$. Since Z is not exactly a normal random vector, we shall investigate whether the above estimator is still better than the usual one.

The risk function of the above estimator is too complicated to be analytically computed. Therefore, we simulate the risk of the above estimator in terms of the following four loss functions:

$$\sum_{i=1}^{k} \left| \theta_i - \delta_i(Z) \right| ; \quad \sum_{i=1}^{k} \left[\theta_i - \delta_i(Z) \right]^2 ; \quad \sum_{i=1}^{k} \left| \hat{\lambda}_i - \lambda_i \right| ; \quad \sum_{i=1}^{k} (\hat{\lambda}_i - \lambda_i)^2$$

In the above four loss functions, $\delta_i(Z)$ is the estimate of θ_i defined in Eq. (68) and $\hat{\lambda}_i$ is the corresponding estimate of λ_i. Based on extensive simulation, the above estimation procedures are also superior to the usual one in the global sense.

Since Theorem 1 was applied to the transformed random vector Z, methods for determining initial estimate presented in Examples 1 through 6 of Section 3 should be operated on the transformed parameter θ_i. The details will be discussed with examples in Part V.

5) Application of Theorem 1 to the k-Variate Binomial Problem

In this section we assume that X_i is a binomial random variable with parameter n_i (sample size) and p_i (unknown proportion). We also assume that $X_1, \ldots, X_k$ are independent. The usual estimator of p_i is

$$Y_i = X_i / n_i \tag{69}$$

It has been shown (2, 7) that the random variable Z_i, defined as

$$Z_i = \sqrt{n_i + 0.5}\ \text{Sin}^{-1}\left[\frac{n_i\,(2Y_i - 1)}{n_i + 0.75}\right] \tag{70}$$

is distributed rather more nearly normally than Y_i if $n_i\, p_i$ is large. Based on our computation, the variance of Z_i is near 1 for $n_i\, p_i \geq 4$. That is, $Z = (Z_1, \ldots, Z_k)'$ is almost a k-variate normal with mean vector $\theta = (\theta_1, \ldots, \theta_k)'$ and identity covariance matrix when $n_i\, p_i$ is fairly large for every i, say ≥ 4. The parameter θ_i is defined as

$$\theta_i = EZ_i$$

$$= \sqrt{n_i + 0.5}\ \sum_{j=0}^{n_i} \text{Sin}^{-1}\left[\frac{2j - n_i}{n_i + 0.75}\right] C_j^{n_i}\, p_i^{\,j}\, (1 - p_i)^{n_i - j} \tag{71}$$

Applying Theorem 1 to the transformed random variable Z, we first estimate θ_i to be

$$\delta_i(Z) = U_i + \left[1 - \frac{\text{Trace }(P) - 2}{(Z - U)'(Z - U)}\right](Z_i - U_i) \tag{72}$$

We then estimate p_i to be the solution of Eq. (71) with θ_i replaced by $\delta_i(Z)$. Another estimate, $\hat{p}_i$, of p_i can be obtained through the inverse function of Eq. (70). That is,

$$\hat{p}_i = \frac{1}{2}\left\{\frac{n_i + 0.75}{n_i}\ \text{Sin}\left[\frac{\delta_i(Z)}{n_i + 0.5}\right] + 1\right\} \tag{73}$$

Again, based on extensive simulation, the above estimation procedure is superior to the usual one.

Methods presented in Examples 1 through 6 of Section 3 should be used on the transformed parameter θ_i to determine initial parameter estimates. The details will be discussed with examples in Part VI.

6) <u>Population Covariance $\sigma^2\Sigma$ with σ^2 Unknown and Σ Known</u>

In this section we assume that $X = (X_1, \ldots, X_k)'$ is a k-variate normal with mean vector $\theta = (\theta_1, \ldots, \theta_k)'$ and covariance matrix $\sigma^2\Sigma$. We also assume that Σ is a known k by k matrix, but σ^2 is unknown. However, we independently observe S^2 distributed as σ^2 times a chi-square with n degree of freedom. We study an estimator of the form

$$\delta(X) = U + \left[1 - \frac{b\, S^2}{(X - U)' \Sigma^{-1} (X - U)} \right] (X - U) \tag{74}$$

If Rank (P) > 2 and Trace (P) > 2, we obtain from Lemma 4 and the independent assumption of X and S^2 that the risk function of the above estimator is minimized at $b = [\text{Trace (P)} - 2] / (n + 2)$. This completes the proof of the following theorem.

Theorem 2 - If Trace (P) > 2 and Rank (P) > 2, the estimator defined in Eq. (74) with $b = [\text{Trace (P)} - 2]/(n + 2)$ is better than the usual one. The risk function of this estimator is

$$R(\theta, \delta) = k - \frac{n}{n + 2} [\text{Trace (P)} - 2]^2 \; E \left[\frac{\sigma^2}{(X - U)' \Sigma^{-1} (X - U)} \right] \tag{75}$$

If P is also symmetrical, $(X - U)' \Sigma^{-1} (X - U)/\sigma^2$ is chi-square with Trace (P) degrees of freedom and noncentrality parameter $\theta' P \Sigma^{-1} P\theta/\sigma^2$.

We now describe a general situation to which the above theorem can be applied to obtain a better estimate of θ. This is as follows: we observe $X_{i1}, \ldots, X_{in_i}$ from the i-th population with mean θ_i and variance σ^2, $i = 1, \ldots, k$. $\theta = (\theta_1, \ldots, \theta_k)'$ and σ^2 are unknown. Furthermore, all X_{ij} are independent. The usual estimator of θ_i is the sample average, $\bar{X}_i$, defined as

$$\bar{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij} \tag{76}$$

If the i-th population is normal, $\bar{X}_i$ is normally distributed. Otherwise, $\bar{X}_i$ is approximately normally distributed with mean θ_i and variance σ^2/n_i when n_i is fairly large. That is, when sample sizes are fairly large, $X = (\bar{X}_1, \ldots, \bar{X}_k)'$ is approximately a k-variate normal with mean vector θ and covariance matrix $\sigma^2 \Sigma$, where Σ is a diagonal matrix whose (ii)-th element is $1/n_i$. In this case, $S^2 = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$ is independent of X, and S^2/σ^2 is a chi-square with $(n_1 + \ldots + n_k - k)$ degrees of freedom. Thus, for any U satisfying $X - U = PX$, we estimate θ_i to be

$$\delta_i(X) = U_i + \left[1 - \frac{\text{Trace}(P) - 2}{\sum_{i=1}^{k} n_i - k + 2} \cdot \frac{S^2}{\sum_{i=1}^{k} (\bar{X}_i - U_i)^2 / n_i} \right] (\bar{X}_i - U_i) \qquad (77)$$

We note that the above estimation procedure should be used with caution when populations are non-normal. The key is to make sure that sample sizes are large enough to guarantee the normality assumption.

We now present a special application of Theorem 2 to linear regression analysis. For this purpose, we denote Y to be the vector of dependent observations, X to be the matrix of independent observations, B to be the vector of parameters, and Σ to be the vector of errors. Then, for the linear model $Y = XB + \Sigma$, the least squares estimate of B is

$$\hat{B} = (X'X)^{-1} X'Y \qquad (78)$$

In ordinary multiple regression analysis, the first column of X is $(1, \ldots, 1)'$ and Σ is normally distributed with null mean vector and covariance matrix $\sigma^2 I$, $\hat{B}$ is normally distributed with mean vector B, and covariance matrix $\sigma^2 (X'X)^{-1}$. Furthermore, the sum of squares of residuals, S^2, defined as

$$S^2 = (Y - X\hat{B})'(Y - X\hat{B}) \qquad (79)$$

is independent of $\hat{B}$ and is σ^2 times a chi-square with N-k degrees of freedom. N is the total number of observations (vector size of Y) and k is the vector size of B. By Theorem 2, the following estimator is better than the usual one for estimating B.

$$\delta(\hat{B}) = U + \left[1 - \frac{\text{Trace}(P) - 2}{N - k + 2} \cdot \frac{S^2}{(\hat{B} - U)'X'X(\hat{B} - U)} \right] (\hat{B} - U) \qquad (80)$$

7) Population Covariance Matrix $\sigma^2 \Sigma$ with σ^2 Known and Σ Unknown

In this section we assume that $X = (X_1, \ldots, X_k)'$ is a k-variate normal with mean vector $\theta = (\theta_1, \ldots, \theta_k)'$ and covariance matrix $\sigma^2 \Sigma$. We also assume that σ^2 is known, but Σ is unknown. However, we independently observe a k by k Wishart matrix S with parameters n and Σ. We study an estimator of the form

$$\delta(X) = U + \left[1 - \frac{b\,\sigma^2}{(X - U)' S^{-1} (X - U)} \right] (X - U) \qquad (81)$$

Using Lemma 4 and the known fact (7) that Y, defined as

$$Y = \frac{(X - U)'\Sigma^{-1}(X - U)}{(X - U)' S^{-1} (X - U)} \tag{82}$$

is chi-square distributed n - k + 1 degrees of freedom and independent of X - U, we find that the risk function of the above estimator is minimized at $b = [\text{Trace}(P) - 2]/(n - k + 3)$ if Trace (P) > 2 and $n > k - 1$. This completes the proof of the following theorem.

Theorem 3 - If Trace (P) > 2, Rank (P) > 2 and $n > k - 1$, the estimator defined in Eq. (81) with $b = [\text{Trace}(P) - 2)]/(n - k + 3)$ is better than the usual one. The risk function of this estimator is

$$R(\theta, \delta) = k\frac{n-k+1}{n-k+3}[\text{Trace}(P) - 2]^2 \cdot E\left[\frac{\sigma^2}{(X-U)'\Sigma^{-1}(X-U)}\right] \tag{83}$$

If P is also symmetrical, $(X - U)'\Sigma^{-1}(X - U)/\sigma^2$ is a chi-square with Trace (P) degrees of freedom and the noncentrality parameter $\theta' P'\Sigma^{-1} p\theta/\sigma^2$.

We now describe a general situation where the above theorem can be applied. We independently observe N sample vectors from a population with mean $\theta = (\theta_1, \ldots, \theta_k)'$ and covariance matrix Σ. Both θ and Σ are unknown. Denote $X_i = (X_{i1}, \ldots, X_{ik})'$ to be the i-th observation vector. Define $X = (\overline{X}_1, \ldots, \overline{X}_k)'$ with

$$\overline{X}_i = \sum_{j=1}^{N} X_{ji}/N \tag{84}$$

If N is fairly large, X is approximately a k-variate normal with mean vector θ and covariance matrix Σ/N. Moreover, $S = (X - X)'(X - X)$ is a Wishart matrix with parameter N - 1 and Σ, and is independent of X. Thus, we estimate θ_i to be

$$\delta_i(X) = U_i + \left[1 - \frac{\text{Trace}(P) - 2}{N(N - k + 2)} \cdot \frac{1}{(X - U)' S^{-1} (X - U)}\right](\overline{X}_i - U_i) \tag{85}$$

Again, the above estimation procedure should be used with caution when the population is non-normal. As before, one should ensure that the sample size N is large enough to satisfy the normality assumption.

We remark that estimators presented in previous sections are of the form defined in Eq. (43). These estimators can be slightly improved by setting the constant c to be 1 if it is greater than 1. By doing so, the new estimate is the weighted average of initial and usual estimates.

IV
GENERAL GUIDELINES FOR USING COMPUTER PROGRAMS OF PARTS V THROUGH VIII TO ESTIMATE PARAMETERS

Four computer programs are written based on theoretical results presented in Part III to handle various types of estimation problems. These computer programs are presented in Parts V through VIII. Thus, for a given problem, the first step is to identify the problem type so that the proper computer program can be selected.

Step 1: Identify the Problem Type

Denote X_i to be the usual estimate of θ_i, $i = 1, \ldots, k$. Computer programs of Parts V through VIII are designed to improve the following types of usual estimates.

A) X_i is Poisson distributed with parameter θ_i, $i = 1, \ldots, k$. Moreover, $X_1, \ldots,$ and X_k are independent. As mentioned in the first section of Part II, accident frequency and traffic counts (of various locations) are of this type. For this kind of problem, the computer program of Part V should be used to estimate Poisson parameters.

B) X_i is the sample proportion based on sample size n_i such that $n_i X_i$ is a binomial random variable with parameters n_i and p_i, $i = 1, \ldots, k$. Moreover, $X_1, \ldots,$ and X_k are independent. As mentioned in the second section of Part II, severity indices (of various fixed object accidents) and proportions of wet accidents (at various locations) are of this type. For this kind of problem, the computer program of Part VI should be used to estimate (proportion) parameters of binomial processes.

C) X_i is the sample average of n_i observations obtained from the i-th population such that X_i is normally distributed with mean parameter θ_i and variance σ^2/n_i, $i = 1, \ldots, k$. That is, k populations have the common unknown variance σ^2. We also independently observe S^2 which is σ^2 times a chi-square with N degrees of freedom. Moreover, $X_1, \ldots,$ and X_k are independent. As mentioned in the third section of Part II, sample averages of peak noise levels (generated by various vehicle types) and friction tests (of various locations) are of this type. For this kind of problem, the computer program of Part VII should be used to estimate mean parameters.

We remark that S^2 is obtained from either previous experiments or current data. In the latter case, S^2 is defined as

$$S^2 = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - X_i)^2$$

where X_{ij} is the j-th observation from the i-th population. In this case, $N = n_1 + \ldots + n_k - k$.

D) $X = (X_1, \ldots, X_k)'$ is the sample average vector of n observation vectors obtained from the population such that X is normally distributed with mean vector $\theta = (\theta_1, \ldots, \theta_k)'$ and covariance matrix Σ/n. In this case, $X_1, \ldots$, and X_k are correlated. We independently observe a Wishart matrix S with parameters N and Σ. As mentioned in Part II, aggregate gradation is of this type. For this kind of problem, the computer program of Part VIII should be used to estimate mean parameters.

We remark that S is obtained from either previous experiments or current data. In the latter case, the (i, j)-th element of S is defined as

$$S_{ij} = \sum_{h=1}^{n} (X_{ih} - X_i)(X_{jh} - X_j)$$

where $(X_{1h}, \ldots, X_{kh})'$ is the h-th observation vector. In this case, $N = n - 1$.

Step 2: Select the Option of Computing Initial Estimates

We have stressed that providing good initial estimates is the key to obtaining good final estimates of parameters. Thus, this is the most important step of the estimation procedure.

Quite often, parameters can be partitioned into groups within which they are closely related in some fashion. With this in mind, each computer program contains many options for computing initial estimates of parameters in each group. These options can be categorized into the following three types.

Type A - Initial estimates are given for a group of parameters.

Type B - For a group of parameters clustering at one point, the (weighted) average of usual estimates is used as the estimate of each parameter in the group.

Type C - For a group of parameters that can be approximated by a linear function of supplemental variables, least squares estimates are used as initial parameter estimates.

In order to utilize these options to obtain good initial estimates, one can use past experience or graphically examine the relationship among usual estimates and supplemental data to properly group parameters. Ideally, the number of groups is small relative to the number of parameters undergoing estimation and, also, parameters within a group are closely related according to one of the above options. Grouping techniques for each type of problem are explained with examples in Parts V through VIII.

Step 3: Run the Computer Program

The final step is to arrange the data into the format specified by the computer program. The input and output formats of computer programs can be found in Parts V through VIII.

V
A COMPUTER PROGRAM FOR SIMULTANEOUSLY ESTIMATING POISSON PARAMETERS BY USING STEIN-LIKE ESTIMATION PROCEDURES

In this part, we provide a FORTRAN Computer Program for simultaneously estimating k Poisson parameters, $\lambda_1, \ldots, \lambda_K$. This program is written based on the theoretical results and grouping techniques presented in Sections 2 and 3 of Part III.

There are eight sections in this part. In Section 1, we describe the basic data required by the program. The estimation procedures are outlined in Section 2. Six methods for computing initial estimates of parameters are discussed in Section 3. The input format of the basic data is presented in Section 4. When the option of transforming final estimates to other estimates is chosen, the user must provide the transformation to the user-supplied subroutine EQN. An example for this purpose is provided in Section 5. The instructions for modifying the program, if needed, to fit a particular problem are given in Section 6. Seven examples are presented in Section 7 which demonstrate the use of various program options and logic for computing initial estimates of parameters. The program listing is presented in the last section.

1) The Basic Data

The essential data for estimating parameters are the independent observations $X_1, \ldots,$ and X_k. The observation X_i is sampled from the i-th population which is Poisson distributed with parameter λ_i.

Other essential data is the information for computing initial estimates of parameters. This information is termed the 'supplemental' information for discussion purposes. Six methods for using the supplemental information to compute initial estimates of parameters are described in Section 3.

Quite often, one is also interested in transforming λ_i to another parameter η_i through the function F. That is,

$$\eta_i = F(\lambda_i) \tag{86}$$

When this option is chosen, the user must provide the functional form of F to the user-supplied subroutine EQN. The information required by F must also be provided. This information is termed the 'auxiliary' information.

2) What the Program Does

The first step of this program is to use Anscombe's transformation (7) to transform X_i to Z_i. That is,

$$Z_i = \sqrt{X_i + 0.375} \tag{87}$$

It has been shown in Part III that, when λ_i is at least 5, Z_i is nearly normally distributed with mean θ_i and variance 1/4, where

$$\theta_i = e^{-\lambda_i} \sum_{j=0}^{\infty} \sqrt{j + 0.375}\, \lambda_i^j / j! \tag{88}$$

The second step is to compute the initial estimate, U_i, of θ_i by the chosen method. In this step, the trace of the idempotent matrix P satisfying the following equation

$$(Z_1, \ldots, Z_k)' - (U_1, \ldots, U_k)' = P\,(Z_1, \ldots, Z_k)' \tag{89}$$

is also computed. The superscript $'$ stands for the transpose of a vector or matrix. The third step is to compute the shrinking factor c defined as

$$c = \frac{\text{Trace (P)} - 2}{4 \sum_{i=1}^{k} (Z_i - U_i)^2} \tag{90}$$

A slightly better estimation procedure is to set c at 1 if it is greater than 1. The fourth step is to compute $\hat{Z}_i$ defined as

$$\hat{Z}_i = U_i + (1 - c)\,(Z_i - U_i),\ i = 1, \ldots, k \tag{91}$$

$\hat{Z}_i$ is the Stein-like estimate of θ_i. The fifth step is to obtain $\hat{X}_i$ which is the λ_i satisfying Eq. (88) with θ_i replaced by $\hat{Z}_i$. $\hat{X}_i$ is also considered a Stein-like estimate of λ_i. If the option of transforming to λ_i specified in Eq. (86) is chosen, the last step is to estimate η_i to be

$$\hat{\eta}_i = F(\hat{X}_i) \tag{92}$$

$\hat{X}_i$ and $\hat{\eta}_i$ are the final estimates of λ_i and η_i, respectively. This program also computes the estimated percentage improvement of the above procedure over the usual one.

3) Methods for Computing Initial Estimates

The key to obtaining good estimates of parameters is to provide good initial estimates $U_1, \ldots, U_k$ in the sense that N is high and ζ is low, where

$$N = \text{Trace}\ (P) \tag{93}$$

and

$$\zeta = 4 \sum_{i=1}^{k} (U_i - \theta_i)^2 \tag{94}$$

We have shown in Part II that, if P is a symmetrical idempotent matrix, N and ζ are, respectively, the number of degrees of freedom and the noncentrality parameter of a noncentral chi-square distribution. For this case, the maximal percentage improvement that can be achieved is $100(N - 2)/k$. Six methods for computing initial estimates of parameters are built in the program. These are:

Method 1: Initial Estimates are Given - Based on past experiments or independent parallel studies, we estimate or guess λ_i to be V_i, $i = 1, \ldots, k$. The initial estimate of θ_i is then obtained from Eq. (88). That is,

$$U_i = e^{-V_i} \sum_{j=0}^{\infty} \frac{\sqrt{j + 0.375}\ V_i^j}{j\,!} \tag{95}$$

In this case, we treat $U_i - \theta_i$ as the parameter to be estimated. Consequently, the matrix P satisfying Eq. (89) is the identity matrix. Thus, N = Trace (P) = k. We note that the maximal trace of P in Eq. (89) is k. The estimation accuracy is the degree of closeness of $(V_1, \ldots, V_k)$ to $(\lambda_1, \ldots, \lambda_k)$. The final estimate of λ_i always lies between the initial estimate V_i and the usual estimate X_i. When the initial estimates are excellent, say $V_i = \lambda_i$ for all i, the final estimate of λ_i is V_i for all i. However, if the initial estimates are poor, i.e., $(V_1, \ldots, V_k)$ is quite distant from $(\lambda_1, \ldots, \lambda_k)$, the final estimate of λ_i will be very close to the usual estimate X_i for every i.

This method is used only when the number of parameters to be estimated is at least 3. When the independent initial estimates are reliably close to the true parameters, this method will provide good parameter estimates.

Method 2: Weighted Average (I) - It is not unusual that the unknown parameter θ_i can be expressed as the product of an unknown parameter b_i and a known variable t_i. That is, $\theta_i = b_i t_i$. In this case, t_i is a supplemental variable. If the variation among $b_1, \ldots,$ and b_k is small, i.e., $(b_1 - b)^2 + \ldots + (b_k - b)^2$ is small for some b, we may write

$$\theta_i = b\, t_i, \; i = 1, \ldots, k \tag{96}$$

The weighted average $\hat{b}$ can then be used to estimate b. That is,

$$\hat{b} = \sum_{i=1}^{k} Z_i \Big/ \sum_{i=1}^{k} t_i \tag{97}$$

We then take $U_i = \hat{b}\, t_i$ as the initial estimate of θ_i. We note that, when all t_i are equal, U_i is the average of $Z_1, \ldots,$ and Z_k for every i.

For this method, we have N = Trace (P) = k - 1. The loss of one degree of freedom is due to the use of $\hat{b}$ as an estimate of b. We note that the above U_i can be rewritten as

$$U_i = w_i \sum_{j=1}^{k} Z_j \tag{98}$$

with

$$w_i = t_i \Big/ \sum_{j=1}^{k} t_j \tag{99}$$

It is obvious that

$$w_1 + w_2 + \ldots + w_k = 1 \tag{100}$$

This method is used only when the number of parameters to be estimated is at least 4. When the supplemental variable t_i is available and, approximately, $\theta_i = b\, t_i$, $i = 1, \ldots, k$, this method will produce good estimates of parameters. In practice, one may use this method if the relationship, $Z_i = b\, t_i$, holds approximately for a supplemental variable t_i, $i = 1, \ldots, k$.

Method 3: Weighted Average (II) - For a given set of numbers, $w_1, \ldots, w_k$, satisfying Eq. (100), the initial estimate of θ_i is taken to be

$U_i = w_i \sum_{j=1}^{k} Z_j$. For this case, N = Trace (P) = k - 1. We note that this is the same form used in Method 2. The only difference is that w_i in Method 2 was computed from the supplemental data, but is given in this method. When $t_1 = . . . = t_k$ and $w_1 = . . . = w_k$, Methods 2 and 3 are identical and use the average of $Z_1, . . . , Z_k$ as the initial estimate of every parameter.

This method is used only when the number of parameters to be estimated is at least 4. When the relationship $\theta_i = b\, w_i$ holds approximately for a given set of numbers satisfying Eq. (100), this method will produce good estimates of parameters. In practice, one may use this method if the linear relationship, $Z_i = b\, w_i$, holds approximately for a set of supplemental numbers $w_1, . . . , w_k$ satisfying Eq. (100). For example, we have approximately $Z_i = b\, t_i$, then $w_i = t_i / \sum_{j=1}^{k} t_j$. We note that, if $t_i = 1$, then $w_i = 1/k$.

Method 4: Weighted Average (III) - For a given set of numbers, $w_1, . . . , w_k$, satisfying Eq. (100), the initial estimate of θ_i is taken to be $U_i = \sum_{j=1}^{k} w_j Z_j$. For this case, N = Trace (P) = k - 1. Methods 3 and 4 are identical when all w_i are equal to 1/k.

This method is used only when the number of parameters to be estimated is at least 4. When all θ_i are almost equal to a constant, this method will produce good estimates of parameters for any set of numbers satisfying Eq. (100). w_i can be interpreted as the weight assigned to Z_i for estimating the common constant. If every X_i or Z_i is equally reliable, we assign $w_i = 1/k$, i = 1, . . . , k. If one has reason to doubt the reliability of a particular observation, say X_1, zero weight may be assigned to Z_1 and an equal weight assigned to the remaining observations. That is, $w_1 = 0$ and $w_i = 1/(k - 1)$, i = 2, . . . , k.

Method 5: Least Squares Estimate (I) - Suppose that the supplemental data $(t_{i1}, . . . , t_{im})$ are available and satisfy the following equation,

$$\theta_i = b_{i1} t_{i1} + . . . + b_{im} t_{im}, \quad i = 1, . . . , k \tag{101}$$

If $b_{1j}, . . . ,$ and b_{kj} are near an unknown common parameter b_j, Eq. (101) can be rewritten as

$$\theta_i = b_1 t_{i1} + . . . + b_m t_{im}, \quad i = 1, . . . , k \tag{102}$$

Denote $(\hat{b}_1, \ldots, \hat{b}_m)$ to be the least squares estimate of $(b_1, \ldots, b_m)$. We then take the initial estimate of θ_i to be

$$U_i = \hat{b}_1 t_{i1} + \ldots + \hat{b}_m t_{im}, \; i = 1, \ldots, k \tag{103}$$

In this case, N = Trace (P) = k - m. The loss of m degrees of freedom is due to the least squares estimates of m linear parameters. We note that U_i can always be improved by increasing the number of supplemental variables, i.e., by increasing m. However, this decreases N which is an undesirable property as previously mentioned. In general, m should be kept small relative to the number of parameters k.

This method is used only when the number of parameters to be estimated is at least m + 3. This method will produce good estimates of parameters if Eq. (102) holds approximately and the number of supplemental variables is small relative to the number of parameters to be estimated. In practice, one may use this method if the following linear relationship holds approximately.

$$Z_i = b_1 t_{i1} + \ldots + b_m t_{im}, \; i = 1, \ldots, k \tag{104}$$

Method 6: Least Squares Estimate (II) - This method is the same as Method 5 except that Eqs. (102) and (103) are, respectively, replaced by

$$\theta_i = b_0 + b_1 t_{i1} + \ldots + b_m t_{im}, \; i = 1, \ldots, k \tag{105}$$

and

$$U_i = \hat{b}_o + \hat{b}_1 t_{i1} + \ldots + \hat{b}_m t_{im}, \; i = 1, \ldots, k \tag{106}$$

In this case, N = Trace (P) = k - m - 1 because of the extra parameter b_0. Thus, this method is used only when the number of parameters to be estimated is at least m + 4. In practice, one may use this method if the following linear relationship holds approximately

$$Z_i = b_0 + b_1 t_{i1} + \ldots + b_m t_{im}, \; i = 1, \ldots, k \tag{107}$$

The k parameters may be arranged into many groups to which different methods are applied to minimize the noncentrality parameter defined in Eq. (94). For example, we may use Method 1 for a group of k_1 parameters, Method 5 for a group of k_2 parameters, Method 2 for a group of k_3 parameters, and again, Method 2 for the group of the remaining ones. In this case, we have $N = \text{Trace}(P) = k_1 + (k_2 - m) + (k_3 - 1) + (k - k_1 - k_2 - k_3 - 1) = k - m - 2$. m is the number of supplemental variables used in Method 5. For this particular case, the number of parameters to be estimated should be at least m + 5. In general, for any method or combination of methods, the number of degrees of freedom should be at least 2. The use of combining the above six methods to compute initial parameter estimates will be clearly demonstrated by examples presented in Section 7.

Every method except the first one suggests that k populations should be arranged into groups in which parameters cluster at a point or can be approximated by a linear function of supplemental variables. The sources for obtaining proper group-method combinations are past experiments, parallel studies, and physical properties of populations. Unfortunately, these sources may not be available or reliable enough in a particular problem. In this circumstance, one may examine the data as suggested in each method to choose the proper method-group combination. That is, the empirical relationship among X_i, Z_i and supplemental variables can be used to determine the method-group combination for computing initial estimates of parameters. Since the program computes the estimated percentage improvement of each chosen method-group combination over the usual method, one may use the one that produces the maximal improvement to estimate parameters. We remark that the method-group combination generated by examining the data may not be the best one and, possibly, could be the worst one for this problem due to random variation of the data. But no matter what the true case is, if the chosen method-group combination is used thereafter for the same problem, the above method is always better than the usual one. The worst situation is that no improvement is made. Based on our experience, qualitative properties of populations are useful bases for grouping purposes.

4) Data Input

The data input of this program is arranged into two portions. The first portion is composed of six cards. The number of cards in the second portion is equal to the number of parameters to be estimated.

The first six cards specify the number of groups and parameters, computational methods, and supplemental and auxiliary variables in each group. Variables used in these cards are defined below.

NP: Number of parameters to be estimated, $1 \leq NP \leq 500$

NGROUP: Number of groups used, $1 \leq NGROUP \leq 20$

NMG(I): Number of parameters in the I-th group, NMG(1) + . . . + NMG (NGROUP) = NP

METHOD(I) = j: The j-th method presented in Section 3 is used to compute initial estimates of parameters in the I-th group, $1 \leq j \leq 6$

NAUX(I): Number of supplemental variables used to compute initial estimates of parameters in the I-th group. This variable is 1 if METHOD(I) $\leq$ 4.

NCOV(I):

- >0: Number of auxiliary variables used to transform parameters in the I-th group, $1 \leq$ NAUX(I) + NCOV(I) ≤ 5
- $= 0$: Means that the supplemental variables also serve as auxiliary variables.
- <0: No transformation.

(DESCPT(I), I = 1, 70): Title (no more than 70 letters).

The second portion is composed of NGROUP subportions or groups. The I-th subportion is composed of NMG(I) cards. Define K = J if I = 1 and K = NMG(1) + . . . + NMG(I - 1) + J if I $>$1. Then, the K-th card of the second portion contains essential and supplemental data for estimating the J-th parameter of the I-th group. We note that the J-th parameter of the I-th group is the parameter of the K-th population. The data input for estimating this parameter is as follows:

NCDV(I)	DATA INPUT
< 0	ID(K),NX(K),(AUX(J,M),M=1,NAUX(I))
> 0	ID(K),NX(K),(AUX(J,M),M=1,NAUX(I)), (BUX(K,M),M=1,NCOV(I))

Variables used in this card are defined below:

ID(K): Identification number of the K-th population such as location number and year, etc. The K-th parameter is the J-th parameter of the I-th group. This number has no effect on the estimation procedure.

NX(K): The K-th sample (the usual estimate of the J-th parameter of the I-th group).

AUX(J, M): The M-th supplemental variable for the K-th parameter.

BUX(K, M): The M-th auxiliary variable for the K-th parameter.

The input deck is presented in Table 17. This deck is also diagrammed in Figure 4 to show the format and logic used. The user can follow this diagram to change, if needed, read statements and formats to fit a particular problem.

