

Detecting Outlying Data in Proficiency Studies with SAS®

Stacey Slone, University of Kentucky Markey Cancer Center

ABSTRACT

Regardless of the industry, one of the first steps in analyzing data from a collaborative or a proficiency study is to detect any outliers in the data. Regulations published by the International Organization for Standardization, ISO, enumerate several tests to consider (ISO 5725-2, 1994; ISO/FDIS 13528, 2015). To identify outlying means in the laboratories participating in proficiency study, the ISO regulations suggest the single Grubb's test and the Mandel's h test. Similarly, the ISO regulations propose using the Cochran's test and the Mandel's k test to identify outlying standard deviations among participating laboratories. ISO 5725-2 provides a partial table of significant values for these four outlier tests (ISO 5725-2, 1994). However, since these collaborative studies typically analyze multiple parameters, a programmatic solution to finding not only the scores, but the critical values, leads to savings in time and effort. In his 2013 paper, Wilrich (Wilrich, 2013) derived the formulas for the critical values for each of the fore mentioned tests.

The included SAS® macro can perform any of these tests on outliers, compare the result with the correct critical values and print a summary or graph as appropriate based on the tests chosen. Options for printing include significant results for the Grubbs or Cochran test, graphs for the Mandel's h and k , with reference lines at the 95% and 99% critical values, and a summary table of results.

INTRODUCTION

Outlier detection is a vital aspect of the statistical analyses in proficiency studies. Regardless of how outlying data are handled, via exclusion or identification, reporting the occurrence rate is an essential metric of data and material reliability. Robust algorithms are considered reliable only if less than 20% of the data are declared outliers (ISO/FDIS 13528, 2015). Hence, even when robust algorithms are used, testing providers must consider the other outlier tests as supporting documentation (ISO 5725-2, 1994) (ISO/FDIS 13528, 2015). While refining the analytic SAS code for the proficiency studies implemented by the Center for Tobacco Research Products (CTRP) at the University of Kentucky, the analytic team needed a programmatic approach to calculate and present the outlier tests. CTRP planned a series of proficiency studies to explore multiple physical and chemical parameters, so relying on the ISO tables would not have been efficacious. The %OUTLIER_DET macro resulted.

CORESTA, the Cooperation Centre for Scientific Research Relative to Tobacco, is a French association that runs proficiency and cooperative studies on tobacco and its derivatives. To illustrate the application of the %OUTLIER_DET macro, an invocation will utilize sample data from a published CORESTA collaborative study (CORESTA, 2015). The example will demonstrate the possible output from each test available in the macro.

OUTLIER TESTS IN PROFICIENCY STUDIES

For proficiency studies, as well as other situations with repeated observations, the consistency of the data collected should be reviewed for outliers in either the means or the standard deviations. To the test for outlying data in the means, one calculates the Mandel's h statistic and the Grubb's test. The Mandel's h statistic compares a laboratory's mean to the mean of the laboratory means divided by the standard deviation of the laboratory means. In essence, it is a standardized difference of each lab's means to the composite mean. Grubb's test has two versions, but the macro focuses on the Single Grubb's test. It compares the absolute value of the standardized difference of a particular lab's mean to the overall mean. To test for highly variable spreads in laboratory values, the Mandel's k and the Cochran's test assess the

data for outliers with respect to the variance of the data. Mandel's k is the standard deviation of a particular laboratory divided by the square root of the mean of the laboratory variances. Cochran's test is the intralaboratory variance divided by the sum of the intralaboratory variances. Although ISO 5725-2 provides a table of the most commonly used critical values for the four tests, the table only provides the 95th percentile and 99th percentile critical values for selected sample sizes (ISO 5725-2, 1994). Typically, values are "stragglers" or issued a "warning" if the value of the test is outside of the 95% critical value. Similarly, values are termed "outliers" or issued an "alert" if the value is outside of the 99% critical value. Being limited to a table to determine these classifications hinders automation of analyses.

However, Wilrich derived the formulas for the critical values, which allowed the computations to be programmed in SAS (Wilrich, 2013) using the SQL procedure and the TINV and FINV functions. Table 1 summarizes the outlier tests available in the macro and the formulas for the corresponding critical values. In the formulae, n is the number of observations for each laboratory while p represents the number of laboratories. For the F-distribution degrees of freedom notation, $v1 = (p - 1)(n - 1)$ and $v2 = (n - 1)$.

