A SAS MACRO FOR PLOTTING "PAIR CHART" AND OBTAINING TWO SAMPLES
MANN-WHITNEY, WILCOXON AND KOLMOGOROV-SMIRNOV STATISTICS

Agam N. Sinha, Kenneth A. Hardy
Social Science Statistical Laboratory
Institute for Research in Social Science
University of North Carolina, Chapel Hill, N.C.

ABSTRACT

The "pair chart" is a useful, graphical representation of ordered data for two-sample problems. In addition, it provides a convenient format for computing common two-sample inferential statistics such as the Wilcoxon, Mann-Whitney and Kolmogorov-Smirnov tests. This paper presents a SAS macro for plotting pair charts on the pen plotter at the UNC Computing Center (or the screen of a Tektronix video terminal) and also calculating the above two-sample test statistics.

Introduction

The nonparametric statistical technique for the classical two-sample comparison problem consists basically in computing either (a) a linear rank statistic, or (b) a statistic based on the empirical distribution functions of the two samples. The classical statistics belonging to the former category are those of Mann-Whitney and Wilcoxon, and statistics belonging to the latter category are those of Kolmogorov-Smirnov and Cramer-von Mises. Quade (1973) has given an ingenious method of comparing two samples pictorially through a diagram that he calls a "pair chart". He has discussed many descriptive uses of pair charts and also explained how to interpret them. The present paper reviews the computational procedures involved with the two-sample Mann-Whitney, Wilcoxon and Kolmogorov-Smirnov statistics, and presents a SAS macro for computing these statistics, and plotting pair charts using a Calcomp plotter or Tektronix video terminal.

Mann-Whitney and Wilcoxon Statistics

Let $X_1, X_2, \ldots, X_M$ be an ordered random sample of $M$ observations from an unknown distribution $F(x)$, and let $Y_1, Y_2, \ldots, Y_N$ be a random sample of $N$ observations from some other unknown distribution $G(y)$, where the two samples are drawn quite independently. The classical two-sample problem consists in testing the null hypothesis

$$
H_0: F(x) = G(y) \quad (1)
$$

$$
H_1: F(x) \neq G(y)
$$

The Mann-Whitney (1947) statistic to test the above hypothesis is computed as follows.

First, all possible pairs of observations $(X_i, Y_j; i = 1, 2, \ldots, M, j = 1, 2, \ldots, N)$ are formed yielding $MN$ such pairs. These pairs are then sorted into three categories as follows:

Type I: pairs in which $X_i > Y_j$;

Type II: pairs in which $X_i < Y_j$; and

Type III: pairs in which $X_i = Y_j$ (ties).

The Mann-Whitney statistics $U_x$ and $U_y$ are given by

$$
U_x = \text{(Number of Type II pairs)} + \frac{1}{2} \cdot \text{(Number of Type III pairs)} \quad (2)
$$

$$
U_y = MN - U_x
$$

To compute the Wilcoxon statistic for testing the null hypothesis (1), first combine the two-sample observations, and then rank order their values keeping track of which sample each observation came from. If $R_i$ is the rank of $X_i$ in the combined samples ($i=1, 2, \ldots, M$), it can be shown that

$$
R_i = i + \text{(Number of Type I pairs)} \quad (3)
$$

where $i = 1, 2, \ldots, M$.

The Wilcoxon statistic is defined as

$$
T_x = \sum_{i=1}^{M} R_i \quad (4)
$$

The relationship between $T_x$ and $U_x$ is given by

$$
T_x = U_x + \frac{M(N+1)}{2} \quad (5)
$$

For purposes of testing significance, the percentile points of the distributions of $U_x$ and $T_x$ under $H_0$ can be found in any text on nonparametric statistics (e.g. Siegel, 1956; Hollander and Wolfe, 1973).

