A SAS program is presented for performing the distribution-free test for ordered alternatives or dose-response trend by A.R. Jonckheere. This program computes Jonckheere's statistics for small sample as well as large sample approximation; for large sample it also provides the p-value for hypothesis testing. This program is especially useful for large samples with unequal sample sizes in each treatment group and with missing values. The number of treatment groups are flexible in this program and examples are given.

**Introduction**

In a dose-response study, an experimenter often wishes to see that the magnitude of response increases or decreases with the quantity of the dosage that is applied. For example, suppose that an experiment which consists of four treatment groups: placebo (0 mg), 10 mg, 50 mg and 75 mg of an experimental drug, is performed to determine the effect of the different dosages upon the reduction of the attacks of a disease. The experimenter, based on his/her experience or theory, expects to see that the number of attacks of this disease decreases with the increasing of dosage of this experimental drug. That is, he/she wishes to test the hypothesis that the four samples have come from the same population against the alternative that the populations are in an order of decreasing value as the dosage increases. Hence a test of significance specifically sensitive to the differences in question is required.

The usual one-way analysis of variance does not satisfy this demand since the F-ratio is independent of the order in which the group means occur. In many experiments, for instance in psychology, the independent variable such as 'stress' is not suitably quantified, it is not possible to use any form of regression analysis. In this case a test procedure proposed by Terpstra [1] and independently by Jonckheere [2], but is known as Jonckheere's test in the literature, can be used.

To obtain Jonckheere's test statistic can be tedious with sample sizes larger than ten in more than two treatment groups. Hence a SAS macro is provided to calculate Jonckheere's test statistic for small sample size or for large sample approximation. The p-value for testing the statistical hypothesis is also calculated for the latter case. Unequal sample sizes or missing values are also properly handled by this macro.

**Assumptions**

1. The basic model is
   \[ X_{ij} = \mu + \tau_j + \epsilon_{ij}, \quad j = 1, \ldots, k, \]
   where \( \mu \) is the (unknown) overall mean, \( \tau_j \) is the (unknown) treatment \( j \) effect, and \( \epsilon_{ij} \) is (random errors) are mutually independent.

2. Each \( \epsilon \) comes from the same continuous population.

3. \( \mu \) and \( \tau_j \) are unknown parameters.

**Procedure**

To test \( H_0: \tau_1 = \ldots = \tau_k \) against

1. Compute \( k(k-1)/2 \) Mann-Whitney counts \( U_{lm} \), \( l < m \), where
   \[ U_{lm} = \sum_{i=1}^{n_l} \sum_{j=1}^{n_m} \phi(X_{ij}, X_{lj}) \]
   and
   \[ \phi(a, b) = \begin{cases} 1 & \text{if } a < b \\ 0 & \text{if } a = b \end{cases} \]

2. Compute \( J = \sum_{l=1}^{k-1} \sum_{m=l+1}^{k} U_{lm} \)

be the sum of these \( k(k-1)/2 \) Mann-Whitney counts.

3. At the \( \alpha \)-level of significance,
   - reject \( H_0 \) if \( J > j(n_1, \ldots, n_k) \),
   - accept \( H_0 \) if \( J < j(n_1, \ldots, n_k) \)

where the constant \( j(n_1, \ldots, n_k) \), which satisfies the equation

\[ P(J > j(n_1, \ldots, n_k)) = \alpha, \]

is obtained from the Table included in Hollander, M. and Wolfe, D.A. (1973).
LARGE SAMPLE APPROXIMATION. Set

\[
J^* = \sqrt{\frac{\text{var}(\hat{\theta})}{\hat{\theta}^2}} \left[ \frac{1}{N(N-1)2} \sum_{i=1}^{N} (x_{i1} - x_{i2})^2 \right]^{3/2}.
\]

When \( H_0 \) is true, the statistic \( J^* \) has an asymptotic \( \text{N}(0,1) \) distribution. The approximate \( \alpha \)-level test is

- reject \( H_0 \) if \( J^* \approx z(\alpha) \)
- accept \( H_0 \) if \( J^* \approx z(\alpha) \).