Test Name	Outlier Tested	Inspection	Formula	Critical Value
Mandel's h Statistic	Mean	Graphical	$h_i = \frac{(\bar{X}_i - \bar{\bar{X}})}{s_{\bar{X}}}$	$h_{p;1-\alpha/2} = \frac{(p-1)t_{p-2;1-\alpha/2}}{\sqrt{p(p-2+t_{p-2;1-\alpha/2}^2)}}$
Single Grubb's Test	Mean	Comparison	$g = \max_i(h_i)$	$g_{p;1-\alpha} \approx \frac{(p-1)t_{p-2;1-\alpha/p}}{\sqrt{p(p-2+t_{p-2;1-\alpha/p}^2)}}$
Mandel's k Statistic	Spread	Graphical	$k_i = \frac{s_i}{\sqrt{\sum_i s_i^2/p}}$	$k_{p,n;1-\alpha} = \sqrt{\frac{p}{(1+(p-1)F_{v1,v2;\alpha})}}$
Cochran's Test	Spread	Comparison	$C_i = \frac{s_i^2}{\sum_i s_i^2}$	$c_{p,n;1-\alpha} \approx (1+(p-1)F_{v1,v2;\alpha/p})^{-1}$

Table 1. Summary of Outlier Tests

MACRO DEFINITION

The %OUTLIER_DET macro expects a raw data set of values with a variable identifying the laboratory or other classification. First, the MEANS procedure calculates the mean, standard deviation and variance from the raw data stratified by the classification variable. An ODS TABLE statement saves the output into a SAS data set. Next, PROC SQL utilizes the data set to perform all of the remaining calculations. The macro supports all four of the described tests for outliers and can be modified to perform any combination of the tests. In addition, the macro can print, if requested, a summary table listing codes for the all laboratories with outlying data from any of the tests. The test codes chosen determine the output from the macro.

The macro call for %OUTLIER_DET has 5 inputs and up to 2 output data sets. The graphs and tables vary according to the tests chosen. The macro call is as follows:

```
%OUTLIER_DET(data_, varlist_, class_, test_=CGHK, table_=Y);
```

Appendix A provides the full code for %OUTLIER_DET.

MACRO CALL INPUTS

The following lists the expected input elements to the macro call:

- DATA_ is the raw data set of values. The macro will perform calculations in the first step using PROC MEANS to summarize the data.

- VARLIST_ is the list of variables to be tested. Each variable should contain the raw data for the corresponding measure.
- CLASS_ is the variable that contains the laboratory or other group identifier. It can be either a numeric or a character field.
- TEST_ is a code referring to the list of statistical tests the macro should calculate. It can include any combination of tests and order the codes are listed is trivial. The codes are G for Grubb's Single test, C for Cochran's test, H for Mandel's h test and K for Mandel's k test. To produce all four tests, the invocation should be GCHK; if the user desires only the two Mandel's graphs, the invocation would be KH. The default is all four tests.
- TABLE_ is Y/N for a summary table to be printed. The table defaults to Y to create and print the table.

MACRO CALL OUTPUT

Grubb's Test

The output from the macro depends on the options chosen. If &TEST_ contains a G, then the single Grubb's test is performed. If any of the laboratories have a mean beyond the 95% and 99% confidence interval from a t-distribution, the PRINT procedure prints the list of laboratories along with the relevant values. A single star in the "Significant" column indicates a straggler, while two stars indicate an outlier. If none of the values fall outside of the confidence intervals, then the macro prints a warning to the log.

Mandel's *h* Test

If &TEST_ contains an H, then the macro calculates the Mandel's *h* statistics and calls the SGPLOT procedure to create a vertical bar graph complete with color coded 95% and 99% confidence bounds from a t-distribution.

Cochran's Test

Similar to the Grubb's test, if &TEST_ contains a C, then Cochran's test is performed. If any of the laboratories have a mean greater than the 95% and 99% critical values from a F-distribution, the PROC PRINT prints the list of laboratories along with the relevant values. A single star in the "Significant" column indicates a straggler, while two stars indicate an outlier. If all of the values are less than the critical values, then the macro prints a warning to the log.

