Income Distribution Within Groups, Among Groups, and Overall:
A Technique of Analysis

by

Sherman Robinson

NOTE: Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. References in publication to Discussion Papers should be cleared with the author(s) to protect the tentative character of these papers. The papers express the views of the author and should not be interpreted to reflect those of the World Bank.
Income Distribution Within Groups, Among Groups, and Overall:
A Technique of Analysis

by
Sherman Robinson

Research Program in Development Studies
Woodrow Wilson School
Princeton University
Princeton, New Jersey
August, 1976
Discussion Paper No. 65

Financial support for this work was provided by the World Bank through the Development Research Center and the Economic Analysis and Projections Department. The Bank does not necessarily agree with any views expressed in this paper.

This work was done as part of a joint project with Irma Adelman to build a model of income distribution in Korea. My thanks to her and also to Gregory Chow for discussions and assistance.

Note: Discussion papers of the Research Program in Development Studies are preliminary material circulated to stimulate discussion and critical comment. Please do not refer to discussion papers without permission of the author.
### TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>II. Aggregating Group Distributions</td>
<td>3</td>
</tr>
<tr>
<td>III. Distribution Statistics</td>
<td>5</td>
</tr>
<tr>
<td>Mean and Variance of Income</td>
<td></td>
</tr>
<tr>
<td>Quantiles of the Distribution</td>
<td></td>
</tr>
<tr>
<td>Lorenz Curve and Gini Coefficient</td>
<td></td>
</tr>
<tr>
<td>Other Distribution Statistics</td>
<td></td>
</tr>
<tr>
<td>IV. Conclusion</td>
<td>15</td>
</tr>
</tbody>
</table>

#### Appendices

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. User's Guide to the Income Distribution Program</td>
<td>16</td>
</tr>
<tr>
<td>Sample Output</td>
<td>19</td>
</tr>
<tr>
<td>B. Programmer's Guide</td>
<td>25</td>
</tr>
<tr>
<td>C. Program Listing</td>
<td>27</td>
</tr>
</tbody>
</table>
I. Introduction

The purpose of this paper is to describe a technique and an associated computer algorithm for deriving the overall distribution of income (the size distribution) given data on the distribution of income between and within aggregate groups in the economy. In particular, the technique can be used to nap from the distribution of income by major factors of production such as labor, capital, and land (the functional distribution) into the overall size distribution. It can also be used with other group definitions such as by regions, sector of production, or socioeconomic categories. Any convenient definitions of groups will do so long as they are complete and mutually exclusive.

The technique was developed for an economy-wide model of income distribution in Korea and has also been applied in a very different model of socio-economic mobility and income distribution. See Adelman and Robinson (1976) and Robinson and Dervis (1974). Related approaches have been used by others -- see Rodgers, Hopkins and Wéry (1976) and Thorbecke and Sengupta (1972). The approach is clearly useful in economic models since it provides a way to take data economy-wide models usually generate such as factor incomes and, with some additional information, generate the overall income distribution.

In addition to its usefulness in models, the basic approach should also be very useful in the preparation of income distribution data for a country. It provides a framework within which one can integrate income distribution data from a number of different sources in order to generate the overall distribution. Using this technique, one can divide
the economy into any convenient complete set of disparate groups and then use data on each separate group to determine its within-group distribution. The data sources used to analyze each separate group might be entirely different. For example, one might use agricultural surveys for rural groups, household surveys for urban workers, tax data for the rich, and so forth. National accounts or input-output data can be used to analyze the between-group distribution and then the algorithm can be used to generate the overall size distribution. It thus can provide a method for reconciling income distribution data from a number of different sources.
II. Aggregating Group Distributions

The technique consists simply of numerically aggregating a number of different within-group income distributions whose functional form and parameters are known. The resulting aggregate distribution is thus completely specified numerically, even though it may not be able to be described by any convenient summary distribution function. All statistics of the aggregate distribution can be generated numerically.

Assume that there are \( n \) different groups and that the distribution of income within each group is given by the probability distribution function:

\[
f_i(y|\theta_i) \quad i = 1, \ldots, n
\]

where \( \theta_i \) is a vector of parameters for each distribution and \( f_i \) reflects the functional form for the distribution within group \( i \) (such as lognormal, Pareto, etc.). The overall distribution of income is given by:

\[
f(y|\theta) = \sum_{i=1}^{n} w_i f_i(y|\theta_i)
\]

where the weight \( w_i \) is the population share of group \( i \) and \( \theta \) is the set of all parameters \( \theta_i \) for all \( i \). Note that since \( 0 \leq w_i \leq 1 \) and \( \sum w_i = 1 \), the function \( f(y|\theta) \) is indeed a probability distribution function. It is simply the weighted average of the separate within-group functions. Note also that since \( f(y|\theta) \) is a sum of distributions and not the distribution of a sum of random variables, central limit theorems do not apply. The distribution \( f(y|\theta) \) may in principal have any shape and be completely intractable analytically.

Given complete knowledge of \( w_i \) and \( \theta_i \), one can generate the overall distribution \( f(y|\theta) \) numerically. While any statistics for the overall distribution are clearly a function of the parameter set \( \theta \) and the population shares, \( w_i \), it is possible to generate them numerically without attempting to deal analytically with the overall distribution function.
For the analysis of income distribution, a number of statistics are of interest which cannot be solved analytically but which must be calculated numerically. There are, of course, special cases of the distribution functions $f_i(y|\theta_i)$ which can yield a convenient and tractable form for the overall distribution function $f(y|\theta)$. The algorithm described here, however, always treats the overall distribution function numerically.

In the computer program described in the appendix, all the $f_i(y|\theta_i)$ distributions are specified as two parameter lognormal distributions. The lognormal distribution is commonly used to represent income distributions and seems to fit income distribution data very well.\(^1\) Note, however, that it is used here only to represent the within-group distribution. The overall distribution is thus a sum of lognormal distributions and may have any shape.\(^2\)

---

\(^1\) See Aitchison and Brown (1957), chapter 11.

\(^2\) If the within-group log variances are all the same and if the group mean incomes are distributed lognormally, then the overall distribution is also lognormal. See Aitchison and Brown (1957), p. 110. Usually, however, it is not reasonable to make either assumption when dealing with the functional distribution.
III. Distribution Statistics

Mean and Variance of Income

Assume that the mean income of each group is given by $\bar{y}_i$. Then the overall mean income is:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} w_i \bar{y}_i$$

The overall variance of income, $s^2$, can be calculated from the decomposition of variance formula:

$$s^2 = \frac{1}{n} \sum_{i=1}^{n} w_i s_i^2 + \frac{1}{n} \sum_{i=1}^{n} w_i (\bar{y}_i - \bar{y})^2$$

where $s_i^2$ is the variance of incomes within group $i$ (assuming it is finite).

The variance of the logarithms of income is a commonly used inequality measure and is especially interesting under the assumption of lognormality since it is one of the parameters of the distribution. Denote the two parameter lognormal cumulative distribution function for group $i$ as $F_i$. Under the lognormal distribution, the variable $x = \log y$ is distributed normally with mean $\mu_i$ and variance $\sigma_i^2$:

$$f_i(y|\theta_i) \ dy = dN(y|\mu_i, \sigma_i^2) = dN(x|\mu_i, \sigma_i^2)$$

when $y$ is distributed lognormally, the within-group mean and variance are given by:

$$\bar{y}_i = \exp (\mu_i + \frac{1}{2} \sigma_i^2)$$

$$s_i^2 = [\exp (2 \mu_i + \sigma_i^2)] [\exp (\sigma_i^2) - 1]$$

---

3 See Aitchison and Brown (1957), p. 8.
The overall mean and variance are found by substituting (4) and (5) into (2) and (3). The overall log mean and log variance are given by:

\[
\mu = \sum_{i=1}^{n} \omega_i \mu_i \tag{6}
\]

\[
\sigma^2 = \sum_{i=1}^{n} \omega_i \sigma_i^2 + \sum_{i=1}^{n} \omega_i (\mu_i - \mu)^2 \tag{7}
\]

There is, of course, no presumption that the overall distribution is log-normal with parameters \( \mu \) and \( \sigma^2 \). Indeed, if one wished to fit a lognormal approximation to the overall distribution, it is not clear that these equations would even be a good way to estimate the parameters of the lognormal approximation. Equation (7), however, is the appropriate formula for calculating the overall log variance for use as an inequality measure.

**Quantiles of the Distribution**

Income distributions are commonly tabulated by quantiles such as quartiles or deciles. Quantiles are defined as the income ranges containing a specified share of the overall distribution, starting from zero. For example, the first decile is defined as the value of \( q_1 \) for which:

\[
\int_{0}^{q_1} f(y|g) \, dy = 0.1
\]

The second decile, \( q_2 \), is defined as:

\[
\int_{q_1}^{q_2} f(y|g) \, dy = 0.1
\]

and so forth to the last decile:

\[
\int_{q_9}^{q_{10}} f(y|g) \, dy = 0.1
\]

It requires nine quantile values to delineate ten decile income ranges.
For the lognormal distribution, the quantiles are easily calculated from the normal distribution. Define \( q^i_k \) as the \( k \)'th quantile for group \( i \) where group \( i \) has a within-group distribution which is lognormal. Define \( p^i_k \) as the share of the density associated with the choice of quantile (say, all equal 0.1 for deciles). The group quantiles can be defined iteratively as:

\[
\int_{q^i_{k-1}}^{q^i_k} \frac{1}{\sigma^i_k \sqrt{2\pi}} e^{-\frac{(\log y - \mu^i_k)^2}{2\sigma^2_k}} \, dy = p^i_k \quad k = 1, \ldots, 9
\]

where \( q^i_0 = 0 \). By a simple change of variable, these quantiles can be expressed in terms of the normal distribution. Let \( x = \log y \) and \( x^i_k = \log q^i_k \). Then

\[
\int_{x^i_{k-1}}^{x^i_k} \frac{1}{\sigma^2_k \sqrt{2\pi}} e^{-\frac{(x - \mu^i_k)^2}{2\sigma^2_k}} \, dx = p^i_k \quad k = 1, \ldots, 9
\]

and \( x^i_0 = -\infty \). Equation (9) is easily solved for the quantiles using standard algorithms for calculating the inverse normal integral.

The quantiles of the overall distribution, however, are not so easily calculated. It is necessary to solve the following equation for \( q^k_k \) given \( q^k_{k-1} \) and \( p^k_k \):

\[
\int_{q^k_{k-1}}^{q^k_k} f(y|\theta) \, dy = p^k_k \quad k = 1, \ldots, 9
\]

where \( q^k_0 = 0 \). There is certainly no simple relationship between \( q^k_k \) and the corresponding within-group quantiles \( q^i_k \). Note, however, that given \( q^k_{k-1} \) and \( q^k_k \), it is easy to use equation (10) to solve for \( p^k_k \) by simply summing the population shares within each group between the two incomes.
Equation (10) can be written as a non-linear algebraic equation in $q_k$:

\begin{equation}
Q(q_k) = \int_{q_{k-1}}^{q_k} f(y|\theta) \, dy - p_k = 0
\end{equation}

which can be evaluated and whose root we seek. This equation can be solved by a number of different techniques. In the computer program described in the appendix, a variant of Newton's method is used in which the derivative of $Q(q_k)$ is calculated numerically. An initial guess for $q_k$ is calculated by taking the weighted geometric mean of the corresponding within-group quantiles, $q_k$. The technique always converged within 5-16 iterations.

**Mean Incomes of Quantiles**

For both the within-group and overall distributions, it is interesting to know the mean income of people falling within given income ranges -- for example, the mean income of those in the lowest decile. Specify two incomes $q_{k-1}$ and $q_k$ corresponding to some quantile of the overall distribution. The mean income of those people whose income falls within the range $q_{k-1} \leq y \leq q_k$ is simply the weighted average of the corresponding mean incomes of those within each group whose incomes are in the same range. Define the mean income of people within some quantile of the overall distribution as:

\begin{equation}
\overline{y}_{1...k} = \frac{1}{p_k} \int_{q_{k-1}}^{q_k} y f(y|\theta) \, dy
\end{equation}

where $q_0 = 0$, and $p_k$ is the share of the density defined by the quantile range:

\begin{equation}
p_k = \int_{q_{k-1}}^{q_k} f(y|\theta) \, dy
\end{equation}

\(^{4}\)See Jarratt (1970) for a survey of the techniques available for solving such equations.
Similarly, for each group:

\begin{equation}
\hat{m}_k^i = \frac{1}{p_k} \int q_k^{q_{k-1}} y f_i(y|\theta) \, dy
\end{equation}

where \(q_{k-1}\) and \(q_k\) define a quantile of the overall distribution, \(q_0 = 0\) and

\begin{equation}
p_i^k = \int q_k^{q_{k-1}} f_i(y|\theta) \, dy
\end{equation}

It follows from the definition of \(f(y|\theta)\) that:

\begin{equation}
\hat{m}_k = \sum_{i=1}^{n} \hat{m}_k^i
\end{equation}

where

\[
W_i = \frac{W_{i \cdot \theta_i}}{\sum_j W_{j \cdot \theta_j}}
\]

Thus, to calculate the mean incomes of those falling within a specified quantile of the overall distribution defined by \(q_{k-1}\) and \(q_k\), one must calculate the mean income of those in each subgroup who fall within the same range.

