



Journal of Statistical Software

April 2008, Volume 25, Issue 13.

<http://www.jstatsoft.org/>

SAS Macros for Analysis of Unreplicated 2^k and 2^{k-p} Designs with a Possible Outlier

John Lawson

Brigham Young University

Abstract

Many techniques have been proposed for judging the significance of effects in unreplicated 2^k and 2^{k-p} designs. However, relatively few methods have been proposed for analyzing unreplicated designs with possible outliers. Outliers can be a major impediment to valid interpretation of data from unreplicated designs. This paper presents SAS macros which automate a manual method for detecting an outlier and performing an analysis of data from an unreplicated 2^k or 2^{k-p} design when an outlier is present. This method was originally suggested by Cuthbert Daniel and is based on the normal or half normal plot of effects. This automated version was shown in simulation studies to perform better than other procedures proposed to do the same thing.

Keywords: normal plot of effects, half normal plot of effects, SAS.

1. Introduction

Motivated by the publication of books such as *Statistics for Experimenters* by Box, Hunter, and Hunter (1978) and the development of user-friendly software, unreplicated 2^k factorial and 2^{k-p} fractional factorial experimental designs have become standard tools used in industry, engineering, and other areas of research. These designs allow researchers to quickly identify important factors and get to the heart of a research question. Application journals and company documents contain ample examples of the use of these designs. Software such as SAS ADX (SAS Institute. 2003), Minitab (Minitab Inc. 2007), Stat Graphics (StatPoint-Inc 2008), Design Ease (Stat-Ease 2008), JMP (JMP 2008), and NCSS (Hintze 2007) are typically used to create the designs and analyze data after completion of the experiments. The most often used run sizes in these experiments are 8, 16 or 32, and the vast majority of these experiments are conducted by practitioners whose primary training is not in statistics.

In unreplicated 2^k factorials and 2^{k-p} fractional factorial designs, factorial main effects are

calculated as the difference in the average response between the high and low coded levels of a factor, and interaction effects are calculated as the difference in the average response between the high and low values of products of coded factor levels. Symbolically the effects can be written as $\hat{E} = \bar{y}_+ - \bar{y}_-$, and the regression coefficients in the linear model are $\hat{\beta} = \hat{E}/2$. For these designs the model is saturated. There are as many coefficients in the model as there are observations. Usually the hypothesis of *effect sparsity* must be assumed where only a fraction of the effects in the model are believed to be *active* or non-zero. These designs are very efficient in maximizing the amount of information obtained from n experiments, but because there are zero degrees of freedom for error, the active effects in these models cannot be identified using the normal theory t or F statistics.

Alternate methods have been proposed in the literature for overcoming this problem. One of the early methods proposed by Daniel (1959) was to make a half normal plot of the absolute effects and qualitatively identify the active effects to be those which appear as outliers on the plot. The rationale for Daniel's proposal was that if none of the effects are active, all the calculated effects will be differences of averages of random errors and approximately normally distributed due to the central limit theorem. Since his proposal there has been a plethora of additional quantitative and graphical methods proposed for detecting the active effects in unreplicated 2^k and 2^{k-p} designs. Refer to Hamada and Balakrishnan (1998) for a review and comparison. Some of the more popular methods have been automated and incorporated into the user-friendly software mentioned earlier.

When there is one or more outliers among the response for an unreplicated 2^k or 2^{k-p} design, there is an additional problem. All the estimated effects are biased, and the power of automatic methods used to detect significant effects in unreplicated designs is compromised. Daniel (1959) believed that outliers or bad values are a major hazard of unreplicated factorial experiments. He said that they are more damaging than missing values because they are not easily identifiable. Normally the presence of outliers can be detected when fitting the general linear model $y = X\beta + \epsilon$ by examination of the residuals $r = y - X\hat{\beta}$. However, for unreplicated 2^k factorials and 2^{k-p} fractional factorials, the residuals are all zero. If there are q outliers in an unreplicated 2^{k-p} design, detecting the outliers and estimating all the effects is equivalent to estimating $q + 2^{k-p}$ parameters with only 2^{k-p} data points.

There have been some methods proposed in the literature for detecting outliers and testing the significance of effects in the presence of outliers in unreplicated designs. However, these methods are relatively few in number compared to the number of methods proposed for detecting significant effects in these designs. John (1978) proposed a method where some terms in the model are assumed *a priori* to be negligible and are left out of the model. Potential outliers are then identified in residual plots and their existence is confirmed by fitting a model to the data that includes dummy variables that are indicators of the observations suspected to be outliers. Carroll (1980) proposed using robust regression to analyze a factorial with possible outliers, but to use his method some terms in the model must again be assumed *a priori* to be negligible. Seheult and Tukey (1982) proposed using robust location estimates to calculate factorial effects in the presence of outliers, but in unreplicated designs there is no robust estimate of the higher order interactions. Therefore these model terms must essentially be assumed negligible. These three methods will work only if the *a priori* assumption that some effects are negligible is correct, and if the number of runs in the design is large.