Kolmogorov-Smirnov Statistic

The two-sample Kolmogorov-Smirnov statistic for testing the null hypothesis (1) is calculated as follows:

Assume that the variables on which observations have been made are continuous (i.e. interval scale variables). Let $F_1(z)$ and $F_2(z)$ be the empirical distribution functions of the samples $X_1, X_2, \ldots, X_M$ and $Y_1, Y_2, \ldots, Y_N$ respectively. Thus

$$
NF_1(z) = \text{Number of observations } X_i \text{ such that } X_i \leq z \quad (6)
$$

$$
NF_2(z) = \text{Number of observations } Y_j \text{ such that } Y_j \leq z \quad (7)
$$

$$
i = 1, 2, \ldots, M, \quad j = 1, 2, \ldots, N.
$$
The one-sided Kolmogorov-Smirnov tests reject $H_0$: $F(x) = G(y)$ for large values of

$$D_x = \max_{1 \leq i \leq M} \left| \frac{F_N(z_i) - G_N(z_i)}{N} \right| \quad (8)$$

$$D_y = \max_{1 \leq j \leq N} \left| \frac{G_N(z_j) - F_M(z_j)}{M} \right| \quad (9)$$

The two-sided test rejects $H_0$: $F(x) = G(y)$ for large values

$$D = \max(D_x, D_y) \quad (10)$$

A convenient way to compute $D_x$, $D_y$ and $D$ is as follows:

$$D_x = \max_{1 \leq i \leq M} \left| \frac{F_N(z_i) - G_N(z_i)}{N} \right|$$

$$D_y = \max_{1 \leq j \leq N} \left| \frac{G_N(z_j) - F_M(z_j)}{M} \right| \quad (11)$$

The percentile points or the distributions of $D$ for the one-sided and two-sided tests are tabulated in any standard text on nonparametric statistics.

**Pair Chart**

Quade (1973) has shown how to get an insight into the two-sample comparison problem by drawing a "pair chart". He has also indicated how to obtain Mann-Whitney, Wilcoxon and Kolmogorov-Smirnov statistics from the pair chart.

The steps involved in drawing a pair chart are as follows:

**Step 1:** Combine the two samples and arrange the observations in ascending order of magnitude keeping track of identity of the variable (X or Y);

**Step 2:** Draw a rectangle of width $M$ units and height $N$ units;

**Step 3:** Starting from the lower left corner of the rectangle (i.e., the origin), draw a line one unit to the right (or upwards) if the smallest observation in the combined sample (as obtained in Step 1) is an X (or a Y);

**Step 4:** Starting from the end of the line drawn in Step 3, draw another line to the right (or upwards) if the second smallest observation is an X (or a Y).

Continue drawing such lines until the line for the largest observation has been drawn.

**SAS Macro PCHART**

The SAS macro PCHART enables user to get the pair chart plotted using CALCOMP plotter (or TERKTRONIX) besides giving the Mann-Whitney, Wilcoxon, and Kolmogorov-Smirnov statistics. PCHART uses PROC PLOTTER (PROC PLOTTEK for Tektronix video display) of SAS. The outputs given by PCHART are:

1. the observations in the two samples given as 'X-values' and 'Y-values';
2. the ranking of observations in the combined sample;
3. Mann-Whitney statistics $U_x$ and $U_y$;
4. Wilcoxon statistics $T_x$ and $T_y$;
5. Kolmogorov-Smirnov statistic $D$ (two-sided); and
6. the pair chart from CALCOMP plotter.

(See APPENDIX for job set up, examples, and macro listing.)