For a lognormal distribution, it is possible to solve for \(\hat{m}_k^i\) analytically. Equation (14) becomes:

\begin{equation}
\hat{m}_k^i = \frac{1}{p_k} \int q_k^{q_{k-1}} y \, d\Lambda(y|u_1, \sigma^2_1)
\end{equation}

\(q_0 = 0\), and:

\begin{equation}
p_i^k = \int q_k^{q_{k-1}} \, d\Lambda(y|u_1, \sigma^2_1)
\end{equation}

This integral can be solved in terms of the normal integral. Change variables:

\[
x = \log y, \quad y = e^x
\]

and

\[
dx = \left( \frac{1}{y} \right) dy, \quad dy = y \, dx
\]

Thus \(x_k = \log q_k\) and \(x_0 = -\infty\).
From the definition of the lognormal distribution, equation (18) becomes:

\[ P_k = \int_{x_{k-1}}^{x_k} dN(x | \mu, \sigma^2) \]

Equation (17) can also be written in terms of the normal distribution. Again, from the definition of the lognormal distribution,

\[ d\Lambda(y | \mu, \sigma^2) = \frac{1}{y \sigma \sqrt{2\pi}} \exp\left[-\frac{1}{2} \left( \frac{\log(y) - \mu}{\sigma} \right)^2 \right] dy \]

Multiply both sides by \( y \), integrate between \( q_{k-1} \) and \( q_k \), change variables on the right hand side, and note that \( y = e^x \). The result is:

\[ \int_{q_{k-1}}^{q_k} y \cdot d\Lambda(y | \mu, \sigma^2) = \frac{1}{\sigma \sqrt{2\pi}} \int_{x_{k-1}}^{x_k} \exp \left[ x - \frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2 \right] dx \]

Consider the term being exponentiated, \( \exp \left[ -x \right] \). Multiply it out and complete the square. The term becomes:

\[ \frac{1}{2} \left( \frac{x - (\mu + \sigma^2)}{\sigma} \right)^2 + (\mu + \frac{1}{2} \sigma^2) \]

Note that the second term does not involve \( x \) and can be factored out of the integral. Furthermore, from equation (4), \( \exp \left( \mu + \frac{1}{2} \sigma^2 \right) = \bar{y}_i \). Thus the integral can be written:

\[ \bar{y}_i \cdot \frac{1}{\sigma \sqrt{2\pi}} \int_{x_{k-1}}^{x_k} \exp \left\{ -\frac{1}{2} \left[ \frac{x - (\mu + \sigma^2)}{\sigma} \right]^2 \right\} dx \]

The integral can be seen to be a normal integral for a variable with mean \( (\mu + \sigma^2) \) and variance \( \sigma^2 \). Thus, the quantile mean in equation (17) equals:

\[ m_k = \frac{\int_{x_{k-1}}^{x_k} dN(x | \mu, \sigma^2)}{\int_{x_{k-1}}^{x_k} dN(x | \mu, \sigma^2)} \bar{y}_i \]
If the within-group distributions are lognormal, equation (20) can be used to calculate the mean income of those with incomes in a specified range for each group. Equation (16) can then be used to calculate the mean income of those with incomes in the specified range for the overall distribution. This is the method used in the computer program described in the appendix.

Lorenz Curve and Gini Coefficient

A common way to present income distribution data is to plot the proportion of income receivers having income less than $y$ along the horizontal axis (this is the cumulative distribution function) against the proportion of total income going to the same income receivers on the vertical axis. The resulting graph is a Lorenz curve, an example of which is given below.

If the distribution of income is perfectly equal, the Lorenz curve is the diagonal straight line. The more unequal is the distribution, the more the curve bows out from the diagonal and the greater is the shaded area. The Gini coefficient is a measure of income inequality based on the Lorenz curve and is equal to the ratio of the shaded area to the area
of the triangle under the diagonal. Formally, for the overall distribution, the definition of the measure is:

$$G = 1 - 2 \int_0^\infty \phi(y) dF(y)$$

where:

$$F(y) = \int_0^y f(t | \theta) dt, \text{ and}$$

$$\phi(y) = \frac{1}{2} \frac{\partial}{\partial y} \left[ \int_0^y t f(t | \theta) dt / \int_0^\infty t f(t | \theta) dt \right]$$

$F(y)$ is simply the cumulative distribution function. Given that one can calculate the mean income of people falling within any income range, then it is straightforward to calculate $\phi(y)$ numerically. Compare equations (12) and (23). The Lorenz diagram simply plots $\phi(y)$ on the vertical axis against $F(y)$ on the horizontal axis.

For the within-group distributions, which are assumed to be log-normal, the Gini coefficients can be derived analytically. They are given by:

$$G_i = 2 \cdot N(\frac{\sigma_i}{\sqrt{2}}) |0,1| - 1 \quad i = 1, \ldots, n$$

For the overall distribution, equation (21) must be evaluated numerically. In the computer program presented in the appendices, the calculation is done in the following steps. First, calculate quantiles (usually deciles) of the overall distribution and the corresponding incomes $q_k$. Second, calculate the mean incomes of people falling within

---


6 See Aitchison and Brown (1957), p. 112.
each quantile. Third, using these means, calculate \( \hat{F}(q_k) \) for all the quantiles. The points \( F(q_k) \) and \( \hat{F}(q_k) \) fall on the Lorenz curve. Since the quantiles are chosen to contain equal proportions of the population (say, deciles), then the points \( F(q_k) \) are equally spaced along the horizontal axis. Fourth, integrate the Lorenz curve numerically based on the calculated points.

It is obvious that if the integration is done by simply connecting the points with straight lines, then the degree of inequality will be underestimated (as the diagram below shows).

In the computer program described in the appendices, the straight line approximation is not used. Instead, an equal-interval quadrature method called Simpson's rule is used which fits a polynomial curve to successive sets of points and then integrates the fitted curve. The technique was tested by calculating the Gini coefficient numerically for a lognormal distribution (based on decile points) and comparing it with the value calculated from equation (24). The values are extremely close, within 0.5 percent of the actual value.

\(^7\) See Arden and Astill (1970), pp. 82-84.
Other Distribution Statistics'.

From the previous equations, one can generate empirically a complete description of all the within-group and overall distributions as well as analyze the group composition of the over/distribution. Using this approach, it is interesting to calculate the composition by groups of quantiles of the overall distribution -- for example, the share of the people in the bottom decile who are rural. It is also interesting to analyze the share of particular groups in quantiles of the over/distribution -- for example, the share of rural workers in the first decile of the overall distribution.

Given that the distribution can be generated empirically, it is easy to compute any aggregate measures of inequality based on it. A number of different measures have been used, generally based either on the frequency function or on the Lorenz curve. For a survey of such measures, see Sen (1973) or Szal and Robinson (1977). Certain aggregate statistics can be decomposed into within-group and between-group contributions. Equations (3) and (7) provide such a decomposition for the variance and log variance. Decompositions have also been devised for the Gini coefficient as well as for other measures.  

---

8 See Sen (1973) and Szal and Robinson (1977).
IV. Conclusion

In the approach described here, the overall income distribution is generated given knowledge about the level and distribution of income within groups in the economy. Any convenient definition of groups can be used so long as it is complete and exclusive. The collection of groups must include the entire population and there can be no overlapping among groups. The approach can be used by model builders who face the problem that economy-wide economic models usually do not generate the size distribution but only the functional distribution. The approach can also be used as a framework for integrating and reconciling income distribution data from a variety of different sources.

In the computer program described in the appendices, all the within-group distributions are given by two-parameter lognormal distributions. The general approach described above certainly does not require this assumption and it would be possible to specify a completely different distribution for each group. The technique for computing the overall distribution does not depend on any particular specification of the within-group distributions.
Appendix A
User's Guide to the Income Distribution Program

This appendix describes how to use the program to aggregate a set of within-group income distributions. The program requires as input some control parameters: group names and a heading for labelling the output; and data on the population, mean income, and log variance for all the groups. In addition, the user may specify a set of absolute income ranges which will be used to generate an analysis of the within-group and overall distributions in the specified ranges.

The input data for each job are summarized in the following table, followed by a detailed discussion. As many jobs as desired say be done in the same run by simply including as many sets of parameter cards as desired, one for each run. After the last set of parameter cards, the user must include one card with the word END in columns 1-3.

Input Data for Each Job

<table>
<thead>
<tr>
<th>Number of cards</th>
<th>Variable</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HDG</td>
<td>A80</td>
</tr>
<tr>
<td>NG/8</td>
<td>NG, NQ, NCL, NCODE</td>
<td>413</td>
</tr>
<tr>
<td>NG</td>
<td>XNAME</td>
<td>8A10</td>
</tr>
<tr>
<td>1</td>
<td>Y, ( \sigma^2 ), w</td>
<td>3F10.0</td>
</tr>
<tr>
<td>1</td>
<td>YCLS</td>
<td>4F10.0</td>
</tr>
</tbody>
</table>

HDG: a heading which will be printed at the top of each page of output. It can be up to 80 characters long (1 card). It must not have END as the first three characters.
SG: **Number** of groups. Must be less than 34.

NQ: **Number** of quantiles desired. Usually set to 10 (deciles). A separate analysis of the top five percent and one percent of the distribution will be generated in any case.

NCL: **Number** of specified income ranges for separate analysis. If NCL > 1, then the user must provide NCL - 1 values of YCLS to define the income ranges. If NCL ≤ 1, then no values of YCLS will be read and the YCLS card must be omitted. Note that it requires SCL - 1 incomes to define NCL quantile ranges. The bottom range is assumed to be from zero to the first value of YCLS and the top range is from the last value of YCLS to infinity. NCL must be less than or equal to 5.

NCODE: An integer code indicating the form in which the log variances are to be read in. If NCODE = 1, the log standard deviations, a, must be provided. If NCODE = 0, then the log variances, a^2, must be provided.

XNAME: Group names, up to 10 characters each. NG such names must be provided and will be read in 8 names per card, using as many cards as necessary.

Y: Mean income for each group.

σ^2: Log variance for each group. If NCODE = 1, the log standard deviation σ, must be provided instead.

w: The population shares of each group. If desired, the absolute populations may be provided instead -- the program will automatically normalize the shares so they sum to one.

Y, σ^2, w: One card for each group (format 3F10.0). NG cards in all.

YCLS: NCL - 1 values of income to define SCL income ranges. If NCL ≤ 1, this card must be omitted.
To run multiple jobs in a single run, simply stack as many sets of parameter cards as desired, one for each run. At the end of the last set of parameter cards, the user must add one card with the word END in the first three columns.

The input cards for a sample program are listed below followed by a listing of the output from the program.

**Sample Job**

<table>
<thead>
<tr>
<th>TEST RUN FOR DISTRIBUTION</th>
<th>PROGRAM 8/9/76</th>
</tr>
</thead>
<tbody>
<tr>
<td>310 4 0</td>
<td></td>
</tr>
<tr>
<td>RICH MIDDLE POOR</td>
<td></td>
</tr>
<tr>
<td>1000. 0.45 15.0</td>
<td></td>
</tr>
<tr>
<td>300. 0.32 35.0</td>
<td></td>
</tr>
<tr>
<td>90. 0.50 50.0</td>
<td></td>
</tr>
<tr>
<td>50.0 200.0 800.0</td>
<td></td>
</tr>
<tr>
<td>END</td>
<td></td>
</tr>
</tbody>
</table>
## Mean Incomes of Quantiles

**Quantiles of 10.00 Percent.**

**Cll 11 is Overall Group Mean Income**

<table>
<thead>
<tr>
<th>Column</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rich</td>
<td>254.47</td>
<td>197.64</td>
<td>507.89</td>
<td>616.93</td>
<td>734.68</td>
<td>869.99</td>
<td>10.36.29</td>
<td>1250.14</td>
<td>1615.02</td>
<td>2706.88</td>
<td>1000.00</td>
</tr>
<tr>
<td>Middle</td>
<td>97.07</td>
<td>141.94</td>
<td>174.51</td>
<td>205.60</td>
<td>238.26</td>
<td>274.76</td>
<td>318.33</td>
<td>375.50</td>
<td>462.79</td>
<td>711.11</td>
<td>300.00</td>
</tr>
<tr>
<td>Poor</td>
<td>21.03</td>
<td>33.62</td>
<td>43.51</td>
<td>53.41</td>
<td>64.20</td>
<td>76.73</td>
<td>92.26</td>
<td>113.39</td>
<td>147.30</td>
<td>254.55</td>
<td>90.00</td>
</tr>
<tr>
<td>Overall</td>
<td>27.32</td>
<td>48.33</td>
<td>69.43</td>
<td>96.22</td>
<td>132.36</td>
<td>180.59</td>
<td>245.26</td>
<td>341.40</td>
<td>528.75</td>
<td>1330.26</td>
<td>300.00</td>
</tr>
</tbody>
</table>

Cols 1 - 10 are shares of income by quantiles of 10.00 percent

Cll 11 is the Gini Coefficient assuming lognormality

Cll 12 is the Gini Coefficient calculated numerically

Cll 13 is the log variance

Cll 14 is the geometric mean income, expressed

Cll 15 is mean income

Cll 16 is group shares in total population

Cll 17 is group shares in aggregate income

Percent of total log variance due to within group variance = 34.790
### Percent Composition by Groups in Quantiles of Overall Distribution

Columns 100 to One Hundred

<table>
<thead>
<tr>
<th>Column</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Rich</td>
<td>0.00</td>
<td>0.01</td>
<td>0.04</td>
<td>0.21</td>
<td>0.81</td>
<td>2.39</td>
<td>5.96</td>
<td>14.55</td>
<td>39.40</td>
<td>86.62</td>
</tr>
<tr>
<td>2 Middle</td>
<td>0.15</td>
<td>1.42</td>
<td>6.03</td>
<td>18.07</td>
<td>39.42</td>
<td>61.82</td>
<td>75.40</td>
<td>76.86</td>
<td>57.74</td>
<td>13.09</td>
</tr>
<tr>
<td>3 Poor</td>
<td>99.85</td>
<td>98.57</td>
<td>93.93</td>
<td>81.72</td>
<td>59.77</td>
<td>35.79</td>
<td>18.65</td>
<td>8.58</td>
<td>2.85</td>
<td>0.28</td>
</tr>
<tr>
<td>4 Overall</td>
<td>38.63</td>
<td>58.26</td>
<td>81.55</td>
<td>152.48</td>
<td>154.24</td>
<td>209.53</td>
<td>285.74</td>
<td>409.56</td>
<td>700.34</td>
<td>0.0</td>
</tr>
</tbody>
</table>

### Percent Distribution of Groups in Quartiles of Overall Distribution

<table>
<thead>
<tr>
<th>Column</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Rich</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td>0.14</td>
<td>0.54</td>
<td>1.59</td>
<td>3.97</td>
<td>9.70</td>
<td>26.27</td>
<td>57.75</td>
</tr>
<tr>
<td>2 Middle</td>
<td>0.04</td>
<td>0.47</td>
<td>1.72</td>
<td>5.16</td>
<td>11.26</td>
<td>17.66</td>
<td>21.54</td>
<td>21.96</td>
<td>16.50</td>
<td>3.74</td>
</tr>
<tr>
<td>3 Poor</td>
<td>19.97</td>
<td>19.71</td>
<td>18.79</td>
<td>16.34</td>
<td>11.95</td>
<td>7.16</td>
<td>3.73</td>
<td>1.72</td>
<td>0.57</td>
<td>0.06</td>
</tr>
</tbody>
</table>
TEST RUN FOR DISTRIBUTION PROGRAM 8/9/76