When the number of runs in an unreplicated fractional factorial design is large, i.e., $n = 2^{k-p} > 16$, such as 32, 64 etc., the design itself is fairly robust to outliers since each calculated

effect is a difference of averages of 16 or more observations. In this case the factorial effects will not be greatly biased by a few outliers and the active effects can usually be identified using the normal or half normal plot of effects (or another method) and included in an initial model. Then, outliers in the data can be identified by examining the residuals from the initial model. Using this approach no effects need be assumed negligible *a priori*. However, if $n = 2^{k-p} = 8$ or 16 all of the calculated effects (active or not) are biased by at least $\pm b/8$ if there is at least one observation that is in error by the amount b . This makes it difficult to determine what effects are active and should be included in the initial model. Therefore it is difficult to detect outliers in the 8 or 16 run designs that are commonly used in practice.

A few additional methods have been proposed in the literature for recognizing the presence outliers, without having to first identify an initial model. Daniel (1959) proposed to recognize the presence of a single outlier from the pattern of points on the half normal plot of effects. He showed that once the presence of a single outlier was recognized, the outlying observation could be identified manually and a correction factor could be calculated from the data. Box (1991) illustrated the use of the full normal plot to accomplish the same thing, and he showed in detail how to identify the outlying observation and calculate the correction factor.

Daniel (1959) and Daniel (1960) also discussed methods to recognize the presence of two or more outliers from the pattern of points on the half normal plot, however these ideas were more subjective and would be more difficult to automate in computer code. Gatlin (2004) wrote SAS code to automate Daniel's proposed method of detecting a single outlier using the full normal plot as illustrated by Box (1991). Box and Meyer (1987) proposed an automatic iterative Bayesian method that is purported to detect outliers in unreplicated factorials and simultaneously provide robust estimates of the effects. Torres (1993) proposed to recognize the presence of an outlier when the large effects identified in the analysis of the response data differed from the large effects identified in the analysis of the ranks of the response data. Lawson and Gatlin (2006) reported the results of a simulation study to compare these three methods.

Although the methods proposed by Box and Meyer (1987) and Torres (1993) theoretically will work with more than one outlier, simulations by Lawson and Gatlin (2006) show that these methods could not even detect a single moderate sized outlier consistently in a 16 run design. In the simulations, Gatlin's automated version of the Daniel-Box manual procedure was more consistent in detecting a single outlier and in general showed a greater increase in the power to detect significant effects in the presence of a single outlier. Although this method will not work with multiple outliers, the chance of multiple outliers in an 8 or 16 run design is small. Daniel (1959) estimated from his experience that the chance an individual observation is in error is 0.01 to 0.1. Thus, the chance of more than one outlier in an 8 or 16 run designs very small.

The software commonly used, in practical applications, for the design and analysis of data from unreplicated 2^k factorials and 2^{k-p} fractional factorial experimental designs do not include procedures for detecting outliers other than through residual analysis, and they do not issue a warning to the user when the analysis could be compromised due to the presence of an outlier. This could be a problem for 8 and 16 run designs that are commonly used in practice, and it would be desirable to have an automated version of the Daniel-Box procedure in the packages that are commonly used for these analyses. Fortunately, some of these packages such as SAS, Minitab and JMP have macro languages that make it possible to add user written procedures. This article presents a SAS macro that performs an automated version of Daniel's procedure

