AN EXAMINATION OF OUTLIERS AND INTERACTION IN A NONREPLICATED TWO-WAY TABLE

by

Barbara R. Kuzmak

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Statistics

APPROVED:

Eric P. Smith, Chair

Robert V. Foutz

Raymond H. Myers

Marion R. Reynolds, Jr.

George R. Terrell

December, 1990

Blacksburg, Virginia

AN EXAMINATION OF OUTLIERS AND INTERACTION IN A NONREPLICATED TWO-WAY TABLE

by

Barbara R. Kuzmak

Eric P. Smith, Chair

Statistics

(ABSTRACT)

The additive-plus-multiplicative model, $Y_{ij} = \mu + \alpha_i + \beta_j + \sum_{p=1}^{k} \lambda_p \tau_{pi} \gamma_{pj}$, has been used to describe multiplicative interaction in an unreplicated experiment. Outlier effects often appear as interaction in a two-way analysis of variance with one observation per cell. I use this model in the same setting to study outliers. In data sets with significant interaction, one may be interested in determining whether the cause of the interaction is due to a true interaction, outliers or both. I develop a new technique which can show how outliers can be distinguished from interaction when there are simple outliers in a two-way table. Several examples illustrating the use of this model to describe outliers and interaction are presented.

I briefly address the topics of leverage and influence. Leverage measures the impact a change in an observation has on fitted values, whereas influence evaluates the effect deleting an observation has on model estimates. I extend the leverage tables for an additive-plus-multiplicative model of rank 1 to a rank k model. Several examples studying the influence in a two-way nonreplicated table are given.

Dedication

This dissertation is dedicated to the two people most responsible for my education, my parents Peter and Dorothy.

Acknowledgements

I am deeply grateful to Eric Smith for his constant guidance throughout this dissertation. I felt privileged to experience his enthusiasm for research. Our frequent conversations made this dissertation exciting, fun and enjoyable. His positive attitude was an invaluable source of encouragement and he always made me feel that I was doing a good job.

I appreciate the effort of my committee members: Dr. Myers, Dr. Terrell, Dr. Foutz and Dr. Reynolds. Their advice and suggestions proved to be priceless during my graduate program and in the development of this dissertation.

I am indebted to the Department of Statistics for being instrumental in the development of my career. I had the opportunity to take many stimulating courses in this department. I appreciate the dedication and hard work of those faculty members who taught them. I value the practical experience I gained through the Consulting Laboratory. And I treasure my experience teaching in the classroom. My students helped me to become a better teacher and love the subject that I taught.

Last, I want to thank the many friends that I have known during my four years in Blacksburg. My colleagues, Donald Mercante, Philip Ramsey, Kathleen O'Brien and Robert Davis, made graduate school a little easier and a lot more fun. The secretaries, Linda Seawell, Dinah Easterly, Leslie Miller, Betty Higginbotham and Cindy Watson, who always found time for a conservation. I have enjoyed the friendship of many people in the Newman Community Over-21 group especially Jennifer Nowicki, Regina Smick-Attisano and James Biter. And finally, I am indebted to Lawrence Vanlieshout whose support made completion of this dissertation possible especially during the loss of Grandma Peckay.

Table of Contents

List of Illustrations

List of Tables

CHAPTER 1 - INTRODUCTION

Nonreplicated experiments may be a statisticians nightmare to analyze. These experiments occur for several reasons. Sometimes a design is not replicated by choice because the cost of obtaining the experimental units or repeating the experimental set up is expensive. Researchers in the area of animal science or veterinary medicine usually pay thousands of dollars for a few large-bodied animals (i.e., cows, horses). If several factors are of interest to the scientist, then true replication is usually prohibited. In other situations, experimental subjects may be readily available, but the cost of an experimental compound may be astronomical to initially produce. Therefore, limited quantities of the material exist for experimentation. The scientist may want to evaluate the performance of the test compound in conjunction with several variables, thus independent repetition of a treatment combination may be sacrificed.

Other times an experiment is not replicated because the investigator does not understand the concept of replication. The scientist may believe that he or she replicated an experiment by recording the same observation over time (repeated measures) or by measuring the same experimental unit several times during a single session (subsample measurement). Unfortunately whether by choice or confusion, the results are the same. No independent observations exist for estimation of the pure error term in the analysis. Therefore, it may not be possible to obtain as complete of an analysis for the nonreplicated experiment as for the replicated design.

Typically, statisticians are not trained to thoroughly analyze unreplicated experiments. Most statisticians are familiar with the analysis of a two-way analysis of variance, and using Tukey's one-degree-of-freedom test to test for nonadditivity of the form $\alpha_i\beta_j$ (where α_i and β_j represent row and column effects, respectively). However, methods do exist to extract more information out of the two-way interaction term.

Mandel (1969, 1971) generalized the decomposition of the two-way interaction term to test for multiplicative interaction of the form $\tau_i\gamma_j$ (where $\tau_i\gamma_j$ is no longer a scalar multiple of main effects). The terms, τ_i and γ_j, are eigenvectors of ZZ′ and Z′Z, respectively, and Z is the residual matrix from the additive portion of the model below. This decomposition recovers more degrees of freedom from error than Tukey's test. Mandel proposed the additive-plus-multiplicative model,

$$Y_{ij} = \mu + \alpha_i + \beta_j + \sum_{p=1}^{k} \lambda_p \tau_{pi} \gamma_{pj}$$

where:

$$k \leq \min\{(r-1, c-1)\}.$$

This model has been employed to analyze an unreplicated two-way table by many researchers (Snee *et al.* 1979, Gauch 1988, Milliken and Johnson 1989). The model possesses a lot of versatility, and it can fit several different forms of multiplicative interaction. However, the use of this model to describe outliers has not been reported in the literature.

The detection of outliers in an unreplicated experiment has been investigated by several researchers. Daniel (1960) studied the identification of a single outlier in factorial designs, and later Stefansky (1972) did the same for a two-way table. Both developed test statistics, however Stefansky showed that they were equivalent. Barnett and Lewis (1984) posed a "contamination" model to describe an outlier in a two-way layout as,

$$Y_{ij} = \mu + \alpha_i + \beta_j + \delta\Delta_{ij} + \varepsilon_{ij}$$

where:

$\Delta_{ij} = 1$ for $i = i', j = j'$ and

$\Delta_{ij} = 0$ otherwise.

This model is similar to Mandel's, and suggests that outliers in a two-way table may be modeled by multiplicative terms in the additive-plus-multiplicative model.

Both outliers and interaction are sources of nonadditivity in a two-way layout. When replication exists, these effects influence separate terms in a model. An interaction term accounts for the interaction, whereas the error term absorbs the outlier effect. However, in the absence of replication, both outliers and interaction enter into the single residual term, η_{ij}, from fitting the additive effects.

For a nonreplicated two-way table, one can test for either outliers or interaction, but not both. If both do exist in a two-way layout with one observation per cell, then testing for either effect may lead to erroneous results. Currently, there does not exist a procedure for detecting both effects at the same time.

I devote a significant portion of this dissertation to studying two-way tables when outliers and interaction are present in an unreplicated design. I consider several facets of this situation. The size of the outlier effect along with the amount of interaction play an important role in the ability to identify both factors. The location of the outlier relative to the interaction in a two-way table can make detection of the outlier range from easy to difficult. Table size also affects the limits of detecting both effects. Lastly, the number of outliers present in the table can complicate the analysis.

I develop a procedure which can locate one or two outliers in the presence of significant interaction. The technique uses Mandel's additive-plus-multiplicative model to detect significant nonadditive effects. If strong nonadditivity is found, then I employ a rotation procedure to uncover the sources of nonadditivity. Finally, I evaluate the performance of this procedure under the different aspects mentioned in the last paragraph.

I complete the study of nonadditivity in a nonreplicated two-way table by briefly considering leverage and influence measures for the additive-plus-multiplicative model. The concepts of leverage and influence are closely related. Leverage determines the impact a change in an observation has on fitted values, whereas influence considers the effect deleting an observation has on model estimates. The current leverage tables are appropriate for the additive-plus-multiplicative model of rank 1 (Emerson *et al.*, 1984). I extend these to a rank k model. I investigate the influence an observation has on the model by applying the techniques of influence analysis used in a multivariate setting to a two-way table.

This dissertation is organized in the following manner. First, a review of the pertinent literature is presented in Chapter 2. I also include an example illustrating how Mandel's model fits a data set in this chapter. In Chapter 3, I extensively study outliers and interaction in an unreplicated two-way table. Included in this chapter is my procedure and an evaluation of its power. The formula for the extended leverage tables is presented in Chapter 4, whereas the influence study is discussed in Chapter 5. Finally in Chapter 6, I make some concluding remarks and speculate on areas of further research.

CHAPTER 2 - BACKGROUND

2.1 THE MODEL

This research uses the additive-plus-multiplicative or Mandel's model to fit data in an unreplicated two-way design. Model details will now be presented along with research related to several aspects of the model.

Mandel (1961, 1969, 1971) postulated various models for a r x c table when the assumption of additivity was questionable. He considered the general nonadditive model,

$$Y_{ij} = \mu + \alpha_i + \beta_j + \eta_{ij} \qquad [2.1.1]$$

where:

i = 1 to r, j = 1 to c

and the usual constraints hold, i.e.,

$$\sum_i \alpha_i = \sum_j \beta_j = \sum_i \eta_{ij} = \sum_j \eta_{ij} = 0.$$

He partitioned the interaction term, η_{ij}, into a sum of multiplicative functions of the form,

$$\eta_{ij} = \sum_{p=1}^{k} \lambda_p \tau_{pi} \gamma_{pj}$$

where:

$$k \leq \min\{(r-1, c-1)\}.$$

He estimated the parameters of model [2.1.1] in two steps. First, an additive two-way model was fit to the data. Second, the residuals from the first step were decomposed using principal components on the matrix Z′Z (or ZZ′) where:

$$\begin{aligned} Z &= [z_{ij}] \\ z_{ij} &= Y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j \\ &= Y_{ij} - Y_{i.} - Y_{.j} + Y_{..}. \end{aligned}$$

This produced the estimates $\hat{\lambda}$, $\hat{\tau}$, $\hat{\gamma}$ where:

$\hat{\lambda}_p^2$ is the largest pth eigenvalue of Z′Z (or ZZ′)

$\hat{\tau}_p$ is the eigenvector corresponding to the pth eigenvalue of ZZ′

$\hat{\gamma}_p$ is the eigenvector corresponding to the pth eigenvalue of Z′Z.

These estimates have the usual constraints,

$$\sum_{i=1}^{r} \hat{\tau}_{pi} = \sum_{j=1}^{c} \hat{\gamma}_{pj} = 0 \quad \forall p$$

$$\sum_{i=1}^{r} \hat{\tau}_{pi}{}^2 = \sum_{j=1}^{c} \hat{\gamma}_{pj}{}^2 = 1 \quad \forall p$$

$$\sum_{i=1}^{r} \hat{\tau}_{pi}\hat{\tau}_{p'i} = \sum_{j=1}^{c} \hat{\gamma}_{pj}\hat{\gamma}_{p'j} = 0 \quad p \neq p'$$

and

$$\hat{\lambda}_1 \geq \ldots \geq \hat{\lambda}_k > 0.$$

Therefore, model [2.1.1] can be written as

$$Y_{ij} = \mu + \alpha_i + \beta_j + \sum_{p=1}^{k} \lambda_p \tau_{pi} \gamma_{pj}. \qquad [2.1.2]$$

This model is frequently called an additive-plus-multiplicative or Mandel's model. Gabriel (1978) proved that this two-step process for fitting an additive-plus-multiplicative model produces least-squares estimates of all parameters in the model. Mandel (1971) constructed an analysis of variance table for this model, Table 1. It resembles the usual ANOVA table for both main effects, but it is novel in its decomposition of the interaction term.

An example to demonstrate how Mandel's model is used in practice to fit both additive and multiplicative effects is given below. The following data came from a nitrogen and phosphorus fertilization study on spring wheat (Black, 1970). The response, yield, is measured in kg/ha. The data are:

		Phosphorus Applied (kg/ha)				
		0	22	45	90	180
Nitrogen Applied (kg/ha)	0	1984	2550	2706	2740	2954
	45	1776	2843	3306	3305	3386
	90	1797	2761	3240	3227	3332

To fit the additive-plus-multiplicative model, I first fit the main effects. The ANOVA table from this step is displayed below:

Table 1. Analysis of variance for Mandel's model

Source	df	SS	MS
α_i	(r-1)	usual	usual
β_j	(c-1)	usual	usual
η_{ij}	(r-1)(c-1)	usual	usual
$\lambda_1\tau_{1i}\gamma_{1j}$	ν_1	$\hat{\lambda}_1^2$	$\hat{\lambda}_1^2/\nu_1$
$\lambda_2\tau_{2i}\gamma_{2j}$	ν_2	$\hat{\lambda}_2^2$	$\hat{\lambda}_2^2/\nu_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\lambda_p\tau_{pi}\gamma_{pj}$	ν_p	$\hat{\lambda}_p^2$	$\hat{\lambda}_p^2/\nu_p$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Note:

$\nu_p = \dfrac{E[\hat{\lambda}_p^2]}{\sigma^2}$ and is calculated by Monte Carlo methods assuming an underlying additive model with independent normally distributed errors.

Source	df	SS	MS	F	Pr
N	2	328075.60	164037.80	5.07	.0378
P	4	3748569.07	937142.27	28.99	.0001
residual	8	258651.73	32331.47		
total	14	4335296.40			

From this table, one can see that both main effects are significant. The estimates of main effects are:

$$\hat{\mu} = 2793.80 \qquad \hat{\alpha} = \begin{bmatrix} -207.00 \\ 129.40 \\ 77.60 \end{bmatrix} \qquad \hat{\beta} = \begin{bmatrix} -941.50 \\ -75.80 \\ 290.20 \\ 296.90 \\ 430.20 \end{bmatrix}.$$

The residuals from step 1 are:

338.70	39.00	−171.00	−143.70	−63.00
−205.70	−4.40	92.60	84.93	32.60
−132.90	−34.60	78.40	58.73	30.40

A decomposition of the residuals produces:

$$\hat{\lambda}_1 = 507.6624 \qquad U_1 = .9964 \;{**}$$
$$\hat{\lambda}_2 = 30.5052$$

$$\hat{\tau} = \begin{bmatrix} -.8123 & -.0825 \\ .4776 & -.6622 \\ .3347 & .7447 \end{bmatrix} \quad \text{and} \quad \hat{\gamma} = \begin{bmatrix} -.8231 & .3050 \\ -.0894 & -.8547 \\ .4124 & .3662 \\ .3485 & -.0214 \\ .1515 & .2048 \end{bmatrix}.$$

I now rewrite this model in matrix form:

$$\begin{array}{cccccccccccc} Y & = & \hat{\mu} & + & \hat{\alpha} & + & \hat{\beta} & + & \hat{\lambda}_1\hat{\tau}_1\hat{\gamma}_1 & + & \hat{\lambda}_2\hat{\tau}_2\hat{\gamma}_2 \end{array}$$

$$\begin{bmatrix} 1984 \\ 2550 \\ 2706 \\ 2740 \\ 2954 \\ 1776 \\ 2843 \\ 3306 \\ 3305 \\ 3386 \\ 1797 \\ 2761 \\ 3240 \\ 3227 \\ 3332 \end{bmatrix} = \begin{bmatrix} 2793.80 \\ 2793.80 \\ 2793.80 \\ 2793.80 \\ 2793.80 \\ 2793.80 \\ 2793.80 \\ 2793.80 \\ 2793.80 \\ 2793.80 \\ 2793.80 \\ 2793.80 \\ 2793.80 \\ 2793.80 \\ 2793.80 \end{bmatrix} + \begin{bmatrix} -207.00 \\ -207.00 \\ -207.00 \\ -207.00 \\ -207.00 \\ 129.40 \\ 129.40 \\ 129.40 \\ 129.40 \\ 129.40 \\ 77.60 \\ 77.60 \\ 77.60 \\ 77.60 \\ 77.60 \end{bmatrix} + \begin{bmatrix} -941.50 \\ -75.80 \\ 290.20 \\ 296.90 \\ 430.20 \\ -941.50 \\ -75.80 \\ 290.20 \\ 296.90 \\ 430.20 \\ -941.50 \\ -75.80 \\ 290.20 \\ 296.90 \\ 430.20 \end{bmatrix} + \begin{bmatrix} 339.43 \\ 36.87 \\ -170.06 \\ -143.71 \\ -62.47 \\ -199.57 \\ -21.68 \\ 99.99 \\ 84.50 \\ 36.73 \\ -139.86 \\ -15.19 \\ 70.07 \\ 59.22 \\ -25.74 \end{bmatrix} + \begin{bmatrix} -.77 \\ 2.15 \\ -.92 \\ .05 \\ -.52 \\ -6.16 \\ 17.27 \\ -7.40 \\ .43 \\ -4.14 \\ 6.93 \\ -19.42 \\ 8.32 \\ -.49 \\ 4.65 \end{bmatrix}.$$

Table 2 displays the complete ANOVA table for the spring wheat data.

Table 2. Analysis of variance table for spring wheat data.

Source	df	SS	MS	F	Pr
N	2	328075.60	164037.80	5.07	.0378
P	4	3748569.07	973142.27	28.99	.0001
N*P	8	258651.73	32331.47		
$\lambda_1 \tau_1 \gamma_1$	6.34	257721.16	40650.03		
$\lambda_2 \tau_2 \gamma_2$	1.66	930.57	560.58		

Johnson and Graybill (1972) tested the significance of λ_1 with the likelihood ratio test and found $U_1 = .9964$ significant at $\alpha = .01$. So they examined only the the first component. They tested all hypotheses of the form H_o: $\tau_i = \tau_{i'}$ and $\gamma_j = \gamma_{j'}$ (where: $i \neq i'$ and $j \neq j'$). Out of the 30 pairs of hypotheses tested only H_o: $\tau_2 = \tau_3$ and $\gamma_3 = \gamma_4$ failed to be rejected. Therefore, they concluded that significant interaction was attributed to cells (1,1), (1,2) and (1,5). Black (1970) concluded that "yield increased dramatically as the level of phosphorus increased up to 45 kg/ha, then yield further increased in the presence of phosphorus with applications of nitrogen at rates of 45 or 90 kg/ha." The control, cell (1,1), may be responsible for some of the interaction because it produced a higher yield than the two treatment combinations with no phosphorus and 45 or 90 kg/ha of nitrogen, cells (2,1) and (3,1), respectively.

Figure 1 is a graph of the N-P fertilization data. It reveals that an interaction occurs between nitrogen rates 0 verses 45 and 90 kg/ha and phosphorus rates 0 and 22 kg/ha. This could be due to a nutrient imbalance (high nitrogen and no phosphorus) which could reduce plant yield. In the absence of any nitrogen application, it appears that some phosphorous, 22 kg/ha, may stimulate plant yield, cell (1,2), however increasing the phosphorus rate to 180 kg/ha did not account for a sizeable increase, cell (1,5). So cell (1,5) may contribute to interaction by not increasing as much as expected for the phosphorus rate applied.

Mandel examined some special cases of the general nonadditive model of rank 1, equation [2.1.2]. He standardized the estimates, $\hat{\alpha}_i$ and $\hat{\beta}_j$, to produce the following model,

$$Y_{ij} = \mu + Rr_i + Gc_j + \lambda\tau_i\gamma_j + \varepsilon_{ij}$$

where:

$$R = \sqrt{\sum \hat{\alpha}_i^2} \qquad r_i = \frac{1}{R}\hat{\alpha}_i$$

and

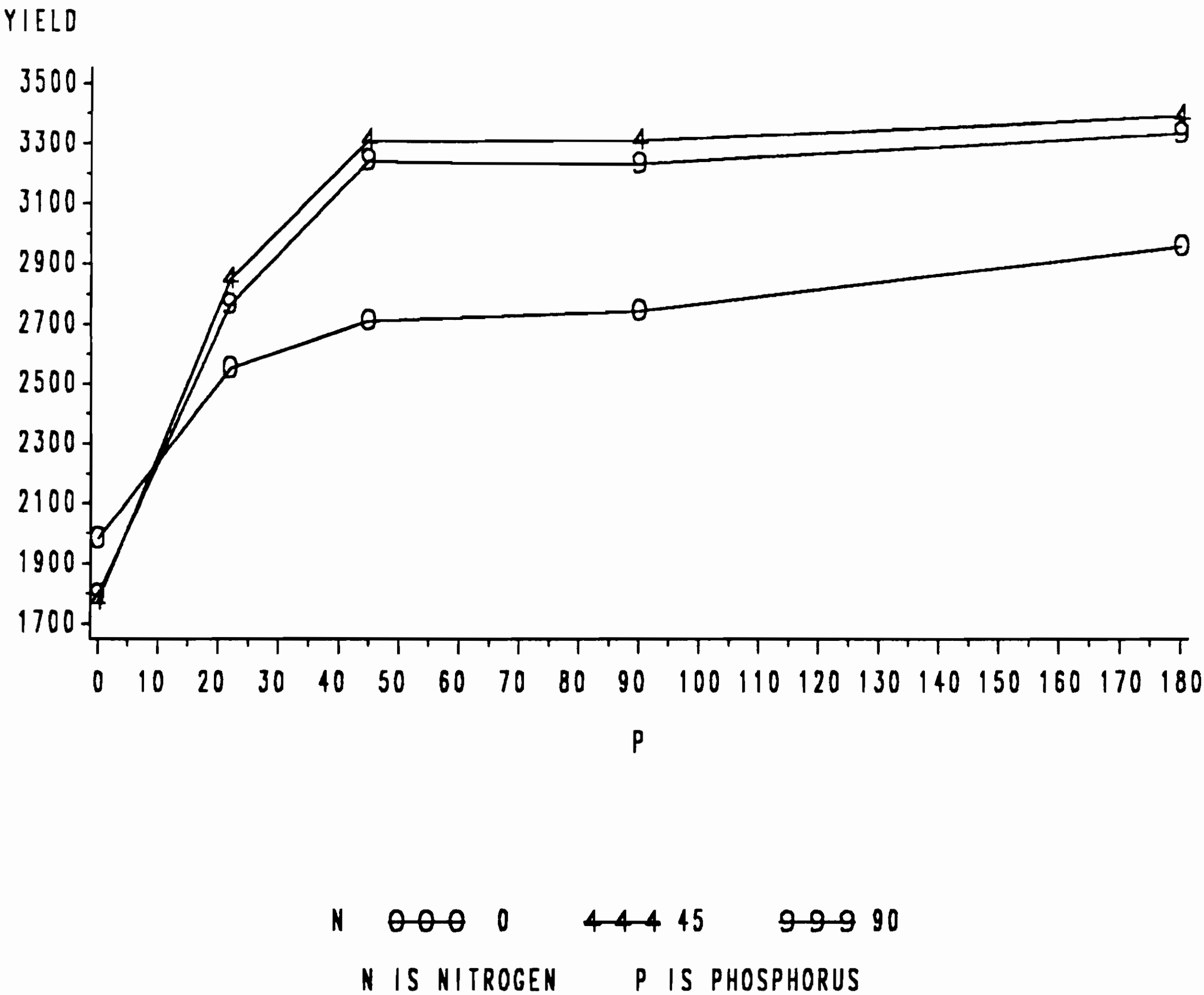

Figure 1. Graph of N-P fertilization on yield of spring wheat.

$$G = \sqrt{\sum \hat{\beta}_j^2} \qquad c_j = \frac{1}{G}\hat{\beta}_j.$$

He considered a model in which the nonadditive term was a linear function of either main effect, i.e.,

$$\tau_i = r_i \qquad \text{"rows linear"}$$

or

$$\gamma_j = c_j \qquad \text{"columns linear"}$$

which generates either,

$$Y_{ij} = \mu + Gc_j + (R + \lambda\gamma_j)r_i + \varepsilon_{ij} \qquad \text{"Rows Linear Model"}$$

or

$$Y_{ij} = \mu + Rr_i + (G + \lambda\tau_i)c_j + \varepsilon_{ij} \qquad \text{"Columns Linear Model."}$$

When the nonadditive term is a linear function of both main effects, i.e., $\tau_i = r_i$ and $\gamma_j = c_j$, then the result is called the "Concurrent Model",

$$Y_{ij} = \mu + Rr_i + Gc_j + \lambda r_i c_j + \varepsilon_{ij}.$$

This is also Tukey's-one-degree-of-freedom test for nonadditivity.

Gollob (1968) developed that same basic model independently of Mandel but used factor analysis instead of principal component analysis to decompose the interaction term. He combined factor analysis with analysis of variance technique, and called this model FANOVA. His method was similar to Mandel's except for the way he partitioned the degrees of freedom for each component. Researchers have accepted Mandel's approach to the degrees of freedom problem, and thus adopted the use of his model rather that Gollob's.

Other researchers have investigated various aspects of Mandel's model. Johnson and Graybill (1972) developed a likelihood ratio test statistic for the inclusion of the first multiplicative component, $\lambda_1\tau_1\gamma_1$, in the two-way additive model. They estimated λ_1 by $\hat{\lambda}_1$, the square root of the leading eigenvalue of Z′Z (or ZZ′), and tested H_o: $\lambda_1 = 0$ vs. H_1: $\lambda_1 \neq 0$ using the following test statistic, U_1, where:

$$U_1 = \frac{\hat{\lambda}_1^2}{\hat{\lambda}_1^2 + \hat{\lambda}_2^2 + \cdots + \hat{\lambda}_k^2} = \frac{\hat{\lambda}_1^2}{\mathrm{tr}(Z'Z)} .$$

They approximated the critical points for U_1 using a Beta approximation. A table of these critical values appears in Johnson and Graybill (1972) and Milliken and Johnson (1989). Schuurmann *et al.* (1973) constructed tables of the exact critical values when (t-b-1)/2 is an integer.

