Robust Designs for The One-way Random Effects Model Using Q-estimators

by

Xiaolong Yang
B.Sc., Nankai University, 1999

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Mathematics and Statistics

University of Victoria

Robust Designs for The One-way Random Effects Model Using Q-estimators

by

Xiaolong Yang
B.Sc., Nankai University, 1999

Supervisory Committee

Dr. J. Zhou, (Department of Mathematics and Statistics)

Supervisor

Dr. W. J. Reed, (Department of Mathematics and Statistics)

Departmental Member

Dr. M. Tsao, (Department of Mathematics and Statistics)

Departmental Member

Dr. M. Wilmut, (School of Earth and Ocean Sciences)

Outside Member

Additional Member

Supervisory Committee

Supervisor

Dr. J. Zhou, (Department of Mathematics and Statistics)

Departmental Member

Dr. W. J. Reed, (Department of Mathematics and Statistics)

Departmental Member

Dr. M. Tsao, (Department of Mathematics and Statistics)

Outside Member

Dr. M. Wilmut, (School of Earth and Ocean Sciences)

Additional Member

ABSTRACT

Robust statistics is an extension of classical parametric statistics, which provides a safeguard against gross errors in experiments. Effectively, robustness properties of Uhlig's Q-estimators are examined and compared with that of Rocke's M-estimators. In particular, the finite-sample implosion and explosion breakdown points are investigated and introduced into constructing robust designs for the one-way random effects model. Optimal robust designs based on Uhlig's Q-estimation are similar to the ones based on Rocke's M-estimation. Ultimately, robust estimation procedures would provide steady and reliable estimates of model parameters in case of the occurrence of outliers.

Contents

List of Tables

List of Figures

Acknowledgements

First, I thank the University of Victoria and the Department of Mathematics and Statistics for providing a stimulating and supportive research environment.

I express my sincere thanks to my supervior, Dr. J. Zhou, for her unrelenting support, guidance, encouragement, and patience extending through my graduate study at the University of Victoria. Her contributions to the completion of my thesis is invaluable.

I thank Dr. W. J. Reed, Dr. M. Tsao, and Dr. M. Wilmut for their constructive comments and suggestions, especially for Dr. M. Wilmut's effective and helpful proofreading assistance. I am also grateful to Jessica Natale for sharing her comprehensive expertise of English writing with me.

Finally, special thanks are due my well-beloved parents and friends for their continued moral support. I know they are always behind me.

Chapter 1

Introduction

During the past decades, extensive research has been conducted on the methods of estimation of model parameters under random effects models. In this chapter, we review several classical methods of estimation and associated experimental designs for the one-way random effects model. We also briefly introduce two types of robust estimators: Rocke's M-estimators and Uhlig's Q-estimators. A thorough study of the two types of estimators will be given in later chapters.

1.1 One-way Random Effects Model

The one-way random effects model has been widely used to study the performance of different test groups, processes, and systems in experimental design. A group, sometimes called a treatment or factor, is a set of relatively homogeneous experimental conditions. The one-way random effects model appears as following:

$$y_{ij} = \mu + \alpha_i + e_{ij} \qquad i = 1, \cdots, I, \qquad j = 1, \cdots, J, \tag{1.1}$$

where y_{ij} is the j^{th} observation from the i^{th} group; μ is a parameter common to all groups (called the overall mean); α_i is a parameter unique to the i^{th} group; e_{ij} is a model error component that incorporates all other sources of variability in the experiment. We usually assume that the α_is and e_{ij}s are independently and normally distributed with mean zero and variance σ_a^2 and σ_e^2, respectively. The variances σ_a^2 and σ_e^2 are called variance components under the model.

In the form of model (1.1), the group is said to be a random factor because all of the group levels are chosen at random, and knowledge about the particular levels is relatively useless for further analysis. Here are some examples: (I) An experiment to determine the concentration of one component in a liquid mixture may require several technicians to perform tests independently, and each technician must implement the test procedure to obtain a number of readings on the concentration. The readings from different technicians may show the variability given the variability of each technician's skills. If one is interested in this effect when analyzing the experimental results, one would consider the technician as a random factor. (II) One pharmaceutical experiment may need a large number of runs, and only a certain number of runs can be made in one day. "Days" could represent a potential source of variability because the effect of different days may account for the variability of test results if the runs are sensitive to weather conditions. In this situation, one could consider the day as a random factor. (III) In an inter-laboratory study determining the amount of nicotine contained on cigarette filter tips, each of experimental samples is divided into sub-samples and sent to different laboratories for tests. Since the laboratories will usually employ different staff, equipments, and techniques, the laboratory can thus be treated as a random factor.

The purpose of experiments under the one-way random effects model is mainly to estimate the overall mean μ. Statisticians are also interested in studying the two variance components σ_a^2 and σ_e^2, since the inference for μ depends, to a certain extent, on knowledge of σ_a^2 and σ_e^2.

1.2 The Analysis of Variance and Optimal Designs

The analysis of variance (ANOVA) is one of the most commonly used methods to estimate variance components σ_a^2 and σ_e^2 under the one-way random effects model. The basic idea of the ANOVA is to partition the total variability of the data into a component that measures the variation between groups and a component that measures the variation within groups. The overall variability is measured by the total corrected sum of squares

$$SST = \sum_{i=1}^{I} \sum_{j=1}^{J} (y_{ij} - \bar{y}_{..})^2 , \tag{1.2}$$

where $\bar{y}_{..}$ represents the overall average of all the observations. Note that SST may be written as

$$\begin{aligned} \sum_{i=1}^{I} \sum_{j=1}^{J} (y_{ij} - \bar{y}_{..})^2 &= \sum_{i=1}^{I} \sum_{j=1}^{J} [(\bar{y}_{i.} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.})]^2 \\ &= J \sum_{i=1}^{I} (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^{I} \sum_{j=1}^{J} (y_{ij} - \bar{y}_{i.})^2 , \end{aligned}$$

where $\bar{y}_{i.}$ represents the average of the observations within the i^{th} group. Then we may write the SST symbolically as

$$SST = SSA + SSE , \tag{1.3}$$

where SSA is called the sum of squares due to groups (i. e. , between groups), and SSE is called the sum of squares due to error (i. e. , within groups). The expected values of SSA and SSE are:

$$E(SSA) = (I-1)J\sigma_a^2 + (I-1)\sigma_e^2 , \tag{1.4}$$

$$E(SSE) = I(J-1)\sigma_e^2 . \tag{1.5}$$

The ANOVA estimates of σ_a^2 and σ_e^2 are obtained by equating the expected values of SSA and SSE to their observed values and solving for the variance components in the following equations:

$$SSA = (I-1)J\sigma_a^2 + (I-1)\sigma_e^2 \,, \tag{1.6}$$

$$SSE = I(J-1)\sigma_e^2 \,. \tag{1.7}$$

Therefore, the estimators of the variance components are

$$\hat{\sigma}_e^2 = \frac{SSE}{I(J-1)} \,, \tag{1.8}$$

$$\hat{\sigma}_a^2 = \frac{SSA}{J(I-1)} - \frac{\hat{\sigma}_e^2}{J} \,. \tag{1.9}$$

The value of $\hat{\sigma}_a^2$ might occasionally be negative, which contradicts the non-negative definition of the variance component. In such cases, one can assume that the sampling variation led to the negative estimate, and set $\hat{\sigma}_a^2 = 0$.

Khuri (2000) gave a review of applications of the ANOVA method and its derivatives in the one-way random effects model. An optimal design based on the ANOVA, roughly speaking, should achieve the best estimates of variances of model parameters under a given design criterion. The design criterion depends on the aim of experimenters. The minimization of the variances is usually set up as design criterion. The choice of design is determined by the number of groups I and the number of observations J_i in each group ($J_1 + J_2 + \cdots + J_I = N$). Balanced designs are often implemented in practice, i. e., each group would be provided with an equal number of observations ($J_1 = J_2 = \cdots = J_I = J$). In general, balanced designs of the ANOVA are most likely to produce minima of the variances in many cases.

Hammersley (1949) showed that for a fixed N (the total number of observations) and $0 \leq I \leq N/2$ (the number of groups), $Var(\hat{\sigma}_a^2)$, the variance of the ANOVA estimator $\hat{\sigma}_a^2$, is minimized by allocating an equal number

$$J = \frac{N\rho + N + 1}{N\rho + 2}$$

observations to each group, where $\rho = \sigma_a^2/\sigma_e^2$. Since J may not yield an integer value, Hammersley suggested that the closest integer be chosen for use. This statement, however, is not precise. For example, if $N = 36$ and $\rho = 0.1$, the calculation gives $J = 7.25$ based on the formula; therefore the closest integers are 7 and 8. Meanwhile, 6 and 9 should also be taken into account since they are very close to 7.25, and they both lead to balanced designs.

Using standard ANOVA estimators with fixed N and I, Anderson and Crump (1967) developed an allocation procedure to minimize $Var(\hat{\sigma}_a^2)$ or $Var(\hat{\rho})$ ($\hat{\rho} = \hat{\sigma}_a^2/\hat{\sigma}_e^2$): $p+1$ observations in each of r groups and p observations in each of $I - r$ groups, where $N = pI + r$, $0 \leq r < I$. In addition, if I is not fixed and the ratio ρ is known, then the optimal number of groups I is conjectured to be the closest integer to

$$\begin{aligned} I_0 &= \frac{N\,(N\,\rho+2)}{N\,\rho+N+1} \quad for\ min\{Var(\hat{\sigma}_a^2)\}, or \\ I_0 &= 1+\frac{(N-5)(N\,\rho+1)}{2\,N\,\rho+N-3} \quad for\ min\{Var(\hat{\rho})\}. \end{aligned}$$

Table 1.1: Variances of the ANOVA estimates of σ_a^2 and ρ

$N = 36$	$Var(\hat{\sigma}_a^2)$		$Var(\hat{\rho})$	
$\rho = 0.1$	$I_0 = 4.97$,	0.0304, if $I = 4$	$I_0 = 4.55$,	0.0350, if $I = 4$
		0.0298, if $I = 5$		0.0349, if $I = 5$
		0.0302, if $I = 6$		0.0361, if $I = 6$
$\rho = 0.2$	$I_0 = 7.49$,	0.0556, if $I = 6$	$I_0 = 6.36$,	0.0703, if $I = 5$
		0.0546, if $I = 7$		0.0683, if $I = 6$
		0.0547, if $I = 8$		0.0686, if $I = 7$
$\rho = 0.5$	$I_0 = 13.09$,	0.1355, if $I = 12$	$I_0 = 9.54$,	0.2060, if $I = 8$
		0.1361, if $I = 13$		0.2017, if $I = 9$
		0.1370, if $I = 14$		0.2029, if $I = 10$
$\rho = 1$	$I_0 = 18.74$,	0.2973, if $I = 17$	$I_0 = 11.92$,	0.5397, if $I = 11$
		0.2925, if $I = 18$		0.5333, if $I = 12$
		0.2954, if $I = 19$		0.5409, if $I = 13$
$\rho = 5$	$I_0 = 30.19$,	2.7630, if $I = 30$	$I_0 = 15.28$,	8.2939, if $I = 15$
		2.7477, if $I = 31$		8.2813, if $I = 16$
		2.7665, if $I = 32$		8.3105, if $I = 17$

Anderson and Crump suggested using I_0, provided I_0 is an integer. If I_0 is not an integer, compute the exact variance for each of the integers just smaller and larger than I_0. Compute the variance for the next adjacent integral value of I for whichever variance is smaller, and continue this process until the smallest variance is obtained. The variances of the ANOVA estimates of σ_a^2 and ρ are reported in Table 1.1 ($N = 36$). In most cases, one of the two original integers is found to give the smallest variance. The optimal designs chosen by Anderson and Crump's method are reported in Table 1.2. Whenever I is a divisor of 36, the design is balanced. If I_0 is close to the I for a balanced design, then the balanced design would most likely be optimal.

Table 1.2: Optimal Designs for N = 36

$N = 36$	$min\{Var(\hat{\sigma}_a^2)\}$	$min\{Var(\hat{\rho})\}$
$\rho = 0.1$	$I_0 = 4.97, I = 5, p = 7, r = 1$	$I_0 = 4.55, I = 5, p = 7, r = 1$
$\rho = 0.2$	$I_0 = 7.49, I = 7, p = 5, r = 1$	$I_0 = 6.36, I = 6, p = 6, r = 0$
$\rho = 0.5$	$I_0 = 13.09, I = 12, p = 3, r = 0$	$I_0 = 9.54, I = 9, p = 4, r = 0$
$\rho = 1$	$I_0 = 18.74, I = 18, p = 2, r = 0$	$I_0 = 11.92, I = 12, p = 3, r = 0$
$\rho = 5$	$I_0 = 30.19, I = 31, p = 1, r = 5$	$I_0 = 15.28, I = 16, p = 2, r = 4$

Thompson and Anderson (1975) further investigated the problem of selecting optimal designs from Anderson and Crump (1967) using standard ANOVA estimators, truncated ANOVA estimators ($\hat{\sigma}_a^2 \geq 0$), maximum likelihood (ML) estimators, and modified maximum likelihood (MML) estimators. For a specified model and error distribution, the method of maximum likelihood selects parameter estimates that maximize the probability of occurrence of the sample results. Milliken and Johnson (1984) provided a nice overview of the method of maximum likelihood applied to experimental design models. The standard ANOVA estimators are given by (1.8) and (1.9). The truncated ANOVA estimators are defined as

$$\begin{aligned} \hat{\sigma}_e^2 &= \min\left\{\frac{SSE}{I(J-1)};\frac{SSA+SSE}{IJ-1}\right\}, && (1.10)\\ \hat{\sigma}_a^2 &= \max\left\{\frac{SSA}{J(I-1)}-\frac{SSE}{IJ(J-1)};0\right\}. && (1.11)\end{aligned}$$

The maximum likelihood estimators are given by:

$$\begin{aligned} \hat{\sigma}_e^2 &= \min\left\{\frac{SSE}{I(J-1)};\frac{SSA+SSE}{IJ}\right\}, && (1.12)\\ \hat{\sigma}_a^2 &= \max\left\{\frac{SSA}{IJ}-\frac{SSE}{IJ(J-1)};0\right\}. && (1.13)\end{aligned}$$

The modified maximum likelihood estimators are slightly different from the ML estimators:

$$\begin{aligned} \hat{\sigma}_e^2 &= \min\left\{\frac{SSE}{I(J-1)};\frac{SSA+SSE}{IJ+1}\right\}, && (1.14)\\ \hat{\sigma}_a^2 &= \max\left\{\frac{SSA}{(I+1)J}-\frac{SSE}{IJ(J-1)};0\right\}. && (1.15)\end{aligned}$$

The minimization of the mean squared error (MSE) was used as the design criterion. Thompson and Anderson concluded that the optimal designs derived from the truncated ANOVA estimators do not differ much from those derived from the standard

ANOVA estimators. The method of maximum likelihood provides estimators of σ_a^2 with smaller mean squared errors than the other methods for an unbalanced design, which consists of one observation in $I - r$ groups and two observations in r groups. For balanced designs, the MML estimator of σ_a^2 is superior than the ANOVA or ML estimator since the optimal design based on the estimator is less sensitive to intra-group correlation τ $(0 < \tau < 0.5)$ than those designs based on minimizing the variance of the standard ANOVA estimator of σ_a^2, where $\tau = \sigma_a^2/(\sigma_a^2 + \sigma_e^2)$.

It is crucial to realize that the implementation of these methods of estimation, such as the ANOVA, ML and MML, rely on the plausibility of the model assumptions. Notably, many studies have shown that these procedures behave poorly under slight violations of the model assumptions. Even one single outlier, namely an observation that is sufficiently far away from the majority of observations, may have a large effect on the estimations. The one-way random effects model belongs to the class of the parametric model. In general, parametric models are only approximations to reality that would not always behave as nicely as described by the model assumptions. Therefore robust procedures that work well not only under strict parametric models, but also in neighborhoods of such models, have been proposed with growing awareness of their importance.

1.3 Robust Procedures and Optimal Designs

Robustness signifies insensitivity to small deviations from the model assumptions (Huber, 1981). A robust procedure is expected to produce efficient and reliable estimates by applying robust estimators under the true underlying model. In other words, small deviations from the model assumptions should impair the performance only slightly, and some larger deviations should not cause a disaster. We will discuss in detail the concepts of robustness and their applications in Chapter 2.

Zhou and Zhu (2003) studied optimal designs for model (1.1) based on Rocke's (1983, 1991) robust M–estimators. The main idea of Rocke's procedure is to replace outliers with pseudo-observations, and then use the standard ANOVA to estimate variance components. A robustness measure of the estimators, the finite-sample breakdown point, was also exhaustively studied in their paper. For choosing optimal designs, the minimization of the mean squared error (MSE) was used as the design criterion subject to a constraint of specified finite-sample breakdown point.

Muller and Uhlig (2001) investigated robustness properties of three variance components estimators that Uhlig (1997) derived from the Q–estimator of Rousseeuw and Croux (1992,1993). The estimators of Uhlig are based on quartiles of possible differences by all observations of a sample and can be calculated non-iteratively. Moreover, they attain relatively high breakdown points in both cases of infinite and finite sample. We follow the idea of Zhou and Zhu (2003) to study robust designs for model (1.1) using Uhlig's estimators and compare the designs with those based on Rocke's estimators and the standard ANOVA estimators. It is interesting to note that Uhlig's estimators and Rocke's estimators behave similarly in finite-sample situations.

1.4 Main Contributions

In this thesis, we make the following contributions to robust designs of variance components analysis in the one-way random effects model.

1. The empirical influence function (IF) and the finite-sample breakdown point are explored in detail. Several classical statistical estimators are compared in terms of the two robustness measurements.
2. Rocke's M-estimators and Uhlig's Q-estimators are introduced, and the related optimal robust designs are constructed according to the minimization of the mean squared error (MSE), which is subject to the constraint of a pre-defined finite-sample breakdown point. Comparisons are made between the derived robust designs and the optimal ANOVA designs.
3. An alternative to constructing robust designs is further developed: minimize the variances of the standard ANOVA estimates under the robust constraint of finite-sample breakdown point. From all the results reported in the thesis,

robust designs derived from this method are very similar to those derived from simulated MSEs.

In Chapter 2, general robustness concepts are introduced. In Chapter 3, we study Rocke's M-estimators and Uhlig's Q-estimators and also investigate their robust properties in finite-sample situation. In Chapter 4, we construct robust designs based on Uhlig's estimators, and then compare them with the designs based on Rocke's M-estimators and standard ANOVA estimators. Concluding remarks are given in Chapter 5.

Chapter 2

Robust Statistics

What is robust statistics? In this chapter, we answer this question by looking at the aims of robust statistics, its properties, and various approaches to its study.

2.1 Robust Statistics

As we know, statistical inferences are based not only on samples, but also prior assumptions about the populations. The assumptions, such as randomness, independence, and model distribution, are commonly made for mathematical tractability and simplification; unfortunately, the underlying situations do not always behave as nicely as described by the assumptions. One reason is the occurrence of gross errors, such as equipment failure, typos, and blunders. The gross errors usually show up as outliers, namely observations which are far away from the pattern formed by the majority of the data, and are the most common deviation for many classical estimation procedures. For instance, a single outlier that is sufficiently far away can ruin a least-squares analysis completely. The frequency of gross errors varies considerably: there may be 1–10% or more gross errors in routine data sets; in high-quality data sets, there is still a tiny fraction of outliers which are sometimes difficult to find. One example considered by Hampel et al. (1986) concerned a quarter million electroencephalographic data, which were automatically recorded by nearly working equipment; the histogram looked nearly normal, except for some extreme values in the tails of the histogram. Analysis revealed that about two dozen extremely large data points were generated when the equipment was switched on. Another example occurred in the 1950 U.S. census data. After careful screening and analysis, Coale and Stephan (1962) showed that a fraction of incorrectly entered data escaped all controls since they produced some pattern deviations in certain rare population subgroups. To avoid misunderstandings, we should emphasize that the above examples do not imply that gross errors can not be prevented. With sufficient care, we can obtain high-quality data free of gross errors; however, the perfect conditions cannot always be secured, and thus we must acknowledge the potential for errors.