Mandel's *k* Test

Analogous to the Mandel's *h*, if &TEST_ contains a K, then the macro calculates the Mandel's *k* statistics and invokes PROC SGPLOT to generate a vertical bar graph complete with color-coded 95% and 99% critical bounds from the F-distribution.

In addition to the tables and figures output from the macro, the user has access to the following two output data sets:

- OUT_ is a data set containing all of the test values and significant values.
- TABLE_ is a data set containing the list of labs by variable with stragglers or outliers (outside the 95% and 99% confidence intervals, respectively).

EXAMPLE

SAMPLE DATA

To illustrate the %OUTLIER_DET macro, the Cooperation Centre for Scientific Research Relative to Tobacco (CORESTA) granted permission to use published data from one of its collaborative studies on tobacco (CORESTA, 2015) which studied multiple chemical smoking properties. Data from the CM 7

monitor cigarette was chosen for the demonstration focusing on the seventeen laboratories submitting data from a linear smoking machine. Each laboratory submitted five replications. The example focuses on the water and nicotine measurements in smoking. Output 1 lists the first eight observations:

Obs	Lab_ Code	make	Conditioned_ Weight	TPM_	Water	Nicotine	NFDPM_	CO	Puff_ count	type
6	2B	L	953.6	16.62	1.46	1.211	13.94	12.57	8.40	ISO
7	2B	L	951.1	16.68	1.46	1.244	13.98	12.49	8.57	ISO
8	2B	L	952.5	16.24	1.44	1.191	13.61	12.31	8.35	ISO
9	2B	L	948.5	16.08	1.44	1.218	13.42	12.44	8.48	ISO
10	2B	L	959.1	16.73	1.53	1.252	13.95	12.79	8.54	ISO
16	3B	L	963.0	16.69	1.58	1.269	13.84	13.39	8.94	ISO
17	3B	L	962.0	16.57	1.42	1.253	13.90	13.04	8.86	ISO
18	3B	L	958.0	16.42	1.37	1.248	13.80	13.09	8.75	ISO

Output 1: Sampling of Example Data

EXAMPLE INVOCATION OF %OUTLIER_DET

To demonstrate the full breadth of the macro %OUTLIER_DATA, the example macro call includes all of the options. The following macro call produces the output shown:

```
%OUTLIER_DET(Example_Data, water nicotine, lab_code, test_=CGHK, table_=Y);
```

Using an iterative loop, the macro performs the tests for the water data, prints any relative output, and then performs the tests for the nicotine data with its output. Lastly, since the macro call includes the request for a summary table, the macro prints the table with each variable as a row.

EXAMPLE OUTPUT

Since in the sample data none of the lab means for water fall outside of the confidence bounds, the macro prints the following message, Output 2, in the SAS log:

```
*****
"Grubb's Test:  No Warnings or Alerts"
*****
```

Output 2. Printout in Log

Similarly, during the second iteration for nicotine values, the same message to the log prints because none of the nicotine values is outside the range either.

Next, Laboratory 15A from the sample data is significantly more variable than the rest according to the Cochran's test. Lab 15A has a Cochran's value of 0.439, which is greater than the 1% critical value of 0.261; hence, lab 15A is an outlier. The macro prints the following table, Table 2, in the listing.

Water (mg/tp) Cochran's Outlier Test

Lab Code	Cochran's	5% Critical Value	1% Critical Value	Significant
15A	0.43871	0.21899	0.26088	**

Table 2. Cochran's Summary Table

However, none of the nicotine lab values are outside of the critical values for Cochran's test, so Output 3 prints in the SAS log:

```

*****
"Cochran's Test:  No Warnings or Alerts"
*****

```

Output 3. Printout in Log

For succinctness, only the Mandel's h and k figures for the water data will be shown. The macro call creates similar graphs for nicotine in the full output. Comparing the mean water values for each lab via the Mandel's h statistic, Lab 28 has a significantly greater mean than the remaining labs. As seen in Figure 1, the bar for Lab 28 crosses the top blue line, which signifies the upper 95% critical value from the t-distribution with 15 degrees of freedom. However, the mean is only a "straggler" since it is below the red line signifying the upper 99% critical value.