PROGRAM INPUT AND OUTPUT

This program consists of two parts. Part I is the program module named 'JONCKHR' which performs the comparisons between observations to obtain the Mann-Whitney counts. It is written for four treatment groups and can be used for any number of treatment groups. The instructions for using the program module are given at the beginning of the program "JONCKHR". Part II is the user's program in which the number of treatment groups can be specified. The user's program includes input data and performs large sample approximation. It is also written for four treatment groups or less. This program will automatically set count4 to zero if only three treatment groups are needed. Jonckheere's statistic for the exact test (small sample size) and for large sample approximation with its p-value will be output. The body of the program is presented in Appendix I.

Three examples are given to illustrate the application of this program and the advantage of using Jonckheere's test. Example 1 is a case when ANOVA failed to show a treatment difference, Jonckheere's test showed a significant dose-response trend. Example 2 shows how this program can be applied easily with any SAS dataset. Example 3 is an example of three treatment groups and data were input directly into the program; in addition, it shows that sometimes this program can be used to get a quick feeling of the result by large sample approximation.

Example 1. The values of four treatment groups are given as follows:

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>0 mg</th>
<th>30 mg</th>
<th>50 mg</th>
<th>75 mg</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>21</td>
<td>40</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>61</td>
<td>99</td>
<td>110</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>80</td>
<td>100</td>
<td>151</td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>129</td>
<td>149</td>
<td>160</td>
<td></td>
</tr>
</tbody>
</table>

Mean 57.25 72.75 97.00 117.50

The experimenter wishes to test the hypothesis that there is no difference in the mean values among the four treatment groups versus the alternative that the four treatment group means are in the same increasing order as the dosages. The F ratio obtained from the analysis of variance equals 1.216 with degrees of freedom of 3 and 12 and the p-value is 0.346. Therefore the experimenter would have accepted the null hypothesis with some confidence. Whereas the Jonckheere's statistic \( J \) equals 71, which gives a p-value of 0.0168. Hence the experimenter would reject the null hypothesis with some certainty. Using large sample approximation the p-value is 0.0159 which is very close to the exact test result.

This example not only showed that Jonckheere's test is sensitive to the differences among means, it is especially useful when the given data are not normally distributed and the ANOVA should not be used.

Example 2. Eighty-three patients were randomly assigned to four dosage regimen: placebo (0 mg), 10 mg, 30 mg, or 50 mg of a test drug and the reductions of frequency of pains from baseline were recorded. The Duncan's multiple range test did not show a clear-cut difference between any two treatment groups; however, from the treatment response means a dose-response trend is clearly visible. Jonckheere's statistic was 3.21 (from normal approximation) which gives a p-value of 0.0007. Therefore, one can conclude that increasing dosage of this test drug will further reduce the frequency of the pains. This example shows that a sample size as large as 83 or larger will not cause any difficulties in the calculation.

Example 3. Student scores (responses) to a particular subject among three schools are compared to see if the amount of preparation given in class before an examination will affect the student scores. School A had given no in-class preparation, School B had given a quick review on the subject, and School C had given a thorough in-class review before the examination. The response scores are listed as follows:

<table>
<thead>
<tr>
<th>SCHOOL</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>42.7</td>
<td>44.7</td>
<td>46.7</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>3.3</td>
<td>4.7</td>
<td>4.5</td>
</tr>
</tbody>
</table>

The ANOVA result showed no differences among the three schools with a p-value of 0.2931. The Jonckheere's statistic was calculated to be 74.5 with a p-value of 0.052, and the large sample approximation resulted a p-value of 0.049. One may conclude that the method
of preparation had a slight effect on the students' scores.

The computer input and output are presented in Appendix II.

REFERENCES

1. Terpstra, T.J. (1952). The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking, Indagationes Mathematicae 14, 327-333.