Given the existence of gross errors, the problem with classical parametric statistics (particularly, those optimized for an underlying normal distribution) is that they are excessively sensitive to minor deviations from the model assumptions. In other words, they cannot provide reliable information when the model assumptions are only approximately valid. Tukey (1960) provided an impressive example. He considered the location model $F(x-\theta)$ with $F(x) = (1-\epsilon)\Phi(x) + \epsilon\Phi(x/3)$, Φ being the standard normal cumulative function and ϵ ranging from 0–10%. This can be considered as a mixture data set with gross errors, in which each single observation could be either a "good" one with probability $1-\epsilon$, or a "bad" one with probability ϵ. Tukey showed that the asymptotic relative efficiency of the classical mean decreases quickly from 1 (for $\epsilon = 0\%$) to about 70% (for $\epsilon = 10\%$). The case for the variance is even worse. As an alternative to classical statistics, robust statistics has been proposed with the growing awareness of its effectiveness. For our purpose, we use the definition of Hampel et al. (1986) for "robust statistics":

Definition 2.1 *In a broad informal sense, robust statistics is a body of knowledge, partly formalized into "theories of robustness", relating to deviations from idealized assumptions in statistics.*

It is clear that one primary goal of robust statistics is to safeguard against gross errors, and to be robust in the sense that other deviations from the model assumptions should not significantly affect estimation procedures. Furthermore, robust statistics is an extension of classical parametric statistics, studying the behavior of statistical procedures not only under strict parametric models, but also under gross-error models.

One frequent argument is whether robust procedures are needed at all since one might first apply some rules of rejection to eliminate outliers, and then apply classical methods to the "cleaned" data. Unfortunately, these steps would not be as applicable as a good robust procedure is, due to a few reasons. First, the outlier detection can not be effectively implemented on high-dimensional and complex data sets. Second, there is no guarantee that the cleaned data will be normal; there may be statistical errors of both kinds, false rejections and false retentions. Third, a Monte Carlo study shows that the best robust procedures can decrease the influence of outliers smoothly, while the best rejection procedures cannot handle their influence as well as the robust procedures do.

2.2 Robustness Measures

Huber and Hampel's pioneer work from the 1960s forms the basis for a comprehensive theory of robustness. Huber (1964) introduced the "gross error model", which assumes that, for a strict parametric model of the form $G(x-\theta)$, a fraction ϵ $(0 \leq \epsilon \leq 1)$ of the data may consist of gross errors with an arbitrary distribution $H(x-\theta)$ (where θ is only introduced to preserve the form of a location model). Thus, the underlying distribution of the model is $F(x-\theta) = (1-\epsilon)G(x-\theta) + \epsilon H(x-\theta)$. Huber's aim is to optimize the worst case over the neighborhood of the model. To fulfill this objective, he also introduced a class of "M–estimators", including the famous Huber-estimator, which are generalized from the classical maximum likelihood estimator. Their properties, such as consistency, asymptotic normality, and asymptotic variance, can be used to describe the behavior of a robust procedure over the neighborhood of the model.

Hampel (1968) introduced the infinitesimal approach, which comprises three central concepts: qualitative robustness, influence function, and breakdown point. For our purpose, we will discuss in detail the use of the influence function and the breakdown point in finite-sample situation.

2.2.1 The Empirical Influence Function

Since the underlying distribution of observations is assumed to lie in some neighborhood of the parametric model, we need a simple and powerful tool to extract the information about the behavior of an estimator in the full neighborhood. The empirical influence function (IF) can satisfy the purpose:

Definition 2.2 *Suppose $T_n(x_1, \cdots, x_n)$ is an estimator based on the sample $(x_1, \cdots, x_n)$. The empirical influence function, or the finite-sample influence function, of the estimator at the sample is*

$$T_n(x_1, \cdots, x_{n-1}, x) \tag{2.1}$$

obtained by replacing an observation, say x_n, with an arbitrary observation x.

The empirical IF describes the effect of one observation in any point x on a statistic T_n. If there exists a constant M such that $|\, T_n(x_1, \cdots, x_{n-1}, x) \,| \leq M$, for $-\infty <$

$x < \infty$, then we say the influence function of T_n is bounded, or T_n is a robust estimator. Let us look at an example.

Example 2.1 *Consider the data on the prolongation of sleep (ten subjects) by means of two drugs A and B (Cushny and Peebles, 1905). The ten pairwise differences (per subject) between drug effects (the number of sleeping hours after taking drug A - the number of sleeping hours after taking drug B) are the following:*

$$0.0, 0.8, 1.0, 1.2, 1.3, 1.3, 1.4, 1.8, 2.4, 4.6\ .$$

We examine the following statistical estimators and their empirical IFs*: three location estimators, including the sample mean* $\bar{X}$*, the sample median* $\tilde{X}$*, and the 10%-trimmed mean* $\bar{X}_{tr}$*, and three scale estimators, including the sample standard deviation* S*, the median absolute deviation* (MAD)*, and the interquartile range* (IQR)*.*

The sample mean is simply the arithmetic average of the data. For the above data,

$$\bar{X}(x_1, \cdots, x_9, x_{10}) = \frac{1}{10}\sum_{i=1}^{10} x_i = 1.58.$$

The empirical IF of the sample mean $\bar{X}$ is given by

$$\begin{aligned}
\bar{X}(x_1, \cdots, x_9, x) &= \frac{x_1 + x_2 + \cdots + x_9 + x}{10} \\
&= \frac{1}{10}\sum_{i=1}^{9} x_i + \frac{x}{10} \\
&= 1.12 + \frac{x}{10}\ .
\end{aligned}$$

The empirical IF of the sample mean is drawn in Figure 2.1(a) together with that of the median (Figure 2.1(b)) and the 10%-trimmed mean (Figure 2.1(c)). The curve in (a) shows that the empirical IF of the sample mean becomes unbounded for $x \rightarrow \pm\infty$. In other words, the sample mean is a non-robust estimator since the extreme values of the replaced observation would break down the estimate.

The sample median is the middle value when the observations are ordered from the smallest to the largest. For the data,

$$\tilde{X}(x_1, \cdots, x_9, x_{10}) = 1.3.$$

The empirical IF of the sample median is given by

$$\tilde{X}(x_1, \cdots, x_9, x) = \begin{cases} 1.25 & \text{if } x \leq 1.2; \\ 0.65 + \frac{x}{2} & \text{if } 1.2 < x < 1.3; \\ 1.3 & \text{if } x \geq 1.3\ . \end{cases}$$

It is clear in Figure 2.1(b) that the empirical IF of the sample median is bounded $(1.25 \leq |\,\tilde{X}(x_1, \cdots, x_9, x)\,| \leq 1.3)$ for all $-\infty < x < +\infty$. The sample median $\tilde{X}$ is thus a robust estimator.

The 10%-trimmed mean $\bar{X}_{tr}$ is defined as the mean of a proportion of the sample by removing 10% of the largest and 10% of the smallest observations. In the aforementioned example, this amounts to deleting the largest observation 4.6, as well as the smallest 0, and computing the average of the remaining eight values: i. e.

$$\bar{X}_{tr}(x_1, \cdots, x_9, x_{10}) = 1.4\ .$$

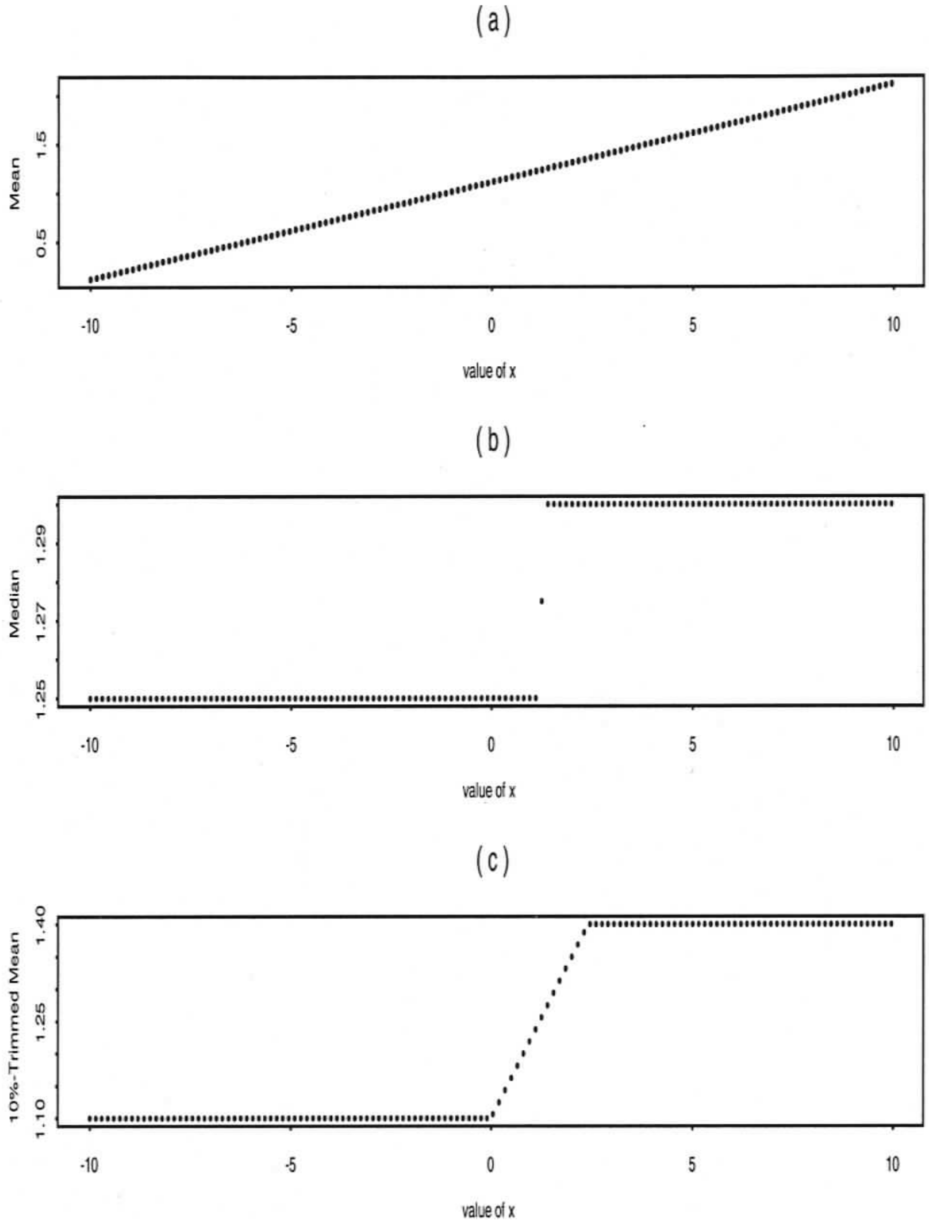


Figure 2.1: The empirical influence functions of (a) the sample mean, (b) the sample median, and (c) the 10%-trimmed mean, where the last observation of the sample has been replaced by an arbitrary value x.

The empirical IF is given by

$$\bar{X}_{tr}(x_1, \cdots, x_9, x) = \begin{cases} 1.1 & \text{if } x \leq 0; \\ 1.1 + \frac{x}{8} & \text{if } 0 < x < 2.4; \\ 1.4 & \text{if } x \geq 2.4 \,. \end{cases}$$

The curve (Figure 2.1(c)) for the 10%-trimmed mean shows that the empirical IF remains constant for both $x \to +\infty$ and $x \to -\infty$, which means that the 10%-trimmed mean is a robust estimator.

The sample standard deviation S is commonly used to estimate the standard deviation of the population. In this example,

$$S(x_1, \cdots, x_9, x_{10}) = \sqrt{\frac{1}{9}\sum_{i=1}^{10}(x_i - \bar{X})^2} = 1.229995 \,,$$

where $\bar{X} = 1.58$ is the sample mean. The empirical IF of the sample standard deviation is given by

$$\begin{aligned} S(x_1, \cdots, x_9, x) &= \sqrt{\frac{1}{9}\Big[\sum_{i=1}^{9} x_i^2 + x^2 - \frac{1}{10}\big(\sum_{i=1}^{9} x_i + x\big)^2\Big]} \\ &= \sqrt{\frac{1}{9}\Big[\sum_{i=1}^{9} x_i^2 - \frac{1}{10}\big(\sum_{i=1}^{9} x_i\big)^2 + \frac{9}{10}x^2 - \frac{2}{10}\big(\sum_{i=1}^{9} x_i\big)x\Big]} \\ &= \sqrt{\frac{1}{9}\Big[17.42 - 0.1(11.2)^2 + 0.9x^2 - 0.2(11.2)x\Big]} \\ &= \sqrt{0.1x^2 - 0.249x + 0.542} \,. \end{aligned}$$

For large positive or negative x, the quadratic term $0.1x^2$ dominates the value of $S(x_1, \cdots, x_9, x)$, which becomes unbounded as x goes to infinity. The sample standard deviation is thus non-robust for any sample with outliers. Figure 2.2(a) shows the empirical IF of S, together with that of the MAD (Figure 2.2(b)), and the IQR (Figure 2.2(c)).

The median absolute deviation (MAD) is defined as the median of the absolute deviations from the median:

$$MAD = 1.4826 med\{|x_i - \tilde{x}|\},$$

where 1.4826 is a tuning constant that makes the estimator consistent when the underlying distribution is the standard normal, and $\tilde{x}$ is the sample median. The MAD is commonly used as the ancillary estimate of scale. Here,

$$MAD(x_1, \cdots, x_9, x_{10}) = 0.593.$$

Its empirical IF is given by

$$MAD(x_1, \cdots, x_9, x) = \begin{cases} 0.519 & \text{if } x \leq 0; \\ C_1(x) & \text{if } 0 < x < 4.6; \\ 0.593 & \text{if } x \geq 4.6 \,. \end{cases}$$

The curve for the MAD in Figure 2.2(b) shows that the value of $C_1(x)$ is finite and lies between 0.29652 and 0.593 no matter where x is positioned in the interval (0, 4.6). Meanwhile, the robust property inherited from the median ensures the empirical IF of the MAD is bounded for both $x \to +\infty$ and $x \to -\infty$.

The interquartile range (IQR) is defined as the difference between the upper and lower (or third and first) quartiles of the sample

$$IQR = \frac{Q_3 - Q_1}{\Phi^{-1}(0.75) - \Phi^{-1}(0.25)},$$

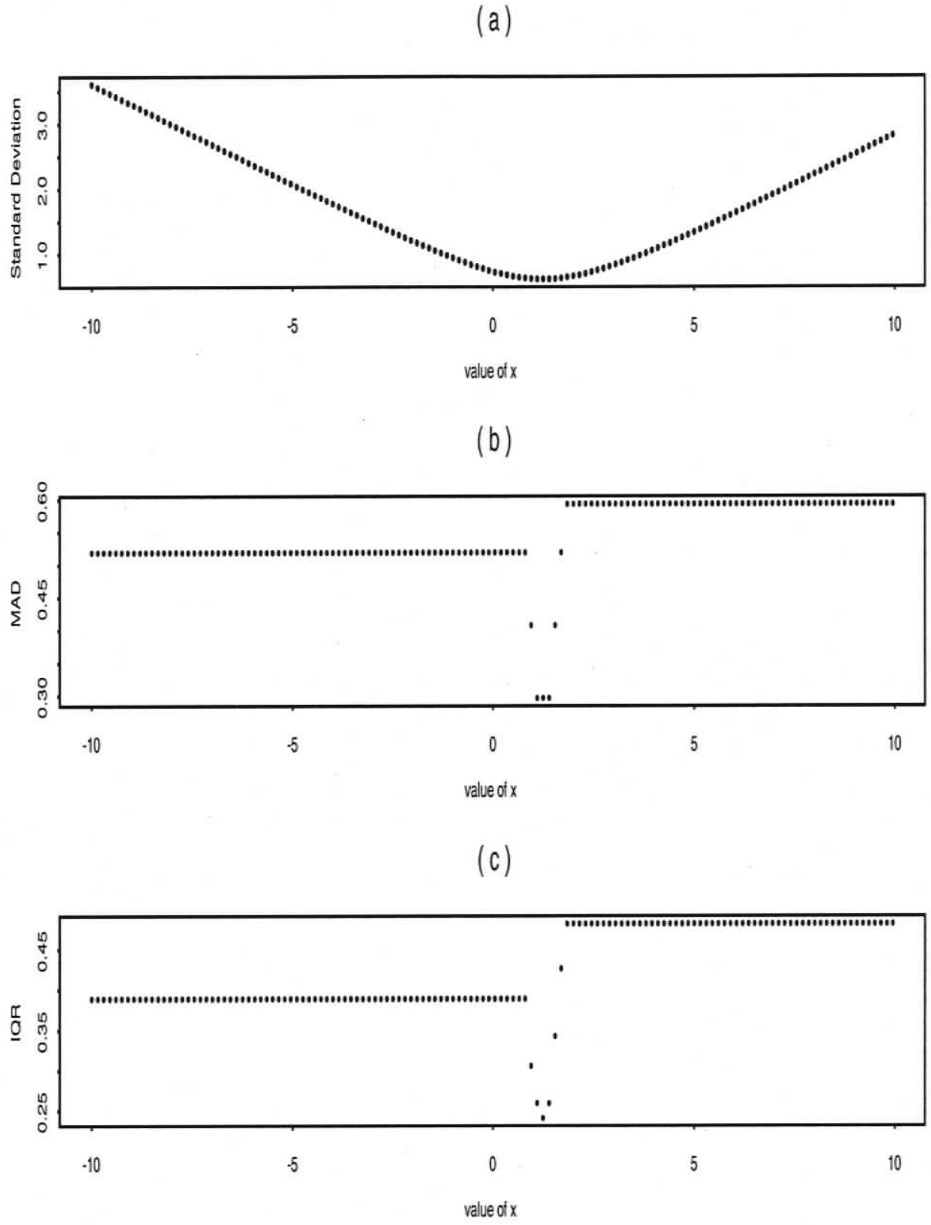


Figure 2.2: The empirical influence functions of (a) S, (b) MAD, and (c) IQR, where the last observation of the sample has been replaced by an arbitrary value x.

where Q_1 and Q_3 are the first quartile and the third quartile of the sample respectively, and where Φ is the standard normal cumulative function. In the example,

$$IQR(x_1, \cdots, x_9, x_{10}) = 0.4818457 \ .$$

The empirical IF of the IQR is given by

$$IQR(x_1, \cdots, x_9, x) = \begin{cases} 0.389 & \text{if } x \leq 0; \\ C_2(x) & \text{if } 0 < x < 4.6; \\ 0.482 & \text{if } x \geq 4.6 \ . \end{cases}$$

The cures for the IQR (Figure 2.2(c)) and for the MAD look very similar. Also the empirical IF of the IQR is bounded for all $-\infty < x < +\infty$ $(0.241 \leq C_2(x) \leq 0.482)$. Thus, the IQR is a robust estimator.

2.2.2 The Finite-Sample Breakdown Point

Note in Example 2.1 that the empirical influence functions of the sample mean and the sample standard deviation become unbounded as x goes to infinity, whereas that of the other estimators remain finite, or in other words, the other estimators can handle at least one outlier. Then we may ask the following questions: how many outliers can an arbitrary estimator handle and remain robust? Do different types of robust estimators deal with different number of outliers? Which estimator has the best performance under specific circumstances or for specific purposes? The robustness measure, namely the finite-sample breakdown point which is the fraction of outliers that can be tolerated by an estimator, would be a heuristic tool to address these questions.