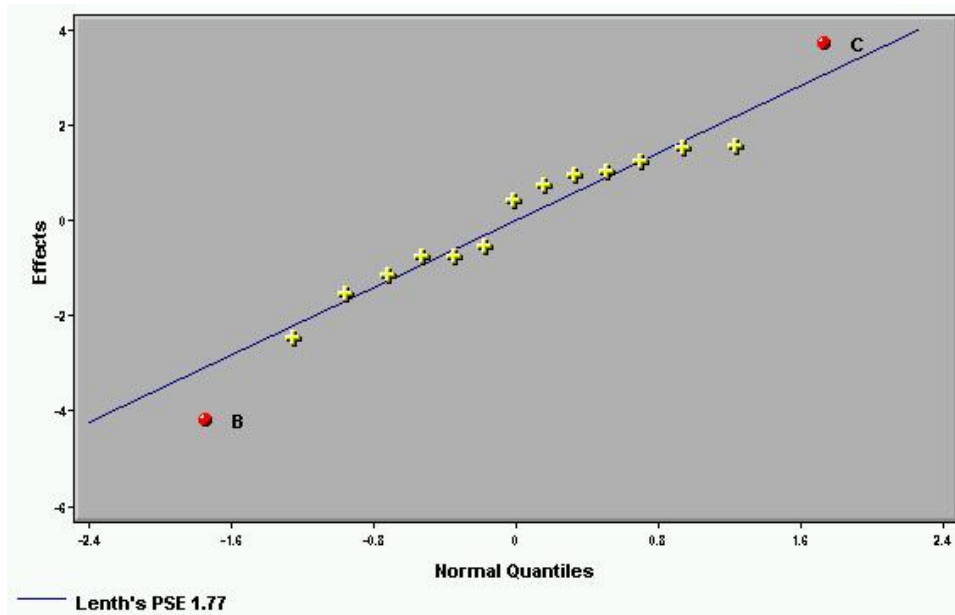


Figure 1: Normal probability plot of effects coefficients demonstrating a gap caused by an outlier.

for detecting a single outlier and correcting a single outlier in an unreplicated 2^k factorial or 2^{k-p} fractional factorial design.

2. Daniel's Method for Detecting an Outlier

Daniel's original suggestion for detecting and correcting outliers in unreplicated factorials consists of four steps. First, the presence of the outlier is recognized by a non-zero intercept on the half normal plot of effects, or as Box explained, by a gap in the full normal plot of regression coefficients. The second step is to identify the outlier by matching the signs of the insignificant effects with the signs for the contrast coefficients for the factor levels and interactions of each observation. The third step is to estimate the magnitude of the discrepancy and correct the outlier. The fourth and final step is to reconstruct the half normal or normal plot of effects calculated from the data with the outlier corrected.

As an example of recognizing the presence of an outlier on the Normal plot of effects, see Figure 1. It was produced by the SAS automated design of experiments package, **ADX**, using the data from the 2^4 experiment described by Box (1991). In the plot the **ADX** system identifies main effects *B* and *C* to be significant using Lenth (1989) method. However, in the center of the plot, where the smaller effects in absolute value tend to fall along a straight line, there is a vertical gap separating the positive and negative effects. As Box (1991) explained, one aberrant observation would cause half the negligible effects to be biased high or positive and half to be biased low or negative. Thus the gap in the center of the negligible effects on the plot.

Gatlin (2004) created an automated version of Daniel-Box procedure for recognizing the presence of an outlier on the Normal plot of effects. To do this, the method of recognizing a

gap on the normal plot was quantified by testing the significance of a gap statistic

$$G = \frac{(\hat{\beta}_S - \hat{\beta}_L)/PSE}{Z_S - Z_L}, \quad (1)$$

where $\hat{\beta}_S$ is the smallest positive estimated regression coefficient and $\hat{\beta}_L$ is the largest negative estimated regression coefficient, PSE is [Lenth \(1989\)](#) pseudo standard error calculated from the set of regression coefficients, and Z_S and Z_L are the normal scores associated with the (S)th and (L)th order statistics from a sample of size $n - 1$ taken from the standard normal distribution. Two passes through the data were required to get an accurate test of G since an outlier in the data not only creates a gap, $(\hat{\beta}_S - \hat{\beta}_L)$, in the normal plot of effects, but also inflates the PSE statistic.

By automating this procedure, its performance could be compared by simulation to other methods of automatically detecting outliers and performing robust tests of the effects. [Lawson and Gatlin \(2006\)](#) showed that this procedure was more powerful for detecting factorial effects and interactions in the presence of outliers than Torres' procedure based on analyzing the ranks of the response data; and, except for the case where the outlier was small relative to the size of the effects, it was more powerful than Box and Meyer's Bayesian method.

SAS macros have been written that implement the automated Daniel-Box procedure, for detecting outliers and performing a robust test in the presence of an outlier. These macros will be described in the next section and the final section will show examples of their use.

3. Description of the macros

This section briefly describes the SAS macros for automating Daniel's method of detecting an outlier and performing analysis of a 2^{k-p} design in the presence of an outlier.

- `%calc`

This macro calculates the regression coefficients and their absolute values, calculates normal scores for the half normal plot, and determines the significance of individual regression coefficients using the LGB test statistic ([Lawson, Grimshaw, and Burt 1998](#)). There are four required inputs for this macro. The first is a SAS data set, `x1`, which contains the saturated orthogonal \mathbf{X} matrix including a column of 1's as coefficients for the intercept term β_0 . The second is a SAS data set, `y`, that contains the column vector of responses. The third is a macro variable, `nobs`, that contains the number of observations or rows in the data sets `x1` and `y`. The fourth is a macro variable, `vname1`, which is a list variable containing names for each of the columns in `x1`. It is used for labeling output and graphs. Macro `%calc` produces no printed output but rather creates several SAS data sets that are used by the other macros.

- `%printeff`

This macro prints the variable labels in `vname1`, the regression coefficients, the absolute value of the regression coefficients, a yes-no indicator of whether each regression coefficient is significant at the $\alpha = 0.05$ significance level, and overall LGB statistic and the 95th percentile of its reference distribution.

