

AD-A132 722

A USER'S GUIDE TO BISAM (BIVARIATE SAMPLE): THE
BIVARIATE DATA MODELING P. (U) TEXAS A AND M UNIV
COLLEGE STATION DEPT OF STATISTICS T J WOODFIELD
AUG 83 TR-A-24 ARO-20140. 4-MA

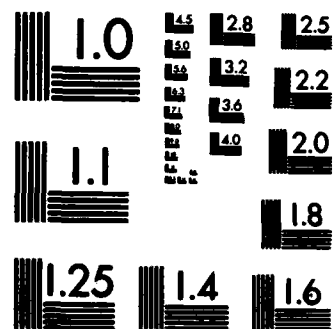
1/1

UNCLASSIFIED

F/G 9/2

NL

END



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

ARO 20140.4-MA

TEXAS A&M UNIVERSITY

COLLEGE STATION, TEXAS 77843-3143

(12)



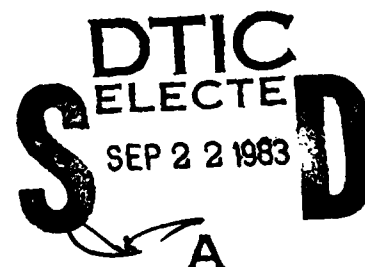
STATISTICS
Phone 409 - 845-3141

A USER'S GUIDE TO BISAM: THE BIVARIATE
DATA MODELING PROGRAM

by Terry J. Woodfield
Department of Statistics
Texas A&M University

Technical Report No. A-24
August 1983

Texas A&M Research Foundation
Project No. 4858



"Functional Statistical Data Analysis and Modeling"

Sponsored by the Army Research Office

Contract DAAG29-83-K-0051

Professor Emanuel Parzen, Principal Investigator

Approved for public release; distribution unlimited.

83 09 20 027

AD-A238722

DTIC FILE COPY

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

association along with some entropy measures that provide diagnostics for model selection and testing for independence. Graphical display of the estimated bivariate dependence density and density-quantile is the responsibility of the user, but a two step procedure utilizing the GCONTOUR and G3D procedures of SAS/GRAPH will be suggested in a later section. Different installations will have different software to provide three dimensional plots, hence the plotting option has been excluded from BISAM. One of the advantages of the modeling approach adopted by the BISAM program is the ability to obtain quickly and efficiently a set of function values for a dense grid of bivariate points promoting the examination of three dimensional representation of functions of interest.

Accession For

NTIS	GRA&I	☒
DTIC TAB		☐
Unannounced		☐
Justification		

DTIC
COPY
RECEIVED

A

S/N 0102- LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

A USER'S GUIDE TO BISAM: The Bivariate Data Modeling Program

I. Introduction

The BISAM (Bivariate SAMple) program is a companion program to the ONESAM (Parzen and Anderson, 1980) and TWOSAM (Prihoda, 1981) programs developed for nonparametric data modeling. The main purpose of BISAM is to perform bivariate data analysis using Fourier expansions and quantile techniques. The motivation and theory behind the data modeling approach incorporated by BISAM is detailed in Woodfield (1982), with foundations provided by Parzen (1979), Tartar and Kronmal (1970, 1976), and Kimeldorf and Sampson (1975).

BISAM is a FORTRAN program composed of a main program and 37 subprograms. Several of the subprograms come from ONESAM and the TIMESBOARD Time Series Subroutine Library (Newton, 1979). The main feature of BISAM is the ability to provide estimates of the bivariate density-quantile function of a set of bivariate data. A univariate analysis is also provided similar to that of ONESAM but with considerably less output. The univariate analysis is not intended to replace a ONESAM analysis if such a detailed analysis is required. The bivariate analysis includes display of various nonparametric measures of association along with some entropy measures that provide diagnostics for model selection and testing for independence. Graphical display of the estimated bivariate dependence density and density-quantile is the responsibility of the user, but a two step procedure utilizing the GCONTOUR and G3D procedures of SAS/GRAPH will be suggested in a later section. Different installations will have different software to provide

three dimensional plots, hence the plotting option has been excluded from BISAM. One of the advantages of the modeling approach adopted by the BISAM program is the ability to obtain quickly and efficiently a set of function values for a dense grid of bivariate points promoting the examination of three dimensional representation of functions of interest.

This document will explain the algorithms employed in generating output and the steps necessary to perform an analysis of a set of bivariate data. The options available and the user input required to execute the BISAM program will also be described in detail. The user should be thoroughly acquainted with the input options before attempting to execute the BISAM program. Incorrectly specified option values may cause errors. Furthermore, BISAM is a very long and complex program that will be fairly expensive to use, and hence improper runs should be avoided.

2. Univariate Analyses

The analysis of each of the paired variables separately is essential to fully understand the nature of the bivariate data set. In particular, the estimation of univariate density-quantile functions should be performed in an optimal manner to insure that the estimated bivariate density-quantile function is appropriate. A univariate analysis will provide diagnostics to help one understand the nature of the univariate density-quantile function and thus guide one in the formulation of the bivariate density-quantile.

BISAM provides output for each variable that is a subset of what may be provided by ONESAM. The univariate analysis is done in two stages, however. In the first stage, descriptive statistics, an Informative Quantile plot, and a plot of $\tilde{D}(u)$ are provided in a goodness-of-fit exploration analysis to guide one in determining the underlying marginal distributions.

In the second stage, a univariate density-quantile function is estimated by the autoregressive method for the null case specified and is then used to form the bivariate density-quantile function as described in section 4. If $\tilde{D}(u)$ in stage one does not approximate a uniform distribution function for the null case specified, then one should specify another null value, or the bivariate density-quantile function estimated will be unreliable.

For a complete description of some of the above concepts, see Parzen (1979).

3. Bivariate Analysis

A bivariate analysis of a set of data includes computation of various nonparametric correlation coefficients, model selection diagnostics, and estimation of the bivariate dependence density and density quantile functions for a dense grid of points. Scatter plots are produced for the original data and the rank transformed data. Obtaining "publication quality" plots depends on the plotting hardware and software available on the system used.

Five correlation coefficients are produced along with various entropy measures of association. We will assume that the input data set is denoted by $(X_1, Y_1), \dots, (X_n, Y_n)$. Pearson's product moment correlation coefficient is defined by

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}},$$

and is labeled "PEARSON" in the BISAM output. Spearman's rank correlation coefficient is given by

$$\rho_n = 1 - 6 \sum_{i=1}^n (Q_i - R_i)^2 / [n(n^2 - 1)],$$

where $Q_i = \text{rank}(X_i)$ and $R_i = \text{rank}(Y_i)$ and is labeled "SPEARMAN" in the BISAM output. Three correlation coefficients related to the concepts of concordance and discordance are also computed, two of which perform corrections for tied observations. Kendall's Tau is defined by

$$\tau_A = (N_C - N_D) / [n(n-1)/2] ,$$

where N_C is the number of concordant pairs and N_D is the number of discordant pairs, and is labeled "KENDALL A" in the output. Kendall's TAU correcting for ties is given by

$$\tau_B = (N_C - N_D) / \sqrt{(N_C + N_D + T_X)(N_C + N_D + T_Y)},$$

where T_X is the number of tied observations in the X variable but not in the Y variable, and T_Y is the number of tied observations in the Y variable but not in the X variable. This measure is labeled "KENDALL B" in the output. Somer's d is defined by

$$d = (N_C - N_D) / (N_C + N_D + T_Y) ,$$

and is seen to be similar to Kendall's Tau-B. This measure of association is labeled "SOMER'S D" in the output.

The ranking procedure employed assigns average ranks for tied observations. Other methods for assigning ranks to tied observations are often employed but are not attempted by BISAM. The presence of a large percentage of tied

observations will weaken the results obtained since underlying continuous distributions are assumed. One should avoid such situations if possible.

Two methods of bivariate density estimation are performed by BISAM, the nearest neighbor technique and the orthogonal expansion technique of Woodfield (1982). However, estimates obtained from the nearest neighbor technique are not displayed, but instead are used as input values for the orthogonal expansion technique. Since the density estimation methods employed are fundamental to the bivariate analysis performed, the next section will discuss these techniques with considerable attention paid to the theory and implementation of such methods. Since the entropy measures of association and model selection diagnostics are so closely related to the density estimation technique employed, the discussion of these quantities will also be withheld until the next section.

4. Nonparametric Density Estimation

Nonparametric density estimation is performed three different ways by BISAM for different purposes. The k-nearest neighbor technique is employed first for $k=6$ to provide an estimate of the dependence density as input to an orthogonal expansion density estimation technique and to produce a raw estimate of the entropy associated with the joint probability density function (p.d.f.) and the marginal p.d.f.'s. The orthogonal expansion method of Woodfield (1982) is then employed to obtain smoothed estimates of the dependence density for various degrees of smoothing. Finally, autoregressive estimates of the marginal p.d.f.'s are obtained and used with the estimated dependence density to produce an estimate of the bivariate density-quantile function. We now elaborate on the details of these procedures.

Let $(X_1, Y_1), \dots, (X_n, Y_n)$ be a bivariate random sample from the random vector (X, Y) with joint cumulative d.f. $F_{X,Y}$, marginals F_X, F_Y , joint p.d.f. $f_{X,Y}$, marginal p.d.f.'s f_X, f_Y , and quantile functions Q_X, Q_Y . Define the dependence distribution function $D(u_1, u_2)$ by

$$D(u_1, u_2) = F_{X,Y}(Q_X(u_1), Q_Y(u_2)), \quad 0 \leq u_1, u_2 \leq 1,$$

and the dependence density function $d(u_1, u_2)$ by

$$d(u_1, u_2) = \frac{\partial^2}{\partial u_1 \partial u_2} D(u_1, u_2) = \frac{f_{X,Y}(Q_X(u_1), Q_Y(u_2))}{f_X(Q_X(u_1)) f_Y(Q_Y(u_2))}.$$

Define the information between two densities f and g by

$$I(f;g) = \int_{-\infty}^{\infty} \left\{ \log \frac{f(x)}{g(x)} \right\} f(x) dx$$

and the entropy of a density f by

$$H(f) = \int_{-\infty}^{\infty} -\{\log f(x)\} f(x) dx.$$

The information inequality states that for two densities f and g ,

$$I(f;g) \geq 0.$$

Furthermore, $I(f;g) = 0$ iff $f=g$ a.e.

It is easy to see that

$$I(f_{X,Y}; f_X f_Y) = -H(d) ,$$

which justifies naming $d(u_1, u_2)$ the dependence density, since by virtue of the information inequality $d(u_1, u_2)$ is indirectly related to a measure of dependence between X and Y . This fact is exploited in obtaining diagnostics for model selection and tests of independence.

Let $\{\phi_k(u)\}_{k=-\infty}^{\infty}$ be a complete orthonormal system of functions in $L^2(0,1)$. Then the system $\{\phi_{jk}(u_1, u_2)\}_{j,k=-\infty}^{\infty}$ defined by

$$\phi_{jk}(u_1, u_2) = \phi_j(u_1)\phi_k(u_2) , \quad 0 \leq u_1, u_2 \leq 1, \quad \text{all } j, k,$$

is a complete orthonormal system in the space of bivariate square integrable functions on the unit square. If $\log d(u_1, u_2)$ is square integrable, then

$$\log d(u_1, u_2) = \sum_{j,k=-\infty}^{\infty} \theta_{jk} \phi_j(u_1) \phi_k(u_2) - \psi(\theta)$$

in the sense of L^2 norm where $\{\theta_{jk}\}_{j,k=-\infty}^{\infty}$ are the Fourier coefficients defined by

$$\theta_{jk} = \int_0^1 \int_0^1 \phi_j(u_1) \phi_k(u_2) \log d(u_1, u_2) du_1 du_2, \quad j, k = -\infty, \dots, \infty .$$

The term $\psi(\theta)$ is included to insure that $d(u_1, u_2)$ integrates to one. For the truncated m -th order model given by

$$\log d_m(u_1, u_2) = \sum_{j,k=-m}^m \theta_{jk} \phi_j(u_1) \phi_k(u_2) - \psi_m(\underline{\theta}) ,$$

it follows that $\log d_m(u_1, u_2) \rightarrow \log d(u_1, u_2)$ as $m \rightarrow \infty$, and hence one calls $d_m(u_1, u_2)$ the m -th order approximation of $d(u_1, u_2)$. One thus seeks an estimator of $d_m(u_1, u_2)$ based on a random sample of bivariate data.

Let

$$u_{1i} = R_i/(n+1), \quad u_{2i} = Q_i/(n+1)$$

where R_i and Q_i are the ranks of X_i and Y_i as defined before. Thus $(u_{11}, u_{21}), \dots, (u_{1n}, u_{2n})$ approximates a random sample distributed uniformly on the unit square under an assumption of independence. Dependence of X and Y suggests alternate uniform distributions that may include a variety of bivariate density shapes from distributions having uniform $(0,1)$ marginals. Next, form $\tilde{d}(u_1, u_2)$ evaluated at these uniform sample points based on the k -nearest neighbor techniques. BISAM uses $k=6$. Form $\log \tilde{d}(u_1, u_2)$ and use this as the dependent variable in a least squares regression routine with independent variables $\phi_j(u_1) \phi_k(u_2)$, $j, k = -m, \dots, m$. Essentially, both $\log \tilde{d}(u_1, u_2)$ and $\phi_j(u_1) \phi_k(u_2)$ represent bivariate functions imbedded into univariate representations for a multiple regression analysis. The design matrix X is an $n \times m^2$ matrix consisting of the m^2 orthogonal combinations evaluated at the n uniform data points. The routine performed by BISAM forms the $(2m+1)^2 \times (2m+1)^2$ correlation matrix* and then uses a SWEEP operator

*The $(0,0)$ term is incorporated into the integration factor and is not included, i.e., the no intercept m^{-1} is employed, hence, the correlation matrix is not $[(2m+1)^2+1] \times [(2m+1)^2+1]$.

on this matrix to obtain least squares estimates of the Fourier coefficients. The resulting estimate is given by

$$\log \hat{d}(u_1, u_2) = \sum_{j,k=-m}^m \hat{\theta}_{jk} \phi_j(u_1) \phi_k(u_2)$$

where $\{\hat{\theta}_{j,k}\}_{j,k=-m}^m$ is the collection of least squares estimates of the expansion parameters. The integration factor $\hat{\psi}_m(\hat{\theta})$ is then derived to insure that $d(u_1, u_2)$ numerically integrates to one. This is accomplished for $\hat{d}(u_1, u_2)$ equally spaced in the u_1 and u_2 directions within the unit square.

Three models are obtained for $m=1,2,3$ which translates into 8 variable, 24 variable, and 48 variable regression models, i.e., the terms with subscripts (i,k) , $i,k=-m,\dots,m$ for $m=1,2,3$ are included. These three models produce three density estimates evaluated at the 40×40 grid of (u_1, u_2) coordinates mentioned above. Using raw Riemann sums as numerical integrals, one then obtains estimates of the entropy $H(d)$ mentioned earlier. A fourth estimate of the entropy of the dependence density is provided by the nearest neighbor estimate using the formula

$$\begin{aligned} H(\tilde{d}) &= \int_0^1 \int_0^1 \log \tilde{d}(u_1, u_2) d D_n(u_1, u_2) \\ &= \frac{1}{n} \sum_{j=1}^n \log \tilde{d}\left(\frac{R_j}{n+1}, \frac{Q_j}{n+1}\right) \end{aligned}$$

where $D_n(u_1, u_2)$ is the empirical c.d.f. of (u_{1i}, u_{2i}) , $i=1, \dots, n$.

As a rule-of-thumb for model selection, $H(\hat{d})$ is compared with $H(\hat{d}_8)$, $H(\hat{d}_{24})$, and $H(\hat{d}_{48})$. An information criterion modeled after Akaike's Information Criterion (AIC) familiar to time series analysts is used to select the "best" model. One forms

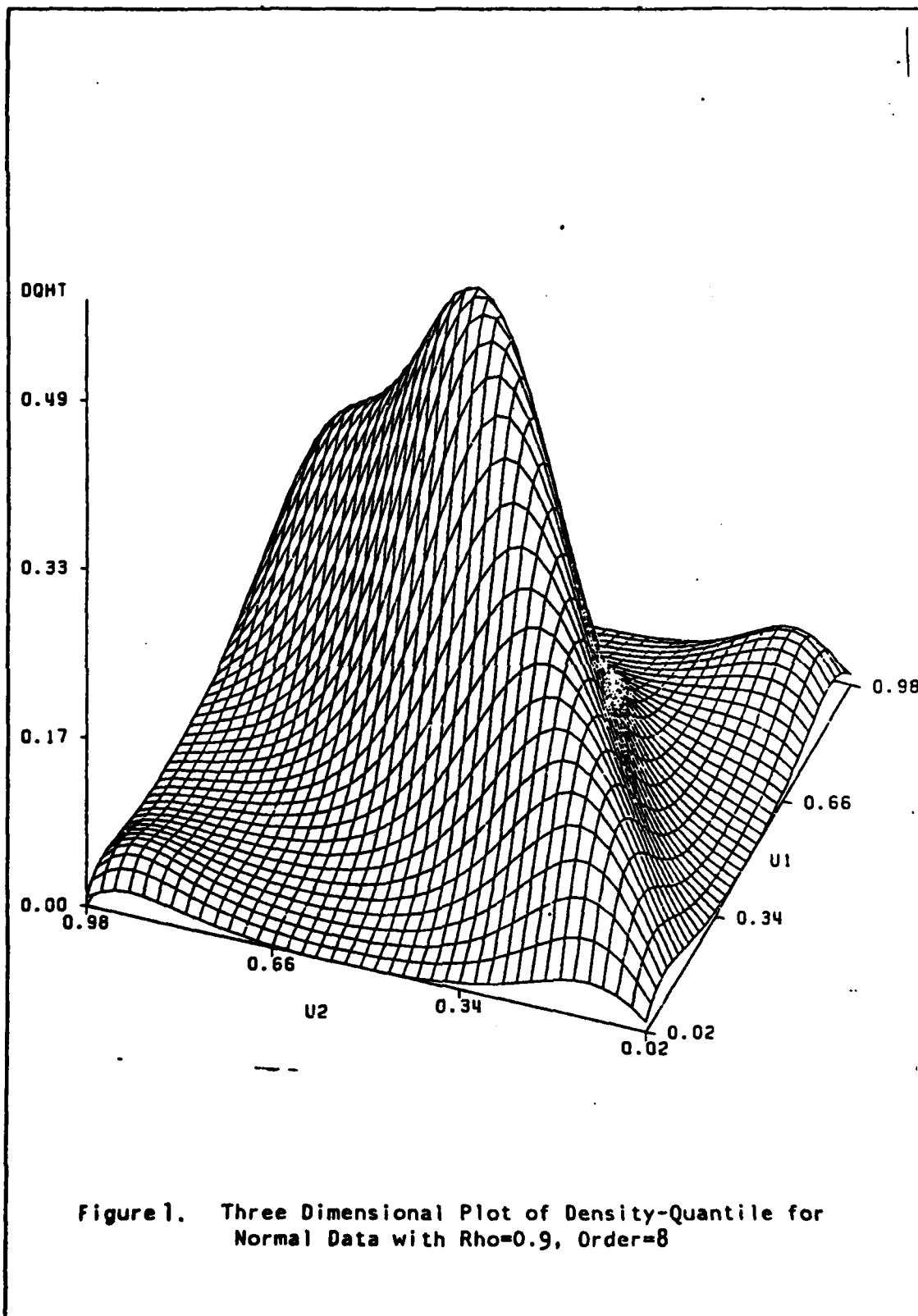
$$AIC(k) = H(\hat{d}_k) - H(\tilde{d}) - 2k/n \quad ;$$

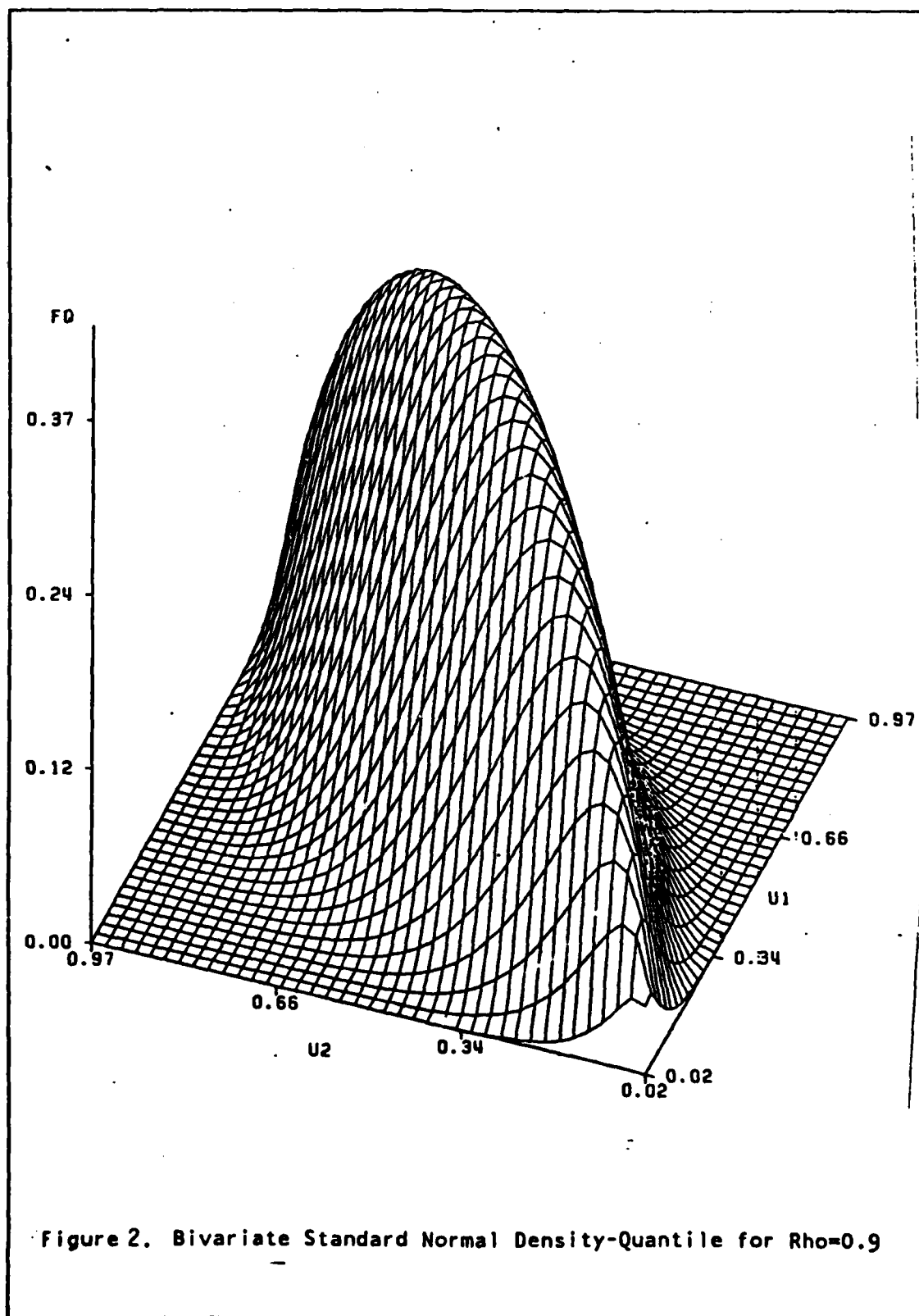
and then selects model k for the k value at which $AIC(k)$ achieves its positive minimum. An alternative approach chooses k for which $|AIC(k)|$ is minimum. Model selection criterion are still under investigation with promising results having been achieved for the criterion employed by BISAM.