**References**


**APPENDIX**

**I. JOB SET UP**

The following JCL and SAS program are needed to execute the macro:

```sas
/* XXXXXX JOB ....................
 // *PL=XXX...
 // EXEC SAS
 //SYSIN DD *

 ....... Deck containing the macro PCHART........
 MACRO AREA XMIN=0 XMAX=nn YMIN=0 YMAX=mm
 XLEN=nn YLEN=mm N=11 XINC=1 YINC=1 %
 * WHERE nn=NO. OF X-OBSNS., mm=NO. OF Y-OBSNS.;
 DATA XDATA;
 INPUT X @@;
 CARDS;
 DATA YDATA;
 INPUT Y @@;
 CARDS;
 PCHART;
 CALPLOT; *FOR PLOTTING WITH PEN PLOTTER;
 OR;
 TEKPLOT; *FOR DISPLAY ON THE SCREEN OF
// TERKTRONIX;

II. EXAMPLE**

We have considered here one of the examples (involving ties) illustrated by Quade (1973).
### Statistical Analysis System

<table>
<thead>
<tr>
<th>XP</th>
<th>COL1</th>
<th>COL2</th>
<th>COL3</th>
<th>COL4</th>
</tr>
</thead>
<tbody>
<tr>
<td>COL5</td>
<td>COL6</td>
<td>COL7</td>
<td>COL8</td>
<td></td>
</tr>
<tr>
<td>X-VALUES</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>XP</th>
<th>COL1</th>
<th>COL2</th>
<th>COL3</th>
<th>COL4</th>
</tr>
</thead>
<tbody>
<tr>
<td>COL5</td>
<td>COL6</td>
<td>COL7</td>
<td>COL8</td>
<td></td>
</tr>
<tr>
<td>Y-VALUES</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

#### Ordered X Values

<table>
<thead>
<tr>
<th>XA</th>
<th>COL1</th>
<th>COL2</th>
<th>COL3</th>
<th>COL4</th>
</tr>
</thead>
<tbody>
<tr>
<td>COL5</td>
<td>COL6</td>
<td>COL7</td>
<td>COL8</td>
<td></td>
</tr>
<tr>
<td>ROW1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>7</td>
</tr>
</tbody>
</table>

#### Ordered Y Values

<table>
<thead>
<tr>
<th>YA</th>
<th>COL1</th>
<th>COL2</th>
<th>COL3</th>
<th>COL4</th>
</tr>
</thead>
<tbody>
<tr>
<td>COL5</td>
<td>COL6</td>
<td>COL7</td>
<td>COL8</td>
<td></td>
</tr>
<tr>
<td>ROW1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>9</td>
</tr>
</tbody>
</table>

#### Ordered X and Y Values: 1 = X-Value, 2 = Y-Value

<table>
<thead>
<tr>
<th>COMB</th>
<th>COL1</th>
<th>COL2</th>
<th>COL3</th>
<th>COL4</th>
<th>COL5</th>
<th>COL6</th>
</tr>
</thead>
<tbody>
<tr>
<td>COL7</td>
<td>COL8</td>
<td>COL9</td>
<td>COL10</td>
<td>COL11</td>
<td>COL12</td>
<td>COL13</td>
</tr>
<tr>
<td>VALUES</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDENTITY</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Kolmogorov-Smirnov Statistic

(TWO SIDED)

<table>
<thead>
<tr>
<th>MAT</th>