Hegemann and Johnson (1976) developed a likelihood ratio test statistic for incorporating the second multiplicative component into an additive-plus-multiplicative model of rank 1. This statistic is an extension of Johnson and Graybill's test for λ_1. The test statistic, L, for the test of H_o: $\lambda_2 = 0$ vs. H_1: $\lambda_2 \neq 0$, is

$$L = \frac{\hat{\lambda}_2^2}{\hat{\lambda}_2^2 + \cdots + \hat{\lambda}_k^2} .$$

Critical values are also tabulated for this test statistic.

Once more, Johnson and Graybill's (1972) test was extended to sequentially test for the inclusion of the pth multiplicative component. Yochmowitz and Cornell (1978) developed the Λ_p^* test statistic for testing H_o: $\lambda_p = 0$ vs. H_1: $\lambda_p \neq 0$. Its form is:

$$\Lambda_p^* = \frac{\hat{\lambda}_p^2}{\hat{\lambda}_p^2 + \cdots + \hat{\lambda}_k^2} .$$

They called this likelihood ratio test the SKC test, since it uses tables in Schuurmann *et al.* (1973) as critical values.

Once a multiplicative component is found to differ significantly from zero, then the cause of the interaction may be investigated. Marasinghe and Johnson (1981) discuss different scenarios which can produce interaction: 1.) significant interaction can be due to a certain row and column interaction, 2.) it can be caused by a particular row reacting with the columns differently than the other rows, or 3.) it can be the result of certain treatment combinations acting synergistically. Outliers would produce the same effect as case three. After the interaction is located, then interaction free subtables can be used to obtain a more reliable estimate of $\hat{\sigma}^2$ and for the testing of main effects.

Like Marasinghe and Johnson, Daniel (1978) recognized that nonadditivity may reside in only a few cells in an unreplicated r x c table. He advocated localizing these disturbances instead of using the entire (r-1)(c-1) degrees of freedom to report a significant interaction. He proposed a technique to estimate the nonaddivity confined to a few cells. However this procedure is based on recognizing patterns in the disturbances in the expected residuals. I defer discussing Daniel's methodology to the next section on outliers, so that I can present some earlier work which deals with the patterns of the expected residuals in the presence of outliers.

2.2 *OUTLIERS*

Hawkins (1984) defines an outlier to be "an observation which deviates so much from other observations as to arouse suspicions that it was generated by a different mechanism." I like to think that an outlier is more extreme than any error attributed to randomness, and therefore appears inconsistent with the rest of the data.

There are several general problems that arise in the detection of outliers. These are discusssed first, then specific issues relevant to a multiway table are addressed. In diagnosing a data set for outliers, an analyst tries to determine the number of outliers present, their location(s) and possibly the magnitude of their perturbation(s). If multiple outliers are suspected, then determining the number of outliers can be difficult. An extreme observation may 'mask' the detection of a more extreme observation as an outlier. The effect of masking is to underestimate the true number of outliers. The opposite phenomenon can also occur. The number of outliers may be overestimated. This is called 'swamping' (Barnett and Lewis 1984, Hawkins 1984).

In a multiway table, the effect of an outlier is spread throughout the table because of the usual constraints on the parameters in the model. These constraints show up in the table of residuals, i.e., the sum of a row or column of residuals is zero and produce correlations among the residuals. This is critical when the number of treatments is small (Gnanadesikan and Kettering, 1972). Outliers may bias the estimate of the grand mean, $\hat{\mu}$, and the main effects, $\hat{\alpha}_i$ and $\hat{\beta}_j$, and may mask its effect on the residuals (Bross 1961, Gnanadesikan and Kettering 1972). Outliers may cause the main effects and interaction terms to be highly significant (Brown, 1975). Bross (1961) called this the 'star burst' effect because the ANOVA table displays an abundance of stars (* and **) to indicate statistical significance.

Gentleman and Wilk (1975a, 1975b) wrote two papers in sequel which dealt with detecting outliers in a two-way additive model. The first paper is critical to the development of this dissertation topic. They determined the expected values of the residuals for one and two outliers. When an outlier is present in a two-way layout, all residuals are affected by its presence to varying degrees. The extent to which the residuals are disturbed depends upon their location with respect to the outlier. The residual in the cell in which the outlier occurs is altered the most, whereas residuals residing in different rows and different columns from the outlier are affected the least. Gentleman and Wilk displayed the expected values of the residuals in which the (1,1) observation has been perturbed by θ in the following table:

$\theta(r-1)(c-1)$	$-\theta(r-1)$	$\dots$	$-\theta(r-1)$
$-\theta(c-1)$	θ	$\dots$	θ
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$-\theta(c-1)$	θ	$\dots$	θ

All cells are divided by rc.

The expected value table for the residuals when two or more outliers are present is simply the sum of the appropriate expected residual tables containing one outlier. Gentleman and Wilk (1975b) developed a technique to determine the number of outliers, K, in a data set. For a hypothetical K, they partitioned a data set $\binom{n}{n-k}$ times. For each partition, they omitted K observations, fitted an additive model to the remaining n-K observations, and calculated the difference between the sum of squares of residuals for the reduced data, n-K, and the full data set, n. They termed the maximum deviation, Q_k^*, "the K most likely outlier subset."

Gentleman and Wilk point out the drawback of working with residuals to detect multiple outliers. Depending upon the configuration of these outliers, they can be arranged in a manner to go undetected in a table of residuals. For example, two observations in the same row perturbed in the same direction will increase the row effect, thereby decreasing the residuals in these two cells. This problem is inherit in any diagnostic procedure or outlier test based upon residuals.

Several statisticians developed procedures to detect outliers in an unreplicated two-way table. Daniel (1960) developed the first outlier test for an unreplicated experiment, a factorial design. He proposed the test statistic, $t_1 = \frac{d_{\max}}{s_1}$, where $d_{\max}$ is the maximum residual. The standard error, s_1, was formulated incorrectly when this paper was published (Daniel, 1960), and I have not be able to find a follow-up paper correcting this term. Since I am not sure how accurate these critical points are for this test statistic, I will not discuss this test any further.

Stefansky (1971, 1972) created a test for a single outlier in a nonreplicated two-way table. The test is based on the maximum normed residual (MNR). The MNR or $|Z|^{(1)}$ is defined to be the largest absolute normed residual,

$$|Z|^{(1)} = \max \{|Z_1|, ..., |Z_n|\}$$

where:

$$\begin{aligned} Z_i &= \text{ith normed residual} \\ &= \frac{e_i}{\sqrt{\sum_{i=1}^{n} e_i^2}}. \end{aligned}$$

The exact critical values for this test are calculated from Bonferroni's upper (or lower) bound under certain conditions. They are tabulated for r, c = 3, . . . , 9 at a 1% and 5% significance level (Stefansky, 1972). Galpin and Hawkins (1981) applied Stefansky's test to a three-way table r, c, t = 3, . . . , 10 at $\alpha = .01, .025, .05, .10$. They also extended Stefansky's tables to include larger tables and 2.5% and 10% significance levels.

Goldsmith and Boddy (1973) criticized Stefansky's test since it relys on fitting a model which included a suscepted outlier. For a given significance level, this would decrease the test's ability to detect a real outlier. Goldsmith and Boddy's procedure is to consider each value as missing, estimate it from the remaining data and calculate the associated mean square residual. The smallest mean square residual is tested to determine if it is significantly different from the overall mean square residual. They proposed the following test statistic:

$$\frac{RSS_c - RSS_m}{RSS_m/(r-1)}$$

where:

RSS_c is the residual sum of squares from fitting all the data to an additive model

r is the degrees of freedom associated with RSS_c

RSS_m is the smallest residual sum of squares.

The distribution of this test statistic under the null hypothesis of no outlier has been approximated using simulation. It was observed that the shape of the simulated test statistic conformed to the shape of an upper F-distribution with 1 and (r-1) degrees of freedom. The experimentwise error rate, P, of the test statistic was approximately 1/4P of the F (1, r-1) distribution. Therefore, they adjusted the critical value to .8P/(r-1) of F (1, r-1) distribution.

Several other researchers devised outlier tests and detection methods for multiple outliers in an unreplicated two-way layout. Brown (1975) developed an approach similar to Goldsmith and Boddy for the identification of possible outliers. He believed that a few cells could be the major source of a significant interaction (this resembles Bross's 'star burst' effect). Like Goldsmith and Boddy, he treated the observation from the most extreme residual as missing, and estimated its expected value under an additive model. Unlike the previous method, the ANOVA was recomputed and the interaction terms were reexamined. If the interactions were still significant, then the process was repeated. This time, however, both cells were simultaneously estimated so that they both had zero residuals. This stepwise routine continued until the interaction terms were no longer significant.

As stated earlier, Daniel (1978) also noticed that nonadditivity in a two-way layout may be caused by a few observations. He observed recognizable patterns in the expected residuals when the disturbances in additivity were confined to less that $G < R/2$ rows and $H < C/2$ columns. He estimated these perturbations by reordering the rows and columns which contain the largest residuals into a G x H subtable in the upper left corner, then calculating the disturbance and its variance for each cell in the subtable by:

$$\overline{D}_{ij} = r_{ij} - \bar{r}_{ij'} - \bar{r}_{i'j} + \bar{r}_{i'j'}$$

and

$$\mathrm{Var}(\overline{D}_{ij}) = \frac{\sigma^2(R\text{-}G+1)(C\text{-}H+1)}{(R\text{-}G)(C\text{-}H)}$$

where:

r_{ij} = residual in (i,j) cell

$\bar{r}_{ij'}$ = average of the residuals in row i, outside of GH

$\bar{r}_{i'j}$ = average of the residuals in column j, outside of GH

$\bar{r}_{i'j'}$ = average of the residuals in neither rows or columns of GH.

An observation would be deleted from the subtable if its estimated disturbance, $\overline{D}_{ij}$, exceeded $3s(\overline{D}_{ij})$. No properties of this decision rule were reported. Daniel illustrated his detection technique with an example of a 7 x 7 matrix. He isolated the suspected residuals into a 4 x 3 subtable, and correctly identified five of the six outliers reported by Bradu. Selecting cells for the subtable involves some subjectivity. Although cells with large residuals are likely candidates for entry into the G x H subtable, other disturbances can be overlooked because their residuals are not outstanding in magnitude. I analyzed this same data set using Daniel's method and missed selecting one row for the subtable, because this row did not contain a large residual. Therefore, I identified four of the six outliers.

John and Draper (1978) developed an outlier test to check for the presence of two or fewer outliers. Draper and John (1980) extended this test to handle three or less outliers. Both tests are based on Gentleman and Wilk's (1975b) Q_k^* statistic. Recall that Q_k^* is the sum of squares of residuals due to K 'outlier' observations. Since residuals in a two-way table are correlated, John and Draper rewrote Q_k^* to be the sum of squares of K successive adjusted normalized uncorrelated residuals. Because I compare the performance of John and Draper's two outlier test to my procedure (Chapter 3.6..3), I discuss their two outlier test instead of their three outlier test. However, their three outlier test is merely an extension of their two outlier test.

When two outliers are suspected, John and Draper test the following hypothesis in stages:

Stage 1- H_o: no outliers vs. H_1 : least one outlier

Stage 2- H_o: one outlier vs. H_1 : at least two outliers

The procedure is to start at stage 1 and proceed to stage 2 as long as the null hypothesis is rejected. At each stage, a different F statistic is calculated. At stage 1, F_1 is defined as:

$$F_1 = \frac{Q_2/2}{RSS(H_1)/(v-2)}$$

where:

Q_2 is the largest residual sum of squares due to the deletion of two residuals, r_1 and r_2

$$Q_2 = \frac{rc(v(r_1^2 + r_2^2) - 2r_1r_2)}{v^2 - 1}$$

$$v = (r-1)(c-1)$$

$RSS(H_1)$ is the sum of squares of residuals when the observations that generated r_1 and r_2 are replaced by their expected values.

Critical points for the F_1 test statistic were simulated using a Monte Carlo simulation. They are tabulated for several table sizes ranging from a 6 x 8 to a 12 x 12 table (John and Draper, 1978).

If the null hypothesis at stage 1 is rejected, then the test statistic F_2 is calculated:

$$F_2 = \frac{Q_1}{RSS(H_1)/(v-2)}$$

where:

$Q_1 = Q_2 - rcr_1^2/v$

$RSS(H_1)$ is the same denominator as F_1.

Critical values for F_2 are tabulated for only an 8 x 12 and a 6 x 8 size table.

2.3 LEVERAGE AND INFLUENCE

The effect of an observation on fitted values and model estimates will now be addressed. Leverage measures the effect a change in an observation has on a fitted value. In a two-way table, all cells are affected by a perturbation in the (i,j) cell, because of model constraints. The size of the change in a fitted value depends upon its location relative to the perturbed cell (Emerson *et al.*, 1984). This change can be written in terms of a "hat table", H(i,j), when the data are additive. The "hat table" measures the effect a change in the (i,j) cell has on all fitted values in a r x c table. H(i,j) is defined as:

$$H(i,j) = e_i e_j' - (e_i - 1/r)(e_j' - 1'/c)$$

where:

e_i denotes a r x 1 vector with a 1 in position i and 0 elsewhere

e_j denotes a c x 1 vector with a 1 in position j and 0 elsewhere.

However under an additive-plus-multiplicative model, a change in a predicted value depends upon the location and the size of the perturbation. Emerson *et al.* measure this change by the generalized leverage table and the Jacobian leverage table for a rank 1 additive-plus-multiplicative model. Leverage tables for the additive-plus-multiplicative model of rank k (where: $k \geq 2$) do not exist. These are developed as part of this dissertation.

Jolliffe (1986), Calder (1986) and Pack *et al.* (1988) examined the influence of an observation on the estimated eigenvalues and eigenvectors in principal component analysis. They defined an influential observation as being an observation whose addition or deletion produced major changes in the analysis. An outlier may or may not be an influential observation. They constructed influence functions for the eigenvalues and eigenvectors of a covariance and a correlation matrix. Since the residuals from an additive model are decomposed using principal component analysis to fit the multiplicative portion of the model, these functions can be adapted to measure the influence of an observation in a two-way table. This avenue of research is pursued in this dissertation.

CHAPTER 3 - OUTLIERS AND INTERACTION

3.1 INTRODUCTION

The purpose of this chapter is to study the use of Mandel's additive-plus-multiplicative model to fit outliers in a nonreplicated two-way table. The reason for this investigation is that the current outlier tests (Daniel 1978, Stefansky 1972, John and Draper 1978, Bradu and Hawkins 1982) assume that the data in a two-way table are additive. The performance of these tests when the model is misspecified and interaction is present is unknown.

Mandel's model is designed to fit both additive and nonadditive effects in an unreplicated two-way layout. The beauty of this model lies in the versatility of the nonadditive term to account for several different arrangements of multiplicative interaction. Marasinghe and Johnson (1981) discuss various scenarios where Mandel's model can be used to describe interaction. They are: 1.) when interaction is present, but there are no main effect differences in either factor, 2.) when all the interaction is due to 1 cell, or 3.) when the interaction is confined to a row (or a column). In the setting of an unreplicated two-way layout, both interaction and outlier effects are sources of nonadditivity. The use of this model to describe outliers has not been reported in the literature.

This chapter is organized in the following manner. I first show how an outlier can be fit by the multiplicative term of Mandel's model in section 3.2. I extend the use of Mandel's model to accommodate two outliers. I then devote a significant portion of this chapter to looking into the use of the additive-plus-multiplicative model to distinguish outliers from interaction when both are present in a two-way table. I demonstrate how the model can easily account for additional sources of nonadditivity in section 3.3. However, using the model in practice to locate outliers when several nonadditive effects are present is a different story. Thus, I develop a rotation technique to aid in the identification of outliers when interaction exists in the data. The rotation procedure is based upon a Givens rotation. In section 3.4, I introduce the Givens rotation, explain where it came from and how I use it to uncover an outlier in the presence of interaction. I then incorporate this rotation technique with the use of Mandel's model. I develop a stepwise procedure to test for significant nonadditive effects, section 3.5, and demonstrate how to use it in two examples. Lastly in section 3.6, I evaluate the power of my procedure to distinguish outliers in the presence of interaction. I compare my procedure's results to the performance of Stefansky's test and John and Draper's test under identical interaction conditions.

3.2 JUSTIFICATION FOR USING MANDEL'S MODEL TO FIT OUTLIERS

An outlier which disturbs the additive structure of a two-way table can be described by the term $\delta\Delta_{ij}$ in Barnett and Lewis's (1984) "contamination" model:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \delta\Delta_{ij} + \varepsilon_{ij}$$

where:

$\Delta_{ij} = 1$ for $i = i', j = j'$ and

$\Delta_{ij} = 0$ otherwise.

However if these same data are fit with the additive-plus-multiplicative model, then the outlier shows up in the multiplicative part of this model. The multiplicative portion is a decomposition of nonadditive effects. In this case, it would be a decomposition of an outlier effect, $\delta\Delta_{ij}$, and the residuals, ε_{ij}. In more complex data sets, it could also include multiplicative interaction. The objective of this section is to show how one or two outliers can be fit by the nonadditive term(s) of the additive-plus-multiplicative model. I demonstrate this for several cases of outliers. These are: 1.) a single outlier, 2.) two outliers in the same row (or column) and 3.) two outliers of the same magnitude in different rows and different columns. To justify the use of this model to fit an outlier(s), it is necessary in all cases to first examine the residuals after fitting an additive model in the presence of an outlier (or outliers).

3.2.1 One Outlier

Gentleman and Wilk (1975a, 1975b) showed that the expected values of residuals from a two-way table with one outlier of magnitude θ_1 added to the (i,j) cell in a r x c table are:

θ_1	$\cdots$	θ_1	$\theta_1(c-1)$	θ_1	$\cdots$	θ_1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
θ_1	$\cdots$	θ_1	$\theta_1(c-1)$	θ_1	$\cdots$	θ_1
$-\theta_1(r-1)$	$\cdots$	$-\theta_1(r-1)$	$-\theta_1(r-1)(c-1)$	$-\theta_1(r-1)$	$\cdots$	$-\theta_1(r-1)$
θ_1	$\cdots$	θ_1	$\theta_1(c-1)$	θ_1	$\cdots$	θ_1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
θ_1	$\cdots$	θ_1	$\theta_1(c-1)$	θ_1	$\cdots$	θ_1

All cells are divided by rc. The table assumes no error.

In an analysis of variance design, all the residuals in the table are affected by an outlier. This is because the usual constraints on the model's parameters spread the outlier's effect throughout the table.

Gentleman and Wilk's expected residual table of rank 1 can be written as a single multiplicative term, $\frac{\theta_1}{rc} R_1 C_1'$ or $\lambda_1 \tau_1 \gamma_1'$, where:

$$R_1 = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ -(r-1) \\ 1 \\ \vdots \\ 1 \end{bmatrix} \qquad C_1 = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ -(c-1) \\ 1 \\ \vdots \\ 1 \end{bmatrix},$$

and

$$\sum_{i=1}^{r} R_{1i} = \sum_{j=1}^{c} C_{1j} = 0.$$

To show the equivalence between $\frac{\theta_1}{rc} R_1 C_1'$ and $\lambda_1 \tau_1 \gamma_1'$, let τ_1 and γ_1 represent the normalized vectors of R_1 and C_1, respectively, then

$$\frac{\theta_1}{rc} R_1 C_1' = \frac{\theta_1}{rc} \sqrt{\sum_{i=1}^{r} R_{1i}^2} \sqrt{\sum_{j=1}^{c} C_{1j}^2} \frac{R_1}{\sqrt{\sum_{i=1}^{r} R_{1i}^2}} \frac{C_1'}{\sqrt{\sum_{j=1}^{c} C_{1j}^2}}$$

$$= \frac{\theta_1}{rc} \sqrt{\sum_{i=1}^{r} R_{1i}^2} \sqrt{\sum_{j=1}^{c} C_{1j}^2}\, \tau_1 \gamma_1'.$$

Let

$$\lambda_1 = \frac{\theta_1}{rc} \sqrt{\sum_{i=1}^{r} R_{1i}^2} \sqrt{\sum_{j=1}^{c} C_{1j}^2} \qquad [3.2.1]$$

or

$$\lambda_1 = \frac{\theta_1}{rc} \sqrt{r(r-1)} \sqrt{c(c-1)}\,, \qquad [3.2.2]$$

then

$$\frac{\theta_1}{rc} R_1 C_1' = \lambda \tau_1 \gamma_1'.$$

The term, $\lambda_1\tau_1\gamma_1'$, is the singular value decomposition of the expected residual table. This enables an outlier effect to be described using the multiplicative part of the additive-plus-multiplicative model. The singular value, λ_1, is the magnitude of the perturbation scaled by constants which depend only on the size of the table. The left and right singular vectors, τ_1 and γ_1, are scaled versions of R_1 and C_1, and denote the row and column location of the outlier, respectively. Therefore in the presence of one outlier and no error, additive data are completely fitted with the following additive-plus-multiplicative model of rank 1:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \lambda_1 \tau_{1i} \gamma_{1j}$$

where:

$$\sum_{i=1}^{r} \alpha_i = \sum_{j=1}^{c} \beta_j = \sum_{i=1}^{r} \tau_{1i} = \sum_{j=1}^{c} \gamma_{1j} = 0 \quad \text{and} \quad \sum_{i=1}^{r} \tau_{1i}^2 = \sum_{j=1}^{c} \gamma_{1j}^2 = 1.$$

The following example demonstrates how Mandel's model fits a single outlier. This example considers a hypothetical data set with no error.

EXAMPLE 1-

Let there exist one outlier, $\theta_1 = 5$, at cell (3,2) in a 5 x 4 table. The observed response, Y, is:

$Y =$

100.25	100.35	100.45	100.55
100.30	100.40	100.50	100.60
100.35	**105.45**	100.55	100.65
100.40	100.50	100.60	100.70
100.45	100.55	100.65	100.75

.

Let R_1 and C_1 denote the row and column location of the outlier, i.e.,

$$R_1 = \begin{bmatrix} 1 \\ 1 \\ -4 \\ 1 \\ 1 \end{bmatrix} \qquad C_1 = \begin{bmatrix} 1 \\ -3 \\ 1 \\ 1 \end{bmatrix}$$

where:

$$\sum_{i=1}^{5} R_{1i}^2 = 20 \quad \text{and} \quad \sum_{j=1}^{4} C_{1j}^2 = 12.$$

Equation [3.2.1] produces the following results:

$$\lambda_1 = \frac{\theta_1 \sqrt{\sum_{i=1}^{r} R_{1i}^2} \sqrt{\sum_{j=1}^{c} C_{1j}^2}}{rc} = \frac{5\sqrt{20}\sqrt{12}}{(5)(4)} = 3.8730$$

$$\tau_1 = \begin{bmatrix} .2236 \\ .2236 \\ -.8944 \\ .2236 \\ .2236 \end{bmatrix} \qquad \gamma_1 = \begin{bmatrix} .2887 \\ -.8660 \\ .2887 \\ .2887 \end{bmatrix}.$$

The values of λ_1, τ_1 and γ_1 were verified using a SAS routine in PROC IML. The singular vectors contain important information about the location of the outlier. The left singular vector, τ_1, is a contrast between row 3 and the remaining rows indicating that the outlier is in row 3. Likewise, the right singular vector, γ_1, is a contrast between column 2 and the other columns denoting that the outlier is in column 2. When the additive-plus-multiplicative model is fit, one looks for contrasts similar to these to indicate the presence of an outlier.

3.2.2 Two Outliers

There are two possible configurations for the presence of two outliers in a r x c table: a.) both outliers in either the same row (or the same column) or b.) both outliers in different rows and different columns. In either arrangement, Gentleman and Wilk demonstrated that the expected values of residuals from two outliers is the sum of the appropriate expected residual matrices of single outliers.

a.) Two outliers in the same row (or the same column)

When two outliers are present in the same row (or column) of a two-way table, then Mandel's model of rank 1 fits these nonadditive effects exactly. To show this, perturb cell (i,j) and (k,j) by θ_1 and θ_2, respectively. The expected values of residuals is a matrix which is formed by:

$$\frac{\theta_1 R_1 C_1' + \theta_2 R_2 C_2'}{rc}.$$

Let R_2 be R_1 except that the quantity -(r-1) is in the kth row instead of the ith row (and ones elsewhere). Since both outliers are in the jth column, then $C_1 = C_2$. This reduces the expected residuals (Theorem 1.2.b, Kolman, 1976) to:

$$\frac{(\theta_1 R_1 + \theta_2 R_2) C_1'}{rc}.$$

Let

$$R = \theta_1 R_1 + \theta_2 R_2$$

$$= \begin{array}{cl} \left[\begin{array}{c} \theta_1 + \theta_2 \\ \vdots \\ \theta_1 + \theta_2 \\ -\theta_1(r-1) + \theta_2 \\ \theta_1 + \theta_2 \\ \vdots \\ \theta_1 + \theta_2 \\ \theta_1 - \theta_2(r-1) \\ \theta_1 + \theta_2 \\ \vdots \\ \theta_1 + \theta_2 \end{array}\right] & \begin{array}{l} \\ \\ \\ \leftarrow \textbf{ith row} \\ \\ \\ \\ \leftarrow \textbf{kth row} \\ \\ \\ \\ \end{array} \end{array}$$

then

$$\frac{\theta_1 R_1 C_1' + \theta_2 R_2 C_2'}{rc} = \frac{RC_1'}{rc}$$

$$= \frac{\sqrt{\sum_{i=1}^{r} R_i^2}\sqrt{\sum_{j=1}^{c} C_{1j}^2}}{rc} \frac{R}{\sqrt{\sum_{i=1}^{r} R_i^2}} \frac{C_1'}{\sqrt{\sum_{j=1}^{c} C_{1j}^2}}$$

$$= \frac{\sqrt{\sum_{i=1}^{r} R_i^2}\sqrt{\sum_{j=1}^{c} C_{1j}^2}}{rc} \tau_1 \gamma_1'.$$

Let

$$\lambda_1 = \frac{\sqrt{\sum_{i=1}^{r} R_i^2}\sqrt{\sum_{j=1}^{c} C_{1j}^2}}{rc}. \qquad [3.2.3]$$

Finally to summarize,

$$\frac{\theta_1 R_1 C_1' + \theta_2 R_2 C_2'}{rc} = \lambda_1 \tau_1 \gamma_1'.$$

Thus the two outliers in the same row are represented by a rank 1 multiplicative component.