Graphical displays of $\hat{d}_k$ suffer from the unruly behavior of $\hat{d}_k$ near the boundary of the unit square, and hence one may prefer to display values of the bivariate density-quantile function given by

$$\hat{fQ}(u_1, u_2) = \hat{d}_k(u_1, u_2) \hat{fQ}_X(u_1) \hat{fQ}_Y(u_2)$$

where $\hat{fQ}_X$ and $\hat{fQ}_Y$ are the autoregressive estimates of the univariate density-quantile functions. Figure 1 contains a three dimensional plot of the bivariate density-quantile function of order 8 for a set of simulated bivariate normal random variables with correlation coefficient $\rho = 0.9$. Figure 2 contains the theoretical bivariate density-quantile function for the simulated data. The plots were produced using SAS/GRAPH. Appendix B outlines a two step FORTRAN-SAS procedure that allows one to produce three dimensional or contour plots of bivariate functions of interest.





5. Using BISAM to Analyze Time Series Data

Although primarily intended to analyze bivariate data in which the data values are uncorrelated, BISAM can also be used to analyze time series data. For a univariate data set one may use option IUNIV to do an analysis of a bivariate data set consisting of the univariate data set paired with lagged values of the data set.

Consider a Gaussian time series $X(t)$ with autocorrelation function $\rho(v)$ given by

$$\rho(v) = \text{Corr}[X(t), X(t+v)], \quad v=0, \pm 1, \dots$$

Define the series $Y(t)$ to be a lagged version of $X(t)$, i.e., $Y(t) = X(t+v)$. Then

$$\rho_{X,Y} = \text{Corr}[X,Y] = \text{Corr}[X(t), X(t+v)] = \rho(v).$$

For a bivariate normal random vector (X,Y) with correlation coefficient $\rho_{X,Y}$, it can be shown that

$$I(f_{X,Y} ; f_X f_Y) = -.05 \log (1-\rho_{X,Y}^2)$$

Consequently, the information between X and lagged values of X is

$$I(v) = -0.5 \log (1-\rho^2(v))$$

In the non-normal case, this equation does not hold so that a nonparametric

estimator may be desired to indicate the relationship between X and lagged values of X .

When one specifies $IUNIV = 1$, values for $KLAG1$ and $KLAG2$ must also be specified. Then a bisam analysis is performed on $(X(t), X(t+v))$ for integer values v satisfying $KLAG1 \leq v \leq KLAG2$.

6. Input Options

The following options are input on the first data card in 11I5 format. They are input in the order listed, and if $NTAPE = 5$, the data set follows this card in the indicated format listed at the end of this section.

- NTAPE - number referring to DD statement describing the input data set.
- IDQX - null distribution for autoregressive smoothing for X input variable.
- IDQY - same as IDQX except for Y input variable.
- MORD - maximum autoregressive order to be used for univariate autoregressive density estimation (≤ 6).
- IPLT1 - 0 for no scatter plots.
1 for scatter plot of data.
2 for scatter plot of rank transformed data.
3 for both scatter plots.
- IPLT2 - 0 for no univariate density plots.
1 for best order AR univariate density plots.
- IDST - 0 for no univariate descriptive statistic.
1 for descriptive statistics for both X and Y .

- KDEL - maximum number of extreme points to exclude from bivariate analysis. An extreme point is located based on the distance of the X or Y coordinate from its median. If KDEL=2 and both the extreme X and Y are paired together, then only that point will be omitted. If KDEL=3, two points with an extreme X value and one point with an extreme Y value will be omitted, i.e., the X-direction received precedence over the Y-direction for odd values of KDEL. This method of deleting outliers works well for nearly linear relationships between X and Y. Its use is questionable for other cases.
- IOUTD - 0 if function values are not to be saved.
1 if function values for the three fitted models are to be written to tapes 1, 2, and 3 respectively.
If IOUTD=1, the JCL must contain three DD cards for FT01F001, FT02F001, and FT03F001 defining permanent dish data files where the function values are to be written. Typically, 10 tracks of storage must be allotted for each file.
- IREG - 0 if no quantile regression performed.
1 if nonparametric estimates of $rQ(u) = E[Y|X = Q(u)]$ and $r(x) = E[Y|X=x]$ are desired.
- IUNIV - 0 if input data set is bivariate data.
1 if input data set is a univariate time series as described in section 5.

If IUNIV=1, a second data card is required containing the values KLAG1 and KLAG2 in 2I5 format as described in section 5.

If IDQX and/or IDQY specify the Pareto to Weibull distributions, a separate data card must be inserted containing estimates of the parameters in those distributions. For the Weibull, the location estimate is given first, with the input format specified as 2F10.0.

The input data set consists of two univariate data sets containing the same number of values. It is always assumed that the X data set is read in first. The values (X,Y) are then paired by the order they are read in. A univariate data set consists of a title card read in 20A4 format, a description card containing the number of data points and the format of the input data read in (I5, 4X, 5A4) format, and the data cards containing the data coded in the format indicated on card 3. If the number of data points indicated for X and Y do not match, the program will terminate.

Example of univariate data set:

YEARLY SNOWFALL IN BUFFALO, 1910-1972

63 (6F10.2)

71.2 69.5 47.8 58.4 29.9 42.5
etc.

The code for BISAM and sample output are contained in the Appendices.

7. Sample Output from BISAM

Following is a listing of the output from BISAM for a typical run with the input options clearly labeled. The JCL for executing BISAM at Texas A&M University is given in Appendix A.

.....
* BISAM - BIVARIATE DATA ANALYSIS USING FOURIER EXPANSIONS
* AND QUANTILE TECHNIQUES
*
* PROGRAMMER: TERRY J. WOODFIELD
*
.....

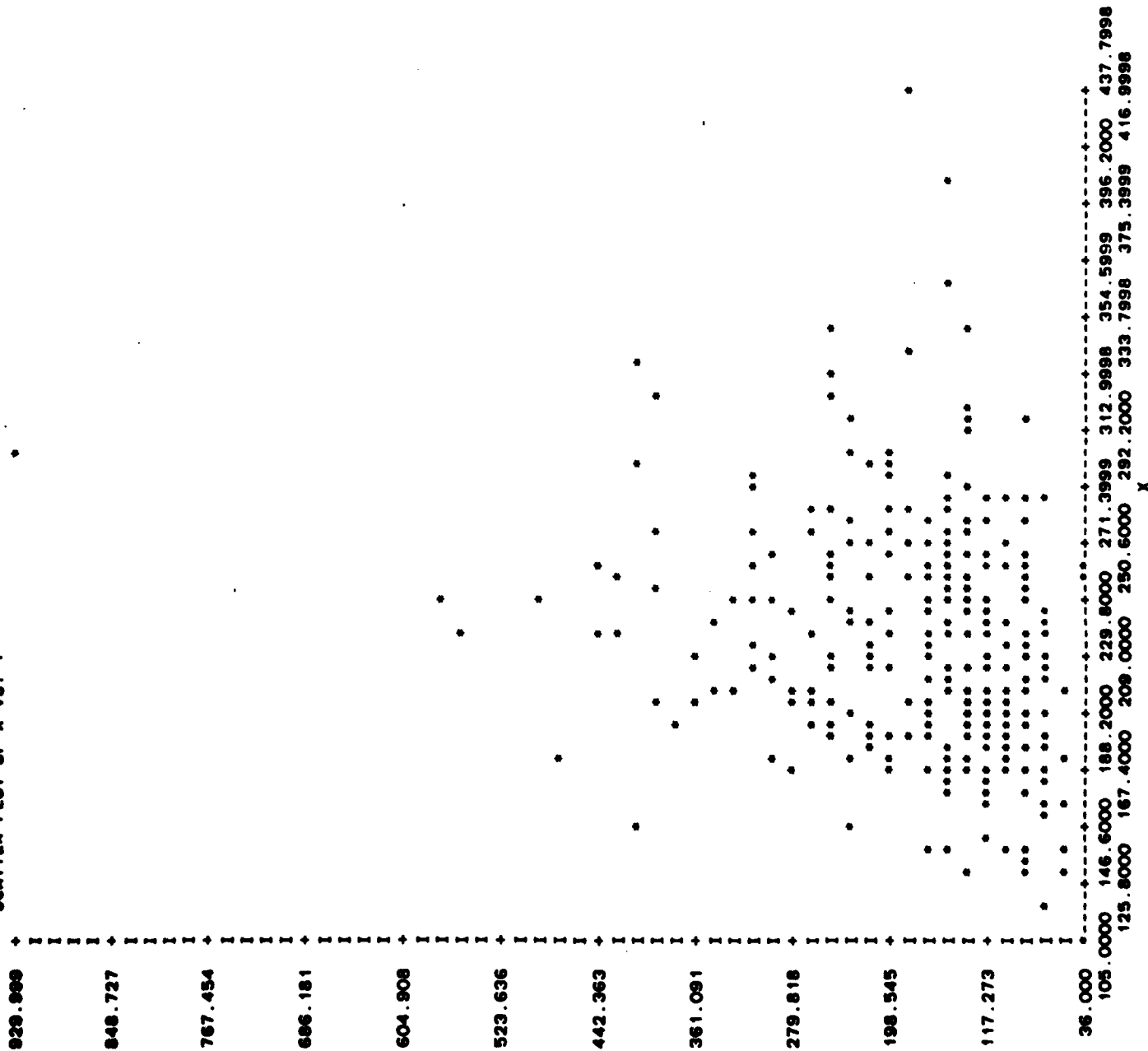
PLASMA CHOLESTEROL - DISEASE IN AT LEAST 1 OF 3 CORONARY ARTERIES
PLASMA TRIGLYCERIDES - DISEASE IN AT LEAST 1 OF 3 CORONARY ARTERIES

SAMPLE SIZE = 320

OPTIONS FOR THIS ANALYSIS:

NTAPE = 13	IDOX = 1	IDQY = 1
MORD = 2	IPLT1 = 3	IPLT2 = 1
IDST = 1	KDEL = 1	IOUTD = 0
IREG = 0	IUNIV = 0	

SCATTER PLOT OF X VS. Y



FULLY NON-PARAMETRIC ANALYSIS

PLASMA CHOLESTEROL - DISEASE IN AT LEAST 1 OF 3 CORONARY ARTERIES
ORIGINAL DATA - X

ORDER STATISTICS IN QUARTERS

SEQUENCE WITHIN QUARTILE *****	FIRST QUARTER *****	SECOND QUARTER *****	THIRD QUARTER *****	FOURTH QUARTER *****
1	105.0000	185.0000	213.0000	242.0000
2	119.0000	185.0000	213.0000	243.0000
3	131.0000	186.0000	214.0000	243.0000
4	131.0000	187.0000	215.0000	243.0000
5	131.0000	187.0000	215.0000	244.0000
6	138.0000	188.0000	216.0000	244.0000
7	139.0000	188.0000	216.0000	245.0000
8	139.0000	189.0000	217.0000	245.0000
9	140.0000	189.0000	217.0000	245.0000
10	140.0000	189.0000	218.0000	245.0000
11	140.0000	190.0000	218.0000	245.0000
12	142.0000	191.0000	218.0000	246.0000
13	144.0000	191.0000	218.0000	247.0000
14	149.0000	191.0000	219.0000	247.0000
15	150.0000	191.0000	219.0000	248.0000
16	151.0000	191.0000	219.0000	248.0000
17	157.0000	191.0000	220.0000	249.0000
18	159.0000	192.0000	220.0000	249.0000
19	160.0000	193.0000	221.0000	250.0000
20	162.0000	193.0000	221.0000	250.0000
21	163.0000	193.0000	221.0000	250.0000
22	164.0000	194.0000	221.0000	251.0000
23	164.0000	194.0000	221.0000	251.0000
24	165.0000	194.0000	222.0000	251.0000
25	165.0000	194.0000	222.0000	251.0000
26	165.0000	194.0000	222.0000	252.0000
27	165.0000	194.0000	222.0000	253.0000
28	167.0000	194.0000	222.0000	254.0000
29	168.0000	194.0000	223.0000	254.0000
30	168.0000	195.0000	223.0000	255.0000
31	168.0000	195.0000	224.0000	257.0000
32	168.0000	196.0000	225.0000	258.0000
33	169.0000	196.0000	226.0000	258.0000
34	169.0000	196.0000	227.0000	258.0000
35	170.0000	197.0000	227.0000	258.0000
36	171.0000	197.0000	227.0000	259.0000
37	171.0000	197.0000	228.0000	260.0000
38	171.0000	197.0000	228.0000	260.0000
39	171.0000	197.0000	228.0000	260.0000
40	171.0000	197.0000	229.0000	261.0000
41	171.0000	198.0000	229.0000	262.0000
42	171.0000	198.0000	230.0000	262.0000
43	172.0000	198.0000	230.0000	263.0000

44	172.0000	198.0000	230.0000	264.0000
45	172.0000	198.0000	230.0000	264.0000
46	172.0000	198.0000	230.0000	264.0000
47	173.0000	200.0000	230.0000	266.0000
48	173.0000	200.0000	230.0000	267.0000
49	174.0000	200.0000	231.0000	267.0000
50	175.0000	201.0000	231.0000	268.0000
51	175.0000	201.0000	231.0000	269.0000
52	175.0000	203.0000	232.0000	270.0000
53	175.0000	204.0000	232.0000	271.0000
54	176.0000	204.0000	232.0000	271.0000
55	176.0000	206.0000	232.0000	273.0000
56	177.0000	206.0000	232.0000	274.0000
57	178.0000	206.0000	233.0000	276.0000
58	178.0000	206.0000	233.0000	278.0000
59	178.0000	207.0000	233.0000	279.0000
60	178.0000	207.0000	233.0000	280.0000
61	178.0000	208.0000	233.0000	283.0000
62	179.0000	208.0000	234.0000	283.0000
63	179.0000	208.0000	235.0000	284.0000
64	180.0000	208.0000	236.0000	285.0000
65	180.0000	208.0000	236.0000	285.0000
66	180.0000	208.0000	237.0000	287.0000
67	181.0000	208.0000	237.0000	297.0000
68	181.0000	209.0000	238.0000	298.0000
69	184.0000	209.0000	239.0000	299.0000
70	184.0000	209.0000	239.0000	304.0000
71	184.0000	210.0000	239.0000	306.0000
72	184.0000	210.0000	239.0000	308.0000
73	184.0000	210.0000	240.0000	313.0000
74	185.0000	211.0000	240.0000	319.0000
75	185.0000	211.0000	240.0000	323.0000
76	185.0000	211.0000	240.0000	331.0000
77	185.0000	211.0000	242.0000	332.0000
78	185.0000	211.0000	242.0000	348.0000
79	185.0000	212.0000	242.0000	386.0000
80	185.0000	212.0000	242.0000	417.0000

SUM 13318.0000 15923.0000 18226.0000 21713.0000

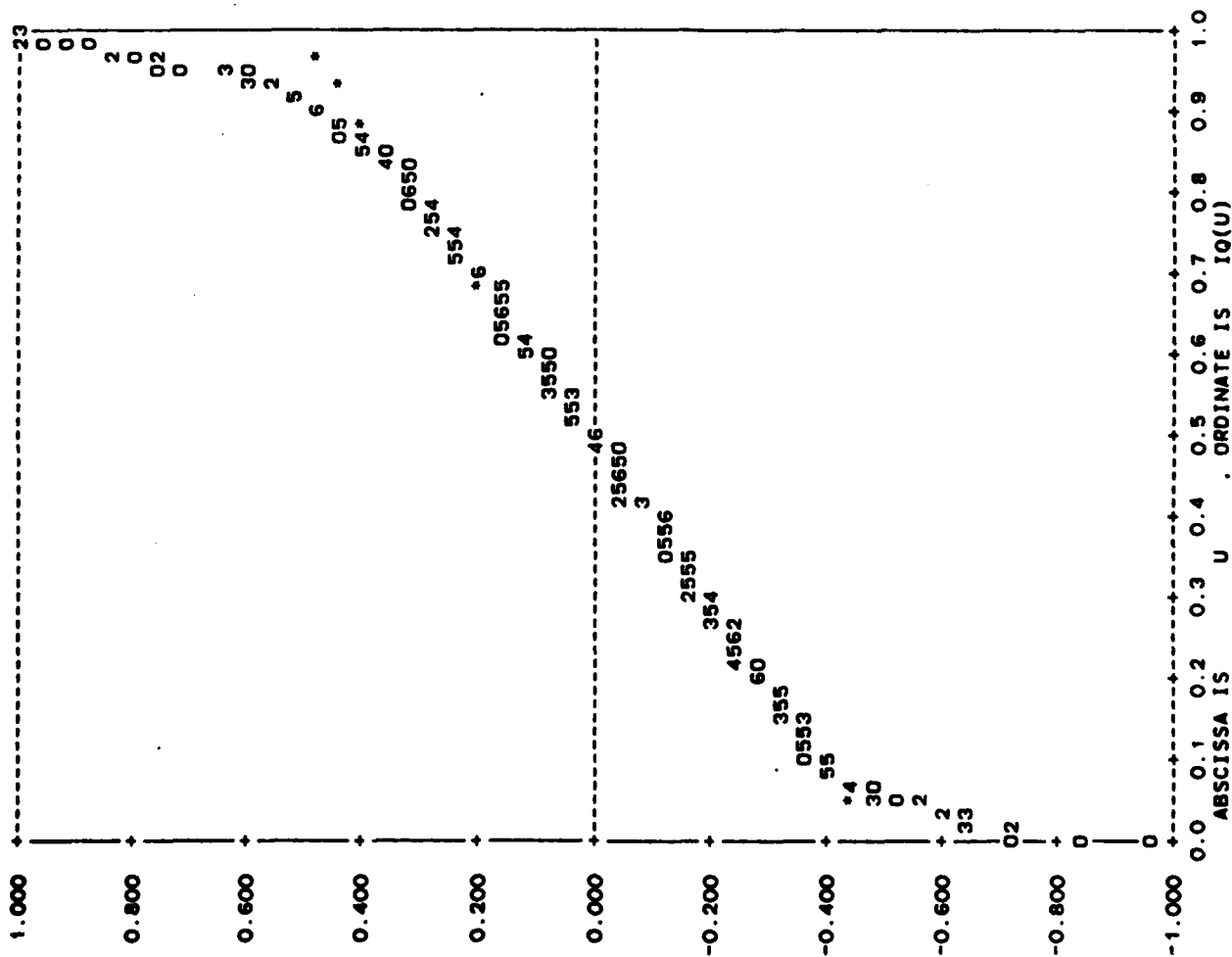
SUM OF SQUARES 2240472.00 3174275.00 4157714.00 5973541.00

DESCRIPTIVE STATISTICS

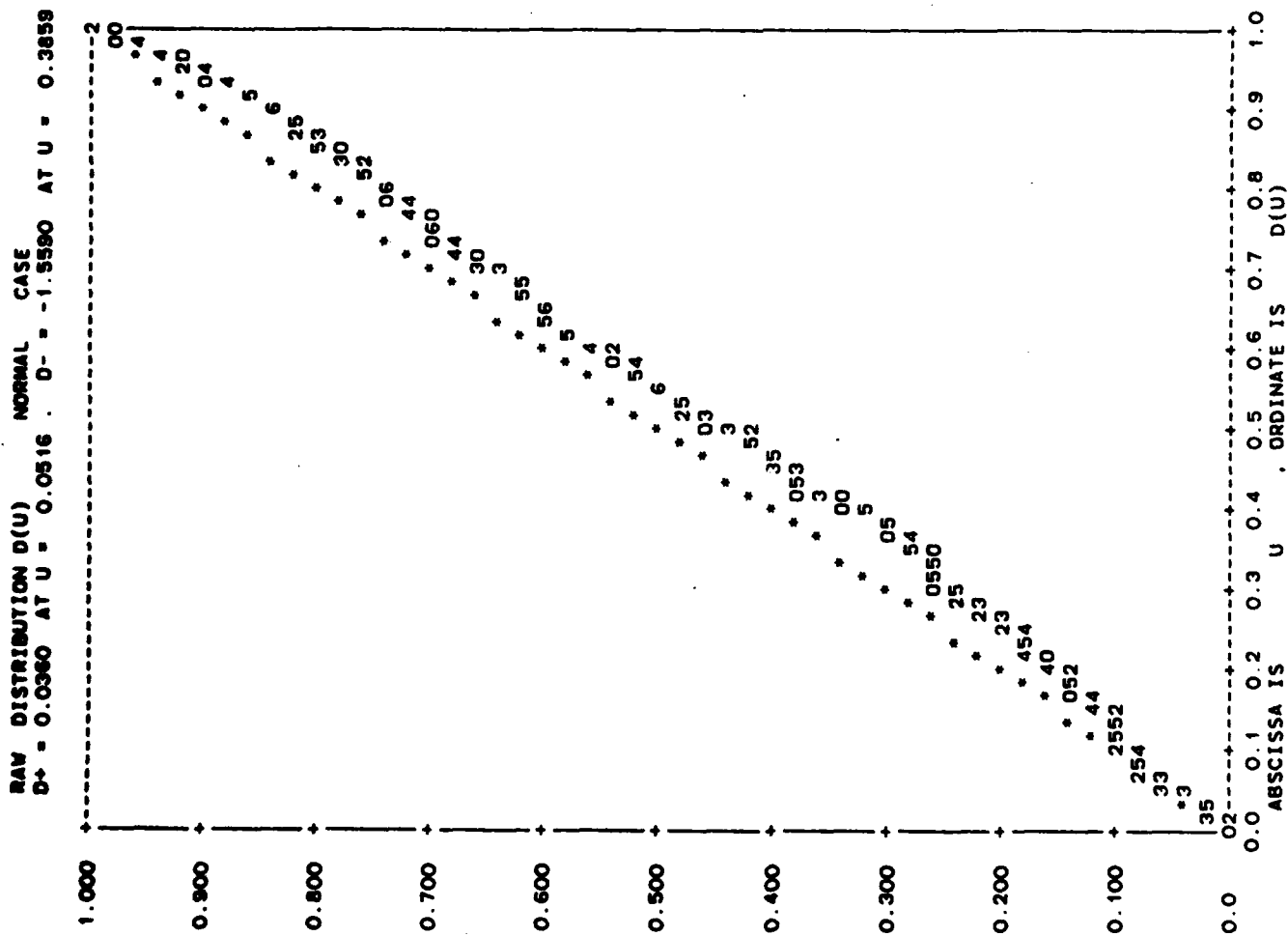
SUMSTAT	SAMPLE SIZE	LOWER QUANTILE	MEDIAN	UPPER QUANTILE	INT QUANTL RANGE	TRIMEAN	GASTWIRTHS ESTIMATE
SUMSTAT	320	185.0	212.5	242.0	57.00	213.0	212.8
SUMSTAT	SUMSQ/N	MEAN	VARIANCE	STD DEV	MEAN IQ	STD DEV IQ	LOG STD IQ
SUMSTAT	.4858E+05	216.2	1850.	43.01	.3235E-01	.3773	-.9747