Table 17
The Input Deck

Card Number	Variables Used in Each Card	Remarks
1	NP , NGROUP	* The Data Input of The First Portion
2	(NMG(I), I=1,...,NGROUP)	
3	(METHOD(I), I=1,...,NGROUP)	
4	(NAUX(I),I=1,NGROUP)	
5	(NCOV(I),I=1,NGROUP)	
6	(DESPT(I),I=1,70)	
7	ID(1),NX(1),(AUX(1,M),M=1,NA), (BUX(1,M),M=1,NC)	* L=NMG(1) * NA=NAUX(1);NC=NCOV(I) * L cards for The First Group
.		
.		
.		
L+6	ID(L),NX(L),(AUX(L,M),M=1,NA), (BUX(L,M),M-1,NC)	
⋮	⋮	⋮
.		* NA=NAUX(I);NC=NCOV(I)
.		* K=NMG(1)+..+NMG(I-1) +J
.		
K+6	ID(K),NX(K),(AUX(J,M),M=1,NA), (BUX(K,M),M=1,NC)	* This card is for the k-th parameter which is the J-th parameter of the I-th group.
.		
.		
⋮	⋮	⋮
N+6	ID(N),NX(N),(AUX(1,M),M=1,NA), (BUX(N,M),M=1,NC)	* L=NMG(NGROUP) * NA=NAUX(NGROUP) * NC=NCOV(NGROUP) * N=NP-L+1 * L Cards for The Last GROUP
.		
.		
.		
.		
NP+6	ID(NP),NX(NP),(AUX(L,M),M=1,NA), (BUX(NP,M),M=1,NC)	

5) <u>The User-Supplied Subroutine (EQN)</u>

The functional form of the transformation F defined in Eq. (86) must be specified between two statements, DO 100 I = N1, N2 and 100 CONTINUE, in the user-supplied subroutine EQN. The input variables for this subroutine are NP, ID, IG, NX, XEST, BUX, N1, and N2. The meaning of

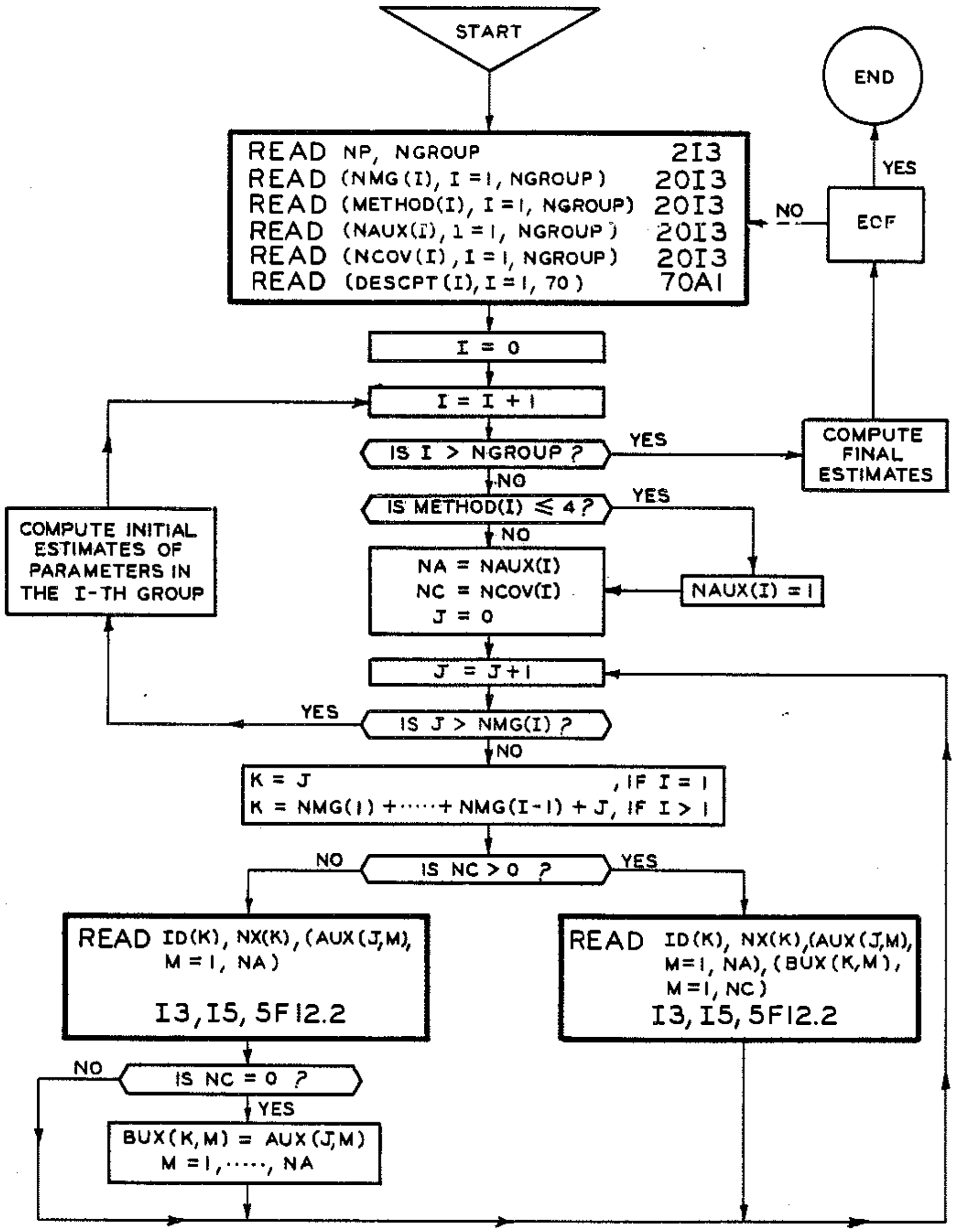

Figure 4. Flow of control for data input.

NP, ID, NX and BUX were explained in Section 4. IG(i) is the group number to which the i-th parameter belongs. XEST(i) is the final estimate of the i-th parameter. N1 and N2 are integers set automatically in the main program. The output of this subroutine is W2(i) and W3(i) for i = N1, . . . , N2, where W2(i) = F[X(i)] and W3(i) = F[XEST(i)].

Example: We are also interested in estimating parameter η_i satisfying the following equations.

$$\lambda_i = \eta_i t_i + (\eta_i t_i)^2, \quad i = 1, \ldots, 10 \tag{108}$$

and

$$\lambda_i = \eta_i M_i, \quad i = 11, \ldots, 100 \tag{109}$$

Thus, t_i is the auxiliary variable for the first 10 populations to convert λ_i to η_i defined in Eq. (108). For the remaining populations, M_i is the auxiliary variable for converting λ_i to η_i defined in Eq. (109). As an example, t_i and M_i can be, respectively, the ADT and number of years involved in measuring the usual estimate X_i. In this case, we have

$$\text{BUX}(i, 1) = \begin{cases} t_i, & i = 1, \ldots, 10 \\ M_i, & i = 11, \ldots, 100 \end{cases} \tag{110}$$

Solving Eqs. (108) and (109), we obtain

$$\eta_i = F(\lambda_i) = \begin{cases} \dfrac{1}{2t_i}\left[\sqrt{4\lambda_i + 1} - 1\right], & i = 1, \ldots, 10 \\ \lambda_i / M_i, & i = 11, \ldots, 100 \end{cases} \tag{111}$$

The usual estimate of η_i is then obtained from Eq. (111) with λ_i replaced by the usual estimate X_i. Similarly, the final estimate of η_i is also obtained from Eq. (111) with λ_i replaced by the final estimate of λ_i. The function F defined in Eq. (111) must be specified in the user-supplied subroutine EQN. For example,

```
      DO 100 I=N1,N2
      IF(I .GT. 10) GO TO 50
      W2(I)=(SQRT(4. * NX(I) +1.)-1.)/(BUX(I,1)*2.)
      W3(I)=(SQRT(4. * XEST(I)+1.)-1.)/(BUX(I,1)*2.)
      GO TO 100
   50 W2(I)=NX(I)/BUX(I,1)
      W3(I)=XEST(I)/BUX(I,1)
  100 CONTINUE
```

6) Limitations of the Program and How to Make Necessary Changes

This program was designed to handle problems where the number of parameters, NP, does not exceed 500. This number can be easily increased to any desired number. However, one must change the dimensions of variables that appear on the DIMENSION STATEMENTS of the main program and subroutines LSE and EQN, i.e., replace each 500 by the desired number.

The number of groups is limited to 20. This number is large enough for most practical problems. However, one can increase this number to any desired number not exceeding NP by properly changing the READ format of the second input card (FORMAT 3).

The total number of supplemental and auxiliary variables is limited to 5. This number should be large enough for practical application. The user can change this number to any desired number, say N, by the following steps:

a) Change the READ format of the second portion of the data input (FORMAT 60).

b) Change dimensions of variables that appear on the DIMENSION STATEMENTS of the main program and subroutines LSE, MTXINV and EQN, i.e., replace 5 and 6 by N and N + 1, respectively.

7) Examples

We shall use the data presented in the first section of Part II to demonstrate the use of various program options for computing the initial estimates of parameters.

Number of accidents and vehicle-miles are available for 24 locations. These locations are coded as 1 through 24. Accident statistics and vehicle-miles of these locations are presented in Table 18. Denote X_i and t_i to be the number of accidents and vehicle-miles of the i-th location, respectively. It is reasonable to assume that X_i is Poisson distributed with parameter λ_i. We are interested in estimating the accident rate (number of accidents per vehicle-mile) of Locations 13 through 24. If we denote η_i to be the accident rate of the i-th location, then

$$\eta_i = F(\lambda_i) = \lambda_i/t_i \tag{112}$$

The question now is what group-method combination should be used to compute initial estimates of parameters. To answer this question, we

Table 18
Accident Statistics And Vehicle-Miles
For Each of 24 Locations

Location	No of Accidents	Vehicle-miles
1	170	3346384.35
2	177	4326633.85
3	177	4741402.00
4	193	4972885.15
5	739	19958427.50
6	895	26901854.00
7	213	3518919.25
8	286	5337649.95
9	237	4397293.20
10	130	3012894.20
11	428	9969259.00
12	634	16111479.60
13	51	1208984.00
14	92	1982933.10
15	103	2363350.60
16	80	1793523.20
17	266	6653599.00
18	354	9773442.80
19	133	1753781.70
20	207	3428029.80
21	107	2373551.00
22	62	976866.20
23	180	3956828.00
24	258	6949258.40

graphically examine the relationship between the supplemental variable and the transformed variable defined in Eq. (87). We see from Figure 5 that the following linear equations hold approximately,

$$Z_i = b_0 + b_1 \sqrt{t_i} \tag{113}$$

$$Z_i = b \sqrt{t_i} \tag{114}$$

or

$$Z_i = b_1 \sqrt{t_i} + b_2 t_i \tag{115}$$

Thus, Methods 2, 5, and 6 can be used to compute the initial parameter estimates. The supplemental variable used in Eqs. (113) and (114) is the square root of t_i. The auxiliary variable for converting λ_i to the accident rate η_i is t_i. If Eq. (115) is used, the supplemental variables become $\sqrt{t_i}$ and t_i. In this case, the second supplemental variable also serves as the auxiliary variable. We present in Table 19 the second portion of the data input for Examples 1 through 5. Four numbers are shown on each card. These are location numbers, number of accidents, square root of vehicle-miles, and vehicle-miles.

Example 1: Using Method 2 with One Group - Since the linear relationship $Z_i = b\sqrt{t_i}$ holds approximately for every i, the weighted average

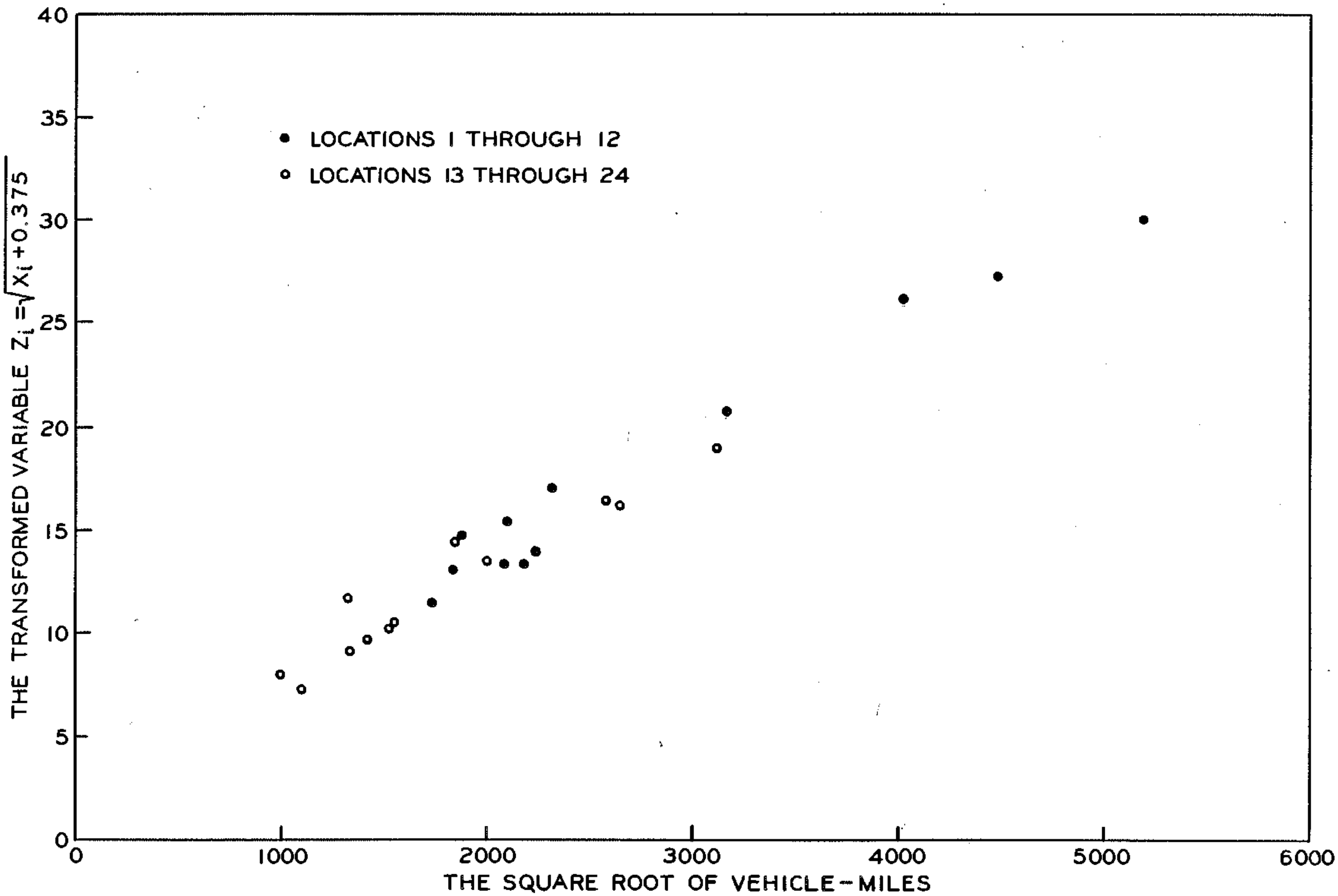

Figure 5. The relationship between the transformed variable Z_i and the square root of vehicle-miles.

Table 19
The Second Portion of The Data Input
of Examples 1 Through 5

Card No.	Column Number 12345678901234567890123456789012345678901234567890			
7	13	51	1099.54	1208984.00
8	14	92	1408.17	1982933.10
9	15	103	1537.32	2363350.60
10	16	80	1339.22	1793523.20
11	17	266	2579.46	6653599.00
12	18	354	3126.25	9773442.80
13	19	133	1324.30	1753781.70
14	20	207	1851.15	3428029.80
15	21	107	1540.63	2373551.00
16	22	62	988.37	976866.20
17	23	180	1989.18	3956828.00
18	24	258	2636.14	6949258.40

(weighted by the square root of t_i) can be used to estimate b. That is, Method 2 is used to compute initial estimates of parameters. The first six cards of the data input for this example are presented below.

CARD NO.	COLUMN NUMBER 12345678901234567890123456789012345678901234567890	
1	12	1
2	12	
3	2	
4	1	
5	1	
6	METHOD 2, 1-GROUP	

The first card sets NP = 12 and NGROUP = 1. Consequently, NMG(1) = 12 which is set in the second card. The third card sets METHOD(1) = 2. The fourth and fifth cards, respectively, set NAUX(1) = 1 and NCOV(1) = 1 Thus, for the only group, AUX(i, 1) = $\sqrt{t_i}$ and BUX(i, 1) = t_i. Since NCOV(1) = 1, the final estimate of λ_i will be converted to η_i defined in Eq. (112). To do this, we supply the following statements to the user-supplied subroutine EQN.

```
      DO 100 I=N1,N2
      W2(I)=NX(I)/BUX(I,1)
      W3(I)=XEST(I)/BUX(I,1)
100   CONTINUE
```

The run results are presented in Table 20.

Example 2: Using Method 2 with Three Groups - Based on the physical properties of these locations, the parameter inhomogeneity can be reduced by separating these locations into the following three groups:

Group 1 - Locations 13 - 18
Group 2 - Locations 19 - 21
Group 3 - Locations 22 - 24

Table 20
Estimated Results of Example 1

```
################################################################################
METHOD 2, 1-GROUP
################################################################################
DATA :
```

ID	X	GROUP NO	SUPPLEMENTAL INFORMATION	
13	51	1	1099.54	1208984.00
14	92	1	1408.17	1982933.10
15	103	1	1537.32	2363350.60
16	80	1	1339.22	1793523.20
17	266	1	2579.46	6653599.00
18	354	1	3126.25	9773442.80
19	133	1	1324.30	1753781.70
20	207	1	1851.49	3428029.80
21	107	1	1540.63	2373551.00
22	62	1	988.37	976866.20
23	180	1	1989.18	3956828.00
24	258	1	2636.14	6949258.40

ESTIMATED RESULTS :

ID	POISSON RATE ACTUAL	POISSON RATE ESTIMATED	CONVERTED RATE ACTUAL	CONVERTED RATE ESTIMATED
13	51	51.65	0.000042184	0.000042720
14	92	92.06	0.000046396	0.000046427
15	103	103.74	0.000043582	0.000043894
16	80	80.41	0.000044605	0.000044835
17	266	269.81	0.000039978	0.000040551
18	354	363.13	0.000036221	0.000037155
19	133	127.06	0.000075836	0.000072450
20	207	201.36	0.000060385	0.000058738
21	107	107.37	0.000045080	0.000045238
22	62	60.26	0.000063468	0.000061683
23	180	180.29	0.000045491	0.000045564
24	258	263.89	0.000037126	0.000037974

```
SHRINKING FACTOR                     =  0.8954
%-IMPROVEMENT OVER USUAL ESTIMATE =  8.2065 %
```

We then use Method 2 to compute initial estimates of parameters in each group. For this case, the first six cards of the data input take the following form.

CARD NO.	COLUMN NUMBER 12345678901234567890123456789012345678901234567890
1	12 3
2	6 3 3
3	2 2 2
4	1 1 1
5	1 1 1
6	METHOD 2,3-GROUPS(13-18,19-21 & 22-24)

The first card sets NGROUP = 3. Consequently, three numbers appear on each of Cards 2 through 5. The second card sets NMG(1) = 6, NMG(2) = 3 and NMG(3) = 3. The user-supplied subroutine EQN is the same as the one used in Example 1. The run results are presented in Table 21.

Table 21
Estimated Results of Example 2

##
METHOD 2, 3-GROUPS(13-18,19-21 & 22-24)
##

DATA :

ID	X	GROUP NO	SUPPLEMENTAL INFORMATION	
13	51	1	1099.54	1208984.00
14	92	1	1408.17	1982933.10
15	103	1	1537.32	2363350.60
16	80	1	1339.22	1793523.20
17	266	1	2579.46	6653599.00
18	354	1	3126.25	9773442.80
19	133	2	1324.30	1753781.70
20	207	2	1851.49	3428029.80
21	107	2	1540.63	2373551.00
22	62	3	988.37	976866.20
23	180	3	1989.18	3956828.00
24	258	3	2636.14	6949258.40

ESTIMATED RESULTS :

ID	POISSON RATE ACTUAL	POISSON RATE ESTIMATED	CONVERTED RATE ACTUAL	CONVERTED RATE ESTIMATED
13	51	50.95	0.000042184	0.000042142
14	92	90.32	0.000046396	0.000045547
15	103	102.16	0.000043582	0.000043227
16	80	79.07	0.000044605	0.000044086
17	266	267.16	0.000039978	0.000040153
18	354	361.85	0.000036221	0.000037024
19	133	127.94	0.000075836	0.000072953
20	207	206.28	0.000060385	0.000060174
21	107	112.67	0.000045080	0.000047470
22	62	58.74	0.000063468	0.000060129
23	180	179.44	0.000045491	0.000045348
24	258	266.30	0.000037126	0.000038320

SHRINKING FACTOR = 0.8281
%-IMPROVEMENT OVER USUAL ESTIMATE = 11.0296 %

Example 3: Using Method 5 with Two Groups - The common scale parameter b in Example 1 can also be estimated by the least squares method. That is, Method 5 can be used to compute initial estimates of parameters. In this case, the first six cards of the data input takes the following form.

CARD NO.	COLUMN NUMBER 1234567890123456789012345678901234567890
1	12 1
2	12
3	5
4	1
5	1
6	METHOD 5, 1-GROUP

The user-supplied subroutine EQN is the same as the one used in Example 1. The run results are presented in Table 22.

Table 22
Estimated Results of Example 3

```
##################################################################
METHOD 5, 1-GROUP
##################################################################

DATA :
```

ID	X	GROUP NO	SUPPLEMENTAL INFORMATION	
13	51	1	1099.54	1208984.00
14	92	1	1408.17	1982933.10
15	103	1	1537.32	2363350.60
16	80	1	1339.22	1793523.20
17	266	1	2579.46	6653599.00
18	354	1	3126.25	9773442.80
19	133	1	1324.30	1753781.70
20	207	1	1851.49	3428029.80
21	107	1	1540.63	2373551.00
22	62	1	988.37	976866.20
23	180	1	1989.18	3956828.00
24	258	1	2636.14	6949258.40

ESTIMATED RESULTS :

ID	POISSON RATE ACTUAL	POISSON RATE ESTIMATED	CONVERTED RATE ACTUAL	CONVERTED RATE ESTIMATED
13	51	51.38	0.000042184	0.000042496
14	92	91.55	0.000046396	0.000046168
15	103	103.19	0.000043582	0.000043661
16	80	79.97	0.000044605	0.000044591
17	266	268.47	0.000039978	0.000040349
18	354	361.47	0.000036221	0.000036984
19	133	126.14	0.000075836	0.000071923
20	207	200.04	0.000060385	0.000058356
21	107	106.79	0.000045080	0.000044990
22	62	59.85	0.000063468	0.000061269
23	180	179.30	0.000045491	0.000045315
24	258	262.65	0.000037126	0.000037796

```
SHRINKING FACTOR                    =  0.8892
%-IMPROVEMENT OVER USUAL ESTIMATE  =  8.7213 %
```

Example 4: Using Method 6 with Two Groups - Figure 5 shows that Eq. (113) is better than Eq. (114) in expressing the linear relationship between Z_i and $\sqrt{t_i}$. That is, the noncentrality parameter ζ defined in Eq. (94) is smaller by using Eq. (113). However, the extra parameter used in Eq. (113) decreases the number of degrees of freedom by one. This is the negative side of using Eq. (113). We shall see later that the percentage improvement has been substantially increased by using Eq. (113). That is, the reduction on ζ by using Eq. (113) is more than enough to compensate for the loss of one degree of freedom. Therefore, Method 6 is better than Method 5 in this problem. The first six cards of the data input are those in

Example 3 with the number '5' in the third card replaced by the number '6'. The user-supplied subroutine is the same as the one used in Example 1. The run results are presented in Table 23.

Table 23
Estimated Results of Example 4

```
##################################################################
METHOD 6, 1-GROUP
##################################################################

DATA :

          GROUP
 ID    X   NO     SUPPLEMENTAL INFORMATION
-----------------------------------------------------------------
 13    51   1      1099.54   1208984.00
 14    92   1      1408.17   1982933.10
 15   103   1      1537.32   2363350.60
 16    80   1      1339.22   1793523.20
 17   266   1      2579.46   6653599.00
 18   354   1      3126.25   9773442.80
 19   133   1      1324.30   1753781.70
 20   207   1      1851.49   3428029.80
 21   107   1      1540.63   2373551.00
 22    62   1       988.37    976866.20
 23   180   1      1989.18   3956828.00
 24   258   1      2636.14   6949258.40

ESTIMATED RESULTS :

       POISSON RATE                 CONVERTED RATE
 ID  ACTUAL ESTIMATED          ACTUAL                ESTIMATED
-----------------------------------------------------------------
 13     51       54.61         0.000042184           0.000045167
 14     92       93.85         0.000046396           0.000047331
 15    103      105.43         0.000043582           0.000044612
 16     80       82.68         0.000044605           0.000046096
 17    266      265.74         0.000039978           0.000039939
 18    354      356.42         0.000036221           0.000036468
 19    133      125.25         0.000075836           0.000071418
 20    207      196.70         0.000060385           0.000057381
 21    107      108.80         0.000045080           0.000045840
 22     62       62.08         0.000063468           0.000063555
 23    180      178.87         0.000045491           0.000045205
 24    258      260.96         0.000037126           0.000037552

 SHRINKING FACTOR                      =   0.8175
%-IMPROVEMENT OVER USUAL ESTIMATE = 13.2790 %
```

<u>Example 5: Using Method 5 with Two Groups</u> - Figure 5 also indicates that the linear relationship can be well expressed by Eq. (115). That is, Method 5 can be used to compute initial estimates of parameters. For demonstration purposes, we shall use Method 2 with two groups: Locations 13 through 18 and 19 through 24. As mentioned before, the supplemental variables for using Eq. (115) are $\sqrt{t_i}$ and t_i. Since t_i also serves as the auxiliary variable, by definition, NCOV(1) and NCOV(2) are set to be 0. The first six cards of the data input take the following form.

CARD NO	COLUMN NUMBER 12345678901234567890123456789012345678901234567890
1	12 2
2	6 6
3	5 5
4	2 2
5	0 0
6	METHOD 5, 2-GROUPS(13-18 & 19-24)

We note that BUX(i, 1) = t_i in Examples 1 through 4. However, in this case, BUX(i, 2) = t_i. Thus, in order to convert λ_i to η_i defined in Eq. (112), the user-supplied subroutine EQN takes the following form.

```
      DO 100 I=N1,N2
      W2(I)=NX(I)/BUX(I,2)
      W3(I)=XEST(I)/BUX(I,2)
  100 CONTINUE
```

The run results are presented in Table 24.

Table 24
Estimated Results of Example 5

```
##################################################################
METHOD 5, 2-GROUPS(13-18 & 19-24)
##################################################################
DATA :
```

ID	X	GROUP NO	SUPPLEMENTAL INFORMATION	
13	51	1	1099.54	1208984.00
14	92	1	1408.17	1982933.10
15	103	1	1537.32	2363350.60
16	80	1	1339.22	1793523.20
17	266	1	2579.46	6653599.00
18	354	1	3126.25	9773442.80
19	133	2	1324.30	1753781.70
20	207	2	1851.49	3428029.80
21	107	2	1540.63	2373551.00
22	62	2	988.37	976866.20
23	180	2	1989.18	3956828.00
24	258	2	2636.14	6949258.40

ESTIMATED RESULTS :

ID	POISSON RATE ACTUAL	POISSON RATE ESTIMATED	CONVERTED RATE ACTUAL	CONVERTED RATE ESTIMATED
13	51	52.60	0.000042184	0.000043506
14	92	91.00	0.000046396	0.000045894
15	103	103.44	0.000043582	0.000043769
16	80	80.32	0.000044605	0.000044781
17	266	264.14	0.000039978	0.000039699
18	354	355.42	0.000036221	0.000036366
19	133	125.49	0.000075836	0.000071556
20	207	197.13	0.000060385	0.000057505
21	107	116.25	0.000045080	0.000048976
22	62	64.06	0.000063468	0.000065581
23	180	184.32	0.000045491	0.000046582
24	258	258.33	0.000037126	0.000037174

```
 SHRINKING FACTOR                     =  0.6862
%-IMPROVEMENT OVER USUAL ESTIMATE = 19.0130 %
```

Example 6: Using Methods 2 and 5 with Two Groups - For demonstration purposes, we separate these locations into two groups:

Group 1 - Locations 13, 16, 19, 22
Group 2 - The remaining locations.

Method 5 with Eq. (115) is then used to compute initial estimates of parameters in the first group as done in Example 5. Thus, NMG(1) = 4, NAUX(1) = 2, and NCOV(1) = 0. For the second group, we use Method 2 to compute initial estimates of parameters as done in Example 1. The data input are presented in Table 25. We see from this table that the first four cards (Cards 7 through 10) of the second portion are the data for estimating parameters in the first group. The order of these four cards has no effect

Table 25
Data Input of Example 6

Card No.	Column Number 12345678901234567890123456789012345678901234567890			
1	12	2		
2	4	8		
3	5	2		
4	2	1		
5	0	1		
6	METHOD 5 ON (13,16,19 & 22), METHOD 2 ON REMAINING LOCATIONS			
7	13	51	1099.54	1208984.00
8	16	80	1339.22	1793523.20
9	19	133	1324.30	1753781.70
10	22	62	98B.37	976866.20
11	14	92	1408.17	1982933.10
12	15	103	1537.32	2363350.60
13	17	266	2579.46	6653599.00
14	18	354	3126.25	9773442.80
15	20	207	1851.49	3428029.80
16	21	107	1540.63	2373551.00
17	23	180	1989.18	3956828.00
18	24	25B	2636.14	6949258.40

on the estimation procedures. Cards 11 through 18 are the data for parameters in the second group. Since BUX(i, 2) = t_i for i = 1, 2, 3, and 4, and BUX(i, 1) = t_i for i $\geq$ 5, the user-supplied subroutine EQN takes the following form.

```
      DO 100 I=N1,N2
      K=1
      IF(I .LE. 4)K=2
      W2(I)=NX(I)/BUX(I,K)
      W3(I)=XEST(I)/BUX(I,K)
100   CONTINUE
```

The run results are presented in Table 26.

Table 26
Estimated Results of Example 6

##
METHOD 5 ON (13,16,19 & 22), METHOD 2 ON REMAINING LOCATIONS
##

DATA :

ID	X	GROUP NO	SUPPLEMENTAL INFORMATION	
13	51	1	1099.54	1208984.00
16	80	1	1339.22	1793523.20
19	133	1	1324.30	1753781.70
22	62	1	988.37	976866.20
14	92	2	1408.17	1982933.10
15	103	2	1537.32	2363350.60
17	266	2	2579.46	6653599.00
18	354	2	3126.25	9773442.80
20	207	2	1851.49	3428029.80
21	107	2	1540.63	2373551.00
23	180	2	1989.18	3956828.00
24	258	2	2636.14	6949258.40

ESTIMATED RESULTS :

ID	POISSON RATE ACTUAL	POISSON RATE ESTIMATED	CONVERTED RATE ACTUAL	CONVERTED RATE ESTIMATED
13	51	52.88	0.000042184	0.000043739
16	80	82.73	0.000044605	0.000046126
19	133	129.23	0.000075836	0.000073686
22	62	61.06	0.000063468	0.000062506
14	92	91.39	0.000046396	0.000046091
15	103	103.04	0.000043582	0.000043598
17	266	268.18	0.000039978	0.000040305
18	354	361.22	0.000036221	0.000036960
20	207	199.52	0.000060385	0.000058202
21	107	106.62	0.000045080	0.000044920
23	180	179.02	0.000045491	0.000045243
24	258	262.45	0.000037126	0.000037767

SHRINKING FACTOR = 0.8846
%-IMPROVEMENT OVER USUAL ESTIMATE = 7.1824 %

Example 7: Using Method 5 with Parallel Data as the Supplemental Variable - The only difference between the i-th and (i-12)-th locations for every i = 13, . . . , 24, is the presence of an intersecting roadway. We now define

$$V_i = \frac{X_{i-12}}{t_{i-12}} t_i, \quad i = 13, \ldots, 24 \tag{116}$$

V_i in Eq. (116) can be interpreted as the number of accidents occurring on the (i-12)-th location when its vehicle-mileage is t_i. We observe graphically that the following linear relationship holds approximately:

$$Z_i = b \sqrt{V_i + 0.375} \tag{117}$$

Table 27
Data Input of Example 7

Card No.	Column Number 12345678901234567890123456789012345678901234567890
1	12 1
2	12
3	5
4	1
5	1
6	METHOD 5 WITH PARALLEL DATA AS SUPPLEMENTAL VARIABLE
7	13 51 7.86 1208984.00
8	14 92 9.03 1982933.10
9	15 103 9.41 2363350.60
10	16 80 8.37 1793523.20
11	17 266 15.71 6653599.00
12	18 354 18.04 9773442.80
13	19 133 10.32 1753781.70
14	20 207 13.61 3428029.80
15	21 107 11.33 2373551.00
16	22 62 6.52 976866.20
17	23 180 13.05 3956828.00
18	24 258 16.55 6949258.40

Table 28
Estimated Results of Example 7

##
METHOD 5 WITH PARALLEL DATA AS SUPPLEMENTAL VARIABLE
##
DATA :

ID	X	GROUP NO	SUPPLEMENTAL INFORMATION	
13	51	1	7.86	1208984.00
14	92	1	9.03	1982933.10
15	103	1	9.41	2363350.60
16	80	1	8.37	1793523.20
17	266	1	15.71	6653599.00
18	354	1	18.04	9773442.80
19	133	1	10.32	1753781.70
20	207	1	13.61	3428029.80
21	107	1	11.33	2373551.00
22	62	1	6.52	976866.20
23	180	1	13.05	3956828.00
24	258	1	16.55	6949258.40

ESTIMATED RESULTS :

ID	POISSON RATE ACTUAL	POISSON RATE ESTIMATED	CONVERTED RATE ACTUAL	CONVERTED RATE ESTIMATED
13	51	56.03	0.000042184	0.000046348
14	92	90.20	0.000046396	0.000045488
15	103	99.92	0.000043582	0.000042281
16	80	78.13	0.000044605	0.000043561
17	266	264.48	0.000039978	0.000039751
18	354	350.90	0.000036221	0.000035903
19	133	125.96	0.000075836	0.000071823
20	207	203.22	0.000060385	0.000059283
21	107	117.00	0.000045080	0.000049294
22	62	55.90	0.000063468	0.000057221
23	180	180.41	0.000045491	0.000045595
24	258	269.15	0.000037126	0.000038731

SHRINKING FACTOR = 0.6498
%-IMPROVEMENT OVER USUAL ESTIMATE = 30.0854 %

Thus, Method 5 with $\sqrt{V_i + 0.375}$ as the supplemental variable can be used to compute initial estimates of parameters. The data input for this case are presented in Table 27. We see from this table that four numbers appear on Cards 7 through 18. These are location number, number of accidents, the square root of $(V_i + 0.375)$ and vehicle-miles. The user-supplied subroutine EQN is the same as the one used in Example 1. The run results are presented in Table 28.

