Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 65(3) 2011 323-330

M-Estimation in Block Designs

Ranjit Kumar Paul[1*] and Lalmohan Bhar[2]
[1]*Central Inland Fisheries Research Institute, Kolkata*
[2]*Indian Agricultural Statistics Research Institute, New Delhi*

Received 26 August 2008; Revised 31 July 2010; Accepted 07 May 2011

SUMMARY

Data generated from the designed experiments is analyzed under the assumptions that the error distribution of observations is normal and homogeneous and data do not contain any outlier. If any of these assumptions is violated, the conclusion drawn from this analysis may be false. In the present paper various M-estimation procedures are applied to designed experiments. Efficiencies of these procedures are measured in terms of average variance. An example is given to illustrate the fact that application of robust method changed the conclusions drawn with analysis of original data. For computation of M-estimation, SAS codes are written in IML and given as Appendix.

Keywords: Block design, M-estimation, Outlier, Robust regression, SAS software package.

1. INTRODUCTION

Data generated from the designed experiment is analyzed under certain assumptions. If any of these assumptions is violated, the conclusion drawn from this analysis may be erroneous. For example, like many other fields data obtained from designed experiments is analyzed assuming that the error distribution of observations is normal and homogeneous. These assumptions are frequently violated in practice. In general many examples of such kind could be quoted in linear regression models. But in particular, it is also a common phenomenon in case of designed experiments. Recently conducted study by Indian Agricultural Statistics Research Institute, New Delhi (Parsad *et al*. 2004) revealed that many of the past experiments conducted in different parts of India have non-normal and heterogeneous distribution of error variances.

Apart from the problem of normality, the data set may contain some outlying observations. Outliers are likely to occur in the data generated from experimental designs due to disease and/or insect attack on some particular plot of the experiment, mistakes creeping in during recording of data, etc. The fact that a small subset of the data can have a disproportionate influence on the estimated parameters or predictions is of concern to users of regression analysis; it is quite possible that the model-estimates are based primarily on this data subset rather than on the majority of the data. If some of the observations are different in some way from the bulk of the data, the overall conclusion drawn from this data set may be wrong. In general, literature on outliers is very vast. A number of statistics are now developed to detect outliers in a data set following linear model. Bhar and Gupta (2001) developed some statistics for detecting outliers in designed experiments. They modified Cook statistic for its application to design of experiments, which is a follow up work of Cook (1977). Once an observation is detected as an outlier, the next question may arise what to do with this outlying observation? Should we discard this observation? Deletion of observation from the existing set is not always recommended. On the other hand, robust

Corresponding author : Ranjit Kumar Paul
E-mail address : ranjitstat@gmail.com

method of estimation is advocated to dampen the effect of an outlying observation. In case of linear regression models, robust regression method is now very popular to tackle the problem of non-normal error variance and the presence of outliers. This approach is designed to employ a fitting criterion that is not as vulnerable as least squares to unusual data. The most common general method of robust regression is M-estimation, introduced by Huber (1964). In this method, the objective function to be minimized to get the parameter estimates is weighted according to the residual of each observation. Literature on robust regression particularly on M-estimation is now vast. A good number of objective functions to be minimized are proposed. Most of these functions are non-linear in nature and therefore, normal equations for solving the parameter estimates are also non-linear in parameters. Iteratively Reweighted Least Squares (IRLS) (Holland and Welsch 1977) methods are employed to solve these equations.

However, not much work on these powerful methods in design of experiments is available in the literature. Carroll (1980) applied this technique to unreplicated factorial experiments and Chi (1994) to Cross-Over Trials. But no work seems to be available in case of block designs. In the present investigation an attempt is made to apply these methods to some existing data set after doing necessary modifications, wherever required. If outlier is present in the data set and we use the usual least squares method of analysis the problem that occur generally is that all the observations including the outlying observations get similar weight and the weight is unity. But if any observation is found to be outlier then it must get some lesser weight than the clean observations. This concept is utilized in the analysis of the design of experiments. For giving appropriate weight to different observations we have used the available functions of M-estimation that are more frequently used in the regression analysis. In block designs, we are generally interested in the estimation of some functions of some sub-set of parameters. This fact was kept in mind while applying this method. In Section 2 M-estimation procedure as applied in linear regression model is discussed followed by its application to designed experiments. In Section 3 M-estimation in designed experiments has been illustrated with an example. Relevant matters regarding M-estimation in designed experiments have been discussed in Section 4.