TRUNCATION POINT	WINSORIZED MEAN	TRIMMED MEAN
0.050	213.9	214.7
0.100	213.5	214.0
0.250	212.1	213.4

PLASMA CHOLESTEROL - DISEASE IN AT LEAST 1 OF 3 CORONARY ARTERIES
INFORMATIVE QUANTILE - ORIGINAL DATA - X



U 0.01000 0.05000 0.10000 0.25000 0.75000 0.90000 0.95000 0.99000
 IQ(U) -0.71481 -0.53684 -0.38947 -0.24123 0.25877 0.47807 0.65260 1.15909



· FULLY NON-PARAMETRIC ANALYSIS

PLASMA TRIGLYCERIDES - DISEASE IN AT LEAST 1 OF 3 CORONARY ARTERIES
ORIGINAL DATA - Y

ORDER STATISTICS IN QUARTERS *****				
SEQUENCE WITHIN QUARTILE *****	FIRST QUARTER *****	SECOND QUARTER *****	THIRD QUARTER *****	FOURTH QUARTER *****
1	36.0000	115.0000	150.0000	220.0000
2	38.0000	116.0000	151.0000	220.0000
3	50.0000	116.0000	151.0000	221.0000
4	54.0000	117.0000	151.0000	222.0000
5	56.0000	117.0000	152.0000	222.0000
6	59.0000	118.0000	152.0000	223.0000
7	61.0000	119.0000	152.0000	227.0000
8	68.0000	120.0000	152.0000	229.0000
9	72.0000	120.0000	153.0000	231.0000
10	73.0000	120.0000	153.0000	232.0000
11	75.0000	120.0000	153.0000	233.0000
12	76.0000	121.0000	153.0000	233.0000
13	77.0000	122.0000	154.0000	237.0000
14	78.0000	123.0000	154.0000	240.0000
15	80.0000	124.0000	154.0000	240.0000
16	80.0000	124.0000	154.0000	242.0000
17	80.0000	124.0000	155.0000	245.0000
18	80.0000	125.0000	156.0000	246.0000
19	82.0000	125.0000	156.0000	248.0000
20	82.0000	125.0000	156.0000	250.0000
21	84.0000	125.0000	156.0000	255.0000
22	84.0000	125.0000	158.0000	256.0000
23	84.0000	126.0000	158.0000	256.0000
24	84.0000	126.0000	158.0000	256.0000
25	84.0000	126.0000	160.0000	257.0000
26	85.0000	126.0000	161.0000	258.0000
27	87.0000	127.0000	161.0000	259.0000
28	87.0000	127.0000	161.0000	259.0000
29	88.0000	128.0000	161.0000	260.0000
30	88.0000	130.0000	162.0000	261.0000
31	89.0000	130.0000	163.0000	262.0000
32	89.0000	130.0000	164.0000	265.0000
33	90.0000	130.0000	164.0000	267.0000
34	90.0000	130.0000	165.0000	268.0000
35	90.0000	131.0000	166.0000	269.0000
36	91.0000	131.0000	166.0000	271.0000
37	91.0000	133.0000	168.0000	272.0000
38	91.0000	133.0000	168.0000	273.0000
39	91.0000	134.0000	169.0000	278.0000
40	91.0000	135.0000	170.0000	284.0000
41	92.0000	135.0000	170.0000	290.0000
42	92.0000	135.0000	170.0000	291.0000

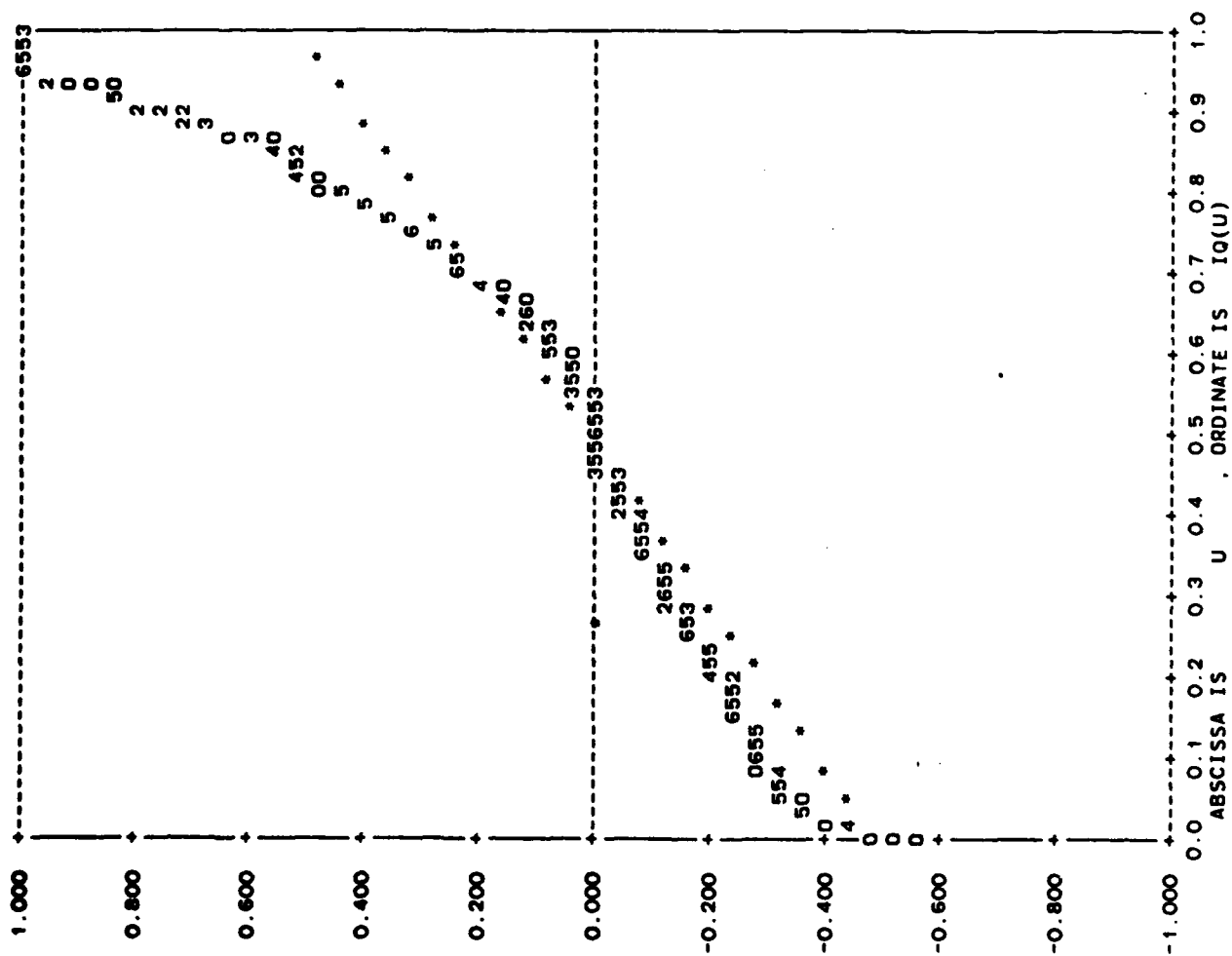
43	92.0000	135.0000	170.0000	296.0000
44	93.0000	136.0000	171.0000	297.0000
45	95.0000	137.0000	172.0000	300.0000
46	96.0000	137.0000	172.0000	304.0000
47	96.0000	137.0000	173.0000	304.0000
48	97.0000	137.0000	174.0000	306.0000
49	98.0000	137.0000	176.0000	312.0000
50	99.0000	140.0000	177.0000	316.0000
51	100.0000	140.0000	179.0000	317.0000
52	100.0000	141.0000	179.0000	322.0000
53	101.0000	141.0000	180.0000	323.0000
54	101.0000	142.0000	181.0000	325.0000
55	101.0000	142.0000	182.0000	327.0000
56	101.0000	142.0000	183.0000	328.0000
57	101.0000	142.0000	184.0000	328.0000
58	101.0000	143.0000	188.0000	333.0000
59	102.0000	144.0000	189.0000	340.0000
60	102.0000	144.0000	192.0000	347.0000
61	103.0000	144.0000	195.0000	348.0000
62	103.0000	145.0000	196.0000	363.0000
63	104.0000	145.0000	196.0000	376.0000
64	105.0000	145.0000	198.0000	390.0000
65	106.0000	146.0000	199.0000	400.0000
66	107.0000	146.0000	199.0000	400.0000
67	108.0000	146.0000	199.0000	402.0000
68	108.0000	146.0000	200.0000	408.0000
69	108.0000	146.0000	201.0000	418.0000
70	109.0000	148.0000	201.0000	424.0000
71	110.0000	148.0000	202.0000	426.0000
72	110.0000	148.0000	202.0000	432.0000
73	111.0000	148.0000	207.0000	441.0000
74	112.0000	148.0000	207.0000	446.0000
75	112.0000	149.0000	208.0000	454.0000
76	112.0000	149.0000	209.0000	489.0000
77	112.0000	149.0000	210.0000	492.0000
78	115.0000	149.0000	217.0000	567.0000
79	115.0000	149.0000	217.0000	583.0000
80	115.0000	150.0000	218.0000	930.0000

SUM 7260.0000 10725.0000 13965.0000 25442.0000

SUM OF SQUARES 682086.0000 1446605.00 2469517.00 9003904.00

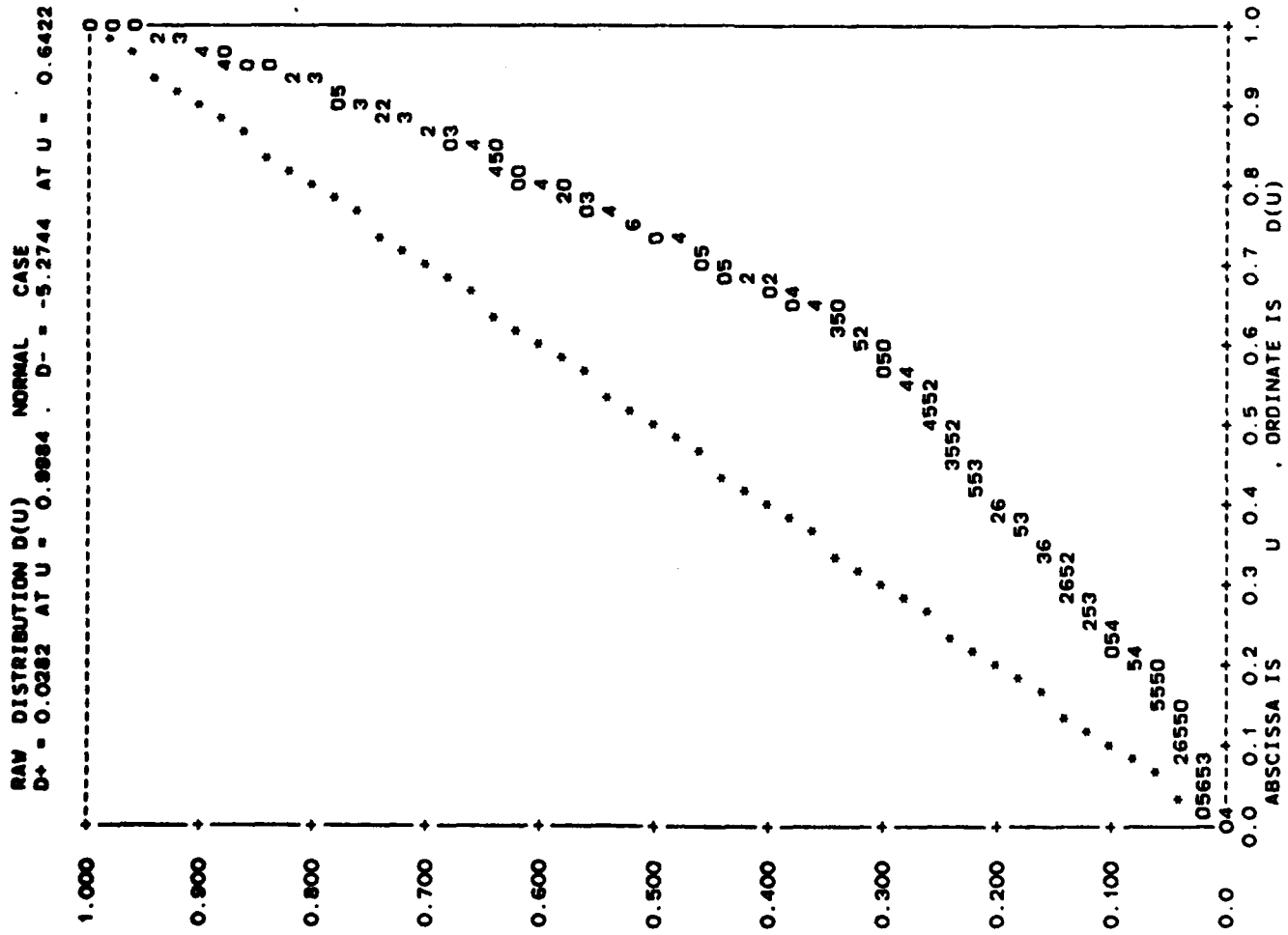
SUMSTAT	SUMSTAT	SAMPLE SIZE	DESCRIPTIVE STATISTICS					GASTWIRTHS ESTIMATE
			LOWER QUANTILE	MEDIAN	UPPER QUANTILE	INT QUANTIL RANGE	TRIMEAN	
SUMSTAT		320	115.0	150.0	219.5	104.5	158.6	152.4
SUMSTAT		SUMSQ/N	MEAN	VARIANCE	STD DEV	MEAN IQ	STD DEV IQ	LOG STD IQ
SUMSTAT	.4251E+05	179.3		.1037E+05	101.8	.1404	.4873	-.7189

TRUNCATION POINT	WINSORIZED MEAN	TRIMMED MEAN
0.050	174.1	168.9
0.100	169.3	163.7
0.250	159.4	154.3



U	0.01000	0.05000	0.10000	0.25000	0.75000	0.90000	0.95000	0.99000
IQ(U)	-0.47445	-0.33493	-0.29139	-0.16746	0.33254	0.77225	1.19378	1.91978

This page is blank.

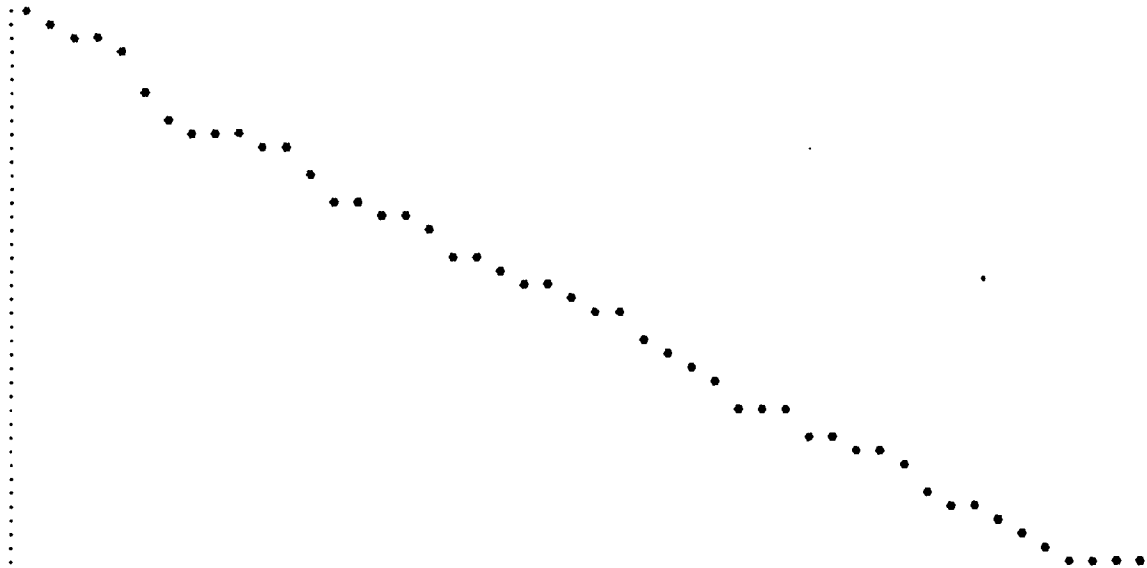


THE FOLLOWING POINTS WERE DELETED FROM THE DATA SET:

320

1 POINTS WERE DELETED LEAVING 319 POINTS IN THE DATA SET.

I	RVAR
1	0.9905
2	0.9799
3	0.9646
4	0.9562
5	0.9457
6	0.9033
7	0.8624
8	0.8471
9	0.8434
10	0.8424
11	0.8382
12	0.8316
13	0.8071
14	0.7780
15	0.7660
16	0.7636
17	0.7621
18	0.7385
19	0.7115
20	0.7092
21	0.7002
22	0.6845
23	0.6702
24	0.6571
25	0.6489
26	0.6402
27	0.6133
28	0.6061
29	0.5885
30	0.5632
31	0.5417
32	0.5353
33	0.5282
34	0.5037
35	0.5025
36	0.4944
37	0.4907
38	0.4792
39	0.4503
40	0.4308
41	0.4229
42	0.4191
43	0.3909
44	0.3737
45	0.3699
46	0.3679
47	0.3655
48	0.3649



UNIVARIATE DENSITY ESTIMATION RESULTS FOR VARIABLE X

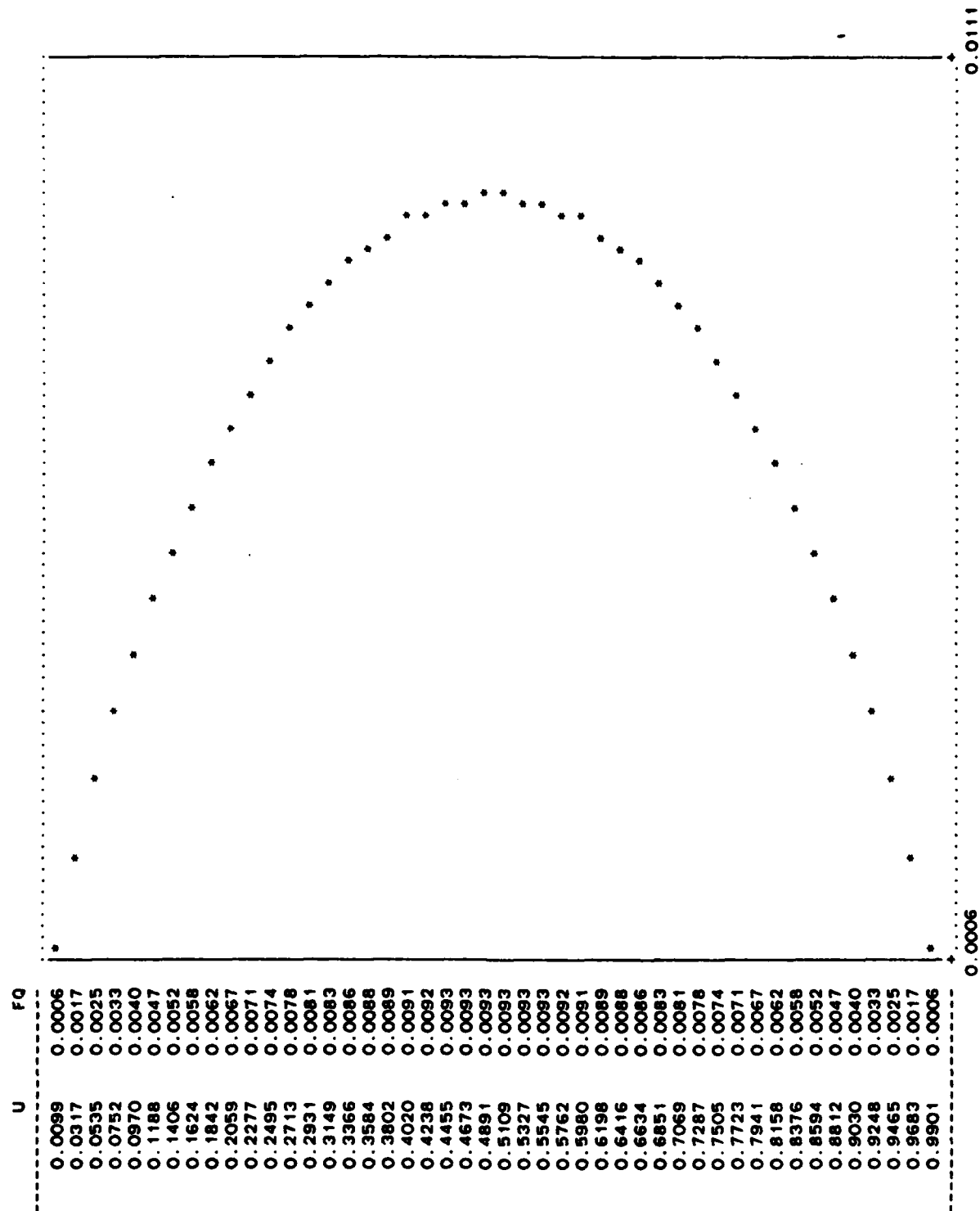
I	RVAR
---	------

1	0.9972
---	--------

2	0.9976
---	--------

SIG0 = 42.6821

UNIVARIATE DENSITY-QUANTILE FOR RANDOM VARIABLE X

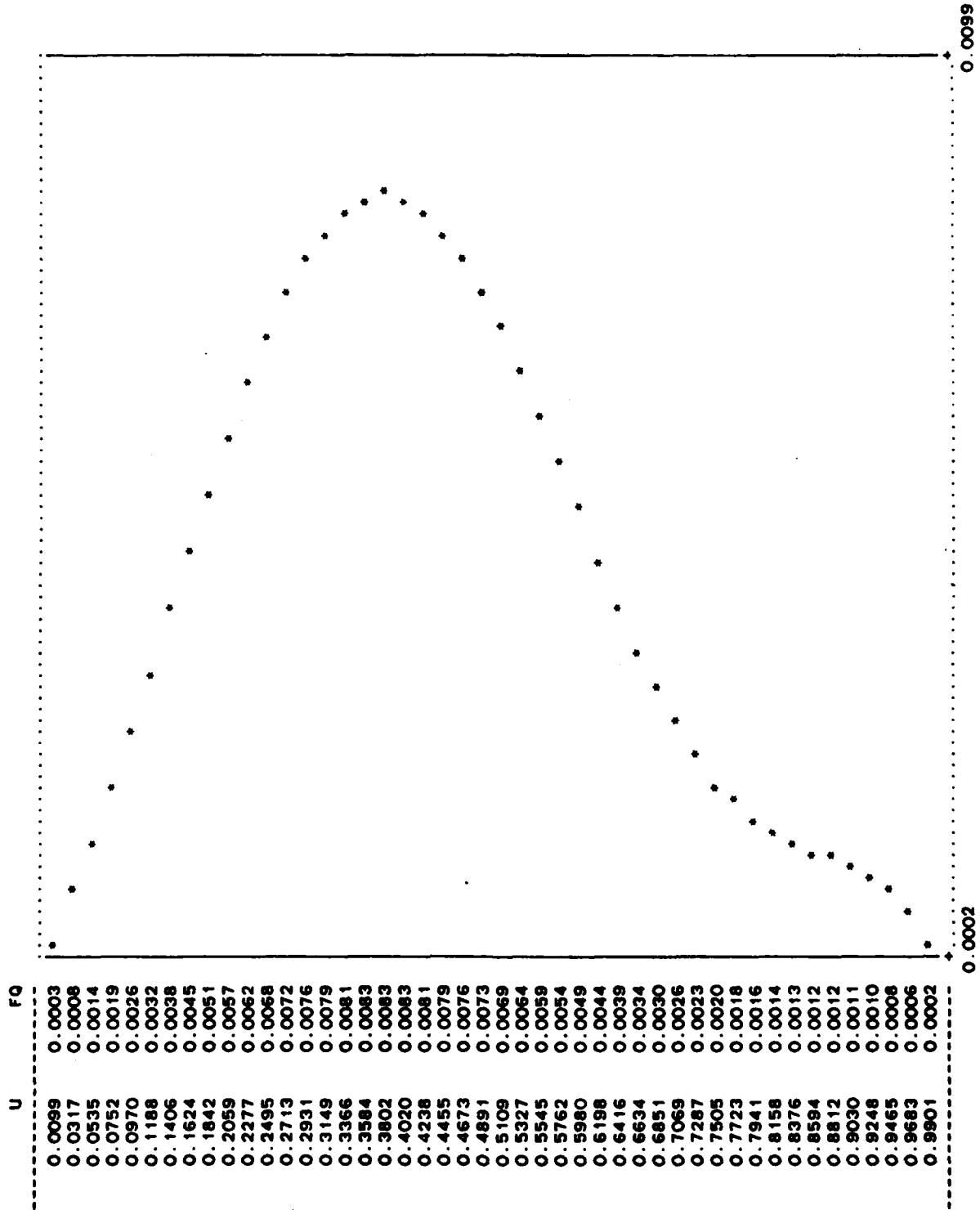


UNIVARIATE DENSITY ESTIMATION RESULTS FOR VARIABLE Y

I	RVAR
1	0.8826
2	0.8788

SIG0 = 91.7530

UNIVARIATE DENSITY-QUANTILE FOR RANDOM VARIABLE Y



UNIVARIATE BEST ORDERS: NVX = 0, NVY = 1

RESULTS FOR ORDER 8 MODEL:

U1	U2	DOHT	DHAT
0.18749994	0.18749994	0.00004772	1.46289539
0.18749994	0.38749993	0.00005591	1.06929111
0.18749994	0.58749992	0.00002575	0.79277152
0.18749994	0.78749990	0.00000923	0.90147328
0.18749994	0.98749989	0.00000250	1.31640625
0.38749993	0.18749994	0.00006505	1.40138245
0.38749993	0.38749993	0.00009387	1.26168442
0.38749993	0.58749992	0.00004522	0.97862923
0.38749993	0.78749990	0.00001353	0.92904085
0.38749993	0.98749989	0.00000313	1.15987206
0.58749992	0.18749994	0.00004023	0.85252440
0.58749992	0.38749993	0.00008944	1.18259430
0.58749992	0.58749992	0.00006110	1.30061340
0.58749992	0.78749990	0.00001472	0.99437904
0.58749992	0.98749989	0.00000210	0.76591164
0.78749990	0.18749994	0.00002302	0.65457690
0.78749990	0.38749993	0.00004399	0.96295321
0.78749990	0.58749992	0.00001111	1.25508970
0.78749990	0.78749990	0.00000138	1.00626469
0.78749990	0.98749989	0.00000358	0.67261958
0.98749989	0.18749994	0.00000569	0.91388923
0.98749989	0.38749993	0.00000361	0.90484190
0.98749989	0.58749992	0.00000117	0.92501044
0.98749989	0.78749990	0.00000021	0.94707549
0.98749989	0.98749989	0.00000001	0.94000223

INTEGRATING FACTOR FOR ORDER 8 IS 0.9728

MAXIMUM VALUE FOR DEPENDENCE DENSITY QUANTILE:

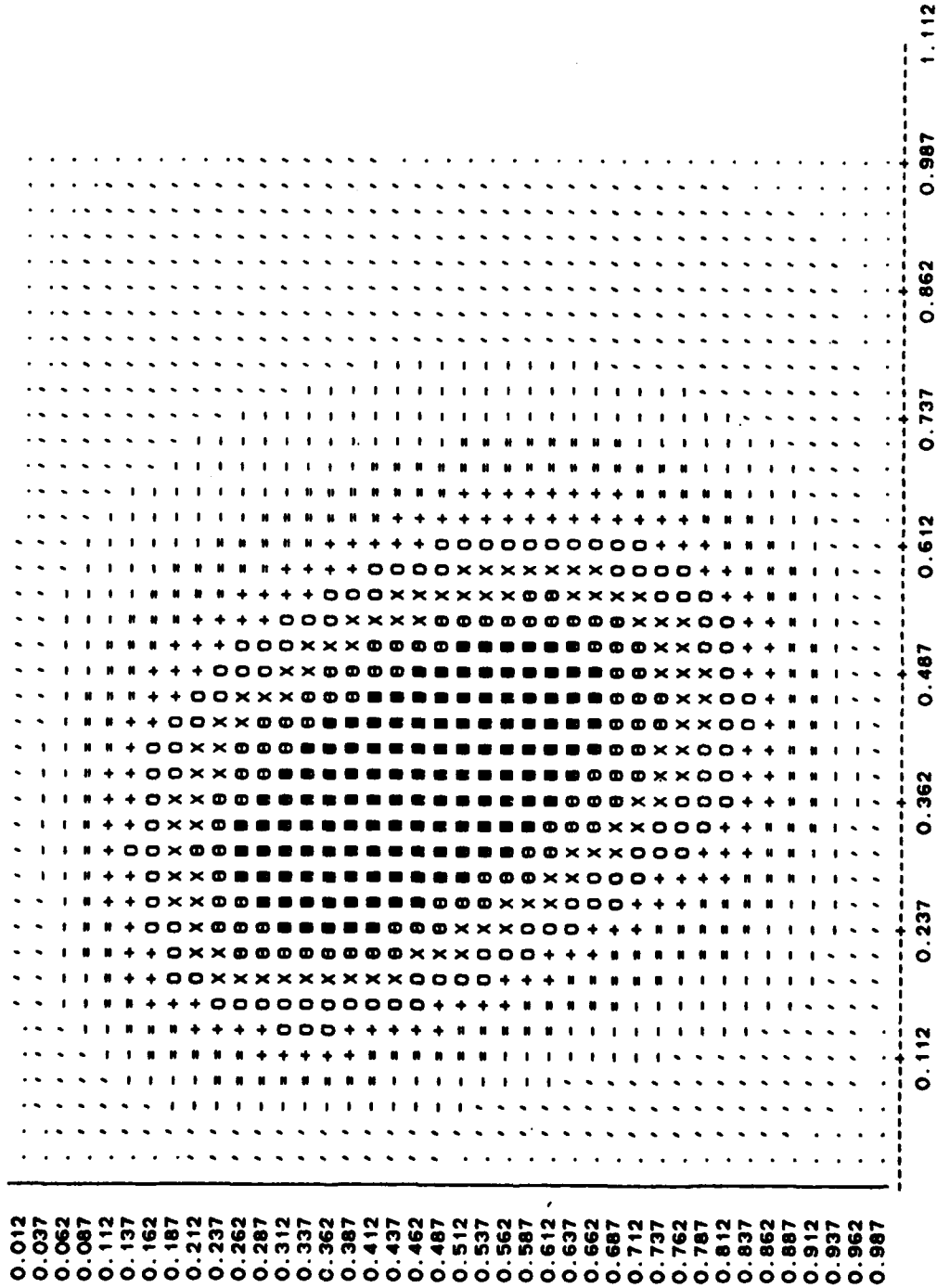
$$D(0.43750, 0.36250) = 0.0001011$$

$$U1^*N = 139.56 \quad U2^*N = 115.64$$

COEFFICIENTS FOR BIVARIATE DEPENDENCE DENSITY

NU1	NU2	REAL(COF)	IMAG(COF)
0	-1	-0.0290	0.0168
-1	0	-0.0341	0.0560
-1	-1	-0.0360	0.0561
0	1	-0.0290	-0.0168
1	0	-0.0341	-0.0560
-1	1	0.0820	0.0736
1	-1	0.0820	-0.0736
1	1	-0.0360	-0.0561

PLASMA CHOLESTEROL - DISEASE IN AT LEAST 1 OF 3 CORONARY ARTERIES
 PLASMA TRIGLYCERIDES - DISEASE IN AT LEAST 1 OF 3 CORONARY ARTERIES
 CONTOUR PLOT FOR BIVARIATE DENSITY QUANTILE - ORDER = 8



ORDINATE IS U1, ABSCISSA IS U2
 U1 CORRESPONDS TO X (FIRST VARIABLE), U2 TO Y

RESULTS FOR ORDER 24 MODEL:

U1	U2	DHAT	DHAT
0.18749994	0.18749994	0.00006591	2.02025032
0.18749994	0.38749993	0.00005909	1.13007641
0.18749994	0.58749992	0.00002053	0.63217348
0.18749994	0.78749990	0.00000841	0.82144260
0.18749994	0.98749988	0.00000127	0.67156988
0.38749993	0.18749994	0.00005310	1.14403343
0.38749993	0.38749993	0.00008828	1.1850246
0.38749993	0.58749992	0.00004822	1.04348946
0.38749993	0.78749990	0.00002334	1.60251522
0.38749993	0.98749989	0.00000298	1.10544777
0.58749992	0.18749994	0.00004044	0.85700732
0.58749992	0.38749993	0.00012231	1.61717510
0.58749992	0.58749992	0.00005357	1.14039707
0.58749992	0.78749990	0.00001090	0.73636955
0.58749992	0.98749989	0.00000271	0.98877615
0.78749990	0.18749994	0.00002065	0.58699453
0.78749990	0.38749993	0.00006526	1.15753841
0.78749990	0.58749992	0.00004041	1.15400219
0.78749990	0.78749990	0.00001366	1.23773098
0.78749990	0.98749989	0.00000131	0.64088982
0.98749989	0.18749994	0.00000330	0.84250641
0.98749989	0.38749993	0.00000408	0.64825362
0.98749989	0.58749992	0.00000514	1.31678867
0.98749989	0.78749990	0.00000102	0.82501429
0.98749989	0.98749989	0.00000027	1.17862511

INTEGRATING FACTOR FOR ORDER 24 IS 1.0001

MAXIMUM VALUE FOR DEPENDENCE DENSITY QUANTILE:

$D(0.61250, 0.41250) = 0.0001278$

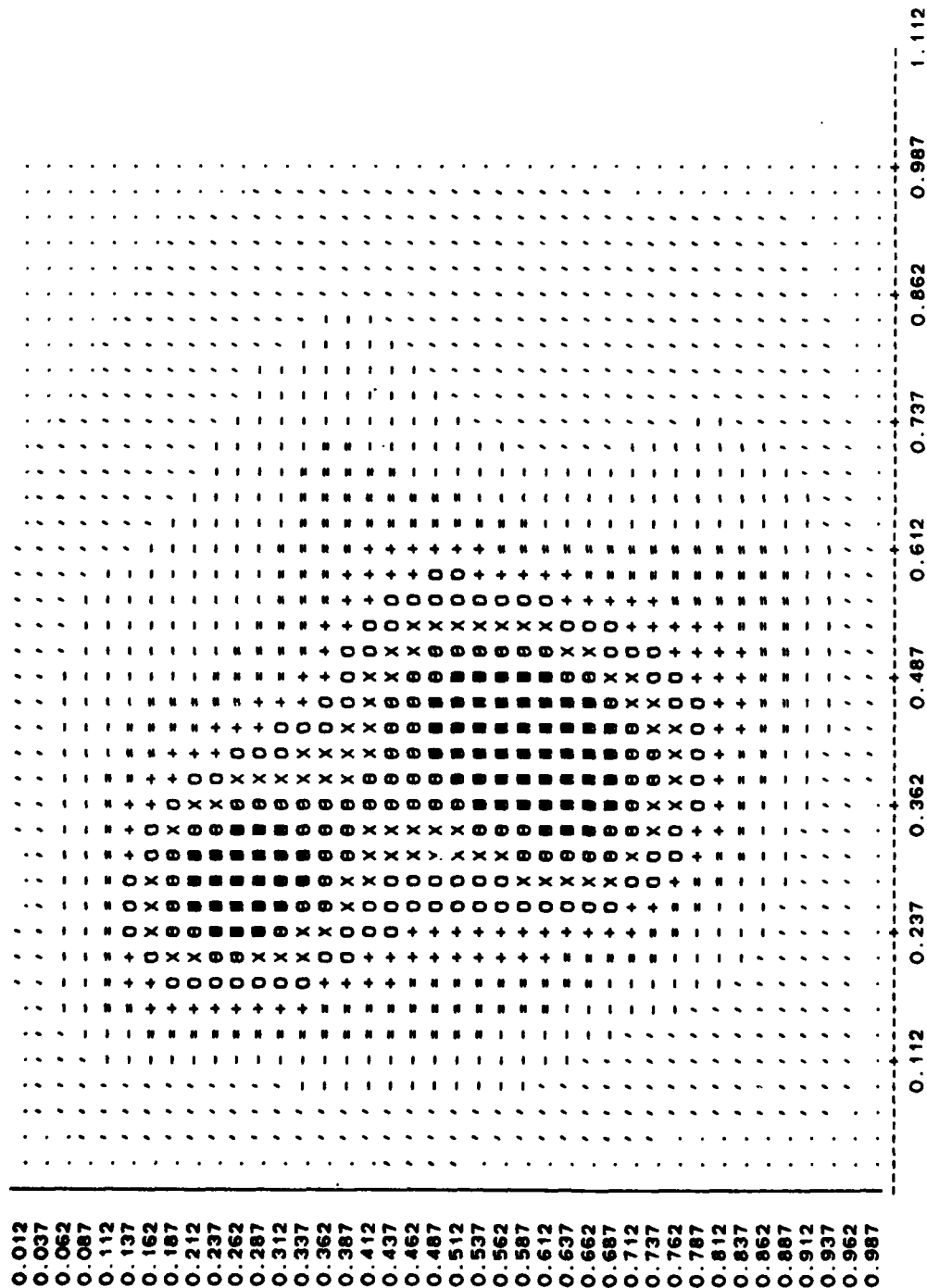
$U1*N = 195.39$ $U2*N = 131.59$

COEFFICIENTS FOR BIVARIATE DEPENDENCE DENSITY

NU1	NU2	REAL(COF)	IMAG(COF)
0	-1	-0.0433	0.0113
-1	0	-0.0550	0.0321
-1	-1	-0.0372	0.0418
0	1	-0.0433	-0.0113
1	0	-0.0550	-0.0321
-1	1	0.0737	0.0619
1	-1	0.0737	-0.0619
1	1	-0.0372	-0.0418
0	-2	-0.0169	-0.0112
-2	0	0.0175	-0.0269
-1	-2	-0.0197	-0.0072
-2	-1	-0.0509	-0.0294
1	-2	0.0178	0.0904

-2	1	0.0935	0.0131
-2	-2	0.0525	0.0351
0	2	-0.0169	0.0112
2	0	0.0175	0.0269
-1	2	0.0178	-0.0904
2	-1	0.0935	-0.0131
1	2	-0.0197	0.0072
2	1	-0.0509	0.0294
-2	2	0.0646	-0.0214
2	-2	0.0646	0.0214
2	2	0.0525	-0.0351

PLASMA CHOLESTEROL - DISEASE IN AT LEAST 1 OF 3 CORONARY ARTERIES
PLASMA TRIGLYCERIDES - DISEASE IN AT LEAST 1 OF 3 CORONARY ARTERIES
CONTOUR PLOT FOR BIVARIATE DENSITY QUANTILE - ORDER = 24



ORDINATE IS U1, ABSCISSA IS U2
U1 CORRESPONDS TO X (FIRST VARIABLE), U2 TO Y

RESULTS FOR ORDER 48 MODEL:

U1	U2	DQHT	DHAT
0.18749994	0.18749994	0.00006638	2.03462219
0.18749994	0.38749993	0.00005525	1.05659676
0.18749994	0.58749992	0.00003006	0.92565721
0.18749994	0.78749990	0.00000834	0.81518406
0.18749994	0.98749989	0.00000080	0.42414057
0.38749993	0.18749994	0.00010054	2.16585255
0.38749993	0.38749993	0.00006806	0.91477740
0.38749993	0.58749992	0.00005214	1.12818909
0.38749993	0.78749990	0.00001949	1.33813477
0.38749993	0.98749989	0.00000247	0.91500378
0.58749992	0.18749994	0.00002933	0.62150997
0.58749992	0.38749993	0.00010528	1.39199924
0.58749992	0.58749992	0.00005599	1.19188213
0.58749992	0.78749990	0.00001759	1.18802643
0.58749992	0.98749989	0.00000212	0.77384871
0.78749990	0.18749994	0.00003154	0.89657325
0.78749990	0.38749993	0.00006950	1.23278046
0.78749990	0.58749992	0.00004922	1.40551186
0.78749990	0.78749990	0.00001147	1.03926563
0.78749990	0.98749989	0.00000134	0.65280330
0.98749989	0.18749994	0.00000217	0.55329788
0.98749989	0.38749993	0.00000563	0.89526796
0.98749989	0.58749992	0.00000369	0.94539171
0.98749989	0.78749990	0.00000111	0.90271294
0.98749989	0.98749989	0.00000029	1.25734138

INTEGRATING FACTOR FOR ORDER 48 IS -1.0443

MAXIMUM VALUE FOR DEPENDENCE DENSITY QUANTILE:

D(0.43750,0.48750) = 0.0001421

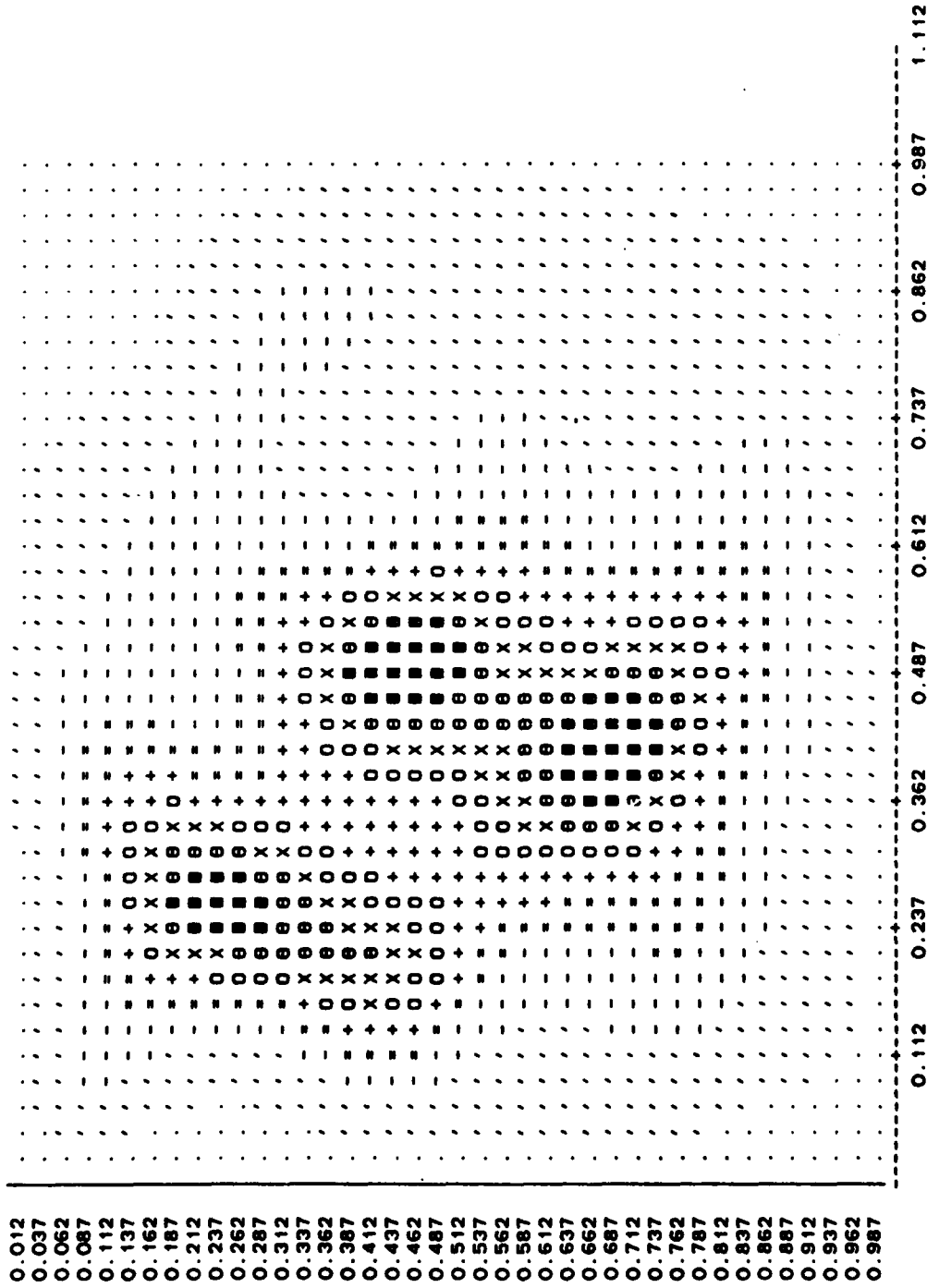
U1*N = 139.56 U2*N = 155.51

COEFFICIENTS FOR BIVARIATE DEPENDENCE DENSITY

NU1	NU2	REAL(COF)	IMAG(COF)
0	-1	-0.0785	0.0079
-1	0	-0.0592	0.0134
-1	-1	-0.0556	0.0353
0	1	-0.0785	-0.0079
1	0	-0.0592	-0.0134
-1	1	0.0687	0.0356
1	-1	0.0687	-0.0356
1	1	-0.0556	-0.0353
0	-2	-0.0196	-0.0085
-2	0	0.0075	-0.0215
-1	-2	-0.0236	-0.0120
-2	-1	-0.0386	-0.0357
1	-2	0.0194	0.1037

-2	1	0.0995	0.0027
-2	-2	0.0487	0.0302
0	2	-0.0196	0.0085
2	0	0.0075	0.0215
-1	2	0.0184	-0.1037
2	-1	0.0995	-0.0027
1	2	-0.0236	0.0120
2	1	-0.0386	0.0357
-2	2	0.0708	-0.0040
2	-2	0.0708	0.0040
2	2	0.0487	-0.0302
0	-3	-0.0462	-0.0077
-3	0	-0.0064	0.0327
-1	-3	0.0733	0.0192
-3	-1	-0.0524	0.0136
1	-3	0.0271	0.0480
-3	1	0.0891	0.0206
-2	-3	-0.0161	0.0730
-3	-2	-0.0033	0.0341
2	-3	0.0262	-0.0330
-3	2	0.0512	0.0774
-3	-3	-0.0111	0.0080
0	3	-0.0462	0.0077
3	0	-0.0064	-0.0327
-1	3	0.0271	-0.0480
3	-1	0.0891	-0.0206
1	3	0.0733	-0.0192
3	1	-0.0524	-0.0136
-2	3	0.0262	0.0330
3	-2	0.0512	-0.0774
2	3	-0.0161	-0.0730
3	2	-0.0033	-0.0341
-3	3	0.0259	-0.0136
3	-3	0.0259	0.0136
3	3	-0.0111	-0.0080

PLASMA CHOLESTEROL - DISEASE IN AT LEAST 1 OF 3 CORONARY ARTERIES
 PLASMA TRIGLYCERIDES - DISEASE IN AT LEAST 1 OF 3 CORONARY ARTERIES
 CONTOUR PLOT FOR BIVARIATE DENSITY QUANTILE - ORDER = 48



ORDINATE IS U1, ABSCISSA IS U2
 U1 CORRESPONDS TO X (FIRST VARIABLE), U2 TO Y

BEST MODEL BY AIC IS ORDER 8 MODEL.