8) Program Listing

```
      PROGRAM POISSON(TAPE1=INPUT,TAPE2=OUTPUT)
      DIMENSION NX(500),U(500),XT(500),NAUX(500),
     1UT(500),XEST(500),NMG(500),IG(500),ID(500),NCOV(500),
     2AUX(500,5),B(6),BUX(500,5),METHOD(500),DESCPT(72)
C
C  THIS PROGRAM SIMULTANEOUSLY ESTIMATES NP POISSON PARAMETERS
C  USING STEIN-LIKE ESTIMATION PROCEDURES. THE ESTIMATED RESULTS
C  ARE BETTER THAN THE USUAL ESTIMATES IN TERMS OF THE SQUARED
C  ERROR LOSS .
C
  1   READ(1,3,END=9999)NP,NGROUP
      READ(1,3) (NMG(I),I=1,NGROUP)
      READ(1,3) (METHOD(I),I=1,NGROUP)
      READ(1,3) (NAUX(I),I=1,NGROUP)
      READ(1,3) (NCOV(I),I=1,NGROUP)
      READ(1,10)(DESCPT(I),I=1,70)
      WRITE(2,15)(DESCPT(I),I=1,70)
      WRITE(2,30)
      SSR=0.
      NDF=0
      K=0
      ICHO=2
      DO 1000 I=1,NGROUP
      NCOUNT=K
      NPT=NMG(I)
      WB1=0.
      WB2=0.
      IF(METHOD(I) .LE. 4)NAUX(I)=1
      DO 500 J=1,NPT
      K=NCOUNT+J
      IG(K)=I
      NA=NAUX(I)
      NC=NCOV(I)
      IF(NC .GT. 0)GO TO 150
      READ(1,60)ID(K),NX(K),(AUX(J,M),M=1,NA)
      WRITE(2,300)ID(K),NX(K),IG(K),(AUX(J,M),M=1,NA)
      IF(NC .LT. 0)GO TO 200
      DO 100 M=1,NA
 100  BUX(K,M)=AUX(J,M)
      GO TO 200
 150  READ(1,60)ID(K),NX(K),(AUX(J,M),M=1,NA),
     1(BUX(K,M),M=1,NC)
      WRITE(2,300)ID(K),NX(K),IG(K),(AUX(J,M),M=1,NA),
     1(BUX(K,M),M=1,NC)
 200  XT(K)=SQRT(0.375+NX(K))
      U(J)=XT(K)
      IF(METHOD(I) .GE. 5)GO TO 500
      IF(METHOD(I) .NE. 1)GO TO 350
      A1=AUX(J,1)
      CALL TRANSF(ICHO,A2,A1)
      UT(K)=A2
      SSR=SSR+(XT(K)-UT(K))**2
      GO TO 500
 350  CONTINUE
```

```
      IF(METHOD(I) .EQ. 4)GO TO 450
      IF(METHOD(I) .EQ. 3)GO TO 400
      WB1=WB1+XT(K)
      WB2=WB2+AUX(J,1)
      GO TO 500
400   WB1=WB1+XT(K)
      GO TO 500
450   WB1=WB1+XT(K)*AUX(J,1)
500   CONTINUE
      IF(METHOD(I) .EQ. 1)GO TO 1000
      IF(METHOD(I) .GE. 5)GO TO 900
      IF(METHOD(I) .EQ. 2)WB1=WB1/WB2
      DO 800 J=1,NPT
      K=NCOUNT+J
      UT(K)=WB1
      IF(METHOD(I) .LE. 3)UT(K)=UT(K)*AUX(J,1)
800   SSR=SSR+(XT(K)-UT(K))**2
      NDF=NDF+1
      GO TO 1000
900   CONTINUE
      NFORCE=METHOD(I)-5
      NIND=NAUX(I)
      NDF=NDF+NIND+NFORCE
      CALL LSE(NFORCE,NPT,NIND,U,AUX,B,XEST,SERQR)
      DO 950 J=1,NPT
      K=NCOUNT+J
      UT(K)=XEST(J)
950   SSR=SSR+(XT(K)-UT(K))**2
1000  CONTINUE
      ICHO=1
      NTRACE=NP-NDF
      IF(NTRACE .GT. 2)GO TO 2000
      WRITE(2,1500)NTRACE
      GO TO 9999
2000  CONTINUE
      SSR=SSR*4
      FACTOR=1.-(NTRACE-2)/SSR
      IF(FACTOR .LT. O.)FACTOR=O.
      DO 2100 J=1,NP
      EST=UT(J)+FACTOR*(XT(J)-UT(J))
      CALL TRANSF(ICHO,EST,WB2)
      XEST(J)=WB2
2100  CONTINUE
      WRITE(2,2200)
      N2=O
      DO 4000 I=1,NGROUP
      N1=N2+1
      N2=N1+NMG(I)-1
      IF(NCOV(I) .LT. O)GO TO 3400
      CALL EQN(NP,ID,IG,NX,XEST,BUX,N1,N2,XT,UT)
      DO 3300 J=N1,N2
      WRITE(2,3200)ID(J),NX(J),XEST(J),XT(J),UT(J)
3300  CONTINUE
      GO TO 4000
3400  DO 3500 J=N1,N2
      WRITE(2,3200) ID(J),NX(J),XEST(J)
3500  CONTINUE
4000  CONTINUE
      SSR=SSR-NTRACE
      IF(SSR .LE. O.)SSR=O.
      CALL EXPECT(NTRACE,SSR,EYY)
      PIMPRO=EYY*(NTRACE-2)**2/NP*100
      WRITE(2,6000)FACTOR,PIMPRO
      GO TO 1
9999  CONTINUE
      LOCK 2
      STOP
C
C  INPUT AND OUTPUT FORMATS FOR THIS PROGRAM
 3    FORMAT(2OI3)
 10   FORMAT(1X,70A1)
```

```
15    FORMAT(/,1X,45H#############################################,
     125H#########################,/,1X,70A1,/,1X,13H#############,
     257H#########################################################)
30    FORMAT(/,1X,4HDATA,//,10X,5HGROUP,/,2X,
     140HID     X    NO     SUPPLEMENTAL INFORMATION,/,1X,
     248H------------------------------------------------,
     320H--------------------)
60    FORMAT(I3,I5,5F12.2)
300   FORMAT(1X,I3,I6,I4,4F12 2,5(/,15X,4F12.2))
1500  FORMAT(//,1X,29H--- ERROR - ERROR - ERROR ---,/,2X,
     19HTRACE(P)=,I4,/,2X,28HTHE TRACE OF P SHOULD BE AT ,
     27HLEAST 2,/,2X,34HCHECK TO SEE WHETHER YOU HAVE USED,
     3/,2X,44HTOO MANY GROUPS OR/AND TOO MANY VARIABLES IN,
     4/,2X,20HTHE REGRESSION LINES,/)
2200  FORMAT(/,1X,17HESTIMATED RESULTS ,///,8X,12HPOISSON RATE,
     115X,14HCONVERTED RATE,/,2X,20HID   ACTUAL ESTIMATED,7X,6HACTUAL,
     213X,9HESTIMATED,/,1X,35H-----------------------------------,
     330H------------------------------)
3200  FORMAT(1X,I3,I7,F11.2,2F20.9)
6000  FORMAT(/,1X,35H SHRINKING FACTOR                  =,F8.4,/,
     11X,35H%-IMPROVEMENT OVER USUAL ESTIMATE =,F8.4,2H %,//)
      END

      SUBROUTINE LSE(MODEL,NPT,NIND,YDATA,XDATA,B,YEST,SEROR)
      DIMENSION YDATA(500),XDATA(500,5),B(6),X(500,6),
     1YEST(500),XX(6,6),XXIXT(6,500)
C
C     SUBROUTINE FOR COMPUTING THE LEAST SQUARES ESTIMATE
C     OF B IN THE LINEAR MODEL,
C
C          YDATA = XDATA * B         , IF MODEL = 0
C     AND
C          YDATA = (I,XDATA) * B   , IF MODEL = 1
C
C     WHERE YDATA IS THE THE (NPT BY 1) VECTOR OF DEPENDENT
C     OBSERVATIONS, XDATA IS THE (NPT BY NIND) MATRIX OF
C     INDEPENDENT OBSERVATIONS, I IS THE COLUMN VECTOR WITH
C     EVERY ELEMENT EQUAL TO ONE, * STANDS FOR THE PRODUCT
C     OF TWO MATRICES, NIND IS THE NUMBER OF INDEPENDENT
C     VARIABLES AND B IS THE VECTOR OF PARAMETERS TO BE
C     ESTIMATED.
C
      IF(MODEL  EQ. 0)GO TO 100
      DO 10 I=1,NPT
10    X(I,1)=1.
      MM=NIND+1
      DO 20 I=2,MM
      II=I-1
      DO 30 J=1,NPT
30    X(J,I)=XDATA(J,II)
20    CONTINUE
      GO TO 200
100   MM=NIND
      DO 120 I=1,MM
      DO 130 J=1,NPT
130   X(J,I)=XDATA(J,I)
120   CONTINUE
200   CONTINUE
      DO 230 I=1,MM
      DO 240 J=1,MM
      XX(I,J)=0.
      DO 250 K=1,NPT
250   XX(I,J)=XX(I,J)+X(K,I)*X(K,J)
240   CONTINUE
230   CONTINUE
      CALL MTXINV(MM,XX,XX)
      DO 300 I=1,MM
      DO 310 J=1,NPT
      XXIXT(I,J)=0.
      DO 320 K=1,MM
320   XXIXT(I,J)=XXIXT(I,J)+XX(I,K)*X(J,K)
310   CONTINUE
```

```
  300 CONTINUE
      DO 400 I=1,MM
      B(I)=0.
      DO 410 J=1,NPT
  410 B(I)=B(I)+XXIXT(I,J)*YDATA(J)
  400 CONTINUE
      SEROR=0.
      DO 500 I=1,NPT
      YEST(I)=0.
      DO 510 J=1,MM
  510 YEST(I)=YEST(I)+X(I,J)*B(J)
      SEROR=SEROR+(YDATA(I)-YEST(I))**2
  500 CONTINUE
      SEROR=SQRT(SEROR/(NPT-MM))
      RETURN
      END

      SUBROUTINE MTXINV(NSIZE,W,WINV)
      DIMENSION ARRAY(6,6),WINV(6,6),W1(6,2),W(6,6)
C
C     SUBROUTINE FOR FINDING THE INVERSE OF AN (NSIZE BY NSIZE)
C     SQUARE MATRIX W BY USING THE PARTITION METHOD. WINV IS THE
C     INVERSE MATRIX OF W.
C
      DO 5 I=1,NSIZE
      DO 5 J=1,NSIZE
    5 ARRAY(I,J)=W(I,J)
      IF(NSIZE .GT. 1) GO TO 10
      WINV(1,1)=1./ARRAY(1,1)
      RETURN
   10 CONTINUE
      MSIZE=NSIZE-1
      DO 15  II=1,MSIZE
      J=II+1
      DO 16  KK=J,NSIZE
      DO  17 M=1,NSIZE
      W1(M,1)=W(M,II)
   17 W1(M,2)=W(M,KK)
      DET=W1(1,1)*W1(2,2)-W1(1,2)*W1(2,1)
      IF(DET .EQ. 0.) GO TO 16
      IF(II .EQ. 1 .AND. KK .EQ. 2) GO TO 19
      DO  18 K=1,NSIZE
      ARRAY(K,1)=W(K,II)
      ARRAY(K,2)=W(K,KK)
      ARRAY(K,II)=W(K,1)
   18 ARRAY(K,KK)=W(K,2)
      GO TO 19
   16 CONTINUE
   15 CONTINUE
   19 CONTINUE
      WINV(1,1)=ARRAY(2,2)/DET
      WINV(2,2)=ARRAY(1,1)/DET
      WINV(1,2)=-ARRAY(1,2)/DET
      WINV(2,1)=-ARRAY(2,1)/DET
      IF(NSIZE .EQ. 2) GO TO 100
      DO 20 I=3,NSIZE
      K=I-1
      DO 21 J=1,K
      W1(J,1)=0.
      W1(J,2)=0.
      DO 22 M=1,K
      W1(J,1)=W1(J,1)+WINV(J,M)*ARRAY(M,I)
   22 W1(J,2)=W1(J,2)+ARRAY(I,M)*WINV(M,J)
   21 CONTINUE
      ELTA=ARRAY(I,I)
      DO 23 J=1,K
   23 ELTA=ELTA-ARRAY(I,J)*W1(J,1)
      WINV(I,I)=1./ELTA
      DO 24 J=1,K
      WINV(J,I)=-W1(J,1)/ELTA
      WINV(I,J)=-W1(J,2)/ELTA
      DO 24 M=1,K
   24 WINV(J,M)=WINV(J,M)+W1(J,1)*W1(M,2)/ELTA
```

```
 20   CONTINUE
 100  CONTINUE
      IF(II .EQ. 1 .AND. KK .EQ. 2) GO TO 888
      DO 401 J=1,NSIZE
      W1(J,1)=WINV(1,J)
 401  W1(J,2)=WINV(2,J)
      DO 402 J=1,NSIZE
      WINV(1,J)=WINV(II,J)
 402  WINV(2,J)=WINV(KK,J)
      DO 403 J=1,NSIZE
      WINV(II,J)=W1(J,1)
 403  WINV(KK,J)=W1(J,2)
 888  CONTINUE
      RETURN
      END

      SUBROUTINE EXPECT(NP,THETA,EYY)
      DOUBLE PRECISION P1,A,EY,P2
C
C     THIS SUBROUTINE COMPUTES THE EXPECTATION OF 1/Y, WHERE Y IS A
C     NONCENTRAL CHI-SQUARE WITH NP DEGREES OF FREEDOM AND NONCENTRALITY
C     PARAMETER THETA. THIS EXPECTATION IS THE SAME AS THE EXPECTATION
C     OF 1/(NP-2+2W), WHERE W IS A POISSON WITH PARAMETER (THETA/2).
C
      ERROR=0.00001
      ERR=ERROR/NP
      A=THETA/2
      P1=DEXP(-A)
      EY=P1/(NP-2)
      P2=P1
      K=0
 10   K=K+1
      P1=P1*A/K
      P2=P2+P1
      EY=EY+P1/(NP-2+K*2)
      CHECK=1.-P2
      IF(CHECK.GE.ERR) GO TO 10
      EYY=EY
      RETURN
      END

      SUBROUTINE TRANSF(ICHO,Y,X)
      DIMENSION THETA(166),EY(166)
C
C     X IS THE POISSON RATE, Y IS THE TRANSFORMED RATE
C     ICHO=1 FOR CONVERTING Y TO X
C     ICHO=2 FOR CONVERTING X TO Y
C
      DATA  (THETA(I),I=1,166)/ 0.0,
     10.1,  0.2,  0.3,  0.4,  0.5,  0.6,  0.7,  0.8,  0.9,  1.0,
     21.1,  1.2,  1.3,  1.4,  1.5,  1.6,  1.7,  1.8,  1.9,  2.0,
     32.1,  2.2,  2.3,  2.4,  2.5,  2.6,  2.7,  2.8,  2.9,  3.0,
     43.1,  3.2,  3.3,  3.4,  3.5,  3.6,  3.7,  3.8,  3.9,  4.0,
     54.1,  4.2,  4.3,  4.4,  4.5,  4.6,  4.7,  4.8,  4.9,  5.0,
     66.0,  7.0,  8.0,  9.0, 10.0, 11.0, 12.0, 13.0, 14.0, 15.0,
     716.0, 17.0, 18.0, 19.0, 20.0, 21.0, 22.0, 23.0, 24.0, 25.0,
     826.0, 27.0, 28.0, 29.0, 30.0, 31.0, 32.0, 33.0, 34.0, 35.0,
     936.0, 37.0, 38.0, 39.0, 40.0, 41.0, 42.0, 43.0, 44.0, 45.0,
     046.0, 47.0, 48.0, 49.0, 50.0, 51.0, 52.0, 53.0, 54.0, 55.0,
     156.0, 57.0, 58.0, 59.0, 60.0, 61.0, 62.0, 63.0, 64.0, 65.0,
     266.0, 67.0, 68.0, 69.0, 70.0, 71.0, 72.0, 73.0, 74.0, 75.0,
     376.0, 77.0, 78.0, 79.0, 80.0, 81.0, 82.0, 83.0, 84.0, 85.0,
     486.0, 87.0, 88.0, 89.0, 90.0, 91.0, 92.0, 93.0, 94.0, 95.0,
     596.0, 97.0, 98.0, 99.0,100.0,105.0,110.0,115.0,120.0,125.0,
     6130.0,135.0,140.0,145.0,150.0,155.0,160.0,165.0,170.0,175.0,
     7180.0,185.0,190.0,195.0,200.0/
      DATA  (EY(I),I=1,166)/ 0.0,
     10.667456, 0.720737, 0.772322, 0.822310, 0.870793, 0.917858,
```

```
     20.963585, 1.008050, 1.051310, 1.093450, 1.134520, 1.174580,
     31.213670, 1.251860, 1.289180, 1.325670, 1.361390, 1.396360,
     41.430630, 1.464220, 1.497160, 1.529490, 1.561230, 1.592410,
     51.623050, 1.653180, 1.682820, 1.711980, 1.740690, 1.768960,
     61.796820, 1.824270, 1.851340, 1.878030, 1.904370, 1.930360,
     71.956010, 1.981340, 2.006370, 2.031090, 2.055520, 2.079670,
     82.103540, 2.127150, 2.150510, 2.173620, 2.196490, 2.219130,
     92.241530, 2.263720, 2.474750, 2.669180, 2.850370, 3.020710,
     03.181940, 3.335390, 3.482080, 3.622830, 3.758310, 3.889080,
     14.015590, 4.138230, 4.257340, 4.373210, 4.486080, 4.596190,
     24.703720, 4.808840, 4.911720, 5.012480, 5.111260, 5.208160,
     35.303300, 5.396760, 5.488620, 5.578980, 5.667890, 5.755430,
     45.841660, 5.926630, 6.010410, 6.093030, 6.174540, 6.255000,
     56.334430, 6.412880, 6.490380, 6.566960, 6.642660, 6.717510,
     66.791540, 6.864760, 6.937220, 7.008920, 7.079900, 7.150170,
     77.219760, 7.288690, 7.356970, 7.424620, 7.491660, 7.558110,
     87.623980, 7.689280, 7.754030, 7.818250, 7.881940, 7.945120,
     98.007810, 8.070010, 8.131730, 8.192980, 8.253790, 8.314140,
     08.374070, 8.433560, 8.492640, 8.551320, 8.609590, 8.667470,
     18.724960, 8.782080, 8.838830, 8.895220, 8.951260, 9.006940,
     29.062280, 9.117290, 9.171970, 9.226320, 9.280360, 9.334080,
     39.387490, 9.440600, 9.493420, 9.545940, 9.598180, 9.650130,
     49.701800, 9.753200, 9.804340, 9.855200, 9.905810, 9.956150,
     510.006200,10.253000,10.494000,10.729600,10.960200,11.185900,
     611.407200,11.624300,11.837400,12.046800,12.252600,12.454900,
     712.654100,12.850100,13.043200,13.233500,13.421100,13.606100,
     813.788600,13.968700,14.146600/
      IF(ICHD .EQ. 2)GO TO 100
      DO 10 I=1,165
      IF(Y .LT. EY(I))GO TO 30
   10 CONTINUE
      X=Y**2-0.375
      RETURN
   30 A=(THETA(I)-THETA(I-1))/(EY(I)-EY(I-1))
      X=THETA(I-1)+A*(Y-EY(I-1))
      RETURN
  100 CONTINUE
      DO 110 I=1,165
      IF(X .LT. THETA(I))GO TO 130
  110 CONTINUE
      Y=SQRT(X+0.375)
      RETURN
  130 A=(EY(I)-EY(I-1))/(THETA(I)-THETA(I-1))
      Y=EY(I-1)+A*(X-THETA(I-1))
      RETURN
      END

      SUBROUTINE EQN(NP,ID,IG,NX,XEST,BUX,N1,N2,W2,W3)
      DIMENSION XEST(500),BUX(500,5),W3(500),W2(500),
     1ID(500),IG(500),NX(500)
C
C     SUBROUTINE FOR CONVERTING THE ACTUAL(USUAL) AND
C     ESTIMATED POISSON RATES TO OTHER RATES BY THE
C     FUNCTION F. NX(I) AND XEST(I) ARE RESPECTIVELY THE
C     ACTUAL AND ESTIMATED POISSON RATES. THE FUNCTIONAL
C     FORM OF F IS TO BE SPECIFIED BETWEEN TWO STATEMENTS :
C          DO 100 I=N1,N2  &    100 CONTINUE
C     W2(I) AND W3(I) ARE RESPECTIVELY THE TRANSFORMED RATES
C     OF NX(I) AND XEST(I). THAT IS, W2(I) = F { NX(I) }
C     AND  W3(I) = F { XEST(I) }.
C
      DO 100 I=N1,N2
      W2(I)=NX(I)/BUX(I,1)
      W3(I)=XEST(I)/BUX(I,1)
  100 CONTINUE
      RETURN
      END
```

VI
A COMPUTER PROGRAM FOR SIMULTANEOUSLY ESTIMATING PROPORTIONS BY USING STEIN-LIKE ESTIMATION PROCEDURES

The probability that an item sampled from the i-th population will possess certain properties under consideration is p_i, $i = 1, \ldots, k$. In practical work, p_i is considered a proportion. For example, p_i may be the severity index of a vehicle colliding with the i-th type fixed object. In this part, we provide a FORTRAN Computer Program for simultaneously estimating k (population) proportions, $p_1, \ldots, p_k$. This program is written based on the theories and grouping techniques developed in Sections 2 and 4 of Part III.

There are eight sections in this part. In Section 1, we describe the basic data required by the program. The estimation procedures are outlined in Section 2. Eight methods for computing initial estimates of parameters are discussed in Section 3. The input format of the basic data is presented in Section 4. When the option of transforming final estimates of proportions to other estimates is chosen, the user must provide the transformation required by the subroutine EQN. An example is presented in Section 5. The instructions for modifying the program, if needed, to fit a particular problem are given in Section 6. Six examples are presented in Section 7 to show the use of various program options and logic to compute initial estimates of parameters. The program listing is presented in the last section.

1) The Basic Data

The essential data for simultaneously estimating k proportions, $p_1, \ldots, p_k$, are the independent observations $(n_1, X_1), \ldots,$ and (n_k, X_k). n_i is the number of items sampled from the i-th population. X_i and $n_i X_i$ are, respectively, the proportion and number of sampled items possessing certain properties under consideration. Since $n_i X_i$ is a binomial random variable with (proportion) parameter p_i, X_i is the usual estimate of p_i.

The other essential data is the information for computing initial estimates of parameters. This information is termed the 'supplemental' information for discussion purposes.

Quite often, one is also interested in transforming p_i to another parameter q_i through the function F. That is,

$$q_i = F(p_i) \tag{118}$$

When this option is chosen, the user will have to provide the functional form of F in the user-supplied subroutine EQN. Any information required by F must also be provided. This is termed the 'auxiliary' information.

2) What the Program Does

The first step of this program is to use Anscombe's transformation (2, 7) to transform X_i to Z_i. That is,

$$Z_i = \sqrt{n_i + 0.5}\ \mathrm{Sin}^{-1}\left[\frac{n_i}{n_i + 0.75}\ (2X_i - 1)\right] \tag{119}$$

It has been shown in Part III that when $n_i\ p_i$ is at least 4, Z_i is nearly normally distributed with mean θ_i and unit variance, where

$$\theta_i = \sqrt{n_i + 0.5}\ \sum_{j=0}^{n_i} \mathrm{Sin}^{-1}\left(\frac{2j - n_i}{n_i + 0.75}\right)\cdot C_j^{n_i}\ p_i^{\ j}\ (1 - p_i)^{n_i - j} \tag{120}$$

The second step is to compute the initial estimate, U_i, of θ_i by the chosen method described in the next section. In this step, the trace of the idempotent matrix P satisfying the following equation

$$(Z_1, \ldots, Z_k)' - (U_1, \ldots, U_k)' = P\ (Z_1, \ldots, Z_k)' \tag{121}$$

is also computed. The superscript $'$ stands for the transpose of a vector or matrix. The third step is to compute the shrinking factor c defined as

$$c = \frac{\mathrm{Trace}\ (P) - 2}{\sum_{i=1}^{k} (Z_i - U_i)^2} \tag{122}$$

A slightly better estimation procedure is to set c to be 1 if it is greater than 1. The fourth step is to compute $\hat{Z}_i$ defined as

$$\hat{Z}_i = U_i + (1 - c)\ (Z_i - U_i),\ i = 1, \ldots, k \tag{123}$$

$\hat{Z}_i$ is the Stein-like estimate of θ_i. The fifth step is to obtain $\hat{X}_i$ which is the p_i satisfying Eq. (120) with θ_i replaced by $\hat{Z}_i$. However, an alternative method of obtaining $\hat{X}_i$ is through the inverse function of Eq. (119). That is,

$$\hat{X}_i = \frac{1}{2}\ \left[\frac{n_i + 0.5}{n_i}\ \mathrm{Sin}\left(\frac{\hat{Z}_i}{n_i + 0.5}\right)\ + 1\right] \tag{124}$$

X_i is also considered a Stein-like estimate of p_i. If the option of transforming p_i to q_i specified in Eq. (118) is chosen, the last step is to estimate q_i:

$$\hat{q}_i = F(\hat{X}_i) \tag{125}$$

$\hat{X}_i$ and $\hat{q}_i$ are the final estimates of p_i and q_i, respectively.

This program also computes the estimated percentage improvement of the above procedure over the usual one.

3) Methods for Computing Initial Estimates

The key to obtaining good estimates of proportions is to provide good initial estimates $U_1, \ldots, U_k$ in the sense that N is high and ζ is low, where

$$N = \text{Trace (P)} \tag{126}$$

and

$$\zeta = \sum_{i=1}^{k} (U_i - \theta_i)^2 \tag{127}$$

We have shown in Part III that if P is symmetrical and idempotent, N and ζ are, respectively, the number of degrees of freedom and the non-centrality parameter of a non-central chi-square distribution. The maximal percentage improvement that can be achieved is $100 (N - 2)/k$. Eight methods for computing initial parameter estimates are built in this program. These are:

Method 1: Initial Estimates are Given - Based on past experiments or independent parallel studies, we estimate or guess p_i to be V_i, $i = 1, \ldots, k$. The initial estimate of θ_i is then obtained from Eq. (119). That is,

$$U_i = \sqrt{n_i + 0.5}\ \text{Sin}^{-1}\left[\frac{n_i}{n_i + 0.75}\ (2V_i - 1)\right] \tag{128}$$

In this case, we treat $U_i - \theta_i$ as the parameter to be estimated. Consequently, the matrix P satisfying Eq. (121) is the identity matrix. Thus, N = Trace (P) = k. We note that the maximal trace of P used in Eq. (121) is k. The estimation accuracy is the degree of closeness of $(V_1, \ldots, V_k)$ to $(p_1, \ldots, p_k)$. The final estimate of p_i always lies between the initial estimate V_i and the usual estimate X_i. When the initial estimates are excellent, say $V_i = p_i$ for all i, the final estimate of p_i is V_i for all i. However, if the initial estimates are poor, i.e., $(V_1, \ldots, V_k)$ is quite distant from $(p_1, \ldots, p_k)$, the final estimate of p_i will be very close to X_i for every i.

This method is used only when the number of proportions to be estimated is at least 3. When the independent initial estimates are reliably close to the true proportions, this method will provide good estimates of parameters.

Method 2: Weighted Average (I) - When n_i is fairly large, $n_i/(n_i + 0.75) \simeq 1$ and, therefore, $Z_i \simeq \sqrt{n_i + 0.5}\ \mathrm{Sin}^{-1}(2X_i - 1)$. Thus, if the variation among proportions is small, all $\theta_i/\sqrt{n_i + 0.5}$ would be near a constant b. The weighted average $\hat{b}$ can then be used to estimate b. That is,

$$\hat{b} = \sum_{i=1}^{k} Z_i \Big/ \sum_{i=1}^{k} \sqrt{n_i + 0.5} \tag{129}$$

We then take $U_i = \hat{b}\sqrt{n_i + 0.5}$ as the initial estimate of θ_i. For this method, we have $N = \mathrm{Trace}\ (P) = k - 1$. The loss of one degree of freedom is due to the use of $\hat{b}$ as an estimate of b. We note that the above U_i can be rewritten as

$$U_i = w_i \sum_{j=1}^{k} Z_j \tag{130}$$

with

$$w_i = \sqrt{n_i + 0.5} \Big/ \sum_{j=1}^{k} \sqrt{n_j + 0.5} \tag{131}$$

It is obvious that

$$w_1 + w_2 + \ldots + w_k = 1 \tag{132}$$

This method is used only when the number of proportions to be estimated is at least 4. When all p_i are near a constant, this method will produce good estimates of proportions. In practice, one may graphically examine the relationship, $Z_i = b\sqrt{n_i + 0.5}$, to ascertain the degree of linearity.

Method 3: Weighted Average (II) - For a given set of numbers, w_1, . . . , w_k, satisfying Eq. (132), the initial estimate of θ_i is taken to be

$$U_i = w_i \sum_{j=1}^{k} Z_j \tag{133}$$

In this case, N = Trace (P) = k - 1. We note that this is the same form used in Method 2. The only difference is that w_i in Method 2 was computed from the sample sizes, but is a given number in this method. When $n_1 = . . . = n_k$ and $w_1 = . . . = w_k$, Methods 2 and 3 are identical and use the average of Z_1, . . . , Z_k as the initial estimate of every θ_i.

This method is used only when the number of proportions to be estimated is at least 4. When the relationship $\theta_i = b\, w_i$ holds approximately for a given set of numbers satisfying Eq. (132), this method will produce good estimates of proportions. In practice, one may graphically examine the relationship, $Z_i = b\, w_i$, to check the linearity. For example, we have approximately $Z_i = b\, t_i$, i = 1, . . . , k, then $w_i = t_i \Big/ \sum_{j=1}^{k} t_j$. Note that, if $t_i = 1$, $w_i = 1/k$.

Method 4: Weighted Average (III) - For a given set of numbers, w_1, . . . , w_k, satisfying Eq. (132), the initial estimate of θ_i is taken to be

$$U_i = \sum_{j=1}^{k} w_j Z_j \tag{134}$$

For this case, N = Trace (P) = k - 1. Methods 3 and 4 are identical when all w_i are equal to 1/k.

This method is used only when the number of proportions to be estimated is at least 4. When all p_i are almost equal to a constant, this method will produce good estimates of proportions for any set of numbers satisfying Eq. (132). In this circumstance, w_i can be interpreted as the weight assigned to Z_i for estimating the common constant. If every X_i or Z_i is equally reliable, equal weight can be assigned to every Z_i. That is, $w_i = 1/k$, i = 1, . . . , k. However, if one has reason to doubt the reliability of a particular observation, say X_1, zero weight may be assigned to Z_1 and equal weight to the remaining observations. That is, $w_1 = 0$ and $w_i = 1/(k - 1)$, i = 2, . . . , k.

Method 5: Least Squares Estimate (I) - Suppose that the supplemental variables $t_{i1}, \ldots,$ and t_{im} are available and satisfy the following equation,

$$\theta_i = \sqrt{n_i + 0.5}\ (b_{i1} t_{i1} + \ldots + b_{im} t_{im}),\ i = 1, \ldots, k \qquad (135)$$

If for every j, $b_{1j}, \ldots,$ and b_{kj} are near an unknown common parameter b_j, Eq. (135) can be rewritten as

$$\theta_i = \sqrt{n_i + 0.5}\ (b_1 t_{i1} + \ldots + b_m t_{im}),\ i = 1, \ldots, k \qquad (136)$$

Denote $(\hat{b}_1, \ldots, \hat{b}_m)$ to be the least squares estimate of $(b_1, \ldots, b_m)$. We then take the initial estimate of θ_i to be

$$U_i = \sqrt{n_i + 0.5}\ (\hat{b}_1 t_{i1} + \ldots + \hat{b}_m t_{im}),\ i = 1, \ldots, k \qquad (137)$$

In this case, N = Trace (P) = k - m. The loss of m degrees of freedom is due to the least squares estimates of m linear parameters. We note that U_i can always be improved by increasing the number of supplemental variables. However, this decreases the number of degrees of freedom, N, which is an undesirable feature as previously mentioned. In general, the number of supplemental variables should be kept small relative to the number of parameters to be estimated.