EXAMPLE 2-

Let there exist two outliers $\theta_1 = 5$ and $\theta_2 = 8$ at (1,1) and (5,1), respectively, in a 5 x 4 table. The response, Y, is:

$$Y = \begin{array}{|c|c|c|c|}\hline \mathbf{105.25} & 100.35 & 100.45 & 100.55 \\ \hline 100.30 & 100.40 & 100.50 & 100.60 \\ \hline 100.35 & 100.45 & 100.55 & 100.65 \\ \hline 100.40 & 100.50 & 100.60 & 100.70 \\ \hline \mathbf{108.45} & 100.55 & 100.65 & 100.75 \\ \hline \end{array},$$

then

$$R_1 = \begin{bmatrix} -4 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad C_1 = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$R_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ -4 \end{bmatrix} \qquad C_2 = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$R = 5R_1 + 8R_2 \qquad C_1 =$$

$$= \begin{bmatrix} 5(-4) + 8 \\ 5 + 8 \\ 5 + 8 \\ 5 + 8 \\ 5 + 8(-4) \end{bmatrix} = \begin{bmatrix} -12 \\ 13 \\ 13 \\ 13 \\ -27 \end{bmatrix} \qquad = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

where:

$$\sum_{i=1}^{5} R_i^2 = 1380 \qquad \sum_{j=1}^{4} C_{1j}^2 = 12.$$

By equation [3.2.3]

$$\lambda_1 = \frac{\sqrt{1380}\,\sqrt{12}}{(5)(4)} = 6.4343$$

and

$$\tau_1 = \begin{bmatrix} \mathbf{-.3231} \\ .3499 \\ .3499 \\ .3499 \\ \mathbf{-.7268} \end{bmatrix} \qquad \gamma_1 = \begin{bmatrix} \mathbf{-.8660} \\ .2887 \\ .2887 \\ .2887 \end{bmatrix}.$$

Again, the values of λ_1, τ_1 and γ_1 were verified using SAS PROC IML. When two outliers exist in the same column (or row), then τ_1 (or γ_1) is a contrast between the two rows containing the outliers verses the remaining rows. If the magnitudes of the outlier effects are different, then the weights of the two rows designating the location of outliers are unequal. The stronger outlier has the larger weight. Since both outliers are in the same column, then γ is a contrast between the column containing both outliers verses the other columns. This pattern holds for several outliers in a row (or column).

b.) Two outliers in different rows and different columns

Results have been obtained for this configuration of outliers when $\theta_1 = \theta_2$ or $\theta_1 = -\theta_2$, but not otherwise. I will present these findings when the perturbations are the same magnitude. When the disturbances are unequal in size, then the singular values and singular vectors are not easily solved in terms of θ's, R's and C's. This is due to the complexity of the residual matrix and the SVD constraint that $\tau_1' \bullet \tau_2 = 0$ and $\gamma_1' \bullet \gamma_2 = 0$. Also, the above theory does not apply to a symmetric matrix when $\theta_1 = -\theta_2$ Multiple singular values are observed which indicate that the decomposition is not unique.

Mandel's model of rank 2 fits the effect of two outliers in different rows and columns. To see this perturb cell (i,j) and (k,l) by θ_1 and θ_2, respectively. Once again, the expected residuals are:

$$\frac{\theta_1 R_1 C_1' + \theta_2 R_2 C_2'}{rc}.$$

When $|\theta_1| = |\theta_2|$, then the residuals can be factored into:

$$\frac{(\theta_1 R_1 - \theta_2 R_2)(|\theta_1| C_1 - |\theta_2| C_2)}{2|\theta_1|} + \frac{(\theta_1 R_1 + \theta_2 R_2)(|\theta_1| C_1 + |\theta_2| C_2)}{2|\theta_1|}.$$

To demonstrate this let

$$R^- = \theta_1 R_1 - \theta_2 R_1 \qquad C^- = |\theta_1| C_1 - |\theta_2| C_2$$

and

$$R^+ = \theta_1 R_1 + \theta_2 R_2 \qquad C^+ = |\theta_1| C_1 + |\theta_2| C_2$$

then

$$\frac{\theta_1 R_1 C_1' + \theta_2 R_2 C_2'}{rc} = \frac{\dfrac{R^- C^{-\prime} + R^+ C^{+\prime}}{2|\theta_1|}}{rc}$$

$$= \frac{\dfrac{\sqrt{\sum_{i=1}^{r}(R_i^-)^2}}{2|\theta_1|}\dfrac{\sqrt{\sum_{j=1}^{c}(C_j^-)^2}}{2|\theta_1|}}{rc}\frac{R^-}{\sqrt{\sum_{i=1}^{r}(R_i^-)^2}}\frac{C^{-\prime}}{\sqrt{\sum_{j=1}^{c}(C_j^-)^2}}$$

$$+ \frac{\dfrac{\sqrt{\sum_{i=1}^{r}(R_i^+)^2}}{2|\theta_1|}\dfrac{\sqrt{\sum_{j=1}^{c}(C_j^+)^2}}{2|\theta_1|}}{rc}\frac{R^+}{\sqrt{\sum_{i=1}^{r}(R_i^+)^2}}\frac{C^{+\prime}}{\sqrt{\sum_{j=1}^{c}(C_j^+)^2}}$$

$$= \frac{\dfrac{\sqrt{\sum_{i=1}^{r}(R_i^-)^2}}{2|\theta_1|}\dfrac{\sqrt{\sum_{j=1}^{c}(C_j^-)^2}}{2|\theta_1|}}{rc}\tau_1\gamma_1'$$

$$+ \frac{\dfrac{\sqrt{\sum_{i=1}^{r}(R_i^+)^2}}{2|\theta_1|}\dfrac{\sqrt{\sum_{j=1}^{c}(C_j^+)^2}}{2|\theta_1|}}{rc}\tau_2\gamma_2'.$$

Let

$$\lambda_1 = \frac{\dfrac{\sqrt{\sum_{i=1}^{r}(R_i^-)^2}}{2|\theta_1|}\dfrac{\sqrt{\sum_{j=1}^{c}(C_j^-)^2}}{2|\theta_1|}}{rc} \qquad [3.2.4A]$$

and

$$\lambda_2 = \frac{\dfrac{\sqrt{\sum_{i=1}^{r}(R_i^+)^2}}{2|\theta_1|}\dfrac{\sqrt{\sum_{j=1}^{c}(C_j^+)^2}}{2|\theta_1|}}{rc}. \qquad [3.2.4B]$$

Therefore,

$$\frac{\theta_1 R_1 C_1' + \theta_2 R_2 C_2'}{rc} = \lambda_1 \tau_1 \gamma_1' + \lambda_2 \tau_2 \gamma_2'.$$

Thus, two outliers in different rows and columns are fit by two multiplicative components.

EXAMPLE 3-

Let there be two outliers, $\theta_1 = 5$ and $\theta_2 = -5$ at (1,1) and (5,4), respectively, in a 5 x 4 table.

$$R_1 = \begin{bmatrix} -4 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad C_1 = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad R_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ -4 \end{bmatrix} \quad C_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -3 \end{bmatrix}$$

$$R^- = 5R_1 - (-5)R_2 = \begin{bmatrix} 5(-4)+5 \\ 5+5 \\ 5+5 \\ 5+5 \\ 5+5(-4) \end{bmatrix} = \begin{bmatrix} -15 \\ 10 \\ 10 \\ 10 \\ -15 \end{bmatrix}$$

$$C^- = |5|C_1 - |-5|C_2 = 5C_1 - 5C_2 = \begin{bmatrix} 5(-3)-5 \\ 5-5 \\ 5-5 \\ 5-5(-3) \end{bmatrix} = \begin{bmatrix} -20 \\ 0 \\ 0 \\ 20 \end{bmatrix}$$

$$\sum_{i=1}^{5}(R^-)^2 = 750 \qquad \sum_{j=1}^{4}(C^-)^2 = 800$$

$$R^{+} = 5R_1 + (-5)R_2 = \begin{bmatrix} -25 \\ 0 \\ 0 \\ 0 \\ 25 \end{bmatrix} \qquad C^{+} = |5|C_1 + |-5|C_2 = 5C_1 + 5C_2 = \begin{bmatrix} -10 \\ 10 \\ 10 \\ -10 \end{bmatrix}$$

$$\sum_{i=1}^{5} (R^{+})^2 = 1250 \qquad \sum_{j=1}^{4} (C^{+})^2 = 400.$$

By equations [3.2.4A] and [3.2.4B]:

$$\lambda_1 = \frac{\sqrt{750/2 \times 5}\,\sqrt{800/2 \times 5}}{(5)(4)} = 3.8730 \quad \text{and} \quad \lambda_2 = \frac{\sqrt{1250/2 \times 5}\,\sqrt{400/2 \times 5}}{(5)(4)} = 3.5350$$

and

$$\tau_1 = \begin{bmatrix} -.5477 \\ .3651 \\ .3651 \\ .3651 \\ -.5477 \end{bmatrix} \quad \tau_2 = \begin{bmatrix} -.7071 \\ 0 \\ 0 \\ 0 \\ .7071 \end{bmatrix} \quad \gamma_1 = \begin{bmatrix} -.7071 \\ 0 \\ 0 \\ .7071 \end{bmatrix} \quad \gamma_2 = \begin{bmatrix} -.5000 \\ .5000 \\ .5000 \\ -.5000 \end{bmatrix}.$$

Again, these values of $\lambda_1, \tau_1, \gamma_1, \lambda_2, \tau_2$ and γ_2 were verified using SAS PROC IML. When two outliers of equal magnitude occur in different rows and different columns, then τ_1 is a contrast between the two rows containing the outliers verses the rows without outliers and τ_2 is a contrast between the two rows possessing the outliers. This pattern is the same for γ_1 and γ_2, except the components switch order. When the outliers are of unequal magnitude, then these contrasts may or may not hold.

In summary, I have shown that a single outlier or two outliers in the same row (or column) can be fit with one multiplicative term of the additive-plus-multiplicative model. Two nonadditive terms are necessary to account for two outliers in different rows and different columns. The singular vectors associated with the multiplicative component(s) reveal specific contrasts which identify the location(s) of the outlier(s) in a two-way table. However, most data sets are more complicated than the examples presented, and can include additional outliers and/or interaction. The versatility of Mandel's model allows one to investigate its usage in fitting outliers in the presence of interaction.

3.3 USE OF MANDEL'S MODEL TO FIT SEVERAL SOURCES OF NONADDITIVITY

We saw in the two outlier case (section 3.2.2) that outlier effects combine simply and become a part of the residuals from the additive part of the model. This result can be extended to include other nonadditive effects as well as outliers. For example, assume that some rank 1 interaction and an outlier are present. The expected residuals from these two effects are:

$$\lambda_1 U_1 V_1' + \frac{\theta_1}{rc} R_1 C_1'$$

where:

$\lambda_1 U_1 V_1'$ represents rank 1 interaction

$\frac{\theta_1}{rc} R_1 C_1'$ denotes an outlier effect.

This equation can be expanded to accommodate interaction of rank k and/or p outliers. The expected residuals will look like:

$$\lambda_1 U_1 V_1' + \cdots + \lambda_k U_k V_k' + \frac{\theta_1 R_1 C_1' + \cdots + \theta_p R_p C_p'}{rc}.$$

This means that under ideal conditions, i.e., no error, up to $p+k$ multiplicative components are needed to fit the $p+k$ nonnadditive effects provided that $p+k < \min\{(r,c)\}$. If all of these effects are unique and orthogonal, then exactly $p+k$ terms are necessary in the model. Usually orthogonality is not the case, and this is when data analysis becomes difficult.

The analysis is more complicated when two or more nonadditive factors are present, because the effects can combine in a manner which disguise the true sources of nonadditivity. In addition, the model uses a least squares criteria to fit the multiplicative components to the nonadditive term. The components are fit orthogonally, and there is no guarantee that the effects will always partition into separate components. When this happens, examining the singular vectors may not reveal enough information to uncover all the sources of nonadditivity in a data set.

To obtain an idea of how this can complicate the analysis, let us examine an example where there are two outliers of unequal magnitude in different rows and columns of a two-way table. I would like to keep this example simple, so I will not include any error in the observed data.

EXAMPLE 4-

Let there exist two outliers, $\theta_1 = 5$ at (1,1) and $\theta_2 = 8$ at (5,4) in a 5 x 4 table. The observed response, Y, is:

Y =

105.25	100.35	100.45	100.55
100.30	100.40	100.50	100.60
100.35	100.45	100.55	100.65
100.40	100.50	100.60	100.70
100.45	100.55	100.65	**108.75**

The decomposition of the residuals produces the following results:

$$\lambda_1 = 6.8731 \quad \text{and} \quad \lambda_2 = 3.1877$$

$$\tau = \begin{bmatrix} -.4924 & .7467 \\ -.1179 & -.3456 \\ -.1179 & -.3456 \\ -.1179 & -.3456 \\ .8461 & .2901 \end{bmatrix} \qquad \gamma = \begin{bmatrix} -.5149 & .6969 \\ -.1566 & -.4748 \\ -.1566 & -.4748 \\ .8282 & .2533 \end{bmatrix}.$$

The vector τ reveals that column 1 is a contrast between rows 1-4 verses 5, and column 2 is a contrast between rows 1 and 5 verses 2-4. As expected, these contrasts did not separate clearly. Row 5 dominates row 1 in the first component of τ, but this pattern is reversed in the second component. Also in the second component, rows 1 and 5 have the same sign. This same arrangement exists for the right singular vector. This suggests that an outlier's effect is "spread" between both components of τ and γ with the larger outlier dominating the first component of both singular vectors. This same result can occur if an outlier and interaction are present in the table instead of two outliers. Therefore, a method is needed to overcome this "spreading" effect. I have developed a technique which uses a Givens rotation to handle this situation. The Givens rotation is discussed in section 3.4.

3.4 GIVENS ROTATION

Givens (1954) developed a 2 x 2 orthogonal rotation which zeros a particular element of a matrix (Kenndey and Gentle, 1980). These rotations are used in the numerical computation of the singular value decomposition (SVD) of a matrix. This common application motivated me to study their use as a rotation technique to uncover the nature of significant nonadditivity in a two-way table.

In this section, I first define a Givens rotation and demonstrate how to calculate this 2 x 2 orthogonal matrix. Second, I show the role that Givens rotations play in the decomposition of a matrix. Third, I discuss the use of these rotations to isolate outliers in a two-way layout by zeroing selected components of the singular vectors. Fourth, I derive this result for a two-way table with two outliers of equal magnitude located in different rows and different columns. However the rotation's effectiveness is not limited to outliers of equal magnitude, and I demonstrate this by applying Givens rotations to the left singular vector in EXAMPLE 4 (section 3.3).

A Givens rotation is designed to introduce a zero into a selected cell of a matrix (or vector). With the appropriate value of θ, the element b of the vector $[a \quad b]$ is zeroed and a new vector is generated, i.e.,

$$[a \quad b] \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = [r \quad 0]. \qquad [3.4.1]$$

This rotation produces two equations which can be solved simultaneously to find $\sin\theta$ and $\cos\theta$:

$$a\cos\theta - b\sin\theta = r \qquad [3.4.2]$$

$$a\sin\theta + b\cos\theta = 0 \qquad [3.4.3]$$

Therefore,

If $b = 0$ then

$$\sin\theta = 0 \quad \text{and} \quad \cos\theta = 1 \qquad [3.4.4]$$

Else

If $|b| > |a|$ then

$$\sin\theta = \frac{b}{\sqrt{a^2 + b^2}} \quad \text{and} \quad \cos\theta = \frac{-a}{\sqrt{a^2 + b^2}} \qquad [3.4.5]$$

Else

$$\text{SIN } \theta = \frac{-\,b}{\sqrt{a^2 + b^2}} \quad \text{and} \quad \text{COS } \theta = \frac{a}{\sqrt{a^2 + b^2}}. \qquad [3.4.6]$$

However distinguishing between equations [3.4.5] and [3.4.6] is not important when these rotations are applied to singular vectors. These equations merely rotate the singular vector in different directions which result in a sign change.

Therefore, substituting the values for SIN θ and COS θ from [3.4.6] into equation [3.4.2] and solving for r yields:

$$r = \sqrt{a^2 + b^2}.$$

I would like to show how these rotations are used in the computation of the SVD for a matrix X. Recall the SVD of X is defined as:

$$X_{mxn} = U_{mxk} D_{kxk} V'_{kxn}$$

where:

$m \geq n$

D is a diagonal matrix of $\lambda_1 \geq \cdots \geq \lambda_k \geq 0$

$k \leq \min \{(m-1, n-1)\}$

$U'U = V'V = I.$

Computing the decomposition of X is a two stage process (Golub and Reinsch, 1971). The first phase is to reduce X to an upper bidiagonal matrix, B, by pre- and postmultiplying X with Householder transformations, i.e,

$$U^n \ldots U^1 X\, V^1 \ldots V^{n-2} = B$$

where:

U^p is a Householder transformation that zeros (m-p) cells in the pth column of X

V^p is a Householder transformation that zeros (n-p-1) cells in the pth column of X.

Since an indepth discussion of Householder transformations is not germane to this section, the reader is referred to Golub and VanLoan (1989) for more information.

The second step zeros the superdiagonal elements of B. This is accomplished by introducing a nonzero element into the (2,1) position of B, then "chasing" this element down the bidiagonal by pre- and postmultiplying B with a series of Givens rotations. These rotations effectively reduce the superdiagonal elements below some acceptable tolerance level to produce:

$$(U_n' \ldots U_2')B(G_2V_3 \ldots V_n) = D$$

where:

G_2 is a Givens rotation to insert a nonzero element in (2,1)

U_p is a Givens rotation to zero the (p,p-1) cell and generate an entry in cell (p-1,p+1)

V_p is a Givens rotation to zero the (p-2,p) cell and produce an entry in cell (p,p-1).

Additional steps are sometimes necessary to zero the superdiagonal. These are not presented, since they are beyond the scope of this discussion. They may be found in Kennedy and Gentle (1980).

To summarize the first part, let

$$U_B = U^1 \ldots U^n \text{ and } V_B = V^1 \ldots V^{n-2}$$

and in the second part, let

$$U_D = U_2 \dots U_n \quad \text{and} \quad V_D = G_2V_3 \dots V_n,$$

then let

$$U = U_BU_D \quad \text{and} \quad V = V_BV_D,$$

so that

$$U'XV = D.$$

Therefore the singular vectors, U and V, are a product of several Householeder transformations and Givens rotations designed to ultimately generate the diagonal matrix D.

Although my technique is to multiply U and V with an additional Givens rotation, my purpose is to introduce a zero into a particular element of U or V. I do this to isolate nonadditive effects, like outliers, into separate columns of a rotated singular vector. This rotation can undo the ″spreading″ effect (due to constraints of SVD) that was seen in EXAMPLE 4 of section 3.3.

I now derive this result for two outliers of equal magnitude in different rows and columns of a two-way table. I have chosen this case because it is the only situation where I can generalize the composition of the two multiplicative components. When cells (i,j) and (k,l) (where $i \neq k$, $j \neq l$) are perturbed by θ_1 and θ_2, respectively, then we saw (section 3.2.2) that two multiplicative components are necessary to fit these outlier effects, i.e.,

$$\frac{(\theta_1 R_1 - \theta_2 R_2)(|\theta_1| C_1 - |\theta_2| C_2)}{2|\theta_1|} \quad \text{and} \quad \frac{(\theta_1 R_1 + \theta_2 R_2)(|\theta_1| C_1 + |\theta_2| C_2)}{2|\theta_1|}.$$

The first factor forms the first multiplicative component, whereas the second forms $\lambda_2 \tau_2 \gamma_2'$. Consider the case where $\theta_1 = \theta_2$. If $\theta_1 = -\theta_2$, then the two multiplicative components switch position.

Let R represent the left singular vector, τ, which has not been normalized. The vector R is composed of:

$$\mathrm{R} = [\ \mathrm{R}^- \quad \mathrm{R}^+\ ].$$

Recall that R^- and R^+ were defined (section 3.2.2) to be:

$$\mathrm{R}^- = \theta_1\mathrm{R}_1 - \theta_1\mathrm{R}_2 \quad \text{and} \quad \mathrm{R}^+ = \theta_1\mathrm{R}_1 + \theta_1\mathrm{R}_1.$$

In this case, R^- and R^+ are:

$$\mathrm{R}^- = \begin{bmatrix} \theta_1 - \theta_2 \\ \vdots \\ \theta_1 - \theta_2 \\ -(r-1)\theta_1 - \theta_2 \\ \theta_1 - \theta_2 \\ \vdots \\ \theta_1 - \theta_2 \\ \theta_1 + (r-1)\theta_2 \\ \theta_1 - \theta_2 \\ \vdots \\ \theta_1 - \theta_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -r\theta_1 \\ 0 \\ \vdots \\ 0 \\ r\theta_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{c} \\ \\ \\ \leftarrow \textbf{ith row} \rightarrow \\ \\ \\ \\ \leftarrow \textbf{kth row} \rightarrow \\ \\ \\ \\ \end{array} \mathrm{R}^+ = \begin{bmatrix} \theta_1 + \theta_2 \\ \vdots \\ \theta_1 + \theta_2 \\ -(r-1)\theta_1 + \theta_2 \\ \theta_1 + \theta_2 \\ \vdots \\ \theta_1 + \theta_2 \\ \theta_1 - (r-1)\theta_2 \\ \theta_1 + \theta_2 \\ \vdots \\ \theta_1 + \theta_2 \end{bmatrix} = \begin{bmatrix} 2\theta_1 \\ \vdots \\ 2\theta_1 \\ -(r-2)\theta_1 \\ 2\theta_1 \\ \vdots \\ 2\theta_1 \\ -(r-2)\theta_1 \\ 2\theta_1 \\ \vdots \\ 2\theta_1 \end{bmatrix}$$

and

$$\sum_{i=1}^{r} (\mathrm{R}_i^-)^2 = 2r^2\theta_1^2 \quad \text{and} \quad \sum_{i=1}^{r} (\mathrm{R}_i^+)^2 = 2r(r-2)\theta_1^2.$$

Therefore standardizing R produces:

$$
\tau = \left[\begin{array}{cc}
0 & \dfrac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\[2ex]
\vdots & \vdots \\[2ex]
0 & \dfrac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\[2ex]
\dfrac{-r\theta_1}{\sqrt{2r^2\theta_1^2}} & \dfrac{-(r-2)\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\[2ex]
0 & \dfrac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\[2ex]
\vdots & \vdots \\[2ex]
0 & \dfrac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\[2ex]
\dfrac{r\theta_1}{\sqrt{2r^2\theta_1^2}} & \dfrac{-(r-2)\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\[2ex]
0 & \dfrac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\[2ex]
\vdots & \vdots \\[2ex]
0 & \dfrac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\[2ex]
\vdots & \vdots \\[2ex]
0 & \dfrac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}}
\end{array}\right]
\begin{array}{l}
\\[2ex]
\\[2ex]
\\[2ex]
\leftarrow \textbf{ith row} \\[2ex]
\\[2ex]
\\[2ex]
\\[2ex]
\leftarrow \textbf{kth row} \\[2ex]
\\[2ex]
\\[2ex]
\leftarrow \textbf{mth row} \\[2ex]
\\[2ex]
\\
\end{array}
$$

Since the ith and kth rows of τ contain outlier effects, three cases of rotation are considered. They are rotations designed for row i (Case 1), row k (Case 2) and neither row i or k, say m, (Case 3).