<th>* X_OBS</th>
<th>* Y_OBS</th>
<th>D_STAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROW1</td>
<td>8</td>
<td>8</td>
<td>0.375</td>
</tr>
</tbody>
</table>
### MANN-WHITNEY STATISTICS

<table>
<thead>
<tr>
<th>UX</th>
<th>COL1</th>
<th>ROW1</th>
<th>30</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>UY</th>
<th>COL1</th>
<th>ROW1</th>
<th>34</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>N1</th>
<th>COL1</th>
<th>ROW1</th>
<th>9</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>N2</th>
<th>COL1</th>
<th>ROW1</th>
<th>8</th>
</tr>
</thead>
</table>

### WILCOXON STATISTIC

<table>
<thead>
<tr>
<th>TX</th>
<th>COL1</th>
<th>ROW1</th>
<th>66</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>N1</th>
<th>COL1</th>
<th>ROW1</th>
<th>8</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>N2</th>
<th>COL1</th>
<th>ROW1</th>
<th>8</th>
</tr>
</thead>
</table>

#### PAIR CHART

![Pair Chart Image]

---

437
MACRO PCHART

PROC MATRX;
FETCH X DATA=XDATA;
FETCH Y DATA=YDATA;
XP=X';YP=Y';
FNAME='X-VALUES';FNAME2='Y-VALUES';
PRINT XP FONAME=FNAME1;
PRINT YP DONAME=FNAME2;
FNAME XP YP DONAME FNAME2;
NX=NCOW(X); NY=NCOW(Y); UX=J(NX,1,2); DO I=1 TO NX; TT=0; TTT=0; DO N=1 TO NY; IF YY(M,1)<XX(I,1) THEN DO; TT=T+1; END;
IF YY(M,1)>XX(I,1) THEN DO;
TTT=TTT+1; END;
UX(X,1)=T+1/4;END;
UY=MXXM(UX,1); UX=MXX2(UX,1);TX=UX(M((N141))#2/2; ID="; INCY=C;
DO I=1 TO N1;
FY=";IY=1;DO K=2 TO N1;
IF XX(I,1)=XX(K,1) THEN DO;
KX=KX+1;END;END;
FY=X;IX=1;
IF YY(I,1)=YY(K,1) THEN DO;
KY=KY+1;END;END;
XI=X+(KX+1)#/2;YI=Y+(KY+1)#/2;NXX=KX+1;
NYY=KY+1;
DO I=1 TO NXX;
DO I=1 TO NYY;
IF XX(I,1)=XX(KX,1) THEN DO;
KXX=KXX+1;END;END;
IF YY(I,1)=YY(KY,1) THEN DO;
KYY=KYY+1;END;END;
IF YY(I,1)>XX(KX,1) THEN DO;
KXX=KXX+1;END;END;
IF XX(I,1)>YY(KY,1) THEN DO;
KYY=KYY+1;END;END;
XX(1)=XXX(1);YY(1)=YYY(1);DEW=XXX(1)+1;";END;END;END;
FREE XX YY;
PRINT XX YY;
PRINT XY;
FREE XY;
FREE XA;
FREE YA;
FREE IA;
FREE J Ax Y; FREE K X Y; FREE L X Y; FREE M X Y; FREE N X Y; FREE O X Y; FREE P X Y; FREE Q X Y; FREE R X Y; FREE S X Y; FREE T X Y; FREE U X Y; FREE V X Y; FREE W X Y; FREE X Y; FREE Y X; FREE X Y; NOTE SKIP=1 ORDERED X VALUES;
PRINT XA;
FREE XA;
NOTE SKIP=1 ORDERED Y VALUES;
PRINT YA;
FREE YA;
NOTE SKIP=2 ORDERED X AND Y VALUES:
1 = X-VALUE, 2 = Y-VALUE;
N = 3fIP;
TFISP = J (N, 3, 0);
K = 0;
DO I = 1 TO IP;
