DECUS
PROGRAM LIBRARY

DECUS NO.	8-350
TITLE	WILCOXON-WHITE TWO SAMPLE RANK TEST
AUTHOR	Jens G. Rosenkrantz, M. D.
COMPANY	Childrens Hospital of Los Angeles Los Angeles, California
DATE	July 20, 1970
SOURCE LANGUAGE	8K FORTRAN

WILCOXON-WHITE TWO SAMPLE RANK TEST

DECUS Program Library Write-up DECUS NO. 8-350

In the statistical analysis of two sets of data to discover differences, the Wilcoxon-White[1,2,3] two sample rank test is quite useful, particularly when there is considerable variation in the measurements. This nonparametric statistical test may be used in comparing unpaired samples and assigns ranks to the pooled measurements, comparing the ranks as ordinal numbers in the two groups. The null hypothesis is that the ranks are randomly distributed between each of the two sets. The smaller of the ranks, T, is calculated and referred to a table for testing of the null hypothesis. Tables are available for unequal group sizes, as well as equal sizes. For values of group size, n_1, and n_2, which lie outside the limits of available tables, the approximate normal deviate, Z is calculated according to the formula:

$$Z = (|m - T| - 1/2)/s$$

where $m = n_1 (n_1 + n_2 + 1)/2$; $s = \sqrt{n_2 m/6}$

n_1 = number of observations in smaller group

n_2 = number of observations in larger group

T = smaller sum of ranks (not necessarily corresponding to n_1).

The calculated value of Z can be referred to tables of normal distribution for the significance probability, P.

From the Department of Surgery, Childrens Hospital of Los Angeles, 4650 Sunset Boulevard, Los Angeles, California 90027. Supported in part by USPHS Grant # HE 12973

DESCRIPTION OF PROGRAM

The program accepts data from each of two groups on the high speed reader. Up to 50 data points may be used for each group. Data acceptable range from -999.999 to 9999.999. Input format is as follows: Number of observations in first group (in I format) CRLF * Ist datum (in F format) CRLF*, and so forth for the first group, then repeat for the second group. To input subsequent pairs of groups, repeat the format above. When finished with the data tape, let number of observations of first group = 100 (or any number above 50). The program runs more quickly if the first group presented is the smaller one.

The program reads the data for the first group, and types back the data on the teletype. It calculates the mean and variance for the first group and then prints them. The data in the second group are then read, analyzed and reported similarly.

The program then ranks the numbers and prints the sum of the ranks of the first group. It considers ties and ranks negative numbers below positive ones. From this, the program calculates the T value by the Wilcoxon-White test and prints this, along with the numbers in the two groups. The program also calculates and prints the corresponding Z value in case the numbers exceed available tables.

In the "If" statements before line 68 and below line 70, there is a statement transferring control to line: 80 B = B+C, prior to assignment of a value to C. Note however that the conditions leading to this transfer cannot occur, since the data have already been arranged in ascending order by the subroutine, Ranks.

* Carriage - Return - Line - Feed

LOADING AND RUNNING THE PROGRAM

1. There are two relocatable binary tapes, the main program and a subroutine, ranks. These are loaded with Linking Loader (DEC # 08-A2 B3), the main program into Field O and the subroutine into Field 1.

2. The program also requires Tape 1 and sections 1-3 of Tape 2 of the Fortran 8K Library.

   Sections 1-2 of Tape 1 is loaded into Field Ø and the remainder of Tape 1 and sections 1-3 of Tape 2 into Field 1.

3. With this loading scheme the rank analysis program begins at starting address 1102 in Field O.

SUMMARY

The program performs the Wilcoxon-White two sample rank test on unpaired observations from two groups, which need not be of equal size. Ties are considered in assignment of ranks. Input is from the DEC high-speed reader; output on the teletype. The program calculates mean and standard deviation for each group, the sum of ranks for the first group, the smaller sum of ranks and the approximate normal deviate. It is written in Fortran 8K.