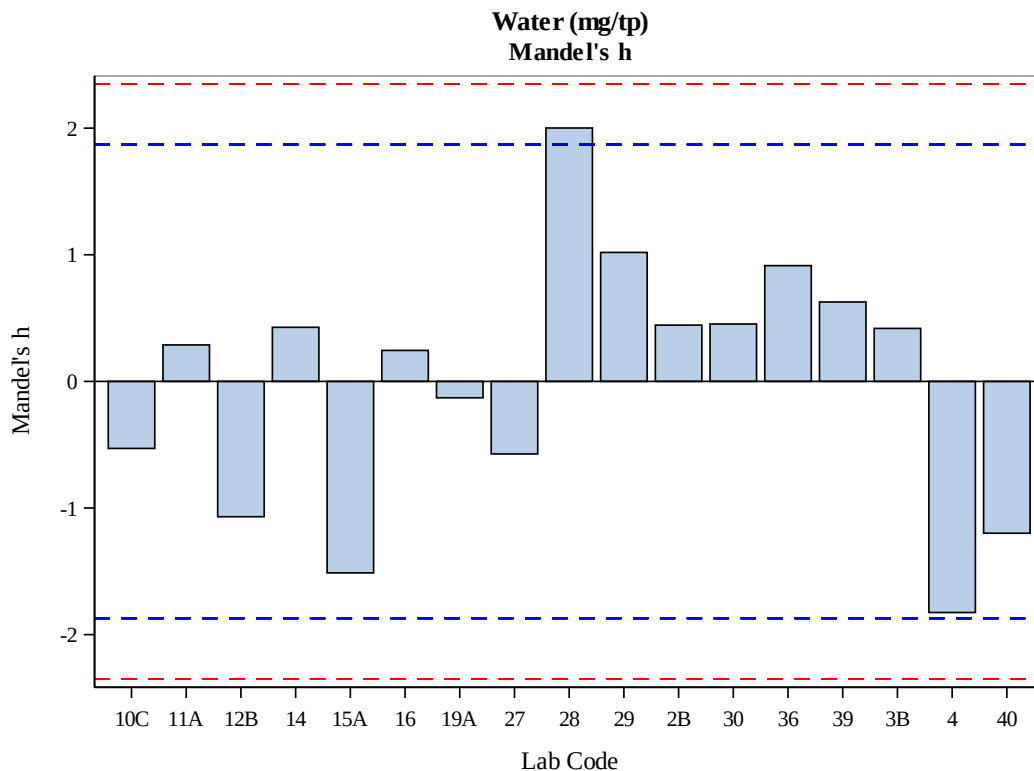


Figure 1. Mandel's h Graph

The macro creates a similar figure for the nicotine data. Next, the macro compares the variation in water values for each lab via the Mandel's k statistic. Both Lab 15A and Lab 28 have significantly more variable than the remaining labs. As seen in Figure 2, the bars for both labs cross the red line, which signifies the 99% critical value from the F-distribution with (64,4) degrees of freedom.