We use the word "breakdown" in a relatively narrow sense; breakdown signifies failures of providing reliable and relevant information of a statistic. In general, an estimator may break down in either of two ways, or both: the situation in which the estimate becomes unbounded is called "explosion"; for a scale estimator, the situation in which the estimate approaches zero is called "implosion".

Definition 2.3 *Suppose* $T_n(x_1, \cdots, x_n)$ *is an estimator for the finite sample* $(x_1, \cdots, x_n)$ *in which all observations are pairwise different. The finite-sample explosion breakdown point* ϵ_n^e *of the estimator* T_n *is given by*

$$\epsilon_n^e(T_n) = \frac{1}{n} \max\{m; \max_{i_1,\cdots,i_m} \sup_{y_1,\cdots,y_m} |T_n(z_1, \cdots, z_n)| < \infty\}, \tag{2.2}$$

and the finite-sample implosion breakdown point ϵ_n^i *of the estimator* T_n *is given by*

$$\epsilon_n^i(T_n) = \frac{1}{n} \max\{m; \min_{i_1,\cdots,i_m} \inf_{y_1,\cdots,y_m} T_n(z_1, \cdots, z_n) > 0\}. \tag{2.3}$$

The sample $(z_1, \cdots, z_n)$ *in (2.2) and (2.3) is obtained by replacing the* m *data points* $x_{i_1}, \cdots, x_{i_m}$ *with the arbitrary values* $y_1, \cdots, y_m$*. For location estimators, the finite-sample breakdown point is*

$$\epsilon_n(T_n) = \epsilon_n^e(T_n);$$

for scale estimators, the finite-sample breakdown point is

$$\epsilon_n(T_n) = \min\{\epsilon_n^e(T_n), \epsilon_n^i(T_n)\}.$$

Basically, the finite-sample breakdown point is the largest fraction of gross errors that can be tolerated by an estimator without any breakdown, and it varies with the sample size n. If we know the breakdown point of an estimator, we can calculate the number of outliers that the estimator can handle, $n \cdot \epsilon_n$. We usually take 0.5 as the cutoff value for the breakdown point, because the contamination of half of the observations in an experiment is considered as the worst-case situation. Once the proportion of outliers goes beyond the cutoff value, the rest of the data may be treated as "current" outliers.

Example 2.2 *Consider those estimators in Example 2.1, and find their finite-sample breakdown points.*

Recall that the sample mean is a non-robust estimator since its empirical IF is unbounded. In other words, one single outlier can completely ruin the estimate. The sample mean thus has a breakdown point of 0.

The sample median can handle up to four outliers for $n = 10$, because the median is the average of the 5^{th} and 6^{th} ordered observations. In other words, any four outliers in the data cannot significantly affect the estimate whether they stand on the same side or not. Notably, five outliers may cause a breakdown, as showed in Figure 2.3 (a). Thus the finite-sample breakdown point is $4/10 = 0.4$. In general, the sample median can deal with as many as $\lfloor (n-1)/2 \rfloor$ outliers, where $\lfloor x \rfloor$ is the largest integer less than or equal to x. The finite-sample breakdown point of the sample median will gradually approach 0.5 as the sample size n grows.

The 10%-trimmed mean is a special case of the α-trimmed mean $(0 < \alpha < 1)$ when α is equal to 0.1. The 10%-trimmed mean has a bounded empirical IF, which means it can deal with at least one outlier. However, if the data contains two outliers, the 10%-trimmed mean could break down, because it removes only one observation on either side. In a case that both outliers lie on the same side, one will remain in the data after trimming. As Figure 2.3 (b) shows, the estimates of the 10% trimmed mean become unbounded when the sample contains two or more outliers. The finite-sample breakdown point of the the 10%-trimmed mean is $1/10 = 0.1$ for $n = 10$. Generally, the α-trimmed mean has the breakdown point α.

The empirical IF of the sample standard deviation becomes unbounded for both $x \to +\infty$ and $x \to -\infty$, which means one single outlier can carry the estimate over the bounds, or say the explosion breakdown point equals to 0 ($\epsilon_{10}^{e} = 0$). Then its finite-sample breakdown point is 0, i. e.

$$\epsilon_{10} = \min\{\epsilon_{10}^{e}, \epsilon_{10}^{i}\} = 0 ,$$

where ϵ_{10}^{i} is non-negative from the definition. The sample standard deviation is non-robust for all n.

The MAD is used to estimate the standard deviation of the population. As shown in Figure 2.4 (a), five outliers may yield an unbounded MAD. In other words, the MAD can stand up to four outliers without explosion ($\epsilon_{10}^{e} = 4/10$). Due to the

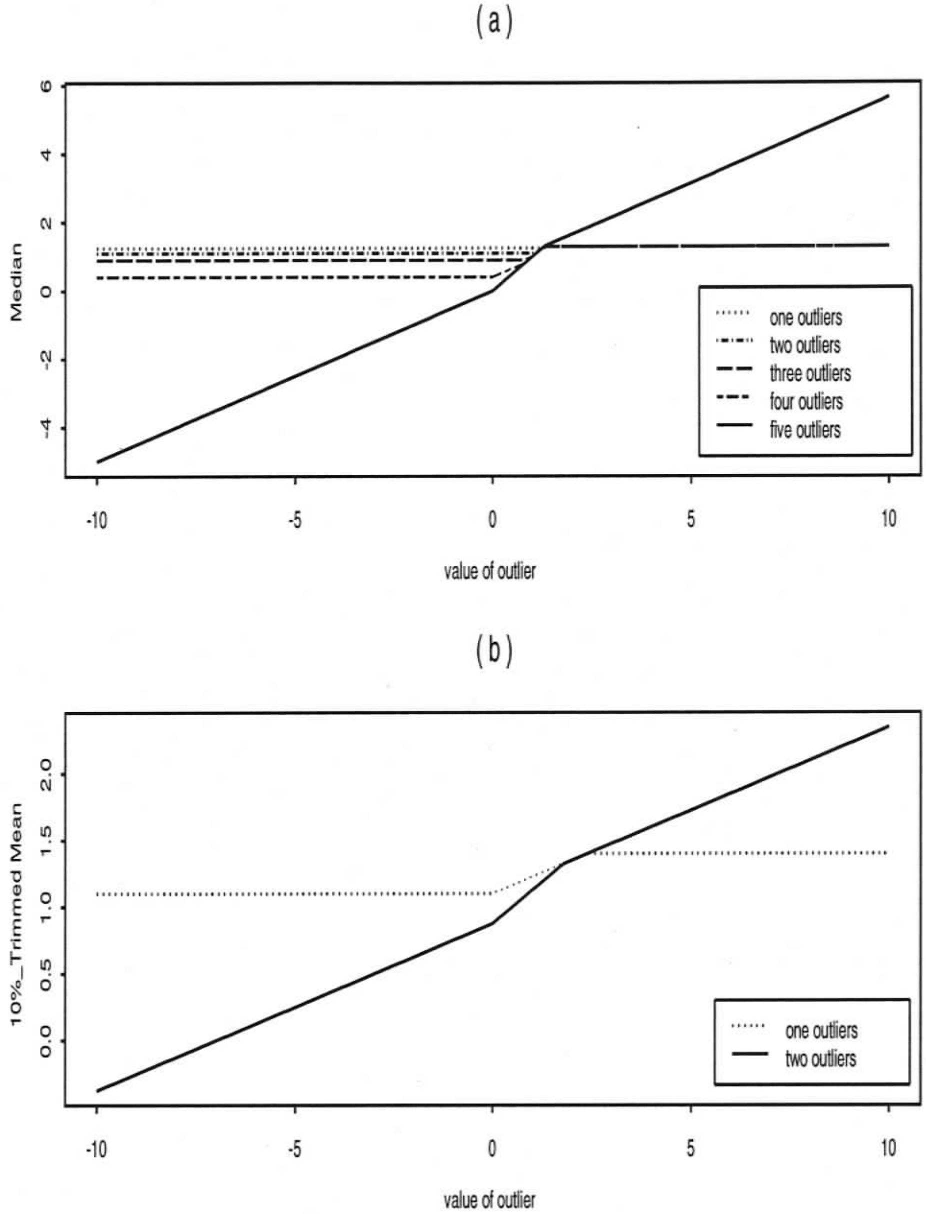


Figure 2.3: Illustration of the finite-sample breakdown points of the sample median and the 10%-trimmed mean for the data of Example 2.1.

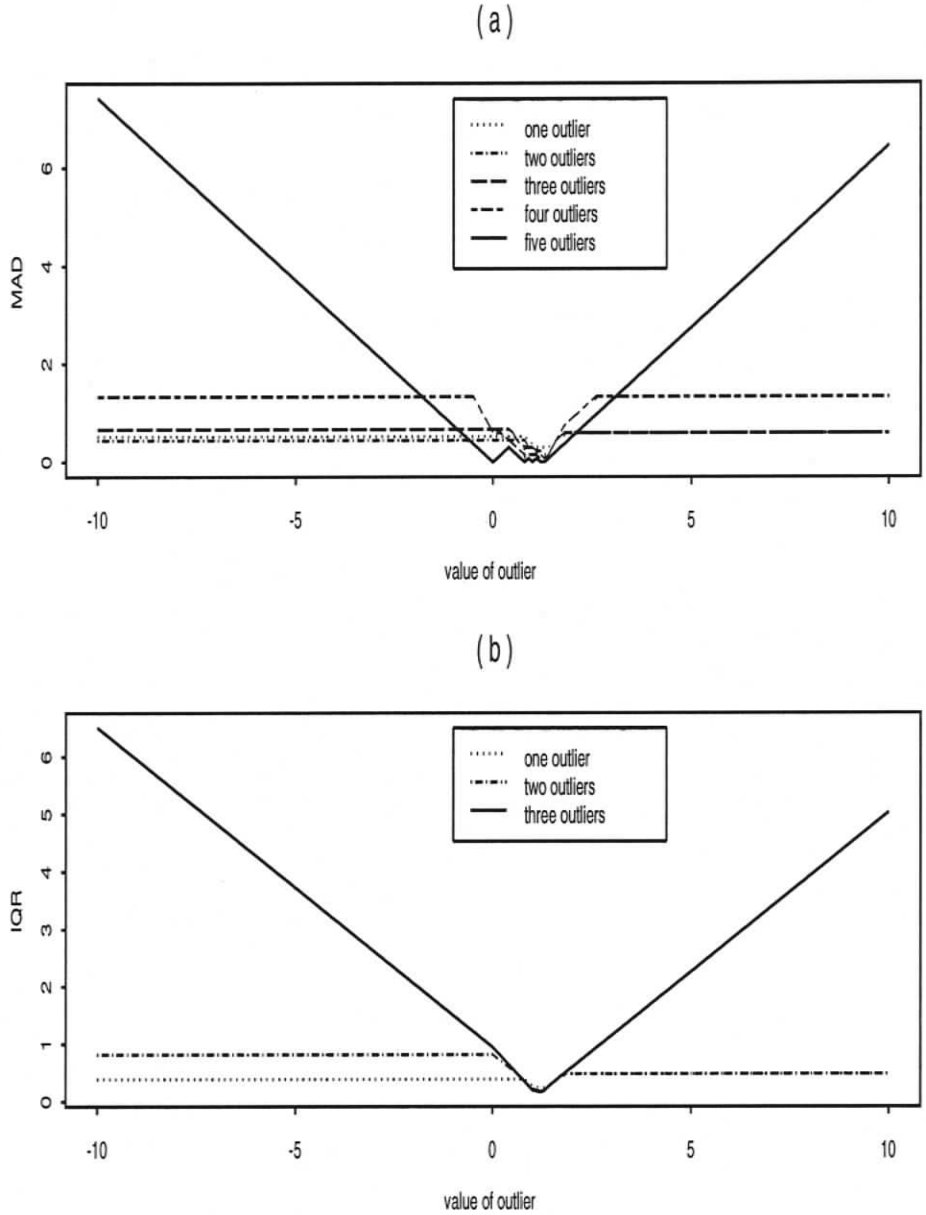


Figure 2.4: Illustration of the finite-sample breakdown points of the *MAD*, and the *IQR* for the data of Example 2.1.

robust property inherited from the median, the MAD can handle $\lfloor (n-1)/2 \rfloor$ outliers without explosion in a general case. On the other hand, the implosion situation becomes a little complicated because of the presence of non-distinct observations in the data. In Example 2.1, the sample consists of ten observations, two of which coincide as 1.3. If we replace four of the other observations with 1.3 (5 outliers in total), the MAD would reach zero, i. e. the implosion breakdown point ϵ_{10}^{i} is equal to 4/10. Hence, the finite-sample breakdown point is

$$\epsilon_{10} = \min\{\epsilon_{10}^{e}, \epsilon_{10}^{i}\} = 0.4 \ .$$

For a sample in which all observations are distinct, the MAD would be non-zero given at most $\lfloor (n-2)/2 \rfloor$ outliers. For this case, the finite-sample breakdown point of the MAD is equal to the implosion breakdown point $\lfloor (n-2)/2 \rfloor / n$, since it is less than or equal to the explosion breakdown point $\lfloor (n-1)/2 \rfloor / n$. It will approach 0.5 gradually as the sample size n grows.

Figure 2.4 (b) shows the curves of the IQR at the sample with different number of outliers. Compared to Figure 2.4 (a), both IQRs and the MADs demonstrate quite similar behaviors given the occurrence of outliers. The IQR becomes unbounded in the case of three outliers, because either the first quartile Q_1 or the third quartile Q_3 will be ruined if three outliers stand on the same side of the sample. The explosion breakdown point is thus $\epsilon_{10}^{e} = 2/10 = 0.2$. In general, the IQR can deal with $\lfloor (n-1)/4 \rfloor$ outliers at most without explosion, which is less than the explosion breakdown point of the MAD ($\lfloor (n-1)/2 \rfloor$). Ultimately, the implosion breakdown point of the IQR is dependent on the composition of a sample. In the data, if we replace the third, fourth, seventh and eighth observations with 1.3 (5 outliers in total), the IQR will reach zero. The implosion breakdown point is thus $\epsilon_{10}^{i} = 4/10 = 0.4$. Intuitively, the implosion breakdown point of the IQR is greater than the explosion breakdown point of the IQR, because the observations located between Q_1 and Q_3 must coincide if $Q_3 - Q_1$ is equal to zero, and the proportion of these observations should be nearly half of the sample size. Therefore, it is unnecessary to determine an formula for calculating the implosion breakdown point. In practice we take the explosion breakdown point $\lfloor (n-1)/4 \rfloor / n$ as the finite-sample breakdown point of the IQR, which will approach 0.25 gradually as the sample size n increases.

We have addressed the questions introduced at the beginning of this section. The finite-sample breakdown point varies with the sample size; while different types of robust estimators would show different tolerances to outliers. However, an estimator with the largest finite-sample breakdown point is not always the best candidate for a certain statistical procedure because other aspects, such as accuracy, efficiency and computational complexity, should also be taken into account. At last, one point we should bear in mind is that sometimes outliers are actually proper observations which are the most valuable ones of all and merit further study.

Chapter 3

Robust Variance Estimators

In Chapter 1, we have reviewed several estimators of variance components for the one-way random effects model, such as standard ANOVA estimators, maximum likelihood estimators, and modified maximum likelihood estimators. These estimators, however, have a common undesirable feature: sensitivity to small deviations from the model assumptions. Our purpose is to construct optimal designs that can tolerate a specified fraction of gross errors. Robust statistics in Chapter 2 has demonstrated the capacity to handle the occurrence of outliers, which may fulfill our purpose upon design construction. In this chapter, we mainly focus on two types of robust estimators of variance components: Rocke's M-estimators and Uhlig's Q-estimators, as well as their robustness properties.

3.1 M-estimators

It is necessary to consider the properties of M-estimators before we study Rocke's M-estimators in detail. Huber (1964) introduced a class of estimators, called "M-estimators".

Definition 3.1 *Given a finite sample* $(x_1, \cdots, x_n)$*, consider a minimum problem of the form*

$$\min_{T_n} \left\{ \sum_{i=1}^{n} \rho(x_i; T_n) \right\} ,$$

or equivalently, an implicit equation

$$\sum_{i=1}^{n} \psi(x_i; T_n) = 0 ,$$

where ρ *is an arbitrary function and* $\psi(x;\theta) = (\partial/\partial\theta)\rho(x;\theta)$*. The* T_n *that minimizes* $\sum_{i=1}^{n} \rho(x_i; T_n)$ *or solves the implicit equation is called an M-estimate.*

Note that the choice $\rho(x;\theta) = -\log f(x;\theta)$ gives the ordinary maximum likelihood estimate. In other words, M-estimators are a generalization of maximum likelihood estimators. Here let us explore Huber's famous estimators. The Huber's estimators were introduced as the solutions (T_n, S_n) of the following equations

$$\begin{aligned} \sum_{i=1}^{n} \psi_c(\frac{x_i - T_n}{S_n}) &= 0 , \\ \sum_{i=1}^{n} \psi_c^2(\frac{x_i - T_n}{S_n}) &= (n-1)\beta , \end{aligned}$$

where

$$\begin{aligned}\psi_c(x) &= \min\{c, \max\{x, -c\}\}\ , \\ &= \begin{cases} -c & x < -c\ , \\ x & -c \le x \le c\ , \\ c & x > c\ , \end{cases}\end{aligned}$$

$$\beta = \int \psi_c^2(x)\, d\Phi(x)\ ,$$

and Φ is the standard normal cumulative function. T_n and S_n are called simultaneous M-estimates of location and scale. The computation procedure starts with initial estimates of location and scale:

$$\begin{aligned}T_n^{(0)} &= med\{x_i\}\ , \\ S_n^{(0)} &= 1.4826 med\left\{|x_i - T_n^{(0)}|\right\}\ .\end{aligned}$$

Simulation (Andrews et al. , 1972) has clearly shown the importance of taking robust initial estimates for the computation; starting with non-robust estimates may lead to poor results. Then we perform iterative Newton steps, where

$$\begin{aligned}[S_n^{(m+1)}]^2 &= \frac{1}{(n-1)\beta} \sum_{i=1}^{n} \psi_c^2\left(\frac{x_i - T_n^{(m)}}{S_n^{(m)}}\right) [S_n^{(m)}]^2\ , \\ T_n^{(m+1)} &= T_n^{(m)} + \frac{S_n^{(m)} \sum_{i=1}^{n} \psi_c\left(\frac{x_i - T_n^{(m)}}{S_n^{(m)}}\right)}{\sum_{i=1}^{n} \psi_c'\left(\frac{x_i - T_n^{(m)}}{S_n^{(m)}}\right)}\ .\end{aligned}$$

The iteration will stop once the termination condition is satisfied. For example,

$$\begin{aligned} \left|S_n^{(m+1)} - S_n^{(m)}\right| &< 0.00001 \ , \\ \left|T_n^{(m+1)} - T_n^{(m)}\right| &< 0.00001 \ . \end{aligned}$$

A convergence proof is given by Huber (1981, Section 7.8).