Case 1. A rotation for row i:

The ith row of τ is:

$$\begin{bmatrix} \dfrac{-r\theta_1}{\sqrt{2r^2\theta_1^2}} & \dfrac{-(r-2)\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \end{bmatrix},$$

therefore,

$$\text{SIN } \theta = \frac{\dfrac{(r-2)\theta_1}{\sqrt{2r(r-2)\theta_1^2}}}{\sqrt{\dfrac{r^2\theta_1^2}{2r^2\theta_1^2} + \dfrac{(r-2)^2\theta_1^2}{2r(r-2)\theta_1^2}}} = \sqrt{\frac{r-2}{2(r-1)}}$$

and

$$\text{COS } \theta = \frac{\dfrac{-r\theta_1}{\sqrt{2r^2\theta_1^2}}}{\sqrt{\dfrac{r^2\theta_1^2}{2r^2\theta_1^2} + \dfrac{(r-2)^2\theta_1^2}{2r(r-2)\theta_1^2}}} = -\sqrt{\frac{r}{2(r-1)}}.$$

Now, rotate τ by the appropriate Givens rotation:

$$= \tau \begin{bmatrix} \text{COS } \theta & \text{SIN } \theta \\ -\text{ SIN } \theta & \text{COS } \theta \end{bmatrix}$$

$$
= \begin{bmatrix}
0 & \frac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\
\vdots & \vdots \\
0 & \frac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\
\frac{-r\theta_1}{\sqrt{2r^2\theta_1^2}} & \frac{-(r-2)\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\
0 & \frac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\
\vdots & \vdots \\
0 & \frac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\
\frac{r\theta_1}{\sqrt{2r^2\theta_1^2}} & \frac{-(r-2)\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\
0 & \frac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\
\vdots & \vdots \\
0 & \frac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\
\vdots & \vdots \\
0 & \frac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}}
\end{bmatrix}
\begin{bmatrix}
-\sqrt{\frac{r}{2(r-1)}} & \sqrt{\frac{r-2}{2(r-1)}} \\
-\sqrt{\frac{r-2}{2(r-1)}} & -\sqrt{\frac{r}{2(r-1)}}
\end{bmatrix}
$$

$$
= \left[\begin{array}{cc}
\dfrac{-1}{\sqrt{r(r-1)}} & \dfrac{-1}{\sqrt{(r-1)(r-2)}} \\
\vdots & \vdots \\
\dfrac{-1}{\sqrt{r(r-1)}} & \dfrac{-1}{\sqrt{(r-1)(r-2)}} \\
\sqrt{\dfrac{r-1}{r}} & 0 \\
\dfrac{-1}{\sqrt{r(r-1)}} & \dfrac{-1}{\sqrt{(r-1)(r-2)}} \\
\vdots & \vdots \\
\dfrac{-1}{\sqrt{r(r-1)}} & \dfrac{-1}{\sqrt{(r-1)(r-2)}} \\
\dfrac{-1}{\sqrt{r(r-1)}} & \sqrt{\dfrac{r-2}{r-1}} \\
\dfrac{-1}{\sqrt{r(r-1)}} & \dfrac{-1}{\sqrt{(r-1)(r-2)}} \\
\vdots & \vdots \\
\dfrac{-1}{\sqrt{r(r-1)}} & \dfrac{-1}{\sqrt{(r-1)(r-2)}} \\
\vdots & \vdots \\
\dfrac{-1}{\sqrt{r(r-1)}} & \dfrac{-1}{\sqrt{(r-1)(r-2)}}
\end{array}\right]
\begin{array}{l}
\vphantom{\dfrac{-1}{\sqrt{r(r-1)}}} \\
\vphantom{\vdots} \\
\vphantom{\dfrac{-1}{\sqrt{r(r-1)}}} \\
\vphantom{\sqrt{\dfrac{r-1}{r}}} \leftarrow \textbf{ith row} \\
\vphantom{\dfrac{-1}{\sqrt{r(r-1)}}} \\
\vphantom{\vdots} \\
\vphantom{\dfrac{-1}{\sqrt{r(r-1)}}} \\
\vphantom{\dfrac{-1}{\sqrt{r(r-1)}}} \leftarrow \textbf{kth row} \\
\vphantom{\dfrac{-1}{\sqrt{r(r-1)}}} \\
\vphantom{\vdots} \\
\vphantom{\dfrac{-1}{\sqrt{r(r-1)}}} \leftarrow \textbf{mth row} \\
\vphantom{\vdots} \\
\vphantom{\dfrac{-1}{\sqrt{r(r-1)}}}
\end{array}
$$

So rotating τ with a Givens rotation designed to zero the second element in row i, produces a r x 2 matrix whose first column is a contrast between the ith row and the remaining rows and whose second column is a contrast between the kth row and the other rows. The (i,2) element is zero by definition, so this element does not provide any useful information about the second column. Rotating the right singular vector is similar to rotating τ and will not be considered. Case 2 is similar to Case 1 with k substituted for i, therefore it is not presented. Case 3 is in Appendix A.

I now apply this technique to τ in EXAMPLE 4 to demonstrate that this rotation also works in cases other than $|\theta_1| = |\theta_2|$. Recall, $\theta_1 = 5$ is in row 1, $\theta_2 = 8$ is in row 5 and

$$\tau = \begin{bmatrix} -.4924 & .7467 \\ -.1179 & -.3456 \\ -.1179 & -.3456 \\ -.1179 & -.3456 \\ .8461 & .2901 \end{bmatrix},$$

therefore

$$\text{SIN } \theta = -.8348 \quad \text{and} \quad \text{COS } \theta = -.5505.$$

Rotating τ by the appropriate rotation produces:

$$
= \begin{bmatrix} -.4924 & .7467 \\ -.1179 & -.3456 \\ -.1179 & -.3456 \\ -.1179 & -.3456 \\ .8461 & .2901 \end{bmatrix} \quad \begin{bmatrix} -.5505 & -.8348 \\ .8348 & -.5505 \end{bmatrix}
$$

$$
= \begin{bmatrix} .8944 & 0 \\ -.2236 & .2887 \\ -.2236 & .2887 \\ -.2236 & .2887 \\ -.2236 & -.8660 \end{bmatrix}.
$$

When r = 5, then the above outcome is equal to the general result derived for $\theta_1 = \theta_2$:

$$
\begin{bmatrix} \sqrt{\frac{r-1}{r}} & 0 \\ \frac{-1}{\sqrt{r(r-1)}} & \frac{1}{\sqrt{(r-1)(r-2)}} \\ \frac{-1}{\sqrt{r(r-1)}} & \frac{1}{\sqrt{(r-1)(r-2)}} \\ \frac{-1}{\sqrt{r(r-1)}} & \frac{1}{\sqrt{(r-1)(r-2)}} \\ \frac{-1}{\sqrt{r(r-1)}} & -\sqrt{\frac{r-2}{r-1}} \end{bmatrix}.
$$

This result makes sense for the rotated τ, since singular vectors denote the location of the outlier and do not depend on the magnitude of θ unless they occur in the same row (or column).

3.5 RECOMMENDED PROCEDURE

I recommend that the following procedure be used in conjunction with examining the residuals to uncover significant sources of nonadditivity in a nonreplicated two-way analysis of variance.

1. Fit the additive part of the additive-plus-multiplicative model and obtain the residuals.

2. Fit the multiplicative part of this model to the residuals and examine the singular vectors for information on the nature of the nonadditivity.

3. Test for significance of the multiplicative parameters, λ_i, using a stepwise procedure. Milliken and Johnson (1989) recommend testing for λ_i until two consecutive test results are nonsignificant. The appropriate tests are several likelihood ratio tests, U_1, U_2 and U_i, developed by Johnson and Graybill (1972), Hegemann and Johnson (1976) and Yochmowitz and Cornell (1978), respectively, where:

$$U_1 = \frac{\hat{\lambda}_1^2}{\hat{\lambda}_1^2 + \cdots + \hat{\lambda}_k^2} \qquad U_2 = \frac{\hat{\lambda}_2^2}{\hat{\lambda}_2^2 + \cdots + \hat{\lambda}_k^2} \qquad \cdots \qquad U_i = \frac{\hat{\lambda}_i^2}{\sum_{j=i}^{k} \hat{\lambda}_j^2}$$

and

$$i = 3, \ldots, k.$$

Critical points for U_1, U_2 and U_i are tabulated in Milliken and Johnson (1989).

4. If at least two multiplicative parameters are significant, then apply a Givens rotation to the left and right singular vectors separately.

Of course, Step 3 is not entirely necessary. If the likelihood ratio tests fail to find any significant test results, then the rotation technique can be used on an exploratory basis. This is especially true if the tables are small, because of the conservative nature of the tests.

When three or more, say q, components are significant, then I recommend sequentially rotating the singular vectors q-1 times. Since the Givens rotation is a 2 x 2 orthogonal rotation, only two components of the singular vector can be rotated at one time. I would start by rotating the information from the last component, q, into the q-1 component. This combines information from the last two components and changes the composition of the q-1 component. I would proceed to rotate the new q-1 and the original q-2 component, and continue until I rotated the second and first components. The last rotation contains information from all the previous rotations and is the one that should be examined to gain an interpretation of the nonadditivity.

Two examples are given to illustrate this procedure. The first example attempts to distinguish an outlier from interaction. This capability is useful when testing for outliers under a misspecified model. Outlier tests assume an additive model and may not perform well in the presence of interaction. I show this by using Stefansky's test to detect the same outlier in the following data. This example resembles a typical run in the simulation of section 3.6.1. The second example looks at using this procedure to identify two outliers and interaction.

EXAMPLE 5-

Let there be one outlier, $\theta_1 = 25$ at (1,4) and linear interaction in a 5 x 4 table of the form:

$$\lambda_1 = 50 \quad \tau_1' = [\,-2\ \ -1\ \ \ 0\ \ \ 1\ \ \ 2] \quad \gamma_1' = [\,-3\ \ -1\ \ \ 1\ \ \ 3] \quad \varepsilon_{ij} \sim N(0,1)$$

where:

τ_1 and γ_1 are normalized so that $\tau_1'\tau_1 = \gamma_1'\gamma_1 = 1$.

The observed response, Y, is:

$$Y = \begin{array}{|c|c|c|c|} \hline 121.80 & 107.90 & 94.573 & \mathbf{105.70} \\ \hline 112.30 & 104.00 & 96.451 & 89.237 \\ \hline 99.999 & 101.20 & 101.10 & 99.654 \\ \hline 88.914 & 97.579 & 103.70 & 111.90 \\ \hline 79.066 & 93.324 & 107.10 & 121.50 \\ \hline \end{array} .$$

Notice that the observation in cell (1,4) does not appear to stand out among the remaining observations. The decomposition of the residuals produces the following results:

$$\hat{\lambda}_1 = 41.6804 \qquad U_1 = .9357 \ ^{**}$$

$$\hat{\lambda}_2 = 10.8955 \qquad U_2 = .9945 \ ^{*}$$

$$\hat{\lambda}_3 = \ 0.8125 \qquad U_3 = 1.000$$

$$\hat{\tau} = \begin{bmatrix} .4469 & .7687 \\ .4815 & -.4714 \\ .0735 & -.4032 \\ -.3262 & -.0438 \\ -.6757 & .1497 \end{bmatrix} \qquad \hat{\gamma} = \begin{bmatrix} .7232 & .2760 \\ .1812 & -.2360 \\ -.3194 & -.6785 \\ -.5850 & .6385 \end{bmatrix}.$$

The first two multiplicative components are significant at the 1% and 5% level, respectively. The third multiplicative component is not tested, since its inclusion would describe the data completely. An examination of the significant components of $\hat{\tau}$ reveals a mild linear interaction effect in the first component, but no trend is uncovered in the second component. A similar pattern exists for $\hat{\gamma}$, however the interaction effect in the first column is slightly stronger. Since the second component is significant but does not have a meaningful interpretation, then the first two components of $\hat{\tau}$ and $\hat{\gamma}$ are rotated. In practice the location of the outlier is unknown, so it is necessary to independently rotate each row of τ and each column of γ. Rotating the left singular vector five times produces:

$$
\begin{array}{ccccc}
\textbf{Row 1} & \textbf{Row 2} & \textbf{Row 3} & \textbf{Row 4} & \textbf{Row 5} \\
\begin{bmatrix} \mathbf{-.8892} & 0 \\ .1655 & .6532 \\ .3116 & .2662 \\ .2018 & -.2600 \\ .2102 & -.6594 \end{bmatrix} &
\begin{bmatrix} -.2184 & .8619 \\ .6738 & 0 \\ .3346 & -.2367 \\ -.2025 & -.2595 \\ -.5876 & -.3657 \end{bmatrix} &
\begin{bmatrix} -.6760 & .5776 \\ .5502 & .3891 \\ .4099 & 0 \\ -.0155 & -.3288 \\ -.2685 & -.6379 \end{bmatrix} &
\begin{bmatrix} .5452 & .7024 \\ .4145 & -.5313 \\ .0193 & -.4094 \\ -.3291 & 0 \\ -.6498 & .2383 \end{bmatrix} &
\begin{bmatrix} .2700 & .8472 \\ .5721 & -.3561 \\ .1590 & -.3777 \\ -.3090 & -.1133 \\ -.6921 & 0 \end{bmatrix}.
\end{array}
$$

To interpret these results I look for the largest weight, **r**, generated by the Givens rotation. The weight, **r**, can be easily recognized since it is in the same row as 0. The largest weight, -.8892, is observed when $\hat{\tau}$ is rotated for Row 1. The first component reveals the presence of an outlier in row 1. When the first component of a rotated singular vector displays an interpretable source of nonadditivity, then the second component of this vector should be examined to uncover the nature of the remaining nonadditivity. In this example, the second component displays a strong linear interaction. Although the linear interaction could be observed in the unrotated $\hat{\tau}$, the outlier effect was not visible.

Rotating the right singular vector produces:

$$
\begin{array}{cccc}
\textbf{Column 1} & \textbf{Column 2} & \textbf{Column 3} & \textbf{Column 4} \\
\begin{bmatrix} .7741 & 0 \\ .0851 & -.2851 \\ -.5403 & -.5201 \\ -.3189 & .8051 \end{bmatrix} &
\begin{bmatrix} .2214 & .7417 \\ .2975 & 0 \\ .3438 & -.6665 \\ -.8627 & -.0752 \end{bmatrix} &
\begin{bmatrix} -.5577 & .5368 \\ .1364 & .2644 \\ .7499 & 0 \\ -.3286 & -.8012 \end{bmatrix} &
\begin{bmatrix} .2850 & .7197 \\ .2964 & -.0258 \\ .2846 & -.6938 \\ \mathbf{-.8660} & 0 \end{bmatrix}.
\end{array}
$$

Examining the outcome of using a Givens rotation for Column 4 reveals that an outlier is present in column 4 of component 1 and that a linear interaction effect is present in the second component. Again, both of these effects were not observed initially in the original data or $\hat{\gamma}$. Notice that a large value, -.8627, is present in the rotation for Column 2. However, this is not the zeroed row and is not to be used to select a rotation.

Now, I use Stefansky's single outlier test on the same two-way table, Y. Since the test statistic is based upon the maximum normed residual, I need to locate the largest residual. The observed residuals from the additive part of the model, Z, are:

Z =

15.77	1.42	−11.64	−5.55
13.22	4.59	−2.79	−15.02
0.93	1.75	1.92	−4.60
−10.18	−1.90	4.42	7.65
−19.74	−5.86	8.09	17.51

Examining the table of residuals reveals that the largest residual resides in cell (5,1), not in cell (1,4). The test does not even identify the correct location of the outlier, therefore I do not calculate the test statistic. This is typical of outlier tests when the outlier is buried in the interaction effect. Although the residual in cell (1,4) appears small relative to the three residuals in the remaining corners of the table, there is no procedure to test residuals other that the largest residual.

The second example uses this procedure to detect two outliers in the presence of significant interaction. The results from the rotations will appear slightly different from the first example, so they need to be discussed. This example represents the methodology used in the simulation of section 3.6.3.

EXAMPLE 6-

Let there be two outliers, $\theta_1 = 12$ at (1,1) and $\theta_2 = 12$ at (2,2), and linear interaction, $\lambda = 15$, in an 8 x 12 table with $\varepsilon_{ij} \sim N(0,1)$.

The decomposition of the residuals produces the following results:

$$\hat{\lambda}_1 = 13.4015 \qquad U_1 = .4601 \; *$$

$$\hat{\lambda}_2 = 10.7431 \qquad U_2 = .5476 \; *$$

$$\hat{\lambda}_3 = 8.0101 \qquad U_3 = .6730 \; *$$

$$\sum_{p=4}^{7} \hat{\lambda}_p = 31.1724$$

$$\hat{\tau} = \begin{bmatrix} .5183 & .5229 & .5038 \\ .0962 & -.8085 & .4449 \\ .3428 & -.0318 & -.2987 \\ .2124 & -.0069 & -.3171 \\ .0442 & -.0798 & -.5794 \\ -.2128 & .0643 & .0448 \\ -.4254 & .2130 & .1141 \\ -.5756 & .1269 & .0877 \end{bmatrix}.$$

The first three multiplicative components of the model are significant at $\alpha = .05$. I present only $\hat{\tau}$ to show how my procedure can detect three nonadditive effects. The same method would be used on $\hat{\gamma}$, however it is omitted to conserve space. The first component of $\hat{\tau}$ is a contrast between rows 1 - 4 verses the remaining rows. Except for the low weight in the second row, .0962, this contrast appears linear. An outlier may be present in row 2 of the second component, but I hesitate to make this statement since rows 3 - 5 have the same sign as row 2. The third component in $\hat{\tau}$ almost seems quadratic, except for the last weight, .0877, which is low. Since the first three components are significant, i.e., q = 3, $\hat{\tau}$ is rotated twice for each of the 8 rows. Rotating the left singular vector produces:

$$
\begin{array}{cccc}
\textbf{Row 1} & \textbf{Row 2} & \textbf{Row 3} & \textbf{Row 4} \\
\begin{bmatrix} -\mathbf{.8921} & 0 \\ .1667 & .2372 \\ -.0118 & .4127 \\ .0598 & .3036 \\ .3483 & .3029 \\ .0607 & -.2181 \\ .0579 & -.4813 \\ .2105 & -.5570 \end{bmatrix} &
\begin{bmatrix} -.1603 & .5378 \\ \mathbf{.9278} & 0 \\ -.0800 & .3530 \\ -.1240 & .2265 \\ -.2037 & .0657 \\ -.0566 & -.2081 \\ -.1750 & -.4095 \\ -.1282 & -.5654 \end{bmatrix} &
\begin{bmatrix} .0231 & -.7600 \\ -.1628 & -.3318 \\ .4558 & 0 \\ .3681 & .0977 \\ .4186 & .4105 \\ -.1939 & .1016 \\ -.4095 & .1781 \\ -.4992 & .3037 \end{bmatrix} &
\begin{bmatrix} .1397 & .7172 \\ .3014 & .3176 \\ -.4395 & .1183 \\ -.3818 & 0 \\ -.5074 & -.2865 \\ .1568 & -.1511 \\ .3353 & -.2875 \\ .3954 & -.4280 \end{bmatrix}
\end{array}
$$

$$
\begin{array}{cccc}
\textbf{Row 5} & \textbf{Row 6} & \textbf{Row 7} & \textbf{Row 8} \\
\begin{bmatrix} .5298 & .5598 \\ .3222 & .1208 \\ -.3253 & .3192 \\ -.3302 & .1880 \\ -.5865 & 0 \\ .0690 & -.2082 \\ .1737 & -.4135 \\ .1473 & -.5661 \end{bmatrix} &
\begin{bmatrix} .2386 & .8520 \\ .2315 & -.3503 \\ .3897 & -.0664 \\ .2639 & -.1022 \\ .1786 & -.3572 \\ -.2268 & 0 \\ -.4821 & .0782 \\ -.5935 & -.0541 \end{bmatrix} &
\begin{bmatrix} .1056 & .8636 \\ .3319 & -.3896 \\ .3815 & .0223 \\ .2616 & -.0306 \\ .2083 & -.2772 \\ -.2235 & -.0374 \\ -.4892 & 0 \\ -.5762 & -.1510 \end{bmatrix} &
\begin{bmatrix} .3151 & .8263 \\ .1996 & -.3731 \\ .3818 & -.1006 \\ .2533 & -.1247 \\ .1450 & -.3702 \\ -.2258 & .0206 \\ -.4730 & .1218 \\ -.5959 & 0 \end{bmatrix}
\end{array}
$$

The largest weights, .9278 and -.8921, are obtained when $\hat{\tau}$ is rotated with the appropriate rotation for Row 2 and Row 1, respectively. The next largest weight in Row 8, -.5959, is not close to the two largest weights, so only rotations for Row 2 and Row 1 are studied in this example. Examining the results of the rotation for Row 2 reveals that an outlier is located in row 2 of the first component and linear interaction is observed in the second component. Studying the contrasts when Row 1 is rotated reveals that a second outlier is present in the first row, the second component appears as a linear contrast, except for the low weight, .2372, that seems out of place in row 2. This contrast appears redundant to the information in the second component of Row 2. Since this procedure successfully detected all three nonadditive effects, the power of this procedure is evaluated in section 3.6.3 under similar conditions.

The two examples in this section illustrate how to use the recommended procedure. I now study the power of this method to distinguish single and multiple outliers from interaction in the next section.

3.6 MONTE CARLO SIMULATION RESULTS

In this section, I illustrate the capability of my procedure to detect different types of nonadditivity by several Monte Carlo simulations. Being able to differentiate an outlier from interaction is novel, so I have studied the performance of my procedure in detail under various conditions. I have investigated my procedure's ability to distinguish one outlier, as well as two outliers, in the presence of interaction. I first study two nonadditive effects: one outlier and interaction (3.6.1) and two outliers (3.6.2), then I extend my procedure to three nonadditive factors: two outliers and interaction (3.6.3). In all cases I assume that the interaction is rank 1. Finally, I compare my method to the current procedures for detecting a single outlier (Stefansky, 1972) and multiple outliers (John and Draper, 1978). These tests assume that the data in a two-way table are additive. The performance of these tests in the presence of interaction is unknown, so the tests are evaluated under the condition of interaction.

3.6.1 One Outlier and Linear Interaction of Rank 1

A Monte Carlo simulation was devised to count the number of times my procedure correctly identified the outlier, the interaction or both under various combinations of θ and λ in a 5 x 5 table. θ represented an outlier effect, whereas λ denoted an interaction effect. The outlier effect was measured in units of standard deviations. The magnitudes of θ and λ ranged from 0 to 50 σ, and the interaction had the following linear form:

$$\tau' = [\ -2\ -1\ \ 0\ \ 1\ \ 2\ ] \qquad \gamma' = [\ -2\ -1\ \ 0\ \ 1\ \ 2\ ]$$

where:

τ and γ were normalized, so that $\tau'\tau = \gamma'\gamma = 1$.

The data were generated using a N(0,1) distribution and the additive-plus-multiplicative model of rank 1. The quantity θ was added to one of three cells designated to be an outlier. This allowed me to study the effect of the outlier's position in respect to the interaction. One location, the (3,3) cell, was orthogonal to the interaction, whereas the (1,5) cell was embedded in the interaction and the (2,4) cell was intermediate. An outlier was orthogonal to the interaction if the vector representation of the outlier was orthogonal to both τ and γ. An outlier was hiddened in the interaction if the inner product of these same vectors was not zero. The simulation was run for a total of 1,000 times for each combination of θ, λ and outlier location. Finally, the power of this procedure was compared to Stefansky's test under identical conditions.

I would like to explain how I simulated these results, then present the findings of this simulation. The simulation was designed to test H_o: $\lambda_i = 0$ vs. H_1: $\lambda_i \neq 0$ in a stepwise fashion using the likelihood ratio test statistic presented in section 3.5. If $\lambda_1 \neq 0$ and $\lambda_2 = 0$, then the procedure studied the first singular vector of $\hat{\tau}$ and $\hat{\gamma}$ to determine the nature of the significance: θ or λ. It compared both singular vectors to the appropriate vector representation of each nonadditive effect, then selected the effect which generated the minimum sum of squared deviations for both $\hat{\tau}$ and $\hat{\gamma}$. If a match did not occur, then the procedure counted a failure.

If $\lambda_2 \neq 0$, then the probable cause for significance was θ and λ. The procedure tested this speculation by independently rotating the first two singular vectors of $\hat{\tau}$ and $\hat{\gamma}$, and locating the maximum weight, i. e., r, for both pairs of rotated vectors. If the locations of these two largest weights corresponded to the row and column location of θ, then the source of significance was identified.

Occasionally λ_2 was significant, but λ_1 was nonsignificant. This occurred when two strong nonadditive effects were present in the table. Usually, this produced two large singular values, $\hat{\lambda}_1$ and $\hat{\lambda}_2$, close in value. When this happened it was difficult for U_1 to be significant due to the presence of $\hat{\lambda}_2{}^2$ in the denominator. However, the procedure recommended by Milliken and Johnson (1989) for sequentially testing λ_i allowed one to continue testing as long as two consecutive nonsignificant test results were not obtained.

Figures 2-4 display the power of my procedure to detect an outlier, interaction or both outlier and interaction, respectively, under various degrees of θ and λ when the outlier is orthogonal to the interaction effect. When no interaction has been introduced into the two-way table, i.e., $\lambda = 0$, the power to correctly identify the source of nonadditivity as a single outlier is high, Figure 2. If $\lambda = 10$ and $\theta \geq 40\sigma$, then the procedure detects just the outlier in approximately 10% of the runs when the outlier is orthogonal to the interaction, i.e., in the (3,3) cell. Once $\lambda \geq 20$, the power of uncovering only a single source of nonadditivity is zero. When interaction, but no outlier effect, exists, then the test detects only interaction about 95% of the time, Figure 3. As the outlier effect increases, the procedure is less likely to conclude that the nonadditivity is due to a single effect. Once $\theta \geq 20\sigma$, the test does not have any ability to detect only interaction. This brings me to evaluating the test's power to identify both sources of nonadditivity, Figure 4. When $\lambda = 10$, the procedure correctly detects both sources about 85% to 90% of the time. Once $\lambda \geq 20$, the power of determining the existence of both outlier and interaction effects is even better. The procedure can identify both sources of nonadditivity when both effects are strong. It can differentiate an outlier from interaction even when the interaction effect, λ, is more than twice the value of the outlier effect, θ.