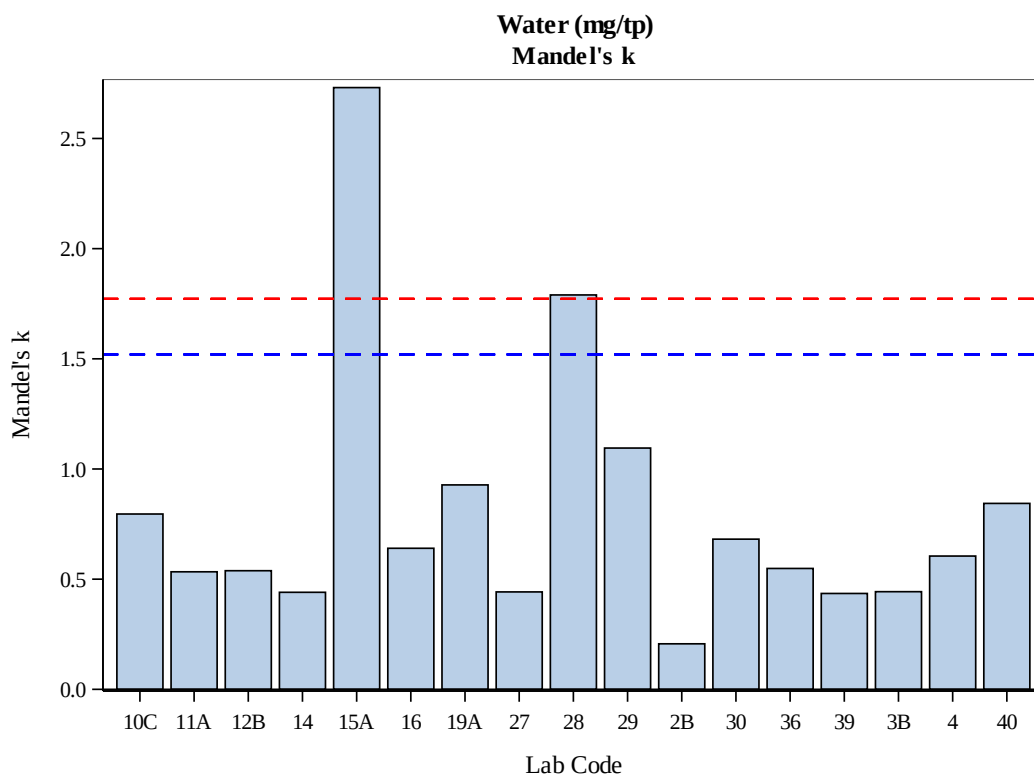


Figure 2. Mandel's k Graph

Finally, the macro produces, as a default, a summary table of stragglers (values between the 95th and the 99th confidence bounds) and the outliers (values outside the 99th confidence bounds). Table 3 illustrates the summary for the sample data for both nicotine and water:

	Stragglers				Outliers			
	Mandel's h	Mandel's k	Cochrans	Grubbs	Mandel's h	Mandel's k	Cochrans	Grubbs
Nicotine (mg/tp)	27							
Water (mg/tp)	28					15A,28	15A	

Table 3. Summary Table of Outliers

CONCLUSION

Outlier detection is a common aspect of proficiency and collaborative testing in a multitude of industries. Identifying outliers in means with the Mandel's *h* and the Grubb's test tells the participating laboratories to investigate their mean values. Similarly, identifying outliers in the spreads with the Mandel's *k* and Cochran's tests imparts the laboratories to review their validations. The %OUTLIER_DET macro fills a void in the world of collaborative and proficiency studies for statisticians and analysts. The macro automates the required testing for outliers in either the means or the standard deviations within laboratories and produces publication ready tables and graphs. The flexibility to choose the tests desired makes %OUTLIER_DET an extremely powerful and useful tool.

REFERENCES

- CORESTA. (2015). *2014-2015 Collaborative Study of CORESTA Monitor Test Pieces CM 7 and CM 8 for the Determination of Test Piece Weight, TPM, Water, Nicotine, NFDPM, Carbon Monoxide and Puff Count Obtained under Mainstream 'ISO' Smoking Regime*.
- ISO 5725-2. (1994). *Accuracy (trueness and Precision) of Measurement Methods and Results – Part 2: Basic Method for the Determination of Repeatability and Reproducibility of a Standard Measurement Method*. Geneva, Switzerland.
- ISO/FDIS 13528. (2015). *Statistical Methods for Use in Proficiency Testing by Interlaboratory Comparison*. Geneva, Switzerland.
- Wilrich, P.-T. (2013). Critical Values of Mandel's h and k, the Grubbs and the Cochran Test Statistic. *AStA Advances in Statistical Analysis*, 97, 1-10.

ACKNOWLEDGMENTS

Funding for this presentation was made possible, in part, by the Food and Drug Administration through grant RFA-FD-14-001. The views expressed in written materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does any mention of trade names, commercial practices, or organization imply endorsement by the United States Government.