Since in practice scale usually occurs as a nuisance parameter in location, Huber also introduced one-step M-estimator of location by solving the equation

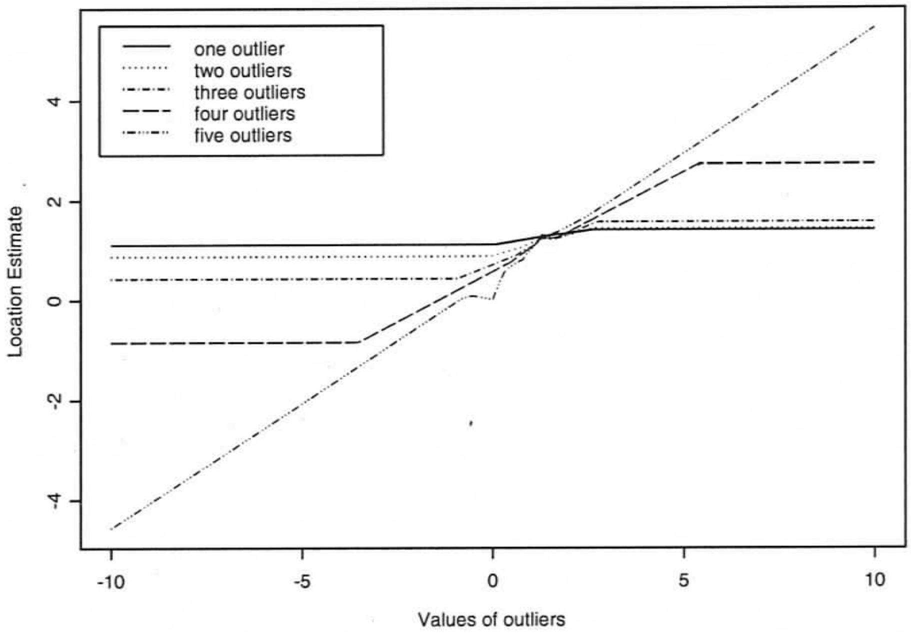
$$\sum_{i=1}^{n} \psi_c\left(\frac{x_i - T_n}{S_n}\right) = 0 \ .$$

Here we begin with $T_n^{(0)}$ and $S_n^{(0)}$ as above. Then perform at least one Newton step, that is, one iteration of

$$T_n^{(m+1)} = T_n^{(m)} + \frac{S_n^{(0)} \sum_{i=1}^{n} \psi_c\left(\frac{x_i - T_n^{(m)}}{S_n^{(0)}}\right)}{\sum_{i=1}^{n} \psi_c'\left(\frac{x_i - T_n^{(m)}}{S_n^{(0)}}\right)} \ .$$

For a finite sample $(x_1, \cdots, x_n)$, the breakdown point of the resulting one-step estimator T_n is $\lfloor \frac{n-1}{2} \rfloor / n$, which will approach 0.5 as the sample size n grows. If we use the data in Example 2.1, T_{10} has the finite-sample breakdown point $\epsilon_{10} = 4/10 = 0.4$. Figure 3.1 shows the behavior of T_{10} for the sample with different numbers of outliers. The estimate becomes unbounded when the sample contains 5 outliers, which agrees with its finite-sample breakdown point of 0.4, or say T_n can handle at most 4 outliers for $n = 10$.

Figure 3.1: The finite-sample breakdown point of T_{10}.



3.2 Rocke's M-estimators

Rocke (1983, 1991) developed a robust method to estimate the variance components σ_a^2 and σ_e^2 for model (1.1). The main idea of this method is to replace potential outliers with pseudo-observations, and then apply the standard ANOVA to the outlier-adjusted sample to estimate the variance components. Zhou and Zhu (2003) summarized the steps of the method as follows:

(R1) Obtain an initial estimate of σ_e. For example,

$$\overline{\mu + \alpha_i} = median_j(y_{ij}), \ \tilde{\sigma}_e = 1.4826\, median_{ij}(|y_{ij} - \overline{\mu + \alpha_i}|).$$

(R2) Obtain M-estimates $\widehat{\mu+\alpha_1}, \cdots, \widehat{\mu+\alpha_I}$ by solving the system of equations

$$\sum_{j=1}^{J} \psi\Big(\frac{y_{ij} - (\widehat{\mu+\alpha_i})}{\tilde{\sigma}_e}\Big) = 0, \qquad i = 1, \cdots, I,$$

where ψ is a bounded antisymmetric function. For example, it could be the Huber's function $\psi_c(x) = x \cdot \min(1,\ c/|x|)$ with a suitable tuning constant c.

(R3) Compute $\hat{\mu}$ by solving the following equation

$$\sum_{i=1}^{I} \psi\Big(\frac{\widehat{\mu+\alpha_i} - \hat{\mu}}{\tilde{\sigma}_a}\Big) = 0,$$

where $\tilde{\sigma}_a$ is a scale estimate based on $\widehat{\mu+\alpha_1}, \cdots, \widehat{\mu+\alpha_I}$, for example, $\tilde{\sigma}_a = MAD(\widehat{\mu+\alpha_i}) = 1.483\, median_i\, |\widehat{\mu+\alpha_i} - median_j(\widehat{\mu+\alpha_j})|$.

(R4) Compute pseudo-observations $\hat{y}_{ij}$.

$$\hat{y}_{ij} = \hat{\mu} + K_1\, \tilde{\sigma}_a\, \psi\Big(\frac{\hat{\alpha}_i}{\tilde{\sigma}_a}\Big) + K_2\, \tilde{\sigma}_e\, \psi\Big(\frac{y_{ij} - (\widehat{\mu+\alpha_i})}{\tilde{\sigma}_e}\Big),$$

where $\hat{\alpha}_i = \widehat{\mu+\alpha_i} - \hat{\mu}$, and where K_1 is the correction factor associated with the location M-estimate applied to the random effects in model (1.1) and K_2 is the correction factor associated with the main M-estimation of the parameters. Following the idea in Huber (1981), Zhou and Zhu set the the correction factors

$$\begin{aligned} K_1 &= 1 + \frac{1}{I}\frac{1 - m_1}{m_1}, \ with \ \ m_1 = \frac{1}{I}\sum_{i=1}^{I} \psi'\Big(\frac{\hat{\alpha}_i}{\tilde{\sigma}_a}\Big), \\ K_2 &= 1 + \frac{1}{IJ}\frac{1 - m_2}{m_2}, \ with \ \ m_2 = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} \psi'\Big(\frac{y_{ij} - (\widehat{\mu+\alpha_i})}{\tilde{\sigma}_e}\Big), \end{aligned}$$

where ψ' denotes the first derivative of ψ function.

(R5) Estimate σ_e^2 and σ_a^2 by applying a standard analysis of variance (ANOVA) to these pseudo-observations $\hat{y}_{ij}$ in $R4$. In summary, the estimates of σ_e^2 and σ_a^2 are computed by

$$\begin{aligned} \hat{\sigma}_e^2 &= \tilde{\sigma}_e^2 K_2^2 \frac{1}{I(J-1)} \sum_{i=1}^{I} \sum_{j=1}^{J} \left[\psi\Big(\frac{y_{ij} - (\widehat{\mu + \alpha_i})}{\tilde{\sigma}_e}\Big)\right]^2, \\ \hat{\sigma}_a^2 &= \tilde{\sigma}_a^2 K_1^2 \frac{1}{I-1} \sum_{i=1}^{I} \left[\psi\Big(\frac{\hat{\alpha}_i}{\tilde{\sigma}_a}\Big)\right]^2 - \frac{\hat{\sigma}_e^2}{J}. \end{aligned}$$

As mentioned above, the breakdown point is an important robustness measure for robust estimators, since it yields the largest proportion of observations, if replaced by outliers, that would not lead to a breakdown. In the one-way random effects model, there are two types of outlier contamination: group contamination and measurement contamination. Group contamination occurs when all observations in one group are contaminated due to systematic bias or error; measurement contamination occurs when observations in different groups are individually contaminated due to measurement error. Usually, both types of contamination can cause an estimator to approach zero or infinity. In that case, the estimator can not provide us with reliable information, i. e. , it breaks down. The breakdown at zero is called implosion, and the breakdown at infinity is called explosion.

The finite-sample breakdown points of Rocke's M-estimators have also been studied by Zhou and Zhu (2003). They presented a theorem for calculating the maximum number of outliers that Rocke's M-estimators can handle without group or measurement explosions. The theorem says that for model (1.1), the finite-sample breakdown points (explosion) of Rocke's estimators are

$$\begin{aligned}\epsilon_R^m &= \frac{1}{IJ}\left\{\left\lfloor\frac{J-1}{2}\right\rfloor + \left\lfloor\frac{J+1}{2}\right\rfloor\cdot\left\lfloor\frac{I-1}{2}\right\rfloor\right\},\\ \epsilon_R^g &= \frac{1}{I}\left\{\left\lfloor\frac{I-1}{2}\right\rfloor\right\},\end{aligned}$$

where $\epsilon_R^m = \min(\epsilon_R^m(\hat{\sigma}_a^2),\ \epsilon_R^m(\hat{\sigma}_e^2))$ is the measurement explosion breakdown point, and $\epsilon_R^g = \min(\epsilon_R^g(\hat{\sigma}_a^2),\ \epsilon_R^g(\hat{\sigma}_e^2))$ is the group explosion breakdown point. According to the calculation of Rocke's M-estimators in $R5$, the breakdowns of $\tilde{\sigma}_a$, $\tilde{\sigma}_e$, K_1, and K_2 are attributed to the breakdowns of $\hat{\sigma}_a^2$ and $\hat{\sigma}_e^2$ since the function ψ is bounded. Notice that the median estimator and the M-estimator can stand at most $\lfloor\frac{IJ-1}{2}\rfloor$ outliers for a sample of size IJ. Therefore, $\tilde{\sigma}_e$ in $R1$ is finite if there are at most $\lfloor\frac{IJ-1}{2}\rfloor$ measurement outliers; in $R2$, $|\widehat{\mu+\alpha_i}|$ is finite if there are at most $\lfloor\frac{J-1}{2}\rfloor$ outliers within each group; in $R3$, $|\hat{\mu}|$ and $\tilde{\sigma}_a$ are finite if there are at most $\lfloor\frac{I-1}{2}\rfloor$ outliers in $|\widehat{\mu+\alpha_i}|$, $i = 1, \cdots, I$. The measurement breakdown point of $\tilde{\sigma}_a$ will be determined by the maximum number of outliers that induce the worst outlier distribution among groups: there are $\lfloor\frac{J-1}{2}\rfloor$ outliers within one group, and $\lfloor\frac{J+1}{2}\rfloor$ outliers within each of other $\lfloor\frac{I-1}{2}\rfloor$ groups, i. e. , $\tilde{\sigma}_a$ can provide a reliable estimate given at most $\lfloor\frac{J-1}{2}\rfloor + \lfloor\frac{J+1}{2}\rfloor\cdot\lfloor\frac{I-1}{2}\rfloor$ outliers occurring in the sample. It is clear that the value of $\lfloor\frac{J-1}{2}\rfloor + \lfloor\frac{J+1}{2}\rfloor\cdot\lfloor\frac{I-1}{2}\rfloor$ is less than the value of $\lfloor\frac{IJ-1}{2}\rfloor$. Meanwhile, the correction factors K_1 and K_2 in $R4$ are finite since the first derivative of Huber's ψ function holds values of 0 or 1. Thus the measurement explosion breakdown point ϵ_R^m is equal to $\frac{1}{IJ}\left\{\lfloor\frac{J-1}{2}\rfloor + \lfloor\frac{J+1}{2}\rfloor\cdot\lfloor\frac{I-1}{2}\rfloor\right\}$. It is easy to work out the group explosion breakdown point since both $\tilde{\sigma}_e$ and $\tilde{\sigma}_a$ can deal with at most $\lfloor \frac{I-1}{2} \rfloor$ contaminated groups. Thus the group explosion breakdown point ϵ_R^g is equal to $\frac{1}{I}\left\{\lfloor \frac{I-1}{2} \rfloor\right\}$.

3.3 Rousseeuw and Croux's Q-estimator

Based on the Q-estimator of Rousseeeuw and Croux (1992, 1993), Uhlig (1997) proposed three new robust variance component estimators for the one-way random effects model. Similar to the approach used in studying Rocke's M-estimators, we first introduce the original Q-estimator of Rousseeuw and Croux.

Rousseeuw and Croux (1992, 1993) constructed several new scale estimators, which have a set of desirable properties: a bounded influence function, a 50% breakdown point, and an affordable computation time. These estimators all have explicit formulas: they can be written as arithmetic operations (such as sums, differences, and absolute values) in combinations with medians and other order statistics. The Q-estimator of a finite sample $(x_1, \cdots, x_n)$ is given by

$$Q_n(x_1, \cdots, x_n) = 2.2219\,\{|x_i - x_j|\}_{(k)}\,, \qquad 1 \le i, j \le n\,, \tag{3.1}$$

where 2.2219 is a tuning constant that makes the estimator consistent for the scale parameter of the standard normal distribution, and $k = \binom{h}{2} \approx \binom{n}{2}/4$, and $h = \lfloor n/2 \rfloor + 1$ is roughly half the number of observations. The estimate is actually the kth order statistic of the $\binom{n}{2}$ pairwise differences from the sample. Recall that the finite-sample breakdown point ϵ_n of an scale estimator is equal to the smaller one of the implosion breakdown point ϵ_n^i and the explosion breakdown point ϵ_n^e: $\epsilon_n = \min\{\epsilon_n^i,\ \epsilon_n^e\}$.

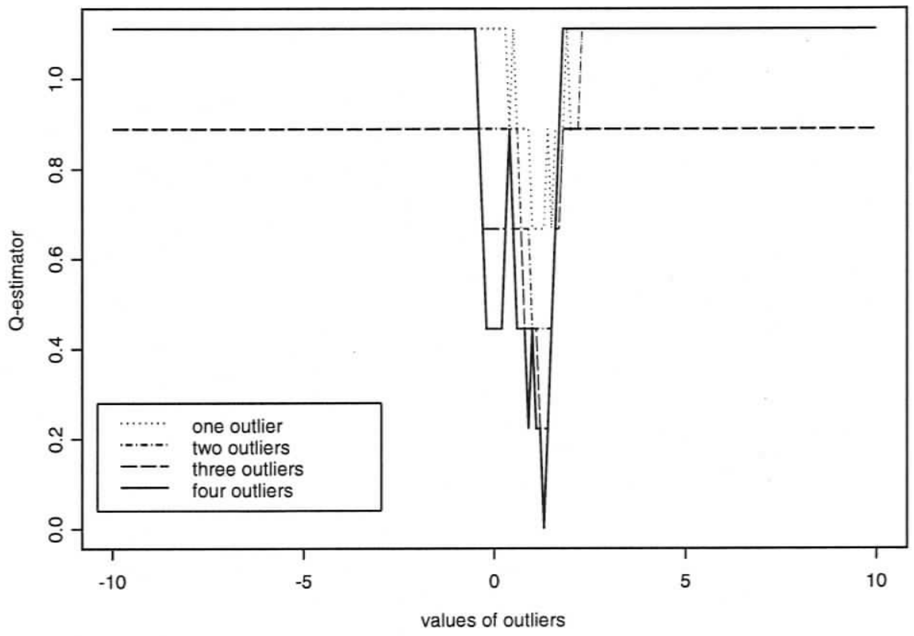
To derive the implosion breakdown point of the Q-estimator, we consider the most extreme case of outliers of a sample $(x_1, \cdots, x_n)$ in which all observations are pairwise different. We assume M observations are corrupted so that $M+1$ observations are equal. This implies that $M \cdot (M+1)/2$ pairwise differences are equal to zero. If we let

$$\frac{M(M+1)}{2} = \binom{h}{2} - 1 \approx \frac{n(n-1)}{8} - 1\,,$$

then the kth ordered pairwise difference will still be a non-zero value. Solving the above equation gives $M = \left\lfloor -0.5 + 0.5\sqrt{n^2 - n - 7} \right\rfloor$. If we use the sample of Example 2.1, M is equal to 4, and then the implosion breakdown point ϵ_{10}^i is equal to 4/10. Figure 3.2 shows the behavior of Q_{10} at the sample with a different number of outliers. The estimate reaches zero when the sample contains four outliers (actually, five outliers because two observations of the sample have exactly the same value of 1.3.), that is to say Q_{10} can deal with at most four outliers without implosion. As the sample size n increases, the implosion breakdown point $\epsilon_n^i = M/n$ would approach 0.5.

Similarly, we derive the explosion breakdown point by assuming M observations are corrupted, i. e. the values of the M observations go to infinity. Then $n - M$ observations are still finite. To ensure that the kth ordered pairwise difference is

Figure 3.2: The implosion breakdown point of Q_{10}

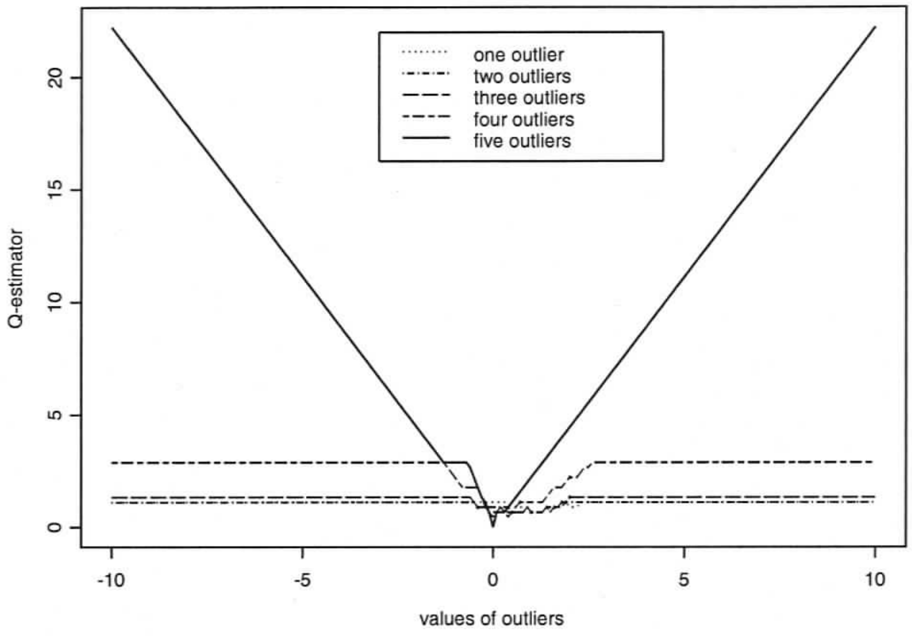


finite, the following equation must be satisfied:

$$\frac{(n-M)(n-M-1)}{2} = \frac{h(h-1)}{2} \approx \frac{n(n-1)}{8}$$

Solving the equation gives $M = \lceil n + 0.5 - 0.5\sqrt{n^2 + 7n + 1} \rceil$. Again, we use the data in Example 2.1 for demonstration. Through calculation, M is equal to 4, and then $\epsilon_{10}^{e} = 4/10$. Figure 3.3 shows the curves of Q_{10} at the sample with different outliers. The estimate becomes unbounded when five outliers are introduced into the data, that is to say Q_{10} can handle at most four outliers without explosion. Unlike the implosion situation, the coincidence of the observations has no significant influence on the capability of the estimator handling explosion outliers. For a large sample size, the explosion breakdown point $\epsilon_n^e = M/n$ would approach 0.5 as well.

Figure 3.3: The explosion breakdown point of Q_{10}



Generally speaking, the Q-estimator belongs to a class of location free estimators, namely estimators that do not rely on any initial location estimate. In contrast to the location-free estimators, the location-based estimators, for example, the MAD, start with computing a location estimate and then consider positive and negative deviations from it with assigned equal importance. Thus, the location-free estimators do not have the drawback of implicitly assuming the underlying distribution to be symmetric, and can be effectively implemented under the condition of an asymmetric distribution. Rousseeuw and Croux (1993) also showed the generalized Q-estimator:

$$Q_n^{(v)} = d_v \{SD(x_{i_1}, \cdots, x_{i_v}); i_1 < i_2 < \cdots < i_v\}_{\binom{h}{v}}\,, \tag{3.2}$$

where d_v is a constant factor, and $SD(x_{i_1}, \cdot, x_{i_v})$ denotes the standard deviation of the v observations $(x_{i_1}, \cdot, x_{i_v})$ and $h = \lfloor n/2 \rfloor + 1$. The Q_n is a special case of (3.2) where $v = 2$. Rousseeuw and Croux expected that more computational effort is required for higher order v.