X = DEL (I, 1); Y = DEL (I, 2); NX = DEL (I, 3); NY = DEL (I, 4);
Y = I + 1;
K = K + 1; KK = K + 1; KKK = KKK + 1;
C1 = K - NY/2; D1 = Y - NY/2;
C2 = X; D2 = Y;
C3 = K + NY/2; D3 = Y + NY/2;
Y = K + 1;
TISP (K, 1: 3) = C1 | D1 | SYM;
TISP (KK, 1: 3) = C2 | D2 | SYM;
TISP (KKK, 1: 3) = C3 | D3 | SYM;
K = KKK;
END;
PL0'T = PLOT /TX;
END;
IND = (IP + 2);
DP = (0 | 1 | (IND) // N1 | N2 | IND);
PLOT = PLOT // DP;
FREE DEL TISP DP C1 C2 C3 D1 D2 D3;
D1 = PLOT (*, 1) // N1; D2 = PLOT (*, 2) // N2;
NOTE PAGE * SKIP = 1 MAX = WHITNEY STATISTICS;
PRINT UX UX N1 N2;
FREE UX UX;
NOTE SKIP = 1 WILCOXON STATISTI;
PEN TX N1 N2;
FREE TX;
DIFF = ABS (D1 - D2);
MAX = DIFF;
MAT = N1 | N2 | D;
CHR = 'X' X OBS 'D_STAT';
NOTE PAGE * SKIP = 1 KOLMOGOROV-SMIRNOV STATISTI;
NOTE * SKIP = 1 (TWO SIDED);
PRINT MAT COLNAME = CNAME;
N22 = 1 * N2;
W1 = (W22 / N22) # (1: N22);
W1 = W1 * 1 (1, N1); W1 = SHAPE (W1, 1);
W2 = (1: N1) * 1 (1, N22); W2 = SHAPE (W2, 1);
PTG = N1 | N22;
S1 = J (N2, 1, 1);
PTG1 = W2 | W1 | S1;
FREE W1 W2 S1;
S11 = 1 * N1;
W3 = (W1 | W1) # (1: N11);
W3 = W3 * 1 (1, N2); W3 = SHAPE (W3, 1);
W4 = (1: N2) * 1 (1, N11); W4 = SHAPE (W4, 1);
PTG = N11 |
S2 = J (N11, 1, C);
PTG2 = W3 | W4 | S2;
FREE W3 W4 S2;
PTG1 = PTG1 / PTG2;
FREE PTG1 PTG2;
PTG = PLOT | // PLOT;
FREE PLOT;
OUTPUT PLOT OUT = PLOT (FILENAME = (C11 = X COL2 = Y COL3 = SYM));

MACRO CALPLOT
PROC FLOTTAG NIP TITLE = 'FAV CHART ' DEFSIZE = .025 AREA ;
GRAPH X Y; ID SYM; FORMAT X Y 3.5;

MACRO TBXPLOT
PROC FLOTTAG NIP TITLE = 'FAV CHART ' DEFSIZE = .025 AREA ;
GRAPH X Y; ID SYM; FORMAT X Y 3.5;

439
<table>
<thead>
<tr>
<th>INDEX #</th>
<th>NAME: MACRO KM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DESCRIPTION: THIS MACRO HAS ONE-HALF PAGE OF CODE FOR KAPLAN-MEIR ESTIMATION OF SURVIVAL DISTRIBUTIONS</td>
</tr>
<tr>
<td></td>
<td>REQUIREMENTS:</td>
</tr>
<tr>
<td></td>
<td>REFERENCE:</td>
</tr>
<tr>
<td></td>
<td>CONTACT: HARRELL, FRANK</td>
</tr>
<tr>
<td></td>
<td>ADDRESS: UNIVERSITY NORTH CAROLINA CHAPEL HILL NC 27514</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>INDEX #</th>
<th>NAME: FUNCTION PUTBIT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DESCRIPTION: ALLOWS USERS TO STORE ANY BIT STRING SUBSET OF A SAS VARIABLE</td>
</tr>
<tr>
<td></td>
<td>REQUIREMENTS: 8 K</td>
</tr>
<tr>
<td></td>
<td>REFERENCE:</td>
</tr>
<tr>