2. M-ESTIMATION

Consider the linear model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \tag{1}$$

where $\mathbf{y}$ is a $n \times 1$ vector of observations, $\mathbf{X}$ is an $n \times p$ design matrix of rank p and $\boldsymbol{\beta}$ is an $n \times 1$ vector of parameters. In general, we may define a class of robust estimators that minimize a function ρ of the errors, *i.e.*,

$$\underset{\beta}{\text{Minimize}} \sum_{i=1}^{n} \rho(e_i) = \underset{\beta}{\text{Minimize}} \sum_{i=1}^{n} \rho(y_i - \mathbf{x}_i'\boldsymbol{\beta}), \tag{2}$$

where $\mathbf{x}_i'$ denotes the i^{th} row of $\mathbf{X}$.

An estimator of $\boldsymbol{\beta}$ from this set up is called an M-estimator. If the method of least squares is used (implying the error distribution is normal), then $\rho(e_i) = (½)\ e_i^2$. Generally instead of $\rho(e_i)$, the function $\rho(e_i/\sigma)$ is minimized, where σ is a scale parameter.

IRLS method is used to obtain the parameter estimates. Suppose that an initial estimate $\hat{\boldsymbol{\beta}}_0$ is available and that s is an estimate of scale. Then the equations for solving for parameter estimates are given as

$$\sum_{i=1}^{n} x_{ij}\psi\left(\frac{y_i - \mathbf{x}_i'\boldsymbol{\beta}}{s}\right) = \sum_{i=1}^{n} \frac{x_{ij}\left\{\psi\left[\left(y_i - \mathbf{x}_i'\boldsymbol{\beta}\right)/s\right]\right\}/\left\{\left(y_i - \mathbf{x}_i'\boldsymbol{\beta}\right)/s\right\}}{\left(y_i - \mathbf{x}_i'\boldsymbol{\beta}\right)/s} = 0 \tag{3}$$

or $$\sum x_{ij} w_{i0}\left(y_i - \mathbf{x}_i'\boldsymbol{\beta}\right)/s = 0, \tag{4}$$

where $\psi = \rho'$, first derivative function of ρ and

$$w_{i0} = \frac{\psi\left[\left(y_i - \mathbf{x}_i'\hat{\boldsymbol{\beta}}_0\right)/s\right]}{\left(y_i - \mathbf{x}_i'\hat{\boldsymbol{\beta}}_0\right)/s} \quad \text{if } y_i \neq \mathbf{x}_i'\hat{\boldsymbol{\beta}}_0 \tag{5}$$

$$= 1 \quad \text{if } y_i = \mathbf{x}_i'\hat{\boldsymbol{\beta}}_0$$

In matrix notation equation (4) becomes

$$\mathbf{X}'\mathbf{W}_0\mathbf{X}\boldsymbol{\beta} = \mathbf{X}'\mathbf{W}_0\mathbf{y}, \tag{6}$$

where $\mathbf{W}_0$ is an $n \times n$ diagonal matrix of "weights" with diagonal elements $w_{10}, w_{20}, \ldots, w_{n0}$ given by equation (5). Now one step estimator is

$$\hat{\boldsymbol{\beta}}_1 = (\mathbf{X}'\mathbf{W}_0\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}_0\mathbf{y}. \tag{7}$$

At the next step the weights are recomputed from equation (5) but using $\hat{\boldsymbol{\beta}}_1$ instead of $\hat{\boldsymbol{\beta}}_0$. This process is continued till the convergence criterion is met.

A number of objective functions that are applied in linear regression model are available in the literature. For example Huber's function, Andrew's function, Hampel's function and Ramsay's function are widely used functions. All these functions have been applied in case of designed experiments. A comparison of the efficiency of these functions has also been done. Some commonly used functions are presented in the Table 1.