The author would like to thank Brent Shelton PhD, Socrates Cañete PhD, and Andrew Shearer for their review of the manuscript and collaboration on the analysis algorithm that served as the inspiration for developing the macro.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Stacey Slone
University of Kentucky Markey Cancer Center
800 Rose Street, CC440
Lexington, KY 40536
859-323-1723
859-257-7715
Stacey.Slone@uky.edu

APPENDIX A

```
%macro outlier_det(data_, varlist_, class_, test_=CGHK, table_=Y);  
%do v=1 %to %sysfunc(countw(&varlist_));  
  
%let var_=%scan(&varlist_, &v);  
  
data _null_;  
  set &data_;  
  call symput ('vlabel_',vlabel(&var_));  
run;  
  
ods select none;  
ods table summary=means_;  
proc means data=&data_ order=data N mean var;  
  class &class_;  
  var &var_;
```

```

run;

data _null_;
  call symput('mhs', ' ');
  call symput('mks', ' ');
  call symput('crs', ' ');
  call symput('grs', ' ');
  call symput('mho', ' ');
  call symput('mko', ' ');
  call symput('cro', ' ');
  call symput('gro', ' ');

ods select all;
proc sql noprint;
  create table out_ as
  select &class_, sqrt(&var_._var)/sqrt(mean(&var_._var)) as k_i
  label="Mandel's k",
  (&var_._mean-mean(&var_._mean))/std(&var_._mean) as h_i label="Mandel's
  h",
  &var_._var/sum(&var_._var) as cochran label="Cochran's",
  abs(&var_._mean-mean(&var_._mean))/std(&var_._mean) as grubbs
  label="Single Grubb's",
  count(unique &class_) as p,

  /* Mandel's k critical values */
  (calculated p / (1+(calculated p-1)*finv(0.05,(calculated p-1)*
  (mean(&var_._N)-1),
  (mean(&var_._N)-1))))*0.5 as Kf05 label="5% Critical# Value",
  (calculated p / (1+(calculated p-1)*finv(0.01,(calculated p-1)*
  (mean(&var_._N)-1), (mean(&var_._N)-1))))*0.5 as Kf01 label="1%
  Critical# Value",

  /* Cochran critical values */
  1 / (1+(calculated p-1)*finv(0.05/calculated p,(calculated p-1)*
  (mean(&var_._N)-1),
  (mean(&var_._N)-1))) as cf05 label="5% Critical# Value",
  1/(1+(calculated p-1)*finv(0.01/calculated p,(calculated p-1)*
  (mean(&var_._N)-1),
  (mean(&var_._N)-1))) as cf01 label="1% Critical# Value",
  case
    when calculated cochran le calculated cf05 then ' '
    when calculated cf01 ge calculated cochran gt calculated cf05 then '*'
    when calculated cf01 lt calculated cochran then '**'
  end as cochran_result label='Significant',

  /* Mandel's h critical values */
  (calculated p-1)*(tinv(1-0.05/2,calculated p-2))*((calculated p*
  (calculated p-2+(tinv(1-0.05/2,calculated p-2)**2))**0.5) as hf05
  label="5% Critical# Value",
  (calculated p-1)*(tinv(1-0.01/2,calculated p-2))*((calculated p*
  (calculated p-2+(tinv(1-0.01/2,calculated p-2)**2))**0.5) as hf01
  label="1% Critical# Value",
  -1*calculated hf05 as nf05, -1*calculated hf01 as nf01,

  /*Grubbs critical values */

```

```

(calculated p-1)*(tinv(1-0.05/calculated p,calculated p-2))*((calculated
p*(calculated p-2+(tinv(1-0.05/calculated p,calculated p-2)**2))**0.5)
as gf05 label="5% Critical# Value",
(calculated p-1)*(tinv(1-0.01/calculated p,calculated p-2))*((calculated
p*(calculated p-2+(tinv(1-0.01/calculated p,calculated p-2)**2))**0.5)
as gf01 label="1% Critical# Value",
case
  when calculated grubbs le calculated gf05 then ' '
  when calculated gf01 ge calculated grubbs gt calculated gf05 then '*'
  when calculated gf01 lt calculated grubbs then '**'
  end as grubbs_result label='Significant'
from means_;

%if &table_ ne %then
%do;
select distinct &class_ into: mhs separated by ',' from out_ where
abs(hf05) le abs(h_i) lt abs(hf01);
select distinct &class_ into: mks separated by ',' from out_ where kf05
le k_i lt kf01;
select distinct &class_ into: crs separated by ',' from out_ where
cochran_result='*';
select distinct &class_ into: grs separated by ',' from out_ where
grubbs_result='*';
select distinct &class_ into: mho separated by ',' from out_ where
abs(h_i) ge abs(hf01);
select distinct &class_ into: mko separated by ',' from out_ where k_i ge
kf01;
select distinct &class_ into: cro separated by ',' from out_ where
cochran_result='**';
select distinct &class_ into: gro separated by ',' from out_ where
grubbs_result='**';
%end;
quit;

%if %index(&test_,C) gt 0 %then %do;
data ctemp_;
set out_;
where cochran_result ne '';
run;
%if &sysnobs=0 %then %do;
%put ;
%put *****;
%put "Cochran's Test: No Warnings or Alerts";
%put *****;
%put ;
%end;
%else %str(
proc print data=ctemp_ label split='#';
id &class_;
var cochran cf05 cf01 cochran_result;
title "&vlabel_";
title2 "Cochran's Outlier Test";
run;);
%end;

%if %index(&test_,G) gt 0 %then %do;
data gtemp_;