<td></td>
<td>CONTACT: HARRELL, FRANK</td>
</tr>
<tr>
<td></td>
<td>ADDRESS: UNIVERSITY NORTH CAROLINA CHAPEL HILL NC 27514</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>INDEX #</th>
<th>NAME: PROC PLOTTER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DESCRIPTION: LINE/PLT PLOTTING FOR CALCOMP PLOTTER</td>
</tr>
<tr>
<td></td>
<td>REQUIREMENTS: LOCALLY WRITTEN BASIC SYMBOL AND LINE PLOTTING SUBROUTIN AND CALCOMP PLOTTER WITH 110K OF CORE.</td>
</tr>
<tr>
<td></td>
<td>REFERENCE:</td>
</tr>
<tr>
<td></td>
<td>CONTACT: HARRELL, FRANK</td>
</tr>
<tr>
<td></td>
<td>ADDRESS: UNIVERSITY NORTH CAROLINA CHAPEL HILL NC 27514</td>
</tr>
</tbody>
</table>

440
INDEX # 4  NAME: MACRO TABULAR

DESCRIPTION: USES TABULAR METHOD OF ANALYSIS TO COMPUTE USER SPECIFIED SINGLE DEGREE OF FREEDOM LINEAR CONTRASTS ON TREATMENT MEANS FROM BALANCED FACTORIAL EXPERIMENTS. EASIER AND MORE CONVENIENT THAN GLM.

REQUIREMENTS: 85 CARD IMAGES

REFERENCE:

CONTACT: CARMER, SAMUEL
E: AGRONOMY DEPT.
ADDRESS: UNIVERSITY ILLINOIS
               URBANA
               IL 61801

INDEX # 5  NAME: MACRO RIDGREGR

DESCRIPTION: CALCULATES COEFF. OF RIDGE REGRESS. AND ALLOWS SELECTION APPROPRIATE COEFF. VALUES. USER SPECIFIES DEP. VARIABLE & RANGE OF K (BIAS) VALUES FOR ANALYSIS. OUTPUTS MEANS, VARIANCES, CORR. MATRIX, VAR. INFLATION FACTORS & OTHER STAT. WITH RIDGE TRACE PLOT.

REQUIREMENTS: USES SAS 76.5 PROC'S FORMAT, MATRIX, AND PLOT

REFERENCE:

CONTACT: ROGERS, ROBERT HILDEBRAND, R.
E: USDA FOREST SERVICE
ADDRESS: UNIVERSITY MISSOURI
               COLUMBIA
               MO 65201

INDEX # 6  NAME: PROC AOYMEAN

DESCRIPTION: PERFORMS A ONE-WAY AOY USING GROUP SIZES, MEANS AND STANDARD DEVIATIONS AS INPUT. T-TESTS FOR 3 TYPES OF COMPARISONS: 1) GROUPS WITH THE FIRST GROUP, ALL POSSIBLE PAIRS OF GROUPS, AND USER SUPPLIED CONTRASTS

REQUIREMENTS: 37.8K

REFERENCE:

CONTACT: TESAR, T. P.
E: UPJOHN COMPANY
ADDRESS: 7293-32-1
               KALAMAZOO
               MI 49001
INDEX # 7
NAME: PROC CCPL0T

DESCRIPTION: PRODUCES A CALCOMP PLOT WITH AS MANY AS FIVE FUNCTIONS (ONE SET OF AXES). EACH FUNCTION IS DEFINED BY A PAIR OF VARIABLES.
USER CONTROLS PLOT SIZE, SYMBOLS FOR EACH FUNCTION, AXIS LABELS, LEGENDS, AND SCALPS.

REQUIREMENTS: 48.4K

REFERENCE:

CONTACT: TESAR, T. P.
E: UPJOHN COMPANY
ADDRESS: 7293-32-1
       KALAMAZOO