In case of M-estimation a number of estimators of σ^2 are proposed. The commonly used estimate of the error mean square is taken as (Huber 1973)

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} w_i\{(y_i - \mathbf{x}_i'\hat{\boldsymbol{\beta}})/s\}^2}{n - \delta(\mathbf{X})}, \qquad (8)$$

where d **(X)** is the rank of **X** matrix.

2.1 M-estimation in Designed Experiments

The model for general block design is given by

$$\mathbf{y} = \Delta'\boldsymbol{\tau} + \mu\mathbf{1} + \mathbf{D}'\theta + \varepsilon \qquad (9)$$

where **y** is a $n \times 1$ vector of observations, Δ' is $n \times v$ incidence matrix of treatments, $\boldsymbol{\tau}$ is a $v \times 1$ vector of treatment effects, $\mathbf{D}'$ is a $n \times b$ incidence matrix of blocks, θ is a $b \times 1$ vector of block effects, **1** is a unit vector of order $n \times 1$ and ε is a $n \times 1$ vector of errors.

We now write down $\mathbf{X} = [\mathbf{X}_1\ \mathbf{X}_2]$, where $\mathbf{X}_1 = \Delta'$ and $\mathbf{X}_2 = [\mathbf{1}\ \mathbf{D}']$ Similarly, $\beta = \begin{bmatrix} \boldsymbol{\tau} \\ \mu \\ \boldsymbol{\theta} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix}$, where

$\beta_1 = [\boldsymbol{\tau}]$ and $\beta_2 = \begin{bmatrix} \mu \\ \boldsymbol{\theta} \end{bmatrix}$.

Now under M-estimation, the normal equations for estimating the parameters in designed experiments are given as

$$\mathbf{X}'\mathbf{W}\mathbf{X}\beta = \mathbf{X}'\mathbf{W}\mathbf{y}$$

Table 1. Commonly used objective functions

Criterion	$\rho(z)$	$\psi(z)$	$w(z)$	Range
Least squares	$(1/2)z^2$	z	1.0	$\lvert z \rvert < \infty$
Huber's t function	$(1/2)z^2$	z	1.0	$\lvert z \rvert \le t$
	$\lvert z \rvert t - (1/2)t^2$	t sign(z)	$t / \lvert z \rvert$	$\lvert z \rvert > t$
Ramsay's function	$a^{-2}\left[\frac{1-\exp(-a\lvert z\rvert)}{(1+a\lvert z\rvert)}\right]$	$z \exp(-a\lvert z \rvert)$	$\exp(-a\lvert z \rvert)$	$\lvert z \rvert < \infty$
Andrews' wave function	$a[1-\cos(z/a)]$	Sin (z/a)	Sin $(z/a)/(z/a)$	$\lvert z \rvert \le a\pi$
	$2a$	0	0	$\lvert z \rvert > a\pi$
	$(½)z^2$	z	1.0	$\lvert z \rvert \le a$
Hampel's function	$a\lvert z \rvert - (½)a^2$	a sign (z)	$a/\lvert z \rvert$	$a < \lvert z \rvert \le b$
	$(a(c\lvert z \rvert - (½)z^2))/(c-b) -$ $7a^2/6$	$(a$ sign $(z)(c-$ $\lvert z \rvert))/(c-b)$	$a(c-\lvert z \rvert)/$ $(\lvert z \rvert (c-b))$	$b < \lvert z \rvert \le c$
	$a(b+c-a)$	0	0	$\lvert z \rvert > c$

Source: Montgomery, Peck and Vining (2001).