3.4 Uhlig's Q-estimators

For the balanced one-way random effects model, Uhlig (1997) derived three robust variance component estimators from the Q-estimator of Rousseeuw and Croux (1992, 1993). The estimators of Uhlig, following the spirit of classical variance partition, split the set of all differences into differences within groups and differences between groups. They have relatively simple structures and can be calculated without iteration. Here, we introduce two of Uhlig's estimators, $S_{N,1}$ and $S_{N,2}$, which are used to estimate σ_e and $\sigma_y = \sqrt{\sigma_a^2 + \sigma_e^2}$ respectively.

Let Φ denote the distribution function for the standard normal distribution. The estimator $S_{N,1}$ of σ_e is given by

$$S_{N,1}(y_{11},\cdots,y_{IJ}) := \frac{1}{2\Phi^{-1}(\frac{5}{8})}\,H_{N,1}^{-1}(0.25) = 1.569\,H_{N,1}^{-1}(0.25),$$

where

$$H_{N,1}(x) = \frac{2}{I(I-1)}\sum_{1\le n<m\le I}\frac{2}{J^2(J-1)^2}\sum_{1\le i<j\le J}\sum_{1\le k<l\le J}[1_{\{|y_{ni}-y_{nj}-y_{mk}+y_{ml}|\le x\}} + 1_{\{|y_{ni}-y_{nj}+y_{mk}-y_{ml}|\le x\}}].$$

The function $H_{N,1}(x)$ can be interpreted as the distribution function of all absolute inter-group differences of the intra-group differences, which are also called second-order differences. In that case, the estimator $S_{N,1}$ corresponds to the lower quartile of these second-order differences.

In model (1.1), the variance of y_{ij} is $Var(y_{ij}) = \sigma_y^2 = \sigma_a^2 + \sigma_e^2$. The estimator $S_{N,2}$ for estimating σ_y is given by

$$S_{N,2}(y_{11},\cdots,y_{IJ}) := \frac{1}{\sqrt{2}\Phi^{-1}(\frac{5}{8})}\,H_{N,2}^{-1}(0.25) = 2.219\,H_{N,2}^{-1}(0.25),$$

where

$$H_{N,2}(x) = \frac{2}{I(I-1)}\sum_{1\le n<m\le I}\frac{1}{J^2}\sum_{i=1}^{J}\sum_{j=1}^{J}1_{\{|y_{ni}-y_{mj}|\le x\}}.$$

The function $H_{N,2}(x)$ can be interpreted as the distribution function of the absolute inter-group differences. Then, σ_a^2 can be estimated by

$$\hat{\sigma}_a^2 = max\{S_{N,2}^2 - \hat{\sigma}_e^2,\ 0\}.$$

The breakdown points of the two estimators have been well studied in Müller and Uhlig (2001). The following notation will be used throughout the rest of the chapters of this thesis. Suppose $S(y_{11}, \cdots, y_{IJ})$ is an estimator based on a sample $(y_{11}, \cdots, y_{IJ})$, where y_{ij} is the jth observation from the ith group, and the total number of observations is N. We denote $L^g_{N,M}(y_{11}, \cdots, y_{IJ})$ as the set of all samples $(y_{11}, \cdots, y_{IJ})$ with at most M contaminated groups among I groups, and $L^m_{N,M}(y_{11}, \cdots, y_{IJ})$ as the set of all samples with at most M contaminated observations (measurement outliers). For an estimator $S(y_{11}, \cdots, y_{IJ})$, we denote

$\epsilon^g_{N,1,S}$ as the group implosion breakdown point,

$\epsilon^g_{N,2,S}$ as the group explosion breakdown point,

$\epsilon^m_{N,1,S}$ as the measurement implosion breakdown point,

$\epsilon^m_{N,2,S}$ as the measurement explosion breakdown point.

The precise definitions for $\epsilon^g_{N,1,S}$, $\epsilon^g_{N,2,S}$, $\epsilon^m_{N,1,S}$, $\epsilon^m_{N,2,S}$ are given below:

$$\begin{aligned}
\epsilon^g_{N,1,S} &:= \frac{1}{I} max\{M, \inf_{L^g_{N,M}(y_{11},\cdots,y_{IJ})} S(y_{11}, \cdots, y_{IJ}) > 0\}, \\
\epsilon^g_{N,2,S} &:= \frac{1}{I} max\{M, \sup_{L^g_{N,M}(y_{11},\cdots,y_{IJ})} S(y_{11}, \cdots, y_{IJ}) < \infty\}, \\
\epsilon^m_{N,1,S} &:= \frac{1}{N} max\{M, \inf_{L^m_{N,M}(y_{11},\cdots,y_{IJ})} S(y_{11}, \cdots, y_{IJ}) > 0\}, \\
\epsilon^m_{N,2,S} &:= \frac{1}{N} max\{M, \sup_{L^m_{N,M}(y_{11},\cdots,y_{IJ})} S(y_{11}, \cdots, y_{IJ}) < \infty\}.
\end{aligned}$$

The group breakdown point and the measurement breakdown point of $S(y_{11}, \cdots, y_{IJ})$ are

$$\epsilon^g_{N,S} = \min\{\epsilon^g_{N,1,S}, \epsilon^g_{N,2,S}\} \quad and \quad \epsilon^m_{N,S} = \min\{\epsilon^m_{N,1,S}, \epsilon^m_{N,2,S}\} \,,$$

respectively.

Müller and Uhlig (2001) presented detailed derivations of the asymptotic breakdown points (as $I \to \infty$) of $S_{N,1}$ and $S_{N,2}$. For any sample in which all differences $|y_{ni} - y_{nj}|$ are pairwise different, the asymptotic group breakdown point of $S_{N,1}$ satisfies $\lim_{I\to\infty} \epsilon^{g}_{N,S_{N,1}} = 0.5$, and the asymptotic measurement breakdown point satisfies $\lim_{I\to\infty} \epsilon^{m}_{N,S_{N,1}} \geq 0.25$; for any sample in which all observations $\{y_{ni}\}$ are pairwise different, the asymptotic group breakdown point of $S_{N,2}$ satisfies $\lim_{I\to\infty} \epsilon^{g}_{N,S_{N,2}} = 0.5$, and the asymptotic measurement breakdown point satisfies $\lim_{I\to\infty} \epsilon^{m}_{N,S_{N,2}} = 0.5$. Moreover, they provided derivations of the finite-sample breakdown points of $S_{N,1}$ and $S_{N,2}$ in their technical report. Since the finite-sample breakdown point will be used as a constraint in the design problems of the next chapter, we briefly summarize the technical report.

The general idea Müller and Uhlig used to compute the finite-sample breakdown points is to maximize the number of intra or inter group differences that go to zero or infinity. According to the report, $S_{N,1}$ has no group implosion if and only if $G \leq \lceil -0.5 + 0.5\sqrt{I^2 - I + 1}\,\rceil$, and has no group explosion if and only if $G \leq \lfloor I - 0.5 - 0.5\sqrt{I^2 - I + 1}\,\rfloor$, where G is the number of contaminated groups that $S_{N,1}$ can handle without group implosion or group explosion. The maximum number of the contaminated groups can be obtained when the equalities hold. Let $G_1 = \lceil -0.5 + 0.5\sqrt{I^2 - I + 1}\,\rceil$, $G_2 = \lfloor I - 0.5 - 0.5\sqrt{I^2 - I + 1}\,\rfloor$. Then the group implosion and explosion breakdown points of $S_{N,1}$ are

$$\epsilon^g_{N,1,S_{N,1}} = \frac{G_1}{I} \quad and \quad \epsilon^g_{N,2,S_{N,1}} = \frac{G_2}{I},$$

respectively.

To present the measurement breakdown points of $S_{N,1}$, we define I_1, I_2, I_3 and I_4 as

$$\begin{aligned} I_1 &= \Big\lceil -1.5 + 0.5\sqrt{1 + \frac{I(I-1)J(J-1)}{2G_1} - 2J(J-1)(G_1-1)}\Big\rceil, \\ I_2 &= \lfloor J - 0.5 - 0.5\sqrt{2J^2 - 2J + 1}\rfloor, \\ I_3 &= J - I_2, \\ I_4 &= \Big\lfloor \frac{II_3 - I_3 + 1}{2} - \sqrt{\frac{(II_3 - I_3 + 1)^2}{4} - \frac{(I-1)II_3^2}{4} + \frac{(I-1)I(J-1)^2J^2}{16(I_3-1)^2}}\Big\rfloor. \end{aligned}$$

$S_{N,1}$ has no measurement implosion if and only if $M \le G_1(J-1) + I_1$ and has no measurement explosion if and only if $M \le II_2 + I_4$, where M is the number of contaminated observations that $S_{N,1}$ can handle without measurement implosion or measurement explosion. The maximum number of the contaminated observations can be obtained when the equalities hold. Let $M_1 = G_1(J-1) + I_1$, $M_2 = II_2 + I_4$. Then the measurement implosion and explosion breakdown points of $S_{N,1}$ are

$$\epsilon^m_{N,1,S_{N,1}} = \frac{M_1}{IJ} \quad and \quad \epsilon^m_{N,2,S_{N,1}} = \frac{M_2}{IJ},$$

respectively.

$S_{N,2}$ has the same group implosion and explosion breakdown points as $S_{N,1}$, i. e.

$$\epsilon^g_{N,1,S_{N,1}} = \frac{G_1}{I} \quad and \quad \epsilon^g_{N,2,S_{N,1}} = \frac{G_2}{I},$$

respectively.

To present the measurement breakdown points of $S_{N,2}$, we define J_1, J_2, and J_3 as

$$\begin{aligned} J_1 &= \left\lceil \frac{J-1}{2} \right\rceil, \\ J_2 &= \left\lceil \frac{-2IJ_1 + 2J_1 - 1}{2} + 0.5\sqrt{(2IJ_1 - 2J_1 - 1)^2 + (I^2 - I)(J - 4J_1^2)} \right\rceil, \\ J_3 &= \left\lfloor \frac{J(I - G_2)}{2} - \frac{(I-1)IJ}{8(I - G_2 - 1)} \right\rfloor. \end{aligned}$$

$S_{N,2}$ has no measurement implosion if and only if $M \leq IJ_1 + J_2 - 1$, and has no measurement explosion if and only if $M \leq G_2 J + J_3$, where M is the number of contaminated observations that $S_{N,2}$ can handle without measurement implosion or measurement explosion. The maximum number of the contaminated observations can be obtained when the equalities hold. Let $M_3 = IJ_1 + J_2 - 1$, $M_4 = G_2 J + J_3$. Then the measurement implosion and explosion breakdown points of $S_{N,2}$ are

$$\epsilon^m_{N,1,S_{N,2}} = \frac{M_3}{IJ} \quad and \quad \epsilon^m_{N,2,S_{N,2}} = \frac{M_4}{IJ},$$

respectively.

Based on the calculation formulas, we compute the finite-sample breakdown points of $S_{N,1}$ and $S_{N,2}$ for two sample sizes 40 (Table 3.1, Table 3.2) and 80 (Table 3.3, Table 3.4). All the designs in the tables are balanced, and the results indicate

(B1) the group explosion breakdown points are smaller than the group implosion breakdown points for $S_{N,1}$ and $S_{N,2}$;

Table 3.1: The finite-sample breakdown points of $S_{N,1}$, $N = 40$.

$(I,\ J)$	$\epsilon^{g}_{N,1,S_{N,1}}$	$\epsilon^{g}_{N,2,S_{N,1}}$	$\epsilon^{m}_{N,1,S_{N,1}}$	$\epsilon^{m}_{N,2,S_{N,1}}$	$\epsilon_{N,S_{N,1}}$	$\epsilon_{N,R}$
(2,20)	0.5	0	0.5	0.275	0	0
(4,10)	0.5	0.25	0.5	0.275	0.25	0.225
(5,8)	0.4	0.4	0.5	0.25	0.25	0.275
(8,5)	0.5	0.375	0.425	0.25	0.25	0.275
(10,4)	0.5	0.4	0.4	0.25	0.25	0.225
(20,2)	0.5	0.45	0.275	0.225	0.225	0.225

Table 3.2: The finite-sample breakdown points of $S_{N,2}$, $N = 40$.

$(I,\ J)$	$\epsilon^{g}_{N,1,S_{N,2}}$	$\epsilon^{g}_{N,2,S_{N,2}}$	$\epsilon^{m}_{N,1,S_{N,2}}$	$\epsilon^{m}_{N,2,S_{N,2}}$	$\epsilon_{N,S_{N,2}}$	$\epsilon_{N,R}$
(2,20)	0.5	0	0.475	0.375	0	0
(4,10)	0.5	0.25	0.475	0.425	0.25	0.225
(5,8)	0.4	0.4	0.475	0.45	0.4	0.275
(8,5)	0.5	0.375	0.475	0.45	0.375	0.275
(10,4)	0.5	0.4	0.475	0.475	0.4	0.225
(20,2)	0.5	0.45	0.475	0.475	0.45	0.225

Table 3.3: The finite-sample breakdown points of $S_{N,1}$, $N = 80$.

$(I,\ J)$	$\epsilon^{g}_{N,1,S_{N,1}}$	$\epsilon^{g}_{N,2,S_{N,1}}$	$\epsilon^{m}_{N,1,S_{N,1}}$	$\epsilon^{m}_{N,2,S_{N,1}}$	$\epsilon_{N,S_{N,1}}$	$\epsilon_{N,R}$
(2,40)	0.5	0	0.5	0.2875	0	0
(4,20)	0.5	0.25	0.5	0.275	0.25	0.2375
(5,16)	0.4	0.4	0.5	0.275	0.275	0.2875
(8,10)	0.5	0.375	0.5	0.275	0.275	0.2375
(10,8)	0.5	0.4	0.475	0.2625	0.2625	0.2875
(16,5)	0.5	0.4375	0.4125	0.2625	0.2625	0.2875
(20,4)	0.5	0.45	0.3875	0.25	0.25	0.2375
(40,2)	0.5	0.475	0.25	0.2375	0.2375	0.2375

Table 3.4: The finite-sample breakdown points of $S_{N,2}$, $N = 80$.

$(I,\ J)$	$\epsilon^{g}_{N,1,S_{N,2}}$	$\epsilon^{g}_{N,2,S_{N,2}}$	$\epsilon^{m}_{N,1,S_{N,2}}$	$\epsilon^{m}_{N,2,S_{N,2}}$	$\epsilon_{N,S_{N,2}}$	$\epsilon_{N,R}$
(2,40)	0.5	0	0.4875	0.375	0	0
(4,20)	0.5	0.25	0.4875	0.4375	0.25	0.2375
(5,16)	0.4	0.4	0.4875	0.45	0.4	0.2875
(8,10)	0.5	0.375	0.4875	0.4625	0.375	0.2375
(10,8)	0.5	0.4	0.4875	0.475	0.4	0.2875
(16,5)	0.5	0.4375	0.4875	0.475	0.4375	0.2875
(20,4)	0.5	0.45	0.4875	0.4875	0.45	0.2375
(40,2)	0.5	0.475	0.4875	0.4875	0.475	0.2375

(B2) the measurement explosion breakdown points are smaller than the measurement implosion breakdown points for $S_{N,1}$ and $S_{N,2}$;

(B3) the group breakdown point of $S_{N,1}(\min\{\epsilon^g_{N,1,S_{N,1}}, \epsilon^g_{N,2,S_{N,1}}\})$ is the same as that of $S_{N,2}(\min\{\epsilon^g_{N,1,S_{N,2}}, \epsilon^g_{N,2,S_{N,2}}\})$;

(B4) the measurement breakdown point of $S_{N,1}(\min\{\epsilon^m_{N,1,S_{N,1}}, \epsilon^m_{N,2,S_{N,1}}\})$ is smaller than that of $S_{N,2}(\min\{\epsilon^m_{N,1,S_{N,2}}, \epsilon^m_{N,2,S_{N,2}}\})$;

To summarize, the finite-sample breakdown points $\epsilon_{N,S_{N,1}}$ and $\epsilon_{N,S_{N,2}}$ are determined by the explosion breakdown points, and $S_{N,2}$ has higher breakdown points than $S_{N,1}$. If one is interested in the finite-sample breakdown point of a robust estimation procedure using S_1 and S_2 to estimate σ_a, σ_e,and σ_y, then it will equal the explosion breakdown point of $S_{N,1}$. For comparison, we also provide the finite-sample breakdown points of Rocke's M-estimators, $\epsilon_{N,R}$, in the tables, which were computed by Theorem 1 of Zhou and Zhu (2003). It is clear from the last two columns of data that $S_{N,1}$ and Rocke's M-estimators have fairly close finite-sample breakdown points for most values of I and J; when $I = 2$, they all have zero breakdown points.

Chapter 4

Optimal Experimental Designs

In Chapter 1, we reviewed several research papers on experimental designs under the one-way random effects model. It includes some of the work in the area of optimal experimental designs associated with the traditional methods of estimation of the variance components, such as the analysis of variance (ANOVA), the maximum likelihood (ML), and the modified maximum likelihood (MML). Generally speaking, a design problem is to obtain accurate estimates of the variance components by choosing the number of groups I and the number of observations J_i ($J_1 + J_2 + \cdots + J_I = N$) within each group. For the most part, the design criteria pertaining to these traditional methods focus on minimizing the variances of the estimates of the variance components σ_a^2, or σ_e^2, or other related model parameters. In this chapter, we introduce adapted design criteria from the perspective of robustness and construct optimal robust designs based on Rocke's M-estimators and Uhlig's

Q-estimators. The resulting designs are compared with the optimal designs based on the standard ANOVA estimators.

4.1 Optimal ANOVA Designs

As we know, the choice of optimal designs depends on some method of estimation as well as design criterion. The setup of a design criterion is directly determined by a specific purpose that the experiment tries to achieve. When the standard ANOVA estimation or other ANOVA derivatives are carried out in the one-way random effects model, the design criterion is commonly set as the minimization of the variances of estimates of the variance components σ_a^2 and σ_e^2 or other related model parameters, such as $(\sigma_a^2 + \sigma_e^2)$, σ_a^2/σ_e^2, and $\sigma_a^2/(\sigma_a^2 + \sigma_e^2)$. The meaningful ratio $\sigma_a^2/(\sigma_a^2 + \sigma_e^2)$ is of interest to experimenters because it reflects the proportion of the variance of an observation ($Var(y_{ij}) = \sigma_a^2 + \sigma_e^2$) that is the result of differences between groups. Let $\hat{\sigma}_a^2$ and $\hat{\sigma}_e^2$ denote the ANOVA estimates of σ_a^2 and σ_e^2, respectively. To demonstrate the choice of optimal ANOVA designs, we derive the expressions for the variances of $\hat{\sigma}_a^2$, $\hat{\sigma}_e^2$, $(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$, and $\hat{\sigma}_a^2/(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$ under the balanced one-way random effects model. Recall that

$$\begin{aligned}\hat{\sigma}_e^2 &= \frac{SSE}{I(J-1)},\\ \hat{\sigma}_a^2 &= \frac{SSA}{J(I-1)} - \frac{\hat{\sigma}_e^2}{J},\end{aligned}$$

where SSA is called the sum of squares due to groups (i. e., between groups), and SSE is called the sum of squares due to error (i. e., within groups). According to Montgomery (1991), $SSA/(\sigma_e^2 + J\sigma_a^2)$ and SSE/σ_e^2 are independently distributed as chi-square random variables, i. e.

$$\begin{aligned} \frac{SSA}{\sigma_e^2 + J\sigma_a^2} &\sim \chi^2_{I-1}\,, \\ \frac{SSE}{\sigma_e^2} &\sim \chi^2_{I(J-1)}\,. \end{aligned}$$

This result gives insight into finding the variances. First,

$$\begin{aligned} Var(\hat{\sigma}_e^2) &= Var\left[\frac{SSE}{I(J-1)}\right] \\ &= Var\left[\frac{\sigma_e^2}{I(J-1)} \cdot \frac{SSE}{\sigma_e^2}\right] \\ &= \frac{\sigma_e^4}{I^2(J-1)^2} \cdot Var\left[\frac{SSE}{\sigma_e^2}\right] \\ &= \frac{\sigma_e^4}{I^2(J-1)^2} \cdot 2I(J-1) \\ &= \frac{2\sigma_e^4}{N-I}\,, \end{aligned} \tag{4.1}$$

where $N = I \cdot J$ is the total number of observations, and where $Var\left(SSE/\sigma_e^2\right) = 2I(J-1)$ since SSE/σ_e^2 is distributed as $\chi^2_{I(J-1)}$. For a fixed number of observations N, the optimal design minimizing $Var(\hat{\sigma}_e^2)$ requires that the number of groups I be the smallest number that the design can implement, since $Var(\hat{\sigma}_e^2)$ is an increasing function of I.