TIES IN X = 388. TIES IN Y = 234

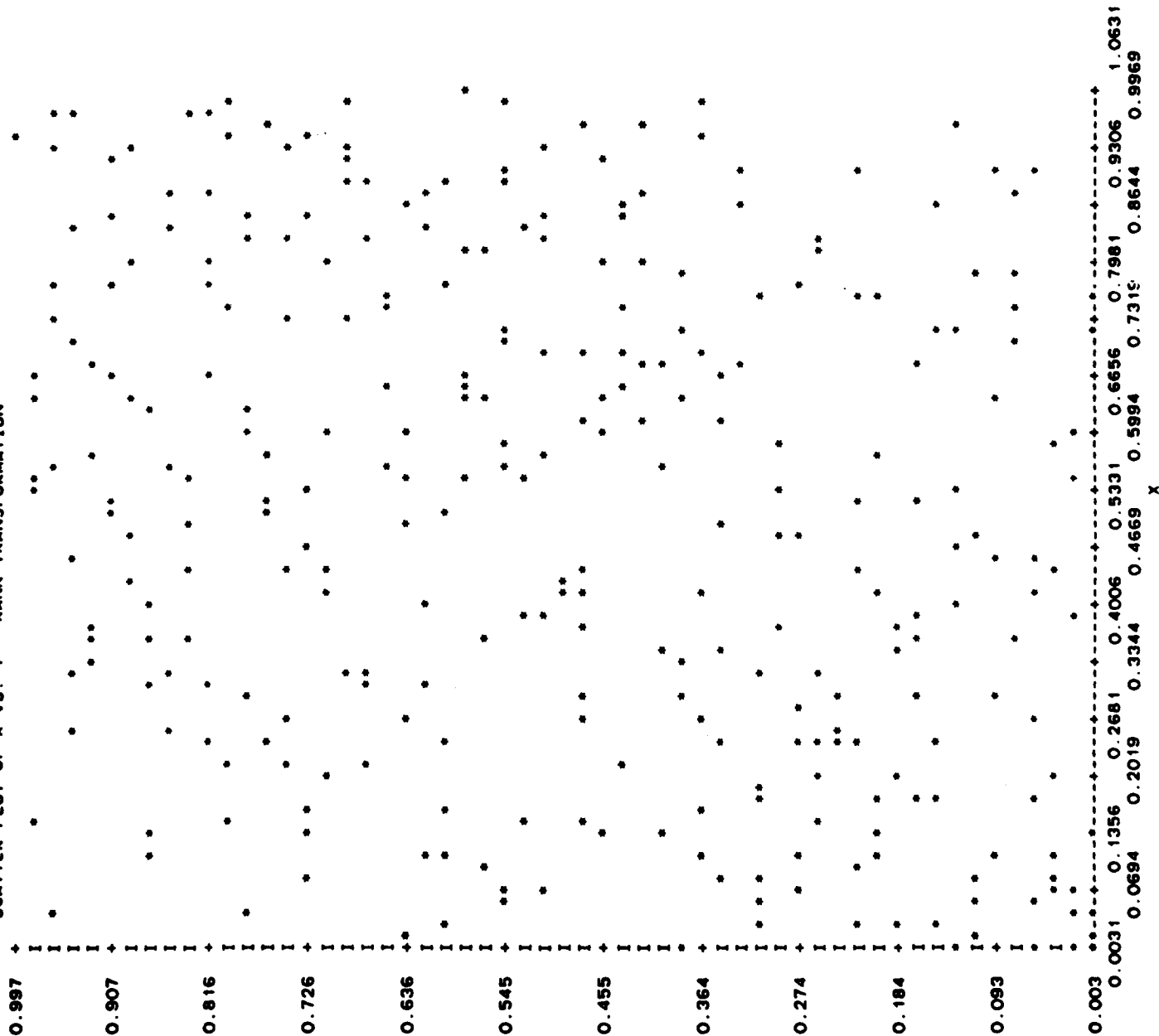
PLASMA CHOLESTEROL - DISEASE IN AT LEAST 1 OF 3 CORONARY ARTERIES
PLASMA TRIGLYCERIDES - DISEASE IN AT LEAST 1 OF 3 CORONARY ARTERIES

SAMPLE SIZE = 319

CORRELATION COEFFICIENT	VALUE	INFORMATION (NORMAL CASE)
PEARSON	0.233332	0.027991
SPEARMAN	0.273826	0.038970
KENDALL A	0.186018	0.017608
KENDALL B	0.187173	0.017831
SOMER'S D	0.187459	0.017887

MODEL	INFORMATION	AIC
D-TILDA	0.03909	0.0
D-HAT 8	0.07118	-0.08224
D-HAT 24	0.09854	-0.20992
D-HAT 48	0.13924	-0.40108

SCATTER PLOT OF X VS. Y - RANK TRANSFORMATION



APPENDIX A - JCL for Executing BISAM

```
// Job Card
//* JES3 Control Cards
//PROCLIB DD DSN=USR.R579.TW.PROCLIB,DISP=SHR
// EXEC BISAM
//SYSIN DD *
```

Parameter Input Card(s)

Data if NTAPE=5

If NTAPE $\neq$ 5, the value of NTAPE is coded as a number nn between 8 and 99 except for 11. (The value 11 is currently used for a scratch file.) A DD card is then inserted defining the input data set. For example,

```
//FT13F001 DD DSN=USR.R579.TW.DATA,DISP=SHR
```

described the dish file containing the input data with NTAPE=13. If IOUDD-1, three such cards will be required as described in section 6.

APPENDIX B - OBTAINING THREE DIMENSIONAL PLOTS

```
1. // JOB CARD
2. //* JES3 CONTROL CARDS
3. //PROCLIB DD DSN=USR.R579.TW.PROCLIB,DISP=SHR
4. // EXEC BISAM
5. //SYSIN DD *
6.
7.
8.
9.
10. *** OPTION CARD FOLLOWED BY DATA IF NTAPE=5 ***
11. //FT01F001 DD DSN=WYL.XX.YYY.STOR1,DISP=OLD
12. //FT02F001 DD DSN=WYL.XX.YYY.STOR2,DISP=OLD
13. //FT03F001 DD DSN=WYL.XX.YYY.STOR3,DISP=OLD
14. //FT13F001 DD DSN=WYL.XX.YYY.DAT1,DISP=OLD
15. /*
16. //STEP2 EXEC SAS,COND=(0,LT)
17. //FILE1 DD DSN=WYL.XX.YYY.STOR1,DISP=OLD
18. //FILE2 DD DSN=WYL.XX.YYY.STOR2,DISP=OLD
19. //FILE3 DD DSN=WYL.XX.YYY.STOR3,DISP=OLD
20. //SYSIN DD *
21. DATA ONE; INFILE FILE1; INPUT U1 U2 DQHT DHAT;
22. TITLE .F=COMPLEX .H=1 FIRST ORDER BIVARIATE DENSITY;
23. PROC G3D DATA=ONE GOUT=A; PLOT U1*U2=DQHT; .
24. PROC G3D DATA=ONE GOUT=B; PLOT U1*U2=DHAT;
25. DATA TWO; INFILE FILE2; INPUT U1 U2 DQHT DHAT;
26. TITLE .F=COMPLEX .H=1 SECOND ORDER BIVARIATE DENSITY;
27. PROC G3D DATA=TWO GOUT=C; PLOT U1*U2=DQHT;
28. PROC G3D DATA=TWO GOUT=D; PLOT U1*U2=DHAT;
29. DATA THREE; INFILE FILE3; INPUT U1 U2 DQHT DHAT;
30. TITLE .F=COMPLEX .H=1 THIRD ORDER BIVARIATE DENSITY;
31. PROC G3D DATA=THREE GOUT=E; PLOT U1*U2=DQHT;
32. PROC G3D DATA=THREE GOUT=F; PLOT U1*U2=DHAT;
33. DATA COMBINE; SET A B C D E F;
34. PROC GREPLAY DATA=COMBINE;
35. /*
```

NOTES: 1. USE CURRENT SAS/GRAPH JCL IN LINE 16.

3. XX.YYY IS USER'S ACCOUNT.

- 5. THE SAS PROCEDURE GCONTOUR MAY BE SUBSTITUTED FOR G3D TO PRODUCE CONTOUR PLOTS INSTEAD OF 3D PLOTS.
- 6. AS INDICATED, PLOTS OF DHAT MAY NOT BE INFORMATIVE, SO DHAT PLOTS INDICATED ABOVE ARE USUALLY OMITTED.
- 7. IN LINES 11-13, THE WYLBUR FILES ARE DUMMY FILES SAVED IN CARD IMAGE THAT WILL BE WRITTEN TO. IN LINE 14, NTAPE=13 HAS BEEN SPECIFIED AND DAT1 CONTAINS THE BIVARIATE DATA SET TO BE ANALYZED.

APPENDIX C - PROGRAM LISTING

```

1. C PROGRAM BISAM
2. C
3. C .....
4. C
5. C DRIVER PROGRAM FOR BIVARIATE DATA ANALYSIS
6. C
7. C INPUT: NTAPE - TAPE WHERE DATA SET RESIDES
8. C (N,V) - BIVARIATE DATA (INDIVIDUALLY, N FIRST)
9. C MORD - MAXIMUM AUTOREGRESSIVE ORDER TO BE USED FOR
10. C UNIVARIATE AR DENSITY ESTIMATION (4-6)
11. C IDOX, IDOY - NULL DISTRIBUTIONS FOR AUTOREGRESSIVE SMOOTHING
12. C IPLT1 - 0 -> NO SCATTER PLOTS
13. C 1 -> SCATTER PLOT OF DATA
14. C 2 -> SCATTER PLOT OF RANK TRANSFORMED DATA
15. C 3 -> BOTH SCATTER PLOTS
16. C IPLT2 - 0 -> NO AUTOREGRESSIVE DENSITY PLOTS
17. C 1 -> BEST ORDER AR DENSITY PLOTS
18. C IDST - 0 -> NO UNIVARIATE DESCRIPTIVE STATISTICS
19. C 1 -> UNIVARIATE DESCRIPTIVE STATISTICS FOR N AND V
20. C KDEL - MAXIMUM NUMBER OF EXTREME POINTS TO EXCLUDE FROM
21. C BIVARIATE ANALYSIS
22. C IOUTD - 1 IF THE 3 MODELS ESTIMATED ARE TO HAVE VALUES
23. C WRITTEN TO TAPES 1, 2, AND 3; 0 O.W.
24. C PTO1001, ETC. DD JCL CARDS MUST BE INCLUDED IF
25. C IOUTD=1.
26. C IREG - 1 IF QUANTILE REGRESSION PERFORMED, 0 O.W.
27. C IUNIV - 1 IF UNIVARIATE DATA SET TO BE READ IN AND A
28. C BIVARIATE DATA SET CREATED AS ORIGINAL DATA
29. C AND LAGGED DATA, 0 O.W.
30. C KLAG1, KLAG2 - RANGE OF LAGS INPUT IF IUNIV=1
31. C
32. C SUBPROGRAMS CALLED: AREST, AUTDEN, AUTORS, CLPLT1, CMPINF, CPLOT,
33. C CPTENT, CSORRG, CSWESP, DATAIN, DESTAT, FCODEA,
34. C FORIER, FOPNC, FTERP, ICODEA, KENDAL, KSD, MAX,
35. C MDNRIS, MIN, MINMAX, ORDS, PARZ, PEARSN, PLOTXY,
36. C PPLST, QFINO, QPLOT, QRES, QTOFO, QUENT, QUICK,
37. C RANK, SPRNG, TRIM, WSPACE
38. C
39. C NOTE: SCRATCH TAPE NUMBER IS SET AT 11 IN FCODEA. A DD CARD
40. C FOR PT11P001 MUST BE INCLUDED IN THE JCL SET UP.
41. C
42. C JULY 1982
43. C PROGRAMMER: TERRY J. WOODFIELD
44. C
45. C .....
46. C
47. C COMMON /DATAR/ X(500),Y(500),RANKX(500),RANKY(500)
48. C COMMON /PARM/ BETAP,BETAW
49. C DIMENSION L(6),LABY(20),LABX(20),Y(500,2),MD(4),AIC(4)
50. C DIMENSION CHAR(6),XNAME(20),YNAME(20),CAPT(20),CRNK(6)
51. C DIMENSION W(1000),LABD(2),LABDX(20),LABDY(20)
52. C EQUIVALENCE (Y(1,1),X(1))
53. C EQUIVALENCE (Y(1,2),Y(1))
54. C DATA NDUT,NIN/6,6/
55. C DATA LAGS/4H N *.4H LAG/
56. C DATA CHAR/1H*,1H*,1H*,1H*,1H*,1H*/
57. C DATA XNAME/10*1H,1H*,9*1H /
58. C DATA YNAME/10*1H,1H*,9*1H /
59. C DATA LABD/4HORIG,4HINAL,4H DAT,4HA -.4HX .15*4H /
60. C DATA LABD/4HORIG,4HINAL,4H DAT,4HA -.4HY .15*4H /
61. C DATA CAPT/4HSCAT,4HYER,4HPLST,4H OF,4H VS,4H, Y,14*4H /
62. C DATA CRNK/4H-BA,4HKK Y,4HTRANS,4HFORM,4HATIO,4HN /
63. C
64. C LACT=0
65. C WRITE(NDUT,1)
66. C 1 FORMAT(1H1)
67. C READ(NIN,10) NTAPE, IDOX, IDOY, MORD, IPLT1, IPLT2, IDST, KDEL, IOUTD,
68. C IREG, IUNIV
69. C 10 FORMAT(12I5)
70. C IF(IUNIV.EQ.1) READ(NIN,10) KLAG1,KLAG2
71. C IF(IDOX.EQ.9) READ(NIN,12) BETAP
72. C IF(IDOY.EQ.10) READ(NIN,12) BETAPX,BETAWX
73. C IF(IDOY.EQ.9) READ(NIN,12) BETAPY
74. C IF(IDOY.EQ.10) READ(NIN,12) BETAPY,BETAWY
75. C 12 FORMAT(2F10,0)
76. C WRITE(NDUT,20)
77. C 20 FORMAT(//,10X,20(4H****),/,10X,*,.78X,*,/,10X,*, BISAM *,
78. C *- BIVARIATE DATA ANALYSIS USING FOURIER EXPANSIONS*,10X,*,
79. C *,.78X,*, AND QUANTILE TECHNIQUES*,44X,*,/,10X,*,
80. C *,.78X,*, PROGRAMMER: TERRY J. WOODFIELD*,750,*,
81. C *,.78X,*,/,10X,20(4H****))
82. C IF(IUNIV.EQ.0) GO TO 28
83. C CALL DATAIN(NTAPE,W,NX,L,LABX)
84. C DO 200 KLAG=KLAG1,KLAG2
85. C N=NX-KLAG
86. C DO 22 1=1,N
87. C X(1)=W(1)
88. C Y(1)=W(1-KLAG)
89. C 22 CONTINUE
90. C DO 24 1=1,20
91. C LABY(1)=LABX(1)
92. C CALL ICODEA(KLAG,6H(14) ,LABY(17))
93. C LABY(15)=LABD(1)
94. C LABY(16)=LABD(2)
95. C GO TO 40
96. C 28 CONTINUE
97. C CALL DATAIN(NTAPE,X,NX,L,LABX)
98. C CALL DATAIN(NTAPE,Y,NV,L,LABY)
99. C N=NV
100. C IF(NX.EQ.NV) GO TO 40
101. C WRITE(NDUT,30) LABX,LABY
102. C 30 FORMAT(1H,10X,20A4/,10X,20A4/,/,10X,SAMPLE SIZES NOT EQUAL.,
103. C *- BIVARIATE ANALYSIS INAPPROPRIATE. EXECUTION TERMINATED.,)
104. C STOP
105. C 40 IF(LAST.EE.1) GO TO 51
106. C WRITE(NDUT,50) LABX,LABY,N,NTAPE, IDOX, IDOY, MORD, IPLT1, IPLT2,
107. C IDST, KDEL, IOUTD, IREG, IUNIV
108. C 50 FORMAT(//,10X,20A4/,10X,20A4/,/,10X,SAMPLE SIZE *,.15,
109. C *,.78X,*,OPTIONS FOR THIS ANALYSIS:*,750,NTAPE *,.12,
110. C *,.78X,*, IDOX *,.12,760, IDOY *,.12,720, MORD *,.12,
111. C *,.78X,*, IPLT1 *,.12,760, IPLT2 *,.12,720, IDST *,.12,
112. C *,.78X,*, KDEL *,.12,760, IOUTD *,.12,720, IREG *,.12,740,
113. C *,.78X,*, IUNIV *,.12)
114. C IF(IPLT1.EQ.1) OR (IPLT1.EQ.3)
115. C CALL PPLST(X,Y,500,N,1,CHAR,CAPT,XNAME,YNAME,0)
116. C 51 WRITE(NDUT,1)
117. C
118. C ORDER BIVARIATE DATA BY X VALUES
119. C
120. C CALL ORDS(T,N,500)
121. C IF(IDST.EQ.0) OR (LACT.EE.1) GO TO 52
122. C BETAP=BETAPX
123. C BETAW=BETAWX

```



```

300 C
301 C.....
302 C
303 DIMENSION NIN(101)
304 DATA NIN,NOUT/5,6/
305 DATA BLANK,DDY,2/1N,1N,1N=7/
306 C
307 IOPYX=0
308 IF(N.GT.1) GO TO 10
309 WRITE(NOUT,11) NAME,N(1)
310 11 FORMAT(10X,A4,'(1)'*,F10.0)
311 GO TO 90
312 10 CONTINUE
313 C
314 INITIALIZE AL:
315 C
316 ON=(MM-1)/2
317 DO 20 J=1,MM
318 AL(J)=DDY
319 WRITE(NOUT,25) NAME,{AL(J),J=1,MM}
320 25 FORMAT(/,14X,1N1,6X,A4/10X,15(1N-),2X,101A1)
321 DO 30 J=1,MM
322 AL(J)=BLANK
323 C
324 FIND MAX AND MIN:
325 C
326 CALL MAX(X,N,NMAX,IND)
327 CALL MIN(X,N,NMIN,IND)
328 NMAX=NMIN
329 IF(NX.LT.1,8-20) IOPYX=1
330 C
331 PLOT:
332 C
333 JJ=INIT
334 DO 40 J=1,N
335 IF(IOPYX.EQ.1) GO TO 36
336 C1=(X(J)-NMIN)/RX
337 C1=2*(C1-.5)
338 GO TO 37
339 C1=0
340 37 K=ON+(C1+.5)*1.5
341 AL(K)=2
342 IF(IOPYX.EQ.1) GO TO 38
343 DO 39 I=1,K
344 AL(I)=2
345 39 CONTINUE
346 WRITE(NOUT,38) JJ,X(J),{AL(I),I=1,MM}
347 38 FORMAT(10X,15,F10.4,2X,101A1)
348 JJ=JJ+1
349 AL(N)=BLANK
350 IF(IOPYX.EQ.1) GO TO 40
351 DO 41 I=1,K
352 AL(I)=BLANK
353 41 CONTINUE
354 C
355 99 CONTINUE
356 RETURN
357 END
358 SUBROUTINE CMPINF(N,MORD,IDX,IDY,IPLY2,ISUD,IRES,LBX,LBY,
359 *ND,AIC)
360 C.....
361 C
362 SUBROUTINE TO COMPUTE COVARIANCE MATRIX OF COMPLEX
363 EXPONENTIAL "SUFFICIENT STATISTICS" TO BE USED IN
364 SEQUENTIAL REGRESSION ROUTINE TO OBTAIN "REGRESSION"
365 MODELS FOR ORDERS 1 THROUGH M-M. SUBROUTINE CPTENT IS USED
366 TO OBTAIN COEFFICIENTS FOR THREE MAXIMUM ENTROPY
367 ESTIMATES OF THE BIVARIATE DEPENDENCE DENSITY. THEN THE
368 BIVARIATE DENSITY QUANTILE IS FORMED BY TAKING THE PRODUCT
369 OF THE ESTIMATED DEPENDENCE DENSITY AND THE UNIVARIATE
370 AUTOREGRESSIVE ESTIMATORS.
371 C
372 INPUT: RANKX,RANKY - VECTORS CONTAINING RANKS OF X AND CORANKS
373 OF Y
374 X,Y - BIVARIATE DATA
375 N - SAMPLE SIZE
376 MORD - MAXIMUM AUTOREGRESSIVE ORDER TO BE USED FOR
377 UNIVARIATE AR DENSITY ESTIMATION (<6)
378 IDOX,IDOY - NULL DISTRIBUTIONS FOR AUTOREGRESSIVE SMOOTHING
379 IPLY2 = 0 -> NO AUTOREGRESSIVE DENSITY PLOTS
380 1 -> BEST ORDER AR DENSITY PLOTS
381 IDST = 0 -> NO UNIVARIATE DESCRIPTIVE STATISTICS
382 1 -> UNIVARIATE DESCRIPTIVE STATISTICS FOR X AND Y
383 ISUD = 1 IF THE 3 MODELS ESTIMATED ARE TO HAVE VALUES
384 WRITTEN TO TAPES 1,2, AND 3; 0 O.W.
385 PTOODI, ETC. DD JCL CARDS MUST BE INCLUDED IF
386 ISUD=1.
387 IRES = 1 IF QUANTILE REGRESSION PERFORMED, 0 O.W.
388 LBX,LBY - LABELS FOR X AND Y
389 C
390 OUTPUT: PHI - COVARIANCE MATRIX
391 FOX,FOY - UNIVARIATE DENSITY QUANTILE FUNCTIONS
392 DONAT - BIVARIATE DENSITY QUANTILE FUNCTION
393 ND - VECTOR OF ENTROPY ESTIMATORS: 1 - ORDER 8 MODEL
394 2 - ORDER 24 MODEL
395 3 - ORDER 48 MODEL
396 4 - RAW (FROM 0-TILOAI)
397 AIC - VALUES OF ENTROPY CRITERION FUNCTION BASED ON
398 AKAIKE'S INFORMATION CRITERION
399 C
400 NOTE: FOX,FOY ARE NOT PASSED BACK TO THE CALLING PROGRAM.
401 ALSO, CRITERION FUNCTIONS ARE PLOTTED BUT NOT PASSED
402 BACK TO THE CALLING PROGRAM.
403 C
404 SUBPROGRAMS CALLED: CSOREG,CPTENT,PLOTXY,PTERP,AUTDEN,MIN
405 C.....
406 C
407 COMMON /DATAR/ X(500),Y(500),RANKX(500),RANKY(500)
408 COMMON /PARAM/ SETAP,SETAW
409 DIMENSION IND(97),RADSO(500),IORD(49),
410 *ND(4),AIC(4),OTIL(500),AMAT(40,40)
411 DIMENSION MENT(3),PSI(2),LABR(20),LABY(20)
412 COMPLEX ARSM(12,12),PHI(50,50),CERP,CONJE,CMPLE,ZARS
413 COMPLEX ALPHX(5),ALPHY(5),COF(97)
414 DATA IORD/25,24,18,17,26,32,19,31,33,23,11,16,10,30,12,9,
415 *27,39,20,38,34,40,13,37,41,32,8,19,3,29,0,8,2,36,0,1,28,
416 *46,21,45,35,47,14,44,42,48,7,43,49/
417 DATA MENT/0.24,46/
418 REAL LABDAT
419 IF(N.GT.20) GO TO 20
420 WRITE(6,10) N
421 10 FORMAT(10X,'SAMPLE SIZE ',12,' IS TOO SMALL. CMPINF SKIPPED.')
422 RETURN
423 C
424 SET VALUES OF CONSTANTS
425 C
426 20 N2=N-2
427 DO 21 I=1,4
428 21 ND(I)=999.0
429 C