APPENDIX I

Part 1 - The Program Module:

```sas
* DESCRIPTION: PROGRAM MODULE COMPARES ALL TREATMENTS AND EVALUATES WHETHER EACH OBSERVATIONS FROM UP TO FOUR DRUGS (OR GROUPS) IS SMALLER, EQUAL, OR GREATER THAN ALL OTHERS IN SUBSEQUENT TREATMENTS. THE RESULT OF EXECUTING THIS MODULE IS A SAS DATASET NAMED "DONCKHR" CONTAINING THE STATISTIC J.

/ * INSTRUCTIONS: *
/ * 
/ * USER'S PROGRAM HERE
/ * 
/ * \ before this module is invoked
/ * 
/ * INCLUDE DOFILE( DONCKHR); *
/ * 
/ * USER'S PROGRAM CONTINUES
/ * 

* DETERMINE NDRUGS- NUMBER OF DRUGS (OR GROUPS)
* NDBS- NUMBER OF OBSERVATIONS / DRUG
* TOTN- TOTAL NUMBER OF DATA ELEMENT

data_analyze; set _last;
proc contents noprint data_analyze out=temp;
data_null; set temp end=EOF;
   if not (nobs = n);
   nobs = n;
   call symput('NDBS', trim(left(put(nobs,4.))));
   call symput('TOTN', trim(left(put(totn,4.))));
   output;
run;
* CONVERT USER'S DATASET TO AN ARRAY WITH TOTN ELEMENTS (CELLS). PROC TRANPOSE IS USED TO PLACE DATA RELATED TO A DRUG ON ONE OBSERVATION.
* proc transpose data=drugs1 out=drugs2;
* proc print data=drugs2;
title3 'Dataset: Drugs1 (transposed dataset)';
data drugs2; set drugs2 end=EOF;
   array drug{*} col1-colnobs;
   array drugs{*} el - @TOTN;
   keep el - @TOTN;
   retain el - @TOTN;
   call symput('TOTN', trim(left(put(totn,4.))));
   element1 = drugs[el] = drugs[nobs];
   do nobs = 1 to nobs;
      element1 = drugs[el] = drugs[nobs];
      if EOF then output;
   end;
   proc print data=drugs2;
title3 'Dataset: Drugs2 (final array)'; run;
```

Part 2 - The User's Program with JCL:

```sas
* PROC CONTENTS NOPRINT DATA=ANALYZE OUT=TEMP;
*   DATA_Null; SET TEMP END=EOF;
*   KEEP NDRUGS NORTN;
*   IF EOF THEN DO;
*      NDRUGS = N;
*      CALL SYMPUT('NDBS', TRIM(LEFT(PUT(NDRUGS,4.))));
*      CALL SYMPUT('TOTN', TRIM(LEFT(PUT(TOTN,4.))));
*      OUTPUT;
*   END;
*   RUN;
*   * CONVERT USER'S DATASET TO AN ARRAY WITH TOTN ELEMENTS (CELLS). PROC TRANPOSE IS USED TO PLACE DATA RELATED TO A DRUG ON ONE OBSERVATION.
*   PROC TRANPOSE DATA_ANALYZE OUT=DRUGS2;
*   PROC PRINT DATA=DRUGS2;
*   TITLE3 'Dataset: Drugs2 (transposed dataset)';
*```

```sas
* REFERENCES
1. Terpstra, T.J. (1952). The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking, Indagationes Mathematicae 14, 327-333.
```

```sas
* DESCRIPTION: PROGRAM MODULE \ EXAMPLE2
* AUTHOR: SUE HUANG DATE CREATED: 09/24/87
* DESCRIPTION: THIS PROGRAM WILL TEST DOSE RESPONSE TRENDS BASED ON A DISTRIBUTION-FREE TEST BY JONCKHEERE
* INCLUDE TESTDATA;
* PROC SORT DATA=ANALYZE;
*   BY DRUG;
* PROC MEANS DATA=ANALYZE CLASS=DRUG MODEL RESPONSE=DRUG MEANS=DUNCAN LSD;
* TITLE1 'ANALYSIS OF VARIANCE ON ';
* TITLE2 'RESPONSE CHANGE DATA';
* RUN;
*```
APPENDIX I - Part 2 (Continued)