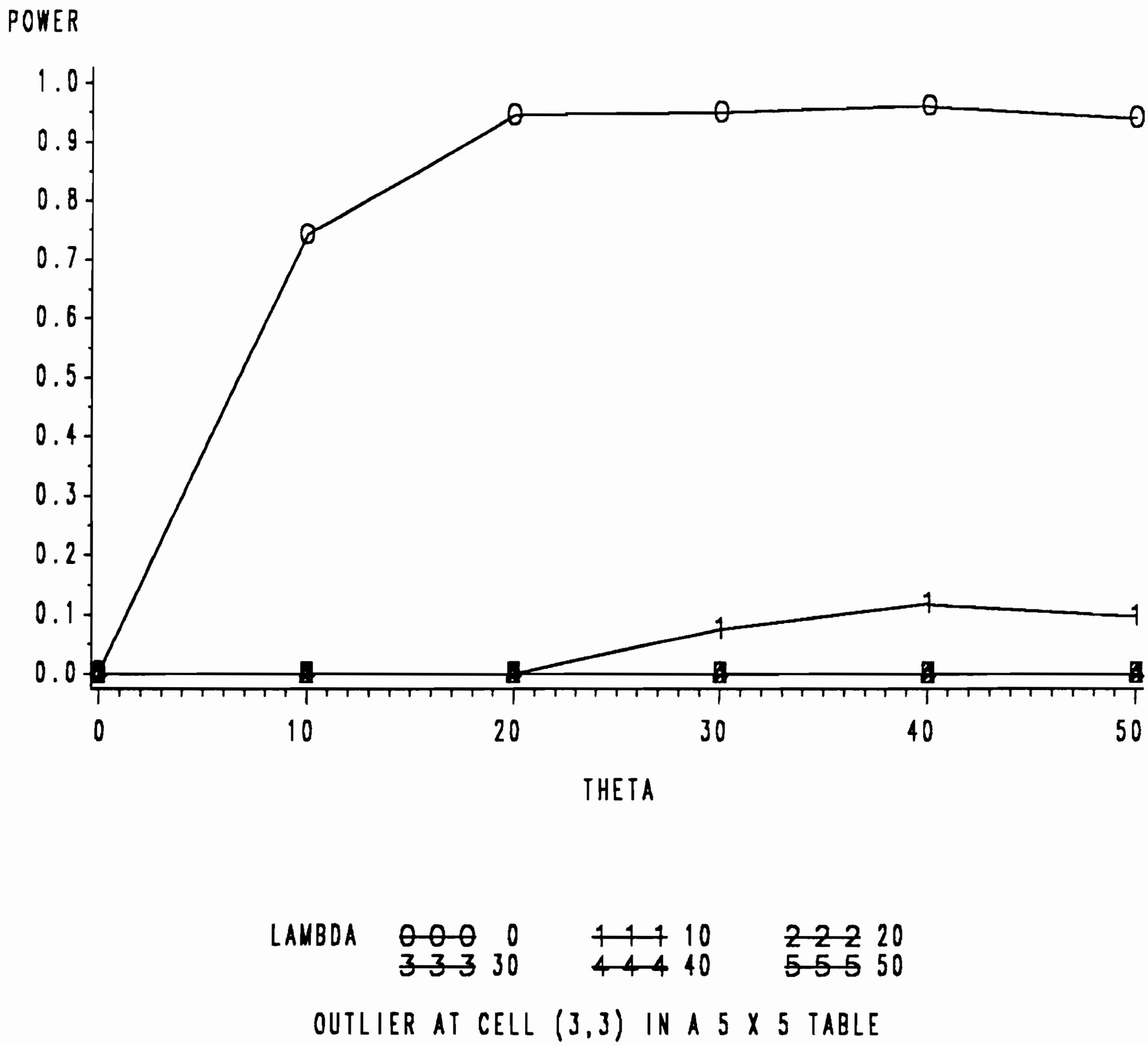

Figure 2. The power of my procedure to detect an outlier, cell (3,3) in the presence of interaction.

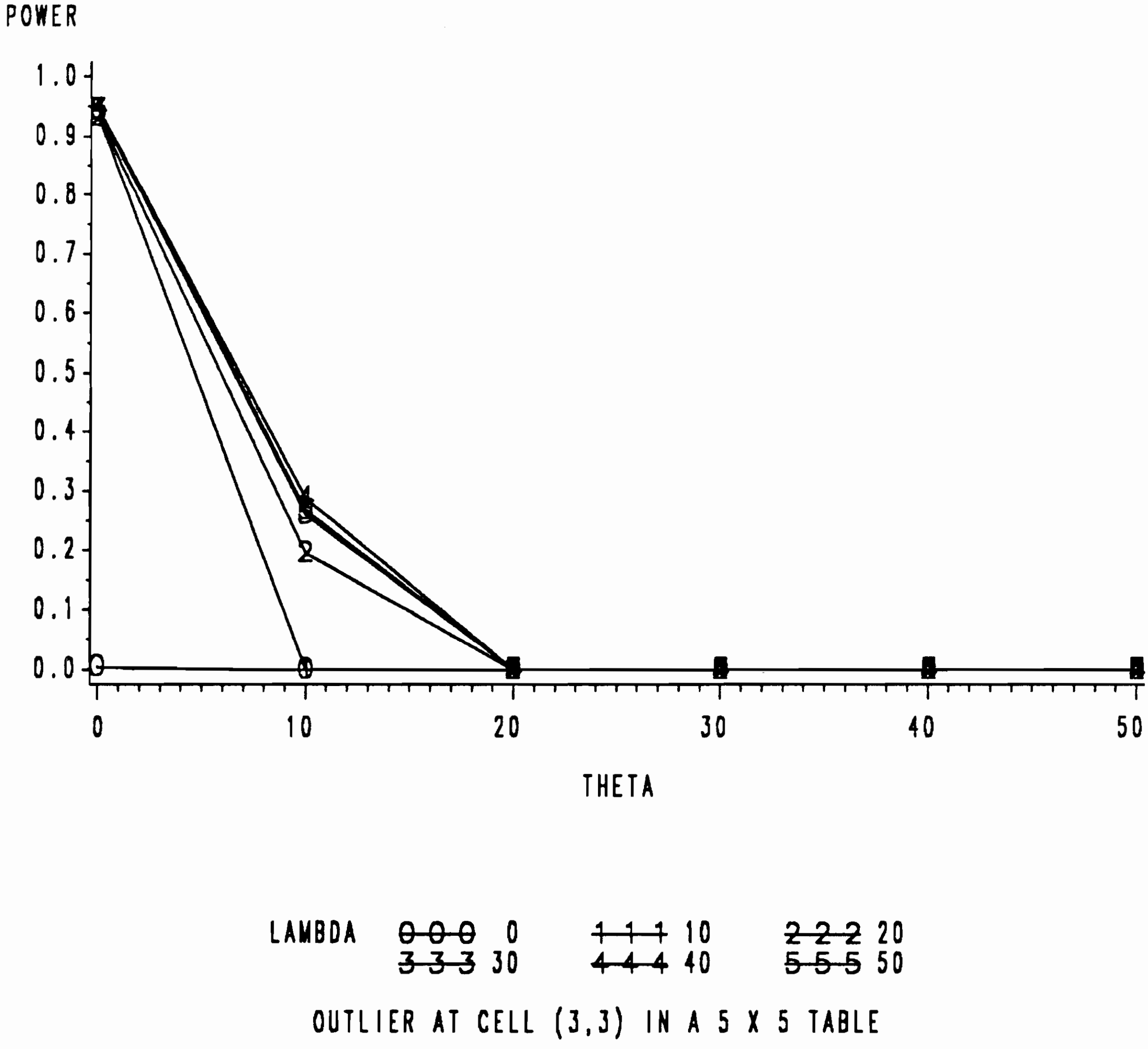

Figure 3. The power of my procedure to detect interaction in the presence of an outlier, cell (3,3).

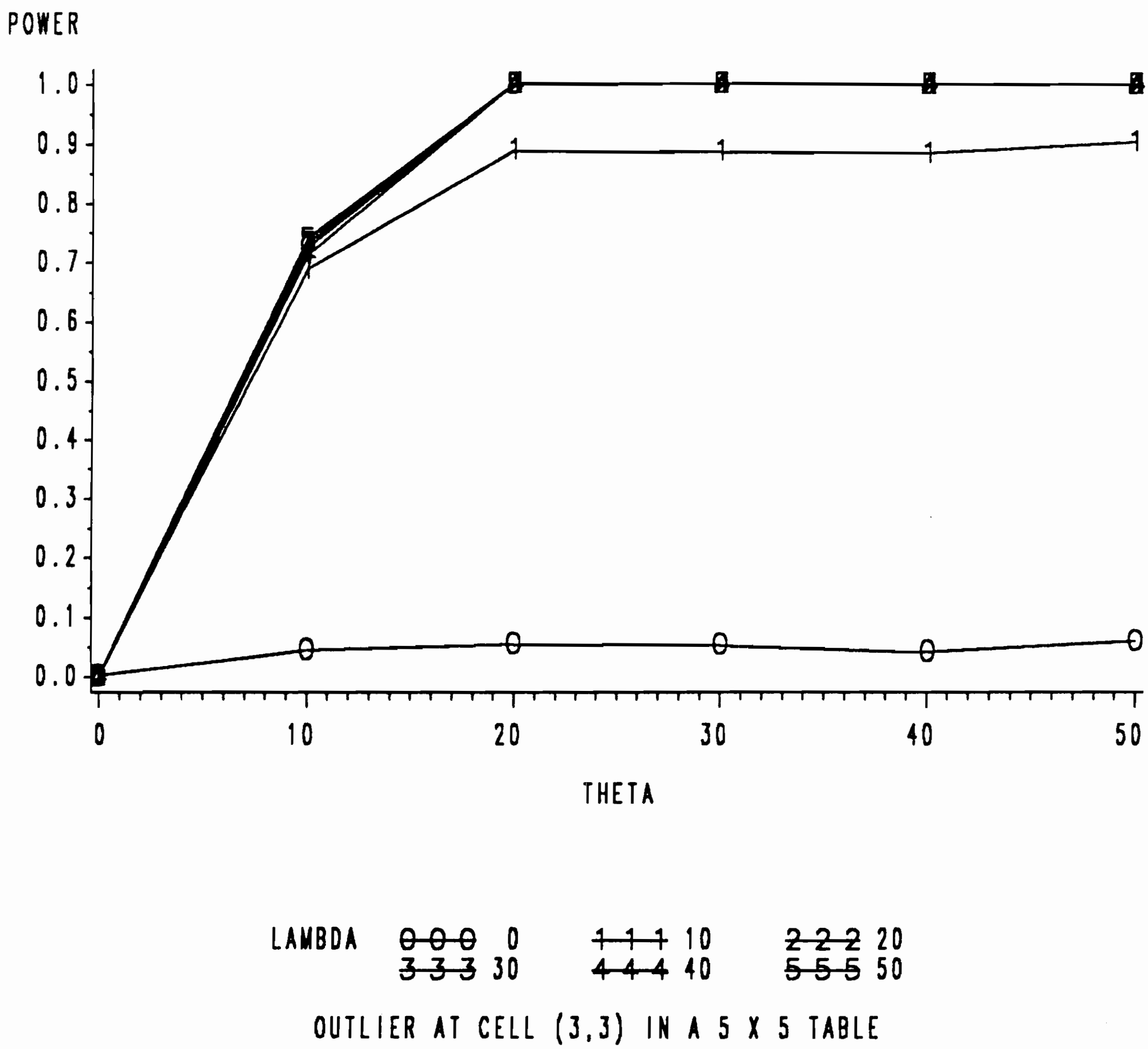

Figure 4. The power of my procedure to detect both outlier, cell (3,3), and interaction.

When interaction is present, the test's power may depend upon the outlier's location in a two-way table. Since orthogonality between the outlier and interaction is not usually the case because very few cell locations are orthogonal to the interaction, I now consider an outlier whose location is embedded in the interaction at cell (1,5), Figures 5-7. This outlier may not stand out when the observed response or the residuals are examined. In Figure 5, the power curve for detecting just the outlier in the presence of interaction is of similar shape as when the outlier is orthogonal to the interaction, i.e., cell (3,3), Figure 2. However, the procedure is more powerful in detecting an outlier at the (1,5) location than in the previous case. This same pattern exists when detecting interaction in the presence of an outlier at cell (1,5), Figure 6. The power curve is of similar form as Figure 3, yet is is more powerful. It appears that the test for a single outlier is stronger when the outlier is in the path of the interaction and vice versa. The reason this occurs is that the outlier actually contributes to the interaction effect, and increases the number of times U_1 is significant. This leads to more vectors being examined, and consequently more samples being tabulated as either a single outlier or interaction. This could be the same reason that the power of detecting a single outlier was higher that expected. For $\lambda \geq 20$, the test's power to identify both sources of nonadditivity is slightly less when the outlier's position is at (1,5), Figure 7, instead of at cell (3,3). However, the power strongly decreases when $\lambda = 10$ and the outlier is in cell (1,5).

The outlier at cell (2,4) is intermediate in location between the orthogonal outlier (3,3) and the embedded outlier (1,5). The power curves for detecting a single outlier (Figure 8), interaction (Figure 9) and both effects (Figure 10) are very similar in shape and power to the respective power curves of the orthogonal outlier. It may be that the test's power is not strongly affected unless the outlier is buried in the interaction.

How reasonable is it to find effects as large as these in real applications so that this procedure can be successfully used? Table size has a lot to do with detecting significant sources of nonadditivity. The bigger the table the easier it is to identify weaker effects (this is demonstrated in 3.6.2). For a closely related table of size 5 x 4, Milliken and Johnson (1989) present two examples of fitting the additive-plus-multiplicative model. One example (p. 4) deals with the height of sorghum plants and

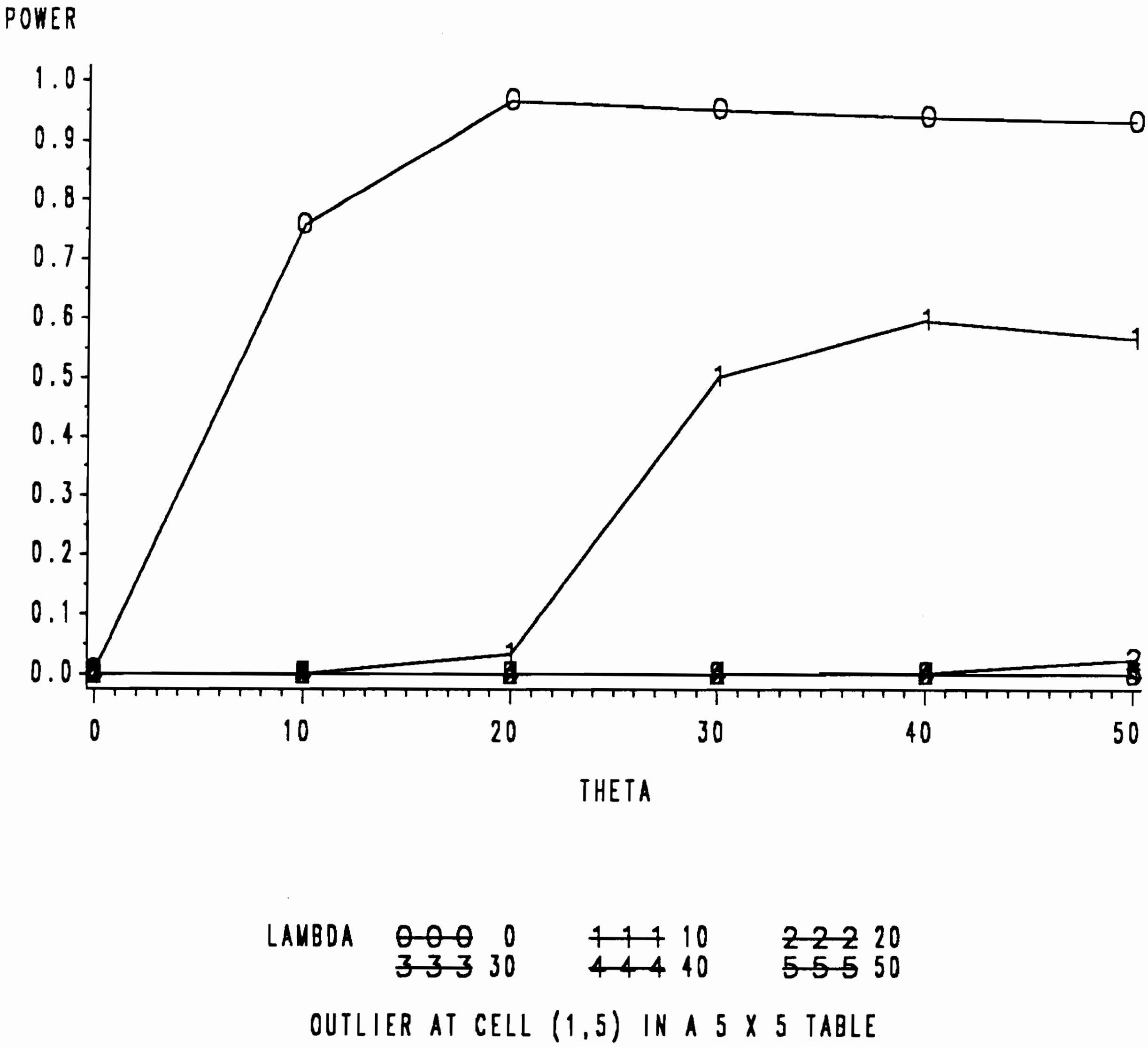

Figure 5. The power of my procedure to detect an outlier, cell (1,5), in the presence of interaction.

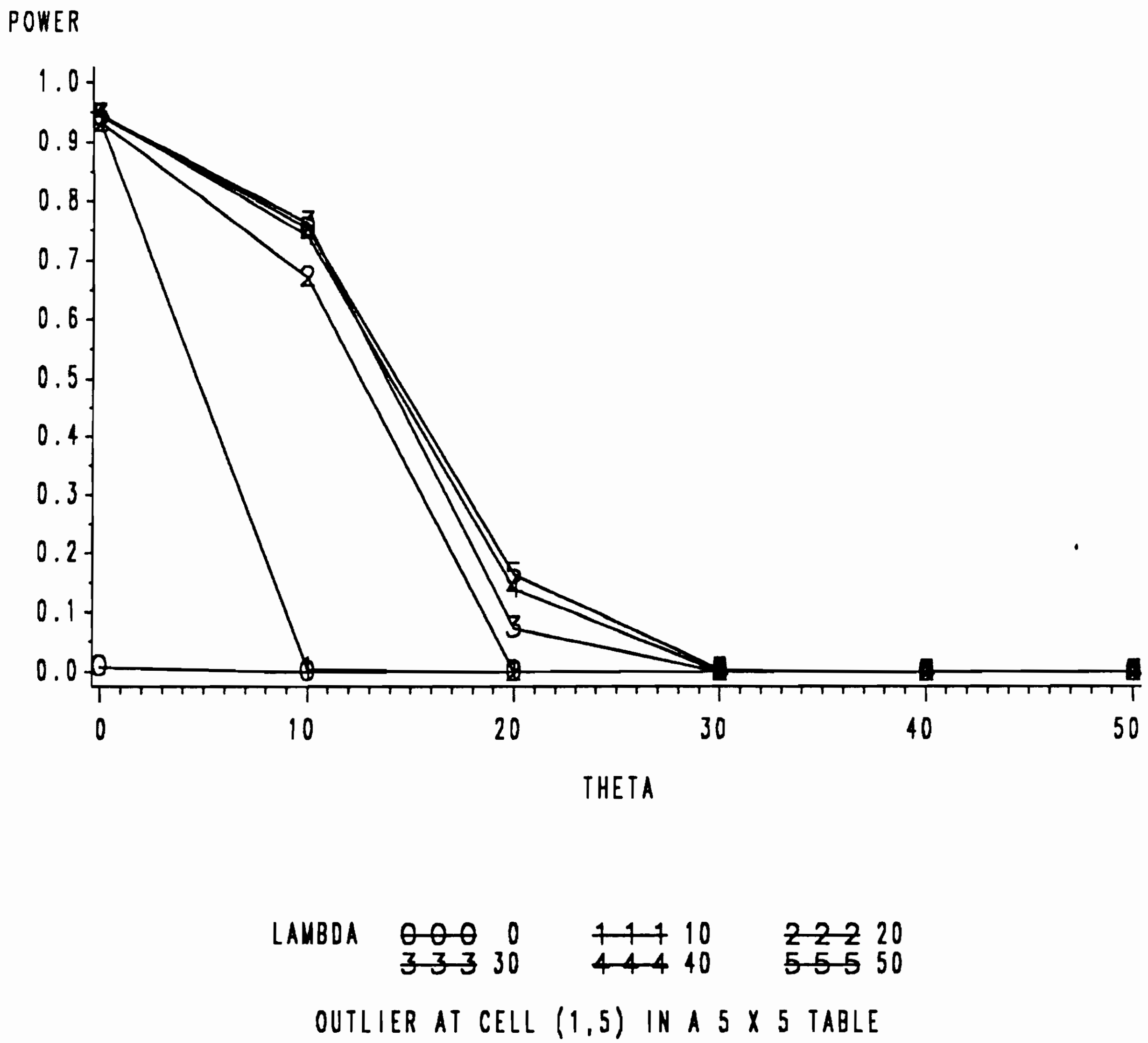

Figure 6. The power of my procedure to detect interaction in the presence of an outlier, cell (1,5).

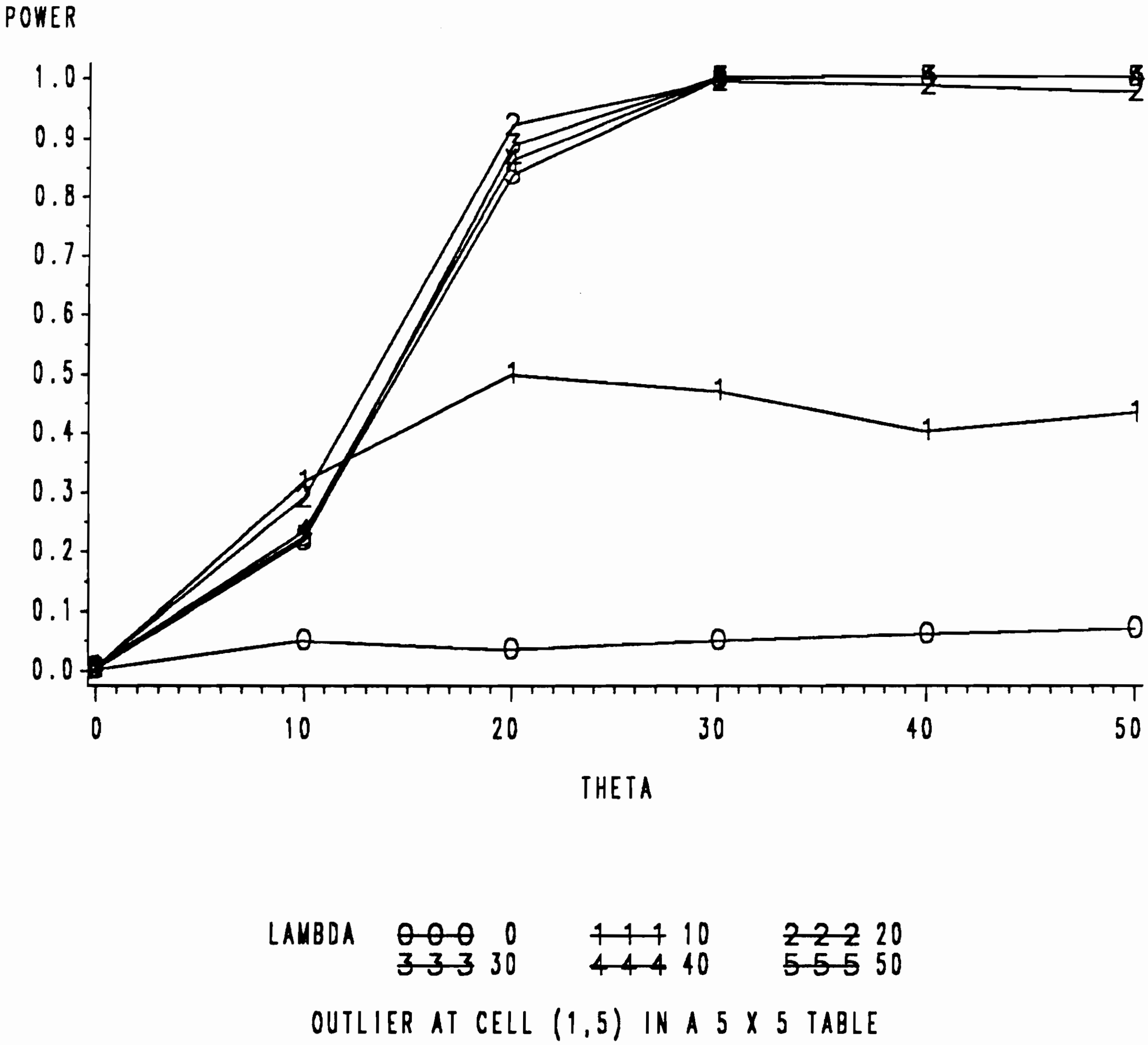

Figure 7. The power of my procedure to detect both outlier, cell (1,5) and interaction.