```

```

030 C FOR THIS VERSION USING COMPLEX SEQUENTIAL REGRESSION THE
031 C MAXIMUM APPROXIMATIVE ORDER IS SET AT 7.
032 C
033 M=7
034 L=MOD(M,2)
035 ML=(M-L)/2
036 IF(L.EQ.0) M=M+1
037 M2=2*M-1
038 MM=M*M
039 M1=MM+1
040 MM1=MM-1
041 DENOM=1.0/FLOAT(M+1)
042 TWOP=1.0*ATAN(1.0)
043 PI=TWOP/2.0
044 C
045 C COMPUTE NEAREST NEIGHBOR DENSITY ESTIMATE AND RAW ESTIMATE
046 C OF ENTROPY
047 C
048 H00=0.0
049 DO 30 I=1,M2
050 DO 25 J=1,M
051 RADSO(J)=RANKX(I)-RANKX(J)**2+(RANKY(I)-RANKY(J))**2
052 CONTINUE
053 DO 25 K=1,M
054 CALL MIN(RADSO,M,RMIN,INDR)
055 RADSO(INDR)=FLOAT(2*M*M)
056 CONTINUE
057 VKJ=RMIN*DENOM-DENOM*PI
058 IF(VKJ.EQ.0.0) VKJ=0.5*DENOM-DENOM*PI
059 C
060 C DTIL IS ALSO(DTIL)
061 C
062 DTIL(I)=ALOG(5.0/((FLOAT(M+1)*VKJ)))
063 H00=H00-DTIL(I)
064 CONTINUE
065 DO 40 I=1,M2
066 H0(I)=H00/((FLOAT(M-4)))
067 C
068 C COMPUTE MATRIX OF EXPONENTIAL CROSS-PRODUCTS TO BE USED FOR
069 C COVARIANCE COMPUTATIONS
070 C
071 DO 60 I=1,M2
072 I1=I-M
073 DO 50 J=1,M2
074 J1=J-M
075 ARG=1.0*CMPLX(0.0,0.0)
076 DO 40 K=1,M2
077 ARG=ARG*(FLOAT(I1)-RANKX(K)-FLOAT(J1)-RANKY(K))-DENOM
078 ZARG=CMPLX(0.0,ARG)
079 ARG=ARG*(ZARG)
080 CONTINUE
081 ARG(I,J)=ARG/(FLOAT(M-4))
082 CONTINUE
083 C
084 C COMPUTE COVARIANCE MATRIX
085 C
086 DO 60 IN=1,MM
087 I1=IN-1
088 I2=MOD(I1,M)
089 I1=(I1-I2)/M*M-ML
090 I2=I2-M*ML
091 DO 55 JN=1,IN
092 J1=JN-1
093 J2=MOD(J1,M)
094 J1=(J1-J2)/M*M-ML
095 J2=J2-M*ML
096 J1=J1-M*ML
097 J2=J2-M*ML
098 PHI(I1,JN)=ARG(I1,J1)-ARG(I1,I2)*CONJG(ARG(J1,J2))
099 PHI(JN,I1)=CONJG(PHI(I1,JN))
100 CONTINUE
101 CONTINUE
102 C
103 C COMPUTE LAST ROW OF COVARIANCE MATRIX
104 C
105 DBAR=0.0
106 DO 70 I=1,M2
107 DBAR=DBAR+DTIL(I)
108 CONTINUE
109 DBAR=DBAR/((FLOAT(M-4)))
110 DO 60 IN=1,MM
111 I1=IN-1
112 I2=MOD(I1,M)
113 I1=(I1-I2)/M*M-ML
114 I2=I2-M*ML
115 PHI(I1,IN)=CMPLX(0.0,0.0)
116 DO 50 K=1,M2
117 ARG=ARG*(FLOAT(I1)-RANKX(K)-FLOAT(I2)-RANKY(K))-DENOM
118 ZARG=CMPLX(0.0,ARG)
119 PHI(I1,IN)=PHI(I1,IN)+DTIL(K)*CONJG(CEXP(ZARG))
120 CONTINUE
121 PHI(I1,IN)=PHI(I1,IN)/((FLOAT(M-4))-DBAR*CONJG(ARG(I1,M,I2*M)))
122 CONTINUE
123 PHI(I1,IN)=CONJG(PHI(I1,IN))
124 DO 100 K=1,M2
125 PHI(MI,MI)=PHI(MI,MI)+DTIL(K)*DTIL(K)
126 CONTINUE
127 PHI(MI,MI)=PHI(MI,MI)/((FLOAT(M-4))-DBAR*DBAR)
128 C
129 C CALL ROUTINE CPYENT TO COMPUTE AND PLOT RESIDUAL VARIANCE AND
130 C DETERMINE THREE MODELS FOR D(U1,U2)
131 C
132 CALL CPYENT(N,M,PHI,1000,'L,COF,MENT)
133 C
134 C COMPUTE UNIVARIATE DENSITY ESTIMATES USING AUTOREGRESSIVE
135 C TECHNIQUE
136 C
137 WRITE(6,134)
138 FORMAT(1H1)
139 CALL AUTDEN(X,N,100X,1PLT2,MORD,ALPHX,RVAX,SIGX,NVX,0.4H X)
140 WRITE(6,134)
141 CALL AUTDEN(Y,N,100Y,1PLT2,MORD,ALPHY,RVAY,SIGY,NVY,1.4H Y)
142 WRITE(6,135) NVX,NVY
143 FORMAT(//,10X,'UNIVARIATE BEST ORDERS: NVX =',13,' , NVY =',13)
144 DO 200 ITER=1,3
145 PSI(ITER)=0.0
146 ENT=0.0
147 WRITE(6,136) MENT(ITER)
148 FORMAT(1H1,10X,'RESULTS FOR ORDER ',12,' MODEL:',
149 //,10X,'U1',17X,'U2',15X,'DONT',15X,'DNAT',17X,
150 //,10X,'')
151 BOMAX=0.
152 LOCIT=1
153 IF(ITER.EQ.2) LOCIT=MENT(1)+1
154 IF(ITER.EQ.3) LOCIT=MENT(1)+MENT(2)+1
155 DO 220 I=1,40
156 W1=(FLOAT(I)-.5)*0.025
157 DO 230 J=1,40
158 U2=(FLOAT(J)-0.5)*0.025
159 C
160 C COMPUTE VALUES OF UNIVARIATE DENSITY-QUANTILE FUNCTIONS

```



```

784 CALL MINMAX(A,IA,M,N,XMIN,XMAX)
785 FACT=(XMAX-XMIN)/10.
786 DO 60 I=1,M
787   IFLAG=0
788   DO 10 J=1,N
789     LINE1(J)=BLANK
790     LINE2(J)=BLANK
791     LINE3(J)=BLANK
792   10 LINE3(J)=BLANK
793   DO 70 J=1,N
794     ITEMP=IPIN([A(I,J)-XMIN]/FACT)*.5+.1
795     INEXT=([ITEMP+.2])/2
796     IF([ITEMP.EQ.1]EQ.1) GO TO 11
797     IF([ITEMP.EQ.1]EQ.1) GO TO 15
798     IF([ITEMP.EQ.1]EQ.1) GO TO 17
799     IF([ITEMP.EQ.1]EQ.1) GO TO 15
800     LINE1(J)=ISYMB(INEXT)
801     GO TO 70
802   11 LINE1(J)=ISYMB(1)
803   GO TO 70
804   15 IFLAG=1
805   IFLAG=1
806   LINE1(J)=ISYMB(6)
807   LINE2(J)=ISYMB(7)
808   LINE3(J)=ISYMB(8)
809   GO TO 70
810   17 IFLAG=1
811   LINE1(J)=ISYMB(6)
812   LINE2(J)=ISYMB(7)
813   GO TO 70
814   15 IFLAG=1
815   LINE1(J)=ISYMB(6)
816   LINE2(J)=ISYMB(3)
817   70 CONTINUE
818   U=(FLDAT(I)-0.5)/DIVM
819   WRITE(6,901)U,LINE1
820   IF(IFLAG.EQ.1)WRITE(6,902)LINE2
821   IF(IFLAG.EQ.1)WRITE(6,903)LINE3
822   80 CONTINUE
823   INDEX=(M+1)/5
824   DO 81 I=1,INDEX
825     XNUM(I)=(5.*FLDAT(I)-0.5)/DIVM
826     NEND=INDEX+10
827     DO 82 J=1,NEND
828       LINE1(J)=ISYMB(3)
829       DO 83 I=1,10,NEND,10
830         LINE1(J)=ISYMB(5)
831       NEND=NEND+1
832       IF(NEND.EQ.110)GO TO 85
833       DO 84 I=NEND,110
834         LINE1(I)=BLANK
835       85 WRITE(6,904)XNUM(I),I+1,INDEX
836       RETURN
837   901 FORMAT(1H .2X,PS.3,2X,"[.1X,50(A1,1X)]")
838   902 FORMAT(1H .11X,50(A1,1X))
839   903 FORMAT(1H .0X,110A1)
840   904 FORMAT(1H .12X,11(5X,PS.3))
841   END
842   SUBROUTINE CPENTIN(M,PHI,IORD,IND,CDF,MENT)
843   C .....
844   C
845   C SUBPROGRAM TO COMPUTE MAXIMUM ENTROPY ESTIMATES OF
846   C THE BIVARIATE DEPENDENT DENSITY.
847   C
848   C THREE FITTED MODELS WILL BE RETURNED WITH COEFFICIENTS
849   C IN CDF, VARIABLE NAMES (INDICES) IN IND, THE FIRST LOCATION
850   C OF COEFFICIENTS AND INDICES FOR THE 2ND MODEL IN MENT(1)+1, I.E.,
851   C COEFFICIENT(1)+1) CONTAINS COEFFICIENT NUMBER ONE OF THE SECOND
852   C REGRESSION MODEL CORRESPONDING TO INDEX IND(MENT(1)+1), ETC.
853   C
854   C INPUT: M,M - SAMPLE SIZE, UNIVARIATE MAXIMUM ORDER (M+2
855   C USED FOR BIVARIATE MAX ORDER)
856   C PHI - COVARIANCE MATRIX
857   C IORD - VECTOR OF ORDERED INDICES FOR SEQUENTIAL REGRESSION
858   C
859   C AUXILIARY: NVAR,NVAR,BEST - VECTORS AND MATRIX
860   C FROM ROUTINE CSOREG
861   C
862   C OUTPUT: CDF,IND - SEE ABOVE
863   C
864   C SUBROUTINES CALLED: CSOREG,CSWEEP,CLPLY1
865   C .....
866   C
867   C
868   C DIMENSION MENT(3),IND(97)
869   C DIMENSION IORD(49),NVAR(49),INDV(1225),NVAR(49)
870   C COMPLEX PHI(50,50),CDF(97),BEST(1225)
871   C MM=M+M
872   C MM1=MM-1
873   C L=MOD(M,2)
874   C ML=(M-L)/2
875   C TWOPI=6.283185307179586
876   C
877   C CALL ROUTINE CSOREG TO PERFORM SEQUENTIAL REGRESSION ON PHI
878   C
879   C CALL CSOREG(PHI,50,MM,IORD,BEST,INDV,1225,NVAR,NVAR,NIVN)
880   C CALL CLPLY1(NVAR,NIVN,1,4,NVAR,41,1)
881   C
882   C PLACE COEFFICIENTS IN CDF FOR EACH ORDER
883   C
884   C
885   C
886   C LOC=1
887   C DO 100 I=1,3
888   C   K=MENT(I)
889   C   IF(K.EQ.0) GO TO 100
890   C   KI=NVAR(K)
891   C   DO 170 KK=1,A
892   C     IND(LOC)=INDV(KI)
893   C     CDF(LOC)=BEST(KI)
894   C     KI=KI+1
895   C   LOC=LOC+1
896   C   170 CONTINUE
897   C   100 CONTINUE
898   C   RETURN
899   C   END
900   C SUBROUTINE CSOREG(A,NDIM,NIV,IORD,BEST,INDV,MDIM,NVAR,NVAR,NIVN)
901   C .....
902   C
903   C SUBPROGRAM TO PERFORM SEQUENTIAL REGRESSION USING COVARIANCE
904   C OR CORRELATION MATRIX A(NIV+1,NIV+1).
905   C
906   C INPUT: A - COVARIANCE MATRIX (COMPLEX)
907   C NDIM - ROW DIMENSION OF A IN CALLING PROGRAM
908   C NIV - NUMBER OF INDEPENDENT VARIABLES
909   C IORD - INTEGER VECTOR CONTAINING INDICES OF VARIABLES
910   C IN THE ORDER THEY ARE TO BE ENTERED INTO THE MODEL
911   C MDIM - DIMENSION OF BEST IN CALLING PROGRAM
912   C
913   C OUTPUT: A - SWEEP COVARIANCE MATRIX
914   C BEST,INDV - VECTORS OF SUBSET INFORMATION
915   C BEST CONTAINS LEAST SQUARES PARAMETER ESTIMATES

```



```

1040 C SUBROUTINE TO PRINT ORDERED ARRAY BY QUANTILES AND COMPUTE
1041 DESCRIPTIVE STATISTICS.
1042 C INPUT:
1043 C   N: ARRAY OF ORDER STATISTICS
1044 C   N: DIMENSION OF ARRAY N
1045 C   NAME: NAME OF DATA SET MUST BE ARRAY OF DIMENSION 20 IN
1046 C   CALLING PROGRAM.
1047 C   IHEAD: HEADING FOR ANALYSIS
1048 C   IUNIT: NUMBER OF UNIT OUTPUT IS DESIRED ON.
1049 C   INIT: 0 FOR FIRST CALL, 1 THEREAFTER
1050 C   IOUT: 1 IF QUANTILES TO BE LISTED, 0 OTHERWISE
1051 C   OUTPUT: PRINTED OUTPUT IS ON IUNIT.
1052 C NO SUBROUTINES CALLED.
1053 C-----
1054 DIMENSION X(N),NAME(20),SUM(4),SUMSQ(4)
1055 DIMENSION ALP(3),L(4)
1056 DIMENSION IHEAD(20)
1057 DATA ALP/.05,.10,.25/
1058 NUNIT = 5
1059 C COMPUTE L, THE ARRAY OF QUANTILE SIZES
1060 IF (INIT .EQ. 0) GOTO 5
1061 IF (L(4) .EQ. N) GOTO 25
1062 L1 = N/4
1063 L2 = L1
1064 L3 = L1
1065 L4 = L1
1066 IIRAN = MOD(N,4) + 1
1067 GOTO (20,11,12,12),IIRAN
1068 11 CONTINUE
1069 L4 = L1 + 1
1070 GO TO 30
1071 12 CONTINUE
1072 L1 = L1 + 1
1073 L4 = L1 + 1
1074 GO TO 20
1075 13 CONTINUE
1076 L2 = L1 + 1
1077 L3 = L1 + 1
1078 L4 = L1 + 1
1079 20 CONTINUE
1080 L(1)=L1
1081 L(2)=L(1) + L2
1082 L(3)=L(2) + L3
1083 L(4)=L(3) + L4
1084 C PRINT DATA ARRAY - ONE COLUMN FOR EACH QUARTER.
1085 25 WRITE(NUNIT,1001) NAME
1086 WRITE(NUNIT,1020) IHEAD
1087 IF (IOUT .EQ. 0) GOTO 35
1088 WRITE(NUNIT,1002)
1089 WRITE(NUNIT,1003)
1090 DO 30 I = 1, L1
1091 WRITE(NUNIT,1004) I, X(I), X(L(1) + I), X(L(2) + I), X(L(3) + I)
1092 WRITE(NUNIT,1005)
1093 IF (L1 .GT. L2) WRITE(NUNIT,1006) X(L(1))
1094 IF (L2 .GT. L3) WRITE(NUNIT,1007) X(L(2))
1095 IF (L3 .GT. L4) WRITE(NUNIT,1008) X(L(3))
1096 IF (L4 .GT. L1) WRITE(NUNIT,1009) L4, X(L(4))
1097 35 IF (INIT .EQ. 1) RETURN
1098 C COMPUTE AND PRINT DESCRIPTIVE STATISTICS.
1099 K = 1
1100 S = 0.
1101 SSO = 0.
1102 DO 40 J = 1, 4
1103 S1 = 0.0
1104 S01 = 0.0
1105 KK = L(J)
1106 DO 40 J = K, KK
1107 S1 = S1 + X(J)
1108 S01 = S01 + X(J)*X(J)
1109 40 CONTINUE
1110 K = 1 + L(1)
1111 S = S + S1
1112 SSO = SSO + S01
1113 SUM(1) = S1
1114 SUMSQ(1) = S01
1115 50 CONTINUE
1116 IF (IOUT .EQ. 0) GO TO 55
1117 WRITE(NUNIT,1010) (SUM (I), I=1,4)
1118 WRITE(NUNIT,1011) (SUMSQ (I), I=1,4)
1119 XBAR = S/FL0AT(N)
1120 VAR = (SSO - S*XBAR)/FL0AT(N-1)
1121 SD = SORT(VAR)
1122 Q25 = OFIND(X,N,.25)
1123 Q50 = OFIND(X,N,.50)
1124 Q75 = OFIND(X,N,.75)
1125 Q10 = Q75 - Q25
1126 XBAR10 = (XBAR - Q50) / (2. + Q10)
1127 SD10 = SD / (2. + Q10)
1128 SD10L = ABS(SD10)
1129 SSON=SSO/FL0AT(N)
1130 TRIMM = (Q25+2.*Q50+Q75)*.25
1131 GASTY = (.2*OFIND(X,N,.25)+.4*Q50+.3*OFIND(X,N,.75))
1132 WRITE(NUNIT,1012)
1133 WRITE(NUNIT,1013) N,Q25,Q50,Q75,Q10,TRIMM,GASTY
1134 WRITE(NUNIT,1014) SSON,XBAR,VAR,SD,XBAR10,SD10L,SD10L
1135 WRITE(NUNIT,1015)
1136 DO 60 I = 1, 3
1137 IC = INT(ALP(I)*FL0AT(N))
1138 M = N - 2*IC
1139 NMSG1 = N - IC - 1
1140 ISP2 = IC + 2
1141 YRN = X(IC+1)+X(N-IC)
1142 WNN = X(IC+1)*FL0AT(IC) + X(N-IC)*FL0AT(IC)
1143 DO 70 J = ISP2, NMSG1
1144 YRN = YRN + X(J)
1145 WNN = WNN + X(J)
1146 70 CONTINUE
1147 YRN = YRN/N
1148 WNN = WNN/FL0AT(N)
1149 WRITE(NUNIT,1017) ALP(I),WNN,YRN
1150 CONTINUE
1151 999 CONTINUE
1152 RETURN
1153 1001 FORMAT(///T20,20A1)
1154 1002 FORMAT(T40,'ORDER STATISTICS IN QUARTERS'/T40,20(1H*))
1155 1003 FORMAT(T20,' SEQUENCE'/T20,' WITHIN'/T20,' QUANTILE'
1156 + 'FIRST QUARTER SECOND QUARTER THIRD QUARTER FOURTH QUARTER'
1157 + '/21,8(1H*),2(2X,12(1H*),2X,14(1H*))')
1158 1004 FORMAT(T20,16 , 4(1X,F10.4))
1159 1005 FORMAT(1H,T20,F10.4)
1160 1006 FORMAT(1H,T40,F10.4)
1161 1007 FORMAT(1H,T61,F10.4)
1162 1008 FORMAT(1H,T20,10,T77,F10.4)
1163 1009 FORMAT(1H,T20,'SUM',5X,4(1X,F10.4))
1164 1010 FORMAT(1H,T20,'SUM OF'/T20,' SQUARES',1X,4(1X,F10.4))
1165 1011 FORMAT(///T40,'DESCRIPTIVE STATISTICS'/T40,23(1H*))
1166 1012 FORMAT(///T20,'SUMSTAT',T20,'SAMPLE',T20,' LOWER',T20,' UPPER'
1167 + 'T60,'INT QUANTL',T60,'GASTYWTHS',T20,'SUMSTAT',T20,' SIZE'
1168 + 'T20,' QUANTILE',T40,' MEDIAN',T20,' QUANTILE'
1169 + 'T60,' RANGE',T70,' TRIMEAN',T60,' ESTIMATE'