---

**THE FOLLOWING IS JONCKHEERE'S TEST FOR DOSE-RESPONSE TREND BY JONCKHEERE'S DISTRIBUTION FREE TEST**

---

- INITIALIZE MACROVARIABLE COUNTERS
  
  \[ \text{LET COUNT1=0; LET COUNT2=0; LET COUNT3=0; LET COUNT4=0; \] 

- PREPARE DATA:
  - SPLIT DATA, ONE PER DATASET
  - ACCUMULATE TOTAL CLASSIFICATION

MACRO PARAMETERS:

- CLASSIFICATION N'S FOR EACH SOURCE DATASET
- OUTPUT DATASETS
- CLASSIFICATION VARIABLE NAME
- CLASSIFICATION VARIABLE VALUE
- RESPONSE VARIABLE NAME
- RESPONSE VARIABLE NAME
- LABEL LABEL TO ASSOCIATE WITH RESPONSE VARIABLE

- EXECUTE MACRO PREPARE(DSET1, OSET2, N, CLASS, VAR, COUNT, LABEL);

%GLOBAL COUNT1 COUNT2 COUNT3 COUNT4;

DATA &OSET2; SET &DSET1;
PROC SORT; BY &CLASS;
DATA &OSET2; SET &OSET2; BY &CLASS;
KEEP &CLASS &CLASS &VAR; 
IF &VAR NE THEN TALLY + 1;
IF LAST.&CLASS THEN 00;
CALL SYMPUT("&COUNT", TRIM(LEFT(PUT(TALLY,4.D));
END;
OUTPUT;
LABEL &VAR=&LABEL; TITLE3 ", LABEL;"
END;
RUN;
SNMEND PREPARE;

%MACRO GENERAL(DSET1, DSET2, N, CLASS, VAR, LABEL);
%PREPARE(DSET1, DSET2, N, CLASS, VAR, COUNT1, LABEL);
%PREPARE(DSET1, DSET2, N, CLASS, VAR, COUNT2, LABEL);
%PREPARE(DSET1, DSET2, N, CLASS, VAR, COUNT3, LABEL);
%PREPARE(DSET1, DSET2, N, CLASS, VAR, COUNT4, LABEL);

%GLOBAL CHANGE, DSET1, DSET2, N, CLASS, VAR, LABEL;
%PREPARE(DSET1, DSET1, N, CLASS, VAR, COUNT1, LABEL);
%PREPARE(DSET1, DSET2, N, CLASS, VAR, COUNT2, LABEL);
%PREPARE(DSET1, DSET3, N, CLASS, VAR, COUNT3, LABEL);
%PREPARE(DSET1, DSET4, N, CLASS, VAR, COUNT4, LABEL);

DATA &DSET1; SET &DSET1;
PROC PRINT; TITLE4 ", LABEL;"
END;
RUN;
SNMEND GENERAL;

---

%GLOBAL CHANGE, DSET1, DSET2, N, CLASS, VAR, LABEL;
%PREPARE(DSET1, DSET1, N, CLASS, VAR, COUNT1, LABEL);
%PREPARE(DSET1, DSET2, N, CLASS, VAR, COUNT2, LABEL);
%PREPARE(DSET1, DSET3, N, CLASS, VAR, COUNT3, LABEL);
%PREPARE(DSET1, DSET4, N, CLASS, VAR, COUNT4, LABEL);

DATA &DSET1; SET &DSET1;
PROC PRINT; TITLE4 ", LABEL;"
END;
RUN;
SNMEND GENERAL;

---

**RECOMBINE THE DATASETS USING MERGE**

DATA DSET1; MERGE DSET1 DSET2 DSET3 DSET4;
KEEP DSET1 DSET2;
PROC PRINT; TITLE4 ", LABEL;"
END;
RUN;