Second,

$$\begin{aligned}
Var(\hat{\sigma}_a^2) &= Var\left[\frac{SSA}{(I-1)J} - \frac{\hat{\sigma}_e^2}{J}\right] \\
&= Var\left[\frac{SSA}{(I-1)J} - \frac{SSE}{I(J-1)J}\right] \\
&= Var\left[\frac{\sigma_e^2 + J\sigma_a^2}{(I-1)J} \cdot \frac{SSA}{\sigma_e^2 + J\sigma_a^2} - \frac{\sigma_e^2}{I(J-1)J} \cdot \frac{SSE}{\sigma_e^2}\right] \\
&= \frac{(\sigma_e^2 + J\sigma_a^2)^2}{(I-1)^2 J^2} \cdot Var\left[\frac{SSA}{\sigma_e^2 + J\sigma_a^2}\right] + \frac{\sigma_e^4}{I^2(J-1)^2 J^2} \cdot Var\left[\frac{SSE}{\sigma_e^2}\right] \\
&= \frac{(\sigma_e^2 + J\sigma_a^2)^2}{(I-1)^2 J^2} \cdot 2(I-1) + \frac{\sigma_e^4}{I^2(J-1)^2 J^2} \cdot 2I(J-1) \\
&= \frac{2(\sigma_e^2 + J\sigma_a^2)^2}{(I-1)J^2} + \frac{2\sigma_e^4}{I(J-1)J^2} \\
&= \frac{2\sigma_e^4}{N^2} \cdot \left[\frac{(I+N\rho)^2}{I-1} + \frac{I^2}{N-I}\right] , \qquad (4.2)
\end{aligned}$$

where $\rho = \sigma_a^2/\sigma_e^2$. Assuming ρ is constant and N is fixed, the optimal design of minimizing $Var(\hat{\sigma}_a^2)$ requires that the number of groups I be the largest integer less than or equal to $N(2+N\rho)/(N\rho+N+1)$, which is the solution to the equation

$$\frac{\partial Var(\hat{\sigma}_a^2)}{\partial I} = \frac{2\sigma_e^4}{N^2}\left[\frac{N^2}{(N-I)^2} - \frac{(1+N\rho)^2}{(I-1)^2}\right] = 0 ,$$

where $\partial Var(\hat{\sigma}_a^2)/\partial I$ is the first derivative of $Var(\hat{\sigma}_a^2)$ with respect to I. The minimum of $Var(\hat{\sigma}_a^2)$ is obtained for this solution since the second derivative of $Var(\hat{\sigma}_a^2)$ with respect to I is

$$\frac{\partial^2 Var(\hat{\sigma}_a^2)}{\partial I^2} = \frac{4\sigma_e^4}{N^2}\left[\frac{N^2}{(N-I)^3} + \frac{(1+N\rho)^2}{(I-1)^3}\right] ,$$

which is greater than zero when I is greater than 1.

Third,

$$\begin{aligned} Var\left(\hat{\sigma}_a^2 + \hat{\sigma}_e^2\right) &= Var\left[\frac{SSA}{(I-1)J} - \frac{\hat{\sigma}_e^2}{J} + \hat{\sigma}_e^2\right] \\ &= Var\left[\frac{SSA}{(I-1)J} + \frac{SSE}{IJ}\right] \\ &= Var\left[\frac{\sigma_e^2 + J\sigma_a^2}{(I-1)J} \cdot \frac{SSA}{\sigma_e^2 + J\sigma_a^2} + \frac{\sigma_e^2}{IJ} \cdot \frac{SSE}{\sigma_e^2}\right] \\ &= \frac{(\sigma_e^2 + J\sigma_a^2)^2}{(I-1)^2 J^2} \cdot 2(I-1) + \frac{\sigma_e^4}{I^2 J^2} \cdot 2I(J-1) \\ &= \frac{2\sigma_e^4}{N^2} \cdot \left[\frac{(I+N\rho)^2}{I-1} + N - I\right] . \end{aligned} \tag{4.3}$$

The optimal design minimizing $Var(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$ requires that the number of groups I be the largest number that the design can implement, since the first derivative of $Var(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$ with respect to I is

$$\frac{\partial Var(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)}{\partial I} = -\frac{2\sigma_e^4}{N^2} \cdot \frac{(1+N\rho)^2}{(I-1)^2} ,$$

which is less than zero if I is greater than 1. In other words, the value of $Var(\hat{\sigma}_a^2+\hat{\sigma}_e^2)$ will decrease as the number of group I increases. Finally,

$$\begin{aligned} Var\left(\frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}_e^2}\right) &= Var\left[\frac{\frac{SSA}{(I-1)J} - \frac{SSE}{I(J-1)J}}{\frac{SSA}{(I-1)J} - \frac{SSE}{I(J-1)J} + \frac{SSE}{I(J-1)}}\right] \\ &= Var\left[1 - \frac{\frac{SSE}{I(J-1)}}{\frac{SSA}{(I-1)J} + \frac{SSE}{IJ}}\right] \\ &= Var\left[\frac{J}{\frac{I(J-1)}{(I-1)} \cdot \frac{SSA}{SSE} + J - 1}\right] \\ &= Var\left[\frac{J}{\frac{\sigma_e^2 + J\sigma_a^2}{\sigma_e^2} \cdot T + J - 1}\right] \\ &= Var\left[\frac{N}{(I+N\rho)T + N - I}\right] , \end{aligned} \tag{4.4}$$

where

$$T = \frac{\frac{SSA}{(I-1)(\sigma_e^2+J\sigma_a^2)}}{\frac{SSE}{I(J-1)\sigma_e^2}}$$

is a random variable distributed as F with $I-1$ and $N-I$ degrees of freedom. It is difficult to determine the explicit expression for the variance. Hence, the minimization of $Var(\hat{\sigma}_a^2/(\hat{\sigma_a^2}+\hat{\sigma_e^2}))$ is achieved in a slightly different way: we numerically compute the variance at various group number Is for a fixed N, and then select the design (or the number of group I) where the variance reaches the minimum.

4.2 Optimal Robust Designs

For the one-way random effects model, the well-known drawback of the standard ANOVA estimation is its sensitivity to model deviations caused by gross errors. Even mild deviations may have a significant impact on the estimation procedure. It can be shown that an ANOVA design that is optimal under the ideal model would not be optimal under the gross-errors model, while a design based on a robust estimation procedure is intent on providing reliable and efficient estimates of the parameters of interest under the gross-errors model. Consequently, a few questions arise: What is an optimality criterion for a robust design? Is a robust optimal design similar to the optimal design based on the ANOVA estimation under certain circumstances? Which design can produce better results? To answer these questions, we investigate robust designs derived from the application of Uhlig's Q-estimators and Rocke's

M-estimators. Let us first consider the design criterion for robust designs.

Since robust estimators of the variance components might be biased, it is desirable to have more general criteria that allow for the variances and the biases of the estimators to be considered. In this section, we denote a parameter of model (1.1) of interest as θ and some function of θ as $\tau(\theta)$.

Definition 4.1 *An estimator* $T_n(x_1, \cdots, x_n)$ *is said to be an unbiased estimator of* $\tau(\theta)$ *if*

$$E(T_n) = \tau(\theta) \ ,$$

where $E(T_n)$ *is the expected value of* $T_n(x_1, \cdots, x_n)$*. Otherwise, we say that* $T_n(x_1, \cdots, x_n)$ *is a biased estimator of* $\tau(\theta)$*, and the bias is given by*

$$b(T_n) = E(T_n) - \tau(\theta) \ .$$

Definition 4.2 *If* $T_n(x_1, \cdots, x_n)$ *is an estimator of* $\tau(\theta)$*, then the mean squared error (MSE) of* $T_n(x_1, \cdots, x_n)$ *is given by*

$$MSE(T_n) = E[T_n - \tau(\theta)]^2 \ .$$

It can be shown that the mean squared error of T_n is actually the sum of the variance of T_n and the square of the bias of T_n, i. e.

$$\begin{aligned} MSE(T_n) &= E[T_n - \tau(\theta)]^2 \\ &= E[T_n - E(T_n) + E(T_n) - \tau(\theta)]^2 \\ &= E[T_n - E(T_n)]^2 + 2[E(T_n) - \tau(\theta)] \times [E(T_n) - E(T_n)] + [E(T_n) - \tau(\theta)]^2 \\ &= Var(T_n) + [b(T_n)]^2 \ . \end{aligned}$$

The MSE is a reasonable measure that incorporates both the variance and the bias of an estimator, and it agrees with the variance criterion if we constrain ourselves to unbiased estimators, since the bias of an unbiased estimator is zero.

Of course, the properties of a robust estimator should be taken into account when we set an optimality criterion for robust designs. The finite-sample breakdown point should play a role of constraint on the design optimization to ensure the selected designs can deal with a specified number of outliers. In this chapter, we use the optimality criterion proposed by Zhou and Zhu (2003) to construct robust designs. For instance, suppose our interest is to obtain a robust optimal design for the variance component σ_a^2 under the one-way random effects model. We denote a robust estimator as $\hat{\sigma}_a^2$, and its finite-sample breakdown point as ϵ. Then we propose the following criterion for a design ξ with components $I, J_1, \cdots, J_I$ $(\Sigma_{i=1}^I J_i = N)$,

$$\begin{cases} \min_\xi MSE(\hat{\sigma}_a^2) \ , \\ s.\ t. \quad \epsilon \geq \gamma_0 \ , \end{cases}$$

where $MSE(\hat{\sigma}_a^2)$ is the mean squared error of the estimate, and where γ_0 is a pre-defined breakdown point that the implemented estimator $\hat{\sigma}_a^2$ is expected to achieve. The minimization of the $MSE(\hat{\sigma}_a^2)$, which is subject to the constraint that the finite-sample breakdown point ϵ must be greater than or equal to γ_0, is carried out under the model assumptions (the random effect α_i and the measurement error e_{ij} are independently normally distributed random variables). In other words, a design that minimizes the $MSE(\hat{\sigma}_a^2)$, but has an undesirable breakdown point, cannot be selected. If we are interested in constructing robust designs for the other model parameters, such as σ_e^2, σ_a^2/σ_e^2, and $\sigma_a^2/(\sigma_a^2+\sigma_e^2)$, the above criterion can be simply modified by minimizing the corresponding MSEs of the model parameters of interest.

However, an exact expression for the MSE of a robust estimator is usually difficult to determine, even under ideal models. Then we resort to a Monte Carlo simulation study for computing the MSE. The Monte Carlo simulation is a statistical method that utilizes sequences of random numbers to perform the simulation. The goal of the Monte Carlo method is to simulate the mathematical (or physical) system by sampling randomly from a specified discrete, or continuous probability distribution, and performing the necessary supplementary computations to describe the system evolution. For example, suppose we are interested in estimating $MSE(\hat{\sigma}_e^2)$, where $\hat{\sigma}_e^2$ is the estimate from a robust estimator. The simulation is performed with 5000 runs under the normal assumptions for α_i and ϵ_{ij}. We first generate 5000 random samples according to model (1.1); then, we compute the estimate $\hat{\sigma}_e^2$ for each of the 5000 samples using the robust estimator. At last, we compute the $MSE(\hat{\sigma}_e^2)$ with the formula:

$$MSE(\hat{\sigma}_e^2) = \frac{1}{5000}\sum_{k=1}^{5000}(\hat{\sigma}_{e,k}^2 - \sigma_{e,0}^2)^2 \ ,$$

where $\hat{\sigma}_{e,k}^2$ is the estimate of σ_e^2 at the kth run, and $\sigma_{e,0}^2$ is the true value of σ_e^2. Similarly, we can carry out simulations to compute $MSE(\hat{\sigma}_a^2)$, $MSE(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$, and $MSE(\hat{\sigma}_a^2/(\hat{\sigma}_a^2 + \hat{\sigma}_e^2))$. We report the simulation results for $MSE(\hat{\sigma}_a^2)$, $MSE(\hat{\sigma}_e^2)$, $MSE(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$, and $MSE(\hat{\sigma}_a^2/(\hat{\sigma}_a^2 + \hat{\sigma}_e^2))$ for two sample sizes of 40 and 80 in a series of tables (Table 4.1 – Table 4.8). To compare the choice of designs based on different types of robust estimators, we compute in simulation the estimates $\hat{\sigma}_a^2$, $\hat{\sigma}_e^2$, $\hat{\sigma}_a^2 + \hat{\sigma}_e^2$, and $\hat{\sigma}_a^2/(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$ by Rocke's M-estimators and Uhlig's Q-estimators (Refer to Chapter 3 for details about the two types of estimators). Without loss of generality, we can set the true value of σ_e^2 as the constant "1" when implementing the simulations under the normal assumption. Hence, given a value of $\rho = \sigma_a^2/\sigma_e^2$, the true value of σ_a^2 is equal to ρ. We also provide the MSEs of the standard ANOVA estimates for each design, among which $MSE(\hat{\sigma}_e^2)$, $MSE(\hat{\sigma}_a^2)$, and $MSE(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$ are calculated by (4.1), (4.2) and (4.3), respectively, because the MSEs are equal to the variances provided that their ANOVA estimators are unbiased. Note that each of the tables consists of three sub-tables, which illustrate the MSEs of the estimates of the model parameters by $A(\bullet)$ (the standard ANOVA estimation), $M(\bullet)$ (Rocke's M-estimation), and $Q(\bullet)$ (Uhlig's Q-estimation). All simulations are implemented under the balanced one-way random effects model; the design groups are the factors of 40 or 80. The finite-sample breakdown points of designs are reported in the last column of each table.

Table 4.1: Simulated $MSE(\hat{\sigma}_a^2)$ for $N = 40$

$A(\hat{\sigma}_a^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{40}
I = 2	0.0073	0.0451	0.6051	2.2051	51.0051	0
I = 4	0.0086	0.0272	0.2406	0.8072	17.3405	0
I = 5	0.0100	0.0262	0.1962	0.6337	13.1337	0
I = 8	0.0151	0.0282	0.1425	0.4139	7.7282	0
I = 10	0.0192	0.0314	0.1292	0.3514	6.1291	0
I = 20	0.0524	0.0629	0.1303	0.2618	3.2092	0

$M(\hat{\sigma}_a^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{40}
I = 2	0.0066	0.0408	0.6069	2.1178	50.8272	0
I = 4	0.0063	0.0244	0.2187	0.7448	17.6603	0.225
I = 5	0.0062	0.0214	0.1920	0.5627	12.3750	0.275
I = 8	0.0116	0.0226	0.1401	0.4171	8.5040	0.275
I = 10	0.0151	0.0305	0.1266	0.3646	5.9860	0.225
I = 20	0.0264	0.0375	0.1148	0.2715	3.5001	0.225

$Q(\hat{\sigma}_a^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{40}
I = 2	0.0296	0.1558	9.2371	65.7101	3512.2007	0
I = 4	0.0193	0.0524	0.6449	2.9406	104.7563	0.25
I = 5	0.0172	0.0459	0.4055	1.6504	60.6091	0.25
I = 8	0.0217	0.0366	0.2523	0.6766	20.4193	0.25
I = 10	0.0266	0.0450	0.2392	0.5442	13.0406	0.25
I = 20	0.0473	0.0589	0.1919	0.4262	5.1527	0.225

Table 4.2: Simulated $MSE(\hat{\sigma}_e^2)$ for $N = 40$

$A(\hat{\sigma}_e^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{40}
I = 2	0.0526	0.0526	0.0526	0.0526	0.0526	0
I = 4	0.0556	0.0556	0.0556	0.0556	0.0556	0
I = 5	0.0571	0.0571	0.0571	0.0571	0.0571	0
I = 8	0.0625	0.0625	0.0625	0.0625	0.0625	0
I = 10	0.0667	0.0667	0.0667	0.0667	0.0667	0
I = 20	0.1000	0.1000	0.1000	0.1000	0.1000	0

$M(\hat{\sigma}_e^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{40}
I = 2	0.0632	0.0658	0.0651	0.0660	0.0667	0
I = 4	0.0725	0.0723	0.0713	0.0697	0.0745	0.225
I = 5	0.0818	0.0887	0.0820	0.0828	0.0836	0.275
I = 8	0.1051	0.1045	0.1008	0.1101	0.1070	0.275
I = 10	0.1131	0.1120	0.1103	0.1111	0.1106	0.225
I = 20	0.1164	0.1058	0.1028	0.1094	0.1124	0.225

$Q(\hat{\sigma}_e^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{40}
I = 2	0.0654	0.0588	0.0620	0.0626	0.0625	0
I = 4	0.0610	0.0666	0.0670	0.0675	0.0668	0.25
I = 5	0.0734	0.0663	0.0690	0.0647	0.0700	0.25
I = 8	0.0748	0.0799	0.0801	0.0756	0.0775	0.25
I = 10	0.0895	0.0863	0.0846	0.0872	0.0868	0.25
I = 20	0.1745	0.1736	0.1607	0.1679	0.1662	0.225

Table 4.3: Simulated $MSE(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$ for $N = 40$

$A(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{40}
I = 2	0.0547	0.0925	0.6525	2.2525	51.0525	0
I = 4	0.0531	0.0717	0.2850	0.8517	17.3850	0
I = 5	0.0529	0.0691	0.2391	0.6766	13.1766	0
I = 8	0.0526	0.0657	0.1800	0.4514	7.7657	0
I = 10	0.0525	0.0647	0.1625	0.3847	6.1625	0
I = 20	0.0524	0.0629	0.1303	0.2618	3.2092	0

$M(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{40}
I = 2	0.0609	0.0978	0.6692	2.1848	50.9225	0
I = 4	0.0654	0.0845	0.3078	0.8298	17.8684	0.225
I = 5	0.0715	0.0986	0.2811	0.6855	12.6547	0.275
I = 8	0.0805	0.0982	0.2271	0.5324	8.7662	0.275
I = 10	0.0831	0.0958	0.2109	0.4759	6.3016	0.225
I = 20	0.0825	0.0820	0.1443	0.3059	3.5609	0.225

$Q(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{40}
I = 2	0.1087	0.2318	9.3560	65.7631	3514.2790	0
I = 4	0.0878	0.1216	0.7198	3.0279	105.1320	0.25
I = 5	0.0879	0.1144	0.4614	1.7320	60.8689	0.25
I = 8	0.0820	0.0999	0.3061	0.7123	20.5417	0.25
I = 10	0.0833	0.1018	0.2907	0.5755	13.2342	0.25
I = 20	0.0831	0.0939	0.1894	0.4126	5.1527	0.225