This method is used only when the number of proportions to be estimated is at least m + 3. This method will produce good estimates of proportions if Eq. (136) holds approximately and the number of supplemental variables is small relative to the number of proportions to be estimated. In practice, one may examine the data to see whether the following linear relationships hold approximately.

$$Z_i / \sqrt{n_i + 0.5} = b_1 t_{i1} + \ldots + b_m t_{im},\ i = 1, \ldots, k \qquad (138)$$

or, when all n_i are fairly large,

$$\text{Sin}^{-1} (2X_i - 1) = b_1 t_{i1} + \ldots + b_m t_{im},\ i = 1, \ldots, k \qquad (139)$$

Method 6: Least Squares Estimate (II) - This method is the same as Method 5 except that Eqs. (136) and (137) are, respectively, replaced by

$$\theta_i = \sqrt{n_i + 0.5}\ (b_0 + b_1 t_{i1} + \ldots + b_m t_{im}),\ i = 1, \ldots, k \qquad (140)$$

and

$$U_i = \sqrt{n_i + 0.5}\ (\hat{b}_0 + \hat{b}_1 t_{i1} + \ldots + \hat{b}_m t_{im}),\ i = 1, \ldots, k \qquad (141)$$

In this case, $N = \text{Trace}(P) = k - m - 1$ because of the extra parameter b_0. Thus, this method is used only when the number of proportions to be estimated is at least $m + 4$. In practice, one may examine the data to see whether the following linear relationships hold approximately.

$$Z_i / \sqrt{n_i + 0.5} = b_0 + b_1 t_{i1} + \ldots + b_m t_{im}, \; i = 1, \ldots, k \tag{142}$$

or when all n_i are fairly large,

$$\sin^{-1}(2X_i - 1) = b_0 + b_1 t_{i1} + \ldots + b_m t_{im}, \; i = 1, \ldots, k \tag{143}$$

Method 7: Least Squares Estimate (III) - This method is the same as Method 5 except that Eqs. (136) and (137) are, respectively, replaced by the following equations.

$$\theta_i = b_1 t_{i1} + \ldots + b_m t_{im}, \; i = 1, \ldots, k \tag{144}$$

and

$$U_i = \hat{b}_1 t_{i1} + \ldots + \hat{b}_m t_{im}, \; i = 1, \ldots, k \tag{145}$$

In practice, one may examine the data to see whether the following linear relationship holds approximately.

$$Z_i = b_1 t_{i1} + \ldots + b_m t_{im}, \; i = 1, \ldots, k \tag{146}$$

Method 8: Least Squares Estimate (IV) - This method is the same as Method 7 except that Eqs. (144) and (145) are, respectively, replaced by the following equations.

$$\theta_i = b_0 + b_1 t_{i1} + \ldots + b_m t_{im}, \; i = 1, \ldots, k \tag{147}$$

and

$$U_i = \hat{b}_0 + \hat{b}_1 t_{i1} + \ldots + \hat{b}_m t_{im}, \; i = 1, \ldots, k \tag{148}$$

In practice, one may examine the data to see whether the following linear relationship holds approximately.

$$Z_i = b_0 + b_1 t_{i1} + \ldots + b_m t_{im}, \; i = 1, \ldots, k \tag{149}$$

The k proportions may be arranged into many groups to which different methods are applied to minimize the noncentrality parameter defined in

Eq. (127). For example, we may use Method 1 for a group of k_1 proportions, Method 5 for a group of k_2 proportions, Method 2 for a group of k_3 proportions, and again, Method 2 for the group of the remaining proportions. In this case, we have $N = \text{Trace}(P) = k_1 + (k_2 - m) + (k_3 - 1) + (k - k_1 - k_2 - k_3 - 1) = k - m - 2$. m is the number of linear parameters used in Method 5. For this particular case, the number of proportions to be estimated should be at least $m + 5$. In general, for any method or combination of methods, the number of degrees of freedom should be at least 2. Combining these eight methods to compute initial parameter estimates will be demonstrated by examples presented in Section 7.

Methods 2 through 8 suggest that k populations should be arranged into groups in which parameters cluster at a point or can be approximated by a linear function of supplemental variables. The physical properties of populations, past experiments and parallel studies are good sources for obtaining the proper group-method combination. Unfortunately, these sources may not be available or sufficiently reliable in a particular problem. In this circumstance, one may examine the data as suggested in each method to choose the proper group-method combination. That is, the empirical relationships among X_i, Z_i, $Z_i/\sqrt{n_i + 0.5}$ and supplemental variables can be used to determine the group-method combination for computing initial parameter estimates. Since the program computes the estimated percentage improvement of each chosen group-method combination over the usual method, one may use the one that produces the maximal improvement to estimate proportions. We remark that the group-method combination generated by examining the data may not be the best one and, possibly, could be the worst one for this problem due to random variation of the data. Nevertheless, if the chosen group-method combination is used thereafter for the same problem, the above procedure is always better than the usual one. The worst situation is that no improvement is made. Based on our experience, the qualitative properties of populations often provide adequate information for grouping purposes.

4) Data Input

The data input of this program is arranged into two portions. The first portion is composed of six cards. The number of cards in the second portion is equal to the number of proportions to be estimated.

The first six cards specify the number of groups and proportions, computational methods, supplemental and auxiliary variables in each group. Variables used in these cards are defined below.

NP: Number of proportions to be estimated, $1 \leq NP \leq 500$

NGROUP: Number of groups used, $1 \leq NGROUP \leq 20$

NMG(I): Number of proportions in the I-th group, NMG(1) + . . . + NMG (NGROUP) = NP

METHOD(I) = j: The j-th method presented in Section 3 is used to compute initial estimates of parameters in the I-th group, $1 \leq j \leq 8$

NAUX(I): Number of supplemental variables used to compute initial estimates of parameters in the I-th group. This variable is 0 if METHOD(I) = 2, and 1 if METHOD(I) = 1, 3, and 4.

NCOV(I):

- >0: Number of auxiliary variables used to transform proportions in the I-th group, $1 \leq$ NAUX(I) + NCOV(I) ≤ 5.
- $= 0$: Means that supplemental variables are also served as auxiliary variables.
- <0: No transformation.

(DESCPT(I), I = 1, 70): Title (no more than 70 letters).

The second portion is composed of NGROUP subportions or groups. The I-th subportion is composed of NMG(I) cards. Define K = J if I = 1 and K = NMG(1) + . . . + NMG(I - 1) + J if I 1. Then, the K-th card of the second portion contains essential and supplemental data for estimating the J-th proportion of the I-th group. We note that the J-th proportion of the I-th group is the proportion of the K-th population. The data input for estimating this proportion are as follows:

METHOD(I)	NCOV(I)	DATA INPUT
2	0	ID(K),X(K),NSAMP(K)
	NOT 0	ID(K),X(K),NSAMP(K),(BUX(K,M),M=1,NCOV(I))
NOT 2	0	ID(K),X(K),NSAMP(K),(AUX(J,M),M=1,NAUX(I))
	NOT 0	ID(K),X(K),NSAMP(K),(AUX(J,M),M=1,NAUX(I)), (BUX(K,M),M=1,NCOV(I))

Variables used in this card are defined below:

ID(K): Identification number of the K-th population such as location number and year, etc. The K-th proportion is the J-th proportion of the I-th group. This number has no effect on the estimation procedure.

X(K): The K-th sample proportion (usual estimate of the J-th proportion of the I-th group).

NSAMP(K): Number of items sampled from the K-th population. That is, the sample size for measuring X(K).

AUX(J, M): The M-th supplemental variable for the K-th proportion.

BUX(K, M): The M-th auxiliary variable for the K-th proportion.

Table 29
The Input Deck

Card Number	Variables Used in Each Card	Remarks
1 2 3 4 5 6	NP , NGROUP (NMG(I), I=1,...,NGROUP) (METHOD(I), I=1,...,NGROUP) (NAUX(I),I=1,NGROUP) (NCOV(I),I=1,NGROUP) (DESPT(I),I=1,70)	* The Data Input of The First Portion
7 . . . L+6	ID(1),X(1),NSAMP(1),(AUX(1,M),M=1,NA),(BUX(1,M),M=1,NC) ID(L),X(L),NSAMP(L),(AUX(L,M),M=1,NA),(BUX(L,M),M-1,NC)	* L=NMG(1) * NA=NAUX(1);NC=NCOV(I) * L cards for The First Group
⋮	⋮	⋮
. . . K+6 . .	ID(K),X(K),NSAMP(K),(AUX(J,M),M=1,NA),(BUX(K,M),M=1,NC)	* NA=NAUX(I);NC=NCOV(I) * K=NMG(1)+..+NMG(I-1)+J * This card is for the k-th parameter which is the J-th parameter of the I-th group.
⋮	⋮	⋮
N+6 . . . NP+6	ID(N),X(N),NSAMP(N),(AUX(1,M),M=1,NA),(BUX(N,M),M=1,NC) ID(NP),X(NP),NSAMP(NP),(AUX(L,M),M=1,NA),(BUX(NP,M),M=1,NC)	* L=NMG(NGROUP) * NA=NAUX(NGROUP) * NC=NCOV(NGROUP) * N=NP-L+1 * L Cards for The Last GROUP

The input deck is presented in Table 29. This deck is also diagrammed in Figure 6 to show the format and logic used. The user can follow this diagram to change, if needed, read statements and formats to fit a particular problem.

5) The User-Supplied Subroutine (EQN)

The functional form of the transformation F defined in Eq. (118) must be specified between two statements, DO 100 I = N1, N2 and 100 CONTINUE, in the user-supplied subroutine EQN. The input variables for this subroutine are NP, ID, IG, NSAMP, X, XEST, BUX, N1, and N2. The meaning of NP, ID, NSAMP, X, and BUX were explained in Section 4. IG(i) is the group number to which the i-th proportion belongs. XEST(i) is the final estimate of the i-th proportion. N1 and N2 are integers automatically set in the main program. The output of this subroutine is W2(i) and W3(i) for i = N1, . . . , N2, where W2(i) = F[X(i)] and W3(i) = F[XEST(i)].

Example: We are also interested in estimating parameters q_i satisfying the following equations.

$$p_i = q_i t_i + (q_i t_i)^2, \quad i = 1, \ldots, 10 \tag{150}$$

and

$$p_i = q_i M_i, \quad i = 11, \ldots, 100 \tag{151}$$

In this case, t_i is the only auxiliary variable for the first 10 populations to convert proportions to q_i defined in Eq. (150), while M_i is the only auxiliary variable for the last 90 populations to convert proportions to q_i defined in Eq. (151). Therefore

$$\text{BUX}(i, 1) = \begin{cases} t_i, & i = 1, \ldots, 10 \\ M_i, & i = 11, \ldots, 100 \end{cases} \tag{152}$$

solving Eqs. (150) and (151), we obtain

$$q_i = F(p_i) = \begin{cases} (\sqrt{p_i + 0.25} - 0.5)/t_i \\ p_i/M_i \end{cases} \tag{153}$$

The t_i and M_i can be, for example, the ADT and the number of years, respectively. The usual estimate of q_i is then obtained from Eq. (153) with p_i replaced by X_i. Similarly, the final estimate of q_i is also obtained from Eq. (153) with p_i replaced by the final estimate of p_i. The function F defined in Eq. (153) must be specified in the user-supplied subroutine EQN.

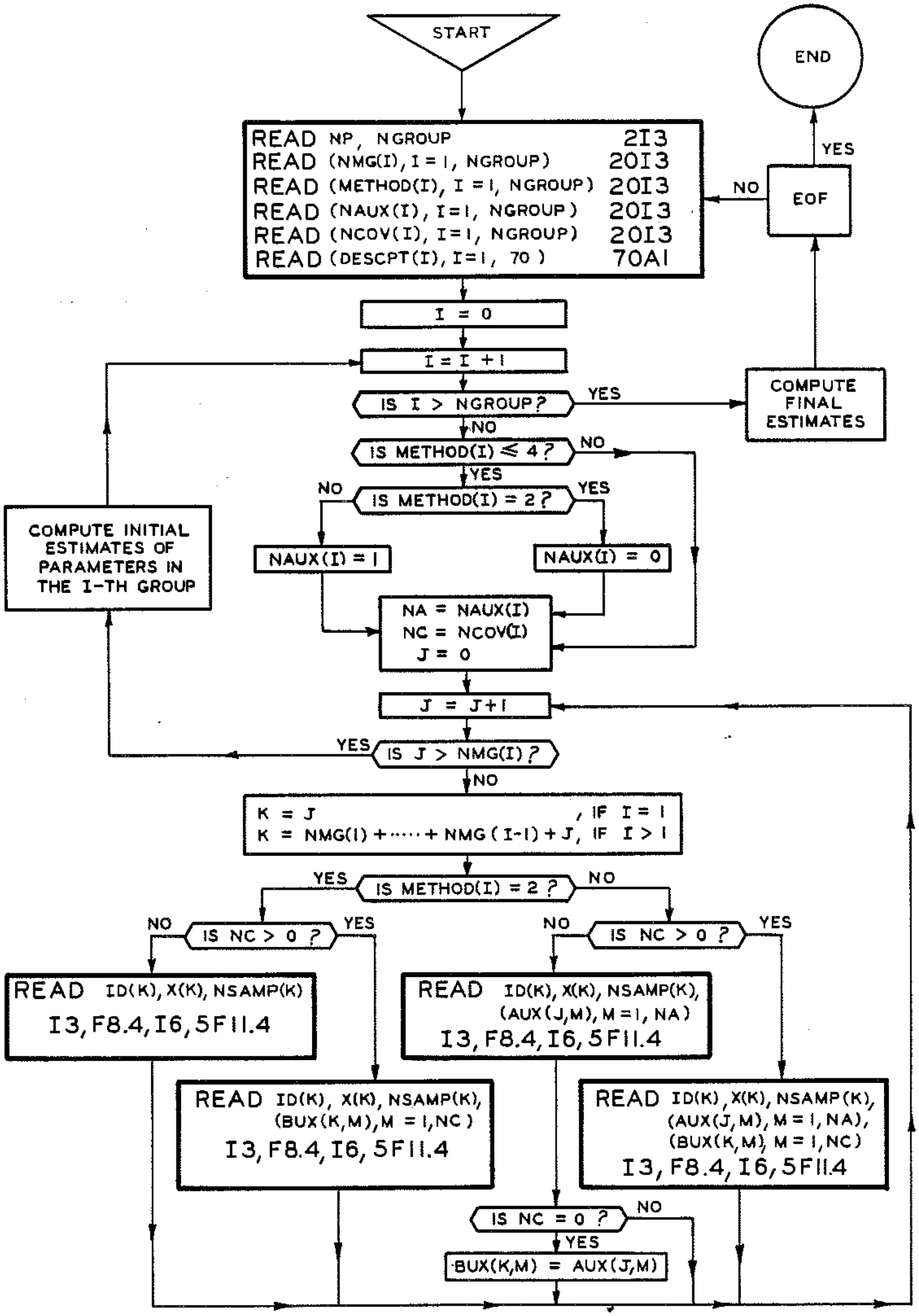

Figure 6. Flow of control for data input.

For example,

```
      DO 100 I=N1,N2
      IF(I .GT. 10) GO TO 50
      W2(I)=(SQRT(4. * X(I) +1.)-1.)/(BUX(I,1)*2.)
      W3(I)=(SQRT(4. * XEST(I)+1.)-1.)/(BUX(I,1)*2.)
      GO TO 100
   50 W2(I)=X(I)/BUX(I,1)
      W3(I)=XEST(I)/BUX(I,1)
  100 CONTINUE
```

6) Limitations of the Program and How to Make Necessary Changes

This program was designed to handle problems where the number of proportions, NP, does not exceed 500. This number can be easily increased to any desired number. However, one must change the dimensions of variables that appear on the DIMENSION STATEMENTS of the main program and subroutines LSE and EQN, i.e., replace each 500 by the desired number.

The number of groups is limited to 20. This number is large enough for most practical problems. However, one can increase this number to any desired number not exceeding NP by properly changing the READ format of the second input card (FORMAT 3).

The total number of supplemental and auxiliary variables is limited to 5. This number is large enough for the practical applications. The user can change this number to any desired number, say N, by performing the following two steps.

a) Change the READ format of the second portion of the data input (FORMAT 60).

b) Change dimensions of variables that appear on the DIMENSION STATEMENTS of the main program and subroutines LSE, MTXINV, and EQN, i.e., replace 5 and 6 by N and N + 1, respectively.

7) Examples

Every 0.2-mile segment of roadway is classified into categories according to certain physical properties. There are 24 categories coded as 1 through 24. Accident statistics and average ADT figures are presented in Table 30. Denote A_{ij} to be the ADT of the j-th segment of the i-th category. Based on the previous study (4), the probability that there is at least

Table 30
Proportion of Segments having Accidents
During A Year And Average ADT in Each Category

Category	Proportion	No. Of Segments	Average ADT
1	0.0658	2355	1420.97
2	0.0601	2695	1605.43
3	0.0854	1850	2562.92
4	0.0518	3455	1439.33
5	0.0560	12385	1611.50
6	0.0745	10900	2468.06
7	0.0629	2975	1182.83
8	0.0807	3345	1595.71
9	0.1178	1740	2527.18
10	0.0578	2060	1462.57
11	0.0729	5270	1891.70
12	0.0954	6070	2654.28
13	0.0600	800	1511.23
14	0.0907	915	2167.14
15	0.1126	835	2830.36
16	0.0661	1120	1601.36
17	0.0686	3700	1798.27
18	0.0826	3860	2531.98
19	0.1020	1245	1408.66
20	C.1041	1690	2028.42
21	0.1133	830	2859.70
22	0.0902	610	1601.42
23	0.0776	2050	1930.16
24	0.1009	2320	2995.37

one accident occurring in a time period on this segment is $q_i A_{ij}$. Furthermore, the expected number of accidents per year for this segment is R_{ij}, defined as

$$R_{ij} = q_i A_{ij} + (q_i A_{ij})^2 \tag{154}$$

We are interested in estimating parameters q_i, $i = 1, \ldots, 24$. For this purpose, we denote n_i to be the number of segments in the i-th category. We also denote t_i to be the average ADT of the i-th category. Define X_i to be the proportion of segments in the i-th category having accidents during a year. If every A_{ij} is near t_i, $n_i X_i$ is approximately binomially distributed with parameter $p_i = q_i t_i$. For demonstration purposes, we shall assume that $n_i X_i$ is binomially distributed with parameter p_i. Thus, the usual estimate of p_i and q_i are X_i and X_i/t_i, respectively. In this section, we present six examples which demonstrate improvement over the usual estimates of p_i and q_i for $i = 13, \ldots, 24$.

After examining the relationships among Z_i, $Z_i/\sqrt{n_i + 0.5}$ and t_i, we see that $Z_i/\sqrt{n_i + 0.5}$ is almost a constant or a linear function of t_i. Thus, Methods 2 and 6 (using the average ADT as the supplemental variable) are applicable to this problem. The second portion of the data input for using these two methods is presented in Table 31. The data in Table 31 are the common data for Examples 1 through 4. We note that the only difference

Table 31
The Second Portion of The Data Input
of Examples 1 Through 4

Card No.	Column Number 12345678901234567890123456789012345678901234567890			
7	13	0.06000	800	1511.23
8	14	0.09070	915	2167.14
9	15	0.11260	835	2830.36
10	16	0.06610	1120	1601.36
11	17	0.06860	3700	1798.27
12	18	0.08260	3860	2531.98
13	19	0.10200	1245	1408.66
14	20	0.10410	1690	2028.42
15	21	0.11330	830	2859.70
16	22	0.09020	610	1601.42
17	23	0.07760	2050	1930.16
18	24	0.10090	2320	2995.37

between segments in the i-th and (i-12)-th categories for every $i = 13, \ldots, 24$, is that segments in one category of each pair are intersected by another roadway. We now define

$$V_i = \frac{X_{i-12}}{t_{i-12}} t_i, \; i = 13, \ldots, 24 \tag{155}$$

V_i can be interpreted as the sample proportion of the (i-12)-th category when the ADT is t_i. We observe graphically that the following linear relationship holds approximately.

$$\text{Sin}^{-1}(2X_i - 1) = b \,\text{Sin}^{-1}(2V_i - 1), \; i = 13, \ldots, 24 \tag{156}$$

Thus, Method 5 with $\text{Sin}^{-1}(2V_i - 1)$ as the supplemental variable is also applicable to this problem. This method will be presented in Example 6.

Example 1: Using Method 2 with One Group - Since all $Z_i/\sqrt{n_i + 0.5}$ or $\text{Sin}^{-1}(2X_i - 1)$ are near a constant b, i.e., $Z_i = b\sqrt{n_i + 0.5}$, we shall use the weighted average to estimate the constant (Method 2). In this case, we do not need the supplemental variable. t_i is the auxiliary variable for converting p_i to q_i defined below.

$$q_i = F(p_i) = p_i/t_i \tag{157}$$

The first six cards of the data input for this case are presented below.

CARD NO.	COLUMN NUMBER 12345678901234567890123456789012345678901234567890
1	12 1
2	12
3	2
4	0
5	1
6	METHOD 2 ON 1 GROUP

The first card sets NP = 12 and NGROUP = 1. Thus, NMG(1) = 12 which is set in the second card. The third card sets METHOD(1) = 2. The fourth and fifth cards, respectively, set NAUX(1) = 0 and NCOV(1) = 1. Thus, BUX(i, 1) = t_i, i = 1, . . . , 12. The sixth card is the title of the run. Since NCOV(1) = 1, the final estimate of p_i will be converted to q_i defined in Eq. (157). To do this, we must specify Eq. (157) in the user-supplied subroutine EQN. For example,

```
      DO 100 I=N1,N2
      W2(I)=X(I)/BUX(I,1)
      W3(I)=XEST(I)/BUX(I,1)
100   CONTINUE
```

The run results are presented in Table 32.

Table 32
Estimated Results of Example 1

```
######################################################################
METHOD 2 ON 1 GROUP
######################################################################

DATA :
```

ID	P	SAMPLE SIZE	GROUP NO	SUPPLEMENTAL AND AUXILIARY INFORMATION
13	0.0600	800	1	1511.230
14	0.0907	915	1	2167.140
15	0.1126	835	1	2830.360
16	0.0661	1120	1	1601.360
17	0.0686	3700	1	1798.270
18	0.0826	3860	1	2531.980
19	0.1020	1245	1	1408.660
20	0.1041	1690	1	2028.420
21	0.1133	830	1	2859.700
22	0.0902	610	1	1601.420
23	0.0776	2050	1	1930.160
24	0.1009	2320	1	2995.370

ESTIMATED RESULTS :

ID	PROPORTION ACTUAL	PROPORTION ESTIMATED	CONVERTED PARAMETER ACTUAL	CONVERTED PARAMETER ESTIMATED
13	0.0600000	0.0635152	0.000039703	0.000042029
14	0.0907000	0.0901448	0.000041852	0.000041596
15	0.1126000	0.1087636	0.000039783	0.000038427
16	0.0661000	0.0688844	0.000041277	0.000043016
17	0.0686000	0.0710930	0.000038148	0.000039534
18	0.0826000	0.0832269	0.000032623	0.000032870
19	0.1020000	0.0997963	0.000072409	0.000070845
20	0.1041000	0.1015904	0.000051321	0.000050083
21	0.1133000	0.1093550	0.000039620	0.000038240
22	0.0902000	0.0896924	0.000056325	0.000056008
23	0.0776000	0.0789018	0.000040204	0.000040878
24	0.1009000	0.0988781	0.000033685	0.000033010

```
SHRINKING FACTOR                    =  0.8584
%-IMPROVEMENT OVER USUAL ESTIMATE = 11.2895 %
```

Example 2: Using Method 2 with Two Groups - Based on physical properties of Categories 13 through 24, the noncentrality parameter can be further reduced by separating 12 categories into two groups: Categories 13 through 18 and 19 through 24. We then use Method 2 to compute initial estimates of parameters in each group. For this case, the first six cards of the data input take the following form.

CARD NO.	COLUMN NUMBER 1234567890123456789012345678901234567890123456789O
1	12 2
2	6 6
3	2 2
4	0 0
5	1 1
6	METHOD 2 ON 2 GROUPS(13-18 & 19-24)

The user-supplied subroutine EQN is the same as the one used in Example 1. The run results are presented in Table 33.

Table 33
Estimated Results of Example 2

##
METHOD 2 ON 2 GROUPS(13-18 & 19-24)
##

DATA :

ID	P	SAMPLE SIZE	GROUP NO	SUPPLEMENTAL AND AUXILIARY INFORMATION
13	0.0600	800	1	1511.230
14	0.0907	915	1	2167.140
15	0.1126	835	1	2830.360
16	0.0661	1120	1	1601.360
17	0.0686	3700	1	1798.270
18	0.0826	3860	1	2531.980
19	0.1020	1245	2	1408.660
20	0.1041	1690	2	2028.420
21	0.1133	830	2	2859 700
22	0.0902	610	2	1601.420
23	0.0776	2050	2	1930.160
24	0.1009	2320	2	2995.370

ESTIMATED RESULTS :

ID	PROPORTION ACTUAL	PROPORTION ESTIMATED	CONVERTED PARAMETER ACTUAL	CONVERTED PARAMETER ESTIMATED
13	0.0600ᴐᴐ0	0.0634044	0.000039703	0.000041956
14	0:0907000	0.0880946	0.000041852	0.000040650
15	0.1126000	0.1052409	0.000039783	0.000037183
16	0.0661000	0.0684100	0.000041277	0.000042720
17	0.0686000	0.0704768	0.000038148	0.000039191
18	0.0826000	0.0817232	0.000032623	0.000032276
19	0.1020000	0.1009989	0.000072409	0.000071699
20	0.1041000	0.1026843	0.000051321	0.000050623
21	0.1133000	0.1099402	0.000039620	0.000038445
22	0.0902000	0.0914944	0.000056325	0.000057133
23	0.0776000	0.0813351	0.000040204	0.000042139
24	0.1009000	0.1001444	0.000033685	0.000033433

SHRINKING FACTOR = 0.8005
%-IMPROVEMENT OVER USUAL ESTIMATE = 14.6214 %

Example 3: Using Method 6 with One Group - We mentioned previously that $\sin^{-1}(2X_i - 1)$ is a linear function of t_i. Thus, Method 6 with t_i serving as the supplemental variable can be used to compute initial estimates of parameters. In this case, t_i also serves as the auxiliary variable for converting p_i. Therefore, NAUX(1) = 1 and NCOV(1) = 0. The first six cards of the data input take the following form.

CARD NO.	COLUMN NUMBER 12345678901234567890123456789012345678901234567890
1	12 1
2	12
3	6
4	1
5	0
6	METHOD 6 ON 1 GROUP

The user-supplied subroutine EQN is the same as the one used in Example 1. The run results are presented in Table 34.

Table 34
Estimated Results of Example 3

```
####################################################################
METHOD 6 ON 1 GROUP
####################################################################
DATA :
```

ID	P	SAMPLE SIZE	GROUP NO	SUPPLEMENTAL AND AUXILIARY INFORMATION
13	0.0600	800	1	1511.230
14	0.0907	915	1	2167.140
15	0.1126	835	1	2830.360
16	0.0661	1120	1	1601.360
17	0.0686	3700	1	1798.270
18	0.0826	3860	1	2531.980
19	0.1020	1245	1	1408.660
20	0.1041	1690	1	2028.420
21	0.1133	830	1	2859.700
22	0.0902	610	1	1601.420
23	0.0776	2050	1	1930.160
24	0.1009	2320	1	2995.370

ESTIMATED RESULTS :

ID	PROPORTION ACTUAL	PROPORTION ESTIMATED	CONVERTED PARAMETER ACTUAL	CONVERTED PARAMETER ESTIMATED
13	0.0600000	0.0626453	0.000039703	0.000041453
14	0.0907000	0.0904847	0.000041852	0.000041753
15	0.1126000	0.1109781	0.000039783	0.000039210
16	0.0661000	0.0680950	0.000041277	0.000042523
17	0.0686000	0.0708171	0.000038148	0.000039381
18	0.0826000	0.0849173	0.000032623	0.000033538
19	0.1020000	0.0972752	0.000072409	0.000069055
20	0.1041000	0.1011423	0.000051321	0.000049863
21	0.1133000	0.1116644	0.000039620	0.000039048
22	0.0902000	0.0881805	0.000056325	0.000055064
23	0.0776000	0.0787956	0.000040204	0.000040823
24	0.1009000	0.1018517	0.000033685	0.000034003

SHRINKING FACTOR = 0.8344
%-IMPROVEMENT OVER USUAL ESTIMATE = 11.9604 %

Example 4: Method 2 for Categories 13 through 18 and Method 6 for Categories 19 through 24 - In this example, we use Method 2 to compute initial estimates of parameters in Categories 13 through 18. Thus, NMG(1) = 6, NAUX(1) = 0 and NCOV(1) = 1 as in Example 1. For Categories 19 through 24, we use Method 6 with t_i as the supplemental variable to compute initial estimates of proportions. Thus, NMG(2) = 6, NAUX(2) = 1, NCOV(2) = 0 and AUX(i, 1) = BUX(i, 1) = t_{i+12} for i = 7, . . . , 12 as in Example 3. The first six cards of the data input take the following form.

```
                          COLUMN NUMBER
CARD NO.  12345678901234567890123456789012345678901234567890
    1       12  2
    2        6  6
    3        2  6
    4        0  1
    5        1  0
    6       METHOD 2 ON CATEGORIES 13-18 & METHOD 6 ON 19-24
```

The user-supplied subroutine EQN is the same as the one used in Example 1. The run results are presented in Table 35.