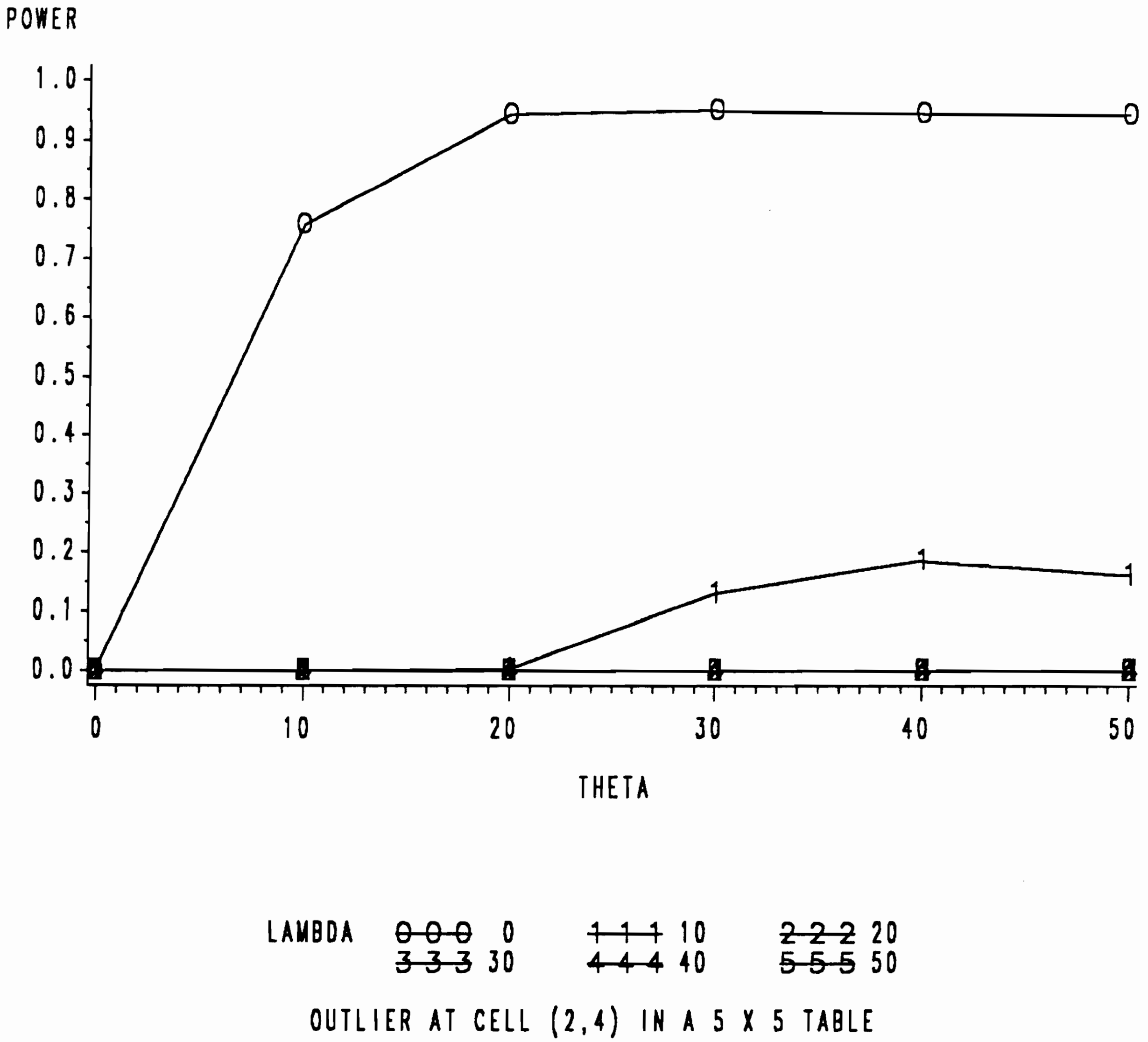

Figure 8. The power of my procedure to detect an outlier, cell (2,4), in the presence of interaction.

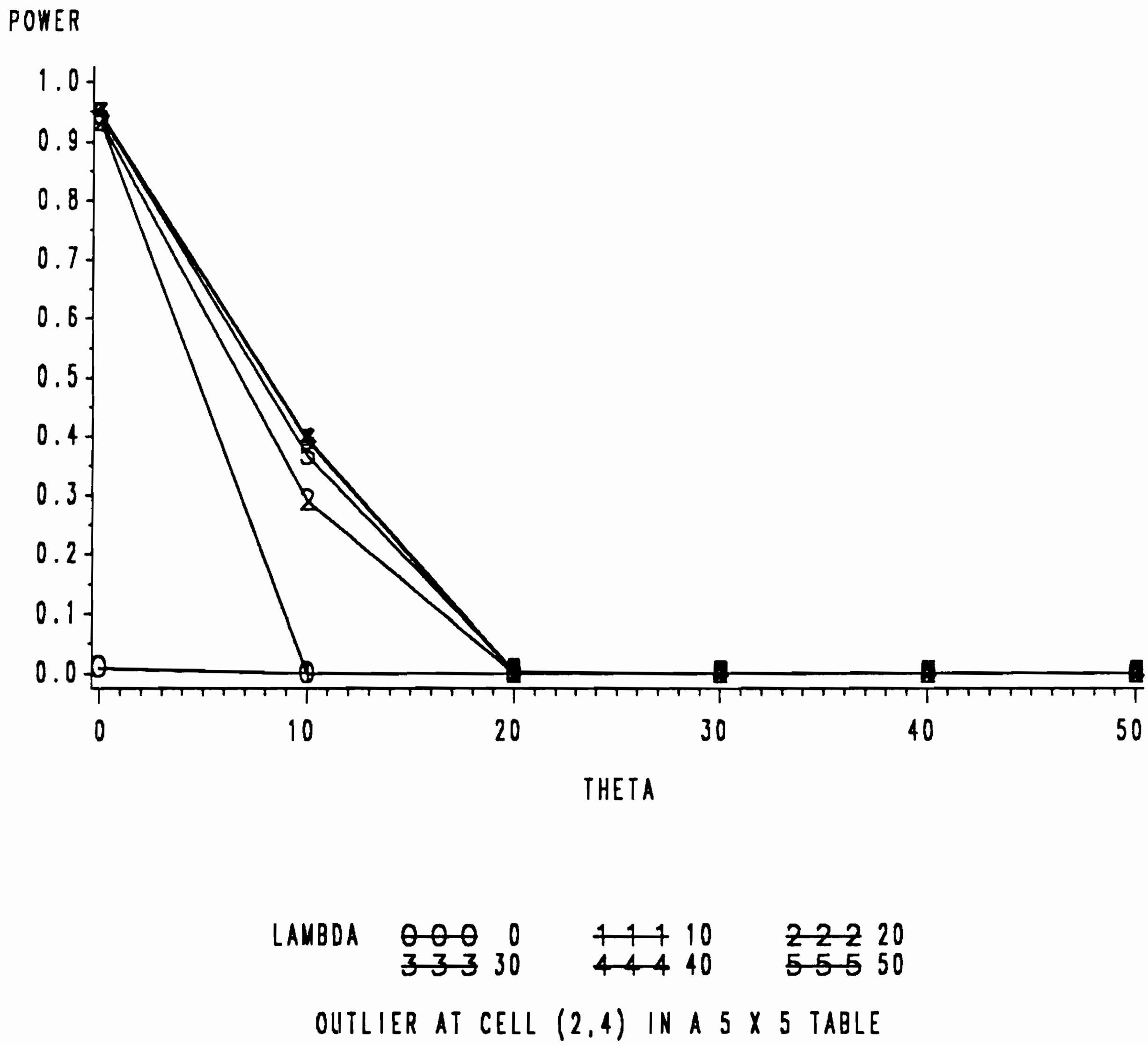

Figure 9. The power of my procedure to detect interaction in the presence of an outlier, cell(2,4).

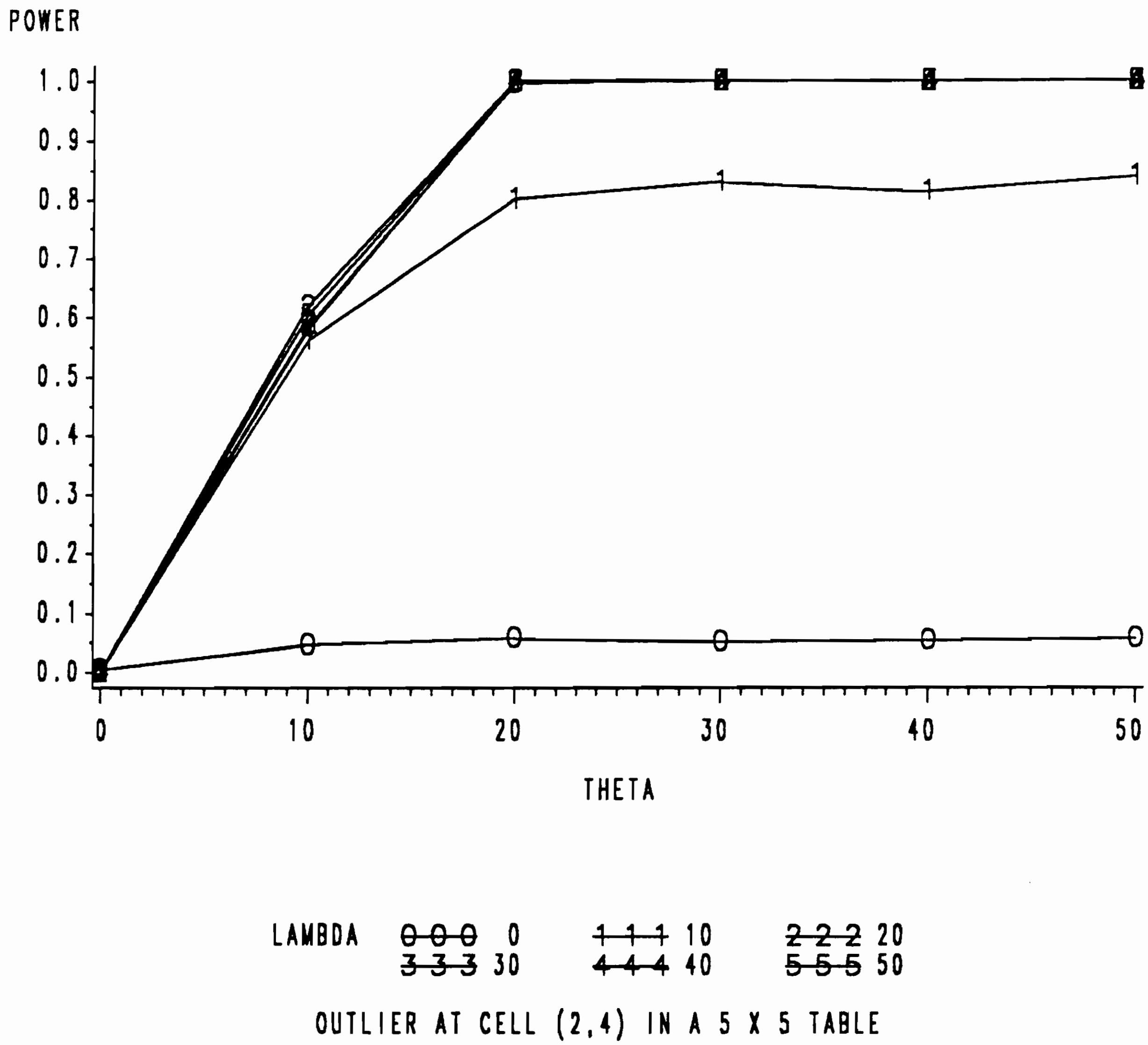

Figure 10. The power of my procedure to detect both outlier, cell (2,4) and interaction.

the second example (p. 54) is taken from Davies (1954, p. 305). Although the examples do not contain outliers, they both have significant multiplicative interaction. The value of $\hat{\lambda}_1$ is approximately 19 in both cases. If a very strong outlier were to appear, it is reasonable to expect that the procedure could identify both sources of nonadditivity.

Lastly, I compare the power of Stefansky's test for detecting a single outlier under the same set of conditions as my procedure. In Appendix B, I show that Stefansky's test is equivalent in expectation to Johnson and Graybill's (1972) U_1 test statistic that used in my procedure. The power curves of Stefansky's test are displayed for an outlier in cell (3,3) (Figure 11), the (1,5) cell (Figure 12) and cell (2,4) (Figure 13). When no interaction is present, Stefansky's test is more powerful in detecting a single outlier than my technique. If interaction exists, Stefansky's test can not identify an outlier until θ exceeds λ. And then it reveals only the outlier not the interaction. As expected, the power of Stefansky's test is greater for an orthogonal outlier than an embedded outlier. In conclusion, it appears for a 5 x 5 table that Stefansky's test performs better in cases of no or little interaction, i.e., $\lambda < 20$, whereas my procedure is more powerful in the presence of stronger interaction, i.e., $\lambda \geq 20$.

As mentioned earlier, table size affects the power of detection. The bigger the table, the easier it is to detect an outlier for a given combination of λ and θ. For tables larger than 5 x 5, it is not necessary to have $\lambda \geq 20$ and $\theta \geq 20\sigma$ before my procedure has strong power. To demonstrate this effect, I evaluate the power of my procedure and Stefansky's test for a 8 x 10 table using a Monte Carlo simulation. I selected this size of table to ensure that critical values existed for both procedures. I study two outlier locations. One position, the (4,6) cell, is nearly orthogonal to the interaction, whereas the (1,1) location is once again embedded in the interaction. Both λ and θ ranged from 0 to 14 in increments of 2. The results of these simulations are presented in Tables 3-6.

These findings confirm the belief that weaker outlier effects can be detected in a larger table. The power of my procedure to locate an outlier in the presence of interaction is quite good when the

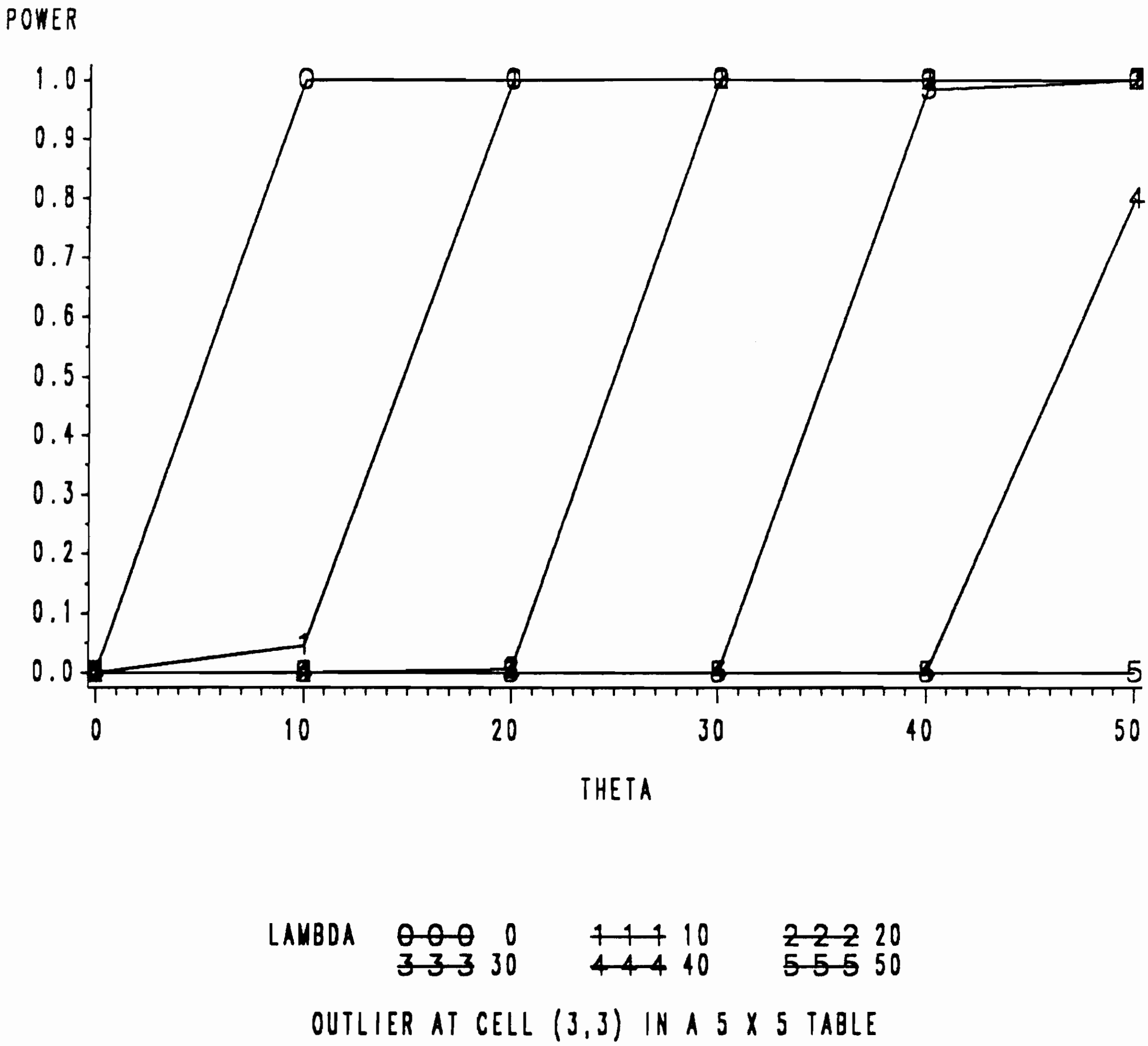

Figure 11. The power of Stefansky's test to detect an outlier, cell (3,3), orthogonal to the interaction.

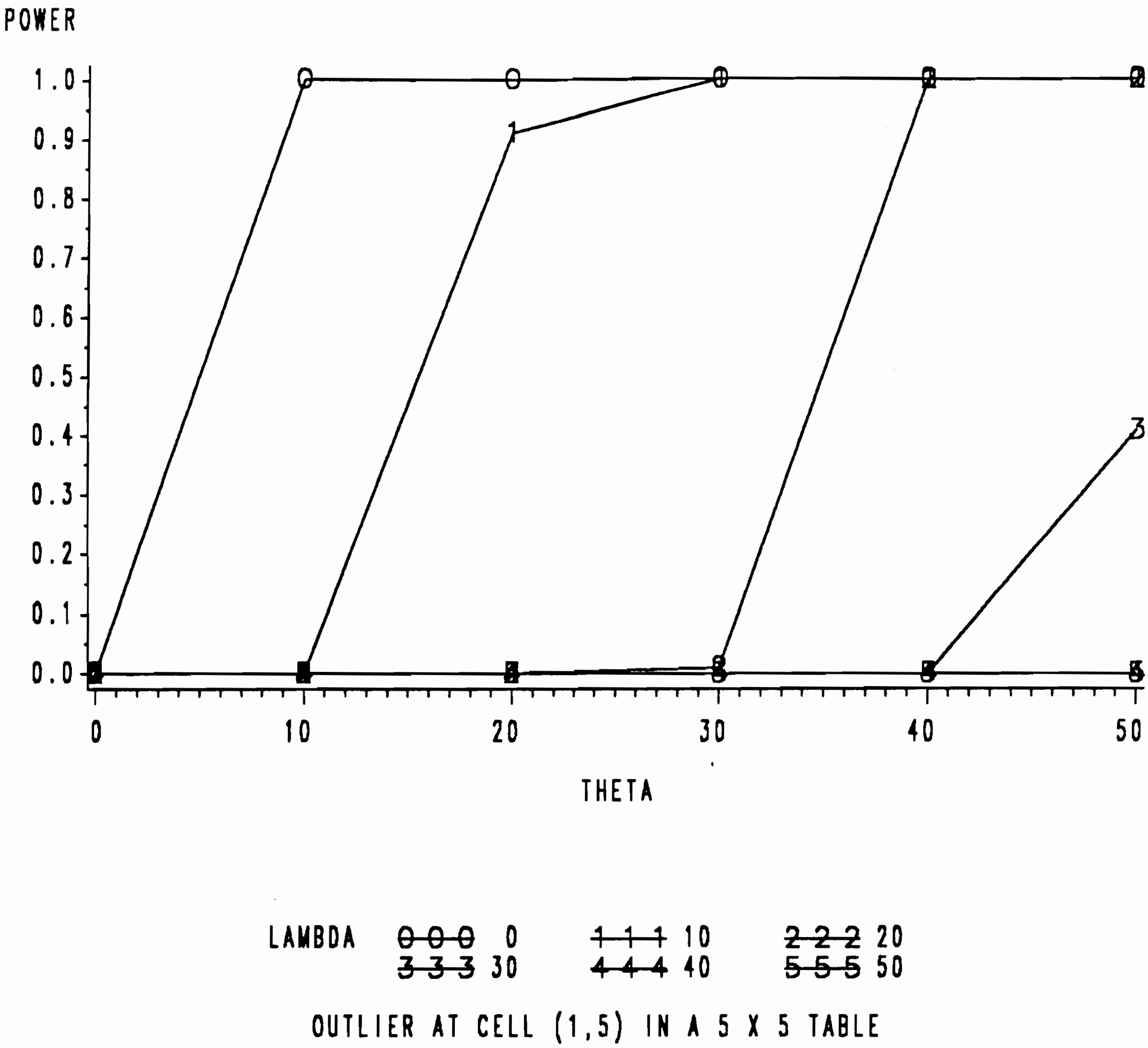

Figure 12. The power of Stefansky's test to detect an outlier, cell (1,5), fully embedded in the interaction.

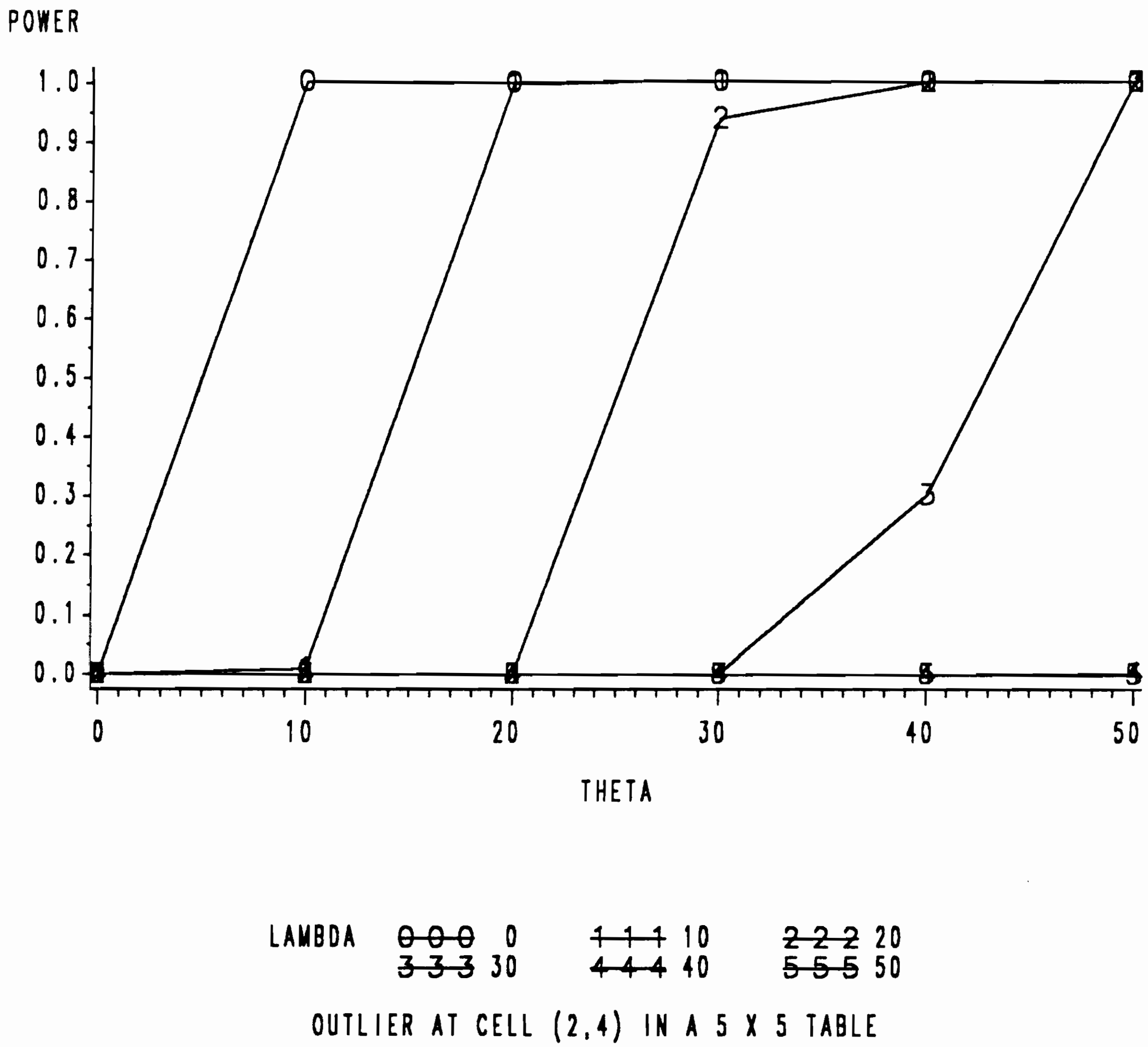

Figure 13. The power of Stefansky's test to detect an outlier, cell (2,4), slightly embedded in the interaction.

(4,6) cell is perturbed by $\theta_1 \geq 8\sigma$, Table 3 (highlighted). Interaction can be identified in the presence of an outlier when $\lambda > 6$ and $\theta_1 \leq 4\sigma$. The procedure performs well in locating an outlier when the (1,1) cell is disturbed by $\theta_1 \geq 10\sigma$, Table 4. Interaction can also be detected when $\lambda > 6$ and $\theta_1 \leq 6\sigma$. In comparison, the power of Stefansky's test is high when $\theta_1 \geq 6\sigma$ and θ_1 is close to or exceeds λ for both the (4,6) position, Table 5, and the (1,1) cell, Table 6. Stefansky's test does not have the capability to detect interaction.

These results also support a major difference between the two procedures. The power of my procedure for a large table is not as highly dependent upon the value of λ as it is for a 5 x 5 table. In either case, my procedure performs well in the presence of interaction. However for Stefansky's test to be effective, θ must be equal to or exceed λ. This same relationship with λ was observed in the case of a 5 x 5 table. This outcome suggests that these two procedures can be used to complement each other. Since Stefansky's test is more powerful than my procedure in the presence of little or no interaction, it should be used first to test for the presence of an outlier. If Stefansky's test fails to uncover an outlier, then my procedure should be employed when interaction is suspected.

3.6.2 Two Outliers

The previous section, 3.6.1, demonstrated the effectiveness of my procedure to distinguish the makeup of a nonadditive term which was composed of an outlier and rank 1 interaction. In this section, I study the power of my procedure to correctly identify a different composition of two nonadditive factors: two outliers. Once again, I investigate my test's effectiveness by means of a Monte Carlo simulation. The outliers are placed in cells (1,1) and (2,2) of an 8 x 12 table, and have magnitudes that range from 0 to 12. These cell locations and outlier effects were selected so that I could compare my results to the published findings of John and Draper's multiple outlier test (1978).

Table 3. Power of my procedure detect interaction and an outlier in cell (4,6) when both are present in an 8 x 10 table.