Table 4.4: Simulated $MSE(\frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2+\hat{\sigma}_e^2})$ for $N = 40$

$A(\frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2+\hat{\sigma}_e^2})$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{40}
I = 2	0.0055	0.0189	0.0706	0.1108	0.1660	0
I = 4	0.0076	0.0159	0.0422	0.0559	0.0455	0
I = 5	0.0092	0.0162	0.0369	0.0468	0.0310	0
I = 8	0.0136	0.0192	0.0321	0.0325	0.0138	0
I = 10	0.0179	0.0218	0.0307	0.0301	0.0106	0
I = 20	0.0494	0.0492	0.0400	0.0312	0.0062	0

$M(\frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2+\hat{\sigma}_e^2})$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{40}
I = 2	0.0059	0.0188	0.0660	0.1070	0.1503	0
I = 4	0.0071	0.0161	0.0459	0.0668	0.0693	0.225
I = 5	0.0080	0.0167	0.0442	0.0597	0.0512	0.275
I = 8	0.0164	0.0217	0.0390	0.0416	0.0189	0.275
I = 10	0.0210	0.0285	0.0382	0.0376	0.0134	0.225
I = 20	0.0302	0.0308	0.0397	0.0387	0.0101	0.225

$Q(\frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2+\hat{\sigma}_e^2})$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{40}
I = 2	0.0126	0.0273	0.0859	0.1269	0.1566	0
I = 4	0.0108	0.0191	0.0514	0.0664	0.0419	0.25
I = 5	0.0101	0.0174	0.0479	0.0571	0.0282	0.25
I = 8	0.0130	0.0177	0.0396	0.0417	0.0162	0.25
I = 10	0.0177	0.0211	0.0391	0.0394	0.0120	0.25
I = 20	0.0355	0.0341	0.0497	0.0445	0.0099	0.225

Table 4.5: Simulated $MSE(\hat{\sigma}_a^2)$ for $N = 80$

$A(\hat{\sigma}_a^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{80}
I = 2	0.0025	0.0313	0.5513	2.1013	50.5013	0
I = 4	0.0025	0.0151	0.2017	0.7351	17.0017	0
I = 5	0.0027	0.0133	0.1583	0.5646	12.8146	0
I = 8	0.0037	0.0117	0.1031	0.3460	7.4317	0
I = 10	0.0045	0.0117	0.0873	0.2817	5.8373	0
I = 16	0.0071	0.0133	0.0666	0.1933	3.6066	0
I = 20	0.0092	0.0150	0.0613	0.1666	2.9034	0
I = 40	0.0258	0.0310	0.0638	0.1279	1.5638	0

$M(\hat{\sigma}_a^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{80}
I = 2	0.0020	0.0291	0.5327	1.9855	45.7686	0
I = 4	0.0018	0.0130	0.2012	0.6872	16.0389	0.2375
I = 5	0.0018	0.0116	0.1529	0.5560	14.3268	0.2875
I = 8	0.0024	0.0096	0.1033	0.3475	7.9505	0.2375
I = 10	0.0029	0.0106	0.0886	0.3075	5.9307	0.2875
I = 16	0.0065	0.0124	0.0685	0.2125	3.8740	0.2875
I = 20	0.0085	0.0141	0.0644	0.1683	3.1146	0.2375
I = 40	0.0130	0.0193	0.0655	0.1366	1.8334	0.2375

$Q(\hat{\sigma}_a^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{80}
I = 2	0.0069	0.1087	7.6589	50.5793	3398.7996	0
I = 4	0.0046	0.0237	0.4486	2.3072	102.3091	0.25
I = 5	0.0054	0.0224	0.3533	1.4092	55.9571	0.275
I = 8	0.0062	0.0164	0.1633	0.5979	15.8687	0.275
I = 10	0.0056	0.0169	0.1387	0.4703	12.3954	0.2625
I = 16	0.0083	0.0167	0.0967	0.2786	5.8539	0.2625
I = 20	0.0093	0.0202	0.0914	0.2423	4.5518	0.25
I = 40	0.0206	0.0303	0.0905	0.1762	2.1727	0.2375

Table 4.6: Simulated $MSE(\hat{\sigma}_e^2)$ for $N = 80$

$A(\hat{\sigma}_e^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{80}
I = 2	0.0256	0.0256	0.0256	0.0256	0.0256	0
I = 4	0.0263	0.0263	0.0263	0.0263	0.0263	0
I = 5	0.0267	0.0267	0.0267	0.0267	0.0267	0
I = 8	0.0278	0.0278	0.0278	0.0278	0.0278	0
I = 10	0.0286	0.0286	0.0286	0.0286	0.0286	0
I = 16	0.0313	0.0313	0.0313	0.0313	0.0313	0
I = 20	0.0333	0.0333	0.0333	0.0333	0.0333	0
I = 40	0.0500	0.0500	0.0500	0.0500	0.0500	0

$M(\hat{\sigma}_e^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{80}
I = 2	0.0359	0.0337	0.0355	0.0343	0.0351	0
I = 4	0.0366	0.0365	0.0395	0.0365	0.0381	0.2375
I = 5	0.0378	0.0388	0.0393	0.0399	0.0384	0.2875
I = 8	0.0462	0.0435	0.0460	0.0450	0.0437	0.2375
I = 10	0.0511	0.0527	0.0518	0.0478	0.0497	0.2875
I = 16	0.0754	0.0761	0.0753	0.0746	0.0772	0.2875
I = 20	0.0777	0.0778	0.0751	0.0774	0.0742	0.2375
I = 40	0.0613	0.0632	0.0574	0.0593	0.0595	0.2375

$Q(\hat{\sigma}_e^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{80}
I = 2	0.0276	0.0274	0.0274	0.0273	0.0309	0
I = 4	0.0276	0.0279	0.0300	0.0297	0.0300	0.25
I = 5	0.0315	0.0300	0.0300	0.0311	0.0311	0.275
I = 8	0.0334	0.0307	0.0310	0.0323	0.0301	0.275
I = 10	0.0317	0.0326	0.0320	0.0344	0.0342	0.2625
I = 16	0.0385	0.0384	0.0365	0.0351	0.0366	0.2625
I = 20	0.0426	0.0429	0.0386	0.0399	0.0404	0.25
I = 40	0.0682	0.0685	0.0737	0.0709	0.0698	0.2375

Table 4.7: Simulated $MSE(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$ for $N = 80$

$A(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{80}
I = 2	0.0268	0.0556	0.5756	2.1256	50.5256	0
I = 4	0.0262	0.0388	0.2254	0.7588	17.0254	0
I = 5	0.0261	0.0366	0.1816	0.5879	12.8379	0
I = 8	0.0260	0.0339	0.1254	0.3682	7.4539	0
I = 10	0.0259	0.0331	0.1087	0.3031	5.8587	0
I = 16	0.0259	0.0320	0.0853	0.2120	3.6253	0
I = 20	0.0259	0.0316	0.0780	0.1832	2.9201	0
I = 40	0.0258	0.0310	0.0638	0.1279	1.5638	0

$M(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{80}
I = 2	0.0350	0.0584	0.5698	2.0300	45.8298	0
I = 4	0.0346	0.0481	0.2405	0.7590	16.2546	0.2375
I = 5	0.0353	0.0502	0.2092	0.6375	14.5157	0.2875
I = 8	0.0408	0.0500	0.1602	0.4215	8.1254	0.2375
I = 10	0.0437	0.0551	0.1497	0.3725	6.1421	0.2875
I = 16	0.0542	0.0679	0.1413	0.3029	4.1466	0.2875
I = 20	0.0521	0.0643	0.1307	0.2435	3.3446	0.2375
I = 40	0.0393	0.0425	0.0828	0.1612	1.9262	0.2375

$Q(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{80}
I = 2	0.0400	0.1453	7.7032	50.5693	3399.9639	0
I = 4	0.0353	0.0559	0.4823	2.3370	102.5121	0.25
I = 5	0.0388	0.0551	0.3837	1.4340	56.0351	0.275
I = 8	0.0388	0.0463	0.1854	0.6271	15.8928	0.275
I = 10	0.0358	0.0472	0.1640	0.5069	12.5054	0.2625
I = 16	0.0365	0.0452	0.1201	0.3000	5.8498	0.2625
I = 20	0.0373	0.0459	0.1046	0.2607	4.5663	0.25
I = 40	0.0350	0.0420	0.0862	0.1722	2.1608	0.2375

Table 4.8: Simulated $MSE(\frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2+\hat{\sigma}_e^2})$ for $N = 80$

$A(\frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2+\hat{\sigma}_e^2})$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{80}
I = 2	0.0020	0.0135	0.0628	0.1028	0.1627	0
I = 4	0.0022	0.0089	0.0341	0.0471	0.0404	0
I = 5	0.0024	0.0083	0.0282	0.0386	0.0276	0
I = 8	0.0034	0.0080	0.0207	0.0246	0.0125	0
I = 10	0.0042	0.0081	0.0180	0.0204	0.0085	0
I = 16	0.0067	0.0098	0.0154	0.0153	0.0049	0
I = 20	0.0088	0.0112	0.0149	0.0139	0.0039	0
I = 40	0.0244	0.0240	0.0201	0.0142	0.0026	0

$M(\frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2+\hat{\sigma}_e^2})$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{80}
I = 2	0.0019	0.0142	0.0622	0.1009	0.1566	0
I = 4	0.0019	0.0090	0.0399	0.0605	0.0627	0.2375
I = 5	0.0019	0.0084	0.0364	0.0531	0.0584	0.2875
I = 8	0.0030	0.0079	0.0264	0.0320	0.0206	0.2375
I = 10	0.0037	0.0095	0.0230	0.0281	0.0131	0.2875
I = 16	0.0102	0.0140	0.0215	0.0207	0.0058	0.2875
I = 20	0.0128	0.0161	0.0215	0.0184	0.0046	0.2375
I = 40	0.0156	0.0182	0.0239	0.0191	0.0036	0.2375

$Q(\frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2+\hat{\sigma}_e^2})$	$\rho = 0.01$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 1$	$\rho = 5$	ϵ_{80}
I = 2	0.0042	0.0192	0.0829	0.1215	0.1408	0
I = 4	0.0033	0.0106	0.0407	0.0553	0.0355	0.25
I = 5	0.0038	0.0104	0.0367	0.0442	0.0256	0.275
I = 8	0.0046	0.0088	0.0260	0.0286	0.0123	0.275
I = 10	0.0043	0.0092	0.0227	0.0262	0.0086	0.2625
I = 16	0.0066	0.0100	0.0198	0.0196	0.0055	0.2625
I = 20	0.0073	0.0121	0.0209	0.0180	0.0049	0.25
I = 40	0.0171	0.0199	0.0273	0.0213	0.0037	0.2375

We call the designs derived from the ANOVA estimation, Rocke's M-estimation, and Uhlig's Q-estimation as the ANOVA design, the M-design and the Q-design, respectively. Now, we can investigate robust designs according to the constrained criterion. For example, let us look at the designs of sample size 40 (Tables 4.1–4.4). If the constraint requires that the breakdown point be greater than or equal to 0.2 ($\gamma_0 = 0.2$), all the designs at $I = 2$ are not applicable because they have zero breakdown points. For $\rho = 0.1$, the M-design and the Q-design that minimize $MSE(\hat{\sigma}_a^2)$ are obtained at $I = 5$ ($\epsilon_{40} = 0.275$), and $I = 8$ ($\epsilon_{40} = 0.25$); the M-design and the Q-design that minimize $MSE(\hat{\sigma}_e^2)$ are obtained at $I = 4$ ($\epsilon_{40} = 0.275$), and $I = 5$ ($\epsilon_{40} = 0.25$); the M-design and the Q-design that minimize $MSE(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$ are obtained at $I = 20$ ($\epsilon_{40} = 0.225$), and $I = 20$ ($\epsilon_{40} = 0.225$); the M-design and the Q-design that minimize $MSE(\hat{\sigma}_a^2/(\hat{\sigma}_a^2 + \hat{\sigma}_e^2))$ are obtained at $I = 4$ ($\epsilon_{40} = 0.225$), and $I = 5$ ($\epsilon_{40} = 0.255$).

As mentioned at the beginning of this chapter, different methods of estimation lead to different designs. To compare the choices of designs from the ANOVA estimation, Rocke's M-estimation, and Uhlig's Q-estimation, we list the optimal designs in Table 4.9 ($N = 40$) and Table 4.10 ($N = 80$). All designs are selected based on the same criterion: the minimization of the mean squared error (MSE). However, for Rocke's M-estimation and Uhlig's Q-estimation, there is a criterion constraint that requires the finite-sample breakdown point of the design to be greater than or equal to 0.2.

Table 4.9: Optimal Designs for N = 40

Parameter	σ_a^2			σ_e^2		
Estimator	Q	M	ANOVA	Q	M	ANOVA
$\rho = 0.01$	I = 5	I = 5	I = 2	I = 4	I = 4	I = 2
$\rho = 0.1$	I = 8	I = 5	I = 5	I = 5	I = 4	I = 2
$\rho = 0.5$	I = 20	I = 20	I = 10	I = 4	I = 4	I = 2
$\rho = 1$	I = 20	I = 20	I = 20	I = 5	I = 4	I = 2
$\rho = 5$	I = 20	I = 20	I = 20	I = 4	I = 4	I = 2

Parameter	$\sigma_a^2 + \sigma_e^2$			$\frac{\sigma_a^2}{\sigma_a^2+\sigma_e^2}$		
Estimator	Q	M	ANOVA	Q	M	ANOVA
$\rho = 0.01$	I = 8	I = 4	I = 20	I = 5	I = 4	I = 2
$\rho = 0.1$	I = 20	I = 20	I = 20	I = 5	I = 4	I = 4
$\rho = 0.5$	I = 20	I = 20	I = 20	I = 10	I = 10	I = 10
$\rho = 1$	I = 20	I = 20	I = 20	I = 10	I = 10	I = 10
$\rho = 5$	I = 20	I = 20	I = 20	I = 20	I = 20	I = 20

Table 4.10: Optimal Designs for N = 80

Parameter	σ_a^2			σ_e^2		
Estimator	Q	M	ANOVA	Q	M	ANOVA
$\rho = 0.01$	I = 4	I = 5	I = 4	I = 4	I = 4	I = 2
$\rho = 0.1$	I = 8	I = 8	I = 10	I = 4	I = 4	I = 2
$\rho = 0.5$	I = 40	I = 20	I = 20	I = 4	I = 5	I = 2
$\rho = 1$	I = 40	I = 40	I = 40	I = 4	I = 4	I = 2
$\rho = 5$	I = 40	I = 40	I = 40	I = 4	I = 4	I = 2

Parameter	$\sigma_a^2 + \sigma_e^2$			$\frac{\sigma_a^2}{\sigma_a^2+\sigma_e^2}$		
Estimator	Q	M	ANOVA	Q	M	ANOVA
$\rho = 0.01$	I = 40	I = 4	I = 40	I = 4	I = 4	I = 2
$\rho = 0.1$	I = 40	I = 40	I = 40	I = 8	I = 8	I = 8
$\rho = 0.5$	I = 40	I = 40	I = 40	I = 16	I = 20	I = 20
$\rho = 1$	I = 40	I = 40	I = 40	I = 20	I = 20	I = 20
$\rho = 5$	I = 40	I = 40	I = 40	I = 40	I = 40	I = 40

The numbers in the tables indicate the following results:

(C1) For σ_a^2, the ANOVA estimation, Rocke's M-estimation, and Uhlig's estimation would give optimal designs with the largest number of groups provided the ratio ρ is greater than or equal to 1.

(C2) For σ_e^2, the ANOVA estimation yields optimal designs with the smallest number of groups, regardless what the value of ρ is, while the other two estimation methods yield optimal designs with the second or third smallest number of groups because of the constraint of the finite-sample breakdown point.

(C3) For $\sigma_a^2 + \sigma_e^2$, the three estimation methods would give optimal designs with the largest number of groups for most values of ρ.

(C4) For $\sigma_a^2/(\sigma_a^2 + \sigma_e^2)$, the number of group of optimal designs would increase as the value of ρ increases.

(C5) Rocke's M-estimation and Uhlig's Q-estimation would yield optimal designs with the same number of groups in most cases, especially when ρ is greater than 0.1.

(C6) Due to a zero breakdown point, both Rocke's M-estimation and Uhlig's Q-estimation cannot yield optimal designs at $I = 2$. If we exclude $I = 2$ from the choice of optimal designs, then the number of groups of optimal designs based on the ANOVA estimation would be very similar to that of optimal robust designs based on Rocke's estimation and Uhlig's estimation. In other words, an alternative to constructing robust designs is to minimize the variance of the ANOVA estimates under the constraint of a non-zero finite-sample breakdown point.

4.3 An Example

Table 4.11 shows the data from a study on determining the amount of arsenic in soil (ISO-5725-2, 1992). In the study, the soil sample was divided, and part of the sample was sent to 23 different laboratories. Each laboratory analyzed the arsenic content twice. The data in Table 4.11 are sorted in ascending order by the first measurement of each laboratory. Note that two data values are less than 1, and three other data values are greater than 25. Approximately 60% of the data values vary within a small range between 1.6 and 2.75, while 28% of the data values vary within a relatively large range between 3.8 and 9. The considerable variation of the data values may be explained by differences in staff members, analytical methods, or analytical equipments.

According to ISO-5725-2 (1992), the standard ANOVA analysis provides the estimates $\hat{\sigma}_e = 0.41$ and $\hat{\sigma}_y = 1.99$ after rejection of the measurements from Laboratories 22 and 23. But the estimate of $\hat{\sigma}_y$ seems to be inconsistent with the fact that 60% data values of the laboratories are within an interval of length 1.15

Table 4.11: Arsenic content (Mg/L) measured by 23 laboratories

Laboratory											
1	2	3	4	5	6	7	8	9	10	11	12
0.19	1.60	1.90	2.00	2.00	2.00	2.00	2.00	2.02	2.10	2.10	2.44
0.21	1.60	2.10	2.50	3.00	1.90	2.00	2.00	2.22	2.00	2.30	2.40
Laboratory											
13	14	15	16	17	18	19	20	21	22	23	
2.50	2.70	2.74	3.80	4.00	4.80	6.20	7.00	8.00	8.98	27.00	
2.40	2.20	2.69	4.00	4.20	5.20	5.80	8.40	7.00	25.83	30.00	

(1.6–2.75) under the assumption of normality. As we know, the estimates from the standard ANOVA are highly dependent on the outlier detection procedure. Uhlig (1997) showed that, if in Laboratory 20 the value 8.40 were replaced by 9.00, the outlier detection procedure of ISO-5725-2 would provide five instead of two outlier laboratories; then the standard ANOVA analysis gives the estimates $\hat{\sigma}_e = 0.23$ and $\hat{\sigma}_y = 1.05$. By applying Rocke's M-estimation and Uhlig's Q-estimation to the data, we obtain the M-estimates: $\hat{\sigma}_a = 1.0056$, $\hat{\sigma}_e = 0.2488$, and $\hat{\sigma}_y = 1.2543$, as well as the Q-estimates: $\hat{\sigma}_a = 1.0643$, $\hat{\sigma}_e = 0.3138$, and $\hat{\sigma}_y = 1.1096$. We summarize the results in Table 4.12. In the table, the ANOVA estimation (2 Laboratories eliminated) yields the largest estimates, whereas the ANOVA (5 Laboratories eliminated) and the other two robust estimation provide very similar results.