Table 35
Estimated Results of Example 4

```
#################################################################
METHOD 2 ON CATEGORIES 13-18 & METHOD 6 ON 19-24
#################################################################
DATA :
          SAMPLE GROUP
 ID  P      SIZE   NO   SUPPLEMENTAL AND AUXILIARY INFORMATION
-----------------------------------------------------------------
 13 0.0600   800    1     1511.230
 14 0.0907   915    1     2167.140
 15 0.1126   835    1     2830.360
 16 0.0661  1120    1     1601.360
 17 0.0686  3700    1     1798.270
 18 0.0826  3860    1     2531.980
 19 0.1020  1245    2     1408.660
 20 0.1041  1690    2     2028.420
 21 0.1133   830    2     2859.700
 22 0.0902   610    2     1601.420
 23 0.0776  2050    2     1930.160
 24 0.1009  2320    2     2995.370

ESTIMATED RESULTS :
          PROPORTION                      CONVERTED PARAMETER
 ID  ACTUAL      ESTIMATED          ACTUAL            ESTIMATED
-----------------------------------------------------------------
 13  0.0600000   0.0630893          0.000039703       0.000041747
 14  0.0907000   0.0883316          0.000041852       0.000040760
 15  0.1126000   0.1059053          0.000039783       0.000037418
 16  0.0661000   0.0681971          0.000041277       0.000042587
 17  0.0686000   0.0703041          0.000038148       0.000039095
 18  0.0826000   0.0818033          0.000032623       0.000032308
 19  0.1020000   0.1000786          0.000072409       0.000071045
 20  0.1041000   0.1027717          0.000051321       0.000050666
 21  0.1133000   0.1115621          0.000039620       0.000039012
 22  0.0902000   0.0906953          0.000056325       0.000056634
 23  0.0776000   0.0808130          0.000040204       0.000041869
 24  0.1009000   0.1016872          0.000033685       0.000033948

 SHRINKING FACTOR                  =  0.8188
%-IMPROVEMENT OVER USUAL ESTIMATE  = 11.6827 %
```

Table 36
Data Input of Example 5

Card No.	Column Number 12345678901234567890123456789012345678901234567890
1	12 2
2	5 7
3	2 2
4	0 0
5	1 1
6	METHOD 2 ON (13,16-18,23) & (14-15,19-22,24)
7	13 0.0600 800 1511.2300
8	16 0.0661 1120 1601.3600
9	17 0.0686 3700 1798.2700
10	18 0.0826 3860 2531.9800
11	23 0.0776 2050 1930.1600
12	14 0.0907 915 2167.1400
13	15 0.1126 835 2830.3600
14	19 0.1020 1245 1408.6600
15	20 0.1041 1690 2028.4200
16	21 0.1133 830 2859.7000
17	22 0.0902 610 1601.4200
18	24 0.1009 2320 2995.3700

Table 37
Estimated Results of Example 5

```
###################################################################
METHOD 2 ON (13,16-18,23) & (14-15,19-22,24)
###################################################################
DATA :
          SAMPLE GROUP
 ID  P      SIZE   NO  SUPPLEMENTAL AND AUXILIARY INFORMATION
-------------------------------------------------------------------
 13 0.0600  800    1    1511.230
 16 0.0661 1120    1    1601.360
 17 0.0686 3700    1    1798.270
 18 0.0826 3860    1    2531.980
 23 0.0776 2050    1    1930.160
 14 0.0907  915    2    2167.140
 15 0.1126  835    2    2830.360
 19 0.1020 1245    2    1408.660
 20 0.1041 1690    2    2028.420
 21 0.1133  830    2    2859.700
 22 0.0902  610    2    1601.420
 24 0.1009 2320    2    2995.370

ESTIMATED RESULTS :

          PROPORTION                   CONVERTED PARAMETER
 ID  ACTUAL      ESTIMATED         ACTUAL             ESTIMATED
------------------------------------------------------------------
 13  0.0600000  0.0668263       0.000039703         0.000044220
 16  0.0661000  0.0696404       0.000041277         0.000043488
 17  0.0686000  0.0708597       0.000038148         0.000039404
 18  0.0826000  0.0768893       0.000032623         0.000030367
 23  0.0776000  0.0747271       0.000040204         0.000038715
 14  0.0907000  0.0969863       0.000041852         0.000044753
 15  0.1126000  0.1065046       0.000039783         0.000037629
 19  0.1020000  0.1020198       0.000072409         0.000072423
 20  0.1041000  0.1029664       0.000051321         0.000050762
 21  0.1133000  0.1068001       0.000039620         0.000037347
 22  0.0902000  0.0966723       0.000056325         0.000060367
 24  0.1009000  0.1016035       0.000033685         0.000033920

 SHRINKING FACTOR                 =  0.4348
%-IMPROVEMENT OVER USUAL ESTIMATE = 46.3151 %
```

Table 38
Data Input of Example 6

Card No.	Column Number 12345678901234567890123456789012345678901234567890				
1	12	1			
2	12				
3	5				
4	1				
5	1				
6	METHOD 5(PARALLEL DATA AS SUPPLEMENTAL VARIABLE)				
7	13	0.0600	800	-1.0338	1511.23
8	14	0.0907	915	-0.9919	2167.14
9	15	0.1126	835	-0.9453	2830.36
10	16	0.0661	1120	-1.0847	1601.36
11	17	0.0686	3700	-1.0651	1798.27
12	18	0.0826	3860	-1.0103	2531.98
13	19	0.1020	1245	-1.0154	1408.66
14	20	0.1041	1690	-0.9182	2028.42
15	21	0.1133	830	-0.8223	2859.70
16	22	0.0902	610	-1.0600	1601.42
17	23	0.0776	2050	-1.0177	1930.16
18	24	0.1009	2320	-0.9018	2995.37

Table 39
Estimated Results of Example 6

```
##################################################################
METHOD 5(PARALLEL DATA AS SUPPLEMENTAL VARIABLE)
##################################################################
DATA :
           SAMPLE GROUP
 ID  P       SIZE   NO  SUPPLEMENTAL AND AUXILIARY INFORMATION
-------------------------------------------------------------------
 13 0.0600  800     1       -1.034   1511.230
 14 0.0907  915     1       -0.992   2167.140
 15 0.1126  835     1       -0.945   2830.360
 16 0.0661 1120     1       -1.085   1601.360
 17 0.0686 3700     1       -1.065   1798.270
 18 0.0826 3860     1       -1.010   2531.980
 19 0.1020 1245     1       -1.015   1408.660
 20 0.1041 1690     1       -0.918   2028.420
 21 0.1133  830     1       -0.822   2859.700
 22 0.0902  610     1       -1.060   1601.420
 23 0.0776 2050     1       -1.018   1930.160
 24 0.1009 2320     1       -0.902   2995.370

ESTIMATED RESULTS :

           PROPORTION                       CONVERTED PARAMETER
 ID  ACTUAL       ESTIMATED            ACTUAL                ESTIMATED
------------------------------------------------------------------
 13  0.0600000  0.0656916          0.000039703          0.000043469
 14  0.0907000  0.0896818          0.000041852          0.000041383
 15  0.1126000  0.1084551          0.000039783          0.000038318
 16  0.0661000  0.0653296          0.000041277          0.000040796
 17  0.0686000  0.0687151          0.000038148          0.000038212
 18  0.0826000  0.0827749          0.000032623          0.000032692
 19  0.1020000  0.0944717          0.000072409          0.000067065
 20  0.1041000  0.1059983          0.000051321          0.000052257
 21  0.1133000  0.1226162          0.000039620          0.000042877
 22  0.0902000  0.0826330          0.000056325          0.000051600
 23  0.0776000  0.0788177          0.000040204          0.000040835
 24  0.1009000  0.1056841          0.000033685          0.000035282

 SHRINKING FACTOR                 =  0.6433
%-IMPROVEMENT OVER USUAL ESTIMATE = 30.7016 %
```

Example 5: Using Method 2 with Two Groups - This example is the same as Example 2 except that group members are different. In this example, 12 categories are arranged into the following two groups:

Group 1 - Categories 13, 16, 17, 18, 23
Group 2 - Categories 14, 15, 19, 20, 21, 22, and 24.

Thus, NMG(1) = 5 and NMG(2) = 7. The data input for this example is presented in Table 36.

As one can see from Table 36, the second portion of the data input for this example was arranged according to the order of group members specified above. Note that the order of members within a group has no effect on the estimation procedure. The user-supplied subroutine is the same as the one in Example 1. The run results are presented in Table 37.

Example 6: Using Parallel Data as a Supplemental Variable - We mentioned at the beginning of the section that $\text{Sin}^{-1}(2V_i - 1)$ with V_i defined in Eq. (155) can serve as the supplemental variable for using Method 5 to compute initial parameter estimates. In this case, AUX(i, 1) = $\text{Sin}^{-1}(2V_{i+12} - 1)$ and BUX (i, 1) = t_{i+12} for i = 1, ... , 12. The data input for this example is presented in Table 38. Note that the fourth and fifth numbers in each of Cards 7 through 18 are, respectively, $\text{Sin}^{-1}(2V_i - 1)$ and t_i. The user-supplied subroutine EQN is the same as the one used in Example 1. The run results are presented in Table 39.

We note that the above procedure can be repeated with X_i in Eq. (155) replaced by an estimate of p_i.

8) Program Listing

```
      PROGRAM PROPORTION(TAPE1=INPUT,TAPE2=OUTPUT)
      DIMENSION NSAMP(500),X(500),U(500),XT(500),NAUX(500),
     1UT(500),XEST(500),NMG(500),IG(500),ID(500),NCOV(500),
     2AUX(500,5),B(6),BUX(500,5),METHOD(500),DESCPT(72),
     3W1(500),W2(500)
C
C     THIS PROGRAM IS FOR SIMULTANEOUSLY ESTIMATING NP
C     PROPORTIONS USING STEIN-LIKE ESTIMATION PROCEDURES.
C
    1 READ(1,3,END=9999)NP,NGROUP,NTITLE
      READ(1,3) (NMG(I),I=1,NGROUP)
      READ(1,3) (METHOD(I),I=1,NGROUP)
      READ(1,3) (NAUX(I),I=1,NGROUP)
      READ(1,3) (NCOV(I),I=1,NGROUP)
      READ(1,10)(DESCPT(I),I=1,70)
      WRITE(2,15)(DESCPT(I),I=1,70)
      WRITE(2,30)
      SSR=0.
      NDF=0
      K=0
      DO 1000 I=1,NGROUP
      NCOUNT=K
```

```
      NPT=NMG(I)
      WB1=O.
      WB2=O.
      IF(METHOD(I) .LE. 4)NAUX(I)=1
      IF(METHOD(I) .EQ. 2)NAUX(I)=0
      DO 500 J=1,NPT
      K=NCOUNT+J
      IG(K)=I
      NC=NCOV(I)
      NA=NAUX(I)
      IF(METHOD(I) .NE. 2)GO TO 110
      IF(NC .GT. 0) GO TO 100
      READ(1,60)ID(K),X(K),NSAMP(K)
      WRITE(2,300)ID(K),X(K),NSAMP(K),IG(K)
      GO TO 150
  100 READ(1,60)ID(K),X(K),NSAMP(K),(BUX(K,M),M=1,NC)
      WRITE(2,300)ID(K),X(K),NSAMP(K),IG(K),(BUX(K,M),M=1,NC)
      GO TO 150
  110 IF(NC .GT. 0)GO TO 130
      READ(1,60)ID(K),X(K),NSAMP(K),(AUX(J,M),M=1,NA)
      IF(NC .LT. 0)GO TO 120
      DO 115 M=1,NA
  115 BUX(K,M)=AUX(J,M)
  120 CONTINUE
      WRITE(2,300)ID(K),X(K),NSAMP(K),IG(K),(AUX(J,M),M=1,NA)
      GO TO 150
  130 READ(1,60)ID(K),X(K),NSAMP(K),(AUX(J,M),M=1,NA),
     1(BUX(K,M),M=1,NC)
      WRITE(2,300)ID(K),X(K),NSAMP(K),IG(K),(AUX(J,M),M=1,NA),
     1(BUX(K,M),M=1,NC)
  150 A=NSAMP(K)
      W1(J)=SQRT(A+0.5)
      BB=A/(A+0.75)
      W2(J)=ARSIN(BB*(X(K)*2-1))
      U(J)=W1(J)*W2(J)
      XT(K)=U(J)
      IF(METHOD(I) .LE. 6)U(J)=W2(J)
      IF(METHOD(I) .GE. 5)GO TO 500
      IF(METHOD(I) .NE. 1)GO TO 350
      UT(K)=W1(J)*ARSIN(BB*(AUX(J,1)*2-1))
      SSR=SSR+(XT(K)-UT(K))**2
      GO TO 500
  350 CONTINUE
      WB1=WB1+XT(K)
      WB2=WB2+W1(J)
      IF(METHOD(I) .EQ. 4)WB1=WB1+AUX(J,1)*XT(K)
  500 CONTINUE
      IF(METHOD(I) .EQ. 1)GO TO 1000
      IF(METHOD(I) .GE. 5)GO TO 900
      IF(METHOD(I) .EQ. 2)WB1=WB1/WB2
      DO 800 J=1,NPT
      K=NCOUNT+J
      UT(K)=WB1*W1(J)
      IF(METHOD(I) .EQ. 3)UT(K)=WB1*AUX(J,1)
      IF(METHOD(I) .EQ. 4)UT(K)=WB1
  800 SSR=SSR+(XT(K)-UT(K))**2
      NDF=NDF+1
      GO TO 1000
  900 CONTINUE
      NFORCE=0
      IF(METHOD(I) .EQ. 6 .OR. METHOD(I) .EQ. 8)NFORCE=1
      NIND=NAUX(I)
      NDF=NDF+NIND+NFORCE
      CALL LSE(NFORCE,NPT,NIND,U,AUX,B,XEST,SERQR)
      DO 950 J=1,NPT
      K=NCOUNT+J
      UT(K)=XEST(J)
      IF(METHOD(I) .LE. 6)UT(K)=UT(K)*W1(J)
  950 SSR=SSR+(XT(K)-UT(K))**2
 1000 CONTINUE
      NTRACE=NP-NDF
      IF(NTRACE .GT. 2)GO TO 2000
      WRITE(2,1500)NTRACE
```

```
      GO TO 9999
 2000 CONTINUE
      FACTOR=1.-(NTRACE-2)/SSR
      IF(FACTOR .LT. O.)FACTOR=O.
      DO 2100 J=1,NP
      EST=UT(J)+FACTOR*(XT(J)-UT(J))
      A=NSAMP(J)
      BB=SQRT(A+0.5)
      A=A/(A+0.75)
      XEST(J)=(SIN(EST/BB)/A+1.)/2.
 2100 CONTINUE
      WRITE(2,2200)
      N2=O
      DO 4000 I=1,NGROUP
      N1=N2+1
      N2=N1+NMG(I)-1
      IF(NCOV(I) .LT. O)GO TO 3400
      CALL EQN(NP,ID,IG,NSAMP,X,XEST,BUX,N1,N2,XT,UT)
      DO 3300 J=N1,N2
      WRITE(2,3200)ID(J),X(J),XEST(J),XT(J),UT(J)
 3300 CONTINUE
      GO TO 4000
 3400 DO 3500 J=N1,N2
      WRITE(2,3200)ID(J),X(J),XEST(J)
 3500 CONTINUE
 4000 CONTINUE
      SSR=SSR-NTRACE
      IF(SSR .LE. O.)SSR=O.
      CALL EXPECT(NTRACE,SSR,EYY)
      PIMPRO=EYY*(NTRACE-2)**2/NP*100
      WRITE(2,6000)FACTOR,PIMPRO
      GO TO 1
 9999 CONTINUE
      STOP
 3    FORMAT(20I3)
 10   FORMAT(1X,70A1)
 15   FORMAT(/,1X,44H############################################,
     126H##########################,/,1X,70A1,/,1X,13H#############,
     257H#########################################################)
 30   FORMAT(/,1X,4HDATA,/,11X,12HSAMPLE GROUP,/,2X,5HID   P,5X,
     149HSIZE    NO   SUPPLEMENTAL AND AUXILIARY INFORMATION,/,1X,
     250H--------------------------------------------------,
     320H--------------------)
 60   FORMAT(I3,F8.4,I6,5F11.4)
 300  FORMAT(1X,I3,F7.4,I5,1X,I4,1X,5F10.3)
 1500 FORMAT(//,1X,29H--- ERROR - ERROR - ERROR ---,/,2X,
     19HTRACE(P)=,I4,/,2X,28HTHE TRACE OF P SHOULD BE AT ,
     27HLEAST 2,/,2X,34HCHECK TO SEE WHETHER YOU HAVE USED,
     3/,2X,44HTOO MANY GROUPS OR/AND TOO MANY VARIABLES IN,
     4/,2X,20HTHE REGRESSION LINES,/)
 2200 FORMAT(/,1X,17HESTIMATED RESULTS,//,11X,10HPROPORTION,20X,
     119HCONVERTED PARAMETER,/,2X,24HID   ACTUAL      ESTIMATED,8X,
     26HACTUAL,12X,9HESTIMATED,/,1X,26H--------------------------,
     340H----------------------------------------)
 3200 FORMAT(1X,I3,2F11.7,2F20.9)
 6000 FORMAT(/,1X,35H SHRINKING FACTOR                  =,F8.4,/,
     11X,35H%-IMPROVEMENT OVER USUAL ESTIMATE =,F8.4,2H %,//)
      END

      SUBROUTINE LSE(MODEL,NPT,NIND,YDATA,XDATA,B,YEST,SEROR)
      DIMENSION YDATA(500),XDATA(500,5),B(6),X(500,6),
     1YEST(500),XX(6,6),XXIXT(6,500)
C
C     SUBROUTINE FOR COMPUTING THE LEAST SQUARES ESTIMATE
C     OF B IN THE LINEAR MODEL,
C
C             YDATA = XDATA * B        , IF MODEL = O
C     AND
C             YDATA = (I,XDATA) * B    , IF MODEL = 1
C
C     WHERE YDATA IS THE THE (NPT BY 1) VECTOR OF DEPENDENT
C     OBSERVATIONS, XDATA IS THE (NPT BY NIND) MATRIX OF
```

```
C     INDEPENDENT OBSERVATIONS, I IS THE COLUMN VECTOR WITH
C     EVERY ELEMENT EQUAL TO ONE, * STANDS FOR THE PRODUCT
C     OF TWO MATRICES, NIND IS THE NUMBER OF INDEPENDENT
C     VARIABLES AND B IS THE VECTOR OF PARAMETERS TO BE
C     ESTIMATED.
C
      IF(MODEL .EQ. 0)GO TO 100
      DO 10 I=1,NPT
 10   X(I,1)=1.
      MM=NIND+1
      DO 20 I=2,MM
      II=I-1
      DO 30 J=1,NPT
 30   X(J,I)=XDATA(J,II)
 20   CONTINUE
      GO TO 200
 100  MM=NIND
      DO 120 I=1,MM
      DO 130 J=1,NPT
 130  X(J,I)=XDATA(J,I)
 120  CONTINUE
 200  CONTINUE
      DO 230 I=1,MM
      DO 240 J=1,MM
      XX(I,J)=0.
      DO 250 K=1,NPT
 250  XX(I,J)=XX(I,J)+X(K,I)*X(K,J)
 240  CONTINUE
 230  CONTINUE
      CALL MTXINV(MM,XX,XX)
      DO 300 I=1,MM
      DO 310 J=1,NPT
      XXIXT(I,J)=0.
      DO 320 K=1,MM
 320  XXIXT(I,J)=XXIXT(I,J)+XX(I,K)*X(J,K)
 310  CONTINUE
 300  CONTINUE
      DO 400 I=1,MM
      B(I)=0.
      DO 410 J=1,NPT
 410  B(I)=B(I)+XXIXT(I,J)*YDATA(J)
 400  CONTINUE
      SEROR=0.
      DO 500 I=1,NPT
      YEST(I)=0.
      DO 510 J=1,MM
 510  YEST(I)=YEST(I)+X(I,J)*B(J)
      SEROR=SEROR+(YDATA(I)-YEST(I))**2
 500  CONTINUE
      SEROR=SQRT(SEROR/(NPT-MM))
      RETURN
      END

      SUBROUTINE MTXINV(NSIZE,W,WINV)
      DIMENSION ARRAY(6,6),WINV(6,6),W1(6,2),W(6,6)
C
C     SUBROUTINE FOR FINDING THE INVERSE OF AN (NSIZE BY NSIZE)
C     SQUARE MATRIX W BY USING THE PARTITION METHOD. WINV IS THE
C     INVERSE MATRIX OF W.
C
      DO 5 I=1,NSIZE
      DO 5 J=1,NSIZE
 5    ARRAY(I,J)=W(I,J)
      IF(NSIZE .GT. 1) GO TO 10
      WINV(1,1)=1./ARRAY(1,1)
      RETURN
 10   CONTINUE
      MSIZE=NSIZE-1
      DO 15 II=1,MSIZE
      J=II+1
      DO 16 KK=J,NSIZE
      DO 17 M=1,NSIZE
```

```
      W1(M,1)=W(M,II)
17    W1(M,2)=W(M,KK)
      DET=W1(1,1)*W1(2,2)-W1(1,2)*W1(2,1)
      IF(DET .EQ. O.) GO TO 16
      IF(II .EQ. 1 .AND. KK .EQ. 2) GO TO 19
      DO  18 K=1,NSIZE
      ARRAY(K,1)=W(K,II)
      ARRAY(K,2)=W(K,KK)
      ARRAY(K,II)=W(K,1)
18    ARRAY(K,KK)=W(K,2)
      GO TO 19
16    CONTINUE
15    CONTINUE
19    CONTINUE
      WINV(1,1)=ARRAY(2,2)/DET
      WINV(2,2)=ARRAY(1,1)/DET
      WINV(1,2)=-ARRAY(1,2)/DET
      WINV(2,1)=-ARRAY(2,1)/DET
      IF(NSIZE .EQ. 2) GO TO 100
      DO 20 I=3,NSIZE
      K=I-1
      DO 21 J=1,K
      W1(J,1)=O.
      W1(J,2)=O.
      DO 22 M=1,K
      W1(J,1)=W1(J,1)+WINV(J,M)*ARRAY(M,I)
22    W1(J,2)=W1(J,2)+ARRAY(I,M)*WINV(M,J)
21    CONTINUE
      ELTA=ARRAY(I,I)
      DO 23 J=1,K
23    ELTA=ELTA-ARRAY(I,J)*W1(J,1)
      WINV(I,I)=1./ELTA
      DO 24 J=1,K
      WINV(J,I)=-W1(J,1)/ELTA
      WINV(I,J)=-W1(J,2)/ELTA
      DO 24 M=1,K
24    WINV(J,M)=WINV(J,M)+W1(J,1)*W1(M,2)/ELTA
20    CONTINUE
100   CONTINUE
      IF(II .EQ. 1 .AND. KK .EQ. 2) GO TO 888
      DO 401 J=1,NSIZE
      W1(J,1)=WINV(1,J)
401   W1(J,2)=WINV(2,J)
      DO 402 J=1,NSIZE
      WINV(1,J)=WINV(II,J)
402   WINV(2,J)=WINV(KK,J)
      DO 403 J=1,NSIZE
      WINV(II,J)=W1(J,1)
403   WINV(KK,J)=W1(J,2)
888   CONTINUE
      RETURN
      END

      SUBROUTINE EXPECT(NP,THETA,EYY)
      DOUBLE PRECISION P1,A,EY,P2
C
C     THIS SUBROUTINE COMPUTES THE EXPECTATION OF 1/Y, WHERE Y IS A
C     NONCENTRAL CHI-SQUARE WITH NP DEGREES OF FREEDOM AND NONCENTRALITY
C     PARAMETER THETA. THIS EXPECTATION IS THE SAME AS THE EXPECTATION
C     OF 1/(NP-2+2W), WHERE W IS A POISSON WITH PARAMETER (THETA/2).
C
      ERROR=0.00001
      ERR=ERROR/NP
      A=THETA/2
      P1=DEXP(-A)
      EY=P1/(NP-2)
      P2=P1
      K=0
10    K=K+1
      P1=P1*A/K
      P2=P2+P1
      EY=EY+P1/(NP-2+K*2)
      CHECK=1.-P2
```

```
      IF(CHECK.GE.ERR) GO TO 10
      EYY=EY
      RETURN
      END

      SUBROUTINE EQN(NP,ID,IG,NSAMP,X,XEST,BUX,N1,N2,W2,W3)
      DIMENSION XEST(500),BUX(500,5),W3(500),W2(500),
     1ID(500),IG(500),X(500),NSAMP(500)
C
C     SUBROUTINE FOR CONVERTING THE ACTUAL(USUAL) AND ESTIMATED
C     PROPORTIONS TO OTHER PARAMETERS THROUGH THE FUNCTION F.
C     X(I) AND XEST(I) ARE RESPECTIVELY THE ACTUAL AND ESTIMATEO
C     PROPORTIONS. THE FUNCTIONAL FORM OF F MUST BE SPECIFIED
C     BETWEEN TWO STATEMENTS : DO 100 I=N1,N2 & 100 CONTINUE.
C     W2(I) ANO W3(I) ARE RESPECTIVELY THE TRANSFORMEO PARAMETERS
C     OF X(I) ANO XEST(I). THAT IS, W2(I) = F ( X(I) ) ANO
C     W3(I) = F ( XEST(I) ).
C
      DO 100 I=N1,N2
      W2(I)=X(I)/BUX(I,1)
      W3(I)=XEST(I)/BUX(I,1)
  100 CONTINUE
      RETURN
      ENO
```

VII
A COMPUTER PROGRAM FOR SIMULTANEOUSLY ESTIMATING INDEPENDENT MEANS BY USING STEIN-LIKE ESTIMATION PROCEDURES

1) The Basic Data

The essential data for estimating population means are the independent observations $\overline{X}_1, \ldots, \overline{X}_k$ and S^2. $\overline{X}_i$ is the average of n_i observations sampled from the i-th population with mean θ_i and unknown variance σ^2. That is, k populations have a common variance. Denote X_{ij} to be the j-th observation from the i-th population. Then,

$$\overline{X}_i = \frac{1}{n_i} \sum_{j=1}^{k} X_{ij} \qquad (158)$$

We assume that either the population is normally distributed or the sample size is large enough to guarantee the normality of sample averages by the Central Limit Theorem. Therefore, $\overline{X} = (\overline{X}_1, \ldots, \overline{X}_k)'$ is a k-variate normal with mean vector $\theta = (\theta_1, \ldots, \theta_k)'$ and covariance matrix $\sigma^2 D$, where the superscript $'$ stands for the transpose of a vector or matrix and D is a diagonal matrix with the (i, i)-th element $1/n_i$. In this case, $\overline{X}$ is the usual estimate of θ.

S^2 is an unbiased estimate of σ^2 such that $m\,S^2/\sigma^2$ is chi-square with m degrees of freedom. This estimate is obtained from either the past experiments or the current data. In the latter case, S^2 is the unbiased sample variance defined as

$$S^2 = \frac{1}{m} \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \overline{X}_i)^2 \qquad (159)$$

and

$$m = n_1 + \ldots + n_k - k \qquad (160)$$

The other essential data is the information for computing initial estimates of mean parameters. This is termed the 'supplemental' information for discussion purposes. Five methods for using the supplemental information to compute initial estimates of mean parameters are discussed in Section 3.

2) What the Program Does

This program is written, based on theoretical results developed in Part III, for combining sample averages and supplemental information to estimate population means, $\theta_1, \ldots, \theta_k$.

The program first computes the initial estimate, U_i, of θ_i by the method chosen. Denote $U = (U_1, \ldots, U_k)'$. In this step, the trace of the idempotent matrix P satisfying the following equation

$$\overline{X} - U = P\,\overline{X} \tag{161}$$

is also computed. The second step is to compute the shrinking factor c defined as

$$c = \left[\text{Trace }(P) - 2\right] \frac{m}{m+2} \cdot \frac{S^2}{\sum_{i=1}^{k} n_i (\overline{X}_i - U_i)^2} \tag{162}$$

A slightly better procedure is to set c equal to 1 if it is greater than 1. The third step is to compute the final estimate of θ_i defined as

$$\hat{\overline{X}}_i = U_i + (1 - c)\,(\overline{X}_i - U_i),\; i = 1, \ldots, k \tag{163}$$

This program also computes the estimated percentage improvement of the above procedure over the usual one.

3) Methods for Computing Initial Estimates

The key to obtaining good estimates of population means is to provide good initial estimates in the sense that N is high and ζ is low, where

$$N = \text{Trace }(P) \tag{164}$$

and

$$\zeta = \sum_{i=1}^{k} n_i (U_i - \theta_i)^2/\sigma^2 \tag{165}$$

We have shown in Part III that, if P is a symmetrical idempotent matrix, N and ζ are, respectively, the number of degrees of freedom and the noncentrality parameter of a noncentral chi-square distribution. For this case, the maximal percentage improvement that can be achieved is $100\, m(N - 2)/\{(m + 2)k\}$. Five methods for computing initial estimates of parameters are built in this program. These are:

Method 1: Initial Estimates are Given - Based on past experiments or parallel studies, we estimate or guess θ_i to be U_i, $i = 1, \ldots, k$. In this case, we treat $U_i - \theta_i$ as the parameter to be estimated. Consequently, the idempotent matrix P satisfying Eq. (161) is the identity matrix. Thus, N = Trace (P) = k. The estimation accuracy is the degree of closeness of U to θ. The final estimate $\hat{\bar{X}}_i$ always lies between U_i and $\bar{X}_i$. When initial estimates are excellent, i.e., U is very close to θ, the final estimate of θ is U. However, if initial estimates are poor, i.e., U is quite distant from θ, the final estimate of θ will be very close to the usual estimate $\bar{X}$.

This method is used only when at least three parameters are to be estimated. When initial estimates are reliably close to the true means, this method will produce good final estimates of population means.

Method 2: Weighted Average (I) - When the population means are almost homogeneous, we may consider that $\theta_i = \bar{\theta}$ for every i. We therefore use the unbiased estimate of $\bar{\theta}$ as the initial estimate of every population mean. That is,

$$U_i = \sum_{j=1}^{k} n_j \bar{X}_j \Big/ \sum_{j=1}^{k} n_j, \quad i = 1, \ldots, k \tag{166}$$

For this method, N = Trace (P) = k - 1. The loss of one degree of freedom is due to the estimation of the unknown parameter $\bar{\theta}$. Note that the above U_i can be written as

$$U_i = \sum_{j=1}^{k} w_j \bar{X}_j, \quad i = 1, \ldots, k \tag{167}$$

with

$$w_i = n_i \Big/ \sum_{j=1}^{k} n_j \tag{168}$$

It is obvious that

$$w_1 + w_2 + \ldots + w_k = 1 \tag{169}$$

This method is used only when the number of population means to be estimated is at least 4. This method will produce good estimates of population means if these means are homogeneous.

Method 3: Weighted Average (II) - For a given set of numbers, $w_1, \ldots, w_k$, satisfying Eq. (169), we take U_i defined in Eq. (167) as the initial estimate of θ_i. For this case, N = Trace (P) = k - 1. We note that this is the same form used in Method 2. The only difference is that w_i in Method 2 is computed from sample sizes, but is given in this method.

This method is used only when the number of population means to be estimated is at least 4. The estimation results will be good if population means are homogeneous.

We remark that w_i can be interpreted as the weight assigned to $\overline{X}_i$ for estimating the common parameter $\overline{\theta}$. If one has reason to doubt the reliability of a particular observation, say $\overline{X}_1$, zero weight may be assigned to $\overline{X}_1$. That is, $w_1 = 0$ and $w_2 + \ldots + w_k = 1$.

Method 4: Least Squares Estimate (I) - Suppose that the supplemental data $(t_{i1}, \ldots, t_{im})$ are available and satisfy the following equation.

$$\theta_i = b_{i1} t_{i1} + \ldots + b_{im} t_{im}, \quad i = 1, \ldots, k \tag{170}$$

If for every j, $b_{1j}, \ldots$, and b_{kj} are near an unknown common parameter b_j, Eq. (170) can be rewritten as

$$\theta_i = b_1 t_{i1} + \ldots + b_m t_{im}, \quad i = 1, \ldots, k \tag{171}$$

Denote $(\hat{b}_1, \ldots, \hat{b}_m)$ to be the least squares estimate of $(b_1, \ldots, b_m)$. We then take the initial estimate of θ_i to be

$$U_i = \hat{b}_1 t_{i1} + \ldots + \hat{b}_m t_{im}, \quad i = 1, \ldots, k \tag{172}$$

In this case, N = Trace (P) = k - m. The loss of m degrees of freedom is due to the least squares estimates of m linear parameters. We note that U can always be improved by increasing the number of supplemental variables, i.e., by increasing m. However, this decreases N which is an undesirable property as previously mentioned. In general, m should be kept small relative to the number of parameters k.

This method is used only when the number of parameters to be estimated is at least m + 3. This method will produce good estimates of parameters if Eq. (171) holds approximately and the number of supplemental variables is small relative to the number of parameters to be estimated.

In practice, one may use this method if the following linear relationship holds approximately.

$$\overline{X}_i = b_1 t_{i1} + \ldots + b_m t_{im}, \; i = 1, \ldots, k \tag{173}$$

Method 5: Least Squares Estimate (II) - This method is the same as Method 4 except that Eqs. (171) and (172) are, respectively, replaced by

$$\theta_i = b_0 + b_1 t_{i1} + \ldots + b_m t_{im}, \; i = 1, \ldots, k \tag{174}$$

and

$$U_i = \hat{b}_0 + \hat{b}_1 t_{i1} + \ldots + \hat{b}_m t_{im}, \; i = 1, \ldots, k \tag{175}$$

In this case, N = Trace (P) = k - m - 1 because of the extra parameter b_0. Thus, this method is used only when the number of parameters to be estimated is at least m + 4. In practice, one may use this method if the following linear relationship holds approximately.

$$\overline{X}_i = b_0 + b_1 t_{i1} + \ldots + b_m t_{im}, \; i = 1, \ldots, k \tag{176}$$

The k parameters may be arranged into many groups to which different methods are applied to minimize the noncentrality parameter defined in Eq. (165). For example, we may use Method 1 for a group of k_1 parameters, Method 4 for a group of k_2 parameters, Method 2 for a group of k_3 parameters, and again, Method 2 for the group of the remaining ones. In this case, we have $N = \text{Trace}(P) = k_1 + (k_2 - m) + (k_3 - 1) + (k - k_1 - k_2 - k_3 - 1) = k - m - 2$. m is the number of supplemental variables used in Method 4. For this particular case, the number of parameters to be estimated should be at least m + 5. In general, for any method or combination of methods, the number of degrees of freedom should be at least 2. Combining the above five methods to compute initial parameter estimates will be demonstrated by examples presented in Section 6.

Every method except the first one suggests that k populations should be arranged into groups in which parameters cluster at a point or can be approximated by a linear function of supplemental variables. Physical properties of populations, past experiments and parallel studies are good sources for obtaining the proper group-method combination. Unfortunately, these sources may not be available or reliable enough in a particular problem. In this circumstance, one may examine the data as suggested in each method to choose the proper group-method combination. That is, the empirical relationship among $\overline{X}_1, \ldots, \overline{X}_k$, and supplemental variables can be used to determine the method-group combination for computing initial estimates of parameters. Since the program computes the estimated percentage improvement of each chosen method-group combination over the usual method, one may use the one that produces the maximal improvement to estimate parameters. We remark that the method-group combination generated by examining the data may not be the best one and, possibly, could be the worst one for this problem due to random variation of the data. Nevertheless, if the chosen group-method combination is used thereafter for the same problem, the above method is always better than the usual one. The worst situation is that no improvement is made. Based on our experience, qualitative properties of populations generally provide adequate information for grouping purposes.

4) Data Input

The data input of this program is arranged into two portions. The first portion is composed of five cards. These cards specify the number of groups and parameters, computational methods, and number of supplemental variables in each group. Variables used in these cards are defined below.