Note: $\rho(z)$ denotes the function of residuals, $\psi(z)$ denotes derivative of $\rho(z)$ and $w(z)$ denotes the weight.

$$\begin{bmatrix} \mathbf{X}_1' \mathbf{W} \mathbf{X}_1 & \mathbf{X}_1' \mathbf{W} \mathbf{X}_2 \\ \mathbf{X}_2' \mathbf{W} \mathbf{X}_1 & \mathbf{X}_2' \mathbf{W} \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1'\mathbf{W}\mathbf{y} \\ \mathbf{X}_2'\mathbf{W}\mathbf{y} \end{bmatrix}, \quad (10)$$

$$\mathbf{X}_1' \mathbf{W}\mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_1' \mathbf{W}\mathbf{X}_2\boldsymbol{\beta}_2 = \mathbf{X}_1' \mathbf{W}\mathbf{y} \quad (11)$$

$$\mathbf{X}_2' \mathbf{W}\mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2' \mathbf{W}\mathbf{X}_2\boldsymbol{\beta}_2 = \mathbf{X}_2' \mathbf{W}\mathbf{y}. \quad (12)$$

From (11), we can get $\boldsymbol{\beta}_2 = (\mathbf{X}_2' \mathbf{W}\mathbf{X}_2)^{-}[\mathbf{X}_2' \mathbf{W}\mathbf{y} - \mathbf{X}_2' \mathbf{W}\mathbf{X}_1\boldsymbol{\beta}_1]$, where $\mathbf{A}^{-}$ is a g-inverse of $\mathbf{A}$, *i.e.*, $\mathbf{A}\mathbf{A}^{-}\mathbf{A} = \mathbf{A}$.

Substituting this $\boldsymbol{\beta}_2$ in (12), we get

$$\mathbf{X}_1' \mathbf{W}\mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_1' \mathbf{W}\mathbf{X}_2(\mathbf{X}_2' \mathbf{W}\mathbf{X}_2)^{-} [\mathbf{X}_2' \mathbf{W}\mathbf{y} - \mathbf{X}_2' \mathbf{W}\mathbf{X}\boldsymbol{\beta}_1] = \mathbf{X}_1' \mathbf{W}\mathbf{y},$$

or

$$[\mathbf{X}_1' \mathbf{W}\mathbf{X}_1 - \mathbf{X}_1' \mathbf{W}\mathbf{X}_2(\mathbf{X}_2' \mathbf{W}\mathbf{X}_2)^{-}\mathbf{X}_2' \mathbf{W}\mathbf{X}_1]\boldsymbol{\beta}_1 = \mathbf{X}_1' \mathbf{W}\mathbf{y} - \mathbf{X}_1' \mathbf{W}\mathbf{X}_2(\mathbf{X}_2' \mathbf{W}\mathbf{X}_2)^{-}\mathbf{X}_2' \mathbf{W}\mathbf{y}$$

The above equations can be written as

$$\mathbf{C}\boldsymbol{\beta}_1 = \mathbf{Q}, \quad (13)$$

where $\mathbf{C} = [\mathbf{X}_1' \mathbf{W}\mathbf{X}_1 - \mathbf{X}_1' \mathbf{W}\mathbf{X}_2(\mathbf{X}_2' \mathbf{W}\mathbf{X}_2)^{-}\mathbf{X}_2' \mathbf{W}\mathbf{X}_1]$ (14)

and $\mathbf{Q} = \mathbf{X}_1' \mathbf{W}\mathbf{y} - \mathbf{X}_1' \mathbf{W}\mathbf{X}_2(\mathbf{X}_2' \mathbf{W}\mathbf{X}_2)^{-}\mathbf{X}_2' \mathbf{W}\mathbf{y}$. (15)

In case of design of experiments, we are generally interested in estimation of treatment contrasts. Let **P** be a (v-1) × v matrix of all (v-1) set of elementary treatment contrasts. Then the M-estimates of this set of contrasts is given as

$$\mathbf{P}\hat{\boldsymbol{\beta}}_1 = \mathbf{P}\hat{\boldsymbol{\tau}} = \mathbf{P}\mathbf{C}^{-}\mathbf{Q}, \quad (16)$$

The variance of $\mathbf{P}\hat{\boldsymbol{\tau}}$ is given as Var($\mathbf{P}\hat{\boldsymbol{\tau}}$) = $\hat{\sigma}^2 \mathbf{P}\mathbf{C}^{-}\mathbf{P}'$, where $\hat{\sigma}^2$ is obtained from equation (8).