ANALYSIS OF TOP 5.00 PERCENT OF INCOME DISTRIBUTION

QUANTILE FOR OVERALL DISTRIBUTION IS 1090.4941
COL 1 IS MEAN INCOMES OF TOP 5.00 PERCENT OF EACH GROUP
COL 2 IS SHARES OF TOP 5.00 PERCENT IN GROUP INCOME
COL 3 IS PERCENT OF GROUP POPULATION IN TOP 5.00 PERCENT OF OVERALL INCOME DISTRIBUTION
COL 4 IS PERCENT COMPOSITION OF TOP 5.00 PERCENT OF OVERALL DISTRIBUTION

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 RICH</td>
<td>3300.36</td>
<td>16.50</td>
<td>32.11</td>
<td>96.33</td>
</tr>
<tr>
<td>2 MIDDLE</td>
<td>841.53</td>
<td>14.02</td>
<td>0.52</td>
<td>3.62</td>
</tr>
<tr>
<td>3 POOR</td>
<td>313.53</td>
<td>17.42</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td>4 OVERALL</td>
<td>1793.58</td>
<td>29.39</td>
<td>5.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

ANALYSIS OF TOP 1.00 PERCENT OF INCOME DISTRIBUTION

QUANTILE FOR OVERALL DISTRIBUTION IS 2164.7547

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 RICH</td>
<td>4900.85</td>
<td>4.89</td>
<td>6.68</td>
<td>99.73</td>
</tr>
<tr>
<td>2 MIDDLE</td>
<td>1744.42</td>
<td>3.91</td>
<td>0.01</td>
<td>0.26</td>
</tr>
<tr>
<td>2 POOR</td>
<td>474.27</td>
<td>5.27</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>4 OVERALL</td>
<td>3045.32</td>
<td>10.15</td>
<td>1.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>
OVERALL INCOME DISTRIBUTION
INCOME INTERVALS ARE 26.93 IN UNITS OF 1.
FREQUENCIES ARE IN PERCENT
SYMBOL * IS ACTUAL DISTRIBUTION
SYMBOL # IS ESTIMATED DISTRIBUTION ASSUMING LOGNORMAL DISTRIBUTION WITH MU = 5.06764 AND SIGMA = 1.11110

ANNUAL INCOME

0.0 60,000 120,000 180,000 240,000 300,000 360,000 420,000 480,000 540,000 600,000

0.0 1.500 3.000 4.500 6.000 7.500 9.000 10.500 12.000

2 2 2 2 2 2 2 2 2 2
### TEST RUN FOR DISTRIBUTION PROGRAM 8/5/76

#### MEAN INCOMES OF GROUPS IN QUANTILE RANGES

Quantile incomes are given in final BCW.

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCW</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Rich</td>
<td>43.55</td>
<td>160.05</td>
<td>516.85</td>
<td>1499.21</td>
</tr>
<tr>
<td>2 Middle</td>
<td>112.87</td>
<td>113.97</td>
<td>356.61</td>
<td>1006.09</td>
</tr>
<tr>
<td>3 Poor</td>
<td>33.58</td>
<td>97.10</td>
<td>282.47</td>
<td>972.69</td>
</tr>
<tr>
<td>4 Overall</td>
<td>33.62</td>
<td>110.26</td>
<td>383.75</td>
<td>1452.71</td>
</tr>
<tr>
<td>Quantiles</td>
<td>50.00</td>
<td>200.00</td>
<td>800.00</td>
<td></td>
</tr>
</tbody>
</table>

#### PERCENT SHARES OF GROUPS IN QUANTILE RANGES

Quantile incomes are given in final BCW.

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCW</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Rich</td>
<td>0.30</td>
<td>1.95</td>
<td>48.16</td>
<td>13.89</td>
</tr>
<tr>
<td>2 Middle</td>
<td>0.20</td>
<td>33.02</td>
<td>64.60</td>
<td>2.19</td>
</tr>
<tr>
<td>3 Poor</td>
<td>31.64</td>
<td>61.45</td>
<td>6.88</td>
<td>0.03</td>
</tr>
<tr>
<td>4 Overall</td>
<td>15.89</td>
<td>42.58</td>
<td>33.27</td>
<td>8.26</td>
</tr>
<tr>
<td>Quantiles</td>
<td>50.00</td>
<td>200.00</td>
<td>800.00</td>
<td></td>
</tr>
</tbody>
</table>

#### PERCENT COMPOSITION OF QUANTILES BY GROUPS

Quantile incomes are given in final BOY.

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCW</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Rich</td>
<td>0.00</td>
<td>0.69</td>
<td>21.71</td>
<td>90.57</td>
</tr>
<tr>
<td>2 Middle</td>
<td>0.43</td>
<td>27.15</td>
<td>57.95</td>
<td>9.26</td>
</tr>
<tr>
<td>3 Poor</td>
<td>99.57</td>
<td>72.17</td>
<td>10.34</td>
<td>0.17</td>
</tr>
<tr>
<td>4 Overall</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Quantiles</td>
<td>50.00</td>
<td>200.00</td>
<td>800.00</td>
<td></td>
</tr>
</tbody>
</table>
Programmer's Guide to the Income Distribution Program

A listing of the program is given in Appendix C. This appendix describes the basic architecture of the program. The program consists of a MAIN program and the following subprograms:

Group 1: STATS
   CLASS

Group 2: ADDER  MEANQ  PAGER
   CUMDF  MEANY  QUANT
   GINI  NORMAL  QUANT2
   LNORM  NORM1  QUANT3
   MATOUT  SOR.??  QUANT4

Group 3: PLOT
   RANKHI
   TSCXLE
   PSSCALE

The Group 1 programs, STATS and CLASS, are the main driving routines for the program. STXTS does the basic analysis of the within-group and overall distributions. CLASS does the analysis of the group and overall distributions using the income ranges specified in the vector YCLS.

The subprograms in group 2 are the main working routines of the program. Subroutine PAGER prints the heading at the top of each page and MATOUT is a utility program for printing a matrix.

Subroutine NORMAL is used to calculate the normal integral and its inverse. NORMAL, and only NORMAL, calls the two routines NORM1 and NORM2. Subprogram NORM1 calculates the standard normal integral and NORM2 calculates the inverse integral. Similar subroutines are usually part of the FORTRAN library at most computer installations. The appropriate library routines for an IBM 360 and a UNIVAC 1108 are indicated in subroutine NORMAL. The algorithms used in NORM1 and NORM2 are crude and it would be better to use a library routine, if available. The sample problem was done using
IBM library routines.

The subprograms in Group 3 are used to plot the distribution function and the Lorenz curve for the overall distribution. These subprograms are only called from subroutine STXTS (which calls subroutine PLOT twice). If no plots are desired, all the programs in Group 3 can be deleted and the two calls to PLOT from STATS must also be deleted. In this case, the parameter IPRINT in STATS should be less than 2 to avoid printing the graph headings.

If it is desired to overlay the program, note that the groups of subprograms are hierarchical. No subprogram in a higher numbered group calls a subprogram from a lower numbered group. In addition, note that the Group 3 subprograms are only called by subroutine STXTS.
Appendix C:

Program Listing
PROGRAM WRITTEN BY SHERMAN BOBGIN, PRINCETON UNIVERSITY,  

MAIN PROGRAM TO RUN STATS, WHICH MAKES UP OVERALL DISTRIBUTION  
FROM A SET OF GROUP DISTRIBUTIONS.

DIMENSION XBAR(35), XMU(35), SIGMA(35), W(35),  
1 XNAME(35, 3), XMEAN(35, 5), POP(35, 5), TCLS(5), SSUM(3)  
DATA SSUM/OVER' 'ALL ' '/, XEND'/END'/

COMMON/PAGE/ IN MAIN, CLASS, MATEST, PAPER, PLOT, QUANT4, STATS  
COMMON/PAGE/ WPAGE, NLINE, HCG(20)
COMMON/WORK/ IN MAIN, CLASS, GINI, QUANT3, QUANT4, XNC STATS  
COMMON /WORK/ NRD, NCD, DUMY(35, 20)

C
C  INITIALIZE FIRST PAGE OF OUTPUT  
NPAGE = 1

DIMENSIONS FOR DUMY, A WORK AREA  
NRD = 35  
NCD = 20

C FOR CARD READER, OR CONVERT TO TAPE UNIT
LR = 5
JB = 0

C READ ONE CAR3 (UP TO 80 CHARS.) WHICH WILL BE A ONE LINE HEADING ON OUTPUT
READ (LR, 1001) NG, NQ, NCL, NCODF

C  READ NUMBER OF GROUPS TO BE AGGREGATED, NUMBER OF QUANTILES, AND  
2 NUMBER OF INCOME CLASSES.
READ(LR,1001) NG, NQ, NCL, NCODF

C  READ LABELS FOR GROUPS, UP TO 10 CHARACTERS
READ (LR, 1002) ((XNAME(I,J), J=1,3), I=1,NG)

C  READ DATA
DO 13 I = 1, NG  
READ (LR, 1003) XBAR(I), SIGMA(I), W(I)

C  READ NCODE .EQ. 1, READ STANDARD DEVIATION = SQRT(LOGVARIANCE) = SIGMA  
C  NCODE .NE. 1, READ LOG VARIANCE, SIGMA ** 2
IF (NCODE .EQ. 1) SIGMA(I) = SQRT(SIGMA(I))
CONTINUE

C  KCL = NCL - 1
IF (NCL .GT. 1) READ(LR,1003) (TCLS(I), I=1,KCL)
TCLS(NCL) = 0.3
DO 20 I = 1, NG  
IF (XBAR(I) .LT. 1.2-8) GO TO 910  
XNU(I) = ALOG(XBAR(I)) - 0.5 * SIGMA(I) * SIGMA(I)
CONTINUE

CALL STATS(XNAME, XBAR, XMU, SIGMA, W, NG, NQ, 2, 3)

1(NCL .GT. 1)
CALL CLASS (YCLS, YMEAN, POP, XBAR, XMU, SIGMA, W, NG, NCL, 1, XNAME)

TO 70 10C

CALL WRITE (6, 2000) JB, NG, NQ, NCL, NCODE

FORMAT (1HO, ' **** JOE', 'I3,' ERROR IN PARAMETERS.  NG = ', I4, '  NQ = ', I4, '  NCL = ', I4, '  NCODE = 1 **** ')

STOP
END

SUBROUTINE STATS (XNAME, XMEAN, XMU, SIGMA, W, NF, N, INDEX, IPRINT)

SUBROUTINE TO CALCULATE DISTRIBUTION STATISTICS AND PRINT THEM
WRITTEN BY S. ROBINSON AUGUST 1976.
INDEX=1 ONLY CALCULATE QUANTILES AND MEAN INCOMES CP QUANTILES
INDEX=2 ALSO CALCULATE QUANTILES AND MEANS OF OVERALL DISTRIBUTION, AS WELL AS NP GROUPS.
IPRINT=1 PRINT THE STATISTICS STORED IN DUMY.
IPRINT=2 IN ADDITION, CALCULATE AND PRINT A FREQUENCY DISTRIBUTION
FOR THE OVERALL DISTRIBUTION.
(N - 1) = NUMBER OF QUANTILES FOR N RANGES. SO N VALUES OF QMEAN ARE CALCULATED. IN DUMY, QUANTILES WILL BE STORED AS SHARES OF TOTAL INCOME RATHER THAN INCOME LIMITS.
QMEAN CONTAINS QUANTILE MEAN INCOMES FOR LAST GROUP.

DIMENSION W (1), QMEAN (35), XNAME (35, 3), XMEAN (1), XMU (1), SIGMA (1), 1 SAVE (35), W (35), IX (50), Y1 (50), Y2 (50), 2 Y3 (50), XX (1), WORK1 (35), WORK2 (35)

COMMON /PAGE/ NPAGE, NLINE, HDG (20)
COMMUN /WORK/ NRD, NCD, DUMY (35, 20)

EQUIVALENCE (XX (1), DUMY (1, 1)), (Y1 (1), DUMY (1, 3)), (Y2 (1), DUMY (1, 5)), 1 (Y3 (1), DUMY (1, 7)), (WORK1 (1), DUMY (1, 19)), (WORK2 (1), DUMY (1, 19))

UNIT = 10
DIMENSION OF DUMY
DIMENSION OF XNAME
INAME = 35
NORMALIZE WEIGHTS SO THAT THEY SUM TO 1.0, IF NECESSARY.