---

**SNMEND GENERAL**

---

**CODE FOR LARGE SAMPLE APPROXIMATIONS**

**(FOUR (4) GROUPS OR LESS)**

---

**DATA JONKHK; SET JONKHK;**

LENGTH N W X Y Z B;
RENAME N=COUNT1 W=COUNT2 X=COUNT3 Y=COUNT4;
N=COUNT1 W=COUNT2 X=COUNT3 Y=COUNT4;
J1 = J - (N*2 - (W*2 + X*2 + Y*2 + Z*2)) / 4;
J2 = (N*2 - (2*W*H)) - (N*2 - (2*W*H));
J3 = (N*2 - (2*W*H)) + (N*2 - (2*W*H));
J4 = (N*2 - (2*W*H)) + (N*2 - (2*W*H));
JSTAR = J1 / (J1 + J3) / J1 **.5;
F_VALUE = 1 - PROBNORM(JSTAR);
APPENDIX II - Example 2 (Continued)

ANALYSIS OF VARIANCE ON RESPONSE CHANGE DATA

GENERAL LINEAR MODELS PROCEDURE

DUNCAN'S MULTIPLE RANGE TEST FOR VARIABLE: RESPONSE

NOTE: THIS TEST CONTROLS THE TYPE I COMPARISON ERROR RATE, NOT THE EXPERIMENTAL ERROR RATE.

\[ \alpha = 0.05 \]

\[ \text{DF} = 75, 145; \text{DF} = 62, 349 \]

WARNING: CELL SIZES ARE NOT EQUAL.

HARMONIC MEAN OF CELL SIZES = 4578

NUMBER OF MEANS = 6

CRITICAL RANGE = 5.04803, 5.30813, 5.47758

MEANS WITH THE SAME LETTER ARE NOT SIGNIFICANTLY DIFFERENT.

DUNCAN GROUPING MEAN

<table>
<thead>
<tr>
<th>DRUG</th>
<th>N</th>
<th>MEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>16</td>
<td>10.864</td>
</tr>
<tr>
<td>B</td>
<td>17</td>
<td>9.048</td>
</tr>
</tbody>
</table>

TESTING DOSE RESPONSE TREND BY JONCKHEERE'S DISTRIBUTION FREE TEST

Dataset jonck: Jonckheere’s Statistic

COUNT 1 COUNT 2 COUNT 3 COUNT 4 TCOUNT

|    | 21 | 22 | 16 | 20 | 79 |

TESTING DOSE RESPONSE TREND BY JONCKHEERE'S DISTRIBUTION FREE TEST

\[ \text{DBS} = 3.20995; \text{P-VALUE} = 0.00066379 \]

APPENDIX II

Computer Input/Output for Example 3:

/*LOAD DATA*/
/*READ IN DATA*/
/*PROC MEANS*/
/*PROC GLM*/
/*DUMP DATA*/
/*CUT DATA*/
/*RUN*/
APPENDIX II - Example 3 (Continued):

**Computer Output**

**ANALYSIS OF VARIANCE ON RESPONSE CHANGE DATA**

**GENERAL LINEAR MODELS PROCEDURE**

**DUNCAN'S MULTIPLE RANGE TEST FOR VARIABLE: RESPONSE**

*NOTE: THIS TEST CONTROLS THE TYPE I COMPARISONWISE ERROR RATE, NOT THE EXPERIMENTWISE ERROR RATE*

<table>
<thead>
<tr>
<th>Alpha</th>
<th>0.05</th>
<th>MSE</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NUMBER OF MEANS</strong></td>
<td>3</td>
<td><strong>CRITICAL RANGE</strong></td>
<td>5.2117 - 5.46707</td>
</tr>
</tbody>
</table>

Means with the same letter are not significantly different.

<table>
<thead>
<tr>
<th>DUNCAN GROUPING</th>
<th>MEAN</th>
<th>N SCHOOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>46.667</td>
<td>6 C</td>
</tr>
<tr>
<td>A</td>
<td>44.067</td>
<td>6 B</td>
</tr>
<tr>
<td>A</td>
<td>42.067</td>
<td>6 A</td>
</tr>
</tbody>
</table>

**TESTING DOSE RESPONSE TREND BY JONCKHEERE'S DISTRIBUTION FREE TEST**

Dataset Jonckhr: Jonckheere's Statistic

<table>
<thead>
<tr>
<th>OBS</th>
<th>JSTAR</th>
<th>P_VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.65733</td>
<td>0.0487267</td>
</tr>
</tbody>
</table>

1108