In linear model the sensitivity of the classical least squares estimates to departures from normality, such as possible presence of outliers, has led to various proposals for robust methods of estimation. Parameter estimation is usually only a first step in the analysis of data arising from a linear model. A classical least squares analysis often focused upon the analysis of variance, which tests simultaneous hypothesis in large subsets of the parameters. Since the terms in a classical analysis of variance are quadratic forms in least squares estimates, one would expect that the sensitivity of the estimates to departures from normality should be inherited by the tests. In fact, for moderate to heavy tailed distributions or in the presence of outliers, it appears that the classical F test does lose power. In view of this fact, many attempts have been made to develop appropriate procedures for testing linear hypotheses. For detail on these testing procedures one may refer Bickel (1976) and Schrader and Hettmansperger (1979).

However, the most natural test is based on the M-estimates directly. In this method first $\hat{\boldsymbol{\beta}}$ is derived from the IRLS algorithm and then by using the final configuration of the weights as fixed and given a priori, a least squares weighted analysis of variance is done. In the present study this approach, *i.e.*, weighted analysis of variance has been adopted.

Now the contrasts sum of squares are given by

$$(\mathbf{P}\hat{\boldsymbol{\tau}})' (\mathbf{P}\mathbf{C}^{-}\mathbf{P}')^{-} (\mathbf{P}\hat{\boldsymbol{\tau}})$$

And the test statistic for testing the significance of the contrasts is

$$\mathrm{F} = \frac{(\mathbf{P}\,\hat{\boldsymbol{\tau}})' (\mathbf{P}\,\mathbf{C}^{-}\,\mathbf{P}')^{-} (\mathbf{P}\,\hat{\boldsymbol{\tau}})}{\hat{\sigma}^2} \rightarrow F_{v-1, error\ df}$$

For testing a particular elementary contrast, say $\mathbf{p}_i' \hat{\boldsymbol{\tau}}$ where $\mathbf{p}_i'$ is the i^{th} row of contrast matrix, the test statistic is

$$\mathrm{F} = \frac{(\mathbf{p}_i' \hat{\boldsymbol{\tau}})' (\mathbf{p}_i' \mathbf{C}^{-} \mathbf{p}_i)^{-} (\mathbf{p}_i' \hat{\boldsymbol{\tau}})}{\hat{\sigma}^2} \rightarrow F_{1, error\ df}$$

The analysis of variance table is given as follows:

Analysis of Variance

Sources of variation	DF	SS
Treatment (adjusted)	v-1	$\mathbf{Q}'\mathbf{C}^{-}\mathbf{Q}$
Block (unadjusted)	b-1	$\mathbf{B}'\mathbf{K}^{-1}\mathbf{B}$ – CF
Error	n-v-b+1	$\mathbf{y}'\mathbf{y} - \mathbf{B}'\mathbf{K}^{-1}\mathbf{B} - \mathbf{Q}'\mathbf{C}^{-}\mathbf{Q}$
Total	n-1	$\mathbf{y}'\mathbf{y}$ – CF

where $\mathbf{B} = \mathbf{D}\mathbf{W}^{1/2}\mathbf{y}$, CF = $(\mathbf{1}'\mathbf{W}^{1/2}\mathbf{y})^2/n$ and $\mathbf{W}^{1/2}$ is a Grammian square root of **W**. A program is written in SAS IML to carry out the analysis. This program is given as an Appendix.

3. ILLUSTRATION

An experiment with 10 treatments was carried out in the randomized complete block (RCB) design with 3 replications at G.K.V.K., Bangalore, India with a view to study the integrated weed management in cowpea (Net plot size: 3.60m × 2.80m). The treatments of the experiment are as follows:

10 Weeding treatments:

T0=Weedy check

T1=Weed free

T2=Sowing at 30 cm row spacing

T3=0.75 Kg a.i/ha of pendimethalin

T4=1.00 Kg a.i/ha of pendimethalin

T5=1.25 Kg a.i/ha of pendimethalin

T6=Hand weeding at 3 weeks after sowing (w.a.s.)