- `%hlnplt`

This macro produces a half normal plot of the absolute regression coefficients, including the fitted line to the plot and an upper prediction limit line. The extreme points are labeled with their names in the macro variable `vname1`.

To use the three macros described above to make a half normal plot of the absolute regression coefficients, with the LGB fitted line and upper prediction limit, and to print a table of the coefficients and significance indicators, call the three macros above in sequence after creating the SAS data sets `x1` and `y`.

- `%gap`

This macro computes preliminary gap information from regression coefficients and ranks stored in temporary files by macro `%calc`. This macro makes no printed output but stores results in temporary files `gap3a` and `gap3b`.

- `%outlier`

This macro accesses the temporary file `effp14g` containing the regression coefficients created by `%calc`. It locates a potential outlier and calculates corrected response data that is in turn stored in a temporary file.

- `%testout(pass)`

This macro tests significance of the gap statistic. If `pass=1`, it compares the gap statistic to the 50th percentile of the reference distribution. If `pass=2`, it compares the gap statistic to the 99th percentile of the reference distribution, as described in [Lawson and Gatlin \(2006\)](#).

- `%pse`

This macro calculates [Lenth \(1989\)](#) *PSE* statistic based on the regression coefficients and associated statistics produced by `%calc`. The output is a SAS data set `pse`.

- `%gapmth`

This macro calls the other macros to perform the automated Daniel-Box method of locating and correcting outliers in unreplicated 16 run or 32 run 2^k or 2^{k-p} experiments. Macros `%calc`, `%printeff`, `%hlnplt`, `%pse`, `%gap`, `%outlier`, and `%testout` are called in succession. If the gap statistic calculated with the initial *PSE* estimate exceeds the 50th percentile of the reference distribution, it prints the initial outlier report and through a second pass recalculates the gap statistic with the corrected data produced by the macro `%outlier`. If the second pass gap statistic exceeds the 99th percentile of the reference distribution, the second pass outlier report is printed along with a table showing the corrected data. The macros `%printeff`, and `%hlnplt` are called again to print the a table of the regression coefficients calculated with the corrected data and make a half normal plot of the absolute coefficients. The details are described in [Lawson and Gatlin \(2006\)](#). Examples in the next section illustrate this.

4. Examples

This section presents some example data sets and illustrates the use of the macros.

4.1. Example 1

The first example illustrates the use of the macro `%gapmth` to analyze the data from an unreplicated 2^4 factorial experiment originally taken from [Box and Meyer \(1987\)](#) and analyzed later by [Box \(1991\)](#) and [Lawson and Gatlin \(2006\)](#). The data are shown in Table 1. A normal plot of the effects calculated with this data was shown in Figure 1.

Run	A	B	C	D	Y
1	-	-	-	-	47.46
2	+	-	-	-	49.62
3	-	+	-	-	43.13
4	+	+	-	-	46.31
5	-	-	+	-	51.47
6	+	-	+	-	48.49
7	-	+	+	-	49.34
8	+	+	+	-	46.10
9	-	-	-	+	46.76
10	+	-	-	+	48.56
11	-	+	-	+	44.83
12	+	+	-	+	44.45
13	-	-	+	+	59.15
14	+	-	+	+	51.33
15	-	+	+	+	47.02
16	+	+	+	+	47.90

Table 1: Example from [Box and Meyer \(1987\)](#).

The SAS commands first create the saturated matrix of contrast coefficients. Next, are the commands to read the response data, create the macro variables `nobs` and `vname1`, and call the macro `%gapmth` for analysis that are shown below.

```
*Create the design matrix of contrast coefficients for a 2^4 design;
data xmatrix16;
do x4=-1 to 1 by 2;
  do x3=-1 to 1 by 2;
    do x2=-1 to 1 by 2;
      do x1=-1 to 1 by 2;
        I=1; A=x1; B=x2; C=x3; D=x4;
        AB=A*B; AC=A*C; AD=A*D; BC=B*C; BD=B*D; CD=C*D;
        ABC=A*B*C; ABD=A*B*D; ACD=A*C*D; BCD=B*C*D; ABCD=A*B*C*D;
        output;
      end;
    end;
  end;
end;
drop x1-x4;
run;
```

```

*Read in the response data for 2^4 design;
data boxmey;
input y @@;
datalines;
47.46 49.62 43.13 46.31 51.47 48.49 49.34 46.10 46.76 48.56 44.83 44.45
59.15 51.33 47.02 47.90
*Defines the nobs and vname1 macro variables needed by the macro;
%let nobs=16;
%let vname1={'I' 'A' 'B' 'C' 'D'
'AB' 'AC' 'AD' 'BC' 'BD' 'CD'
'ABC' 'ABD' 'ACD' 'BCD' 'ABCD'};
*Invokes the macro;
data y; set boxmey;
data x1; set xmatrix16;
%gapmth

```

The first output is an analysis of the effects assuming no outliers. In the table below, it can be seen that no significant effects are found in the first pass through the data using the LGB statistic. A half normal plot was also produced, but it is not shown here.