NP:	Number of parameters to be estimated, $1 \leq NP \leq 500$
NGROUP:	Number of groups used, $1 \leq NGROUP \leq 20$
MDEG:	Number of degrees of freedom for estimating S^2. (When this number is positive, the unbiased estimate of σ^2 is supplied. Otherwise, the program will use Eq. (159) to compute this estimate.)
VAR:	The unbiased estimate, S^2, of σ^2. (When MDEG is negative, the input value of this variable is meaningless and will be computed by the program.)
NMG(I):	Number of parameters in the I-th group, MNG(1) + . . . + NMG (NGROUP) = NP.
METHOD(I) = j:	The j-th method presented in Section 3 is used to compute initial estimates of parameters in the I-th group, $1 \leq j \leq 5$
NAUX(I):	Number of supplemental variables used to compute initial estimates of parameters in the I-th group. This variable is 1 if METHOD(I) = 1 or 3, and is 0 if the METHOD(I) = 2.
(DESCPT(I), I = 1, 70):	Title (no more than 70 letters).

The second portion is composed of NGROUP subportions or groups. The input format of this portion is determined by the input value of MDEG. This is explained below.

When MDEG is positive, the I-th subportion is composed of NMG(I) cards. Define K = J if I = 1 and K = NMG(1) + . . . + NMG(I - 1) + J if I > 1. Then, the K-th card of the second portion contains essential and supplemental data for estimating the J-th parameter of the I-th group. We note that the J-th parameter of the I-th group is the parameter of the K-th population. The data input for estimating this parameter is as follows:

ID(K), NSAMP(K), X(K), (AUX(J, M), M = 1, NAUX(I))

Variables used in this card are defined below:

ID(K): Identification number of the K-th population such as location number and year, etc. The K-th parameter is the J-th parameter of the I-th group. This number has no effect on the estimation procedure.

NSAMP(K): Number of observations from the K-th population.

X(K): The K-th sample average (the usual estimate of the K-th population mean).

AUX(J, M): The M-th supplemental variable for the K-th parameter.

When MDEG is negative, the data input for estimating the K-th population mean is

ID(K), NSAMP(K), (AUX(J, M), M = 1, NAUX(I))
(XT(M), M = 1, NSAMP(K))

XT(M) is the M-th observation from the K-th population. These observations are arranged into cards such that each card, except the last one, has 10 observations. In this case, X(1), . . . , X(NP), VAR and MDEG will be computed according to Eqs. (158) through (160).

The input deck is presented in Table 40 for the case that MDEG is positive, and in Table 41 for the case that MDEG is negative. The input deck is also diagrammed in Figure 7 to show the format and logic used. The user can follow this diagram to change, if needed, read statements and formats to fit a particular problem.

5) Limitations of the Program and How to Make Necessary Changes

This program was designed to handle problems where the number of parameters, NP, does not exceed 500. This number can be easily increased

Table 40
Input Deck Used When An Unbiased Estimate of The Population Variance Is Available
(The Input Value Of MDEG Is Positive)

Card Number	Variables Used in Each Card	Remarks
1	NP , NGROUP	
2	(NMG(I), I=1,...,NGRDUP)	
3	(METHOD(I), I=1,...,NGROUP)	* The Data Input of The First Portion
4	(NAUX(I),I=1,NGROUP)	
5	(DESPT(I),I=1,70)	
6	ID(1),NSAMP(1),X(1),(AUX(1,M),M=1,NA)	* L=NMG(1) * NA=NAUX(1) * L Cards for The First Group
L+5	ID(L),NSAMP(L),X(L),(AUX(L,M),M=1,NA)	
⋮	⋮	⋮
		* K=NMG(1)+..+NMG(I-1) +J ; NA=NAUX(I)
K+5	ID(K),NSAMP(K),X(K),(AUX(J,M),M=1,NA) ←	This card is for the k-th parameter which is the J-th parameter of the I-th group
⋮	⋮	⋮
N+5	ID(N),NSAMP(N),X(N),(AUX(1,M),M=1,NA)	* L=NMG(NGROUP) * NA=NAUX(NGROUP) * N=NP-L+1 * L Cards for The Last * Group of The Second Portion
NP+5	ID(NP),NSAMP(NP),X(NP),(AUX(L,M),M=1,NA)	

Table 41
Input Deck Used When An Unbiased Estimate of The Population Variance Is Not Available
(The Input Value Of MDEG Is Negative)

Card Number	Variables Used in Each Card	Remarks
1	NP,NGROUP,MDEG,VAR	* The Data Input of The First Portion
2	(NMG(I),I=1,...,NGROUP)	
3	(METHDD(I),I=1,...,NGROUP)	
4	(NAUX(I),I=1,NGROUP)	
5	(DESPT(I),I=1,70)	
6		* Cards for The First Group of The Second Portion
⋮	⋮	⋮
		* NA=NAUX(I) * L=(NSAMP(K)-1)/10+1 * K=NMG(1)+..+NMG(I-1)+J
	ID(K),NSAMP(K),X(K),(AUX(J,M),M=1,NA)	← There are (L+1) cards for the K-th parameter which is the J-th parameter of the I-the group. The last L cards with 10 numbers each, except the last one.are sample observations. ←
	XT(1), ..., XT(10)	
	XT(11), ..., XT(20)	
	. . . . , XT(NSAMP(K))	
⋮	⋮	⋮
		* Cards for the Last Group of The Second Portion

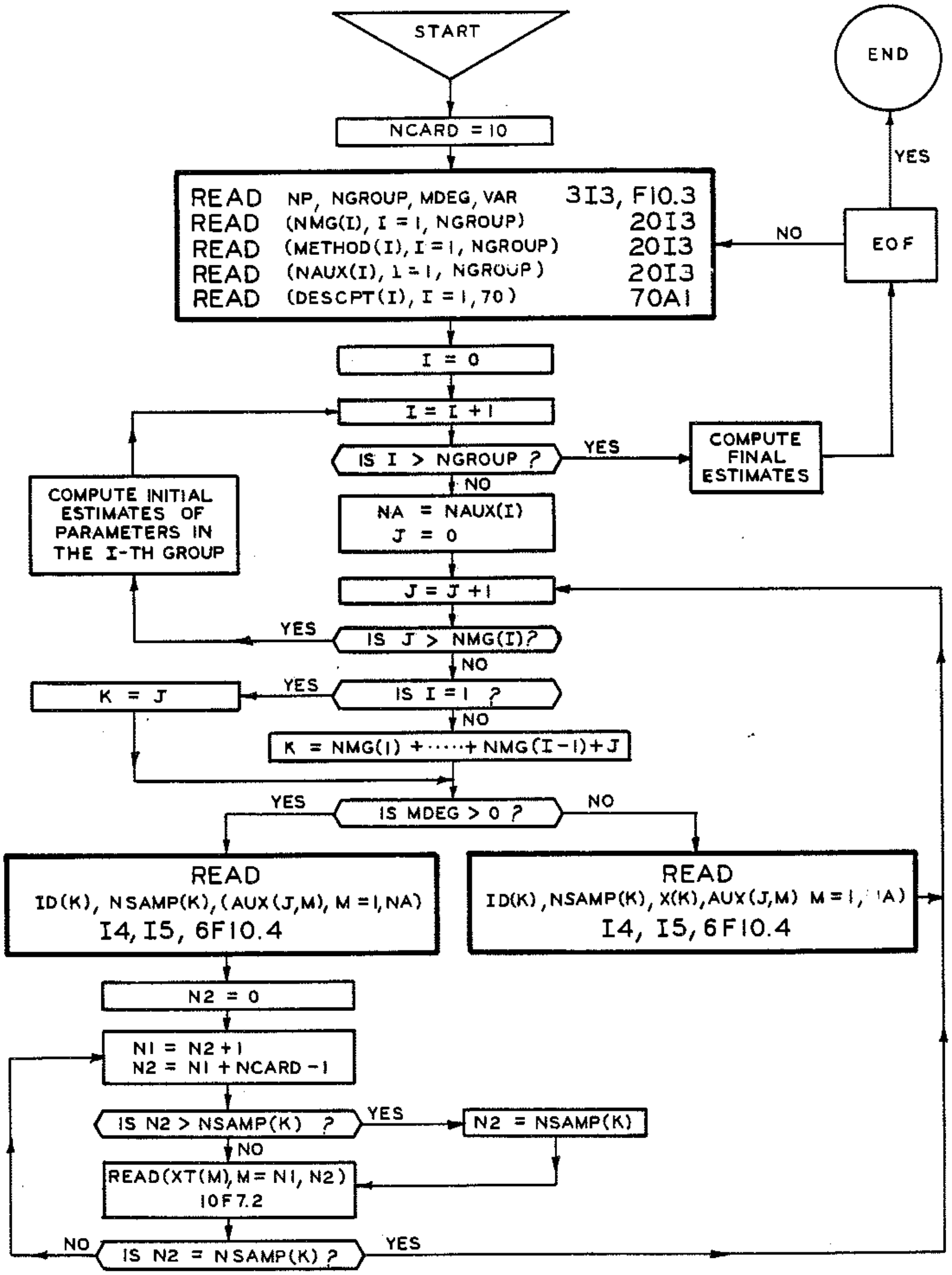

Figure 7. Flow of control for data input.

to any desired number. However, one must change the dimensions of variables that appear on the DIMENSION STATEMENTS of the main program and subroutines LSE, i.e., replace every 500 by the desired number.

When the input value of MDEG is negative, the individual observations sampled from each population are the input data. The number of observations for each population is limited to 2000. This is sufficiently large for this type of estimation problem. However, one can increase this number to any desired number by changing the dimension of the input variable, XT, which can be located in the DIMENSION STATEMENT of the main program. These observations are arranged into cards such that each card, except the last one, has 10 observations. One can easily reset this number through the first statement, NCARD = 10, of the main program. Of course, one would also have to properly change the input format (FORMAT 61) to cope with the corresponding change.

The number of groups is limited to 20. This number is large enough for most practical problems. However, one can increase this number to any desired number not exceeding NP by properly changing the READ format of the second input card (FORMAT 3).

The total number of supplemental variables is limited to 5. This number should be large enough for the practical applications. The user can change this number to any desired number, say N, by the following two steps:

a) Change the READ format of the second portion of the data input (FORMAT 60).

b) Change dimensions of variables that appear on the DIMENSION STATEMENTS of the main program and subroutines LSE and MTXINV, i.e., replace 5 and 6 by N and N + 1, respectively.

6) Examples

For an isolated commercial vehicle passing through a checkpoint, the peak noise level generated by this vehicle was measured, its traveling speed and number of axles were also observed. Two-hundred twenty-six vehicles were observed. We are interested in estimating peak noise levels of various vehicle types (classified according to number of axles) traveling at various speeds. Initial investigation of this data set reveals that:

a) The variance of the peak noise level is constant over vehicle type and speed.

b) The peak noise level can be well approximated by a linear function of speed and number of axles. This relationship will be used to obtain initial estimates of mean noise levels.

We now provide three examples using this set of data to show how to improve usual estimates of mean noise levels.

Example 1: Using Method 2 with Ten Groups - We arrange the full data set into 10 groups according to the number of axles. Every vehicle with (i + 1) axles belongs to the i-th group, i = 1, . . . , 10. Group averages are then used as initial estimates of mean noise levels. Due to the length of the data set, we shall only present the partial data in Table 42 to show the input format. The first card sets NP = 226, NGROUP = 10, MDEG = -1,

Table 42
Data Input of Example 1

```
                                        Column  Number
1234567890123456789012345678901234567890123456789012345678901234567890

226  10  -1        0.000
 42  28  28  40  20  11  12  10   7  28
  2   2   2   2   2   2   2   2   2   2
  0   0   0   0   0   0   0   0   0   0
 WEIGHTED AVERAGES ON 10 GROUPS ACCORDING TO NO OF AXLES
 |||      |
 223      1
  66.80
 |||      |.
 230      5
  69.80   70.30   74.40   67.50   70.70
 |||      |
 235     11
  70.10   72.00   73.80   72.50   67.10   76.70   77.20   72.50   79.00   80.70
  74.50
 |||      |
 353      5
  81.10   83.90   85.[illegible]0   83.00   86.30
 |||      |
 455      7
  86.70   85.90   81.90   82.50   86.60   86.60   82.90
 |||      |
 554     18
  83.10   87.30   90.80   84.50   84.60   84.60   81.90   83.60   84.80   82.90
  91.10   82.80   83.30   83.50   87.20   85.60   85.60   84.20
 |||      |
 664      1
  86.70
 720      1
  69.90
 |||      |
 841      2
  82.30   83.30
 957      2
  92.10   85.50
 |||      |
1057      3
  90.10   86.10   87.70
 |||      |
1157      7
  88.10   92.40   90.00   86.80   89.10   87.20   88.30
```

Table 43
Estimated Results of Example 1

##
WEIGHTED AVERAGES ON 10 GROUPS ACCORDING TO NO OF AXLES
##

DATA :

ID	SAMPLE SIZE	SAMPLE AVERAGE	GROUP NO	SUPPLEMENTAL VARIABLES
223	1	66.8000	1	
230	5	70.5400	1	
235	11	74.1909	1	
353	5	84.0200	2	
455	7	84.7286	3	
554	18	85.0778	4	
664	1	86.7000	5	
720	1	69.9000	6	
841	2	82.8000	7	
957	2	88.8000	8	
1057	3	87.9667	9	
1157	7	88.8429	10	

STRATIFIED SAMPLE VARIANCE = 7.0510
NO OF DEGREES OF FREEDOM = 512

ESTIMATED RESULTS :

ID	SAMPLE AVERAGE	NEW ESTIMATE
223	66.8000	68.4949
230	70.5400	71.5852
235	74.1909	74.6020
353	84.0200	83.6398
455	84.7286	84.5943
554	85.0778	84.8974
664	86.7000	86.4655
720	69.9000	72.5542
841	82.8000	83.4936
957	88.8000	88.5158
1057	87.9667	88.1259
1157	88.8429	88.3730

SHRINKING FACTOR = 0.8263
%-IMPROVEMENT OVER USUAL ESTIMATE = 16.4971 %

and VAR = 0. Consequently, there are 10 numbers in each of Cards 2 through 4. In this case, NMG(I) is the number of different speeds travelled by vehicles that have I + 1 axles. The third card instructs that the weighted average of noise levels in a group is the initial estimate of every parameter belonging to this group. Thus, no supplemental variable is required by the program. This is set in the fourth card. Since the input value of MDEG is negative, the common population variance is to be estimated from sample observations. Therefore, the data input for the K-th parameter in the second portion is as follows:

```
ID(K), NSAMP(K)
XT(1), ............, XT(10)
XT(11), ..........., XT(20)
  ......................
  ..............,XT(NSAMP(K))
```

NSAMP(K) is the number of observations sampled from the K-th population identified by ID(K). In this example, ID(K) is a four-digit number. The first and last two digits stand for the number of axles and speed, respectively, of a vehicle. These observations are arranged into many cards such that each card, except the last one, contains 10 observations. Again, the estimated results are partially presented in Table 43 for the same reason as before.

Example 2: Using Method 5 with One Group - As previously mentioned, the noise level is almost a linear function of speed and number of axles. Thus, we shall use Method 5 with speed and number of axles as supplemental variables to compute initial estimates. The input data for this method is partially presented in Table 44. The second number of the first card in this table sets NGROUP to be one. Therefore, there is only one number in each of Cards 2 through 4. The third card commands the program using Method 5 to compute initial estimates. The number of supplemental variables needed for Method 5 is set in the fourth card. The estimated results are partially presented in Table 45. Comparing this example with Example 1, we see that the percentage improvement has been increased from 16.49 to 46.71 percent. This can be credited to the better method of computing initial estimates.

Example 3: Using Method 5 with Five Groups of the Combined Data - We see from Table 42 that some sample sizes are very small. If the peak noise level generated by a vehicle is not normally distributed, the normality assumption of sample average would not be satisfied. Consequently, the above estimated results would be invalid. One way to remedy this problem is to increase the sample size. For this purpose, we list noise levels according to the rank of vehicle speeds within each vehicle type (number of axles). At 4 mph intervals, we consecutively take averages of noise levels and speeds within each vehicle type. In so doing, we have reduced 226 members (combinations of speed and number of axles) to 64 members. We

Table 44
Data Input of Example 2

```
                              Column  Number
12345678901234567890123456789012345678901234567890123456789012345678901234567890

226  1 -1      0.000
226
  5
  2
 SPEED & NO OF AXLES ARE SUPPLEMENTAL VARIABLES, ONE GROUP
 |||     |
223      1      2.0000    23.0000
  66.80
 |||     |
230      5      2.0000    30.0000
  69.80  70.30  74.40  67.50  70.70
 |||     |
235     11      2.0000    35.0000
  70.10  72.00  73 80  72.50  67.10  76.70  77.20  72.50  79.00  80.70
  74.50
 |||     |     ||||||   |||||||
353      5      3.0000    53.0000
  81.10  83.90  85.80  83.00  86.30
 |||     |     ||||||   |||||||
455      7      4.0000    55.0000
  86.70  85.90  81.90  82.50  86.60  86.60  82.90
 |||     |     ||||||   |||||||
554     18      5.0000    54.0000
  83.10  87.30  90.80  84.50  84.60  84.60  81.90  83.60  84.80  82.90
  91.10  82.80  83.30  83.50  87.20  85.60  85.60  84.20
 |||     |     ||||||   |||||||
664      1      6.0000    64.0000
  86.70
720      1      7.0000    20.0000
  69.90
 |||     |     ||||||   |||||||
841      2      8.0000    41.0000
  82.30  83.30
957      2      9.0000    57.0000
  92.10  85.50
 |||     |    |||||||   |||||||
1057     3     10.0000    57.0000
  90.10  86.10  87.70
 |||     |    |||||||   |||||||
1157     7     11.0000    57.0000
  88.10  92.40  30.00  86.80  89.10  87.20  88.30
```

Table 45
Estimated Results of Example 2

##
SPEED & NO OF AXLES ARE SUPPLEMENTAL VARIABLES, ONE GROUP
##

DATA :

ID	SAMPLE SIZE	SAMPLE AVERAGE	GROUP NO	SUPPLEMENTAL VARIABLES	
223	1	66.8000	1	2.000	23.000
230	5	70.5400	1	2.000	30.000
235	11	74.1909	1	2.000	35.000
353	5	84.0200	2	3.000	53.000
455	7	84.7286	3	4.000	55.000
554	18	85.0778	4	5.000	54.000
664	1	86.7000	5	6.000	64.000
720	1	69.9000	6	7.000	20.000
841	2	82.8000	7	8.000	41.000
957	2	88.8000	8	9.000	57.000
1057	3	87.9667	9	10.000	57.000
1157	7	88.8429	10	11.000	57.000

STRATIFIED SAMPLE VARIANCE = 7.0510
NO OF DEGREES OF FREEDOM = 512

ESTIMATED RESULTS :

ID	SAMPLE AVERAGE	NEW ESTIMATE
223	66.8000	69.6742
230	70.5400	72.5419
235	74.1909	75.1048
353	84.0200	83.0000
455	84.7286	84.0399
554	85.0778	84.5042
664	86.7000	87.0558
720	69.9000	72.9646
841	82.8000	82.8601
957	88.8000	88.4856
1057	87.9667	88.4575
1157	88.8429	89.3275

SHRINKING FACTOR = 0.5254
%-IMPROVEMENT OVER USUAL ESTIMATE = 46.7176 %

Table 46
Data Input of Example 3

```
                                  Column  Number
1234567890123456789012345678901234567890123456789012345678901234567890
 64   5512        7.051
  9   8   8  10  29
  5   5   5   5   5
  1   1   1   1   2
 METHOD 5 ON FIVE GROUPS (2,3,4,5 & 6-11 AXLES) OF THE COMBINED DATA
 201     5    72.2400    25.2000
 202    19    72.1700    30.0000
 203    60    73.7600    35.5000
 204    34    75.3410    40.0000
 205    22    75.8520    44.6000
 206    13    79.2080    50.2000
 207    32    80.2930    55.3000
 208    27    82.0480    59.3000
 209     4    83.4750    65.5000
 301     2    79.6000    25.5000
 302     9    76.7460    31.7000
 303     7    79.3290    37.0000
 304    11    78.9290    41.5000
 305     2    82.7000    46.5000
 306    14    85.1780    54.1000
 307    19    84.2080    58.0000
 308     3    83.8000    65.7000
 401     3    77.9000    27.7000
 402     5    78.9200    35.4000
 403     5    79.3200    41.8000
 404     2    81.9500    48.0000
 405    22    84.7830    53.5000
 406    19    85.3480    57.9000
 407     7    86.0160    62.3000
 408     2    89.7000    66.0000
 501     4    77.4000    27.8000
 502    18    77.3670    32.3000
 503    22    79.5330    36.7000
 504     9    80.6000    42.0000
 505     1    81.0000    45.0000
 506    39    84.5640    52.9000
 507    83    85.6910    56.9000
 508    55    85.9250    61.4000
 509     8    85.1890    66.3000
 510     2    88.8500    71.0000
 601     2    78.5000    27.0000    6.0000
 602     4    79.3250    36.3000    6.0000
 603     2    80.9500    41.0000    6.0000
 604     8    85.7740    52.8000    6.0000
 605    12    86.8500    56.8000    6.0000
 606    16    86.9230    61.4000    6.0000
 701     1    69.9000    20.0000    7.0000
 702     1    80.2000    36.0000    7.0000
 703     2    82.3000    46.0000    7.0000
 704     6    87.5670    57.3000    7.0000
 705     5    87.5200    62.0000    7.0000
 801     2    82.7500    35.5000    8.0000
 802     2    82.8000    41.0000    8.0000
 803     5    88.8000    54.0000    8.0000
 804     4    88.3000    60.3000    8.0000
 805     1    86.8000    65.0000    8.0000
 901     3    82.4670    33.3000    9.0000
 902     4    88.4250    55.3000    9.0000
 903     4    89.4250    59.8000    9.0000
1001     6    88.0350    55.3000   10.0000
1002     6    89.7330    59.8000   10.0000
1101     5    81.3820    24.0000   11.0000
1102     7    81.8710    31.4000   11.0000
1103    11    85.0460    40.3000   11.0000
1104     4    86.5250    45.0000   11.0000
1105     8    88.2380    53.3000   11.0000
1106    14    88.6910    57.9000   11.0000
1107     3    88.7670    61.7000   11.0000
1108     1    89.8000    66.0000   11.0000
```

Table 47
Estimated Results of Example 3

##
METHOD 5 ON FIVE GROUPS (2,3,4,5 & 6-10 AXLES) OF THE COMBINED DATA
##

DATA :

ID	SAMPLE SIZE	SAMPLE AVERAGE	GROUP NO	SUPPLEMENTAL VARIABLES	
201	5	72.2400	1	25.200	
202	19	72.1700	1	30.000	
203	60	73.7600	1	35.500	
204	34	75.3410	1	40.000	
205	22	75.8520	1	44.600	
206	13	79.2080	1	50.200	
207	32	80.2930	1	55.300	
208	27	82.0480	1	59.300	
209	4	83.4750	1	65.500	
301	2	79.6000	2	25.500	
302	9	76.7460	2	31.700	
303	7	79.3290	2	37.000	
304	11	78.9290	2	41.500	
305	2	82.7000	2	46.500	
306	14	85.1780	2	54.100	
307	19	84.2080	2	58.000	
308	3	83.8000	2	65.700	
401	3	77.9000	3	27.700	
402	5	78.9200	3	35.400	
403	5	79.3200	3	41.800	
404	2	81.9500	3	48.000	
405	22	84.7830	3	53.500	
406	19	85.3480	3	57.900	
407	7	86.0160	3	62.300	
408	2	89.7000	3	66.000	
501	4	77.4000	4	27.800	
502	18	77.3670	4	32.300	
503	22	79.5330	4	36.700	
504	9	80.6000	4	42.000	
505	1	81.0000	4	45.000	
506	39	84.5640	4	52.900	
507	83	85.6910	4	56.900	
508	55	85.9250	4	61.400	
509	8	85.1890	4	66.300	
510	2	88.8500	4	71.000	
601	2	78.5000	5	27.000	6.000
602	4	79.3250	5	36.300	6.000
603	2	80.9500	5	41.000	6.000
604	8	85.7740	5	52.800	6.000
605	12	86.8500	5	56.800	6.000
606	16	86.9230	5	61.400	6.000
701	1	69.9000	5	20.000	7.000
702	1	80.2000	5	36.000	7.000
703	2	82.3000	5	46.000	7.000
704	6	87.5670	5	57.300	7.000
705	5	87.5200	5	62.000	7.000
801	2	82.7500	5	35.500	8.000
802	2	82.8000	5	41.000	8.000
803	5	88.8000	5	54.000	8.000
804	4	88.3000	5	60.300	8.000
805	1	86.8000	5	65.000	8.000
901	3	82.4670	5	33.300	9.000
902	4	88.4250	5	55.300	9.000
903	4	89.4250	5	59.800	9.000
1001	6	88.0350	5	55.300	10.000
1002	6	89.7330	5	59.800	10.000
1101	5	81.3820	5	24.000	11.000
1102	7	81.8710	5	31.400	11.000
1103	11	85.0460	5	40.300	11.000
1104	4	86.5250	5	45.000	11.000
1105	8	88.2380	5	53.300	11.000
1106	14	88.6910	5	57.900	11.000
1107	3	88.7670	5	61.700	11.000
1108	1	89.8000	5	66.000	11.000

GIVEN ESTIMATE OF VARIANCE = 7.0510
NO OF DEGREES OF FREEDOM = 512

ESTIMATED RESULTS :

ID	SAMPLE AVERAGE	NEW ESTIMATE
201	72.2400	71.5582
202	72.1700	72.3850
203	73.7600	74.0301
204	75.3410	75.4931
205	75.8520	76.5270
206	79.2080	78.9276
207	80.2930	80.2904
208	82.0480	81.7369
209	83.4750	83.4387
301	79.6000	78.4848
302	76.7460	77.9640
303	79.3290	79.6169
304	78.9290	79.9371
305	82.7000	82.0538
306	85.1780	83.9119
307	84.2080	83.9290
308	83.8000	84.5924
401	77.9000	77.1553
402	78.9200	78.9177
403	79.3200	80.1956
404	81.9500	82.3702
405	84.7830	84.5082
406	85.3480	85.5078
407	86.0160	86.5505
408	89.7000	88.7316
501	77.4000	77.1774
502	77.3670	77.8476
503	79.5330	79.4212
504	80.6000	80.6725
505	81.0000	81.2956
506	84.5640	83.9852
507	85.6910	85.0640
508	85.9250	85.8457
509	85.1890	86.2831
510	88.8500	88.5268
601	78.5000	77.7200
602	79.3250	79.6234
603	80.9500	81.0899
604	85.7740	85.0827
605	86.8500	86.2026
606	86.9230	87.0041
701	69.9000	73.3564
702	80.2000	80.3406
703	82.3000	82.8938
704	87.5670	86.9879
705	87.5200	87.7560
801	82.7500	81.7240
802	82.8000	82.6667
803	88.8000	87.3519
804	88.3000	88.1989
805	86.8000	88.3600
901	82.4670	81.6390
902	88.4250	87.8151
903	89.4250	88.9870
1001	88.0350	88.0542
1002	89.7330	89.5177
1101	81.3820	80.4311
1102	81.8710	81.8756
1103	85.0460	84.6935
1104	86.5250	86.0991
1105	88.2380	88.2058
1106	88.6910	89.1660
1107	88.7670	89.8346
1108	89.8000	90.9868

SHRINKING FACTOR = 0.4177
%-IMPROVEMENT OVER USUAL ESTIMATE = 47.9383 %

note that the average noise level is still the usual estimate of the true noise level due to the sectionally linear relationship between noise level and traveling speed. Thus, the above method for combining data to increase sample sizes is valid in this problem.

In general, the proposed estimation method works better when large numbers of parameters (populations) are to be estimated. Thus, combining data to reduce parameters should not be done unless necessary. Because sample sizes are increased, new sample averages are more reliable than those in Examples 1 and 2. Consequently, it will be more difficult for the method to improve new sample averages.

For this combined data set, we shall arrange sample averages into five groups with members defined below:

Group 1 - Vehicles with 2 axles
Group 2 - Vehicles with 3 axles
Group 3 - Vehicles with 4 axles
Group 4 - Vehicles with 5 axles
Group 5 - Vehicles with 6 - 11 axles.

Again, the linear relationships among peak noise level, speed and number of axles are used to compute initial estimates. Thus, traveling speed is the only supplemental variable for each of Groups 1 through 4. For the fifth group, the supplemental variables are speed and number of axles. Since the combined data are no longer suitable for estimating the variance of the peak noise level, we shall supply the estimated variance obtained from the original sample observations (see Table 43 or 45). The data input for the above model specification is presented in Table 46. We see from this table that the first card sets NP = 64, NGROUP = 5, MDEG = 512, and VAR = 7.051. The second card indicates that there are 9, 8, 8, 10, and 29 members in Groups 1, 2, 3, 4, and 5, respectively. The third card instructs that the linear least squares estimates are to be used as initial estimates. The fourth card specifies that only one supplemental variable is to be used for the first four groups, but two for the last group. The estimated results are presented in Table 47.