```

```

1100      *      SUMSTAT',T20,15,T20,6(511.4,1X))
1101 1014 FORMAT(///,SUMSTAT',T20,'SUNDO/N',,MEAN',T40,' VARIANCE',T52,
1102      *      STD DEV',T60,' MEAN',T20,' STD DEV',T60,'
1103      *      LOC STD',T20,' SUMSTAT',T10,T10(511.4,1X))
1104 1015 FORMAT(///T40,' TRUNCATION',T52,' UNBORDERED',T64,' TRIMMED',
1105      *      T40,' POINT',T52,' MEAN',T64,' MEAN',T64)
1106 1017 FORMAT(T40,P7.3,T52,2(511.4,1X))
1107 1020 FORMAT(T20,20A6)
1108      END
1109      SUBROUTINE PCDBEA(N,IFORM,NAME)
1110      C-----
1111      C SUBROUTINE TO CONVERT REAL VARIABLE X
1112      C WHICH HAS 4 CHARACTER V-FORMAT IFORM
1113      C TO 8 CHARACTER ALPHAMERIC ARRAY NAME WHICH IS
1114      C IN 6 FORMAT
1115      C INPUT : NSCREEN : SCRATCH TAPE NUMBER
1116      C          N : IFORM
1117      C OUTPUT : NAME(1),NAME(2) : 4 CHARACTERS EACH
1118      C-----
1119      DIMENSION NAME(2),IFORM(2)
1120      NSCREEN = 11
1121      REWIND NSCREEN
1122      WRITE(NSCREEN,IFORM)X
1123      REWIND NSCREEN
1124      READ(NSCREEN,10)NAME
1125      10 FORMAT(2A4)
1126      RETURN
1127      END
1128      SUBROUTINE FORIER(F,U,N,A,MA)
1129      C-----
1130      C SUBROUTINE TO COMPUTE THE FOURIER TRANSFORM
1131      C PHI(V) OF A DENSITY DEFINED ON (0,1) FOR V=0,1,...,M
1132      C INPUT :
1133      C          F,U,N : VECTORS OF LENGTH N CONTAINING F(U),U
1134      C          MA : MAXIMUM VALUE OF V FOR WHICH PHI(V) IS COMPUTED
1135      C OUTPUT : A : COMPLEX-VALUED VECTOR CONTAINING THE PHI'S
1136      C SUBROUTINE CALLED : NONE
1137      C-----
1138      DIMENSION F(N),U(N)
1139      COMPLEX A(MA),Z
1140      TWOPI=6.2831853
1141      FM=FLOAT(M)
1142      DO 30 IM=1,MA
1143      FIM=IM-1
1144      A(IM)=CMPLX(0.,0.)
1145      DO 10 J=1,N
1146      Z=CMPLX(F(J),TWOPI*(FIM-U(J)))
1147      A(IM)=A(IM)+F(J)*Z*EXP(Z)
1148      10 CONTINUE
1149      A(IM)=A(IM)/FLOAT(N)
1150      20 CONTINUE
1151      A(1)=CMPLX(1.,0.)
1152      RETURN
1153      END
1154      FUNCTION POPNC(X,100H)
1155      C-----
1156      C ROUTINE TO COMPUTE THE VARIOUS DENSITY-QUANTILE FUNCTIONS
1157      C INPUT :
1158      C          X : VALUE AT WHICH THE FUNCTION IS TO BE COMPUTED
1159      C          100H : INDICATOR FOR THE DESIRED FUNCTION.
1160      C          MUST BE IN THE EXCLUSIVE RANGE 1-11.
1161      C-----
1162      COMMON /PARM/ BETAP,BETAW
1163      GO TO(1,2,3,4,5,6,7,8,9,10,11),100H
1164      C COMPUTE THE NORMAL
1165      1 CONTINUE
1166      IF(X .LT. .001) GO TO 101
1167      IF(X .GT. .999) GO TO 101
1168      PI=4.*ATAN(1.)
1169      CALL MORRIS(X,T,18R)
1170      P000=EXP(-0.5*T**2)/(SORT(2.*PI))
1171      POPNC = P000
1172      GO TO 99
1173      101 P000=0.
1174      POPNC = P000
1175      GO TO 99
1176      C COMPUTE THE EXPONENTIAL
1177      2 CONTINUE
1178      P000=1.-X
1179      POPNC = P000
1180      GO TO 99
1181      C COMPUTE THE LOGISTIC
1182      3 CONTINUE
1183      POPNC = X/(1. + X)
1184      GO TO 99
1185      C COMPUTE THE DOUBLE EXPONENTIAL
1186      4 CONTINUE
1187      IF(X .LE. .5) POPNC = X
1188      IF(X .GT. .5) POPNC = 1. - X
1189      GO TO 99
1190      C COMPUTE THE UNIFORM RECIPROCAL
1191      5 CONTINUE
1192      POPNC = (1. - X)**2
1193      GO TO 99
1194      C COMPUTE THE CAUCHY
1195      6 CONTINUE
1196      PI = 4.*ATAN(1.)
1197      POPNC = SIN(PI*X)/(1.-2*PI
1198      GO TO 99
1199      C COMPUTE THE EXTREME VALUE
1200      7 CONTINUE
1201      IF(X .GE. 1.) GO TO 102
1202      POPNC = (X - 1.)*ALOG(1. - X)
1203      GO TO 99
1204      102 CONTINUE
1205      POPNC = 0.
1206      GO TO 99
1207      C COMPUTE THE LOG NORMAL
1208      8 CONTINUE
1209      IF(X .GE. 1.) GO TO 103
1210      IF(X .GE. 0.) GO TO 103
1211      CALL MORRIS(X,T,18R)
1212      PI = 4.*ATAN(1.)
1213      POPNC = (-0.5*T**2-T)/(SORT(2.*PI))
1214      POPNC = EXP(POPNC)
1215      GO TO 99
1216      103 POPNC = 0.0
1217      GO TO 99
1218      C COMPUTE THE PARETO
1219      9 CONTINUE
1220      POPNC = (1.-X)**(1.+BETAP)/BETAP
1221      GO TO 99
1222      C COMPUTE THE WEIBULL
1223      10 CONTINUE
1224      IF(X .GE. 1.) GO TO 104
1225      POPNC = (1.-X)/(1.-ALOG(1.-X))**BETAW/BETAW
1226      GO TO 99
1227      104 CONTINUE
1228      POPNC = 0.
1229      GO TO 99
1230      C COMPUTE THE HALF LOGISTIC
1231      11 CONTINUE
1232      POPNC=1.-X**2

```

```

1312      95 RETURN
1313      END
1314      SUBROUTINE FTERP(U,V,X,F,M)
1315      C-----
1316      C
1317      C SUBROUTINE TO PERFORM LINEAR INTERPOLATION ON V TO
1318      C OBTAIN F AT THE M X VALUES
1319      C
1320      C INPUT: U - VECTOR OF VALUES AT WHICH V EVALUATED
1321      C          V - FUNCTION VALUES TO INTERPOLATE
1322      C          X - VALUES AT WHICH INTERPOLATED FUNCTION TO BE
1323      C             EVALUATED
1324      C          M - DIMENSION OF VECTORS U AND V
1325      C          N - DIMENSION OF VECTORS X AND F
1326      C
1327      C NOTE: ALL ABSCISSA VECTORS MUST BE ORDERED
1328      C
1329      C OUTPUT: F - INTERPOLATED FUNCTION VALUES
1330      C
1331      C-----
1332      DIMENSION U(N),V(N),X(M),F(M)
1333      IF(N.GE.M) GO TO 100
1334      I=1
1335      DO 30 I=1,M
1336      10 IF(X(I)-U(1))20,40,50
1337      20 IF(X(I).NE.1) GO TO 30
1338      F(I)=V(1)+(V(2)-V(1))*(X(I)-U(1))/(U(2)-U(1))
1339      GO TO 60
1340      30 F(I)=V(1)+(V(2)-V(1))*(X(I)-U(1))/(U(1)-U(2))
1341      GO TO 60
1342      40 F(I)=V(1)
1343      GO TO 60
1344      50 I=1+1
1345      IF(I.LY.N) GO TO 10
1346      I=N
1347      GO TO 30
1348      60 CONTINUE
1349      100 RETURN
1350      END
1351      SUBROUTINE ICDEA(K,IFORM,NAME)
1352      C-----
1353      C SUBROUTINE TO CONVERT INTEGER VARIABLE K
1354      C WHICH HAS 8 CHARACTER I-FORMAT IFORM
1355      C TO 8 CHARACTER ALPHANERIC ARRAY NAME WHICH IS
1356      C IN A-FORMAT.
1357      C INPUT : NSCRCH : SCRATCH TAPE NUMBER
1358      C          K : IFORM
1359      C OUTPUT : NAME(1),NAME(2) : 4 CHARACTERS EACH
1360      C-----
1361      DIMENSION NAME(2),IFORM(2)
1362      NSCRCH=11
1363      REWIND NSCRCH
1364      WRITE(NSCRCH,IFORM)K
1365      REWIND NSCRCH
1366      READ(NSCRCH,10)NAME
1367      10 FORMAT(2A4)
1368      RETURN
1369      END
1370      SUBROUTINE KENDAL(N,TAUA,TAUB,SOMER,NC,ND,NIND,NDEP,NPAIRS)
1371      C
1372      C-----
1373      C
1374      C SUBROUTINE TO COMPUTE KENDALL'S TAU-A AND TAU-B (FOR TIED
1375      C RANKS), AND SOMER'S D.
1376      C
1377      C INPUT: RANKX - THE VECTOR OF RANKS OF THE INDEPENDENT VARIABLE
1378      C          RANKY - THE VECTOR OF RANKS OF THE DEPENDENT VARIABLE
1379      C          N - NUMBER OF PAIRED OBSERVATIONS.
1380      C
1381      C OUTPUT: TAUA - KENDALL'S TAU-A ASSUMING NO TIES.
1382      C          TAUB - KENDALL'S TAU-B FOR TIED RANKS.
1383      C          SOMER - SOMER'S D FOR TIED RANKS IN THE DEPENDENT
1384      C                 VARIABLE.
1385      C          NC - NUMBER OF CONCORDANT PAIRS
1386      C          ND - NUMBER OF DISCORDANT PAIRS
1387      C          NIND - NUMBER OF PAIRS WITH TIES IN THE INDEPENDENT
1388      C                 VARIABLE BUT NOT IN THE DEPENDENT VARIABLE
1389      C          NDEP - NUMBER OF PAIRS WITH TIES IN THE DEPENDENT
1390      C                 VARIABLE BUT NOT IN THE INDEPENDENT VARIABLE
1391      C          NPAIRS - NUMBER OF PAIRS OF BIVARIATE OBSERVATIONS
1392      C
1393      C SUBROUTINES CALLED: ORB2
1394      C-----
1395      C
1396      COMMON /DATA/ X(500),Y(500),RANKX(500),RANKY(500)
1397      DIMENSION T1(500,2)
1398      DO 10 I=1,N
1399      10 T1(I,1)=RANKX(I)
1400      T1(I,2)=RANKY(I)
1401      10 CONTINUE
1402      C
1403      C INITIALIZE NC,ND,NIND,NDEP AND BEGIN COUNTING PROCEDURE
1404      C
1405      NC=0
1406      ND=0
1407      NIND=0
1408      NDEP=0
1409      I=1
1410      20 K=I+1
1411      30 IF(K.GT.N) GO TO 60
1412      C
1413      C CHECK IF RANKS ARE EQUAL
1414      C
1415      C TOLERANCE FOR REAL ARITHMETIC: ASSUME N=0 IF ABS(X).LE.0.E-6
1416      C
1417      TEST1=ABS(T1(I,1)-T1(K,1))
1418      IF(TEST1.GT.0.E-6) GO TO 40
1419      TEST2=ABS(T1(I,2)-T1(K,2))
1420      IF(TEST2.GT.0.E-6) NIND=NIND+1
1421      K=K+1
1422      GO TO 30
1423      C
1424      C COUNT NUMBER OF CONCORDANT AND DISCORDANT PAIRS
1425      C
1426      40 DO 50 J=K,N
1427      IF(T1(J,2)-T1(I,2).GT.0.E-6) NC=NC+1
1428      IF(T1(J,2)-T1(I,2).LT.-0.E-6) ND=ND+1
1429      50 CONTINUE
1430      GO TO 1+1
1431      IF(I.LY.N) GO TO 20
1432      C
1433      C SWITCH COLUMNS OF T1
1434      C
1435      DO 70 I=1,N
1436      T1(I,2)=RANKX(I)
1437      T1(I,1)=RANKY(I)
1438      70 CONTINUE
1439      C
1440      C ORDER BY Y VALUES AND COMPUTE NDEP
1441      C
1442      CALL ORB2(T1,N,500)

```



```

1700. SUBROUTINE PAR2(RVAR,M,N,CAT,NORD)
1701. C-----
1702. C SUBROUTINE TO DETERMINE THE ORDER OF AN AUTOREGRESSIVE
1703. C PROCESS BY PARZEN'S CAT CRITERIA
1704. C INPUT :
1705. C   M,RVAR(1),...,RVAR(M) : STANDARDIZED RES VAR
1706. C   FOR ORDERED 1 THRU M.
1707. C   N : SAMPLE SIZE
1708. C OUTPUT :
1709. C   NORD : DETERMINED ORDER
1710. C   CAT(1),...,CAT(M)
1711. C SUBROUTINES CALLED : MIN
1712. C-----
1713. C DIMENSION RVAR(M),CAT(M)
1714. C   ON=FLOAT(N)
1715. C   DO 1 I=1,M
1716. C     C=0.
1717. C     DO 2 J=1,I
1718. C       C=C+1.-((FLOAT(J)/ON))/RVAR(J)
1719. C       C=C/ON
1720. C     CAT(I)=C-((1.-((FLOAT(I)/ON))/RVAR(I))
1721. C     CALL MIN(CAT,M,CATM,NORD)
1722. C     IF(CATM.GT.-1.-3./ON) NORD=0
1723. C     RETURN
1724. C   END
1725. C SUBROUTINE PEARSH(N,R)
1726. C-----
1727. C SUBROUTINE TO COMPUTE PEARSON'S PRODUCT MOMENT CORRELATION
1728. C COEFFICIENT.
1729. C INPUT : X - THE VECTOR OF VALUES OF THE INDEPENDENT VARIABLE
1730. C   Y - THE VECTOR OF VALUES OF THE DEPENDENT VARIABLE
1731. C   N - THE NUMBER OF DIVARIATE OBSERVATIONS
1732. C OUTPUT : R - PEARSON'S PRODUCT MOMENT CORRELATION COEFFICIENT
1733. C-----
1734. C COMMON /DATA/ X(500),Y(500),RANKX(500),RANKY(500)
1735. C INITIALIZE VALUES
1736. C
1737. C   SUMX=0.
1738. C   SUMY=0.
1739. C   SUMXY=0.
1740. C   SUMX2=0.
1741. C   SUMY2=0.
1742. C   SUMXY=0.
1743. C COMPUTE SUMS OF SQUARES AND CROSS PRODUCTS
1744. C
1745. C   DO 10 I=1,N
1746. C     SUMX=SUMX+X(I)
1747. C     SUMY=SUMY+Y(I)
1748. C     SUMXY=SUMXY+X(I)*Y(I)
1749. C     SUMX2=SUMX2+X(I)*X(I)
1750. C     SUMY2=SUMY2+Y(I)*Y(I)
1751. C   10 CONTINUE
1752. C COMPUTE R
1753. C
1754. C   SX=SUMX-SUMX*SUMX/FLOAT(N)
1755. C   SY=SUMY-SUMY*SUMY/FLOAT(N)
1756. C   SVY=SUMYV-SUMY*SUMY/FLOAT(N)
1757. C   R=SVX/SQRT(SX*SVY)
1758. C   RETURN
1759. C END
1760. C SUBROUTINE PLINE(NR,NC,NL,STRT,END,CHARL,PARRAY,VMIN,VMAX,VINC)
1761. C DIMENSION STRT(NL),END(NL),CHARL(NL),PARRAY(NR,NC)
1762. C   NCHI = NC - 1
1763. C   PNC = 1. / FLOAT(NCHI)
1764. C   NRP1 = NR + 1
1765. C
1766. C   DO 10 IL = 1, NL
1767. C     IF ( STRT(IL) - END(IL) ) 20,20,40
1768. C     20 STRT(IL) = AMAX1(STRT(IL),VMIN)
1769. C     END(IL) = AMIN1(END(IL),VMAX)
1770. C     GOTO 40
1771. C     40 STRT(IL) = AMIN1(STRT(IL),VMAX)
1772. C     END(IL) = AMAX1(END(IL),VMIN)
1773. C     60 DS = (END(IL) - STRT(IL)) * PNC * VINC
1774. C     S = (STRT(IL) - VMIN ) * VINC + 1.5
1775. C     IND = INT(S)
1776. C     DO 60 I = 2,NCHI
1777. C       S = S + DS
1778. C       IF ( INT(S) .EQ. IND ) GOTO 60
1779. C       IND = INT(S)
1780. C       PARRAY(NRP1-IND,I) = CHARL(IL)
1781. C     CONTINUE
1782. C     GOTO 10
1783. C     30 IF ( STRT(IL) .LE. VMIN .OR. STRT(IL) .GE. VMAX) GOTO 10
1784. C     IND = (STRT(IL) - VMIN) * VINC + 1.5
1785. C     DO TO I = 2,NCHI
1786. C       PARRAY(NRP1-IND,I) = CHARL(IL)
1787. C     CONTINUE
1788. C     10 RETURN
1789. C   END
1790. C SUBROUTINE PLOTXY(X,V,N,CAPT,NAMX,NAMY,IOPY)
1791. C-----
1792. C SUBROUTINE TO PRINT AND PRINTER PLOT THE N-VECTOR V AS A
1793. C FUNCTION OF X.
1794. C INPUT : N,N,V - N IS ORDERED ON INPUT AND V(1)=V(X(1))
1795. C   CAPT - LITERAL CONSTANT FOR TITLE OF PLOT IN 2044 FORMAT
1796. C   NAMX,NAMY : 0 CHARACTER LITERAL CONSTANTS GIVING
1797. C   LABELS FOR X AND Y
1798. C   IOPY : 1,2 (POINT OR BAR PLOT)
1799. C SUBROUTINES CALLED : FTERP,MAX,MIN
1800. C-----
1801. C DIMENSION X(N),V(N),T(46),VI(46),CAPT(20),AL(101)
1802. C   DATA HOUT/6/
1803. C   DATA BLANK,OUT,2,SL,PLUS/1N,1N,1N,1N,1N+/
1804. C
1805. C   NHI=0
1806. C   IOPY=0
1807. C   IF(N.GT.10) GO TO 11
1808. C   WRITE(ROUT,10) N
1809. C   10 FORMAT(10X,'SAMPLE SIZE OF ',12,' IS TOO SMALL TO PERFORM ',
1810. C   'INTERPOLATION IN PLOTXY.')
1811. C   GO TO 100
1812. C   11 CONTINUE
1813. C   WRITE(ROUT,12) CAPT
1814. C   12 FORMAT(1N1,32X,2044,/)
1815. C
1816. C   CREATE Y VECTOR OF EQUALLY SPACED X AND INTERPOLATE TO OBTAIN
1817. C   CORRESPONDING Y VALUES

```

```

1040 C      DEC=(X(N)-X(1))/45.0
1041      DO 10 J=1,45
1042      Y(1)=X(1)+FLOAT(J-1)*DEC
1043      CONTINUE
1044      CALL PTERP(X,V,Y,VI,N,45)
1045
1046 C      INITIALIZE AL :
1047 C
1048 C      ON=(NM-1)/2
1049      DO 30 J=1,NM
1050      AL(J)=DOT
1051      WRITE(INDUT,25) NAME,NAMY,(AL(J),J=1,NM)
1052      FORMAT(/,10X,A4,6X,A4/10X,20(1N-1,2X,101A1))
1053      DO 30 J=1,NM
1054      AL(J)=BLANK
1055      AL(1)=BL
1056      AL(NM)=BL
1057
1058 C      FIND MAX AND MIN :
1059 C
1060      CALL MAX(V,45,VMAX,IND)
1061      CALL MIN(V,45,VMIN,IND)
1062      RV=1.2*(VMAX-VMIN)
1063      IF(RV.LT.1.E-20) ISPTV=1
1064
1065 C      PLOT :
1066 C
1067      DO 40 J=1,45
1068      IF(ISPTV.EQ.1) GO TO 36
1069      C1=(V(J)-VMIN)/RV
1070      C1=2.*(C1-.5)
1071      GO TO 37
1072
1073      36 C1=0.
1074      37 K=ON+(C1+.1)*2.5
1075      AL(K)=2
1076      IF(ISPTV.EQ.1) GO TO 38
1077      DO 30 J=1,K
1078      AL(J)=2
1079      CONTINUE
1080      WRITE(INDUT,30) Y(J),V(J),(AL(I),I=1,NM)
1081      FORMAT(10X,F10.4,1X,F9.4,2X,101A1)
1082      AL(K)=BLANK
1083      IF(ISPTV.EQ.1) GO TO 40
1084      DO 41 J=2,K
1085      AL(J)=BLANK
1086      CONTINUE
1087      DO 50 J=1,NM
1088      AL(J)=DOT
1089      AL(1)=PLUS
1090      AL(NM)=PLUS
1091      WRITE(INDUT,60) (AL(I),I=1,NM)
1092      FORMAT(10X,20(1N-1,2X,101A1))
1093
1094 C      VMAX=RV+VMIN
1095      WRITE(INDUT,70) VMIN,VMAX
1096      FORMAT(27X,F10.4,70X,F10.4)
1097      70
1098      CONTINUE
1099      RETURN
1100      END
1101
1102 C-----
1103 C      SUBROUTINE PPLLOT(X,V,IV,N,NFUN,CHAR,CAPT1,XNAME,YNAME,ISPT)
1104 C-----
1105 C      SUBROUTINE PPLLOT PLOTS UP TO 5 FUNCTIONS ON THE SAME AXIS USING
1106 C      A DIFFERENT SYMBOL FOR EACH.
1107 C
1108 C      INPUT: X - VECTOR OF X VALUES, LENGTH=N
1109 C      V - VECTOR OR MATRIX OF Y VALUES, SIZE=N BY NFUN
1110 C      IV - ROWS ALLOCATED TO V IN CALLING PROGRAM
1111 C      N - NUMBER OF X-VALUES
1112 C      NFUN - NUMBER OF FUNCTIONS TO BE PLOTTED (1-5)
1113 C      CHAR - VECTOR OF LENGTH 5 CONTAINING CHARACTERS TO BE
1114 C      USED IN THE PLOT. (CHAR(1) IS USED FOR THE FIRST
1115 C      FUNCTION, CHAR(2) FOR THE SECOND, ETC.)
1116 C      CAPT1 - VECTOR OF LENGTH 20 TO BE USED FOR
1117 C      CAPTION ABOVE THE PLOT (20A1 FORMAT)
1118 C      XNAME,YNAME - 2 VECTORS OF LENGTH 20 TO BE USED FOR
1119 C      LABELS ON THE X AND Y AXES (20A1 FORMAT)
1120 C      ISPT=0 --> SMALLEST POINT WILL BE DIRECTLY ON AXIS
1121 C      ISPT=1 --> SMALLEST POINT WILL BE SLIGHTLY AWAY
1122 C      FROM THE AXIS
1123 C
1124 C      SUBROUTINES CALLED: MINMAX
1125 C
1126 C      PROGRAMMER: PHIL SPECTOR
1127 C-----
1128 C-----
1129      DIMENSION PARRAY(55,75),XJ(15),YJ(15),INDY(5),CHAR(5),CAPT1(20),
1130      XNAME(20),YNAME(20),YNAME1(50),XIN),V(IV,NFUN)
1131      DATA CAP1,PLUS,DASH,XM,BLANK/INT,INT,INT,INT,INT,INT /
1132      DO 1 I=1,15
1133      1 YNAME1(I)=BLANK
1134      DO 2 I=1,30
1135      2 YNAME1(10+I)=YNAME(I)
1136      DO 3 I=1,10
1137      3 YNAME1(30+I)=BLANK
1138      LN=75
1139      LV=55
1140      LVP1=LV+1
1141      LVN1=LV-1
1142      CALL MINMAX(V,IV,N,NFUN,VMIN,VMAX)
1143      CALL MINMAX(X,N,1,NMIN,NMAX)
1144      IF(ISPT.EQ.1)VMIN=VMIN-1.1*((VMAX-VMIN)/FLOAT(LV-1))
1145      IF(ISPT.EQ.1)NMIN=NMIN-1.1*((NMAX-NMIN)/FLOAT(LN-1))
1146      VINC=(VMAX-VMIN)/FLOAT(LV-1)
1147      NINC=(NMAX-NMIN)/FLOAT(LN-1)
1148      IF(VINC.EQ.0.0)GO TO 55
1149      DO 10 I=1,LV
1150      PARRAY(I,1)=CAP1
1151      DO 6 J=2,LN
1152      PARRAY(I,J)=BLANK
1153      CONTINUE
1154      10 PARRAY(1,1)=PLUS
1155      DO 11 I=1,LN
1156      11 PARRAY(1,I)=DASH
1157      J=0
1158      DO 12 I=1,LN,5
1159      J=J+1
1160      XJ(J)=NMIN+FLOAT(J)*5.*NINC
1161      12 PARRAY(1,I)=PLUS
1162      J=0
1163      DO 13 I=1,LV,5
1164      J=J+1
1165      YJ(J)=VMIN+FLOAT(J)*5.*VINC
1166      13 PARRAY(I,1)=PLUS
1167      DO 45 K=1,N
1168      INDX=(X(K)-NMIN)/NINC+1.0
1169      DO 44 L=1,NFUN
1170      44 INDY(L)=IV(K,L)-VMIN)/VINC+1.0
1171      PARRAY(INDY(1),INDX)=CHAR(1)