Effect Report

Label	Effect	Abs(Effect)	Sig(at 0.05)
A	-0.4	0.4	no
B	-2.11	2.11	no
C	1.855	1.855	no
D	0.505	0.505	no
AB	0.455	0.455	no
AC	-1.245	1.245	no
AD	-0.29	0.29	no
BC	-0.4	0.4	no
BD	-0.59	0.59	no
CD	0.745	0.745	no
ABC	0.6	0.6	no
ABD	0.36	0.36	no
ACD	0.2	0.2	no
BCD	-0.79	0.79	no
ABCD	0.76	0.76	no

```

Lawson, Grimshaw & Burt Rn Statistic=    1
95th percentile of Rn= 1.201

```

The next two sections of output are printed by the macro %gapmth. First is the result of the testing the significance of the gap statistic $G = (\hat{\beta}_S - \hat{\beta}_L)/PSE)/(Z_S - Z_L)$ during the first pass through the data. This report was printed because the gap statistic exceeded the 50th percentile of its reference distribution.

Initial Outlier Report from test at 50th Percentile

Gap	PSE	Stand-Gap	Critical Value	Significance
0.49	0.885	3.35323	1.788	yes

Next is the report of testing the significance of the gap statistic in the second pass through the data. This report and all further reports are only printed when the gap statistic exceeds the 99th percentile of its reference distribution.

Final Outlier Report from test at 99th Percentile

Gap	PSE	Stand-Gap	Critical Value	Significance
0.49	0.225	13.1894	5.1	yes

The next report is the corrected data report. Here it can be seen that observation 13 was identified as an outlier, and the correction factor of 6.4 was subtracted from this observation to get the corrected value of 52.75.

Corrected Data Report

Response	Corrected Response	Detect Outlier
47.46	47.46	no
49.62	49.62	no
43.13	43.13	no
46.31	46.31	no
51.47	51.47	no
48.49	48.49	no
49.34	49.34	no
46.1	46.1	no
46.76	46.76	no
48.56	48.56	no
44.83	44.83	no
44.45	44.45	no
59.15	52.75	yes
51.33	51.33	no
47.02	47.02	no
47.9	47.9	no

The final page of output is the analysis of the effects computed from the data with the corrected observation. In this table it can be seen that the LGB statistic discovers contrasts B , C , and the AC interaction significant after correcting the data. Recall the AC interaction was not discovered in Figure 1 before the outlier was identified and corrected. Figure 2, shows a graphical representation of the results in the table below.

Effect Report

Label	Half Effect	Abs(Effect)	Sig(at 0.05)
A	0	0	no
B	-1.71	1.71	yes
C	1.455	1.455	yes
D	0.105	0.105	no
AB	0.055	0.055	no
AC	-0.845	0.845	yes
AD	0.11	0.11	no
BC	0	0	no
BD	-0.19	0.19	no
CD	0.345	0.345	no
ABC	0.2	0.2	no
ABD	-0.04	0.04	no
ACD	0.6	0.6	no
BCD	-0.39	0.39	no
ABCD	0.36	0.36	no