P = 1.0C / FLOAT (N)
SUMW = 0.0
I? (NF, LE, 6) GO TO 50
DO 4 IX = 1, NP
4 SUMW = SUMW + W (IX)
1? (ABS (SUMW), LT, 1.0E-6) GO TO 50
DO 5 IX = 1, NF
2
5 1 = W (IX) / SUMW

CALCULATE QUANTILES FOR EACH GROUP

NN = 4 * 1
CALL PAPER (100)
WRITE (6, 900) I, NN, (JJ, JJ = 1, NM)

FORMAT (1H8, 5X, 'MEAN INCOMES OF QUANTILES', '/X, 1X,
1 'QUANTILES OF ', F5.2, ' PERCENT.' ,
2 '/X, 1X, 'COL ', I2, ' IS OVERALL GROUP MEAN INCOME',
3 '/X, 2X, 7HCOLUMN', 7X, (11 (I5, 4X)))
WRITE (6, 901)
1 FORMAT (1H, 'N X'),
LINE = LINE + 1
IF(N.LE.1) GO TO 50
XMEAN(NP+1) = Q.0
DO 10 K=1,NF
  XMEAN(NP+1) = XMEAN(NP+1) + WW(K) * XMEAN(K)
  KK = K
  CALL QUANT2(SAVE,N,XMU(K),SIGMA(K),XXX)
  CALL MEANY(QMEAN,SAVE,XMU(K),SIGMA(K),XMEAN(K),N)
  CALL GINI(QMEAN,XMU(K),SIGMA(K),K,KK)
  DO 3 K=1,N
  QMEAN(K) = QMEAN(K) * UNIT
  XMEAN = XMEAN(K) * UNIT
  WRITE (6,902) K,(XNAME(K,JJ),JJ=1,3), (QMEAN(J),J=1,N),XMEAN
  DO 3 K=1,N
  CALL QUANT(SAVE,N,XMU,SIGMA,WW,NF,XXX)
  CALL LNORM(XMU,YSIGMA,XMU, SIGMA,WW,NF,RATIO)
  Q1 = 0.0
  SUM = C.3
  DO 73 I=1,NF
    Q2 = SAVE(I)
    IF(I.EQ.N) Q2 = 0.0
    CALL MEANY2(QMEAN(I),Q1,Q2,XMU,SIGMA,WW,NF,WORK1,WORK2)
    SUM = SUM + QMEAN(I)/FLCAT(N)
    QMEAN(I) = QMEAN < WW/I
  70 Q1 = Q2
  XMEAN = XMEAN(NP+1) * UNIT
  KK = NF+1
  WRITE (6,902) KK,(XNAME(KK,JJ),JJ=1,3), (QMEAN(J),J=1,N),XMEAN
  LINE = LINE + 1
  SUMY = C.3
  DO 76 I = 1,NF
    DUMMY(I,N+5) = XMEAN(I)
    DUMMY(I,N+6) = WW(I)
  75 SUMY = SUMY + WW(I) * XMEAN < I
  DO 77 I = 1,NF
    DUMMY(I,N+7) = 1.0 + WW(I) * XMEAN(I) / SUMY
    DUMMY(NF+1,N+5) = XMEAN(NF+1)
    DUMMY(NF+1,N+6) = 1C.3
    DUMMY(NF+1,N+7) = 1C.3
    CALL GINI(QMEAN,YMU,YSIGMA,N,NF+1).
  100 CONTINUE
  IF (I .NE. 1) RETURN
  NF2 = 0
  IF(NYXX.EQ.2) NF2 = N1 = N+1
  N2 = N + 7
  NUMLIN = NUMLIN + 2 * NF + N2
  CALL PAGER...(NUMLIN)
  IF (NL IN < > 0) WRITE(...,'(5')
  1014 FORMAT (1H0)
  WRITE (6,1015) N,P,(I,Y..1 S..,RATIO)
1010 FORMAT(1H , 'COLS 1 - ' , I2 , ' ARE SHARES OF INCOME BY QUANTILES OF ' , F5.2 , ' PERCENT' , / , 1X , 2 'COL ', I2 , ' IS THE GINI COEFFICIENT ASSUMING LOGNORMALITY' , / , 1X , 3 'COL ', I2 , ' IS THE GINI COEFFICIENT CALCULATED NUMERICALLY' , / , 1X , 4 'COL ', I2 , ' IS THE LOG VARIANCE' , / , 1X , 5 ' IS THE GEOMETRIC MEAN INCOME, EXP(XMU)' , / , 1X , 6 'COL ', I2 , ' IS MEAN INCOME' , / , 1X , 7 'COL ', I2 , ' IS GROUP SHARES IN TOTAL POPULATION' , / , 1X , 8 'PERCENT OF TOTAL LOG VARIANCE DUE TO WITHIN GROUP ' , X 'VARIANCE = ' , F6.2 , ' )  
NLINE = NLINE + 10  
CALL MATOUT(DUMY,XNAME,NF,N2,NPD,NCD,INAME,4,NF2)  
IS (INDEX,EQ.1) GO TO 110  
NNF = NF + 1  
call QUANT3 (SAVE,XMU,SIGMA,W,NF,N)  
call PAGER(160)  
WRITE((5,1305) NNF,P)  
1.5 FORMAT(1H , ' PERCENT COMPOSITION BY GROUPS IN QUANTILES OF ' , 1 'OVERALL DISTRIBUTION' , / , 1X , 'COLUMNS SUM TO ONE HUNDRED' , / , 2 'X , 'ROW 2 'IS INCOMES OF QUANTILES OF ' , F5.2 , ' PERCENT' , / )  
NLINE = NLINE + 4  
call MATOUT(DUMY,XNAME,NF,N,NRD,NCD,INAME,2,1)  
call ADDER HERE TO GET ROW SUMS, NOT COL SUMS  
call ADDER (DUMY,NNF,N,NRD,NCD)  
do 30 I = 1, NF  
do 30 J = 1, N  
30 DUMY (I,J) = DUMY (I,J ) * 100.0 / DUEY(I,N+1)  
NUMIN = NLINE + NF + 7  
call PAGER(NUMIN)  
if (NLINE .LE. 2) GO TO 61  
write (6,903)  
nline = nline + 1  
61 write (6,1015)  
1015 FORMAT(1H , '/1X,' , 1X , 'PERCENT DISTRIBUTION OF GROUPS IN QUANTILES OF OVERALL DISTRIBUTION' , / , 1X , 'ROWS SUM TO ONE HUNDRED' , / )  
nline = nline + 4  
call MATOUT(DUMY,XNAME,NF,N,NRD,NCD,INAME,2,0)  
110 continue  
calculate and print distribution statistics for to? quantiles  
call QUANT4 (XNAME,XMEAN,XMU,SIGMA,WW,NF,INAME)  
if (IPRINT .LT. 2) RETURN  
calculate and print a density function of the overall distribution
Y = (ALCG(XX(I)) - YMU) / YSIGMA
CALL NORMAL (F2, Y, IER)
Y2(I) = (2 - 1) * 100.0
F1 = F2
CALL CUMDF (F2, XX(I), YMU, SIGMA, WW, NF, 1)
Y1(I) = (F2 - F1) * 100.0
20 F1 = F2

RESCALE INCOME SO IT IS MEASURED IN UNITS.
30 25 I = 1, NNN
25 XX(I) = XX(I) * UNIT
DDY = DX * UNIT

INSERT FOLLOWING CARD TO CMIT GRAPH OF FREQUENCY FUNCTION.
1? (IPRINT .GE. 2) GO TC 209

CALL PAGEP (100)
WRITE (6, 1010) DDX, UNIT, YMU, YSIGHA
1310 FORMAT ('H, 2D, Y, 'OVERALL INCOME DISTRIBUTION', '/Y, 16X, 'INCOME ',
1 'INTERVALS ARE ', F10.2, ' IN UNITS OF: ', 6.0, '/Y, 16X, 
2 'FREQUENCIES ARE IN PERCENT', '/Y, 16X, 
3 'SYMBOL IS ACTUAL DISTRIBUTION', '/Y, 16X, 
4 'SYMBOL IS ESTIMATED DISTRIBUTION ASSUMING LOG NORM', '/Y, 16X, 
5 'DISTRIBUTION WITH MU = ', F10.5, ' AND SIGMA = ', F10.5, '/Y)

CALL PLOT (41, 101, XX, NNN, 2, Y1, Y2, Y1, Y1, 4, 1, X1, X2, X3, X4)
WRITE (6, 1011)
1011 FORMAT (1X, '/Y, 35X, 'ANNUAL INCOME' )
NLNE = 100

209 CALL QUANT (Y3, 20, YMU, SIGMA, WW, NF, PROB)
 SUM = 0.0
 Q1 = 3.0
 DO 210 I = 1, 20
 Q2 = Y3(I)
 I? (I, EQ. 20) Q2 = 0.0
 CALL MEANY2 (Y2(I), Q1, Q2, YMU, SIGMA, WW, NF, WORK1, WCRK2)
 Q1 = Q2
 XX(I) = 5.0 * PLOT(I)
 SUM = SUM + Y2(I)
 Y1(I) = Y2(I)
 IF (I .GT. 1) Y1(I) = 1 (1) + 1 (I - 1)
210 CONTINUE
 DO 220 I = 1, 20
 Y1 (I) = 100.0 * Y1(I) / SUM
220 Y2(I) = XX(I)
 XMAX = 100.0
 XMIN = 0.0

USE THE FOLLOWING CARDS FOR A SMALL GRAPH

CALL PAGE9 (100)
WRITE (6, 1025)
ALL PLOT (36, 61, XX, 20, Y2, Y1, Y3, 1.1, XMAX, XMIN, XMAX, XMIN)
WRITE (6, 1021)
ZALL PAGE9 (100)
WRITE (6, 1025)
1025 FORMAT ('H, 25X, 'LORENZ CURVE FOR THE OVERALL DISTRIBUTION', 'Y, 
1 1X, 'PERCENT OF INCOME' )
CALL PLOT(51, 101, XX, 20, 2, Y2, Y1, Y3, Y3, 4, 1, XMAX, XMIN, XMAX, XMIN)
WRITE(6, 1921)
1021 FORMAT(1H0, 35X, 'PERCENT OF HCRUSZONLDS')
NLINE = 100
RETURN
5 WRITE(6, 1001) N, S, SUMW
1331 FORMAT(1H1, '***ERRO IN PARAMETER IN STATS***', /, 5X, 'H? = ', I3, ' N? = ', I3, ' SUMW = ', E12.5)
NLINE = NLINE + 2
RETURN
END
SUBROUTINE CLASS(YCLS, YMEAN, POP, XMEAN, XMU, SIGMA, W, N, NCL, IPRINT, 1 XNAME)
SUBROUTINE TO CALCULATE POPULATION SHARE AND
MEAN INCOME OF GROUPS PALLING WITHIN SPECIFIC
INCOME RANGES GIVEN IN YCLS FOR NG GROUPS
NG = NUMBER OF GROUPS
XMU = XMU (NG) LOG MEANS
SIGMA = SIGMA (NG) LOG STANDAED DEVIATIONS
W = GROUP POPULATION OR SHARES
NCL = NUMBER OF INCOME CLASSES BY INCOME RANGE
YCLS = ARITHMETIC INCOMES OF INCOME CLASS YCLS (NG-1)
YMEAN = YMEAN (NG+1, YC) OUTPUT MEAN INCOME O? PEOPLE
PALLING IN INCOME CLASSES
POP = POP (NG+1, N) POPULATION SHARES IN EACH CLASS. SHARES OF GROUP PO
PRINT: PRINTING CCDE. PRINT TABLE IF = 1
DIMENSION WW(35), YMEAN (35, 5), POP (35, 5), XNAME (35, 3), 1 YCLS (NCL), XMU (NG), SIGMA (NG), W (NG), YMEAN (NG)
COMMON /PAGE/ NPAGE, NLINE, HDG (20)
COMMON /WORK/ NRD, NCP, SUMY (35, 20)
SET DIMENSIONS
NRP IS NC. ROWS FCE YMEAN, FCP
NCP IS NO. COLUMNS PCA YMEAN, POP
NRP = 35
NCP = 5
INAME = NRP
I? (NG, LT, 1, OR, HG, GT, (NRP - 2)) IC TO 900
IF (IC, LT, 1, OR, NCL, GT, NCP) GO IC 900
SUM = 0.0
DO 5 I = 1, NG
5 SUM = SUM + W (I)
DO 6 I = 1, NG
6 WW (I) = W (I) / SUM
YCLS (NCL) = 0.0
NG = NG + 1
Y1 = 3.0
DO 100 IC = 1, NCL
100 Y2 = YCLS (IC)
I IF 1 .EQ. HCL) Y2 = 0.0
IF (IC, LT, NCL, AND, Y2 .LE. Y1) GO TO 10
CALL MEANY2 (X, Y1, Y2, XMU, SIGMA, WW, NG, YMEAN (1, IC), POP (1, IC))
YMEAN (NG, IC) = X
F1 = 0.6
F2 = 1.0
IF (Y1, GT, 1.0) CALL CUMDF (F1, Y1, XMU, SIGMA, WW, NG, 1)
IF (Y2, GT, 1.0) CALL CUMDF (F2, Y2, XMU, SIGMA, WW, NG, 1)
POP (NG, IC) = F2 - F1
10 Y1 = Y2
IF (IPRINT, LT, 1) RETURN
NCLL = NCL - 1
FORCE HEADER
CALL PAGER (100)
WRITE (6,1000)
1000 FORMAT (1H8,5X,'MEAN INCOMES OF GROUPS IN QUANTILE RANGES',//,6X,
1 'QUANTILE INCOMES ARE GIVEN IN FINAL ROW',//)
NLINE = NLINE + 3
CALL MATOUT (XNAME,NG,NCL,NRP,NCP,INAME,2,1)
WRITE (6,1001) (YCLS(K),K=1,NCLL)
1331 FORMAT (1H0,' QUANTILES ',10F10.2,/) 
YLINE = NLINE + 3
DO 10 I = 1, HNG
DO 13 J = 1, NCL
10 DUMMY(I,J) = 100.0 * POP(I,J)
NUMLIN = 2 * (NLINE - 2)
CALL PAGES (NUMLIN)
I? (NLINE.GT. 2) WRITE (6,1004)
1004 FORMAT (1H0)
WRITE (6,'(I102)
1002 FORMAT (1H8,5X,'PERCENT SHARES OF GROUPS IN QUANTILE RANGES',//,6X,
1 'QUANTILE INCOMES ARE GIVEN IN FINAL ROW',//)
NLINE = NLINE + 3
CALL MATOUT (DUMMY,XNAME,NG,NCL,NRP,NCD,INAME,2,1)
WRITE (6,1301) (PCLS(K),K=1,NCLL)
NLIYE = NLINE + 3
DO 25 J = 1, HCL
SUM = 0.0
DO 15 I = 1, NG
DUMMY(I,J) = WW(I) * POP(I,J)
15 SUM = SUM + DUMMY(I,J)
DO 20 I = 1, NG
20 DUMMY(I,J) = 100.9 * DUMMY(I,J) / SUM
25 CONTINUE
NUMLIN = NLINE + 11 + NG
CALL PAGES (NUMLIN)
I? (NLIYS.GT. 2) WRITE (6,1004)
WRITE (6,1003)
1003 FORMAT (1H8,5X,'PERCENT COMPOSITION OF QUANTILES BY GROUPS',//,6X,
1 'QUANTILE INCOMES ARE GIVEN IN FINAL ROW',//)
NLINE = NLINE + 3
CALL ADDER (DUMMY,NG,NCL,NRP,NCD)
CALL MATOUT (DUMMY,XNAME,NG,NCL,NRP,NCD,INAME,2,1)
WRITE (6,1001) (YCLS(K),K=1,NCLL)
NLINE = NLINE + 3
RETURN
900 WRITE (6,1010) NJ,YCL
1010 FORMAT (1H0,' **** ERROR IN CLASS. NG = ',I5)
NLINE = NLINE + 2
RETURN
910 WRITE (6,1010) NG,NCL
WRITE (6,1001) (YCLS(K),K=1,NCLL)
NLINE = YLINE + 5
RETURN
END
SUBROUTINE ADDER (Z,NROW,NCOL,KROW,KCOL)
DOUBLE PRECISION ROWSUM,COLSUM,SUM
THE ONLY DCUBLP PRECISION STATEMENT
DIMENSION Z(KROW,KCOL)
I? (NCOL.EQ. 1) GO TO 5
DO 60 I=1,NROW
FOWSUM = 0.
CALCULATE SUN OF COLUMNS SUMS

SUM = 0.
DO 3 J=1,NROW
30 SUM = SUM + Z(I,NCOL+1)
Z(NROW+1,NCOL+1) = SUM
RETURN
END

SUBROUTINE CMDF(F,X,XMU,SIGMA,W,NF,INDEX)
DIMENSION XMU(NF),SIGMA(NF),W(NF)

PROGRAM TO CALCULATE CUMULATIVE FREQUENCY OF SUM OF NF LOGNORMAL DISTRIBUTIONS GIVEN INCOME X. THE FREQUENCIES ARE WEIGHTED BY W.