Table 4.12: ANOVA estimates, M-estimates, and Q-estimates, $N = 46$

Parameter	$\hat{\sigma}_a$	$\hat{\sigma}_e$	$\hat{\sigma}_y$
ANOVA (2 Labs eliminated)	1.9473	0.4100	1.9900
ANOVA (5 Labs eliminated)	1.0245	0.2300	1.0500
Rocke's M-estimates	1.0056	0.2488	1.2543
Uhlig's Q-estimates	1.0643	0.3138	1.1096

However, it is not reasonable to discard any laboratory from the population, since the experimental data came from different laboratories. Our task is to provide more realistic conclusions with robust methods of estimation that can deal with the occurrence of outliers, or to incorporate the information from outliers with that from non-outliers. To further illustrate the robustness behaviors of Rocke's M-estimators and Uhlig's Q-estimators in different contamination situations, we have carried out a simple simulation study, in which we assume that the error component ϵ_{ij} has a "contaminated" distribution which is a mixture of a normal distribution ($N(0, \sigma_e^2)$) and a double exponential distribution. Each randomly generated sample consists of 95% "good" observations from the normal distribution and 5% "bad" observations from the double exponential distribution. If the mean of the double exponential distribution is zero, we call it "symmetric" contamination; otherwise, if the mean of the double exponential is non-zero (we set the mean as 10 in the simulation), we call it "asymmetric" contamination. We implement the simulation with 1000 runs and report the estimates of σ_a^2 and σ_e^2 and their variances in Table 4.13. In the simulation, we set both σ_a^2 and σ_e^2 equal to 1.

Table 4.13: Simulated Estimates for N = 40

$I = 5$	$\hat{\sigma}_a^2$			$\hat{\sigma}_e^2$		
Estimator	Q	M	ANOVA	Q	M	ANOVA
symmetric	1.2828	0.8756	1.0487	1.0779	0.8803	1.044
asymmetric	1.3033	0.8346	.9375	1.6422	1.0612	6.0625

$I = 5$	$Var(\hat{\sigma}_a^2)$			$Var(\hat{\sigma}_e^2)$		
Estimator	Q	M	ANOVA	Q	M	ANOVA
symmetric	1.5547	0.6098	0.6369	0.0779	0.0673	0.0774
asymmetric	2.1473	0.6093	1.0679	0.1716	0.1012	1.1650

Note that the "symmetric" contamination does not have an significant impact on the inference for σ_a^2 and σ_e^2 because all three methods of estimation yield relatively consistent estimates for each of the variance components. However, the "asymmetric" contamination affects significantly on the ANOVA estimate of σ_e^2, and mildly on the Q-estimate of σ_e^2. Similarly, the variances of the estimates are significantly influenced in the "asymmetric" contamination case, but not in the "symmetric" case. To summarize, the ANOVA estimation is more sensitive to the contaminations than Rocke's M-estimation and Uhlig's Q-estimation. Here we suggest using

Rocke's M-estimation since it would yield more accurate and reliable parameter estimates than Uhlig's Q-estimation in both the "symmetric" and "asymmetric" contamination cases.

Based on the data from the laboratories, we have also studied the problem of design selection through the standard ANOVA estimation, Rocke's M-estimation, and Uhlig's Q-estimation. Suppose we have all the necessary resources to analyze a sample 50 times, i. e. $N = 50$. The problem of design selection becomes a question of how many laboratories are needed to obtain accurate estimates of model parameters under the one-way random effects model. The design criterion is still set as the minimization of MSE of the estimates of model parameters. For Rocke's M-estimation and Uhlig's Q-estimation, the criterion constraint requires that the finite-sample breakdown point of design be greater than or equal to 0.2. The MSEs of the estimates of the model parameters have been computed by simulation under the balanced one-way random effects model (Table 4.14 and Table 4.15). The ratio ρ is assigned the values 16.34 and 11.50, which are calculated from Rocke's M-estimates and Uhlig's Q-estimates in Table 4.12. The selected optimal designs are reported in Table 4.16. Once again, the results from Table 4.16 illustrate that the three methods of estimation would yield very similar optimal designs under the balanced model. To reduce the computation complexity of the simulation, an alternative to construct robust designs is to minimize the variance of the ANOVA estimates under a constraint of non-zero finite-sample breakdown points.

Table 4.14: Simulated MSE for $N = 50$

$A(\hat{\sigma}_a^2)$	$\rho = 16.34$	$\rho = 11.50$	$A(\hat{\sigma}_e^2)$	$\rho = 16.34$	$\rho = 11.50$	ϵ_{50}
I = 2	2.05521	2.58371	I = 2	0.00016	0.00040	0
I = 5	0.51757	0.65267	I = 5	0.00017	0.00043	0
I = 10	0.23285	0.29511	I = 10	0.00019	0.00049	0
I = 25	0.09059	0.11660	I = 25	0.00031	0.00078	0

$M(\hat{\sigma}_a^2)$	$\rho = 16.34$	$\rho = 11.50$	$M(\hat{\sigma}_e^2)$	$\rho = 16.34$	$\rho = 11.50$	ϵ_{50}
I = 2	2.00912	2.76023	I = 2	0.00020	0.00051	0
I = 5	0.49890	0.63639	I = 5	0.00024	0.00061	0.28
I = 10	0.24486	0.30227	I = 10	0.00036	0.00094	0.28
I = 25	0.10129	0.13246	I = 25	0.00034	0.00087	0.24

$Q(\hat{\sigma}_a^2)$	$\rho = 16.34$	$\rho = 11.50$	$Q(\hat{\sigma}_e^2)$	$\rho = 16.34$	$\rho = 11.50$	ϵ_{50}
I = 2	182.42480	228.40280	I = 2	0.00018	0.00046	0
I = 5	3.01190	3.24723	I = 5	0.00020	0.00048	0.26
I = 10	0.60048	0.70106	I = 10	0.00024	0.00061	0.26
I = 25	0.15502	0.18368	I = 25	0.00049	0.00122	0.24

Table 4.15: Simulated MSE for $N = 50$

$A(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$	$\rho = 16.34$	$\rho = 11.50$	$A(\frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2+\hat{\sigma}_e^2})$	$\rho = 16.34$	$\rho = 11.50$	ϵ_{50}
I = 2	2.05535	2.58408	I = 2	0.13431	0.14617	0
I = 5	0.51771	0.65301	I = 5	0.00974	0.01429	0
I = 10	0.23296	0.29541	I = 10	0.00195	0.00336	0
I = 25	0.09059	0.11660	I = 25	0.00069	0.00128	0

$M(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$	$\rho = 16.34$	$\rho = 11.50$	$M(\frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2+\hat{\sigma}_e^2})$	$\rho = 16.34$	$\rho = 11.50$	ϵ_{50}
I = 2	2.00961	2.75865	I = 2	0.12159	0.13715	0
I = 5	0.50172	0.64218	I = 5	0.02660	0.03325	0.28
I = 10	0.24725	0.30817	I = 10	0.00359	0.00497	0.28
I = 25	0.10279	0.13431	I = 25	0.00099	0.00195	0.24

$Q(\hat{\sigma}_a^2 + \hat{\sigma}_e^2)$	$\rho = 16.34$	$\rho = 11.50$	$Q(\frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2+\hat{\sigma}_e^2})$	$\rho = 16.34$	$\rho = 11.50$	ϵ_{50}
I = 2	182.4318	228.4322	I = 2	0.11462	0.12942	0
I = 5	3.0150	3.25037	I = 5	0.00818	0.01267	0.26
I = 10	0.60179	0.70336	I = 10	0.00215	0.00367	0.26
I = 25	0.15596	0.18483	I = 25	0.00101	0.00187	0.24

Table 4.16: Optimal Designs for N = 50

Parameter	σ_a^2			σ_e^2		
Estimator	Q	M	ANOVA	Q	M	ANOVA
$\rho = 16.34$	I = 25	I = 25	I = 25	I = 5	I = 5	I = 2
$\rho = 11.50$	I = 25	I = 25	I = 25	I = 5	I = 5	I = 2

Parameter	$\sigma_a^2 + \sigma_e^2$			$\frac{\sigma_a^2}{\sigma_a^2+\sigma_e^2}$		
Estimator	Q	M	ANOVA	Q	M	ANOVA
$\rho = 16.34$	I = 25	I = 25	I = 25	I = 25	I = 25	I = 25
$\rho = 11.50$	I = 25	I = 25	I = 25	I = 25	I = 25	I = 25

Chapter 5

Conclusion

The experimental designs for the one-way random effects model have been exhaustively studied from the perspective of non-robustness and robustness. The design implementations are based on the standard ANOVA estimation, Rocke's M-estimation, and Uhlig's Q-estimation.

The method of the standard ANOVA estimation is not a robust procedure because one single outlier may significantly influence the estimates of the group mean and the overall mean. Due to its non-robust property, the standard ANOVA estimation performs well under ideal parametric models, but behaves poorly under gross-error models. Usually, the design criterion for the standard ANOVA estimation is set as the minimization of the variances of estimates of the variance components σ_a^2 and σ_e^2, or other related model parameters, such as $(\sigma_a^2 + \sigma_e^2)$, σ_a^2/σ_e^2, and $\sigma_a^2/(\sigma_a^2 + \sigma_e^2)$. The expressions for the variances are derived in Chapter 4, explicitly showing the choice of design groups that satisfy the criterion.

Uhlig's Q-estimators are derived from the Rousseeuw and Croux's Q-estimator. The estimators have simple and explicit formulas and can be applied to asymmetric distributions. Their finite-sample breakdown points are thoroughly examined with detailed computation formulas in Chapter 3. The design criterion for Uhlig's M-estimation is set as the minimization of the mean squared errors (MSE) of estimates of the model parameters subject to the constraint of a non-zero finite-sample breakdown point. The MSE is a reasonable measure since it incorporates both the variance and the bias of a robust estimator given it may be biased. The constraint of finite-sample breakdown point ensure that the selected designs can deal with a specified number of outliers. In practice, it may help experimenters to choose suitable design groups (neither too few or too many). The Monte Carlo simulation has been implemented for computing the MSEs of the estimates since the exact expressions for the MSEs are difficult to obtain analytically. The resulting optimal designs are reported in Chapter 4.

Rocke's M-estimation is a robust procedure estimating the variance components σ_a^2 and σ_e^2 for the one-way random effects model. The main idea is to replace the actual data with pseudo-data, then apply the standard ANOVA to the pseudo-data to estimate the variance components. The pseudo-data are equal to or close to the original data except for outliers which would be replaced by quite different pseudo-data. The aspect of robustness of Rocke's M-estimation is reflected from the finite-sample breakdown point. The finite-sample breakdown point is an important robustness measure, providing the information on the proportion of contaminated observations that a robust estimator can stand without breakdown. The two types of explosion breakdown point (group and measurement) have been calculated in finite samples according to the theorem developed by Zhou and Zhu (2003). Further study is needed for the group and measurement implosion breakdown points, since the implosion situation in which the estimates approach zero turns out to be more complicated than the explosion situation, especially when the intermediate parameters are implicitly estimated through a system of equations. Similarly, the computation of the MSEs of the estimates of the model parameters has been carried out through the Monte Carlo simulation. The optimal designs are reported based on the simulated MSEs and the finite-sample breakdown points of the estimators.

Furthermore, Uhlig's Q-estimators are compared with Rocke's M-estimators in terms of the finite-sample breakdown point and the choice of robust design. The comparisons show that the finite-sample breakdown points are quite close in most of the studied design groups; the choices of optimal design are slightly different at a lower level ratio ($\rho <= 0.1$) and are nearly the same at a higher level ratio ($\rho >= 0.5$).

With robust designs, the choices of methods of estimation are flexible: for high-quality data or"pre-cleaned" data, one may still use the standard ANOVA estimation; for data with no prior knowledge or information about the gross errors, one may use robust estimation. When implementing ANOVA optimal designs in experiments, one should not choose the one with a zero breakdown point in case of the occurrence of outliers, because the implementation of robust estimators cannot save the estimation procedure without tolerance to gross errors. The drawback of robust designs come from the much needed effor to compute the MSEs. An alternative to constructing robust designs is to minimize the variance of the ANOVA estimates under a constraint of non-zero finite-sample breakdown points. From all examples studied, the robust designs derived from this method are very similar to the ones using simulated MSEs.

Extending robust design aspects of the one-way random effects model to the two-way random effects model may be of special interest for further study. The two-way random effects model is commonly given by

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \qquad \begin{cases} i = 1, \cdots, I \\ j = 1, \cdots, J \\ k = 1, \cdots, K\ , \end{cases} \tag{5.1}$$

where μ is an overall mean effect; α_i is the effect of the ith level of the row factor A; β_j is the effect of the jth level of column factor B; $(\alpha\beta)_{ij}$ is the effect of the interaction between α_i and β_j; e_{ijk} is a random error component. We assume that α_i, β_j, $(\alpha\beta)_{ij}$, and e_{ijk} are independent random variables having the normal distribution with zero means and variances σ_a^2, σ_b^2, σ_{ab}^2, and σ_e^2, respectively. Model (5.1) is called a reduced two-way model if it does not include the interaction term.

Most of the published work devoted to design aspects of variance components analysis dates back to the 1960s and 1970s, or even earlier. Yates (1934) introduced two methods of estimation for the two-way model: weighted squares of means and fitting constants. Henderson (1953) developed three methods for model (5.1), namely, Method 1, Method 2, and Method 3. Then Gaylor (1960) considered the problem of optimal designs to estimate σ_a^2 using Henderson's Method 3, and he proposed the L-design and the balanced disjoint design. Bush and Anderson (1963) compared three estimation procedures for estimating model (5.1)'s variance components, namely, Henderson's Method 1 and 3, and Yates' weighted squares of means. The comparison was based on the variances of estimates of the variance components. Mostafa (1967) used the method of un-weighted squares of means (Yates, 1934) to obtain unbiased estimates of the variance components and proposed two optimal designs, D_1 and D_2. Using asymptotic maximum likelihood (ML) results, Muse and Anderson (1978) compared several designs proposed by Gaylor (1960) and Bush (1962) for model (5.1) with no interaction. A comparison of small sample ML results and asymptotic results was also carried out using computer simulation. Effectively, the estimation procedures mentioned above are accomplished by computing least-squares estimates or maximum- likelihood estimates, or equating expected sums of squares to certain computed sums of squares in the standard analysis of variance and then solving a set of equations for the estimates of the variance components. As discussed in previous chapters, these classical methods of estimation are very sensitive to the presence of outliers: even one single outlier may have a large effect on the resulting estimates. In order to overcome this drawback, robust methods of estimation have been proposed as alternatives. Although much work has been done on robust statistical methods since the 1970s, the relevant research on the two-way random effects model is somewhat limited. This is partially because the random effects models play a relatively small part in the area of experimental design. In this thesis, we introduced two types of robust estimators, Rocke's M-estimators (Rocke, 1983, 1991)and Uhlig's Q-estimators (Uhlig, 1997), and various connected robust designs for model (1.1). Rocke (1983) suggested that his M-estimators may also be applied to problems with more than two variance components. For example, if replicate observations are available in experiments, the random interaction effects may also be estimated. Uhlig's Q-estimators are all based on differences between the single measurements. As the ANOVA estimators split the overall variance into the variance within groups and the variance between groups, they partition the set of all differences into differences within groups and differences between groups. Thus these estimators have a simple structure and can be calculated non-iteratively. Moreover, they attain relatively high breakdown points in both finite and asymptotic situations. These desirable properties of robustness may motivate the extension of Uhlig's Q-estimators to the two-way random effects model.

Hopefully, what we have accomplished in this thesis will contribute to further research in the area of robust design of variance components analysis.

Bibliography

[1] Anderson, R. L. and Crump, P. P. (1967). Comparisons of designs and estimation procedures for estimating parameters in a two-stage nested process. *Technometrics*, **9**, 499-516.

[2] Andrews, D. F., Bickel, P. J., Hampel, F. R., Huber, P. J., Rogers, W. H., and Tukey, J. W. (1972). *Robust Estimates of Location: Survey and Advances*, Princeton University Press, Princeton, N.J.

[3] Bush, N. and Anerson, R. L. (1963). A comparison of three different procedures for estimating variance components. *Technometrics*, **5**, 421-440.

[4] Coale, A. J. and Stephan, F. F. (1962). The case of the Indians and the teen-age widows. *Journal of American Statistical Association*, **57**, 338-347.

[5] Cushny, A. R. and Peebles, A. R. (1905). The action of optical isomers. II. Hyoscines. *Journal of Physiology*, **32**, 501-510.

[6] Gaylor, D. W. (1960). The construction and evaluation of some designs for the estimation of parameters in random models. Ph.D. Thesis, North Carolina State University, Raleigh, North Carolina.

[7] Hammersley, J. M. (1949). The unbiased estimate and standard error of the interclass variance. *Metron*, **15**, 189-205.

[8] Hampel, F. R. (1968). Contributions to the theory of robust estimation. Ph.D. Thesis, University of California, Berkeley.

[9] Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986). *Robust Statistics - The Approach Based on Influence Functions*, Wiley.

[10] Henderson, C. R. (1953). Estimation of variance and covariance components. *Biometrics*, **9**, 226-252.

[11] Herrendörfer, G. (1979). Optimum designs for estimating the variance components in a simple variance component model (Model II), I. *Biometrical Journal*, **21**, 11-16.

[12] Huber, P. J. (1964). Robust estimation of a location parameter. *Annals of Mathematical Statistics*, **35**, 73-101.

[13] Huber, P. J. (1981). *Robust Statistics*, Wiley.

[14] ISO-5725-2 (1992). *Accuracy (Trueness and Precision) of Test Measurements.* Geneva: International Organisation for Standardisation.

[15] Khuri, A. I. (2000). Designs for variance components estimation: past and present. *International Statistical Review*, **68**, 311-322.

[16] Milliken, G. A. and D. E. Johnson (1984). *Analysis of Messy Data, Vol. 1.* Van Nostrand Reinhold, New York.

[17] Montgomery, D. C. (1991). *Design and Analysis of Experiments*, Wiley.

[18] Mostafa, M. G. (1967). Designs for the simultaneous estimation of functions of variance components from two-way crossed classifications. *Biometrika*, **54**, 127-131.

[19] Müller, C. H. and Uhlig, S. (2001). Estimation of variance components with high breakdown point and high efficiency. *Biometrika*, **88**, 353-366.

[20] Muse, H. D. and Anderson, R. L. (1978). Comparison of designs to estimate variance components in a two-way classification model. *Technometrics*, **20**, 159-166.

[21] Rocke, D. M. (1983). Robust statistical analysis of interlaboratory studies. *Biometrika*, **70**, 421-431.

[22] Rocke, D. M. (1991). Robustness and balance in the mixed model. *Biometrics*, **47**, 303-309.

[23] Rousseeuw, P. J. and Croux, C. (1992). Explicit scale estimators with high breakdown point. In L_1*-Statistical Analysis and Related Methods*, Ed.Y. Dodge, pp. 77-92. Amsterdam: North-Holland.

[24] Rousseeuw, P. J. and Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of American Statistical Association*, **88**, 1273-83.

[25] Thompson, W. O. and Anderson, R. L. (1975). A comparison of designs and estimators for the two-stage nested random model. *Technometrics*, **17**, 37-44.

[26] Tukey, J. W. (1960). A survey of sampling from contaminated distributions. In *Contributions to Probability and Statistics*, I. Olkin (ed.). Stanford University Press, Stanford, Calif., pp. 448-485.

[27] Uhlig, S. (1997). Robust estimation of variance components with high breakdown point in the 1-way random effect model. In *Industrial Statistics. Aims and Computational Aspects*, Ed. C. P. Kitsos and L. Edler, pp. 65-73. Heidelberg: Physica-Verlag.

[28] Yates, F. (1934). The analysis of multiple classifications with unequal numbers in the different classes. *Journal of the American Statistical Association*, **29**, 51-66.

[29] Zhou, J. and Zhu H. (2003). Robust estimation and design procedures for the random effects model. *Canadian Journal of Statistics*, **31**, 99-110.