Lawson, Grimshaw & Burt Rn Statistic=1.626

95th percentile of Rn= 1.201

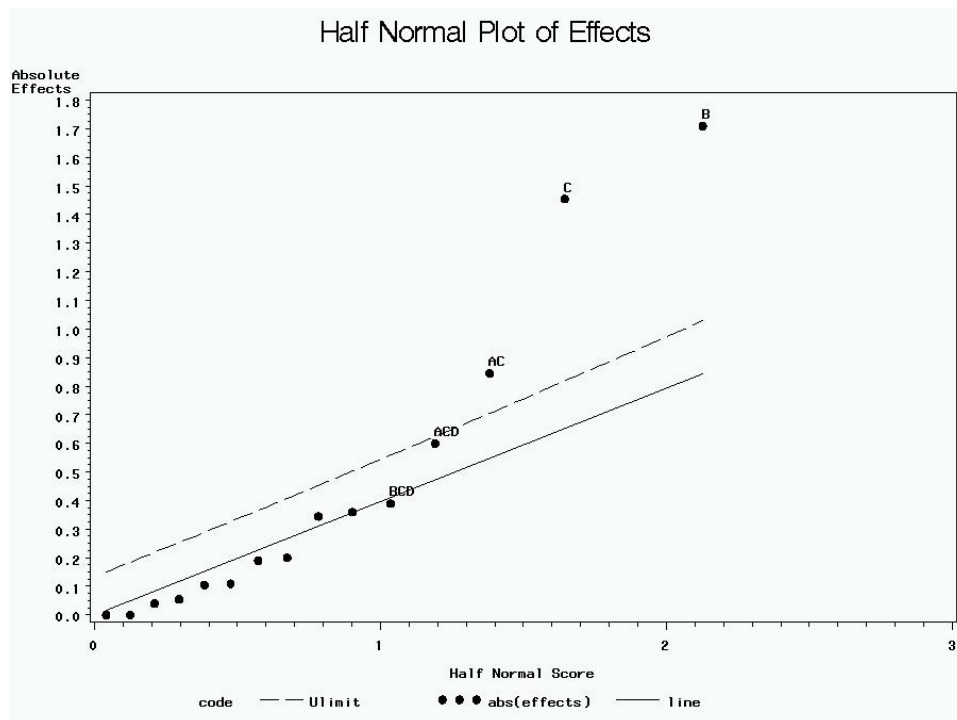


Figure 2: Half normal plot of regression coefficients after correcting data.

4.2. Example 2

The second example illustrates the use of the macro `%gapmth` to analyze data from an eight run fractional factorial experiment with an outlier. The data from this experiment is shown in Table 2.

Obs	A	B	C	D	E	F	AF	y
1	-	-	-	+	+	+	-	1.299
2	+	-	-	-	-	+	+	1.601
3	-	+	-	-	+	-	+	1.359
4	+	+	-	+	-	-	-	1.461
5	-	-	+	+	-	-	+	1.338
6	+	-	+	-	+	-	-	1.486
7	-	+	+	-	-	+	-	1.330
8	+	+	+	+	+	+	+	1.470

Table 2: 2^{6-3} fractional factorial with defining relation $I = ABD = ACE = BCF = BCDE = ACDF = ABEF = DEF$.

The SAS commands below create the saturated matrix of contrast coefficients and read in the response data. This is a resolution III design and the saturated set of contrasts is composed of the six main effects and one interaction contrast. In the commands to create the contrast matrix, one interaction AF is created to represent the last contrast in the saturated set since it is not confounded with any main effect. In the macro variable `vname1` this contrast was given the name `'AF+BE+CD'` since these three two factor interactions are confounded.

```
*test outlier in 8 run design;
Data xmatrix8;
do x3= -1 to 1 by 2;
  do x2 = -1 to 1 by 2;
    do x1 = -1 to 1 by 2;
      I=1; A=x1; B=x2; C=x3; D=A*B; E= A*C; F= B*C; AF=A*F;
      output;
    end;
  end;
end;
drop x1 x2 x3;
data ff;
input y @@;
datalines;
1.299 1.601 1.359 1.461 1.338 1.486 1.330 1.470
data y; set ff;
data x1; set xmatrix8;
%let nobs=8;
%let vname1={'I' 'A' 'B' 'C' 'D' 'E' 'F' 'AF+BE+CD'};
%gapmth
```

Below is the output created by %gapmth. Here it can be seen that in the analysis of the raw data, none of the contrasts were found to be significant using the LGB statistic. However, an outlier was detected at observation 2, and after this outlier is corrected, main effect *A* is found to be significant. Figure 3 shows a graphical representation of the results from the final table.

Effect Report

Label	Half Effect	Abs(Effect)	Sig(at 0.05)
A	0.0865	0.0865	no
B	-0.013	0.013	no
C	-0.012	0.012	no
D	-0.026	0.026	no
E	-0.015	0.0145	no
F	0.007	0.007	no
AF+BE+CD	0.024	0.024	no

Lawson, Grimshaw & Burt Rn Statistic=1.526
95th percentile of Rn= 1.534

Initial Outlier Report from test at 50th Percentile

Gap	PSE	Stand-Gap	Critical Value	Significance
0.019	0.0206	2.61015	1.788	yes

Final Outlier Report from test at 99th Percentile

Gap	PSE	Stand-Gap	Critical Value	Significance
0.019	0.0056	9.57057	5.1	yes

Corrected Data Report

Response	Corrected Response	Detect Outlier
1.299	1.299	no
1.601	1.508	yes
1.359	1.359	no
1.461	1.461	no
1.338	1.338	no
1.486	1.486	no
1.33	1.33	no
1.47	1.47	no

Effect Report

Label	Half Effect	Abs(Effect)	Sig(at 0.05)
A	0.0749	0.0749	yes
B	-0.001	0.0014	no
C	-37E-5	0.0004	no
D	-0.014	0.0144	no
E	-0.003	0.0029	no
F	-0.005	0.0046	no
AF+BE+CD	0.0124	0.0124	no