T7=Interculturing at 3 w.a.s

T8=T3+ Hand weeding at 3 w.a.s

T9=T3+ Interculturing at 3 w.a.s

The data on grain yield per plot in quintals for different treatments is given in Table 2

Table 2. Yield of cowpea in quintal/plot

Treatments	Replication		
	1	2	3
1	0.36	0.68	1.52
2	1.35	1.50	1.35
3	1.15	1.31	0.48
4	0.97	1.10	0.59
5	1.15	1.40	1.05
6	0.75	1.25	0.80
7	0.88	1.30	0.67
8	0.80	1.15	0.60
9	1.10	1.45	1.41
10	0.95	1.72	0.98

We first conducted usual analysis of the data. The analysis of variance table is given in Table 3. From the table it is observed that the treatment effects are not significant even at 5% level of significance, whereas block effects are significant at 5% level of significance.

Table 3. Analysis of variance with the original data

Sources of variation	DF	SS	MS	F	Significance Level
Treatment	9	1.136	0.126	1.50	0.223
Block	2	0.772	0.386	4.58	0.024
Error	18	1.520	0.084		
Total	29	3.428			

By using the Cook statistic for detection of outlying observation, as given by Bhar and Gupta (2001) we see that the observation corresponding to the treatment number 1 and block number 3 is an outlier. The value of Cook-statistics is given in Table 4. This outlying observation is deleted and analysis of variance is done again. The result is presented in Table 5. It is now seen that the treatment effects are now significant at 5% level of significance.

Table 4. Cook-statistics

Yield	Treatment	Replication	Cook-Statistic
0.36	1	1	0.1586
1.35	2	1	0.0043
1.15	3	1	0.0878
0.97	4	1	0.0422
1.15	5	1	0.0043
0.75	6	1	0.0054
0.88	7	1	0.0020
0.80	8	1	0.0043
1.10	9	1	0.0125
0.95	10	1	0.0259
0.68	1	2	0.1757
1.50	2	2	0.0176
1.31	3	2	0.0116
1.10	4	2	0.0002
1.40	5	2	0.0007
1.25	6	2	0.0088
1.30	7	2	0.0165
1.15	8	2	0.0058
1.45	9	2	0.0103
1.72	10	2	0.0837
1.52	1	3	**0.6684**
1.35	2	3	0.0044
0.48	3	3	0.1634
0.59	4	3	0.0365
1.05	5	3	0.0014
0.80	6	3	0.0004
0.67	7	3	0.0302
0.60	8	3	0.0202
1.41	9	3	0.0456
0.98	10	3	0.0165

Table 5. Analysis of variance after deleting the outlier

Sources of variation	DF	SS	MS	F	Significance Level
Treatment	9	1.784	0.198	6.69	0.0004
Block	2	0.920	0.460	15.53	0.0001
Error	17	0.503	0.029		
Total	28	3.207			

3.1 Robust Analysis through M–estimation

We now applied various M-estimation procedures to this data and obtained the analysis of variance. Different values of the tuning constants used in the objective functions are determined as follows.

For Huber's function the value of the constant $t = 1.5$

For Andrew's function $a = 1.339$

For Ramsay's function $a = 0.3$

For Hampel's function $a = 1.7$, $b = 3.4$, $c = 8.5$

3.2 Weighted Analysis of Variance

The results are given in the following tables.

Table 6. Analysis of variance (Huber's function)

Sources of variation	DF	SS	MS	F	Significance Level
Treatment	9	1.505	0.167	3.866	0.0070
Block	2	1.007	0.503	11.642	0.0005
Error	18	0.778	0.043		
Total	29	3.290			

Average variance for the set of treatment contrasts is obtained as 0.034.

Table 7. Analysis of variance (Andrew's function)

Sources of variation	DF	SS	MS	F	Significance Level
Treatment	9	1.668	0.185	3.758	0.0080
Block	2	1.296	0.648	13.141	0.0003
Error	18	0.887	0.049		
Total	29	3.851			

Average variance for the set of treatment contrasts is obtained as 0.043.

Table 8. Analysis of variance (Ramsay's function)

Sources of variation	DF	SS	MS	F	Significance Level
Treatment	9	1.562	0.173	3.604	0.0090
Block	2	1.288	0.644	13.416	0.0002
Error	18	0.868	0.048		
Total	29	3.718			

Average variance for the set of treatment contrasts is obtained as 0.040.