7) Program Listing

```
      PROGRAM IMEAN(TAPE1=INPUT,TAPE2=OUTPUT)
      DIMENSION NSAMP(500),X(500),U(500),XT(2000),NAUX(500),
     1XEST(500),NMG(500),IG(500),ID(500),AUX(500,5),B(6),
     2METHOD(500),DESCPT(72),XL(500)
C
C     THIS PROGRAM IS FOR SIMULTANEOUSLY ESTIMATING NP INDEPENDENT
C     POPULATION MEANS BY USING STEIN-LIKE ESTIMATION PROCEDURES.
C
      NCARD=10
    1 READ(1,2,END=9999)NP,NGROUP,MDEG,VAR
      READ(1,3) (NMG(I),I=1,NGROUP)
      READ(1,3) (METHOD(I),I=1,NGROUP)
```

```
      READ(1,3) (NAUX(I),I=1,NGROUP)
      READ(1,10)(DESCPT(I),I=1,70)
      WRITE(2,15)(DESCPT(I),I=1,70)
      WRITE(2,30)
      SSR=0.
      NDF=0
      SSR1=0.
      NDF1=0
      K=0
      DO 1000 I=1,NGROUP
      NCOUNT=K
      NPT=NMG(I)
      WB1=0.
      WB2=0.
      DO 500 J=1,NPT
      K=NCOUNT+J
      IG(K)=I
      NA=NAUX(I)
      IF(MDEG .LE. 0)GO TO 111
      IF(METHOD(I) .NE. 2)GO TO 110
      READ(1,60)ID(K),NSAMP(K),X(K)
      GO TO 220
110   READ(1,60)ID(K),NSAMP(K),X(K),(AUX(J,M),M=1,NA)
      GO TO 200
111   CONTINUE
      IF(METHOD(I) .NE. 2)GO TO 112
      READ(1,60)ID(K),NSAMP(K)
      GO TO 113
112   READ(1,60)ID(K),NSAMP(K),(AUX(J,M),M=1,NA)
113   N3=NSAMP(K)
      IF(N3 .GT. 1)GO TO 116
      READ(1,61)X(K)
      IF(METHOD(I) .EQ. 2)GO TO 220
      GO TO 200
116   N2=0
      SX=0.
      SXX=0.
117   N1=N2+1
      N2=N1+NCARD-1
      IF(N2 .GT. N3)N2=N3
      READ(1,61)(XT(M),M=N1,N2)
      DO 118 M=N1,N2
      SX=SX+XT(M)
11    SXX=SXX+XT(M)**2
      IF(N2 .NE. N3)GO TO 117
      X(K)=SX/N3
      SSR1=SSR1+SXX-X(K)**2*N3
      NDF1=NDF1+N3-1
      IF(METHOD(I) .EQ. 2)GO TO 220
200   WRITE(2,210)ID(K),NSAMP(K),X(K),IG(K),(AUX(J,M),M=1,NA)
      IF(METHOD(I) .LT. 4)GO TO 213
      XL(J)=X(K)
      GO TO 500
213   CONTINUE
      IF(METHOD(I) .EQ. 3)GO TO 215
      U(K)=AUX(J,1)
      SSR=SSR+(X(K)-U(K))**2*NSAMP(K)
      GO TO 500
215   WB1=WB1+AUX(J,1)*X(K)
      GO TO 500
220   WRITE(2,210)ID(K),NSAMP(K),X(K),IG(K)
      WB1=WB1+X(K)*NSAMP(K)
      WB2=WB2+NSAMP(K)
500   CONTINUE
      IF(METHOD(I) .EQ. 1)GO TO 1000
      IF(METHOD(I) .GE. 4)GO TO 900
      IF(METHOD(I) .EQ. 2)WB1=WB1/WB2
      DO 800 J=1,NPT
      K=NCOUNT+J
      U(K)=WB1
```

```
800  SSR=SSR+(X(K)- U(K))**2*NSAMP(K)
     NDF=NDF+1
     GO TO 1000
900  CONTINUE
     NFORCE=0
     IF(METHOD(I) .EQ. 5)NFORCE=1
     NIND=NAUX(I)
     NDF=NDF+NIND+NFORCE
     CALL LSE(NFORCE,NPT,NIND,XL,AUX,B,XEST,SERQR)
     DO 950 J=1,NPT
     K=NCOUNT+J
     U(K)=XEST(J)
950  SSR=SSR+(X(K)-U(K))**2*NSAMP(K)
1000 CONTINUE
     IF(MDEG .LE  0)GO TO 1020
     WRITE(2,1015)VAR,MDEG
     GO TO 1050
1020 MDEG=NDF1
     VAR=SSR1/MDEG
     WRITE(2,1025)VAR,MDEG
1050 NTRACE=NP-NDF
     IF(NTRACE .GT. 2)GO TO 2000
     WRITE(2,1500)NTRACE
     GO TO 9999
2000 CONTINUE
     FACTOR=1.-VAR*MDEG*(NTRACE-2)/(SSR*(MDEG+2))
     IF(FACTOR .LT. O.)FACTOR=O.
     WRITE(2,2200)
     DO 2100 J=1,NP
     XEST(J)=U(J)+FACTOR*(X(J)-U(J))
     WRITE(2,3200)ID(J),X(J),XEST(J)
2100 CONTINUE
     SSR=SSR/VAR-NTRACE
     IF(SSR .LE. O.)SSR=O.
     CALL EXPECT(NTRACE,SSR,EYY)
     PIMPRO=EYY*(NTRACE-2)**2*MDEG/(MDEG+2)/NP*100
     WRITE(2,6000)FACTOR,PIMPRO
     GO TO 1
9999 CONTINUE
     STOP
2    FORMAT(3I3,F10.3)
3    FORMAT(20I3)
10   FORMAT(1X,70A1)
15   FORMAT(/,1X,45H#############################################,
    125H#########################,/,1X,70A1,/,1X,13H#############,
    257H#########################################################)
30   FORMAT(//,1X,4HDATA,//,5X,21HSAMPLE  SAMPLE  GROUP,/,2X,
    147HID  SIZE  AVERAGE   NO   SUPPLEMENTAL VARIABLES,/,1X,
    250H--------------------------------------------------,
    320H--------------------)
60   FORMAT(I4,I5,6F10.4)
61   FORMAT(10F7.2)
210  FORMAT(1X,I4,I6,F10.4,I3,2X,5F9.3)
1015 FORMAT(/,1X,28HGIVEN ESTIMATE OF VARIANCE =,F10.4,/,1X,
    128HNO OF DEGREES OF FREEDOM   =,I5,/)
1025 FORMAT(/,1X,28HSTRATIFIED SAMPLE VARIANCE =,F10.4,/,1X,
    128HNO OF DEGREES OF FREEDOM   =,I5,/)
1500 FORMAT(//,1X,33H--- WOO ----- ERROR ----- WOO ---,/,2X,
    113HTRACE( P ) = ,I4,/,2X,28HTHE TRACE OF P SHOULD BE AT ,
    27HLEAST 2,/,2X,38HCHECK TO SEE WHETHER YOU HAVE TOO MANY,
    37H GROUPS,/ 2X,34HOR/AND TOO MANY PAREMETERS IN THE ,
    416HREGRESSION LINES,/)
2200 FORMAT(//,1X,17HESTIMATED RESULTS,//,2X,18HID  SAMPLE AVERAGE,
    113H NEW ESTIMATE,/,1X,33H---------------------------------)
3200 FORMAT(1X,I4,3X,F10.4,4X,F10.4)
6000 FORMAT(/,1X,35H SHRINKING FACTOR                  =,F8.4,/,1X,
    135H%-IMPROVEMENT OVER USUAL ESTIMATE =,F8.4,2H %,/)
     END
```

```
      SUBROUTINE LSE(MODEL,NPT,NIND,YDATA,XDATA,B,YEST,SEROR)
      DIMENSION YDATA(500),XDATA(500,5),B(6),X(500,6),
     1YEST(500),XX(6,6),XXIXT(6,500)
C
C     SUBROUTINE FOR COMPUTING THE LEAST SQUARES ESTIMATE
C     OF B IN THE LINEAR MODEL,
C
C             YDATA = XDATA * B        , IF MODEL = 0
C     AND
C             YDATA = (I,XDATA) * B   , IF MODEL = 1
C
C     WHERE YDATA IS THE THE (NPT BY 1) VECTOR OF DEPENDENT
C     OBSERVATIONS, XDATA IS THE (NPT BY NIND) MATRIX OF
C     INDEPENDENT OBSERVATIONS, I IS THE COLUMN VECTOR WITH
C     EVERY ELEMENT EQUAL TO ONE, * STANDS FOR THE PRODUCT
C     OF TWO MATRICES, NIND IS THE NUMBER OF INDEPENDENT
C     VARIABLES AND B IS THE VECTOR OF PARAMETERS TO BE
C     ESTIMATED.
C
      IF(MODEL .EQ. 0)GO TO 100
      DO 10 I=1,NPT
 10   X(I,1)=1.
      MM=NIND+1
      DO 20 I=2,MM
      II=I-1
      DO 30 J=1,NPT
 30   X(J,I)=XDATA(J,II)
 20   CONTINUE
      GO TO 200
 100  MM=NIND
      DO 120 I=1,MM
      DO 130 J=1,NPT
 130  X(J,I)=XDATA(J,I)
 120  CONTINUE
 200  CONTINUE
      DO 230 I=1,MM
      DO 240 J=1,MM
      XX(I,J)=0.
      DO 250 K=1,NPT
 250  XX(I,J)=XX(I,J)+X(K,I)*X(K,J)
 240  CONTINUE
 230  CONTINUE
      CALL MTXINV(MM,XX,XX)
      DO 300 I=1,MM
      DO 310 J=1,NPT
      XXIXT(I,J)=0.
      DO 320 K=1,MM
 320  XXIXT(I,J)=XXIXT(I,J)+XX(I,K)*X(J,K)
 310  CONTINUE
 300  CONTINUE
      DO 400 I=1,MM
      B(I)=0.
      DO 410 J=1,NPT
 410  B(I)=B(I)+XXIXT(I,J)*YDATA(J)
 400  CONTINUE
      SEROR=0.
      DO 500 I=1,NPT
      YEST(I)=0.
      DO 510 J=1,MM
 510  YEST(I)=YEST(I)+X(I,J)*B(J)
      SEROR=SEROR+(YDATA(I)-YEST(I))**2
 500  CONTINUE
      SEROR=SQRT(SEROR/(NPT-MM))
      RETURN
      END
```

```
      SUBROUTINE MTXINV(NSIZE,W,WINV)
      DIMENSION ARRAY(6,6),WINV(6,6),W1(6,2),W(6,6)
C
C     SUBROUTINE FOR FINDING THE INVERSE OF AN (NSIZE BY NSIZE)
C     SQUARE MATRIX W BY USING THE PARTITION METHOD. WINV IS THE
C     INVERSE MATRIX OF W.
C
      DO 5 I=1,NSIZE
      DO 5 J=1,NSIZE
 5    ARRAY(I,J)=W(I,J)
      IF(NSIZE .GT. 1) GO TO 10
      WINV(1,1)=1./ARRAY(1,1)
      RETURN
 10   CONTINUE
      MSIZE=NSIZE-1
      DO 15  II=1,MSIZE
      J=II+1
      DO 16  KK=J,NSIZE
      DO  17 M=1,NSIZE
      W1(M,1)=W(M,II)
 17   W1(M,2)=W(M,KK)
      DET=W1(1,1)*W1(2,2)-W1(1,2)*W1(2,1)
      IF(DET .EQ. 0.) GO TO 16
      IF(II .EQ. 1 .AND. KK .EQ. 2) GO TO 19
      DO  18 K=1,NSIZE
      ARRAY(K,1)=W(K,II)
      ARRAY(K,2)=W(K,KK)
      ARRAY(K,II)=W(K,1)
 18   ARRAY(K,KK)=W(K,2)
      GO TO 19
 16   CONTINUE
 15   CONTINUE
 19   CONTINUE
      WINV(1,1)=ARRAY(2,2)/DET
      WINV(2,2)=ARRAY(1,1)/DET
      WINV(1,2)=-ARRAY(1,2)/DET
      WINV(2,1)=-ARRAY(2,1)/DET
      IF(NSIZE .EQ. 2) GO TO 100
      DO 20 I=3,NSIZE
      K=I-1
      DO 21 J=1,K
      W1(J,1)=0.
      W1(J,2)=0.
      DO 22 M=1,K
      W1(J,1)=W1(J,1)+WINV(J,M)*ARRAY(M,I)
 22   W1(J,2)=W1(J,2)+ARRAY(I,M)*WINV(M,J)
 21   CONTINUE
      ELTA=ARRAY(I,I)
      DO 23 J=1,K
 23   ELTA=ELTA-ARRAY(I,J)*W1(J,1)
      WINV(I,I)=1./ELTA
      DO 24 J=1,K
      WINV(J,I)=-W1(J,1)/ELTA
      WINV(I,J)=-W1(J,2)/ELTA
      DO 24 M=1,K
 24   WINV(J,M)=WINV(J,M)+W1(J,1)*W1(M,2)/ELTA
 20   CONTINUE
 100  CONTINUE
      IF(II .EQ. 1 .AND. KK .EQ. 2) GO TO 888
      DO 401 J=1,NSIZE
      W1(J,1)=WINV(1,J)
 401  W1(J,2)=WINV(2,J)
      DO 402 J=1,NSIZE
      WINV(1,J)=WINV(II,J)
 402  WINV(2,J)=WINV(KK,J)
      DO 403 J=1,NSIZE
      WINV(II,J)=W1(J,1)
 403  WINV(KK,J)=W1(J,2)
 888  CONTINUE
      RETURN
      END
```

```
      SUBROUTINE EXPECT(NP,THETA,EYY)
      DOUBLE PRECISION P1,A,EY,P2
C
C     THIS SUBROUTINE COMPUTES THE EXPECTATION OF 1/Y, WHERE Y IS A
C     NONCENTRAL CHI-SQUARE WITH NP DEGREES OF FREEDOM AND NONCENTRALITY
C     PARAMETER THETA. THIS EXPECTATION IS THE SAME AS THE EXPECTATION
C     OF 1/(NP-2+2W), WHERE W IS A POISSON WITH PARAMETER (THETA/2).
C
      ERROR=0.00001
      ERR=ERROR/NP
      A=THETA/2
      P1=DEXP(-A)
      EY=P1/(NP-2)
      P2=P1
      K=0
10    K=K+1
      P1=P1*A/K
      P2=P2+P1
      EY=EY+P1/(NP-2+K*2)
      CHECK=1.-P2
      IF(CHECK.GE.ERR) GO TO 10
      EYY=EY
      RETURN
      END
```

VIII
A COMPUTER PROGRAM FOR SIMULTANEOUSLY ESTIMATING CORRELATED MEANS BY USING STEIN-LIKE ESTIMATION PROCEDURES

1) The Basic Data

The essential data for estimating population mean vector are $\overline{X} = (\overline{X}_1, \ldots, \overline{X}_k)'$ and S. The superscript $'$ stands for the transpose of a vector or matrix. $\overline{X}$ is the sample average of n observation vectors from the population with mean vector $\theta = (\theta_1, \ldots, \theta_k)$ and unknown covariance matrix Σ. The size of mean vector is k which is the number of mean parameters to be estimated. Denote $X_h = (X_{h1}, \ldots, X_{hk})$ to be the h-th observation vector. Then,

$$\overline{X}_i = \frac{1}{n} \sum_{h=1}^{n} X_{hi} \tag{177}$$

We assume that either the population is normally distributed or the sample size is large enough to guarantee the normality of the vector of sample averages by the Central Limit Theorem. Therefore, $\overline{X} = (\overline{X}_1, \ldots, \overline{X}_k)'$ is a k-variate normal with mean vector $\theta = (\theta_1, \ldots, \theta_k)'$ and covariance matrix Σ/n. In this case, $\overline{X}$ is the usual estimate of θ.

S is an unbiased estimate of Σ such that m S is a Wishart matrix with parameters n and Σ. This matrix is obtained from either past experiments or current data. In the latter case, S is the unbiased sample covariance matrix with the (i, j)-th element defined as

$$S_{ij} = \frac{1}{n-1} \sum_{h=1}^{n} (X_{hi} - \overline{X}_i)(X_{hj} - \overline{X}_j) \tag{178}$$

In this case, we have

$$m = n - 1 \tag{179}$$

Other essential data is the information for computing initial estimates of mean parameters. This is termed the 'supplemental' information for discussion purposes. Five methods for using the supplemental information to compute initial parameter estimates are discussed in Section 3.

2) What the Program Does

This program is written, based on theoretical results developed in Part III, for combining the usual estimate $\overline{X}$ and supplemental information to estimate the population mean vector θ.

The program first computes the initial estimate, U_i, of θ_i by the chosen method. Denote $U = (U_1, \ldots, U_k)'$. In this step, the trace of the idempotent matrix P satisfying

$$\overline{X} - U = P\,\overline{X} \tag{180}$$

is also computed. The second step is to compute the shrinking factor c defined as

$$c = [\text{Trace}(P) - 2]\ \frac{m}{m-k+3} \cdot \frac{1}{n(X-U)'S^{-1}(X-U)} \tag{181}$$

A slightly better procedure is to set c to be 1 if it is greater than 1. The third step is to compute the final estimate of θ_i defined as

$$\hat{\overline{X}}_i = U_i + (1-c)(\overline{X}_i - U_i),\ i = 1, \ldots, k \tag{182}$$

This program also computes the estimated percentage improvement of the above procedure over the usual one.

3) Methods for Computing Initial Estimates

The key to obtaining good estimates of mean parameters is to provide good initial estimates in the sense that N is high and ζ is low, where

$$N = \text{Trace}(P) \tag{183}$$

and

$$\zeta = n(U-\theta)'\Sigma^{-1}(U-\theta) \tag{184}$$

We have shown in Part II that if P is a symmetrical idempotent matrix, N and ζ are, respectively, the number of degrees of freedom and the noncentrality parameter of a noncentral chi-square distribution. For this case, the maximal percentage improvement that can be achieved by this method is $100\ [(m-k+1)(N-2)]\ /\ [(m-k+3)k]$. Five methods for computing initial estimates of mean parameters are built into this program. These methods are presented below.

Method 1: Initial Estimates are Given - Based on past experiments or parallel studies, we estimate or guess θ_i to be U_i, $i = 1, \ldots, k$. In this case, we treat $U_i - \theta_i$ as the parameter to be estimated. Consequently, the idempotent matrix P satisfying Eq. (180) is the identity matrix. Thus, N = Trace (P) = k. The estimation accuracy is the degree of closeness of U to θ. The final estimate $\hat{\bar{X}}_i$ always lies between U_i and $\bar{X}_i$. When initial estimates are excellent, i.e., U is very close to θ, the final estimate of θ is U. However, if initial estimates are poor, i.e., U is quite distant from θ, the final estimate of θ will be very close to the usual estimate $\bar{X}$.

This method is used only when the number of parameters to be estimated is at least 3. When initial estimates are reliably close to the true means, this method will produce good final estimates of mean parameters.

Method 2: Weighted Average (I) - When mean parameters are almost homogeneous, we may consider that $\theta_i = \bar{\theta}$ for every i. We therefore use an unbiased estimate of $\bar{\theta}$ as the initial estimate of every mean parameter. That is,

$$U_i = \sum_{j=1}^{k} \bar{X}_j/n, \quad i = 1, \ldots, k \tag{185}$$

For this method, N = Trace (P) = k - 1. The loss of one degree of freedom is due to the estimation of the unknown parameter $\bar{\theta}$.

This method is used only when the number of parameters to be estimated is at least 4. This method will produce good estimates of mean parameters, if these mean parameters are nearly homogeneous.

Method 3: Weighted Average (II) - For a given set of numbers, $w_1, \ldots, w_k$, satisfying $w_1 + \ldots + w_k = 1$, we define

$$U_i = \sum_{j=1}^{k} w_j \bar{X}_j, \quad i = 1, \ldots, k \tag{186}$$

For this case, N = Trace (P) = k - 1. We note that this method is the same as Method 2 when $w_1 = \ldots = w_k$.

This method is used only when the number of mean parameters to be estimated is at least 4. The estimation results will be good if the mean parameters are almost homogeneous.

We remark that w_i can be interpreted as the weight assigned to $\overline{X}_i$ for estimating the common parameter $\overline{\theta}$. If one has reason to doubt the reliability of a particular observation, say $\overline{X}_1$, zero weight may be assigned to $\overline{X}_1$. That is, $w_1 = 0$ and $w_2 + \ldots + w_k = 1$.

Method 4: Least Squares Estimate (I) - Suppose that the supplemental data $(t_{i1}, \ldots, t_{im})$ are available and satisfy the following equation,

$$\theta_i = b_{i1} t_{i1} + \ldots + b_{im} t_{im}, \; i = 1, \ldots, k \tag{187}$$

If for every j, $b_{1j}, \ldots$, and b_{kj} are near an unknown common parameter b_j, Eq. (187) can be rewritten as

$$\theta_i = b_1 t_{i1} + \ldots + b_m t_{im}, \; i = 1, \ldots, k \tag{188}$$

Denote $(\hat{b}_1, \ldots, \hat{b}_m)$ to be the least squares estimate of $(b_1, \ldots, b_m)$. We then take the initial estimate of θ_i to be

$$U_i = \hat{b}_1 t_{i1} + \ldots + \hat{b}_m t_{im}, \; i = 1, \ldots, k \tag{189}$$

In this case, N = Trace (P) = k - m. The loss of m degrees of freedom is due to the least squares estimates of m linear parameters. We note that U_i can always be improved by increasing the number of supplemental variables, i.e., to increase m. However, this decreases N which is an undesirable property as previously mentioned. In general, m should be kept small relative to the number of parameters k.

This method is used only when the number of mean parameters to be estimated is at least m + 3. This method will produce good estimates of mean parameters if Eq. (188) holds approximately and the number of supplemental variables is small relative to the number of parameters to be estimated. In practice, one may use this method if the following linear relationship holds approximately.

$$\overline{X}_i = b_1 t_{i1} + \ldots + b_m t_{im}, \; i = 1, \ldots, k \tag{190}$$

Method 5: Least Squares Estimate (II) - This method is the same as Method 4 except that Eqs. (188) and (189) are, respectively, replaced by

$$\theta_i = b_0 + b_1 t_{i1} + \ldots + b_m t_{im}, \; i = 1, \ldots, k \tag{191}$$

and

$$U_i = \hat{b}_0 + \hat{b}_1 t_{i1} + \ldots + \hat{b}_m t_{im}, \; i = 1, \ldots, k \tag{192}$$

In this case, $N = \text{Trace}(P) = k - m - 1$ because of the extra parameter b_0. Thus, this method is used only when the number of mean parameters to be estimated is at least $m + 4$. In practice, one may use this method if the following linear relationship holds approximately

$$\overline{X}_i = b_0 + b_1 t_{i1} + \ldots + b_m t_{im}, \quad i = 1, \ldots, k \tag{193}$$

The k components of the mean vector may be arranged into many groups to which different methods are applied to minimize the noncentrality parameter defined in Eq. (184). For example, we may use Method 1 for a group of k_1 parameters (components), Method 4 for a group of k_2 parameters, Method 2 for a group of k_3 parameters, and again, Method 2 for the group of the remaining ones. In this case, we have $N = \text{Trace}(P) = k_1 + (k_2 - m) + (k_3 - 1) + (k - k_1 - k_2 - k_3 - 1) = k - m - 2$. m is the number of supplemental variables used in Method 4. For this particular case, the number of mean parameters to be estimated should be at least $m + 5$. In general, for any method or combination of methods, the number of degrees of freedom should be at least 2. Combining these five methods to compute initial parameter estimates will be demonstrated by examples presented in Section 6.

Every method except the first one suggests that k parameters should be arranged into groups in which parameters cluster at a point or can be approximated by a linear function of supplemental variables. Physical properties of populations, past experiments and parallel studies are good sources for obtaining the proper group-method combination. Unfortunately, these sources may not be available or reliable enough in a particular problem. In this circumstance, one may examine the data as suggested in each method to choose the proper group-method combination. That is, the empirical relationship among $\overline{X}_1, \ldots, \overline{X}_k$, and supplemental variables can be used to determine the group-method combination for computing initial estimates of parameters. Since the program computes the estimated percentage improvement of each chosen method-group combination over the usual method, one may use the one that produces the maximal improvement to estimate parameters. We remark that the group-method combination generated by examining the data may not be the best one and, possibly, could be the worst one for this problem due to random variation of the data. Nevertheless, if the chosen group-method combination is used thereafter for the same problem, the above method is always better than the usual one. The worst situation is that no improvement is made. Based on our experience, qualitative properties of populations often provide adequate information for grouping purposes.

4) Data Input

The data input of this program is arranged into three portions. The first portion is composed of five cards. These cards specify the number of groups and parameters, computational methods, and number of supplemental variables in each group. Variables used in these cards are defined below.

NP:	Number of parameters to be estimated (the size of mean vector), $1 \leq$ NP ≤ 50
NGROUP:	Number of groups, $1 \leq$ NGROUP ≤ 20
NSAMP:	Number of observation vectors sampled from the population to obtain usual estimate of population mean vector.
MDEG:	Number of degrees of freedom for estimating Σ. (When this number is positive, the unbiased estimate of Σ is supplied. Otherwise, the program will use Eq. (178) to compute the sample covariance matrix as an estimate of Σ.)
NMG(I):	Number of mean parameters in the I-th group, NMG(1) + . . . + NMG (NGROUP) = NP
METHOD(I) = j:	The j-th method presented in Section 3 is used to compute initial estimates of mean parameters in the I-th group, $1 \leq j \leq 5$
NAUX(I):	Number of supplemental variables used to compute initial estimates of mean parameters in the I-th group. This variable is 1 if METHOD(I) = 1 or 3, and is 0 if METHOD(I) = 2.
(DESCPT(I), I = 1, 70):	Title (no more than 70 letters).

The input format of the second portion is determined by the input value of MDEG. This is explained below.

a) MDEG is positive. An unbiased estimate of the population covariance matrix is supplied. The input variable of the (I, J)-th element of this matrix is COVAR(I, J). Since this matrix is symmetrical, only the lower triangular portion of the estimated covariance matrix is needed. Each row is arranged into cards such that each card, except the last one, has 10 numbers. Thus, the input format of the I-th row is as follows:

```
COVAR(I,1), ......., COVAR(I,10)
COVAR(I,11),......., COVAR(I,20)
.............................
.............................
...........COVAR(I,I)
```

The number of cards for the I-th row is L if $10 (L - 1) < I \leq 10L$.

b) MDEG is negative. An unbiased estimate of the population covariance matrix is not supplied and is to be obtained from sample observation vectors (see Eq. (178)). This portion consists of NSAMP subportions. Each subportion contains one observation vector. The data input format of the I-th subportion is as follows:

```
XT(1), .................., XT(10)
XT(11),.................., XT(20)
  ..........................
  ..........................
  ............, XT(NP)
```

XT(J) is the J-th component of the I-th observation vector. The number of cards in this subportion is L if $10 (L - 1) < NP \leq 10L$. Thus, the I-th observation vector is arranged into L cards such that each card, except the last one, has 10 numbers. We note that NSAMP observation vectors will be converted to the usual estimate by Eq. (177) and the sample covariance matrix S by Eq. (178).

The third portion is composed of NGROUP subportions or groups. We define K = J if I = 1 and K = NMG(1) + . . . + NMG(I - 1) + J if I > 1. Then, the K-th card of the third portion contains essential and supplemental data for estimating the J-th parameter of the I-th group. We note that the J-th parameter of the I-th group is the K-th parameter. The data input for estimating this parameter is, when MDEG is positive,

ID(K), X(K), (AUX(J, M), M = 1, NAUX(I))

and is, when MDEG is negative,

ID(K), (AUX(J, M), M = 1, NAUX(I))

The input variables of the above statements are defined below.

ID(K): Identification of the K-th component (parameter) such as sieve number. The K-th parameter is the J-th parameter of the I-th group. This number has no effect on the estimation procedure.

X(K): The usual estimate of the K-th mean parameter defined in Eq. (177).

AUX(J, M): The M-th supplemental variable for the K-th mean parameter.

Table 48
Input Deck Used When An Unbiased Estimate of The Covariance Matrix Is Supplied (The Input Value of MDEG Is Positive)

Card Number	Variables Used in Each Card	Remarks
1	NP,NGROUP,NSAMP,MDEG	
2	(NMG(I),I=1,NGROUP)	* The Data Input of
3	(METHOD(I),I=1,NGROUP)	The First Portion
4	(NAUX(I),I=1,NGROUP)	
5	(DESCPT(I),I=1,70)	
6	COVAR(1,1)	* There are NP subportions
7	COVAR(2,1),COVAR(2,2)	in the 2nd portion of
.	....	the data input
.	....	
.	COVAR(I,1), ...,COVAR(I,10)	← This is the I-th
.	COVAR(I,11),.........,COVAR(I,20)	. subportion which is
.	....	. composed of L cards if
.	..............,COVAR(I,I)	← 10(L-1) < I < 10L
.	....	
.	....	
.		* NMG(1) Cards for The 1st
.		Group of The 3rd Portion
⋮	⋮	⋮
.		* NMG(I) Cards for The
.		I-th Group ; NA=NAUX(I)
.		* K=NMG(1)+..+NMG(I-1)+J
.		
.	ID(K),X(K),(AUX(J,M),M=1,NA)	← This card is for the
.		K-th parameter which
.		is the J-th parameter
.		of the I-th Group
.		
⋮	⋮	⋮
.		* NMG(NGROUP) Cards for
.		The Last Group of The
.		Third Portion

The input deck is presented in Table 48 for the case that MDEG is positive, and in Table 49 for the case that MDEG is negative. The input deck is also diagrammed in Figure 8 to show the format and logic used. The user can follow this diagram to change, if needed, read statements and formats to fit a particular problem.

Table 49
Input Deck Used When An Unbiased Estimate
of The Covariance Matrix Is Not Supplied
(The Input Value of MDEG Is Negative)

Card Number	Variables Used in Each Card	Remarks
1	NP,NGROUP,NSAMP,MDEG	* The Data Input of The First Portion
2	(NMG(I),I=1,NGROUP)	
3	(METHOD(I),I=1,NGROUP)	
4	(NAUX(I),I=1,NGROUP)	
5	(DESCPT(I),I=1,70)	
.		* The 2nd portion of the data input has NSAMP subportions
.	XT(1), ..., XT(10)	← This is the I-th subportion which is
.	XT(11),..............., XT(20)	composed of L cards
.	....	
.	..............,XT(NP)	← if 10(L-1) < I < 10L
.		* NMG(1) Cards for The 1st Group of The 3rd Portion
⋮	⋮	⋮
.		* NMG(I) Cards for The I-th Group ; NA=NAUX(I)
.		* K=NMG(1)+..+NMG(I-1)+J
.	ID(K),(AUX(J,M),M=1,NA)	← This card is for the K-th parameter which is the J-th parameter of the I-th group
⋮	⋮	⋮
.		* NMG(NGROUP) Cards for The Last Group of The Third Portion

5) Limitations of the Program and How to Make Necessary Changes

This program was designed to handle problems where the number of mean parameters, NP, does not exceed 50. This number can be easily increased to any desired number. However, one must change dimensions of variables that appear on the DIMENSION STATEMENTS of the main program and subroutines LSE, i.e., replace every 50 by the desired number.

When the input value of MDEG is negative, NSAMP observation vectors sampled from the population are the input data for computing the vector of sample averages and sample covariance matrix. Each observation vector is arranged into L cards, if $10\,(L - 1) < NP \leq 10L$, such that each card, except the last one, has 10 numbers. One can easily reset this number

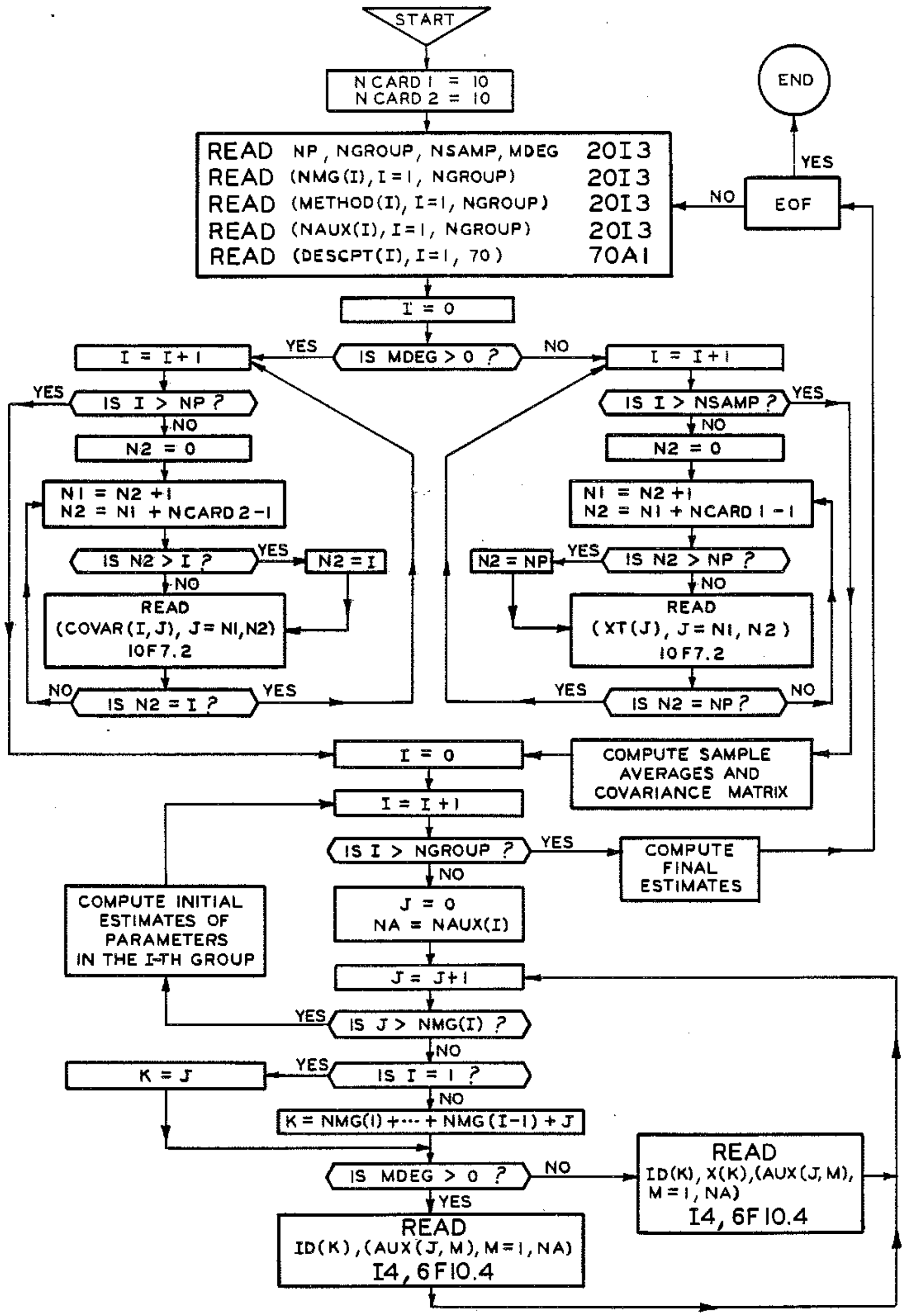

Figure 8. Flow of control for data input.

through the first statement, NCARD1 = 10, of the main program. That is, to replace the number '10' by the desired one. Of course, the corresponding READ statement (FORMAT 60) should also be changed to fit the new arrangement.

When the input value of MDEG is positive, the lower triangular portion of the estimated covariance matrix is supplied to the program. The I-th row of the lower triangular matrix is arranged into L cards, if $10(L-1) < I \leq 10L$, such that each card, except the last one, has 10 numbers. Again, this number can be reset through the second statement, NCARD2 = 10, of the main program. That is, to replace the number '10' by the desired number. One would also have to change the corresponding READ statement (FORMAT 61) to fit the new arrangement.

The number of groups is limited to 20. This number is large enough for most practical problems. However, one can increase this number to any desired number not exceeding NP by properly changing the READ format of the second input card (FORMAT 3).

The total number of supplemental variables is limited to 5. This number is large enough for most practical work. The user can change this number to any desired number, say N, by the following two steps.

a) Change the READ format of the second portion of the data input (FORMAT 62).

b) Change dimensions of variables that appear on the DIMENSION STATEMENTS of the main program and subroutine LSE, i.e., to replace 5 by N. We note that N should not exceed 49.

6) Examples

If aggregate inspection is to be shifted from the production site to the construction site, specification limits should be adjusted for aggregate degradation due to transporting and compaction. For the purpose of estimating aggregate degradation, 30 samples each were obtained from production and construction sites. For each pair of samples, the differences in aggregate percentage passing 3/4-in., 1/2-in., 3/8-in., No. 4, No. 8, No. 16, No. 30, No. 50, No. 100 and No. 200 sieves, and percent loss-by-washing were measured. We shall use this set of data to demonstrate how to improve the sample average vector.