Lawson, Grimshaw & Burt Rn Statistic=2.437
 95th percentile of Rn= 1.534

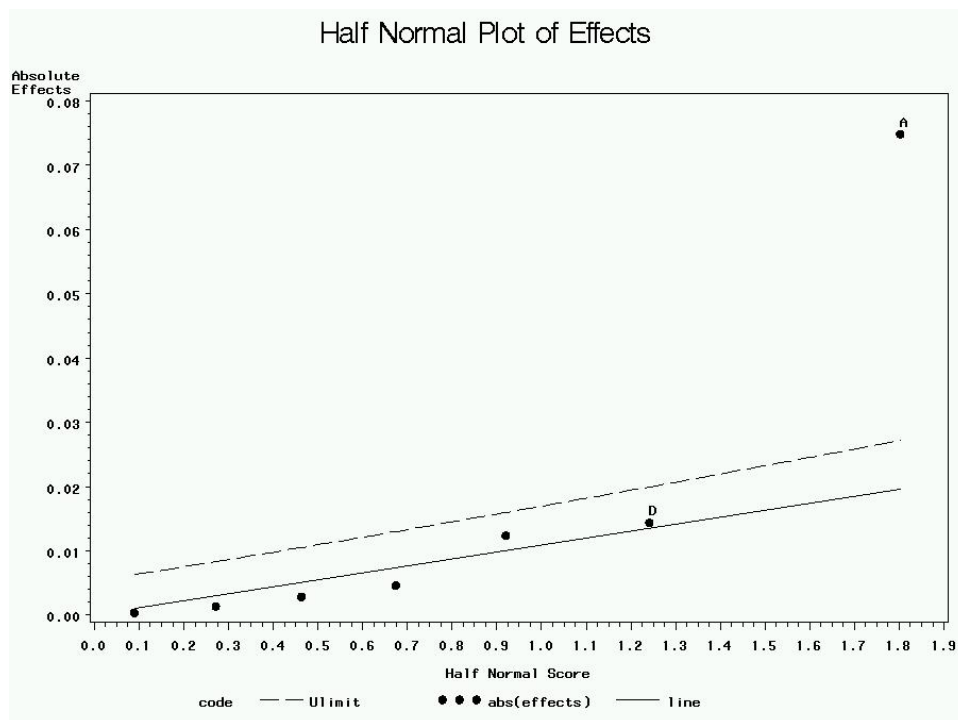


Figure 3: Half normal plot of regression coefficients after correcting data.

4.3. Example 3

The third example shows the use of the macro `%gapmth` to analyze the data from [Davies \(1971\)](#) 2^5 experiment to study the effects of concentrations of components of a nutrient medium on the yield of *Penicillium chrysogenum* grown in a surface culture. The data was reanalyzed by [Daniel \(1976\)](#) who identified an outlier in the 16th run. Daniel used $y = \text{yield} - 130$ as the

response and concentrated on the first 16 experiments where factor E was held constant at its low level. The data from these first 16 experiments are shown in Table 3.

Run	A	B	C	D	Y
1	-	-	-	-	12
2	+	-	-	-	-16
3	-	+	-	-	-1
4	+	+	-	-	-21
5	-	-	+	-	55
6	+	-	+	-	32
7	-	+	+	-	70
8	+	+	+	-	42
9	-	-	-	+	18
10	+	-	-	+	-22
11	-	+	-	+	16
12	+	+	-	+	-35
13	-	-	+	+	70
14	+	-	+	+	34
15	-	+	+	+	85
16	+	+	+	+	-12

Table 3: First 16 runs from Davies 2^5 on penicillin.

The SAS commands to read the data and define the macro variables `nobs` and `vname1` are shown below. The saturated matrix of contrast coefficients `xmatrix16` is the same as that created for Example 1.

```
*Davies 2^5 on Penicillin with E = - ;
%let nobs=16;
%let vname1={'I' 'A' 'B' 'C' 'D'
'AB' 'AC' 'AD' 'BC' 'BD' 'CD'
'ABC' 'ABD' 'ACD' 'BCD' 'ABCD'};
data Daviestp;
input y @@;
datalines;
12 -16 -1 -21 55 32 70 42 18 -22 16 -35 70 34 85 -12
data y; set Daviestp;
data x1; set xmatrix16;
%gapmth
```

The last two tables of the output are shown below. An outlier was detected at observation 16. This is the same observation Daniel (1976) determined to be a bad value after an insightful examination of the data. After correcting for the outlier, main effects A and C were found to be significant along with interactions AD and BC . The interactions were not found to be significant in the first pass through the data when the outlier was ignored, and the outlier

is not obvious from a plot of the residuals versus predicted or a normal plot of the residuals from a model involving the main effects and all two factor interactions (the default model fit by SAS ADX).