       MI 49001

INDEX # 8
NAME: MACRO HIST

DESCRIPTION: PROVIDES FOR VERT. & HOR. HISTOGRAMS WITH SINGLE AXIS LABELING AND TITLE INFORMATION. USES PUT STATEMENTS TO FORMAT THE GRAPHS. FREQUENCIES ARE ALSO PLOTTED.

REQUIREMENTS: 1K

REFERENCE:

CONTACT: THARP, M. L. STRAND, R. H.
E: ENVIRONMENTAL SCI. DIV.
ADDRESS: P.O. BOX X, BLDG. 1505, ORNL
       OAK RIDGE
       TN 37830

INDEX # 9
NAME: PROC DISPLAY

DESCRIPTION: PLOTS 2-DIMENSIONAL DATA USING SAS ON A CALCOMP PLOTTER. LINEAR OR LOG. SCALES, UP TO 500 PTS., AXES LABELS, TITLING, AND UP LINES PER GRAPH ARE AVAILABLE.

REQUIREMENTS: DISSPLA SOFTWARE & CALCOMP PLOTTER 4K

REFERENCE:

CONTACT: OLSON, R. J.
E: ENVIRONMENTAL SCI. DIV.
ADDRESS: P.O. BOX X, BLDG. 1505, ORNL
       OAK RIDGE
       TN 37830
INDEX # 10  NAME: PROC WILCOX

DESCRIPTION: DISTRIBUTION-FREE ESTIMATES OF THE RATIO OF TWO RANDOM VARIABLES. CONFIDENCE INTERVALS ARE ALSO ESTIMATED. PRODUCES WILCOXON T-STATISTIC FOR PAIRED SAMPLES.

REQUIREMENTS: 4K

REFERENCE:

CONTACT: KUMAR, DEVA
       & : ENVIRONMENTAL SCI. DIV.
ADDRESS: P.O. BOX X, BLDG. 1505, ORNL
        OAK RIDGE, TN 37830

INDEX # 11  NAME: FUNCTION GETBIT

DESCRIPTION: ALLOWS USERS TO RETRIEVE ANY BIT STRING SUBSET OF A SAS VARIABLE

REQUIREMENTS: 4K

REFERENCE:

CONTACT: HARRELL, FRANK
       & : BIOSTATISTICS DEPT.
ADDRESS: UNIVERSITY NORTH CAROLINA
        CHAPEL HILL, NC 27514

INDEX # 12  NAME: MACRO RECODE

DESCRIPTION: RECODES A SPECIFIED VALUE OF INDICATED VARIABLES ON ALL OBSERVATIONS ON A SAS DATA SET TO A SECOND SPECIFIED VALUE (12 CARD IMAGES)

REQUIREMENTS:

REFERENCE:

CONTACT: HENDERSON, DON
       & : DATA SYSTEMS APPLICATION DIV.
ADDRESS: ARS, NATIONAL AGRIC. LIB. BLDG.
        BELTSVILLE, MD 20705
INDEX # 13

NAME: MACRO KSISAMP

DESCRIPTION: KOLMOGOROV-SMIRNOV ONE SAMPLE TEST. OUTPUTS SAMPLE SIZE, TEST STATISTIC DSDP, AND, WHEN N GE 30, SELECTED ASYMPTOTIC CRITICAL VALUES FOR DETERMINING P-VALUE.

REQUIREMENTS:

REFERENCE: SUGI 77 PROCEEDINGS

CONTACT: GJERTSEN, W. R. HARRELL, P. E.
        E: SAS INSTITUTE, INC.
        ADDRESS: P.O. BOX 10066
        RALEIGH, NC 27605

INDEX # 14

NAME: MACRO KSISAMP

DESCRIPTION: KOLMOGOROV-SMIRNOV TWO SAMPLE TEST. OUTPUTS SAMPLE SIZE N1 AND N2, TEST STATISTIC DSDP, AND, WHEN N1 AND N2 GE 30, SELECTED ASYMPTOTIC CRITICAL VALUES FOR DETERMINING P-VALUE.