```

```

        set out_;
        where grubbs_result ne '';
run;
%if &sysnobs=0 %then %do;
    %put ;
    %put *****;
    %put "Grubb's Test:  No Warnings or Alerts";
    %put *****;
    %put ;
%end;
%else %str(
    proc print data=gtemp_ noobs label split='#';
        id &class_;
        var grubbs gf05 gf01 grubbs_result;
        title "&vlabel_";
        title2 "Grubb's Outlier Test";
    run;);
%end;

ods graphics on / border=off;
%if %index(&test_,H) gt 0 %then %str(
proc sgplot data=out_ noautolegend;
    vbar &class_ /response=h_i stat=mean nostatlabel;
    reflate hf05 / noclip lineattrs=(color=blue pattern=dash);
    reflate hf01 / noclip lineattrs=(color=red pattern=dash);
    reflate nf05 / noclip lineattrs=(color=blue pattern=dash);
    reflate nf01 / noclip lineattrs=(color=red pattern=dash);
    title "&vlabel_";
    title2 "Mandel's h";
run;
);

%if %index(&test_,K) gt 0 %then %str(
proc sgplot data=out_ noautolegend;
    vbar &class_ /response=k_i stat=mean nostatlabel;
    reflate kf05 / noclip lineattrs=(color=blue pattern=dash);
    reflate kf01 / noclip lineattrs=(color=red pattern=dash);
    title "&vlabel_";
    title2 "Mandel's k";
run;
);

%if %upcase(&table_)=Y | %upcase(&table_)=YES %then %do;

data ttable_;
    length var $ 25;
    var=vlabel(&var_);
    %if %index(&test_,H) gt 0 %then %do;
        mandelhs="&mhs";
        mandelho="&mho";
    %end;
    %if %index(&test_,K) gt 0 %then %do;
        mandelks="&mks";
        mandelko="&mko";
    %end;
    %if %index(&test_,C) gt 0 %then %do;
        cochrans="&crs";

```

```

        cochrano="&cro";
    %end;
    %if %index(&test_,G) gt 0 %then %do;
        grubbss="&grs";
        grubbso="&gro";
    %end;
run;

proc append base=table_ data=ttable_ force;
run;
%end;
proc datasets nolist;
    delete ttable_ means_ out_ gtemp_ ctemp_;
    quit;
%end;

%if %upcase(&table_)=Y | %upcase(&table_)=YES %then %do;
proc report data=table_ nowindows;
    column var ('Stragglers' (mandelhs mandelks cochrans grubbss))
            ('Outliers' (mandelho mandelko cochrano grubbso));
    define var / order order=internal "";
    %if %index(&test_,H) gt 0 %then %str(
define mandelhs / display "Mandel's h";
define mandelho / display "Mandel's h"););

    %if %index(&test_,K) gt 0 %then %str(
define mandelks / display "Mandel's k";
define mandelko / display "Mandel's k"););

    %if %index(&test_,C) gt 0 %then %str(
define cochrans / display "Cochrans";
define cochrano / display "Cochrans"););

    %if %index(&test_,G) gt 0 %then %str(
define grubbss / display "Grubbs";
define grubbso / display "Grubbs"););

    title "Summary Table";
run;
%end;

%mend outlier_det;

```