```

```

1072 IF (NPNUM.EQ.1) GO TO 40
1073 DO 40 I=1,NPNUM
1074 IF (PARRAY(INDV(I),INDX).NE.BLANK) GO TO 40
1075 PARRAY(INDV(I),INDX)=CHAR(I)
1076 GO TO 40
1077 40 PARRAY(INDV(I),INDX)=NM
1078 40 CONTINUE
1079 40 CONTINUE
1080 WRITE(6,902)CAPT1
1081 K=0
1082 KOUNT=1
1083 DO 50 J=1,LVM1
1084 K=K+1
1085 IF (K.EQ.5) GO TO 50
1086 WRITE(6,905)VNAME1(1:1),(PARRAY(LVP1-1,J),J=1,LN)
1087 GO TO 50
1088 50 WRITE(6,906)VNAME1(1:1),VJ(12-KOUNT),(PARRAY(LVP1-1,J),J=1,LN)
1089 KOUNT=KOUNT+1
1090 K=0
1091 55 CONTINUE
1092 WRITE(6,908)BLANK,VMIN,(PARRAY(1,J),J=1,LN)
1093 WRITE(6,910)VMIN,(NJ(1),I=2,10,2)
1094 WRITE(6,910)(NJ(1),I=1,10,2)
1095 WRITE(6,920)NAME
1096 GO TO 100
1097 90 WRITE(6,901)
1098 100 CONTINUE
1099 901 FORMAT(1H,'ERROR IN PLOT: RANGE OF V VALUES TOO SMALL')
1100 902 FORMAT(1H,20X,20A4)
1101 905 FORMAT(2X,A1,10X,100A1)
1102 906 FORMAT(2X,A1,3X,F10.3,2X,100A1)
1103 910 FORMAT(12X,12(P9.4,1X))
1104 916 FORMAT(10X,9(P9.4,1X))
1105 920 FORMAT(40X,20A1)
1106 RETURN
1107 END
1108 FUNCTION OPIND(O,N,QUANT)
1109 C-----
1110 C PURPOSE: TO FIND THE PERCENTILE VALUE OF Q
1111 C AT QUANT.
1112 C INPUT:
1113 C O - VECTOR OF SIZE N
1114 C N - NUMBER OF VALUES IN V
1115 C QUANT - QUANTILE VALUE
1116 C-----
1117 DIMENSION O(N)
1118 P = FLOAT(N+1) * QUANT
1119 I = INT(P)
1120 P = P-I
1121 IF (I.NE.0) GO TO 1
1122 OPIND = O(1)
1123 RETURN
1124 IF (I.LT.N) GO TO 2
1125 OPIND = O(N)
1126 RETURN
1127 2 OPIND = (1.-P) * O(I) + P * O(I+1)
1128 RETURN
1129 END
1130 SUBROUTINE PLOT(X,V,NV,ICHAR,XZ,Z,NZ,IZ,ZCHAR,IORD,VMIN,VMX,
1131 NLINE,STRT,END,CHARL,LX,LV,LC,LI)
1132 C-----
1133 C ROUTINE TO DISPLAY PRINTER QUANTILE-BOX PLOT
1134 C INPUT:
1135 C X - VECTOR CONTAINING THE VALUES J/(NV+1) WHERE
1136 C J=1,2,3,...,NV
1137 C V - VECTOR OF SIZE NV TO BE PLOTTED
1138 C ICHAR - CHARACTER FOR V IN PLOT
1139 C LX,LV - VECTORS OF SIZE 2 CONTAINING THE LABELS
1140 C FOR X AND Y RESPECTIVELY
1141 C LC - VECTOR OF SIZE 20 CONTAINING THE CAPTION FOR
1142 C THE PLOT
1143 C Z - OPTIONAL VECTOR OF SIZE NZ TO BE PLOTTED
1144 C XZ - ABSCISSA FOR Z
1145 C NLINE,STRT,END - VALUES FOR SUB. PLINE WHEN Z IS A LINE
1146 C VMIN,VMX - MIN AND MAX VALUES FOR ORIGINATE OF PLOT
1147 C IORD - EQUAL ZERO IF BOX PLOTS ARE NOT WANTED
1148 C IZ - EQUAL ZERO IF VECTOR Z IS TO BE PLOTTED
1149 C ZCHAR - PLOTTING CHARACTER TO BE USED FOR THE
1150 C Z VECTOR. MUST BE DIFFERENT FROM O.
1151 C LI - OPTIONAL LABEL TO FOLLOW CAPTION
1152 C IORD = 1 IF V IS ORDERED FROM MIN TO MAX
1153 C = 0 IF V IS NOT ORDERED
1154 C = 2 FOR HORIZONTAL ZERO-LINE (FOR 10 PLOT)
1155 C-----
1156 DIMENSION PARRAY(51,61),CHAR(10),X(NV),V(NV),LX(2),LV(2),
1157 LC(20),XZ(1),Z(NZ),LI(20),XZ(NZ),STRT(1),END(1),CHARL(1)
1158 DATA BLANK,DASH,PLUS,CAPT,ZERO / ' ','+', '-', '0' /
1159 DATA CHAR,CHARL / 'O', 'O', '2', '2', '4', '6', '6', '7', '8', '9', 'M' /
1160 DATA EPS / 1.E-10 /
1161 JINC = 0
1162 KINC = 0
1163 NR = 1 + 10 * JINC
1164 NC = 1 + 10 * KINC
1165 NRP1 = NR + 1
1166 NRM1 = NR - 1
1167 NCM1 = NC - 1
1168 PC = FLOAT(NCM1)
1169 NUNIT = 0
1170 CHAR(1) = ICHAR
1171 IF (CHAR(1) EQ. BLANK) CHAR(1) = CHAR1
1172 IF (IORD=1) 2,3,4
1173 2 CALL MIN(V,NV,VMIN,IER)
1174 CALL MAX(V,NV,VMX,IER)
1175 IF (IZ.EQ.0) GO TO 1
1176 CALL MIN(Z,NZ,ZMIN,IER)
1177 CALL MAX(Z,NZ,ZMAX,IER)
1178 IF (ZMIN.LT.VMIN) VMIN=ZMIN
1179 IF (ZMAX.GT.VMX) VMX=ZMAX
1180 GO TO 1
1181 3 VMIN = V(1)
1182 VMX = V(NV)
1183 IF (IZ.EQ.0) GO TO 1
1184 ZMIN = Z(1)
1185 ZMAX = Z(NZ)
1186 IF (ZMIN.LT.VMIN) VMIN = ZMIN
1187 IF (ZMAX.GT.VMX) VMX = ZMAX
1188 GO TO 1
1189 4 VMIN = VNM
1190 VMX = VMN
1191 CONTINUE
1192 RANGE = VMX - VMIN
1193 IF (RANGE.LT.EPS) GO TO 99
1194 VINC = FLOAT(NRM1) / RANGE
1195 DO 10 I=1,NRM1
1196 PARRAY(I,1)=CAPT
1197 PARRAY(I,NC)=CAPT
1198 DO 5 J=2,NCM1
1199 PARRAY(I,J)=BLANK
1200 CONTINUE
1201 10 CONTINUE
1202 DO 20 J=1,NC
1203 PARRAY(1,J) = DASH

```



```

2360 20 ON(1)+V(1)
2361 TIO=1./((X75-X25))
2362 DO 30 I=1,N
2363 ON(1)=ON(1)-X50)+TIO
2364 WK(1)=U(1)-.5
2365 30 WK2(1)=ON(1)
2366 DO 40 I=1,N
2367 IF(WK2(1)).GE.-1.) GO TO 50
2368 40 WK2(1)=.1
2369 DO 60 I=1,N
2370 IF(WK2(N-1+I)).LE.1.) GO TO 70
2371 60 WK2(N-1+I)=.1
2372 70 CONTINUE
2373 CALL OPLOT(U,WK2,NOP1,BLK,O.,O.,1,O,BLK,2,-1.,1.,2,STRT,FND,
2374 & CHARL,NAMU,NAMIO,NAME,LAB1)
2375 DO 75 I = 1,2
2376 EXIO(1) = OFIND(ON,NOP1,EXU(1))
2377 75 CONTINUE
2378 WRITE(UNIT,60) EXU,EXIO
2379 60 FORMAT(///TIO, ' U ',3X,5(P6.5,1X),
2380 & /TIO, ' IO(U)',3X,5(P6.5,1X),///)
2381 KASE=INUL+2
2382 ICASE(1)=LCASE(KASE-1)
2383 ICASE(2)=LCASE(KASE)
2384 C COMPUTE AND PLOT RAW SPACINGS (10) AND RAW FO (11/0)
2385 CALL STOPS(ON,U,NOP1,WK1,SPEPAC)
2386 C COMPUTE AND PLOT WEIGHTED SPACINGS FOR CASE ICASE
2387 DO 100 I = 1,NO
2388 WK2(1) = POPNC(U(1+I),INUL)
2389 100 CONTINUE
2390 CALL WSPACE(WK5,D,NOP1,WK1,WK2,U,SW)
2391 C PLOT CUMULATIVE WEIGHTED SPACINGS WITH D+ AND D-
2392 CALL KSD(D,U,NOP1,DM,UM,OP,UP)
2393 CALL PCODEA(OP,5H(P7.4) ,LAB9(2))
2394 CALL PCODEA(UP,5H(P7.4) ,LAB9(5))
2395 CALL PCODEA(DM,5H(P7.4) ,LAB9(10))
2396 CALL PCODEA(UM,5H(P7.4) ,LAB9(14))
2397 LAB10(1)=ICASE(1)
2398 LAB10(5)=ICASE(2)
2399 CALL OPLOT(U,D,NOP1,BLK,O.,O.,1,O,BLK,2,O.,1.,1,O.,1.,ASTER,
2400 & NAMU,NAMCWS,LAB10,LAB9)
2401 WRITE(UNIT,130)
2402 130 FORMAT(////////)
2403 RETURN
2404 END
2405 SUBROUTINE QUICK(N,T)
2406 C-----
2407 C QUICK SORT. THIS ALGORITHM IS ALSO REFERRED TO AS A PARTITIONED
2408 C EXCHANGE SORT. EXPECTED RUNTIME IS PROPORTIONAL TO N*LOG2(N)
2409 C ALTHOUGH THE WORST CASE IS PROPORTIONAL TO N**2.
2410 C REFERENCE: DONALD E. KNUTH- THE ART OF COMPUTER PROGRAMMING VOL 3.
2411 C INPUT :
2412 C X,N : VECTOR TO BE SORTED OF LENGTH N
2413 C OUTPUT :
2414 C X : SORTED VECTOR
2415 C SUBROUTINES CALLED : NONE
2416 C-----
2417 REAL T(N),V
2418 INTEGER IP,LV(16),IV(16),LP,IUP
2419 LV(1)=1
2420 IV(1)=N
2421 IP=1
2422 10 IF(IP.LT.1) GO TO 75
2423 15 IF((IV(IP)-LV(IP)).LT.1) GO TO 20
2424 GO TO 35
2425 20 IP=IP-1
2426 GO TO 10
2427 25 LP=LV(IP)-1
2428 IUP=IV(IP)
2429 V=V(IP)
2430 IF((IUP-LP).LT.2) GO TO 45
2431 LP=LP+1
2432 IF(T(LP).LE.V) GO TO 30
2433 T(IUP)=T(LP)
2434 35 IF((IUP-LP).LT.2) GO TO 40
2435 IUP=IUP-1
2436 IF(T(IUP).GE.V) GO TO 35
2437 T(LP)=T(IUP)
2438 GO TO 30
2439 40 IUP=IUP-1
2440 45 T(IUP)=V
2441 IF((IUP-LV(IP)).LT.(IV(IP)-IUP)) GO TO 55
2442 GO TO 60
2443 55 LV(IP+1)=LV(IP)
2444 IV(IP+1)=IUP-1
2445 LV(IP)=IUP+1
2446 GO TO 70
2447 60 LV(IP+1)=IUP+1
2448 IV(IP+1)=IV(IP)
2449 IV(IP)=IUP-1
2450 70 IP=IP+1
2451 GO TO 15
2452 75 RETURN
2453 END
2454 SUBROUTINE RANK(X,N,RR)
2455 C-----
2456 C SUBROUTINE TO RANK THE N-VECTOR X WITH RANKS PLACED IN
2457 C THE N-VECTOR RR.
2458 C TIED VALUES ARE GIVEN AVERAGE RANKS.
2459 C SUBROUTINE CALLED: GRD2
2460 C-----
2461 DIMENSION X(N),RR(N),W1(500,2),W2(500,2),Y(500)
2462 C CREATE MATRIX W1 WHOSE FIRST COLUMN CONTAINS THE VECTOR X
2463 C AND WHOSE SECOND COLUMN CONTAINS THE OBSERVATION NUMBER
2464 C
2465 DO 10 I=1,N
2466 W1(I,1)=X(I)
2467 W1(I,2)=I*FLOAT(1)
2468 10 CONTINUE
2469 CALL GRD2(W1,N,500)
2470 C CREATE MATRIX W2 WHOSE FIRST COLUMN CONTAINS SORTED X VALUES
2471 C AND WHOSE SECOND COLUMN CONTAINS THE RANKS OF THE X VALUES
2472 C BEFORE CORRECTING FOR TIES.
2473 C
2474 DO 20 I=1,N
2475 W2(I,1)=W1(I,1)
2476 W2(I,2)=I*FLOAT(1)
2477 20 CONTINUE
2478 C CORRECT FOR TIES BY REPLACING RANKS OF TIED VALUES BY
2479 C AVERAGE RANK
2480 C
2481 I=1
2482 30 K=I+1

```



```

0001 05 V(1)=V(1) 01SAM
0002 CALL ORD2(DELM,KDEL,20) 01SAM
0003 DO 70 I=1,KHALF 01SAM
0004   IDEL(KHALF+1)=IFIN(DELM(KHALF+1,2)+0.5) 01SAM
0005 70 CONTINUE 01SAM
0006   KM1=KHALF+1 01SAM
0007   DO 80 J=1,KDIM 01SAM
0008   DO 80 J=KM1,KDIM 01SAM
0009   IF(IDEL(I).NE.IDEL(J)) GO TO 80 01SAM
0010   IDEL(J)=N+1 01SAM
0011   KDEL=KDEL+1 01SAM
0012 80 CONTINUE 01SAM
0013   DO 82 I=1,KDIM 01SAM
0014   KDEL(I)=FLOAT(IDEL(I)) 01SAM
0015   CALL QUICK(KDIM,KDEL) 01SAM
0016   DO 84 I=1,KDIM 01SAM
0017   IDEL(I)=IFIN(IDEL(I)+0.5) 01SAM
0018 84 CONTINUE 01SAM
0019   IF(KENK.EQ.1) KDEL=1 01SAM
0020   NM1=N-1 01SAM
0021   DO 100 I=1,KDEL 01SAM
0022   IND=IDEL(KDEL-I+1) 01SAM
0023   NEWN=NEWN-1 01SAM
0024   IF(IND.EQ.N) GO TO 92 01SAM
0025   DO 90 J=IND,NM1 01SAM
0026   X(J)=X(J+1) 01SAM
0027   V(J)=V(J+1) 01SAM
0028 90 CONTINUE 01SAM
0029   X(NEWN+1)=0.0 01SAM
0030   V(NEWN+1)=0.0 01SAM
0031   NM1=NEWN 01SAM
0032 100 CONTINUE 01SAM
0033   WRITE(6,110) 01SAM
0034 110 FORMAT(0,'THE FOLLOWING POINTS WERE DELETED FROM THE DATA SET:') 01SAM
0035   WRITE(6,120) (IDEL(I),I=1,KDEL) 01SAM
0036 120 FORMAT(10X,615) 01SAM
0037   WRITE(6,130) KDEL,NEWN 01SAM
0038 130 FORMAT(10X,15,' POINTS WERE DELETED LEAVING ',15,' POINTS ', 01SAM
0039   ' IN THE DATA SET.') 01SAM
0040   RETURN 01SAM
0041   END 01SAM
0042   SUBROUTINE WSPACE(WKS,CWKS,NOP1,XS,FOOO,U,SIG0,AVLWX) 01SAM
0043 C----- 01SAM
0044 C SUBROUTINE TO COMPUTE D(U), CUMULATIVE D'S, AND SIGMAO 01SAM
0045 C FOR THE MODEL O(U)=MU+SIGMA*OO(U) 01SAM
0046 C INPUT : 01SAM
0047 C   XS,N0 : VECTOR OF LENGTH N0 CONTAINING LITTLE O(U) 01SAM
0048 C   FOOO : HYPOTHEZIZED DENSITY QUANTILE FUNCTION 01SAM
0049 C   U : VECTOR OF LENGTH N0 CONTAINING U VALUES 01SAM
0050 C OUTPUT : 01SAM
0051 C   WKS : VECTOR OF LENGTH N0 CONTAINING D(U) 01SAM
0052 C   CWKS : VECTOR OF LENGTH N0 CONTAINING THE 01SAM
0053 C   CUMULATIVE D'S 01SAM
0054 C   SIG0 : COMPUTED VALUE OF SIGMAO = CWKS(N0) 01SAM
0055 C 01SAM
0056 C SUBROUTINES CALLED : NONE 01SAM
0057 C----- 01SAM
0058   DIMENSION FOOO(NOP1),XS(NOP1),U(NOP1),WKS(NOP1),CWKS(NOP1) 01SAM
0059   NO=NOP1-1 01SAM
0060   CWKS(1)=0. 01SAM
0061   DO 10 I=1,NO 01SAM
0062   WKS(I)=FOOO(U(I)) 01SAM
0063   CWKS(I)=CWKS(I)+WKS(I) 01SAM
0064 10 CONTINUE 01SAM
0065   FN=FLOAT(N0) 01SAM
0066   D1=1./CWKS(NOP1) 01SAM
0067   D2=D1*FN 01SAM
0068   DO 20 I=1,NO 01SAM
0069   WKS(I)=WKS(I)+D1 01SAM
0070   CWKS(I)=CWKS(I)+D1 01SAM
0071 20 CONTINUE 01SAM
0072   CWKS(NOP1)=1. 01SAM
0073   SIG0=1./D2 01SAM
0074   RETURN 01SAM
0075   END 01SAM

```

REFERENCES

- Kimeldorf, George, and Sampson, Allen R. (1975), "Uniform Representations of Bivariate Distributions," Comm. Stat., 4, 617-628.
- Newton, H. Joseph (1979), "The TIMESBOARD Time Series Analysis Computing Library," Institute of Statistics, Texas A&M University, College Station, Texas.
- Parzen, Emanuel (1979), "Nonparametric Statistical Data Modeling," JASA, 74, 105-131.
- Parzen, Emanuel, and Anderson, Scott (1980), "ONESAM, A Computer Program for Nonparametric Data Analysis and Density Estimation," Technical Report B-1, Institute of Statistics, Texas A&M University, College Station, Texas.
- Prihoda, Thomas J. (1981), "A Generalized Approach to the Two Sample Problem: The Quantile Approach," unpublished Ph.D. dissertation, Institute of Statistics, Texas A&M University.
- SAS/GRAPH User's Guide (1981), SAS Institute Inc., Cary, North Carolina.
- Tartar, M. E., and Kronmal, R. A. (1970), "On Multivariate Density Estimates Based on Orthogonal Expansions," Annals of Math. Stat., 41, 718-722.
- ____ (1976), "An Introduction of the Implementation and Theory of Nonparametric Density Estimation," American Statistician, 30, 105-112.
- Woodfield, Terry J. (1982), "Statistical Modeling of Bivariate Data," unpublished Ph.D. dissertation, Institute of Statistics, Texas A&M University.