IF INDEX=0, PROGRAM TAKES F AND DELIVERS GEOMETRIC MEANS OF THE QUANTILES FROM THE YF DISTRIBUTIONS.
IF INDEX=1, PROGRAM TAKES X AND DELIVERS F.

IF (INDEX .EQ. 0) GO TO 100
F = 3.6
IF (X .GE. 1.E-8) GO TO 100
WRITE(6,1001) X
1001 FORMAT(1H0,'**** ERROR IN CMDF, X = ',F13.6)
RETURN
10 Y = ALOG(X)
DC 100 I=1,NF
YY = (Y-XMU(I))/SIGMA(I)
CALL NORMAL (YY,I,YY,IER)
100 ? = F + W(I)*YY
RETURN
110 CONTINUE
X = .0
DO 200 I=1,NF,
CALL NORMAL (F,YY,IER)
IF (IEF .EQ. 1) WRITE(6,1000) F, YY
1000 FORMAT(1H0,'**** ERROR IN CMDF F = ',F13.6,YY = ',F13.6, 1 **')
Y = YY + SIGMA(I) + XMU(I)
200 X = X + W(I)*Y
X = EXP(X)
RETURN
END

SUBROUTINE GINI (X,XMU,SIGMA,N,K)
DIMENSION X(N)
COMMON /WORK/ NCD,NCD,DUMY(35,20)

SUBROUTINE TAKES N MEAN INCOMES IN QUANTILES (IN X) AND CALCULATES SHAPES OF TOTAL INCOME BY QUANTILES AND STORES THE RESULTS IN ROW X OF DUMY. THE GINI COEFFICIENT IS ALSO CALCULATED AND STORED IN COLUMN N+1 OF DUMY. SIGMA**2 IS STOSED IN COLUMN N+2.
XMU IS STOPEED IN COLUMN N+3

SUM = 0.0
DO 10 I=1,N
10 SUM = SUM + X(I)
DO 20 I=1,N
20 DUMY(K,I) = 0.5 * X(I) / SUM

CALCULATE GINI COEFFICIENT ASSUMING LOGNORMAL

Y = SIGMA/SQRT(2.0)
CALL NORMAL (F',Y,IER
DUMY(K,N+1) = 2.0 * FY - 1.0
DUMY(K,N+3) = SIGMA*SIGMA
DUMY(K,N+4) = EXP(DUMY)

CALCULATE GINI COEFFICIENT
INTEGRATE LOGENZ CURVE NUMERICALLY USING EQUAL INTERPOLATION
POLYNOMIAL.
NOTE CORRECTION FOR ODD NUMBER OF INTERVALS.

DX = 1.0/FLOAT(N)
NN = 2*INT(FLOAT(N)/2.0)
SUM = 0.0
CUM = 0.0
DO 70 I=1,N
CUM = CUM + 0.01*DUMY(I)
ADD = 0.0
IF(NN .NE. N) GO TO 60
II = I
55 CONTINUE
IF(MOD(I,2) .EQ. 0) ADD=2.0*CUM
IF(MOD(I,2) .NE. 0) ADD=4.0*CUM
IF(I .EQ. N) ADD=CUM
GO TO 70
60 II = I + 1
IF(I .GT. 1) GO TO 55
ADD = 0.5*CUM*DX + CUM
70 SUM = SUM + ADD
CUM = CUM + DX/3.0
DUMY(K,N+2) = 1.0 - 2.0*SUM
RETURN
END
SUBROUTINE LNQMN(XMU2,SIGMA2,XMU,SIGMA,W,NF,RATIO)
DIMENSION X(1),XMUM(NF),SIGMA(NF),W(NF)

SUBROUTINE TO ESTIMATE MEAN AND VARIANCE OF A POOL OF NF
DISTRIBUTIONS WITH MEANS XM && VARIANCES SIGMA**2

XMU2 = 0.0
SIGMA3 = 0.0
SUMW = 0.0
XMUSQ = 0.0
DO 10 I=1,NF
SUMW = SUMW + W(I)
XMU2 = XMU2 + W(I)*XMU(I)
XMUSQ = XMUSQ + W(I)*XMU(I)**2
SIGMA3 = SIGMA3 + W(I)**2*SIGMA(I)
10 SIGMA3 = SIGMA3 / SUMW
XMUSQ = XMUSQ / SUMW
SIGMA2 = SIGMA3 / SUMW + XMU2**2 - XMUSQ
RETURN
END
ENTRY LNOPM1 (XMU2, SIGMA2, XMU, SIGMA, W, NF, RATIO)

PROGRAM TO ESTIMATE PARAMETERS OF A LOGNORMAL DIST, XMU AND SIGMA
BY METHOD OF QUANTILES. SEE, AITCHISON AND BICYN, PAGES 40-42
TWO SETS OF TWO QUANTILES ARE NEEDED. FOR XMU, QUANTILES OF
THE OVERALL CIST. ARE CALCULATED AT P=.27 AND .73. FOR SIGMA,
THEY ARE P=.93 AND .07.

CALL QUANT(X, 1, XMU, SIGMA, W, NF, .73)
X1 = X(1)
CALL QUANT(X, 1, XMU, SIGMA, W, NF, .27)
X2 = X(1)
XMU2 = 0.5*(ALOG(X2) + ALOG(X1))

CALL NORINV (.93, ETA, IER)
I? (IER = EQ. 1) SO TO 25
ETA = 2.9 * ETA
CALL QUANT(X, 1, XMU, SIGMA, W, NF, .07)
X1 = X(1)
CALL QUANT(X, 1, XMU, SIGMA, W, NF, .93)
X2 = X(1)
SIGMA2 = (ALOG(X2) - ALOG(X1))/ETA
RETURN

WRITE(6, 1000)
1000 FORMAT(1H, '***FAILURE TO FIND ETA IN LNOPM***')
XMU2 = 0.0
SIGMA2 = 0.0
RETURN
END

SUBROUTINE MATOUT(Z, SNAKE, NBOW, NCOL, KROW, KCOL, KS, KK, LS)

DIMENSION Z(KROW, KCOL), SNAME(KS, 3)
COMMON/PAGE/, NPAGE, NLINE, HDG(20)

Z = SINGLE PRECISION MATRIX TO BE PRINTED
SNAME(I, J) = MATRIX OF ROW NAMES. I FOR ROW, J=1,3 FOR 3 WORDS TO
A 12 CHARACTER STRING
NBOW, NCOL = NO. OF ROWS AND COLS TO BE PRINTED
KROW, KCOL = DIMENSIONS OF Z IN MAIN PROGRAM
KS = DIMENSION OF SNAKE IN MAIN PROGRAM
KK = NO. OF DIGITS TO RIGHT OF DECIMAL POINT
LS = OPTION FOR SKIPPING LINE AND PRINTING COL SUMS
IF LS=.3E. 1, (NBOW+1) TH ROW IS PRINTED AFTER SKIPPING A LINE.
NP COL = NO. OF COLUMNS ACROSS PRINTED PAGE. SET TO 13 FOR BATCH
OUTPUT. SET TO 5 FOR TERMINAL OUTPUT.

NP COL = 10
W= 6
MADD = 5
KP = KK + 1
IR(KP, GT, 7) KP =
M = 1
M = NCOL

N (MM, GT, NCOL) MM = NCOL
LSTART = 1
SEND = KROW
LSkip = LS
COL. NCS. ACFOSS PAGE
WRITE(IW,1011) (I,I=M,MM)
1311 FORMAT(1H,' 1X,'COLUMN',7X,10(I5,5X))
WRITE(IW,1012)
1012 FORMAT(1H,'ROW')
NLIN = NLIN + 1
8 DO 30 I = LSTART, LEND
NLIN = NLIN + 1
GO TO (10, 11, 12, 13, 14, 15, 16), KP
10 WRITE(IW,1000) I, (SNAME(I,J), J=1,3), (Z(I,J), J=M,MM)
1000 FORMAT(1H,' 1X, 2A4,A2,10 (F10.0))
GO TO 30
11 WRITE(IW,1001) I, (SNAME(I,J), J=1,3), (Z(I,J), J=M,MM)
1001 FORMAT(1H,' 1X, 2A4,A2,10 (F10.1))
GO TO 30
12 WRITE(IW,1002) I, (SNAME(I,J), J=1,3), (Z(I,J), J=M,MM)
1002 FORMAT(1H,' 1X, 2A4,A2,10 (F10.2))
GO TO 30
13 WRITE(IW,1003) I, (SNAME(I,J), J=1,3), (Z(I,J), J=M,MM)
1323 FORMAT(1H,' 1X, 2A4,A2,10 (F10.3))
GO TO 30
14 WRITE(IW,1004) I, (SNAME(I,J), J=1,3), (Z(I,J), J=M,MM)
1004 FORMAT(1H,' 1X, 2A4,A2,10 (F10.4))
GO TO 30
15 WRITE(IW,1005) I, (SNAME(I,J), J=1,3), (Z(I,J), J=M,MM)
1035 FORMAT(1H,' 1X, 2A4,A2,10 (F10.5))
GO TO 30
16 WRITE(IW,1006) I, (SNAME(I,J), J=1,3), (Z(I,J), J=M,MM)
1006 FORMAT(1H,' 1X, 2A4,A2,10 (F10.6))
CONTINUE
30 IF(LS平坦IE.O) GO TO 35
WRITE(IW,2000)
2000 FORMAT(1H)
NLIN = NLIN + 1
LSKIP = 0
LSTART = NROW + 1
LEN = LSTART
GO TO 8
CCC
DO WE HAVE MORE THAN HPCCL COLUMNS TO PRINT?
35 IF (MM .GE. HCOL) GO TO 75
CCC
WILL THEY FIT ON SAME OUTPUT PAGE?
NLIN = NLIN + NADD + HROW
CALL PAGE(NLIN)
WRITE(IW,2001)
2001 FORMAT(1H0)
NLIN = NLIN + 2
73 MM = MM + NPCOL
GO TO 5
75 RETURN
END

SUBROUTINE MEANQ(X,Q1,Q2,XMU,SIGMA,XMEAN,PROB)
C SUBROUTINE TO CALCULATE MEAN INCOME CP PEOPLE BETWEEN QUANTILES
C Q1 AND Q2 FOR A LOG-NORMAL DISTRIBUTION.
I (Q2 .GE. 1. E-3 AND .Q2 .LE. Q1) GO TO 900
I? (ABS(XMEAN) .LE. 1. E-8) XMEAN = EXP(XMU + 0.5*SIGMA*SIGMA)
XMU2 = XMU + SIGMA * SIGMA
X1 = 0.0
X2 = 0.0
I? (Q1 .GE. 1. E-8) X1 = (ALOG(Q1) - XMU) / SIGMA
IF (Q2 .GT. 1.E-8) X2 = (ALOG(Q2) - XMU) / SIGMA
F1 = C.C
F2 = 1.0
IF (Q1 .GT. 1.E-3) CALL NORMAL(F1,X1,IER)
IF (Q2 .GT. 1.E-8) CALL NORMAL(F2,X2,IER)
PROB = F2 - F1
X1 = 3.0
X2 = 6.0
IF (Q1 .GT. 1.E-9) X1 = (ALOG(Q1) - XMU2) / SIGMA
IF (Q2 .GT. 1.E-8) X2 = (ALOG(Q2) - XMU2) / SIGMA
F1 = 0.0
F2 = 1.3
IF (Q1 .GT. 1.E-9) CALL NORMAL(F1,X1,IER)
IF (C2 .GT. 1.E-8) CALL NORMAL(F2,X2,IER)
PROB2 = F2 - F1
X = 3.0
I? (ABS(PROB) .GT. 1.E-8) X = XMEAN * FCFB2 / PROB
RETURN
WRITE(6,1000) Q1,Q2
1000 FORMAT(1HO,'**** ERROR IN MEANQ. Q1 = ',E13.6,' Q2 = ',
1 E13.6,' ****')
RETURN
END
SUBROUTINE MEANQ(X,Q,XMU,SIGMA,XMEAN,N)
DIMENSION X(N),Q(N),P(20)

SUBROUTINE TO CALCULATE MEAN INCOMES OF QUANTILES OF LOGNORMAL DIST. GIVEN IN Q

IF (N .GT. 20) GO TO 910
Q1 = 0.0
Q(N) = 0.0
SUMP = 0.0
SUMX = 0.0
DO 10 I = 1, N
Q2 = Q(I)
I? (Q2 .LE. Q1 .AND. Q2 .GT. 1.E-8) GC TO 910
CALL MEANQ(X(I),Q1,Q2,XMU,SIGMA,XMEAN,F(I))
1 = Q2
SUMP = SUMP + P(I)
10 SUMX = SUMX + X(I)*P(I)
IF (ABS(SUMP-1.3) .LT. 0.001 .AND. ABS(SUMX-XMEAN) .LT. 0.01*XMEAN)
1 RETURN
WRITE(6,2000) SUMP, SUMX, XMEAN
2000 FORMAT(1HO,'*** IN MEANY, SUMP = ',E13.6,' SUMX = ',E13.6,
1 XMEAN = ',E13.6)
RETURN
0 WRITE(6,1000) (Q(K),K=1,N)
1000 FORMAT(1HO,'*** ERROR IN MEANY. C(I) = ',(5E13.6))
RETURN
910 WRITE(6,1001) N
1001 FORMAT(1HO,'*** ERROR IN MEANY, N = ',I5)
RETURN
ENC
SUBROUTINE MEANQ2(X,Q1,Q2,XMU,SIGMA,W,NF,XX,PP)
DIMENSION XRU(NF),SIGMA(NF),W(NF),XX(NF),PR(NF)

SUBROUTINE TO CALCULATE MEANS OF QUANTILES OF OVERALL DISTRIBUTION.
Q1 AND Q2 ARE TWO QUANTILES. QUANTILE MEAN OF DIST BETWEEN THE
TWO QUANTILES IS CALCULATED BY SUMMING MEANS OF INDIVIDUAL DISTRI-
BETWEEN THE SAME TWO INCOMES. IF Q1=0.0, RANGE IS FROM MINUS INFINITY. IF Q2=0.0, RANGE IS TO PLUS INFINITY.