---

1. Wilcoxon, F.: Biometrics Bul. 1: 80, 1945.

2. White, C.: Biometrics 8:33, 1952.

3. Mann, H.B. and Whitney, D.R.: Ann. Math. Statistics 18: 50, 1947.

```
C       WILCOXON-WHITE TWO SAMPLE RANK TEST.
C       ACCEPTS UP TO 50 ENTRIES IN EACH OF 2 GROUPS.
C       FROM -999 TO +9999 WITH 3 DECIMEL PLACES.
C       HIGH SPEED READER.
C       USES FLOAT, ABS AND SORT LIBRARY SUBROUTINES AND
C       1 SUBPROGRAM, RANKS.
        COMMON A,N
        DIMENSION A(50,2),N(2)
        WRITE (1,2)
2       FORMAT ('RANK TEST')
9       WRITE (1,10)
10      FORMAT (///5X,'DATA ARE:')
        CALL RANKS
        B=0
        I=1
        M=1
52      IF (A(I,1)-A(M,2)) 60,65,54
54      IF (M-N(2))      56,58,58
56      M=M+1
        GOTO 52
58      JD=I+M
        GOTO 62
60      JD=I+M-1
62      C=FLOAT(JD)
        I=I+1
        GOTO 80
65      N3=N(1)-I
        N4=N(2)-M
        DO 68 K=1,N3
        IF (A(I,1)-A(I+K,1))     70,68,80
68      CONTINUE
        K=N3+1
70      DO 74 L=1,N4
        IF (A(M,2)-A(M+L,2))     76,74,80
74      CONTINUE
        L=N4+1
76      NU=K*(2*I+2*M+K+L-3)
        XN=FLOAT(NU)
        C=XN/2.
        I=I+K
        IF (M+L-N(2)) 78,78,79
78      M=M+L
        GOTO 80
79      M=M+L-1
80      B=B+C
        IF (I-N(1)) 52,52,82
```

```
 82       IF (N(1)-N(2))  84,84,86
 84       Z1=FLOAT(N(1))
          Z2=FLOAT(N(2))
          T1=B
          GOTO 88
 86       Z1=FLOAT(N(2))
          Z2=FLOAT(N(1))
          T1=(Z1+Z2)*(Z1+Z2+1.)/2.-B
 88       T2=Z1*(Z1+Z2+1.)-T1
          IF (T1-T2) 9 ,94,92
 90       T=T1
          GOTO 94
 92       T=T2
 94       XM=Z1*(Z1+Z2+1.)/2.
          F=Z2*XM/6.
          WRITE (1,96)
 96       FORMAT (/5X,'RESULTS ARE:')
          G=SQRT(F)
          Y=T-XM
          YZ=ABS(Y)
          Z=(YZ-.5)/G
          WRITE (1,100) B
 100      FORMAT(/19HSUM RANKS GROUP #1:F6.1)
          WRITE(1,102) T,Z1,Z2
 102      FORMAT(/2HT=F6.1/3HN1=F4.0/3HN2=F4.0)
          WRITE (1,104)
 104      FORMAT(5X,'IF NUMBERS ABOVE EXCEED TABLE,USE Z BELOW')
          WRITE (1,106) Z
 106      FORMAT(2HZ=F .3)
          GOTO 9
          END
          SUBROUTINE RANKS
 C        READS IN DATA.
          COMMON A,N
          DIMENSION A(5 ,2),N(2)
          DO 49 J=1,2
          READ (2,12) N(J)
 12       FORMAT (I3)
          N1=N(J)
          IF (5 -N1) 5 ,13,13
 13       S=0
          SS=0
          DO 18 I=1,N1
          READ (2,14) A(I,J)
 14       FORMAT (F8.3)
          S=S+A(I,J)
 18       SS=SS+A(I,J)**2
          K=N1+1
          L=N1+4
          DO 20 I=K,L
```

```
20      A(I,J)=111111.11111
        ZN=FLOAT(N(J))
        X=S/ZN
        S1=(SS-S**2/ZN)/(ZN-1.)
        G=SQRT(S1)
        IND=1
        WRITE (1,21) J
21      FORMAT (/5HGROUP #:I1)
22      DO 24 I=1,N1,5
23      FORMAT (5(2X,F8.3))
24      WRITE(1,23)A(I,J),A(I+1,J),A(I+2,J),A(I+3,J),A(I+4,J)
        IF (IND) 49,49,26
26      IND=IND-1
        WRITE (1,28) X,G,N1
28      FORMAT(/5X,5HMEAN=F10.3/5X,7HST DEV=F8.3/5X,2HN=I9/)
        DO 49 I=2,N1
        M=I-1
33      IF (A(M,J)-A(M+1,J)) 42,42,40
40      TEM=A(M,J)
        A(M,J)=A(M+1,J)
        A(M+1,J)=TEM
42      M=M-1
        IF (M) 49,49,33
49      CONTINUE
        RETURN
50      WRITE (1,51)
51      FORMAT ('FINISHED')
        STOP
        END

*
```

MAIN	01142 ←
OPEN	13525
WRITE	03352
IOH	05142
RANKS	11053
FLOT	07153
STO	06444
SUBSC	14230
IFAD	07116
CHS	07211
FAD	06410
CLEAR	07227
FLOAT	07034
MPY	13000
FDV	06711
FMP	06623
FSB	06000
SQRT	15411
ABS	13236
READ	03271
FIPOW	15076
ISTO	07061
CKIO	13521
SETERR	13600
ERROR	13703
TTYOUT	13427
HSOUT	13455
TTYIN	13400
HSIN	13445
DIV	13045
IREM	13216
FIX	06510
IFIX	06556
IABS	13270
IRDSW	13313
EXIT	13542
CLRERR	13631
IIPOW	15060
IFPOW	15062
FFPOW	15250
EXP	14652
ALOG	14547
0061	
0010	

LOADING SCHEME:

MAIN PROGRAM - FIELD 0

SUBROUTINE RANKS - FIELD 1

SECT 1-2, TAPE 1 LIBRARY - FIELD 0

SECT 3-5, TAPE 1
SECT 1-3, TAPE 2 LIBRARY } - FIELD 1