Corrected Data Report

Response	Corrected Response	Detect Outlier
12	12	no
-16	-16	no
-1	-1	no
-21	-21	no
55	55	no
32	32	no
70	70	no
42	42	no
18	18	no
-22	-22	no
16	16	no
-35	-35	no
70	70	no
34	34	no
85	85	no
-12	23.5	yes

Effect Report

Label	Half Effect	Abs(Effect)	Sig(at 0.05)
A	-17.97	17.969	yes
B	-0.219	0.2188	no
C	28.781	28.781	yes
D	1.0313	1.0313	no
AB	-2.094	2.0938	no
AC	-0.594	0.5938	no
AD	-5.594	5.5938	yes
BC	3.9063	3.9063	yes
BD	-1.094	1.0938	no
CD	0.6563	0.6563	no
ABC	-1.719	1.7188	no
ABD	-2.469	2.4688	no
ACD	-0.219	0.2188	no
BCD	-1.469	1.4688	no
ABCD	-0.094	0.0938	no

Lawson, Grimshaw & Burt Rn Statistic=3.108
95th percentile of Rn= 1.201

5. Conclusions

8 run and 16 run unreplicated 2^k and 2^{k-p} designs are widely used in research where the data is normally analyzed by user-friendly software packages. These packages include methods for detecting significant effects, but they do not include methods for detecting outliers or performing an analysis of the data in the presence of outliers. A single outlier can seriously reduce the power of detecting significant effects in an unreplicated design. The SAS macros illustrated in this paper automate the Daniel-Box method of analyzing an unreplicated 2^k and 2^{k-p} design with a single outlier. This method has been shown to be effective in simulation studies and is recommended whenever an experimenter may suspect an outlier.

References

- Box GEP (1991). "George's Column: Finding Bad Values in Factorial Designs." *Quality Engineering*, **3**, 249–254.
- Box GEP, Hunter WG, Hunter JS (1978). *Statistics for Experimenters*. John Wiley & Sons, New York.
- Box GEP, Meyer RD (1987). "Analysis of Unreplicated Factorials Allowing for Possibly Faulty Observations." In CL Mallows (ed.), "Design Data and Analysis by Some Friends of Cuthbert Daniel," pp. 1–12. John Wiley & Sons, New York.
- Carroll RJ (1980). "Robust Methods for Factorial Experiments with Outliers." *Applied Statistics*, **29**, 246–254.
- Daniel C (1959). "Use of Half-Normal Plots in Interpreting Factorial Two-Level Experiments." *Technometrics*, **1**, 311–341.
- Daniel C (1960). "Locating Outliers in Factorial Experiments." *Technometrics*, **2**, 149–156.
- Daniel C (1976). *Applications of Statistics to Industrial Experimentation*. John Wiley & Sons, New York.
- Davies OL (1971). *The Design and Analysis of Industrial Experiments*. John Wiley & Sons, New York.
- Gatlin JL (2004). *Performance of a Method of Identifying Outliers in Unreplicated Factorial Experiments*. Master's thesis, Brigham Young University. Department of Statistics.
- Hamada M, Balakrishnan N (1998). "Analyzing Unreplicated Factorial Experiments: A Review with Some New Proposals." *Statistica Sinica*, **8**, 1–41.
- Hintze J (2007). *NCSS 2007*. Kaysville, UT. URL <http://www.ncss.com/>.
- JMP (2008). "JMP 7." URL <http://www.jmp.com/>.
- John JA (1978). "Outliers in Factorial Experiments." *Applied Statistics*, **27**, 111–119.
- Lawson J, Gatlin J (2006). "Finding Bad Values in Factorials - Revisited." *Quality Engineering*, **18**, 491–501.

- Lawson J, Grimshaw S, Burt J (1998). “A Quantitative Method for Identifying Active Contrasts in Unreplicated Factorial Experiments Based on the Half Normal Plot.” *Computational Statistics and Data Analysis*, **26**, 425–436.
- Lenth RV (1989). “Quick and Easy Analysis of Un-replicated Factorials.” *Technometrics*, **31**, 469–473.
- Minitab Inc (2007). *Minitab, Version 15*. State College, PA. URL <http://www.minitab.com/>.
- SAS Institute (2003). *The SAS System, Version 9.1*. Cary, NC. URL <http://www.sas.com/>.
- Seheult A, Tukey JW (1982). “Some Resistant Procedures for Analyzing 2^n Factorial Experiments.” *Utilitas Mathematica*, **21B**, 57–98.
- Stat-Ease (2008). *Design Ease, Version 7.1*. URL <http://www.statease.com/>.
- StatPoint-Inc (2008). *Stat Graphics Centurion XV*. URL <http://www.statgraphics.com/>.
- Torres VA (1993). “A Simple Analysis of Unreplicated Factorials with Possible Abnormalities.” *Journal of Quality Technology*, **25**, 183–187.

Affiliation:

John Lawson
Department of Statistics
Brigham Young University
Provo, UT, United States of America
E-mail: lawson@byu.edu