1 = (Q2 .GT. 1.0 .AND. Q2 .LT. Q1) GO TO 900

SUM = 0.0
DO 10 I = 1, NF
XMEAN = 0.3
CALL MEANQ (XX(I), Q1, Q2, XMEU(I), SIGMA(I), XMEAN, PR(I))
10 SUM = SUM + W(I) * PR(I)
X = 0.0
IF (SUM .LT. 1.2-8) RETURN
30 20 I = 1, NF
20 X = X + W(I) * PR(I) * XX(I) / SUM
RETURN

SUBROUTINE MEANQ (XX(I), Q1, Q2, XMEU(I), SIGMA(I), XMEAN, PR(I))

SUBROUTINE TO CALCULATE NORMAL INTEGRAL FCIF MINUS INFINITY TO X AND RETURN THE VALUE IN FY
ENTRY NORMAL TAKE3 FY AHC RETURNS X
IF = 1 IF THERE IS AN ERFCF.

THE ROUTINE CAN USE IN-HOUSE LIBRARY ROUTINES FOR THE STANDARD NORMAL AND INVERSE NORMAL INTEGRALS. THEY ARE NORMAL AND MRIS FOR IEM 360

SUBROUTINE NORMAL (FY, X, IER)

ENTRY NORM (X)
FY = 0.5 * ERFC (-SQRT(0.5) * X)
CALL NCNORM (X, FY)
RETURN

SUBROUTINE TO CALCULATE NORMAL INTEGRAL FCIF MINUS INFINITY TO X AND RETURN THE VALUE IN FY
ENTRY NORMAL TAKE3 FY AHC RETURNS X
IF = 1 IF THERE IS AN ERFCF.

THE ROUTINE CAN USE IN-HOUSE LIBRARY ROUTINES FOR THE STANDARD NORMAL AND INVERSE NORMAL INTEGRALS. THEY ARE NORMAL AND MRIS FOR IEM 360

SUBROUTINE NORMAL (FY, X, IER)

ENTRY NORM (X)
FY = 0.5 * ERFC (-SQRT(0.5) * X)
CALL NCNORM (X, FY)
RETURN

SUBROUTINE TO CALCULATE NORMAL INTEGRAL FCIF MINUS INFINITY TO X AND RETURN THE VALUE IN FY
ENTRY NORMAL TAKE3 FY AHC RETURNS X
IF = 1 IF THERE IS AN ERFCF.

THE ROUTINE CAN USE IN-HOUSE LIBRARY ROUTINES FOR THE STANDARD NORMAL AND INVERSE NORMAL INTEGRALS. THEY ARE NORMAL AND MRIS FOR IEM 360

SUBROUTINE NORMAL (FY, X, IER)

ENTRY NORM (X)
FY = 0.5 * ERFC (-SQRT(0.5) * X)
CALL NCNORM (X, FY)
RETURN
\( b = \text{INPUT VALUE} \)

\( d = \text{VALUE OF INTEGRAL} \)

ACCURACY IS \( 7.10^{-7} \)

\( d = \text{PROBABILITY DENSITY AT } x \)

\[
\text{DATA A1,A2,A3,A4,A5,A6,A7} / 0.2316415,0.31933915,-0.356638, \\
1.51478,-1.024256,1.332274,0.3989423/}
\]

\( a = \text{AES}(y) \)

\( x = 0.0 / (1.0 + a * x) \)

\( d = A7 * \exp(-x * x/2.0) \)

\( p = 1.0 - d * \exp((((A6 * x + A5) * x + A4) * x + A3) * x + A2) \)

\( \text{IF } (x) 5.10^{-10} \)

\( p = 1.0 - p \)

19 RETURN

2ND

SUBROUTINE NORM2(X, P, IER)

SUBROUTINE TO CALCULATE INVERSE NORMAL INTEGRAL

\( p = \text{VALUE OF INTEGRAL} \quad \text{O.E. } p \leq 1 \)

\( x = \text{OUTPUT VALUE OF } x \)

\( \text{IER = ERROR CONDITION = C IF NO ERROR} \)

\( = 1 \quad \text{IF ERROR} \)

AX = \text{ERROR IS .CLJ45}

\( s = \text{PROBABILITY DENSITY AT } x \)

\[
\text{DATA 0,1,2,3,4,5,6,7} / 2.515517,0.602053,0.010329,1.432788, \\
0.189259,0.001399,0.3589423/}
\]

\( x = 0.0 \)

\( D = x \)

IF \( (? \) 11,14,12

11 IER = 1

30 TO 50

12 IF \( (p - 1.0) 17,15,11

14 \( x = -1.0 * e^{-70} \)

15 D = 0.0

G0 TO 50

17 D = P

IF \( (D - C.5) 19,19,18

18 D = 1.0 - D

13 T2 = ALOG1(1.0 / \( (D * 3) \))

T = \text{SORT}(T2)

Y = \( x = (B1 + B2 + B3 + 12) / (1.0 + E4*T + B5*T2 + E6*T*T2) \)

\( = (B3 - 3.5) \quad 21,20,21

20 Y = -x

21 D = 37 * \exp(-x * x/2.0)

50 RETURN

END

SUBROUTINE PAPER(N)

COMMON/PAGE/NPAGE,NLINE,HDG(2C)

IF \( (N = 1,17.55) \) RETURN

WRITE(6,1000) (HDG(J),J=1,2C),NPAG

1000 FORMAT(1H1,5X,20A4,5X,NPAGE,T,13)

WRITE(6,1001)

1001 FORMAT(1H1)

NLINE = 2

NPAGE = NPAG + 1

RETURN

ELSE

SUBROUTINE QUANT(X,N,XMU,SIGMA,W,EFCE)

\( \text{DIMENSION } y(x), xmu(nf), sigma(nf), w(nf) \)

PROGRAM TO ESTIMATE N-1 QUANTILES OF X DEFINING N RANGES EACH
MAXTRY = 25
TEST = 0.0001
NN = N-1
IF ( 1 ) NN=1
P = 0.0
DO 203 K=1,NN
P = P + 1.0/PLOAT(N)
IF (N.EQ.0) P = PROB
C MAKE INITIAL GUESS
CALL CUMDF(P,XX,XMU,SIGMA,W,NP,0)
XX1 = 0.0
P1 = 0.0
ITRY = 1
10 CONTINUE
TRY OUT THE VALUE 0? XX
CALL CUMDF(P2,XX,XMU,SIGMA,W,NP,1)
IF (ABS(P-P2).LT.TEST) GO TO 200
IF (ITRY.GT.MAXTRY) GO TO 190
C CALCULATE THE NUMERICAL DERIVATIVE AND ITERATE
ITRY = ITRY + 1
IF (ITRY.EQ.2) DERIV=XX/PP
IF (ABS(Deriv).GT.1.E-6) DERIV=(XX-XX1)/(PP-P1)
XX1 = XX
P1 = PP
DDX = DERIV*(P-PP)
IF (ABS(DDX).GT.1.E-6) DDX=0.25*XX*SIGN(1.0,DDX)
XX = XX + DDX
GO TO 10
190 WRITE(6,1) K,ITRY
1000 FORMAT(1H, '***QUANTILE NO ', I2,' DID NOT CONVERGE AFTER ', I2, ' ITERATIONS IN QUAN***')
200 X(K) = <X
RETURN
END
SUBROUTINE QUAN2(X,N,XMU,SIGMA,PROB)
* DIMENSION X(N)*
* PROGMA TO CALCULATE N-1 INCOME QUANTILES FOR LOGNORMAL DISTRIBUTION.*
* EACH QUANTILE CONTAINS 1/N OF THE PROBABILITY. IF N=1, ONLY ONE* QUANTILE IS CALCULATED, USING PROB.*
* SEE ATCHISON AND BROWN, PAGES 8-9.
* NN = N-1
* IF (N.EQ.1) NN=1
* P = 0.0
* DO 33 K=1,NN
* P = P + 1.0/PLOAT(N)
* IF (N.EQ.1) P = PROB
* CALL NQPINV(P,YY,IER)
* IF (IER.EQ.1) GO TO 95
* Y = YY*SIGMA + XMU
* X(K) = EXP(Y)
* GO TO 100
* 95 WRITE(6,1) K
100 FORMAT(1H,'*** IN QUANT2, QUANTILE ',I2,' FAILED TO SOLVE***')
X(K) = 0.0
100 CONTINUE
RETURN
END

SUBROUTINE QUANT3(X,XMU,SIGMA,W,NP,N)
DIMENSION X(I),XMU(I),SIGMA(I),W(1)
COMMON / WORK/ NRD,NCD,DUMY(35,20)

SUBROUTINE TO CALCULATE SHARES C2 TOTAL DISTRIBUTION QUANTILES
GIVEN IN X THAT COME FROM VARIOUS GROUPS. RESULTS ARE
STORED IN DUMY

NM = N - 1
DO 13 I=1,NP
IL = I
F1 = J, J
33 5 J=1,N
JJ = J
TF (ABS(X(J)).LT.1.E-3) GC TC 27C
Y = (ALCG(X(J)) - XMU(I))/SIGMA(I)
CALL NORMD(P2,Y,IER)
DUMY (I,J) = F2 - F1
5 P1 = P2
DUMY(I,J) = 1.0 - F1
10 CONTINUE

SUBROUTINE TO CALCULATE SHARES C2 TOTAL DISTRIBUTION QUANTILES
GIVEN IN X THAT COME FROM VARIOUS GROUPS. RESULTS ARE
STORED IN DUMY

USING SHARES IN ANGROUS, RELOC CELLS IN DUMY AS FREQUENCIES,
THEN AS SHARES

15 DUMY (I,J) = DUMY (I,J)*W(I)
CALL ADDER (DUMY,NF,N,NFD,NCD)
DO 20 I=1,NP
DO 20 J=1,N
DUMY(I,J) = 100.*DUMY(I,J)/DUMY(NP+1,J)

PUT QUANTILES IN NP+1 ROW OF DUMY

25 DUMY(NP+1,J) = UNIT*X(J)
DUMY(NP+1,N) = 0.0
RETURN

200 WRITE(6,1000) II, JJ
1000 FORMAT(1H,'*** ERROR IN QUANT3. GROUP #',12,' QUANTILE ',I2,1' ***')
RETURN
END

SUBROUTINE QUANT4(XNAME,XMEAN,XMU,SIGMA,W,NP,INAME)
DIMENSION XNAME(INAME,3),XMEAN(1),XMU(1),SIGMA(1),W(1),Q(1),
1 WORKI(35),WORK2(35)
COMMON / PAGE/ NPAGE,NLINES,HLC(20)
COMMON / WORK/ NRD,NCD,DUMY(25,20)
EQUIVALENCE (DUMY(1,1),WORK1(1)), (DUMY(1,2),WORK2(1))

SUBROUTINE TO CALCULATE ASCII PRINT DISTRIBUTION STATISTICS FOR
TOP 5 PERCENT AND TOP 1 PERCENT OF DISTRIBUTION
UNIT = 1,0
Q2 = 0.0
DO 100 IJ=1,2
IF (IJ .EQ. 1) P = .95
IF (IJ .EQ. 2) P = .99
P? = 1.0 - P
PPP = PP*100.
DO 10 I=1,NP
II = I
CALL QUANT2 (Q1, XMU(I), SIGMA (I), P)
Q1 = Q1
I? (Q1 .LT. 1.E-6) GO TO 900
CALL MEANQ (XX, Q1, Q2, XMU (I), SIGMA (I), XMEAN (I), P3)
IF (ABS (P? - P3) GT 0.005) WRITE (6, 10C0) I, P, P3, PPP
101 FORMAT (1H1, '*** II: QUANT4, I = ', I4, ' P = ', E13.6, ' P3 = ',
1 13.5, ' PP = ', E13.6)
IF (ABS (PP - P3) GT 0.005) NLINE = NLINE + 2
DUMY (I, 1) = XX*UNIT
10 DUMY (I, 2) = 100.*XX*PP/XMEAN (I)

CALL QUANT (O, 1, XMU, SIGMA, W, NF, P)
Q1 = Q(1)
CALL MEANY2 (XX, Q1, Q2, XMU, SIGMA, W, NF, WORK1, WORK2)
DUMY (NF+1, 2) = 100.*XX*PP/XMEAN (NF+1)
DUMY (NP+1, 1) = UNIT*XX
SUM = 0.0
DO 20 I=1,NF
XI = (ALGS (Q(1)) - XMU (I)) / SIGMA (I)
CALL NORMAL (XX1, X1, IEP)
DUMY (I, 3) = 100.* (1.0-XX1)
DUMY (I, 4) = DUMY (I, 3) * W (I)
20 SUM = SUM + DUMY (I, 4)
DO 25 I=1,NF
25 DUMY (I, 4) = 100.*DUMY (I, 4)/SUM

DUMY (NF+1, 3) = P2P
DUMY (NP+1, 4) = 100.
NUMLIN = NLINE + 8 + NF
IF (IJ .EQ. 1) NUMLIN = 10C
CALL PAGE (NUMLIN)
IF (NLINE .LE. 2) GO TO 50
WRITE (6, 10C1)
1001 FORMAT (1H1)
NLINE = NLINE + 1
50 WRITE (6, 1102) PPP, Q1
1002 FORMAT (1H1, 'ANALYSIS OF TOTAL PERCENT OF INCOME',
1 'DISTRIBUTION', '1, X1, QUANTILE POE OVERALL DISTRIBUTION IS'
2 'P12.4' )
NLINE = NLINE + 3
IF (IJ .EQ. 21) GO TO 55
WRITE (6, 10C0) PPP, PPP, PEE, PPP
1003 FORMAT (1H1, 'COL 1 IS MEAN INCOME OF TOP',
1 2 ' PERCENT OF EACH GROUP', '/1X',
2 'COL 2 IS SHARES OF TOP', 'P4.2', ' PERCENT IN GROUP INCOME',
3 '/1X', 'COL 3 IS PERCENT OF GROUP POPULATION IN TOP', 'P4.2',
4 ' PERCENT OF OVERALL INCOME DISTRIBUTION', '/1X',
5 'COL 4 IS PERCENT COMPOSITION OF TO', 'P4.2', ' PERCENT OF ',
6 'OVERALL DISTRIBUTION')
NLINE = NLINE + 4
55 WRITE (6, 10C1)
CALL MATCUT (DUMY,NAME,NF,4,NRD,NCD,INAME,2,1)

100 CONTINUE
RETURN
900 WRITE (6,1004) Q1
1004 FORMAT (1HO,*** ERROR IN QUANT4. C1 = ',E13.6 )
NLINE = NLINE + 2
RETURN
END

SUBROUTINE PLOT(LENGTH,NWIDTH,X,NPTS, NY, Y1, Y2, Y3, Y4, KEY, NOPUT,
1 XMAX, XMIN, XMAX, XMIN)
COMMON /PAGE/NPAGE,NLINE,HDG(20)

SUBROUTINE PLOT SCALES AND DFAYS PRINTER PLOTS, A MAXIMUM 2 PAGE GRAPE.
WRITTEN SPRING TERM 1975 BY ALICE ANNE NAVIN AND SHERMAN ROBINSON.
PRINCETON UNIVERSITY. REVISED SUMMER 1976 TO INCLUDE SUB. PAGE
THE ROUTINE PANKS AND SORTS UP TO 4 Y VECTORS, THE DEPENDENT VARIABLES.
ASSIGNS THEM TO A 2 DIMENSIONAL ARRAY, SAVING ORIGINAL SUBSCRIPTS.