	θ_1							
λ	0	2	4	6	8	10	12	14
0	.023 .017	.011 .043	.001 .196	.000 .607	.000 **.939**	.000 **.998**	.000 **1.000**	.000 **1.000**
2	.058 .003	.037 .025	.013 .150	.004 .527	.000 **.907**	.000 **.997**	.000 **1.000**	.000 **1.000**
4	.275 .000	.200 .010	.115 .072	.030 .395	.003 **.785**	.000 **.984**	.000 **1.000**	.000 **1.000**
6	**.734** .000	.673 .002	.483 .064	.228 .322	.032 **.734**	.000 **.959**	.000 **.998**	.000 **1.000**
8	**.961** .000	**.952** .005	**.859** .064	.527 .364	.140 **.754**	.020 **.950**	.001 **.996**	.000 **1.000**
10	**.983** .000	**.976** .003	**.907** .073	.588 .393	.196 **.793**	.027 **.966**	.000 **.999**	.000 **1.000**
12	**.976** .000	**.976** .004	**.891** .083	.637 .352	.174 **.824**	.026 **.972**	.001 **.999**	.000 **1.000**
14	**.984** .000	**.972** .005	**.899** .085	.611 .381	.188 **.812**	.022 **.978**	.001 **.999**	.000 **1.000**

The top number is the power of the procedure to detect interaction.

The bottom number is the power to detect either an outlier or both outlier and interaction.

Table 4. Power of my procedure detect interaction and an outlier in cell (1,1) when both are present in an 8 x 10 table.

λ	θ_1: 0	2	4	6	8	10	12	14
0	.025	.014	.000	.001	.000	.000	.000	.000
	.017	.044	.209	.585	**.945**	**.996**	**1.000**	**1.000**
2	.055	.041	.016	.001	.000	.000	.000	.000
	.007	.019	.114	.452	**.847**	**.990**	**1.000**	**1.000**
4	.278	.170	.087	.034	.004	.000	.000	.000
	.004	.007	.040	.257	.692	**.954**	**.987**	**.999**
6	**.727**	.590	.385	.197	.038	.002	.000	.000
	.003	.001	.020	.140	.459	**.862**	**.987**	**.999**
8	**.960**	**.931**	**.847**	.583	.222	.038	.002	.000
	.002	.006	.025	.152	.435	**.772**	**.963**	**.999**
10	**.983**	**.985**	**.949**	**.813**	.462	.111	.012	.002
	.001	.004	.030	.152	.465	**.830**	**.962**	**.993**
12	**.976**	**.980**	**.956**	**.849**	.537	.154	.018	.001
	.001	.001	.025	.135	.457	**.838**	**.980**	**.996**
14	**.984**	**.980**	**.967**	**.852**	.531	.187	.034	.002
	.000	.003	.021	.140	.464	**.812**	**.966**	**.998**

The top number is the power of the procedure to detect interaction.

The bottom number is the power to detect either an outlier or both outlier and interaction.

Table 5. Power of Stefansky's test to detect an outlier in cell (4,6) when linear interaction is present in an 8 x 10 table.

	θ_1							
λ	0	2	4	6	8	10	12	14
0	.001	.033	.476	**.940**	**1.000**	**1.000**	**1.000**	**1.000**
2	.000	.035	.470	**.930**	**1.000**	**1.000**	**1.000**	**1.000**
4	.000	.015	.319	**.897**	**.999**	**1.000**	**1.000**	**1.000**
6	.000	.005	.174	**.763**	**.994**	**1.000**	**1.000**	**1.000**
8	.000	.001	.065	.586	**.969**	**1.000**	**1.000**	**1.000**
10	.000	.000	.024	.359	**.881**	**.998**	**1.000**	**1.000**
12	.000	.000	.004	.131	**.730**	**.978**	**1.000**	**1.000**
14	.000	.000	.001	.034	.452	**.925**	**1.000**	**1.000**

Table 6. Power of Stefansky's test to detect an outlier in cell (1,1) when linear interaction is present in an 8 x 10 table.

λ	θ_1 0	2	4	6	8	10	12	14
0	.000	.043	.484	**.950**	**.999**	**1.000**	**1.000**	**1.000**
2	.002	.008	.244	**.839**	**.997**	**1.000**	**1.000**	**1.000**
4	.005	.000	.049	.550	**.964**	**1.000**	**1.000**	**1.000**
6	.008	.000	.003	.183	**.790**	**.989**	**1.000**	**1.000**
8	.008	.000	.000	.022	.376	**.910**	**.997**	**1.000**
10	.012	.000	.000	.000	.069	.572	**.963**	**.999**
12	.004	.000	.000	.000	.001	.165	**.720**	**.985**
14	.007	.000	.000	.000	.001	.011	.257	**.861**

Table 7 displays the power of my procedure for detecting the outliers in an 8 x 12 table. For each combination of θ_1 and θ_2, the power is presented for correctly locating two outliers (top row) and one outlier (bottom row). My procedure performs well in the multiple outlier case when $\theta_1 \geq 8\sigma$ and $\theta_2 \geq 8\sigma$, and in the single outlier case when $\theta_1 \geq 7\sigma$ and $\theta_2 \geq 7\sigma$. However, John and Draper obtained similar power for detecting two outliers when $\theta_1 \geq 5\sigma$ and $\theta_2 \geq 5\sigma$, Table 8, and for identifying one outlier when $\theta_1 \geq 6\sigma$ or $\theta_2 \geq 6\sigma$. Again, it seems that tests designed to detect a single outlier (Stefansky) or multiple outliers (John and Draper) are more powerful than my procedure in the absence of interaction. I now compare how John and Draper's test performs in the presence of interaction.

3.6.3 Two Outliers and Linear Interaction of Rank 1

The previous sections, 3.6.1 and 3.6.2, demonstrated the capability of my procedure to identify two effects in the nonadditive term. In this section, I extend the application of my technique to include three nonadditive factors, i.e., two outliers and rank 1 interaction. I have two objectives for this section. First, to investigate the performance of John and Draper's test under the same conditions as their published results, but now in the presence of interaction. Since John and Draper's test for two outliers achieved high power when $\theta_1 \geq 5\sigma$ and $\theta_2 \geq 5\sigma$, I reexamine its performance for θ_1, θ_2 and λ ranging from 5 to 8. Second, to compare the power of both procedures under a wider set of values, i.e., θ_1 , θ_2 and λ from 6 to 15. In both simulations, the interaction has the following linear form:

$$\tau' = [\ -7\ -5\ -3\ -1\ \ 1\ \ 3\ \ 5\ \ 7\ ] \qquad \gamma' = [\ 11\ \ 9\ \ 7\ \ 5\ \ 3\ \ 1\ -1\ -3\ -5\ -7\ -9\ -11\ ]$$

where:

τ and γ are normalized, so that $\tau'\tau = \gamma'\gamma = 1$.

Table 7. Power of my procedure to detect two outliers in cells (1,1) and (2,2) of an 8 x 12 table.

	θ_1												
θ_2	0	1	2	3	4	5	6	7	8	9	10	11	12
12	.001	.000	.004	.018	.038	.117	.260	.472	**.710**	**.854**	**.958**	**.985**	.997
	.035	.045	.072	.100	.180	.321	.530	**.728**	**.906**	**.973**	**.997**	**.998**	**1.000**
11	.000	.000	.006	.011	.042	.129	.237	.486	**.714**	**.874**	**.962**	**.985**	
	.051	.040	.054	.086	.176	.331	.505	**.741**	**.894**	**.968**	**.996**	**.999**	
10	.001	.001	.000	.015	.043	.115	.257	.445	**.714**	**.855**	**.953**		
	.040	.045	.047	.097	.150	.325	.501	**.711**	**.906**	**.971**	**.994**		
9	.001	.002	.004	.018	.054	.124	.257	.494	**.724**	**.884**			
	.050	.047	.046	.090	.169	.324	.535	**.754**	**.900**	**.978**			
8	.001	.002	.008	.014	.052	.140	.270	.492	.698				
	.046	.039	.051	.108	.161	.340	.510	**.749**	**.877**				
7	.001	.004	.007	.019	.049	.102	.270	.466					
	.041	.054	.057	.100	.161	.320	.530	**.728**					
6	.002	.002	.002	.017	.054	.114	.247						
	.050	.046	.041	.094	.188	.314	.504						
5	.002	.000	.003	.016	.037	.115							
	.053	.050	.059	.098	.163	.300							
4	.001	.002	.005	.016	.038								
	.037	.040	.058	.086	.153								
3	.000	.001	.004	.018									
	.049	.045	.062	.098									
2	.001	.005	.006										
	.047	.062	.043										
1	.000	.000											
	.044	.052											
0	.001												
	.050												

The top number is the power to detect two outliers.

The bottom number is the power to detect either outlier.

Table 8. Power of John and Draper's test (1978; p.77) to detect two outliers in cells (1,1) and (2,2) of an 8 x 12 table.

	θ_1								
θ_2	0	1	2	3	4	5	6	7	8
8	.000	.005	.037	.209	.506	.817	.961	.997	.999
	.986	.989	.996	.999	.998	1.000	1.000	1.000	1.000
7	.001	.003	.039	.192	.487	.792	.947	.990	
	.925	.938	.951	.972	.990	.997	.998	1.000	
6	.001	.003	.040	.176	.493	.777	.924		
	.740	.770	.832	.898	.956	.987	.994		
5	.001	.002	.031	.169	.369	.646			
	.448	.484	.544	.715	.821	.938			
4	.000	.000	.022	.102	.247				
	.164	.163	.291	.445	.657				
3	.000	.001	.007	.038					
	.037	.056	.094	.215					
2	.001	.000	.003						
	.005	.013	.020						
1	.000	.000							
	.000	.001							
0	.000								
	.000								

The top number is the power to detect two outliers.

The bottom number is the power to detect either outlier.

I also investigate the effect of the outliers′ position with respect to the interaction by perturbing either two cells orthogonal to the interaction, cells (4,7) and (5,6) or two cells embedded in the interaction, (1,1) and (2,2).

The results of the first simulation of John and Draper's test are contained in Table 9 for θ_1 and θ_2 added to cells (4,7) and (5,6). As expected for both outlier/interaction configurations, their test performs best in the least amount of interaction, $\lambda = 5$, and declines in power as λ approaches 8. When the outliers are orthogonal to the interaction, Table 9, the test can successfully detect both outliers as long as the magnitudes of θ_1 and θ_2 are close to or greater than the amount of interaction. For example, their test does well when $\lambda = 6$, $\theta_1 \geq 7\sigma$ and $\theta_2 \geq 7\sigma$. When θ_1 and θ_2 are embedded in the interaction, Table 10, then their test can detect only outliers whose magnitudes are much greater than the amount of interaction. For example, their test can successfully locate both outliers when $\lambda = 6$, $\theta_1 \geq 8\sigma$ and $\theta_2 \geq 8\sigma$, but it does not perform well when $\lambda = 8$, $\theta_1 = 8\sigma$ and $\theta_2 = 8\sigma$.

Before I present the findings of the second simulation, I would like to explain how I simulated the results. A Monte Carlo simulation similar to the simulation used in section 3.6.1 was developed to study my procedure's performance when two outliers and interaction were present in a two-way table. When the nonadditive effect was composed of three factors, there were seven ways to generate significant nonadditivity. Significance could have been the result of: 1.) θ_1, 2.) θ_2, 3.) λ, 4.) θ_1 and θ_2, 5.) θ_1 and λ, 6.) θ_2 and λ, or 7.) θ_1, θ_2 and λ. This meant that the following pairs of hypotheses, i.e., H_o: $\lambda_i = 0$ vs. H_1: $\lambda_i \neq 0$ (where: $i = 1, \ldots, 3$), were successively tested three times using the likelihood ratio test statistic (see section 3.5).

Based upon the outcome of these statistical tests, the simulation determined the nature of the nonadditivity. The following chart related the result of the likelihood tests to probable cause of significance:

Table 9. Power of John and Draper's test to detect two outliers in cells (4,7) and (5,6) of an 8 x 12 table.

$\lambda = 5$

θ_2 \ θ_1	5	6	7	8
8	.550 .881	.797 .978	.865 1.000	.780 .999
7	.609 .873	.837 .974	.933 .992	.848 .997
6	.532 .837	.772 .935	.821 .971	.783 .980
5	.375 .699	.545 .830	.626 .893	.569 .877

$\lambda = 7$

θ_2 \ θ_1	5	6	7	8
8	.388 .802	.677 .962	.828 .993	.785 .999
7	.427 .798	.704 .934	.857 .978	.824 .994
6	.340 .698	.591 .870	.698 .920	.688 .955
5	.200 .530	.371 .696	.412 .774	.360 .792

$\lambda = 6$

θ_2 \ θ_1	5	6	7	8
8	.458 .853	.725 .966	.837 .997	.759 .999
7	.504 .830	.801 .953	.908 .993	.853 .996
6	.449 .765	.693 .913	.780 .960	.733 .974
5	.297 .628	.460 .789	.510 .845	.463 .865

$\lambda = 8$

θ_2 \ θ_1	5	6	7	8
8	.293 .758	.564 .931	.794 .981	.749 .995
7	.309 .717	.625 .904	.780 .980	.781 .985
6	.232 .606	.442 .789	.631 .913	.618 .928
5	.122 .398	.281 .620	.302 .723	.305 .764

The top number is the power to detect two outliers.

The bottom number is the power to detect either outliers.

Table 10. Power of John and Draper's test to detect two outliers in cells (1,1) and (2,2) of an 8 x 12 table.

$\lambda = 5$

θ_2 \ θ_1	5	6	7	8
8	.176	.434	.698	**.913**
	.484	**.780**	**.941**	**.990**
7	.148	.385	**.715**	**.864**
	.442	**.718**	**.914**	**.970**
6	.132	.325	.501	.645
	.379	.618	**.811**	**.861**
5	.059	.177	.259	.323
	.217	.431	.550	.663

$\lambda = 7$

θ_2 \ θ_1	5	6	7	8
8	.028	.129	.372	.651
	.206	.437	**.716**	**.909**
7	.028	.107	.289	.529
	.155	.372	.645	**.816**
6	.014	.058	.161	.298
	.087	.242	.416	.618
5	.005	.028	.058	.092
	.033	.135	.238	.359

$\lambda = 6$

θ_2 \ θ_1	5	6	7	8
8	.067	.266	.566	**.808**
	.307	.629	**.846**	**.969**
7	.043	.258	.512	**.727**
	.263	.570	**.814**	**.907**
6	.048	.186	.343	.477
	.197	.435	.653	**.768**
5	.017	.072	.151	.186
	.101	.255	.430	.515

$\lambda = 8$

θ_2 \ θ_1	5	6	7	8
8	.008	.065	.213	.486
	.076	.311	.559	**.792**
7	.006	.043	.163	.331
	.069	.205	.432	.659
6	.004	.017	.086	.151
	.026	.103	.267	.444
5	.000	.006	.021	.052
	.008	.041	.120	.212

The top number is the power to detect two outliers.

The bottom number is the power to detect either outliers.

Test Outcome	$\lambda_1 \neq 0$ $\lambda_2 = \lambda_3 = 0$	$\lambda_2 \neq 0$ $\lambda_3 = 0$	$\lambda_3 \neq 0$
Cause of **Nonadditivity**	θ_1 or θ_2 or λ	θ_1 and θ_2 or θ_1 and λ or θ_2 and λ or θ_1, θ_2 and λ	θ_1, θ_2 and λ

If only $\lambda_1 \neq 0$, then the procedure examined the first singular vector of $\hat{\tau}$ and $\hat{\gamma}$ to decide upon the cause of significance: θ_1, θ_2 or λ. Like the earlier simulation of an outlier and interaction (section 3.6.1), this simulation compared the left and right singular vectors to a vector representation of each effect. It then selected the effect with the minimum sum of squared deviations for both $\hat{\tau}$ and $\hat{\gamma}$. This was the only case where the singular vectors were not rotated.

However if λ_2 was significant, then the cause of significance may be due to: θ_1 and θ_2, θ_1 and λ, or θ_2 and λ. (The case of θ_1, θ_2 and λ will be considered shortly.) The methodology was to independently rotate the first two vectors of $\hat{\tau}$ and $\hat{\gamma}$, and find the rotation with the largest weight, i.e., r. The location of these two weights must have agreed in row and column location of one of the outliers before the procedure continued. Because the rotation favored isolating the largest outlier in the first component, the second component was examined for presence of a second outlier or interaction using a minimum sum of squares deviation criteria. The effect with the minimum sum of squares for both $\hat{\tau}$ and $\hat{\gamma}$ was selected. If an effect was not found at this point, then the simulation proceeded to locate the second largest weight of the rotated left and right singular vectors. If these two weights correctly identified the row and column location of the second outlier, θ_1 or θ_2, then the cause of significance was found.

Occasionally all three effects caused the significance of λ_2, but not λ_3. This was especially true when the outliers were embedded in the interaction. Because my procedure rotated data, this case could be detected. First, the procedure must have correctly identified two effects where one effect was interaction. Next, the rotation with the second largest weight was identified for both $\hat{\tau}$ and $\hat{\gamma}$. If the location of these weights corresponded to the correct location for the second outlier in both row and column, then it was assumed that all three effects were detected.

Usually, these three nonadditive factors caused the significance of λ_3. When $\lambda_3 \neq 0$, the procedure independently rotated the first three vectors of $\hat{\tau}$ and $\hat{\gamma}$ in two steps. (See section 3.5 for a more detailed discussion of rotating q vectors.) The simulation found the two largest weights for both row and column effect. If these corresponded to the row and column location for θ_1 and θ_2, then a correct identification was made.

Up to now, I have recommended using a stepwise procedure for testing H_o: $\lambda_i = 0$ vs. H_1: $\lambda_i \neq 0$ along with examining consecutive test results for significance (section 3.5). However in this last simulation of three effects, there were several parameter combinations in which the likelihood ratio test frequently resulted in the following outcome: $\lambda_1 = 0$, $\lambda_2 = 0$ and $\lambda_3 \neq 0$. For example when $\theta_1 = \theta_2 = \lambda = 9$ and the outliers were in positions orthogonal to the interaction, U_1 and U_2 were not significant 413 times out of the 779 times that U_3 was significant. According to the recommended procedure of section 3.5, one would conclude that the data were completely additive if U_1 and U_1 were nonsignificant. When in reality, the data were anything but additive.

This occurs for the same reason that λ_2, but not λ_1, is significant in section 3.6.1. In this case, the three nonadditive effects are about equally spread among the first three components. This is supported by the result that the first three singular values are usually very close in value. When this happens, it is difficult for the likelihood ratio test statistic to be significant for U_1 and U_2 when $\hat{\lambda}_1{}^2$, $\hat{\lambda}_2{}^2$ and $\hat{\lambda}_3{}^2$, and $\hat{\lambda}_2{}^2$ and $\hat{\lambda}_3{}^2$ are in the denominators of U_1 and U_2, respectively. In light of this situation, a more appropriate test might be H_o: $\lambda_1 = \lambda_2 = \lambda_3 = 0$ vs. H_1: not H_o. Although the test statistic already exists for this hypothesis, critical points are not available (Corsten and van Eijnsbergen, 1972).

For the purpose of this simulation, sample results in which only U_3 is significant are included in the power calculation, since it is known that three effects were introduced into the data. However, in practice deciding how to handle this situation is not simple. One could plot the values of the k singular values against the order of the components to establish the relative closeness of the λ_i 's. If the first three singular values appear to clump together, then one might want to examine the three associated singular vectors to determine if they reveal any information about the nonadditivity.

The results of the second simulation are presented in Tables 11-14. Tables 11 and 12 contain the power of my procedure under the condition that the outliers are orthogonal to the interaction and embedded in the interaction, respectively. For each combination of θ_1, θ_2 and λ studied, there are seven numbers which represent the procedure's ability to detect each of the seven different ways to obtain significant nonadditivity (see chart earlier in this section). However, I present only the procedure's power to detect one or two outliers. As expected, the procedure performs better in detecting both outliers when they are orthogonal to the interaction instead of being buried in the interaction. The test achieves good power for correctly locating both outliers when $\theta_1 \geq 9\sigma$ and $\theta_2 \geq 9\sigma$, Table 11. The procedure has high power for detecting at least one outlier under both outlier/interaction configurations in about almost all combinations of θ_1, θ_2 and λ. The procedure's ability to detect θ_1 and θ_2 increases with an increased interaction effect. When the interaction component is strong, e.g., $\lambda = 15$, the procedure still has the capability of detecting two outliers even if the outliers are concealed in the interaction and their effects are less than 15. This result is also observed in the simulation of two nonadditive factors (section 3.6.1).

It is necessary to explain why it is easier for the procedure to detect outlier effects as the interaction effect increases. When $\lambda = 6$, $\theta_1 = \theta_2 = 9\sigma$ and the outliers are embedded in the interaction, (Table 12), the combined nonadditive effect mainly results in the significance of λ_1, i.e., U_1 was found significant 498 times, U_2 was significant 227 times and U_3 was significant 110 times out of 1000 runs. There were also 165 nonsignificant outcomes. Therefore, most of the significance is due to a single outlier effect, i.e., θ_1 or θ_2. This is confirmed by the procedure successfully detecting θ_1 and θ_2 312 times and either θ_1 or θ_2 711 times. As the interaction effect increases, the significance of U_2

Table 11. Power of my technique to detect two outliers in cells (4,7) and (5,6) of an 8 x 12 table.

$\lambda = 6$

θ_2	θ_1 = 6	9	12	15
15	.339 .982	.774 .993	.985 .998	.999 1.000
12	.302 .950	.741 .945	.958 .988	.979 .999
9	.245 .643	.671 .792	.741 .951	.801 .997
6	.163 .274	.260 .618	.314 .945	.318 .979

$\lambda = 12$

θ_2	θ_1 = 6	9	12	15
15	.392 .984	.929 .996	.999 1.000	1.000 1.000
12	.420 .978	.903 .982	.993 .995	1.000 1.000
9	.350 .798	.817 .904	.889 .989	.929 .997
6	.194 .363	.347 .792	.388 .972	.386 .986

$\lambda = 9$

θ_2	θ_1 = 6	9	12	15
15	.356 .984	.910 .982	.993 .999	.998 .999
12	.354 .954	.872 .963	.986 .990	.995 .998
9	.337 .750	.797 .859	.878 .964	.901 .987
6	.199 .351	.343 .734	.375 .941	.361 .981

$\lambda = 15$

θ_2	θ_1 = 6	9	12	15
15	.374 .985	.914 .996	1.000 1.000	1.000 1.000
12	.367 .971	.914 .984	.999 1.000	.999 1.000
9	.363 .816	.836 .908	.927 .989	.927 .997
6	.193 .373	.421 .815	.401 .973	.402 .992

The top number is the power to detect two outliers.

The bottom number is the power to detect either outliers.

Table 12. Power of my technique to detect two outliers in cells (1,1) and (2,2) of an 8 x 12 table.

$\lambda = 6$	θ_1			
θ_2	6	9	12	15
15	.109 .858	.507 .988	.901 .999	.990 .997
12	.088 .945	.417 .973	.793 .961	.922 .994
9	.084 .712	.312 .711	.486 .945	.566 .996
6	.051 .218	.120 .610	.152 .960	.149 .983

$\lambda = 12$	θ_1			
θ_2	6	9	12	15
15	.149 .979	.406 .973	.809 .974	.961 .986
12	.241 .971	.554 .931	.844 .930	.894 .975
9	.391 .855	.647 .809	.689 .904	.582 .979
6	.296 .432	.456 .781	.279 .948	.158 .985

$\lambda = 9$	θ_1			
θ_2	6	9	12	15
15	.074 .954	.386 .952	.761 .982	.927 .983
12	.092 .916	.387 .892	.697 .901	.815 .977
9	.206 .697	.388 .639	.471 .854	.485 .983
6	.165 .308	.215 .583	.171 .919	.120 .968

$\lambda = 15$	θ_1			
θ_2	6	9	12	15
15	.256 .989	.510 .989	.875 .995	.984 .992
12	.396 .982	.694 .977	.928 .970	.954 .984
9	.492 .853	.762 .876	.836 .959	.675 .988
6	.323 .423	.521 .778	.404 .956	.235 .988

The top number is the power to detect two outliers.

The bottom number is the power to detect either outliers.

occurs more frequently. When $\lambda = 15$ (and $\theta_1 = \theta_2 = 9\sigma$), U_1 was significant only 49 times, U_2 was significant 752 times, U_3 was significant 197 times and 2 outcomes were nonsignificant. This increased interaction effect results in more samples being rotated, which in turn gives rise to more outliers being identified. In this case, both outliers were correctly distinguished 762 times and either outlier was recognized 876 times.

The results of John and Draper's test for detecting two outliers are tabulated in Table 13 when the outliers are orthogonal to the interaction and Table 14 when they are buried in the interaction. Two numbers are presented for each combination of θ_1, θ_2 and λ. The first number is the power of their test to detect two outliers, whereas the second number is the test's ability to locate at least one outlier. In Table 13, their test performs very well in recognizing both outliers once $\theta_1 \geq 9\sigma$ and $\theta_2 \geq 9\sigma$. However if the outliers are concealed in the interaction, then their test achieves high power only when both outlier effects are close to or exceed the interaction effect, Table 14. When $\lambda >> \theta$, their test does not do an adequate job of detecting even a single outlier. So as the amount of interaction increases, the power of their test decreases.

I now compare the power of John and Draper's test to my procedure. When the outliers are orthogonal to the interaction, John and Draper's test performs slightly better than my procedure given $\theta_1 \geq 9\sigma$ and $\theta_2 \geq 9\sigma$. If $\theta_1 < 9\sigma$ and $\theta_2 < 9\sigma$, then their test procedure is more powerful for $\lambda = 6$, whereas my procedure performs better at $\lambda = 15$. When the outliers are hidden in the interaction, the procedures perform well in different domains. John and Draper's test performs satisfactorily in the presence of the least amount of interaction, $\lambda = 6$, by contrast my procedure produces favorable results in the presence of the most interaction, $\lambda = 15$. This result is similar to the performance of Stefansky's test in the presence of interaction. Recall that Stefansky's test performed well only when $\theta > \lambda$ (see section 3.6.1).