Table 9. Analysis of variance (Hampel's function)

Sources of variation	DF	SS	MS	F	Significance Level
Treatment	9	1.659	0.184	3.408	0.0120
Block	2	1.350	0.675	12.487	0.0003
Error	18	0.973	0.054		
Total	29	3.982			

Average variance for the set of treatment contrasts is obtained as 0.046.

It is clear from the above tables that due to presence of outliers the analysis with the original data set resulted in wrong conclusions. That means the treatment effects are non significant in the least squares analysis with the original data. Whereas due to application of M-estimation in the original data we have seen that the analysis results in a valid conclusion.

4. DISCUSSION

Different M-estimation procedures as available in the literature were applied to the above data. In this data set treatment effects are not significant with the original data. But robust analysis revealed that the treatment effects are actually significant at 5% level of significance. That is, inferences to be drawn are reversed through robust analysis. Actually this experiment contains outliers. Similarly, in regard to the significance of the elementary contrasts, as with the original data the treatment effects are non significant so the contrasts are also non significant. But after deletion of outlying observation from the data set if we go for contrast analysis, we see that almost every pair of independent elementary contrasts are highly significant. The result of contrast analysis is nearly same for all the functions adopted for the robust analysis of the designed experiments.

The reasons for giving wrong results through original data might be the presence of outlying observations. Once these outlying observations are deleted, results tallying with that obtained through robust analysis. However, from statistical point of view it is not advised to delete any observation, because every observation carries some information that should be exploited. Robust analysis actually gives small weights to those outlying observations, thus extracting some information from that observation.

However, a question may arise in our mind that which M-estimation procedure should we use? It is difficult to answer this question, because these procedures depend on weights and weights are determined by observations. From experiences of analyzing a good number of experiments, it is observed that Huber's function performs well. This is observed from the fact that the average variance of the set of elementary treatment contrasts is small for most of the experiments for this function comparing to other M-estimation functions. However this is empirically true, there is no theoretical evidence.

In those experiments where no outlier is present, there may be little difference between the analysis with original data and analysis through robust regression. Off course the levels of significance may be changed a little bit. It is therefore, generally advised to carryout analysis through ordinary least squares (OLS), if we are sure that the data do not contain any outlying observation and the errors are normal, because OLS estimates posses some good statistical properties. But in general we do not know the form of distribution of the errors in advance. It is therefore, suggested to apply robust analysis always. Even the error distribution is normal; we may not loose much efficiency.

ACKNOWLEDGEMENT

Authors are grateful to the referee for the valuable suggestions.

REFERENCES

Bhar, L. and Gupta, V.K. (2001). A useful statistic for studying outliers in experimental designs. *Sankhya,* **B63**, 338-350.

Bickel, P.J. (1976). Another look at robustness. *Scand. J. Statist*., **3**, 145-148.

Carroll, R.J. (1980). Robust methods for factorial experiments with outliers. *Appl. Stat.*, **29**, 246-251.

Chi, E.M. (1994). M-estimation in cross-over trials. *Biometrics*, **50**, 486-493.

Cook, R.D. (1977). Detection of influential observation in linear regression. *Technometrics*, **19**, 15-18.

Holland, P.W. and Welsch, R.E. (1977). Robust regression using iteratively reweighted least squares. *Comm. Statist.-Theory Method*, **A6**, 813-827.

Huber, P.J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.*, **35**, 73-101.

Huber, P.J. (1973). Robust regression: Asymptotics, conjectures, and Monte carlo. *Ann. Stat*., **1**, 799-821.

Montgomery, D.C., Peck, E.A. and Vining, G.G. (2001). *Introduction to Linear Regression Analysis.* 3rd edition, John Wiley and Sons, Inc, New York.

Parsad, R., Gupta, V.K., Srivastava, R., Batra, P.K., Kaur, A. and Arya, P. (2004). *A diagnostic study of design and analysis of field experiments*. Technical Report, Indian Agricultural Statistics Research Institute, New Delhi.