INTEGER BLANK
3 YLABEL DIMENSION NOW 21. IF LENGTH MAX. GREATER THAN 101 BE SURE TO INCRE
3 DIMENSION X(1),Y1(1),Y2(1),Y3(1),Y4(1),
3 DIMENSION X(NPTS), Y1(NPTS), Y2(NPTS), Y3(NPTS), Y4(EPTS).
1 LOT(4), NMARK(4), LCOLA(3), LROWS (2),
2 NWATCH(1), XLABEL(21), LINE(101), Ji (1C1), XLABEL (11)
C MEMORY SAVINGS CAN BE ACHIEVED BY PASSING THE FCLCUIY ARRAYS
AS ARGUMENTS AND GIVING THEM VARIABLE DIMENSIONS.
C IF MAX. NPTS CHANGED TO GREATER THAN 100, SUBROUTINE RANKHI
C HAS TWO DIMENSIONS TO CHANGE--IRK(100), NYY (100) ----2/19/75
C DIMENSION CUMY(130), RANKY(4,100), IPTY(4,100), IPT(100)
CCC UNIVAC NEEDS UPPER CASE I IN LIU CE OF BAR. CCC
C DATA NMARK/*','*','*','0','/BLANK'/',
1 NWATCH/*' , '2', '3', '4', '5', '6', '7', '8', '9', '0'/,
2 LCOLA/*-----' +*','/LROWS/++','/C

C PORTRAN UNIT LINE PRINTER
LW = 6
C MULTIPLE NUMBER FOR ROWS C0YN, YLABELS
NP = 5
C MAX. ROWS PER PAGE
NRP = 51
C LIMIT FOR TIC NOTATIONS WHERE F0? TIES MORE THAN 9, WE INSERT 0
NINE = 9
C MULTIPLE ?0? COLUMNS ACROSS, X LABELS
NTEM = 10
C A A A LENGTH OF ROWS IS 101, A 2 PAGE GRAPH, 2/19/75
MAXLEN = 101
IT (LENGTH, LE. 1) GC TC 199
C A MAX. WIDTH OF COLS. ALSO IS 101, TC FIT 132 PRINTERS SPACES.
MAXWID = 101
IP (WIDTH, LE. 1) GO TO 199
IP (NY, LE.0) OR Y.GT. 4) GC TC 199
IP (NPTS.LE. 0) GO TO 199
IP (KEY.LE. 3, OP. KEY 4) GO TO 199
C CCC IF ALL CK, PROCEED WITH FLC?
C CONTINUE
C ROUND THE WIDTH UP TO A MULTIPLE OF 13 COLUMNS + 1 FOR XLABEL(NW)
CCC KEY = 3 KEY LOWER THESE VALUES
WIDTH = MINO(WIDTH, MAXWID)
WID = (WIDTH - 1) / NTEN
IF (MOD ((WIDTH - 1), NTEN).NE. 0) WID = WID + 1
WIDTH = WID * NTEN + 1

C ROUND THE LENGTH UP TO A MULTIPLE OF 5 ROWS, PLUS 1 FOR BOTTOM LINE WITH
CCC KEY = 3 MAY LOWER THESE VALUES
LENGTH = MINO(LENGTH, MAXLEN)
LEN = (LENGTH - 1) / NR
IF (MOD(LENGTH - 1, NR).NE. 0) LEN = ZEN + 1
LENGTH = (LEN * NR) + 1

C WILL USE FOR YAXIS LABELS, I.E. EACH + LINE.
NL = LEN + 1

3 RANK Y'S HIGH TO LOW
DO 17 J = 1, NY
DO 15 I = 1, NPTS
GO TO (11, 12, 13, 14), J
11 DUMMY(I) = Z1(I)
GO TO 15
12 DUMMY(T) = Y2(T)
GO TO 15
13 DUMMY(I) = Y3(I)
GO TO 15
14 DUMMY(I) = Y4(I)
15 CONTINUE
C
C IF ( ' ' I . EQ. 0) GC TO 10
WRITE(6W, 3)
3 FORMAT ('THO', J ', 1X, ' T ', 1X, 'DUMMY(I) ', 3X,
')
TPT(J, I), 1X, RANKY(J, I), 1X, 'X (TPT(J, I))'/)
CALL BAYSHI (DUMY, NPTS, IPT)
C TPT(J, I) IS TRUE! WHOSE PAYK IS I.
C RANKY(J, I) IS EACH Y IN DESCENDING ORDER OF RANK(AVG VALUE).
DO 16 T = 1, NPTS
RANKY(J, I) = DUMMY(IPT(T))
TPT(J, I) = IPT(T)
IF (NPRINT .GT. 0) GC TO 16
WRITE(6W, 4) J, I, DUMMY(I), IPTY(J, I), RANKY(J, I), X (IPTY(J, I))
4 FORMAT (1H, T2, 1X, I3, 1X, F6.2, 9X, I3, 4X, F6.2, 2X, F6.2)
16 CONTINUE
17 CONTINUE
C
C IF (KEY .NE. 1) GO TO 20
C
C USER-SUPPLIED MIN/MAXES, BUT SAVE AND TEST DATA TO BE SURE.
SAYMAX = XMAX
SAYMIN = XMIN
SAYMAX = YMAX
SAYMIN = YMIN
CCC SCAN THROUGH THE DATA GETTING MIN, MAX VALUES.
C PURPOSELY RIDICULOUS FIGURES
20 XMAX = 1.2 - 37
XMIN = 1.2 + 37
YMAX = 1.2 - 37
YMIN = 1.2 + 37
C X VECTOR IS IN ORIGINAL ORDER, NO? RANKED
DO 22 I = 1, NPTS
XMAX = AMAX1(XMAX, X(I))
22 XMIN = AMIN1(XMIN, X(I))
Y VALUES IN RANK ORDER IN MATRIX, IN DESCENDING ORDER, YHI TO YLO
DO 24 J = 1, NY
YMAX = A MAX 1(RANKY(J,1),YMAX)
YMIN = A MIN 1(RANKY(J,NPTS),YMIN)

25 IT = (KEY . . NE. 1) GO TO 110

WE DO SUBSTITUTE OUR MIN/MAX FOR USER'S WHEN IN OBSERVED VALUE IS CLOSER
YMIN = A MIN 1(YMIN,SAYMIN)
YMAX = A MAX 1(YMAX,SAYMAX)

3 ALL OPTIONS REQUIRE RANGES, BUT TSCALE MAY CHANGE MIN/MAX VALUES.

PHONE'S COMPARISONS AFTER TSCALE

110 SAYMIN = XMIN
SAYMAX = XMAX

0 THESE ARGUMENTS TO PROVIDE TSCALE
SAYMIN = YMIN
SAYMAX = YMAX

2 ALI OPTIONS REQUIRE RANGES
RANGE FOR X IS A PLUS INCREMENT ACROSS THE ROW COL, BY COL, XLO TO XHI
DIFFX = XMAX - XMIN

RANGE FOR Y IS A PLUS INCREMENT DOWN EACH ROW FROM YHI TO YLO. NOTE MIN
STMIN IN THE FOLLOWING STATEMENT.

DIFFY = -(YMAX - YMIN)
IF (N OprNT .GT. 0) GO TO 123
WRITE(LW,1050) XMAX,YMAX,XMIN,YMIN

1050 FORMAT (1HC,10X,'XMAX = ','G15.4,' YMAX = ','G15.4/1H,10X,'XMIN = ',
'G15.4,' YMIN = ','G15.4)
WRITE(LW,1051) DIFFX, DIFFY, KEY

1051 FORMAT (1HC,10X,'XSPAN = ','G15.4,' YSPAN = ','G15.4,' KEY = ',I1)
WRITE(LW,1043) NWID, NU, LEN, FL

1049 FORMAT (1HC,10X,'NWID, NW = ','2I5,' LEN, NL = ','2I5)

123 IF (KEY .EQ. 1 .OR. KB Y .EQ. 2) GOTO 143

CALL TSCALE (XMIN,XMAX,XLADEL, NW,NWID)

C

NOTE THAT XMIN, XMAX MAY BE ADJUSTED BY TSCALE
SPAN = XMAX - XMIN
IF (SPAN .LE. .1E-1C) GO TO 499

C

IF (N OprNT .GT. 0) GO TO 141
WRITE(XMAX,NE,SAVMAX - XMIN,NE,SAVMIN)

1052 FORMAT (1HC,10X,'S SCALED XMAX AND SCALED XMIN = ','2G15.4)

141 RANGEX = SPAN/PLCAT(NWIDTH - 1)
TESTX = 0.5 * RANGEX
IF (KEY = 1 . OR. KEY .EQ. 2) GO TO 150

TSSCALE STARTS WITH MIN, VALUES, ADD 'EFFETTS UP' LABELS FO? X AND Y.
OPTIONAL CODE HERE TO SHRINK TSCALE GRAPH AT RIGHT SICE OA X AXIS(XLABEL)
SEE STMT. 146 F., BEOF FOR Y AXIS.
SET KEY = 3 IN CALLING PROGRAM IF WANT TO SHRINK GRAPH, ELSE
IF (KEY .EQ. 4) GO TO 146

IF LARGEST OBSERVED GREATER THAN OR EQUAL TO NEXT-TO-LAST XLABEL PLUS TEST
CANNOT SHRINK.
IF (SAVMAX .GE. (XLADEL(NWID) + TESTX)) GO TO 146

ELSE WE CAN SHRINK NWIDTH OF GRAPH, AT RIGHT SIDE OF X AXIS
CHANGES IN XMAX DO NOT AFFECT GRAPH, BUT MUST USE REVISED XMIN FOR (LO
POP STMT. 529 BELOW
NWID = NWID - 1
NW = NWID * 1
NWIDTH = NWID * NTEN + 1
ABOVE CODE TO ELIM. LAST AXIS BLOCK WHERE XMAX > SAVMAX*FLOAT(NTEN)*FANGEX

146 CALL TSCALE(SAVNIY, SAVMAY, XLABZL, NL, LEN)

150 SPAN = - (SAVMAX - SAVMIY)
IF (-SPAN .LE. .1E-10) GC TC 498
IF (HOPF NT .GE. 0) GO TO 151
151 RANGEY = SPAN / FLOAT(LENGTH - 1)
TESTY = .5 * RANGEY

GO TO (251, 251, 245, 250), KEY

CODE TO SHRUNK TSCALE GRAPH AT TOP OF Y AXIS (XLABZL(NL))
WHERE LEN = 13 PADD PAGE, NL = LEN + 1 (FCF YLABELS)
RECALL, TSCALE YLABS STILL IN LOW TO HIGH ORDER
COMPARE WITH TSSCALE'S NEXT TO LAST LABEL, THE OBSERVED YMAX.

245 IF (YMAX .GE. (YLABE(LLEN) + TESTY)) GC TC 250
SAVMAY = YLABE(LLEN)
LEN = LEN - 1,
LENGTH = LEN * NR + 1

250 CALL REVERS(YLABEL, NL)

NOT IF THE TSSCALE SHRUNK MUST DO THIS BEFORE CALL REVERS
IF (KEY .EQ. 3) NL = LEN + 1
ABOVE WILL ONLY CHANGE IN VALUE IF LEN HAS BEEN DIMINISHED

251 IF (HOPF NT .GE. 0) GC TC 3CC
151 WRITE(LW, 1053) SAVMAY, SAVMIY

WRITE(LW, 1054) RANGEX, RANGEY, KEY
1054 FORMAT('1HO, 10X, RANGEX = ', G15.4, ', RANGEY = ', G15.4, ', KEY = ', I1)

GETTING READY FOR Y AND X LABELS, TSSCALE'S AND CURS.