Example 1: Method 5 with One Group - There is no supplemental information available for computing initial estimates of aggregate changes measured by various sieve sizes. However, if we code the sieve sizes from 1 through 11, the aggregate change is approximately a third-degree polynomial function of sieve size. That is, we have approximately,

$$\overline{X}_i = a + b\,i + c\,i^2 + d\,i^3, \; i = 1, \ldots , 11$$

where $\overline{X}_i$ is the aggregate change measured by the sieve size with index i (1 for 3/4-in., 2 for 1/2-in., . . . , etc.). The input data for using the above relationship to estimate initial and final aggregate changes measured by various sieve sizes is presented in Table 50. The first card sets NP = 11, NGROUP = 1, NSAMP = 30, and MDEG = -1. Because NGROUP = 1, only one number appears on each of cards 2 through 4, and the number of members in the only group is 11 which is set in the second card. The third and fourth cards instruct that Method 5 with three supplemental variables is to be used to compute initial estimates of aggregate changes due to transporting and compaction. Since NSAMP = 30 and MDEG is negative, 30 observation vectors comprise the data input of the second portion. Since the

Table 50
Data Input of Example 1

```
                                 Column  Number
12345678901234567890123456789012345678901234567890123456789012345678901234567890

11   1 30 -1
11
 5
 3
L.S.E. OF A POLINOMIAL OF THE 3RD ORDER OF THE SIEVE SIZE
  4.75   11.75   12.20   12.04   10.31    7.39    4.34    2.71    2.05    1.94
  1.72
  ..................................................................
  ....
  ..................................................................
  ....
  0.15   -2.85   -3.40   -4.36   -2.49   -0.71   -0.46   -1.19    0.05    0.54
  0.42
  1       1.0000     1.0000     1.0000
  2       2.0000     4.0000     8.0000
  3       3.0000     9.0000    27.0000
  4       4.0000    16.0000    64.0000
  5       5.0000    25.0000   125.0000
  6       6.0000    36.0000   216.0000
  7       7.0000    49.0000   343.0000
  8       8.0000    64.0000   512.0000
  9       9.0000    81.0000   729.0000
 10      10.0000   100.0000  1000.0000
 11      11.0000   121.0000  1331.0000
```

vector size is 11, each observation vector is arranged into two cards with 10 and 1 numbers, respectively. These 11 numbers are the aggregate changes measured by various sieve sizes. Due to the large data set, we only present the first and thirtieth observation vectors in Table 50 to show the data input format. The data of the second portion will be converted into sample averages and sample covariance matrix. The third portion is composed of 11 cards. The i-th card of this portion is the identification and supplemental variables for the aggregate change measured by the sieve with index i. In this example, this card contains i, i, i^2, and i^3. The estimated results are presented in Table 51. The total percentage improvement is only 8.6 percent. This is probably because the number of supplemental variables (3) is too large relative to the number of parameters (11) to be estimated in this problem. Better supplemental information is needed in order to further improve these results.

We note that sample averages and the sample covariance matrix can be the data input of the program. The data input for this option is presented in Table 52. We see from this table that MDEG in this case is 29 (NSAMP - 1). Therefore, the second portion is the lower triangular portion of the sample covariance matrix (see Table 51). The first 10 rows of the lower triangular portion of the sample covariance matrix are the first 10 cards of the second portion. However, the last row is arranged into two cards with 10 and 1 numbers, respectively, and is presented in the last two cards of the second portion. The third portion is composed of 11 cards. The i-th card contains i, $\overline{X}_i$, i, i^2, and i^3, i = 1, . . . , 11.

Example 2: Method 2 with Two Groups - Based on sample averages, we shall use group averages of the following two groups to compute the initial parameter estimates.

Group 1 - Sieve sizes indexed by 2 through 7
Group 2 - Sieve sizes indexed by 8 through 11, and by 1.

The data input is presented in Table 53. We see from this table that the first card sets NP = 11, NGROUP = 2, NSAMP = 30, and MDEG = -1. Thus, each of cards 2 through 4 has two numbers. The second card sets 6 and 5 parameters in Groups 1 and 2, respectively. The third card specifies the use of group averages as initial estimates. Therefore, no supplemental information is required. This is set in the fourth card. The second portion of the data input is obtained from the second portion of Example 1 by rearranging the sample covariance matrix to match with the new arrangement of group parameters. The third portion is similarly arranged. The estimated results are presented in Table 54. This table shows that this method is not as good as the one used in Example 1. This means that the above group arrangement is not a good way to obtain initial and, consequently, final estimates of mean parameters.

Example 3: Using Previous Results as Supplemental Information - Aggregate degradation due to transporting and compaction has been previously established for a different type of aggregate. We observe graphically that

Table 51
Estimated Results of Example 1

##
L.S.E. OF A POLINOMIAL OF THE 3RD ORDER OF THE SIEVE SIZE
##

DATA :

ID	SAMPLE AVERAGE	GROUP NO	SUPPLEMENTAL VARIABLES		
1	1.9933	1	1.000	1.000	1.000
2	3.3967	1	2.000	4.000	8.000
3	4.0433	1	3.000	9.000	27.000
4	4.0333	1	4.000	16.000	64.000
5	3.8300	1	5.000	25.000	125.000
6	3.5200	1	6.000	36.000	216.000
7	3.1200	1	7.000	49.000	343.000
8	2.4667	1	8.000	64.000	512.000
9	1.7833	1	9.000	81.000	729.000
10	1.5200	1	10.000	100.000	1000.000
11	1.2750	1	11.000	121.000	1331.000

SAMPLE SIZE (FOR SAMPLE AVERAGE) = 30
NO OF DEGREES OF FREEDOM (FOR COVARIANCE)= 29
THE UNBIASED ESTIMATE OF COVARIANCE MATRIX IS

13.93										
15.08	31.05									
14.09	30.39	34.93								
11.58	26.66	25.87	36.08							
8.34	19.40	19.59	27.94	22.90						
5.92	11.66	11.27	19.55	17.00	14.22					
4.06	7.12	6.01	14.67	13.25	12.05	11.21				
2.49	3.79	1.10	10.93	9.63	9.09	9.19	8.96			
1.08	1.42	0.10	4.32	3.84	3.84	3.84	3.81	1.84		
1.05	1.53	0.95	2.49	2.27	2.21	2.13	1.92	0.92	0.62	
0.92	1.48	1.68	1.17	1.30	1.28	1.15	0.77	0.41	0.42	0.49

ESTIMATED RESULTS :

ID	SAMPLE AVERAGE	NEW ESTIMATE
1	1.9933	2.0167
2	3.3967	3.3670
3	4.0433	4.0129
4	4.0333	4.0543
5	3.8300	3.8636
6	3.5200	3.5308
7	3.1200	3.0930
8	2.4667	2.4529
9	1.7833	1.8027
10	1.5200	1.5048
11	1.2750	1.2829

SHRINKING FACTOR = 0.7500
%-IMPROVEMENT OVER USUAL ESTIMATE = 8.5980 %

Table 52
Data Input of The Second Option in Example 1

```
                                    Column  Number
1234567890123456789012345678901234567890123456789012345678901234567890
 11   1 30 29
 11
  5
  3
L.S.E. OF A POLINOMIAL OF THE 3RD ORDER OF THE SIEVE SIZE
  13.93
  15.08   31.05
  14.09   30.39   34.93
  11.58   26.66   25.87   36.08
   8.34   19.40   19.59   27.94   22.90
   5.92   11.66   11.27   19.55   17.00   14.22
   4.06    7.12    6.01   14.67   13.25   12.05   11.21
   2.49    3.79    1.10   10.93    9.63    9.09    9.19    8.96
   1.08    1.42    0.10    4.32    3.84    3.84    3.84    3.81    1.84
   1.05    1.53    0.95    2.49    2.27    2.21    2.13    1.92    0.92    0.62
   0.92    1.48    1.68    1.17    1.30    1.28    1.15    0.77    0.41    0.42
   0.49
  1     1.9933     1.0000      1.0000      1.0000
  2     3.3967     2.0000      4.0000      8.0000
  3     4.0433     3.0000      9.0000     27.0000
  4     4.0333     4.0000     16.0000     64.0000
  5     3.8300     5.0000     25.0000    125.0000
  6     3.5200     6.0000     36.0000    216.0000
  7     3.1200     7.0000     49.0000    343.0000
  8     2.4667     8.0000     64.0000    512.0000
  9     1.7833     9.0000     81.0000    729.0000
 10     1.5200    10.0000    100.0000   1000.0000
 11     1.2750    11.0000    121.0000   1331.0000
```

Table 53
Data Input of Example 2

```
                                    Column  Number
1234567890123456789012345678901234567890123456789012345678901234567890
 11  2 30 -1
  6  5
  2  2
  0  0
USING 2-GROUP AVERAGES( SIZES 2-7 & OTHERS )
  11.75   12.20   12.04   10.31    7.39    4.34    2.71    2.05    1.94    1.72
   4.75
  ..................................................................
  ....
  ..................................................................
  ....
 -2.85   -3.40   -4.36   -2.49   -0.71   -0.46   -1.19    0.05    0.54    0.42
  0.15
  2
  3
  4
  5
  6
  7
  8
  9
 10
 11
  1
```

Table 54
Estimated Results of Example 2

##
USING 2-GROUP AVERAGES(SIZES 2-7 & OTHERS)
##

DATA :

ID	SAMPLE AVERAGE	GROUP NO	SUPPLEMENTAL VARIABLES
2	4.0433	1	
3	4.0333	1	
4	3.8300	1	
5	3.5200	1	
6	3.1200	1	
7	2.4667	2	
8	1.7833	2	
9	1.5200	2	
10	1.2750	2	
11	1.9933	2	
1	3.3967	1	

SAMPLE SIZE (FOR SAMPLE AVERAGE) = 30
NO OF DEGREES OF FREEDOM (FOR COVARIANCE)= 29
THE UNBIASED ESTIMATE OF COVARIANCE MATRIX IS

31.05									
30.39	34.93								
26.66	25.87	36.08							
19.40	19.59	27.94	22.90						
11.66	11.27	19.55	17.00	14.22					
7.12	6.01	14.67	13.25	12.05	11.21				
3.79	1.10	10.93	9.63	9.09	9.19	8.96			
1.42	0.10	4.32	3.84	3.84	3.84	3.81	1.84		
1.53	0.95	2.49	2.27	2.21	2.13	1.92	0.92	0.62	
1.48	1.68	1.17	1.30	1.28	1.15	0.77	0.41	0.42	0.49
15.08	14.09	11.58	8.34	5.92	4.06	2.49	1.08	1.05	0.92
13.93									

ESTIMATED RESULTS :

ID	SAMPLE AVERAGE	NEW ESTIMATE
2	4.0433	4.0382
3	4.0333	4.0283
4	3.8300	3.8277
5	3.5200	3.5218
6	3.1200	3.1272
7	2.4667	2.4578
8	1.7833	1.7837
9	1.5200	1.5238
10	1.2750	1.2821
11	1.9933	1.9908
1	3.3967	3.4002

SHRINKING FACTOR = 0.9866
%-IMPROVEMENT OVER USUAL ESTIMATE = 0.5609 %

this information is linearly related to the current data. We therefore use Method 4 to compute initial estimates. The data input is presented in Table 55. Explanations of the first and second portions of this table are the same as those in Example 1. Each card of the third portion is sieve index (identification) and the available aggregate degradation. The estimated results are presented in Table 56. We see from this table that the percentage improvement is substantially increased, but differences between usual and proposed estimates are very small. This means that the relationship between aggregate degradations of the two types of aggregates is indeed linear. This information can be used for future application.

Table 55
Data Input of Example 3

```
                                 Column  Number
1234567890123456789012345678901234567890123456789012345678901234567890

11   1 30 -1
11
 4
 1
SUPPLEMENTAL VARIABLE IS OBTAINED FROM PREVIOUS EXPERIMENT
  4.75   11.75   12.20   12.04   10.31    7.39    4.34    2.71    2.05    1.94
  1 72
  ........... ..................................................
  ....
  ..............................................................
  ....
  0.15   -2.85   -3.40   -4.36   -2.49   -0.71   -0.46   -1.19   0 05    0.54
  0.42
  1      1.7500
  2      3.1000
  3      3.6000
  4      3.6500
  5      3.4000
  6      3.1500
  7      2.8500
  8      2.2500
  9      1.6200
 10      1.3500
 11      1.2000
```

Table 56
Estimated Results of Example 3

##
SUPPLEMENTAL VARIABLE IS OBTAINED FROM PREVIOUS EXPERIMENT
##

DATA :

ID	SAMPLE AVERAGE	GROUP NO	SUPPLEMENTAL VARIABLES
1	1.9933	1	1.750
2	3.3967	1	3.100
3	4.0433	1	3.600
4	4.0333	1	3.650
5	3.8300	1	3.400
6	3.5200	1	3.150
7	3.1200	1	2.850
8	2.4667	1	2.250
9	1.7833	1	1.620
10	1.5200	1	1.350
11	1.2750	1	1.200

SAMPLE SIZE (FOR SAMPLE AVERAGE) = 30
NO OF DEGREES OF FREEDOM (FOR COVARIANCE)= 29
THE UNBIASED ESTIMATE OF COVARIANCE MATRIX IS

13.93										
15.08	31.05									
14.09	30.39	34.93								
11.58	26.66	25.87	36.08							
8.34	19.40	19.59	27.94	22.90						
5.92	11.66	11.27	19.55	17.00	14.22					
4.06	7.12	6.01	14.67	13.25	12.05	11.21				
2.49	3.79	1.10	10.93	9.63	9.09	9.19	8.96			
1.08	1.42	0.10	4.32	3.84	3.84	3.84	3.81	1.84		
1.05	1.53	0.95	2.49	2.27	2.21	2.13	1.92	0.92	0.62	
0.92	1.48	1.68	1.17	1.30	1.28	1.15	0.77	0.41	0.42	0.49

ESTIMATED RESULTS

ID	SAMPLE AVERAGE	NEW ESTIMATE
1	1.9933	1.9618
2	3.3967	3.4265
3	4.0433	4.0150
4	4.0333	4.0468
5	3.8300	3.7960
6	3.5200	3.5066
7	3.1200	3.1492
8	2.4667	2.4875
9	1.7833	1.7936
10	1.5200	1.5070
11	1.2750	1.3119

SHRINKING FACTOR = 0.3629
%-IMPROVEMENT OVER USUAL ESTIMATE = 36.5558 %

7) Program Listing

```
      PROGRAM CMEAN(TAPE1=INPUT,TAPE2=OUTPUT)
      DIMENSION X(50),U(50),XT(50),NAUX(50),XEST(50),NMG(50),
     1IG(50),ID(50),AUX(50,5),B(50),METHOD(50),DESCPT(72),XL(50),
     2COVAR(50,50),COINV(50,50)
C
C     THIS PROGRAM IS FOR SIMULTANEOUSLY ESTIMATING NP
C     CORRELATED MEANS BY USING STEIN-LIKE ESTIMATION PROCEDURES.
C
      NCARD1=10
      NCARD2=10
    1 READ(1,3,END=9999)NP,NGROUP,NSAMP,MDEG
      READ(1,3) (NMG(I),I=1,NGROUP)
      READ(1,3) (METHOD(I),I=1,NGROUP)
      READ(1,3) (NAUX(I),I=1,NGROUP)
      READ(1,10)(DESCPT(I),I=1,70)
      WRITE(2,15)(DESCPT(I),I=1,70)
      WRITE(2,30)
      IF(MDEG .GT. 0)GO TO 50
      DO 32 I=1,NP
      DO 32 J=1,NP
   32 COVAR(I,J)=0.
      DO 40 I=1,NSAMP
      N2=0
   33 N1=N2+1
      N2=N1+NCARD1-1
      IF(N2 .GT. NP)N2=NP
      READ(1,60)(XT(J),J=N1,N2)
      IF(N2 .NE. NP) GO TO 33
      DO 38 J=1,NP
      X(J)=X(J)+XT(J)/NSAMP
      DO 36 K=1,J
   36 COVAR(J,K)=COVAR(J,K)+XT(J)*XT(K)
   38 CONTINUE
   40 CONTINUE
      DO 44 I=1,NP
      DO 42 J=1,I
      COVAR(I,J)=(COVAR(I,J)-X(I)*X(J)*NSAMP)/(NSAMP-1)
   42 COVAR(J,I)=COVAR(I,J)
   44 CONTINUE
      GO TO 70
   50 CONTINUE
      DO 55 I=1,NP
      N2=0
   51 N1=N2+1
      N2=N1+NCARD2-1
      IF(N2 .GT. I)N2=I
      READ(1,61)(COVAR(I,J),J=N1,N2)
      IF(N2 .NE. I) GO TO 51
      DO 52 J=1,I
   52 COVAR(J,I)=COVAR(I,J)
   55 CONTINUE
   70 SSR=0.
      NDF=0
      K=0
      DO 1000 I=1,NGROUP
      IF(METHOD(I) .GE. 2 .AND. METHOD(I) .LE. 3)NDF=NDF+1
      NCOUNT=K
      NPT=NMG(I)
      WB1=0.
      DO 500 J=1,NPT
      K=NCOUNT+J
      IG(K)=I
      NA=NAUX(I)
      IF(MDEG .LE. 0)GO TO 111
      IF(METHOD(I) .NE. 2)GO TO 110
      READ(1,62)ID(K),X(K)
      GO TO 220
```

```
110   READ(1,62)ID(K),X(K),(AUX(J,M),M=1,NA)
      GO TO 200
111   CONTINUE
      IF(METHOD(I) .NE. 2)GO TO 112
      READ(1,62)ID(K)
      GO TO 220
112   READ(1,62)IO(K),(AUX(J,M),M=1,NA)
200   WRITE(2,210)ID(K),X(K),IG(K),(AUX(J,M),M=1,NA)
      IF(METHOD(I) .LT. 4)GO TO 213
      XL(J)=X(K)
      GO TO 500
213   CONTINUE
      IF(METHOD(I) .EQ. 3)GO TO 215
      U(K)=AUX(J,1)
      GO TO 500
215   WB1=WB1+AUX(J,1)*X(K)
      GO TO 500
220   WRITE(2,210)ID(K),X(K),IG(K)
      WB1=WB1+X(K)/NPT
500   CONTINUE
      IF(METHOD(I) .EQ. 1)GO TO 1000
      IF(METHOD(I) .GE. 4)GO TO 900
      DO 800 J=1,NPT
      K=NCOUNT+J
800   U(K)=WB1
      GO TO 1000
900   CONTINUE
      NFORCE=0
      IF(METHOD(I) .EQ. 5)NFORCE=1
      NIND=NAUX(I)
      NDF=NDF+NIND+NFORCE
      CALL LSE(NFORCE,NPT,NIND,XL,AUX,B,XEST,SERQR)
      DO 950 J=1,NPT
      K=NCOUNT+J
950   U(K)=XEST(J)
1000  CONTINUE
      IF(MDEG .LE. 0)MDEG=NSAMP-1
      WRITE(2,1015)NSAMP,MDEG
      DO 1020 I=1,NP
      N2=0
      WRITE(2,1021)
1022  N1=N2+1
      N2=N1+NCARD2-1
      IF(N2 .GT. I)N2=I
      WRITE(2,1025) (COVAR(I,J),J=N1,N2)
      IF(N2 .NE. I) GO TO 1022
1020  CONTINUE
      NTRACE=NP-NDF
      IF(NTRACE .GT. 2)GO TO 1070
      WRITE(2,1500)NTRACE
      GO TO 9999
1070  NDEG=NP-1
      IF(MDEG .GT. NDEG)GO TO 2000
      WRITE(2,1505)NDEG
      GO TO 9999
2000  CONTINUE
      NNN=NP
      CALL MTXINV(NNN,COVAR,COINV)
      DO 2010 I=1,NP
      DO 2005 J=1,NP
2005  SSR=SSR+(X(I)-U(I))*COINV(I,J)*(X(J)-U(J))
2010  CONTINUE
      SSR=SSR*NSAMP
      XT(1)=(NTRACE-2)*MDEG
      FACTOR=1.-XT(1)/(MDEG-NP+3)/SSR
      IF(FACTOR .LT. 0.)FACTOR=0.
      WRITE(2,2200)
      DO 2100 J=1,NP
      XEST(J)=U(J)+FACTOR*(X(J)-U(J))
      WRITE(2,3200)ID(J),X(J),XEST(J)
```

```
2100 CONTINUE
      SSR=SSR-NTRACE
      IF(SSR .LE. O.)SSR=O.
      CALL EXPECT(NTRACE,SSR,EYY)
      MDEG=MDEG-NP+1
      PIMPRO=EYY*(NTRACE-2)**2*MDEG/(MDEG+2)/NP*100
      WRITE(2,6OOO)FACTOR,PIMPRO
      GO TO 1
9999 CONTINUE
      LOCK 2
      STOP
3     FORMAT(20I3)
5     FORMAT(///,1X)
10    FORMAT(1X,70A1)
15    FORMAT(/,1X,45H#############################################,
     125H#########################,/,1X,70A1,/,1X,13H#############,
     257H#########################################################)
30    FORMAT(//,1X,4HDATA,//,8X,12HSAMPLE GROUP,/,2X,
     141HIO    AVERAGE   ND  SUPPLEMENTAL VARIABLES,/,1X,
     250H--------------------------------------------------,
     320H--------------------)
60    FORMAT(10F7.2)
61    FORMAT(10F7.2)
62    FORMAT(I4,6F10.4)
210   FORMAT(1X,I4,F10.4,I4,1X,5F10.3)
1015 FORMAT(///,1X,42HSAMPLE SIZE (FOR SAMPLE AVERAGE)          =,I4,
     1/,1X,42HNO OF DEGREES OF FREEDOM (FOR COVARIANCE)=,I4,/,1X,
     245HTHE UNBIASED ESTIMATE OF COVARIANCE MATRIX IS,/)
1021 FORMAT(/)
1025 FORMAT(1X,10F7.2)
1500 FORMAT(//,1X,33H--- WOO ----- ERROR ----- WOO ---,/,2X,
     113HTRACE( P ) = ,I4,/,2X,28HTHE TRACE OF P SHOULD BE AT ,
     27HLEAST 2,/,2X,38HCHECK TO SEE WHETHER YOU HAVE TOO MANY,
     37H GROUPS,/,2X,34HOR/AND TOO MANY PAREMETERS IN THE ,
     416HREGRESSION LINES,/)
1505 FORMAT(//,1X,33H--- WOO ----- ERROR ----- WOO ---,/,2X,
     158HNO OF DEGREES OF FREEDOM FOR COVARIANCE SHOULD BE AT LEAST,I5)
2200 FORMAT(//,1X,17HESTIMATED RESULTS,//,2X,18HID  SAMPLE AVERAGE,
     113H NEW ESTIMATE,/,1X,33H---------------------------------)
3200 FORMAT(1X,I4,3X,F10.4,4X,F10.4)
6000 FORMAT(/,1X,35H SHRINKING FACTOR                  =,F8.4,/,1X,
     135H%-IMPROVEMENT OVER USUAL ESTIMATE =,F8.4,2H %,/)
      END

      SUBROUTINE LSE(MODEL,NPT,NIND,YDATA,XDATA,B,YEST,SEROR)
      DIMENSION YDATA(50),XDATA(50,5),B(50),X(50,50),YEST(50),
     1XX(50,50),XXIXT(50,50)
C
C     SUBROUTINE FOR COMPUTING THE LEAST SQUARES ESTIMATE
C     OF B IN THE LINEAR MODEL,
C
C             YDATA = XDATA * B       , IF MODEL = O
C     AND
C             YDATA = (I,XDATA) * B   , IF MODEL = 1
C
C     WHERE YDATA IS THE THE (NPT BY 1) VECTOR OF DEPENDENT
C     OBSERVATIONS, XDATA IS THE (NPT BY NIND) MATRIX OF
C     INDEPENDENT OBSERVATIONS, I IS THE COLUMN VECTOR WITH
C     EVERY ELEMENT EQUAL TO ONE, * STANDS FOR THE PRODUCT
C     OF TWO MATRICES, NIND IS THE NUMBER OF INDEPENDENT
C     VARIABLES AND B IS THE VECTOR OF PARAMETERS TO BE
C     ESTIMATED.
C
      IF(MODEL .EQ. O)GO TO 100
      DO 10 I=1,NPT
10    X(I,1)=1.
      MM=NIND+1
      DO 20 I=2,MM
      II=I-1
      DO 30 J=1,NPT
```

```
30    X(J,I)=XDATA(J,II)
20    CONTINUE
      GO TO 200
100   MM=NIND
      DO 120 I=1,MM
      DO 130 J=1,NPT
130   X(J,I)=XDATA(J,I)
120   CONTINUE
200   CONTINUE
      DO 230 I=1,MM
      DO 240 J=1,MM
      XX(I,J)=0.
      DO 250 K=1,NPT
250   XX(I,J)=XX(I,J)+X(K,I)*X(K,J)
240   CONTINUE
230   CONTINUE
      CALL MTXINV(MM,XX,XX)
      DO 300 I=1,MM
      DO 310 J=1,NPT
      XXIXT(I,J)=0.
      DO 320 K=1,MM
320   XXIXT(I,J)=XXIXT(I,J)+XX(I,K)*X(J,K)
310   CONTINUE
300   CONTINUE
      DO 400 I=1,MM
      B(I)=0.
      DO 410 J=1,NPT
410   B(I)=B(I)+XXIXT(I,J)*YDATA(J)
400   CONTINUE
      SEROR=0.
      DO 500 I=1,NPT
      YEST(I)=0.
      DO 510 J=1,MM
510   YEST(I)=YEST(I)+X(I,J)*B(J)
      SEROR=SEROR+(YDATA(I)-YEST(I))**2
500   CONTINUE
      SEROR=SQRT(SEROR/(NPT-MM))
      RETURN
      END

      SUBROUTINE MTXINV(NSIZE,W,WINV)
      DIMENSION ARRAY(50,50),WINV(50,50),W1(50,2),W(50,50)
C
C     SUBROUTINE FOR FINDING THE INVERSE OF AN (NSIZE BY NSIZE)
C     SQUARE MATRIX W BY USING THE PARTITION METHOD. WINV IS THE
C     INVERSE MATRIX OF W.
C
      DO 5 I=1,NSIZE
      DO 5 J=1,NSIZE
5     ARRAY(I,J)=W(I,J)
      IF(NSIZE .GT. 1) GO TO 10
      WINV(1,1)=1./ARRAY(1,1)
      RETURN
10    CONTINUE
      MSIZE=NSIZE-1
      DO 15  II=1,MSIZE
      J=II+1
      DO 16  KK=J,NSIZE
      DO  17 M=1,NSIZE
      W1(M,1)=W(M,II)
17    W1(M,2)=W(M,KK)
      DET=W1(1,1)*W1(2,2)-W1(1,2)*W1(2,1)
      IF(DET .EQ. 0.) GO TO 16
      IF(II .EQ. 1 .AND. KK .EQ. 2) GO TO 19
      DO  18 K=1,NSIZE
      ARRAY(K,1)=W(K,II)
      ARRAY(K,2)=W(K,KK)
      ARRAY(K,II)=W(K,1)
18    ARRAY(K,KK)=W(K,2)
      GO TO 19
```

```
 16    CONTINUE
 15    CONTINUE
 19    CONTINUE
       WINV(1,1)=ARRAY(2,2)/DET
       WINV(2,2)=ARRAY(1,1)/DET
       WINV(1,2)=-ARRAY(1,2)/DET
       WINV(2,1)=-ARRAY(2,1)/DET
       IF(NSIZE .EQ. 2) GO TO 100
       DO 20 I=3,NSIZE
       K=I-1
       DO 21 J=1,K
       W1(J,1)=0.
       W1(J,2)=0.
       DO 22 M=1,K
       W1(J,1)=W1(J,1)+WINV(J,M)*ARRAY(M,I)
 22    W1(J,2)=W1(J,2)+ARRAY(I,M)*WINV(M,J)
 21    CONTINUE
       ELTA=ARRAY(I,I)
       DO 23 J=1,K
 23    ELTA=ELTA-ARRAY(I,J)*W1(J,1)
       WINV(I,I)=1./ELTA
       DO 24 J=1,K
       WINV(J,I)=-W1(J,1)/ELTA
       WINV(I,J)=-W1(J,2)/ELTA
       DO 24 M=1,K
 24    WINV(J,M)=WINV(J,M)+W1(J,1)*W1(M,2)/ELTA
 20    CONTINUE
 100   CONTINUE
       IF(II .EQ. 1 .AND. KK .EQ. 2) GO TO 888
       DO 401 J=1,NSIZE
       W1(J,1)=WINV(1,J)
 401   W1(J,2)=WINV(2,J)
       DO 402 J=1,NSIZE
       WINV(1,J)=WINV(II,J)
 402   WINV(2,J)=WINV(KK,J)
       DO 403 J=1,NSIZE
       WINV(II,J)=W1(J,1)
 403   WINV(KK,J)=W1(J,2)
 888   CONTINUE
       RETURN
       END

       SUBROUTINE EXPECT(NP,THETA,EYY)
       DOUBLE PRECISION P1,A,EY,P2
C
C      THIS SUBROUTINE COMPUTES THE EXPECTATION OF 1/Y, WHERE Y IS A
C      NONCENTRAL CHI-SQUARE WITH NP DEGREES OF FREEDOM AND NONCENTRALITY
C      PARAMETER THETA. THIS EXPECTATION IS THE SAME AS THE EXPECTATION
C      OF 1/(NP-2+2W), WHERE W IS A POISSON WITH PARAMETER (THETA/2).
C
       ERROR=0.00001
       ERR=ERROR/NP
       A=THETA/2
       P1=DEXP(-A)
       EY=P1/(NP-2)
       P2=P1
       K=0
 10    K=K+1
       P1=P1*A/K
       P2=P2+P1
       EY=EY+P1/(NP-2+K*2)
       CHECK=1.-P2
       IF(CHECK.GE.ERR) GO TO 10
       EYY=EY
       RETURN
       END
```

REFERENCES

1. James, W. and Stein, C., "Estimation with Quadratic Loss," Proceedings, Fourth Berkeley Symposium Math. Statist. Prob. 1, pp. 361-379, 1961.

2. Effron, B. and Morris, C., "Data Analysis Using Stein Estimation and Its Generalization," OEO Research Report R-1394, March 1974.

3. Effron, B. and Morris, C., "Stein's Rule and Its Competitors - An Empirical Bayes Approach," Journal of American Stat. Assoc., Vol. 68, No. 341, March 1973.

4. Kuo, W. H., "An Accident Frequency Prediction Model for Selected State Trunkline Classifications," Michigan Department of Transportation, Research Report R-1068, December 1977.

5. Michaels, R. M., "Two Simple Techniques for Determining the Significance of Accident-Reducing Measures," Public Roads, A Journal of Highway Research, Vol. 30, No. 10, pp. 238-239, October 1959.

6. Gerlough, D. L. and Schuhl, A., "Poisson and Traffic," Eno Foundation for Highway Traffic Control, Saugatuck, CN, p. 75, 1955.

7. Anscombe, R. J., "The Transformation of Poisson, Binomial and Negative-Binomial Data," Biometrika, Vol. 35, pp. 246-254, 1948.

8. Ross, S. M., "Introduction to Probability Models," Academic Press, New York, 1972.

9. A User's Guide to Positive Guidance, U. S. Department of Transportation, Federal Highway Administration, Office of Traffic Operation, June 1977.

10. Glennon, J. C., "Roadside Safety Improvement Programs on Freeways - A Cost Effectiveness Approach," Transportation Research Board, NCHRP Report 148, pp. 10-24, 1974.

11. Michigan Traffic Accident Facts, Prepared by the Michigan Department of State Police.

12. Holbrook, L. F., "Accident Rates and Surface Properties - An Investigation of Relationships," Michigan Department of Transportation, Research Report R-994, October 1976.

13. "Noise Prediction Model, STAMINA 1.0," Federal Highway Administration, Office of Research and Development, Washington, Research Report FHWA-RD-78-138, May 1979.

14. Kuo, W. H., "Precisions of the Aggregate Sample Splitter and Testing Method," Michigan Department of Transportation, Research Report R-1133, November 1980.

15. Kuo, W. H., "Aggregate Gradation Quality Control," Michigan Department of Transportation, Research Report R-1024, November 1976.

16. Stein, C., "Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution," Third Berkeley Symposium Math. Statist. Prob. 1, p. 197, 1956.

17. Stein, C., "Confidence Sets for the Mean of a Multivariate Normal Distribution," J. Royal Statist. Soc. B, Vol. 24, pp. 265-296, 1962.

18. Hudson, H. M., "A Natural Identity for Exponential Families with Application in Multivariate Estimation," Ann. Statist., Vol. 6, No. 3, pp. 473-484, 1978.