Schrader, R.M. and Hettmansperger, T.P. (1979). Robust analysis of variance based upon a likelihood ratio criterion. *Biometrika*, **67**, 93-102.

APPENDIX

SAS Code for computing different M-estimation

```
data ran;
input trt blk y;
datalines;
;
proc iml;
    use ran;
            read all into X;
                    mrowX=nrow(X);
                    ncolX=ncol(X);
                    v=max(X[,1]);
            x1=j(mrowX,v,0);
    X2=j(mrowX,1,1);
    y=X[,ncolX];
/*Create Delta*/
do i=1 to mrowX;
    x1[i,X[i,1]]=1;
end;
/*create the matrix for the nuisance factor*/
do j=2 to ncolX-1;
    order=max(X[,j]);
    D1=J(mrowX,order,0);
    intB=max(X[,2]);
/*  create the diagnal matrix of block size*/
    K=J(intB,intB,0);
    do i=1 to mrowX;
            D1[i,X[i,j]]=1;
            K[X[i,2],X[i,2]]=K[X[i,2],X[i,2]]+1;
    end;
    X2=X2||D1;
end;
x=x1||x2;
betahat0=j(ncol(x),1,0);
betahat0=ginv(x'*x)*x'*y;
yhat=x*betahat0;
resid=y-yhat;
n=nrow(resid);
rank=round(trace(ginv(x)*x));
s1=median(abs(resid-median(resid)))/0.6745;
/*calculation of the standarized residuals*/
z1=j(n,1,0);
z1=resid/s1;
a=1;
    w0=j(n,1,0);
    w=j(n,n,0);
/*calculation of weight matrix*/
do while(a>0.005);
    do i=1 to n;
            z1[i,1]=abs(z1[i,1]);
            if z1[i,1]<=2 then
            w0[i,1]=1;
            else w0[i,1]=2/z1[i,1];
    end;
    w=diag(w0);
    betahat1=j(ncol(x),1,0);
    betahat1=ginv(x'*w*x)*(x'*w*y);
    yhat1=x*betahat1;
    resid2=(y-yhat1);
    s2=median(abs(resid2-median(resid2)))/
    0.6745;
    z2=(y-yhat1)/s2;
    a=abs((s2-s1)/s1);
    z1=z2;
    s1=s2;
end;
ww=sqrt(w);
C=(x1'*w*x1)(x1'*w*x2)*(ginv(x2'*w*x2))*(x2'*w*x1);
Q=x1'*w*yx1'*w*x2*(ginv(x2'*w*x2))*x2'*w*y;
/*Q is the treatment total vector*/
tauhat=ginv(c)*Q;
TSS=Q'*ginv(c)*Q;
B=D1'*ww*Y;/*B is the block total vector*/
one=j(n,1,1);
cf=(one'*ww*y)*(one'*ww*y)/n;
BSS=B'*inv(k)*B-cf;
bms=bss/(intB-1);
TMS= TSS/(v-1);
totss=(y'*w*y)-cf;
ess=totss-tss-bss;
ems=ess/(n-rank);
FT=TMS/ems;
FB=bms/ems;
pvalt=1 - probf(ft,v-1,n-rank);
pvalb=1 - probf(fb,intB-1,n-rank);
print "Huber function data set" 25;
print pvalt pvalb ft fb;
/*calculation of the all possible elementary
contrasts*/
p=j((v-1),v,0);
do i=1 to v-1;
    p[i,1]=1;
    j=i+1;
    p[i,j]=-1;
end;
print p;
conout=j((v-1),4,0);
contss=j((v-1),1,0);
do i=1 to nrow(p);
pi=p[i,];
/*calculation of t statistics*/
contst=((pi*tauhat)'*ginv((pi*ginv(c)*pi'))*(pi*tauhat))/
ems;
conpval=1 - probf(contst,1,n-rank);
conout[i,3]=contst;
conout[i,4]=conpval;
end;
varcon=j(v-1,v-1,0);
varcon=p*ginv(c)*p'*ems;
abgvar=trace(varcon)/(v-1);
print'    trt       vs' 'trt ' ' F Value    ' 'Pr    > F ';
print varcon abgvar;
run;
quit;
```