JYL = 0
KEY = 1 OR = 2
HE?D ORIGINAL, OBSERVED YMAX FOR YHT = EBLOW.
SEE ALSO, STMT. 110 ABECU
IF KEY = 3 OR = 4 XMIN = XMIN POSSIBLY REVISED BY TSSCALE

YHT = SAVMAY
IF (KEY .EQ. 3 OR. KEY .EQ. 4) GC TC 3C6

IF KEY = 1 OR 2, SET XLABELS HERE
XLABEL (1) = XMAX
DO 305 I = 2, NW
I1 = I - 1
LAST XLABEL WILL CONTAIN XMAX VALUE
XLABEL(I) = XLABEL(I-1) + (FLOAT(WTEN) * SANGE)

305 CONTINUE

306 IF (NOTPRT .EQ. 1) GC TC 310

CCC THIS IS PAGE PRINTED WHEN NOTPRT = 2 (CR = 0), BUT NOT WHEN NOTPRT = 1

CCC FORCE HEADERS
CALL PAGER (100)
WRITE(LW, 6001)

6001 FORMAT (1H , 10X, 'THIS PROGRAM WAS WRITTEN IN FORTRAN IV FOR THE IBM
1 360/31 AT/1H, 6X, 'PRINCETON UNIVERSITY, SPRING TERM 1975, BY AL
2 ICE ANNIE NAVIN AND/1H, 6X, 'SHERMAN FOEINSON, REVISIONS TO SCALE/PSCA
3LE ROUTINES WRITTEN 11/70 '/1H, 6X, 'BY B. N. REIL, UNIVERSITY OF MARYLAND.' /
51H, 10X, 'IT REQUIRED REGION=18K AND T= .16 MINUTES FOR THE MAXIMUM'
6/1H, 6X, 'NPTS, NY (4), NWIDTH (101), AND LENGTH (UP TO 101 ROWS)
7 FOR A '/1H, 6X, 'TWO PAGE GRAPH.' /)

WRITE(LW, 6003) NWIDTH, LENGTH

6003 FORMAT (1H, 10X, 'THE FOLLOWING GRAPH IS ', i3, 'COLUMNS WIDE ON THE'
1HORIZONTAL, ', X'/1H, 6X, 'AXIS, AND ', i4, 'ROWS LONG ON THE VERTICAL,
2 Y AXIS. ')

IPLOT = NY = NPTS

WRITE(LW, 6004) NPTS, NY, IPLOT, KEY

6004 FORMAT (1H0, 10X, 'THE NUMBER OF OBSERVATIONS, NPTS, ARE ', i5, 'AND '
1 THE NUMBER '/1H, 6X, 'OF Y'S, NY, ARE ', i1, ', NY=NPTS (PLOTTED FOR'
2 NPTS) SHOULD BE ', i31, '. '/1H, 6X, 'THE OPTION KEY WAS SET TO ', i3, '
3.' )

CALL PAGER (100)

110 CONTINUE

CCC 11 IS MAX. NO FOR + FOR XLABELS, PER PAGE IE. 132 CCLS . PCB NWIDTH =
CCC 101 INCLUSIVE.
WRITE(LW, 6010) ((LCOLA(K), K=1, 3), N=1, NWID)

6010 FORMAT (1H , 17X, 1H+, 11, (2A4, A2) )

CCC THE BIG LOOP IS DO 100, IE. PCB BY PCB START TO END
CCC TOP ROW BY ALL COLUMNS, XLC TO XHI, 2ND PCB BY ALL CCLS., ETC.....

DO 1000 JRCW = 1, LENGTH

COUNT LINES, ETC. PER PAGES, LABELS, ETC.
IF (MOD(JFOW,NRP) .EQ. 0 .AND. JRCW .EQ. 11. LENGTH) CALL PAGER (100)
JBP = JPCW + NR - 1

ONLY INCREMENT EVERY FIFTH ROW POP YLABELS
IF (MOD(JB, 9) .EQ. 0) JBP = 1
IF (MOD(JB, 9) .NE. 0) JBP = 2

SEE STMT. 39C ABOVE, JYLP = 0.
JYLB = JYLB - JBP + 2

DO 512 NCOL = 1, NWIDTH
LINE (NCOL) = BLANK
JL (NCOL) = 0

512 CONTINUE
IF (JPOW .GT. 1) PHI = YHI + RANGEY

DO 700 J = 1, NY
   NWATCH(1) = NMARK(J)
   IF (JPOW .EQ. 1) LOI(J) = 0
   NUM1 = LOI(J) + 1
   IF (NUM1 .GE. NPTS) NUM1 = NPTS

DO 603 I = NUM1, NPTS
   RECALL, PANKY'S ALREADY IN DESCENDING ORDER, HIGH TO LOW.
   RECALL, IFIY = (YMAX - YMIN)
   IF (PANKY(J, I) .GT. (YHI - TESTY) .GE.
      1 PANKY(J, I) .LE. (YHI + TESTY)) GO TO 700

LOI(J) = I

AT THE START OF EACH NEW RCW, FIRST COLUMN
X50 = XMIN

DO 550 NCOL = 1, NWIDTH
   COLUMN BY COLUMN POSITIVE INCREMENT FOR X.
   IF (NCOL .GT. 1) XLO = XLO + RANGEX
   IF (X(IPTY(J, I)) .LT. (KLE - TESTX) .GE.
      1 X(IPTY(J, I)) .GE. (XLO + 'IESTX)) GO TO 550

WE HAVE A FCNHT
CCC MUST P überhaupt ANY AND ALL TIES FOR THIS (JRCW, NCCL)
   JL(NCOL) = JL(NCOL) + 1
CCC IF MORE THAN NINE TIES, WE MARK THE GRAPH WITH A ZER
   IF (JL(NCOL) .GT. NINE) JL(NCOL) = 10
   LINE(NCCL) = NWATCH(JL(NCOL))

550 CONTINUE

630 CONTINUE

700 CONTINUE

NOW, TO PRINT THIS LINE OF BLANKS AND/OR MARKS, WITH APPROPRIATE YLABEL:
AND/OR Y AXIS NOTATION.
CCC JJR IS SET ABOVE STMT. 512 ABOVE FOR + CP | AT START OF EAC
   IF (JJP .NE. 1) GO TO 982
CCC JYLB ALSO SET ABOVE STMT. 512 ABOVE.
   IF (KEY .EQ. 1 .OR. KEY .EQ. 2) YLABEL(JYLB) = YEL
   WRITE(LW, 985) YLABEL(JYLB), LRCWA(JJR), (LINE(N), N=1, NWIDTH).
   1 LRCWA(JJR)
985 FORMAT (1H, 1X, G15.4, A1, 101A1, A1)
   GO TO 1000

982 WRITE(LW, 983) LRCWA(JJP), (LINE(N), N=1, NWIDTH),
   LRCWA(JJR)
983 FORMAT (1H, 16X, A1, 101A1, A1)
CCC END OF A ROW IN BIG LOOP, DC 1000 JRCW = 1, LENGTH
1000 CONTINUE
ALL ROWS PLOTTED, NEW XAXIS AND XLABELS PRINTED AT BOTTOM OF GRAPH
WRITE(LW,6010) ((LCOLA(K), K=1,3), N=1,NWID)

WILL THE XLABELS FIT IN 10 MAX. SPACES ??
CANNOT KEEP LABELS IN G11.4 FORMAT IF ICHK = C
ICHK = 0
ICHK2 = 0
DO 950 I = 1, NW
IF (ABS(XLABEL(I)) .LT. 0.1) ICHK2 = ICHK2 + 1
IF (ABS(XLABEL(I)) .GE. 10000.0) ICHK = 1
950 CONTINUE
IF (ICHK .EQ. 2) ICHK = 1
IF (ICHK .EQ. 0) GC TC 953

NOTE WE PRINT THE XLABELS ON TWO LINES IN A STAGGERED WAY.
WRITE(LW,6011) (XLABEL(I), I=1,NW,2)
WRITE(LW,6014) (XLABEL(I), I=2,NW,2)
GO TO 954

WRITE(LW,6013) (XLABEL(I), I=1,NW)
WRITE(LW,6013) (XLABEL(I), I=1,NW)
GO TO 954
CONTINUE
GRAPH IS COMPLETED

SHOW GRAPH NOTATIONS PCR EACH OF U POSSIBLE Y'S (WHEN NO TIES).
WRITE(LW,6005) (J,MARK(J), J=1,4)
6005 FORMAT(1H0, 6X,4('Y(1,4)', 1S '1, A1, 3Y') )
GO TO 999

ERROR MESSAGE FETUFS ***

CONTINUE
WRITE(LW,198)
WRITE(LW,199) LENGTH, WIDTH, NY, NPTS, KEY
WRITE(LW,195) CALLING SEQUENCE ERROR IN CALL TC PLOT/
1 1H , 'SPECIFIED LENGTH = ', I11/
2 1H , 'SPECIFIED WIDTH = ', I11/
3 1H , 'NO. OF ORDINATES = ', I11/
4 1H , 'DATA POINTS = ', I11/
5 1H , 'OPTION CODE', KEY = ', I1
WRITE(LW,199) 1
GO TO 999

WRITE(LW,198)
WRITE(LW,195) SAVMAX,SAVMIN,XMAX,XMIN,DIFFY
WRITE(LW,195) 'ERROR IN SPAN, SAVMAX = ', X1PE10.3, ' SAVMIN = ', X1PE10.3, ' DIFFY = ', X1PE10.3/
WRITE(LW,198) 1
GO TO 999

WRITE(LW,198)
WRITE(LW,195) YMAX,YMIN,SAVMAX,SAVMIN,DIFFY
WRITE(LW,195) 'ERROR IN SPAN, YMAX = ', G15.4, ' YMIN = ', G15.4, ' SAVMAX = ', G15.4, ' SAVMIN = ', G15.4, ' DIFFY = ', G15.4/
WRITE(LW,198) 1
GO TO 999

RETURN
SUBROUTINE RANKHT(Y,NPTS,IPT)

A REVISED COPY OF A RANK LOW TO HIGH ROUTINE BY LARRY E. WESTPHAL.

SUB.RANKHT YIELDS RANKING FROM HIGH TO LOW OF ITEMS BASED ON VECTOR Y VA.

IRK(I) IS RANK OF I'TH ITEM

IPT(I) IS THE ITEM WHOSE RANK IS I

2/13/75, SUB. PLOT DOES NOT USE TBP

REAL IRK

SINCE MAX. NPTS = 100, DIMENSION ITY(100), IRK(100) 2/19/75

DIMENSION Y(NPTS),IPT(NPTS),IRK(100),ITY(100)

DO 2 I = 1, NPTS
   IRK(I) = 0.0
2    IPT(I) = 0
   I = 1
14   R = - 1000.0000
   LT = 0
   IY = 0
   DO 20 L = 1, NPTS
20   ITY(L) = 0
   DO 10 L = 1, NPIS
      IF (IPK(L) .NE. 0.0) GO TO 10
   IF (R - Y(L)) = 11, 12, 10
11   LT = L
   R = Y(L)
   IY = 0
   DO 21 L1 = 1, NPTS
21   ITY(L1) = 0
   GO TO 10
12   IY = IY + 1
   ITY(IY) = L
10 CONTINUE

ITY (LT .EQ. 0) GO TO 998
RI = .FLOAT(I) + (.FLOAT(IY) / 2.0)
IPK(LT) = RI
IPT(I) = LT
   IF (IY .EQ. 0) GO TO 22
   DO 23 L1 = 1, L
      IFK(ITY(L1)) = RI
      I = I + 1
22    I = I + 1
23   IPT(I) = ITY(L1)
   IF (I .LE. NPTS) GO TO 14
   GO TO 209
998 WRITE(6,997)
997 FORMAT(1HO,9H**********/1H, 'ERCR IN RANKHT'/)
      11H,*H**********/;
999 RETURN

END

SUBROUTINE REVERS(YLABEL,NL)

WHERE YLABELS FROM TSCALE ARE IN ASCENDING ORDER, CALL REVERSE(YLABEL,NL)!

CC Works for unmatched pairs as well as matched pairs, since center would stay put.

CCC WRITTEN SPRING 1975, AAN/NS.

DIMENSION YLABEL(NL)

NTEST = NL / 2
K = NL - 1
SAVE = YLABEL(J)
YLABEL(J) = YLABEL(K)
YLABEL(P) = SAVE

349 CONTINUE
RETURN
END

SUBROUTINE ISCALE(VMIN, VMAX, VSCA, NDIM, NAXIS)

C SUBROUTINE TO SCALE A PLOT.
C WRITTEN BY B. - RIBD, UNIV. OF MARYLAND, 1975. REVISED 2/75 SR/AAN
C
C THE INPUT VARIABLES
C
C VMIN AND VMAX CONTAIN THE OBSERVED VALUES OF MINIMUM/MAXIMUM
C VALUES OF THE VARIABLE TO BE SCALeD. ON OUTPUT, THEY WILL
C CONTAIN THE ADJUSTED VALUES TO BE USED FOR 'NICE' NUMBERS
C ALONG THE AXIS OF THE PLOT.
C
C VSCA IS THE ARRAY WHICH WILL CONTAIN AXIS LABELS. NAXIS IS
C THE NUMBER OF AXIS BLOCKS, SO THAT THESE WILL BE 'NAXIS+1'
C
C ENTRIES MADE IN VSCA.
C
NDIM = NAXIS + 1
DIMENSION VSCA(NDIM), XSCF(12)
DATA XSCF/C,, 1., 1.25, 1.5, 2., 3., 4., 5., 6., 7.5, 8., 10. /
NLOOP = C
SPAN = 0.0
NSCPF = 12
NSM1 = NSCPF - 1
ZFC = VM T

650 SCP = (VMAX - VMIN) / FLCAT(NAXIS)
MINSIG = ISIGN(1, I?1X(VMIN))
DX1 = 0.
IF (SC? .LE. 0.0) GO TO 120
XNC = ALOS10(SC?)
INC = XNC
IF (XNC .LE. 0.) INC = INC - 1
II(INC.G'1.0) XNC10 = 10.*INC
IF (INC .LE. 0.) XNC10 = -(10.*ABS(FLCAT(IVC))))
DX = ?SCALE(SCF, -XNC10)
DO 651 TSC = 1, NSM1
651 IF (DX .GT. XSCF(ISC)) DX1 = XSCF(ISC + 1)

C A PROBABLE SCALE FACTOR HAS BEEN COMPUTED. LET US
C NOW TRY TO FIND A MINIMUM. (THIS CODE IS TAKE 3 OF
C MINIMUM-FINDING.)

120 CONTINUE
IF (VMIN .LE. ZFC) GO TO 129
VMIN = VMIN - DX1
GO TO 120
129 CONTINUE
SPAN = VMIN + FLAT(NAXIS) * DX1
IF (SPAN .GE. VMAX) GO TO 6U9
C
C WE HAVE MOVED THE MINIMUM DOWN SO FAR THAT THE INCREMENT
C DOESN'T WORK ANY MORE

NLOCP = NLCF + 1
IF (NLOOP .LE. 2) GO TO 650

120 WRITE(6, 124)
124 FORMAT(1HC, 9H**********/1H, 'ERROR IN ISCALE ROUTINE'/,
1 1H, 9H**********/)

649 CONTINUE
DX = DX1
VMAX=SPAN
DO 655 IH = 1, NDIM
VSCA(IH) = VMIN + FLOAT(IH - 1) * DX
655 CONTINUE
RETURN
END

FUNCTION PSSCALE(MANTIS,CHARAC)
C FUNCTION TO SCALE BY EXPONENTIAL SCALE FACTOR
C WRITTEN BY B.K. EID, UNIV. OF MARYLAND, 1970.
REAL YANTIS
IF(CHARAC.GT.0) PSSCALE=MANTIS*CHARAC
IF(CHARAC.LT.0) PSSCALE=MANTIS/ABS(CHARAC)
RETURN
END
References