REQUIREMENTS:

REFERENCE: SUGI 77 PROCEEDINGS

CONTACT: GJERTSEN, W. R. HARRELL, P. E.
        E: SAS INSTITUTE, INC.
        ADDRESS: P.O. BOX 10066
        RALEIGH, NC 27605

INDEX # 15

NAME: MACRO DATALOG

DESCRIPTION: A SERIES OF MACROS FOR UNIVARIATE DESCRIPTIVE STATISTICS PROVIDING: A) LISTING OF MOMENT AND PERCENTILE INFORMATION, B) OUTLIER (FAR-OUT VALUES) REPORT, AND C) ONE OR MORE PLOTS. OUTPUT IS FLEXIBLY CONTROLLED BY THE USER.

REQUIREMENTS:

REFERENCE: SUGI 77 PROCEEDINGS

CONTACT: GJERTSEN, W. R.
        E: SAS INSTITUTE, INC.
        ADDRESS: P.O. BOX 10066
        RALEIGH, NC 27605
### SAS Macro INDEX BY KEYWORD

<table>
<thead>
<tr>
<th>INDEX #</th>
<th>KEYWORD</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>ANOVA</td>
</tr>
<tr>
<td>5</td>
<td>BIASED</td>
</tr>
<tr>
<td>2</td>
<td>BIT</td>
</tr>
<tr>
<td>11</td>
<td>BIT</td>
</tr>
<tr>
<td>3</td>
<td>CALCOMP</td>
</tr>
<tr>
<td>7</td>
<td>CALCOMP</td>
</tr>
<tr>
<td>4</td>
<td>CONTRASTS</td>
</tr>
<tr>
<td>10</td>
<td>DISTRIBUTION</td>
</tr>
<tr>
<td>1</td>
<td>DISTRIBUTIONS</td>
</tr>
<tr>
<td>1</td>
<td>ESTIMATION</td>
</tr>
<tr>
<td>5</td>
<td>ESTIMATION</td>
</tr>
<tr>
<td>10</td>
<td>ESTIMATION</td>
</tr>
<tr>
<td>4</td>
<td>FACTORIAL</td>
</tr>
<tr>
<td>9</td>
<td>GRAPHICS</td>
</tr>
<tr>
<td>8</td>
<td>GRAPHS</td>
</tr>
<tr>
<td>8</td>
<td>HISTOGRAMS</td>
</tr>
<tr>
<td>1</td>
<td>KAPLAN-MEIR</td>
</tr>
<tr>
<td>13</td>
<td>KOLMOGOROV-SMIRNOV</td>
</tr>
<tr>
<td>14</td>
<td>KOLMOGOROV-SMIRNOV</td>
</tr>
<tr>
<td>2</td>
<td>MANIPULATION</td>
</tr>
<tr>
<td>11</td>
<td>MANIPULATION</td>
</tr>
<tr>
<td>15</td>
<td>MOMENTS</td>
</tr>
<tr>
<td>15</td>
<td>OUTLIERS</td>
</tr>
<tr>
<td>7</td>
<td>PLOT</td>
</tr>
<tr>
<td>15</td>
<td>PLOT</td>
</tr>
<tr>
<td>3</td>
<td>PLOTTER</td>
</tr>
<tr>
<td>3</td>
<td>PLOTTING</td>
</tr>
<tr>
<td>9</td>
<td>PLOTTING</td>
</tr>
<tr>
<td>10</td>
<td>RANDOM</td>
</tr>
<tr>
<td>12</td>
<td>RECODING</td>
</tr>
<tr>
<td>5</td>
<td>REGRESSION</td>
</tr>
<tr>
<td>5</td>
<td>RIDGE</td>
</tr>
<tr>
<td>10</td>
<td>SAMPLING</td>
</tr>
<tr>
<td>2</td>
<td>SUBSTRING</td>
</tr>
<tr>
<td>11</td>
<td>SUBSTRING</td>
</tr>
<tr>
<td>1</td>
<td>SURVIVAL</td>
</tr>
<tr>
<td>4</td>
<td>TABULAR</td>
</tr>
<tr>
<td>12</td>
<td>TRANSFORMATION</td>
</tr>
<tr>
<td>15</td>
<td>UNIVARIATE</td>
</tr>
</tbody>
</table>