I conclude this chapter with the finding that tests devised to detect outliers in an additive two-way table perform well in the absence of interaction or the presence of mild interaction. If a strong interaction effect exists in the table, then the current outlier tests may not perform adequately. This

Table 13. Power of John and Draper's test to detect two outliers in cells (4,7) and (5,6) of an 8 x 12 table.

$\lambda = 6$

θ_2 \ θ_1	6	9	12	15
15	.821 .970	1.000 1.000	1.000 1.000	1.000 1.000
12	.836 .974	1.000 1.000	1.000 1.000	1.000 1.000
9	.841 .970	1.000 1.000	1.000 1.000	1.000 1.000
6	.687 .902	.835 .969	.826 .967	.834 .967

$\lambda = 12$

θ_2 \ θ_1	6	9	12	15
15	.236 .754	.965 .999	1.000 1.000	1.000 1.000
12	.238 .757	.961 .999	1.000 1.000	1.000 1.000
9	.227 .731	.932 .992	.954 .999	.970 .998
6	.072 .297	.234 .725	.245 .783	.247 .754

$\lambda = 9$

θ_2 \ θ_1	6	9	12	15
15	.588 .902	.998 1.000	1.000 1.000	1.000 1.000
12	.583 .891	.994 1.000	1.000 1.000	1.000 1.000
9	.561 .900	.990 1.000	.997 1.000	.996 1.000
6	.348 .705	.585 .898	.581 .902	.586 .916

$\lambda = 15$

θ_2 \ θ_1	6	9	12	15
15	.052 .540	.804 .990	1.000 1.000	1.000 1.000
12	.043 .509	.791 .988	1.000 1.000	1.000 1.000
9	.034 .385	.646 .958	.797 .991	.804 .991
6	.004 .042	.029 .398	.047 .523	.041 .527

The top number is the power to detect two outliers.

The bottom number is the power to detect either outliers.

Table 14. Power of John and Draper's test to detect two outliers in cells (1,1) and (2,2) of an 8 x 12 table.

$\lambda = 6$

θ_2 \ θ_1	6	9	12	15
15	.271 .630	**.978** **.999**	**1.000** **1.000**	**1.000** **1.000**
12	.267 .618	**.976** **.996**	**1.000** **1.000**	**1.000** **1.000**
9	.282 .627	**.975** **.993**	**.948** **1.000**	**.992** **.999**
6	.167 .417	.494 **.828**	.539 **.840**	.568 **.841**

$\lambda = 12$

θ_2 \ θ_1	6	9	12	15
15	.001 .008	.133 .557	**.905** **.987**	**1.000** **1.000**
12	.000 .010	.122 .525	**.901** **.983**	**.999** **1.000**
9	.000 .008	.075 .335	.538 **.847**	.553 **.873**
6	.000 .000	.001 .022	.007 .147	.007 .144

$\lambda = 9$

θ_2 \ θ_1	6	9	12	15
15	.017 .163	.665 **.934**	**.997** **.999**	**1.000** **1.000**
12	.018 .176	.662 **.927**	**.996** **.998**	**1.000** **1.000**
9	.017 .149	.651 **.888**	**.913** **.983**	**.931** **.982**
6	.005 .029	.102 .383	.126 .498	.150 .500

$\lambda = 15$

θ_2 \ θ_1	6	9	12	15
15	.000 .000	.001 .104	.374 **.833**	**.988** **1.000**
12	.000 .000	.001 .090	.325 **.770**	**.893** **.985**
9	.000 .000	.001 .010	.047 .285	.111 .552
6	.000 .000	.000 .000	.000 .005	.000 .007

The top number is the power to detect two outliers.

The bottom number is the power to detect either outliers.

is especially true if the outliers are embedded in the interaction effect and are not visible upon examination of the residuals. The procedure that I have developed for detecting outliers performs best in the presence of strong interaction for a given outlier effect. This procedure can be used in conjunction with the appropriate single or multiple outlier tests. If Stefansky's test or John and Draper's test fails to identify the outlier(s), then my procedure should be used when interaction is suspected. Therefore, the analyst now has a technique at his or her disposal for identifying outliers in a two-way table when the true underlying model is additive-plus-multiplicative of rank 1.

CHAPTER 4 - LEVERAGE TABLES

The goal of this study is to extend the leverage tables developed by Emerson *et al.* (1984) for an additive-plus-multiplicative model from from rank 1 to rank k. Leverage measures the impact a change in an observation has on a fitted value. In a two-way table, all cells are influenced by a disturbance in the (i,j) cell (Emerson *et al.*, 1984). When the model is additive, the change in the magnitude of a predicted value depends upon its location relative to the perturbed (i,j) cell. This change can be written as a "hat table", H(i,j),

$$\mathrm{H(i,j)} = e_i e_j' - (e_i - 1/r)(e_j' - 1'/c)$$

where:

e_i denotes a r x 1 vector with a 1 in position i and 0 elsewhere

e_j denotes a c x 1 vector with a 1 in position j and 0 elsewhere.

Under an additive-plus-multiplicative model, a change in a fitted value not only depends upon its location with respect to the disturbed cell, but also the magnitude of the perturbation. Emerson *et al.* derive the generalized leverage table and the Jacobian leverage table for a rank 1 additive-plus-multiplicative model. The generalized leverage table is defined to be the difference between predicted values for the perturbed, $\overline{Y}(\theta; i,j)$, and the unperturbed, $\overline{Y}$, tables under the additive-plus-multiplicative model divided by the amount of the perturbation θ, i. e.,

$$\mathrm{G}_{Y,\theta}(i,j) = \frac{\overline{Y}(\theta;i,j) - \overline{Y}}{\theta}.$$

where:

$$\overline{\mathrm{Y}}(\theta;\, i,j) = \hat{\mathrm{Y}} + \theta \mathrm{H(i,j)} + \hat{\lambda}_{1'}\hat{\tau}_{1'}\hat{\gamma}_{1'}{}'$$

$$\overline{\mathrm{Y}} = \hat{\mathrm{Y}} + \hat{\lambda}_1\hat{\tau}_1\hat{\gamma}_1{}'$$

$$1 \neq 1'.$$

The Jacobian leverage table, $G_Y(i,j)$, is the limit of the generalized leverage table as θ approaches zero and does not depend on θ, i.e.,

$$\mathrm{G}_Y(i,j) = \lim_{\theta \to 0} \mathrm{G}_{Y,\theta}(i,j).$$

They call this latter leverage table Jacobian, because $\mathrm{G}_Y(i,j)$ is a column of the Jacobian matrix which maps Y into $\overline{Y}$, i.e.,

$$\mathrm{G}_Y(i,j) = \left(\frac{\partial \overline{y}_{kl}}{\partial y_{ij}} \right).$$

The computational form of $\mathrm{G}_Y(i,j)$ is

$$\mathrm{G}_Y(i,j) = e_i e_j{}' - (e_i - 1/r - \hat{\tau}_i\hat{\tau})(e_j{}' - 1'/c - \hat{\gamma}_j\hat{\gamma}').$$

Leverage tables for an additive-plus-multiplicative model of rank 2 or greater do not exist. These are developed as part of this dissertation.

The easiest way to accomplish this extension is to develop the generalized leverage table for a rank k model, then take the limit of θ as it approaches zero to obtain the Jacobian leverage table. Throughout this chapter, I assume that the unperturbed table contains nonadditive effects that are of rank k. When the (i,j) cell is disturbed by θ, the rank of the table becomes k+1. This can be written as follows:

$$\text{true model (rank k)} \xrightarrow{\theta} \text{perturbed model (rank k + 1).}$$

Let the notation, $G_Y(i,j)_k$, represent the generalized leverage for a model that is rank k in its unperturbed state, therefore

$$\begin{aligned} G_Y(i,j)_k &= \lim_{\theta \to 0} G_{Y,\theta}(i,j)_k \\ &= \lim_{\theta \to 0} \frac{\overline{Y}(\theta; i,j)_k - \overline{Y}_k}{\theta} \end{aligned}$$

where:

$\overline{Y}(\theta; i,j)_k$ represents the fitted value for the perturbed table of rank k

$\overline{Y}_k$ represents the fitted value for the unperturbed table of rank k.

The fitted value for the unperturbed table, $\overline{Y}_k$, is equal to the sum of the additive and the multiplicative fits:

$$\overline{Y}_k = \hat{Y} + \hat{\lambda}_1\hat{\tau}_1\hat{\gamma}_1' + \cdots + \hat{\lambda}_k\hat{\tau}_k\hat{\gamma}_k',$$

whereas the least squares fit to the table perturbed by θ in cell (i,j) is:

$$\overline{Y}(\theta; i,j)_k = \hat{Y} + \theta H(i,j) + \hat{\lambda}_{1'}\hat{\tau}_{1'}\hat{\gamma}_{1'}' + \cdots + \hat{\lambda}_{k'}\hat{\tau}_{k'}\hat{\gamma}_{k'}'$$

where:

$$\hat{\lambda}_i \neq \hat{\lambda}_{i'}, \quad \hat{\tau}_i \neq \hat{\tau}_{i'}, \quad \hat{\gamma}_i \neq \hat{\gamma}_{i'}, \quad i = 1, \dots, k.$$

Therefore,

$$\begin{aligned} G_{Y,\theta}(i,j)_k &= \frac{\overline{Y}(\theta; i,j)_k - \overline{Y}_k}{\theta} \\ &= \frac{(\hat{Y} + \theta H(i,j) + \hat{\lambda}_{1'}\hat{\tau}_{1'}\hat{\gamma}_{1'}' + \cdots + \hat{\lambda}_{k'}\hat{\tau}_{k'}\hat{\gamma}_{k'}') - (\hat{Y} + \hat{\lambda}_1\hat{\tau}_1\hat{\gamma}_1' + \cdots + \hat{\lambda}_k\hat{\tau}_k\hat{\gamma}_k')}{\theta} \\ &= \frac{(\theta H(i,j) + \hat{\lambda}_{1'}\hat{\tau}_{1'}\hat{\gamma}_{1'}' + \cdots + \hat{\lambda}_{k'}\hat{\tau}_{k'}\hat{\gamma}_{k'}') - (\hat{\lambda}_1\hat{\tau}_1\hat{\gamma}_1' + \cdots + \hat{\lambda}_k\hat{\tau}_k\hat{\gamma}_k')}{\theta}. \end{aligned}$$

As mentioned before, when a table is perturbed by θ the rank of the nonadditive part increases by 1. In this case, the data can be fit exactly by using a k + 1 model or by adding an additional k + 1 component to the model. Therefore, $\overline{Y}(\theta; i,j)_k$ now becomes $\overline{Y}(\theta; i,j)_{k+1}$ or

$$\overline{Y}(\theta; i,j)_{k+1} = \hat{Y} + \theta H(i,j) + \hat{\lambda}_{1'}\hat{\tau}_{1'}\hat{\gamma}_{1'}' + \cdots + \hat{\lambda}_{k+1'}\hat{\tau}_{k+1'}\hat{\gamma}_{k+1'}'.$$

I use the overfit model in the calculation of the generalized leverage for the perturbed model. The notation, $G_{Y,\theta}(i,j)^{+}_{k+1}$, is used in this case because of the atypical way the generalized leverage is determined, i.e.,

$$G_{Y,\theta}(i,j)^{+}_{k+1} = \frac{(\hat{Y} + \theta H(i,j) + \hat{\lambda}_{1'}\hat{\tau}_{1'}\hat{\gamma}_{1'}' + \cdots + \hat{\lambda}_{k+1'}\hat{\tau}_{k+1'}\hat{\gamma}_{k+1'}') - (\hat{Y} + \hat{\lambda}_{1}\hat{\tau}_{1}\hat{\gamma}_{1}' + \cdots + \hat{\lambda}_{k}\hat{\tau}_{k}\hat{\gamma}_{k}')}{\theta}$$

$$= \frac{(\theta H(i,j) + \hat{\lambda}_{1}'\hat{\tau}_{1'}\hat{\gamma}_{1'}' + \cdots + \hat{\lambda}_{k+1'}\hat{\tau}_{k+1'}\hat{\gamma}_{k+1'}') - (\hat{\lambda}_{1}\hat{\tau}_{1}\hat{\gamma}_{1}' + \cdots + \hat{\lambda}_{k}\hat{\tau}_{k}\hat{\gamma}_{k}')}{\theta}.$$

This produces a 1 in the perturbed (i,j) cell and 0 elsewhere, ie.,

$$G_{Y,\theta}(i,j)^{+}_{k+1} = \begin{cases} 0 & (i,j) \neq (l,m) \qquad l = 1, \ldots, r \\ 1 & (i,j) = (l,m) \qquad m = 1, \ldots, c. \end{cases}$$

Let's consider the two cases:

i.) When (i,j) $\neq$ (l,m), then

$$0 = \frac{(\theta H(i,j) + \hat{\lambda}_{1'}\hat{\tau}_{1'}\hat{\gamma}_{1'}' + \cdots + \hat{\lambda}_{k+1'}\hat{\tau}_{k+1'}\hat{\gamma}_{k+1'}') - (\hat{\lambda}_{1}\hat{\tau}_{1}\hat{\gamma}_{1}' + \cdots + \hat{\lambda}_{k}\hat{\tau}_{k}\hat{\gamma}_{k}')}{\theta}$$

or

$$\frac{-\hat{\lambda}_{k+1'}\hat{\tau}_{k+1'}\hat{\gamma}_{k+1'}'}{\theta} = (\theta H(i,j) + \hat{\lambda}_{1'}\hat{\tau}_{1'}\hat{\gamma}_{1'}' + \cdots + \hat{\lambda}_{k'}\hat{\tau}_{k'}\hat{\gamma}_{k'}') - (\hat{\lambda}_{1}\hat{\tau}_{1}\hat{\gamma}_{1}' + \cdots + \hat{\lambda}_{k}\hat{\tau}_{k}\hat{\gamma}_{k}').$$

Thus, the generalized leverage of a rank k table for cells other than the perturbed (i,j) cell is equal to the k + 1 component (overfit term) divided by θ, i.e.,

$$G_{Y,\theta}(i,j)_k = \frac{-\hat{\lambda}_{k+1'}\hat{\tau}_{k+1'}\hat{\gamma}_{k+1'}'}{\theta}. \qquad [4.1]$$

ii.) When (i,j) = (l,m) then

$$1 = \frac{(\theta H(i,j) + \hat{\lambda}_{1'}\hat{\tau}_{1'}\hat{\gamma}_{1'}' + \cdots + \hat{\lambda}_{k+1'}\hat{\tau}_{k+1'}\hat{\gamma}_{k+1'}') - (\hat{\lambda}_1\hat{\tau}_1\hat{\gamma}_1' + \cdots + \hat{\lambda}_k\hat{\tau}_k\hat{\gamma}_k')}{\theta}$$

or

$$1 - \frac{\hat{\lambda}_{k+1'}\hat{\tau}_{k+1'}\hat{\gamma}_{k+1'}'}{\theta} = \frac{(\theta H(i,j) + \hat{\lambda}_{1'}\hat{\tau}_{1'}\hat{\gamma}_{1'}' + \cdots + \hat{\lambda}_{k'}\hat{\tau}_{k'}\hat{\gamma}_{k'}') - (\hat{\lambda}_1\hat{\tau}_1\hat{\gamma}_1' + \cdots + \hat{\lambda}_k\hat{\tau}_k\hat{\gamma}_k')}{\theta}.$$

Therefore, the change in the fitted value for the perturbed cell is 1 minus the change in the other cells divided by θ:

$$G_{Y,\theta}(i,j)_k = \frac{\theta - \hat{\lambda}_{k+1'}\hat{\tau}_{k+1'}\hat{\gamma}_{k+1'}'}{\theta}. \qquad [4.2]$$

In summary, the generalized leverage of a rank k model for cells other than the perturbed (i,j) cell is equal to the additional k+1 component (divided by θ) used to fit the perturbed table exactly. For the (i,j) cell, the generalized leverage is 1 minus the k+1 component (divided by θ):

$$G_{Y,\theta}(i,j)_k = \begin{cases} \dfrac{-\hat{\lambda}_{k+1'}\hat{\tau}_{k+1'}\hat{\gamma}_{k+1'}'}{\theta} & (i,j) \neq (l,m) \\[2ex] \dfrac{\theta - \hat{\lambda}_{k+1'}\hat{\tau}_{k+1'}\hat{\gamma}_{k+1'}'}{\theta} & (i,j) = (l,m). \end{cases}$$

I verified these formulas using a rank 1 model. I calculated $G_{Y,\theta}(i,j)_1$ and it was equivalent to $G_{Y,\theta}(i,j)$. I took the limit of $G_{Y,\theta}(i,j)_1$ as $\theta \to 0$ and it approximated the Jacobian leverage table nicely. Emerson *et al.* developed a computational form for the Jacobian leverage table which is not based on θ for the rank 1 case, but I have not been able to accomplish the same for the rank k case.

In practice, the leverage of an observation in the additive-plus-multiplicative model of rank k is not useful, because it requires knowledge of the true magnitude of θ to perform the calculation. However, additional research in this area may produce a computational form that is not based on θ.

CHAPTER 5 - STUDY OF INFLUENCE

To complete the picture on the examination of outliers and interaction in a nonreplicated two-way table, it is necessary to look at the influence each observation has on the analysis. Influential observations are commmonly equated with outliers, but the two concepts are different. Calder (1986) defines an influential observation as "one whose deletion from, or addition to, the data set leads to an unusually 'large' change in some aspect of the analysis." An outlier may or may not be influential. An influential observation may reveal important information about the system under investigation. It can also be used to direct the investigator to points of interest if replication were possible at a future date. A major segment of this dissertation has been devoted to outliers, so now I want to focus attention on influence analysis.

Calder (1986) extensively studied influence in a multivariate setting. Her approach was to delete an observation and estimate the change in the eigenvalues and eigenvectors. She determined an observation's influence on the eigenvalues by finding the difference between:

$$\hat{\lambda}_p^2 - \hat{\lambda}_p^{2*}$$

where:

$\hat{\lambda}_p^2$ is the pth eigenvalue of the entire data set

$\hat{\lambda}_p^{2*}$ is the pth eigenvalue of the data set with the ith observation removed.

The influence that an observation has on an eigenvector is measured by the angle the eigenvector rotates when the observation is omitted. It is calculated by:

$$\cos\ \theta = 1 - \frac{1}{2}(\hat{\alpha}_p - \hat{\alpha}_p^*)^2$$

where:

$\hat{\alpha}_p$ is the pth eigenvector corresponding to the pth eigenvalue of the complete data set

$\hat{\alpha}_p^*$ is the pth eigenvector corresponding to the pth eigenvalue of the data set with the observation deleted.

In a multivariate framework, measuring influence is straightforward. An observation is omitted and the corresponding change in the eigenvalues and eigenvectors is assessed. This is repeated for each observation in the data set. The observations are then ranked in order from the most influential, i.e., the observation that produced the greatest change, to the least influential, i.e., the observation that generated the least change. Lastly, observations that are highly influential on the eigenvalues are compared to those that are highly influential on the eigenvectors.

Throughout the development of this dissertation, multivariate techniques have been applied to a two-way table, i.e., principal components were used to fit the multiplicative part of the model. It is my intention to extend the approach of evaluating influence in a multivariate context to a two-way table.

In a two-way analysis of variance layout, assessing influence is more difficult. The reason is that one can not simply delete a row of a matrix to omit an observation like in a multivariate or regression setting. Nor can one merely delete a value in a particular cell to exclude an observation.

The way I omit a cell is to designate the deleted cell as missing and use the additive-plus-multiplicative model of rank k (where k is equal to the rank of the complete table) to fit the missing cell. This is accomplished in two steps. First, replace the observation to be removed with its predicted value under the additive model. This generates a zero in the residual matrix, Z, for this cell. Second, fit multiplicative part of the model to the residuals using weighted least squares. The weights, $w_{i,j}$, are all 1 except for the missing cell which receives a 0. Gabriel (1978) has shown that fitting a multiplicative model is equivalent to fitting the matrix AB′ where A is r x k and B is c x k. Then, minimize the following equation:

$$\Phi\ (A,B) = \left\{ \sum_{i=1}^{r} \sum_{j=1}^{c} w_{i,j}(z_{i,j} - a_i'b_j)^2 \right\}.$$

Gabriel and Zamir (1979) developed several iterative routines to minimize this equation and to fit a multiplicative model when a missing cell is present. This technique fits the remaining cells under the additive-plus-multiplicative model. Therefore, the missing cell does not contribute any sum of squares to $\Phi(A,B)$ or influence the fit of the multiplicative portion of Mandel's model. Once deletion is accomplished, then the analysis proceeds in a similar manner as for a multivariate data set. The only difference in the study of influence between a two-way layout and a multivariate setup is that two singular vectors need to be examined instead of one eigenvector for each component in the decomposition of the original data.

I studied influence in two agricultural data sets. The first example is the spring wheat data that was presented in section 2.1. The second example uses data from an experiment which investigated the effect of temperature and humidity on sorghum height in a growth chamber.

Table 15 compares the actual and percent change in the first singular value for several influential observations in the spring wheat data. The three most influential cells on $\hat{\lambda}_1$ are (1,1), (2,1) and (1,3), whereas the three least influential observations are (2,2), (3,2) and (1,2). The three most in-

Table 15. Comparison of the change in the first singular value for the most and the least influential cells of the spring wheat data.

		$\hat{\lambda}_1 = 507.66$		
		Actual Change	Percent Change	Cell
Most Influential	1	307.50	60.57%	(1,1)
	2	93.50	18.40%	(2,1)
	3	57.50	11.26%	(1,3)
	⋮	⋮	⋮	⋮
Least Influential	3	2.60	.4973%	(1,2)
	2	1.60	.3070%	(3,2)
	1	.40	.0670%	(2,2)

fluential cells on the first singular value are also strongly influential on the corresponding singular vectors, Table 16. This table displays the most and the least influential observations on the left and right singular vectors. The most influential observations for $\hat{\tau}$ all occur in column 1 (i.e., no phosphorus treatment). This may be due to the result that yield actually decreased in column 1 as nitrogen rates increased (i.e., the effect of a nitrogen-phosphorus imbalance). The highly influential cells for $\hat{\gamma}$ exist in row 1 (i.e., no nitrogen treatment). The control in the experiment, cell (1,1), stands out as extremely influential on $\hat{\lambda}_1$ and it is the only cell that is highly influential on both singular vectors. Observation (2,1) is influential on the row vector, $\hat{\tau}$, whereas cell (1,3) is influential on the column vector, $\hat{\gamma}$. Two of the cells that are influential on $\hat{\tau}$ and $\hat{\gamma}$, (3,1) and (1,4), respectively, do not rank among the top 3 influential observations on $\hat{\lambda}_1$, however both of them rank within the top 5 influential cells.

The spring wheat data has been analyzed by Johnson and Graybill (1972) and Milliken and Johnson (1989). Recall that the decomposition of residuals produced:

$$\hat{\lambda}_1 = 507.6624$$
$$\hat{\lambda}_2 = 30.5052$$

$$\hat{\tau} = \begin{bmatrix} .8123 \\ -.4776 \\ -.3347 \end{bmatrix} \quad \text{and} \quad \hat{\gamma} = \begin{bmatrix} .8231 \\ .0894 \\ -.4124 \\ -.3485 \\ -.1515 \end{bmatrix}$$

They found $\lambda_1 \neq 0$, so they tested all hypotheses of the form H_o: $\tau_i = \tau_{i'}$ and $\gamma_j = \gamma_{j'}$ (where: $i \neq i'$ and $j \neq j'$). Therefore, they concluded that significant interaction was attributed to cells (1,1), (1,2) and (1,5). Of the three cells responsible for the interaction, only one cell is influential, observation (1,1), the control. Cell (1,2) is among the least influential observation on the first singular value. It ranks fourth from the bottom of least influential observations for $\hat{\tau}$. The observation in

Table 16. Comparison of the angle that the first singular vectors rotate for the most and the least influential cells of the spring wheat.

		ϕ_1			
		$\hat{\tau}_1$		$\hat{\gamma}_1$	
		Angle	Cell	Angle	Cell
Most Influential	1	29.89	(2,1)	67.36	(1,1)
	2	19.27	(3,1)	27.29	(1,3)
	3	7.17	(1,1)	22.68	(1,4)
	⋮	⋮	⋮	⋮	⋮
Least Influential	3	.1381	(1,5)	1.871	(3,5)
	2	.0286	(2,2)	.6918	(3,1)
	1	.0267	(1,4)	.3941	(2,2)

cell (1,5) is also noninfluential on $\hat{\tau}$. It appears that two of the three cells responsible for significant interaction are not influential. The reason for their lack of influence may be related to the weights in $\hat{\tau}$ and $\hat{\gamma}$. The three most influential observations on $\hat{\gamma}$ are ranked as (1,1), (1,3) and (1,4). This is the same order as the product of the biggest weight in row 1 of $\hat{\tau}$, .8123, and the three largest weights in $\hat{\gamma}$. Columns 2 and 5 of $\hat{\gamma}$ are not influential because the magnitudes of their weights are low. This may also explain why the three most influential cells on $\hat{\tau}$ all exist in column 1, i.e., .8231 is the largest weight in $\hat{\gamma}$.

In closing, I would like to mention that my study of influence on the spring wheat data only involved an examination of the first multiplicative component. The reason for this is that the data's influence on the second component of $\hat{\tau}$ are identical to the first component. I believe that this does not reveal any information about the influence on $\hat{\tau}_2$, but merely reflects a dimension constraint on the minimum size of the table, $r = 3$. This constraint does not affect $\hat{\gamma}$. This same pattern was observed on the minimum dimension of the next example, $c = 4$. Since three components are generated in the decomposition of the residuals, the influence of only 2 components is unique for $\hat{\gamma}$.

The second example is a study of the effect of five temperatures and four humidity levels on the height of sorghum in a growth chamber (Milliken and Johnson, 1989; p.4). The data are displayed below:

Humidity, %

Temperature, °F	20	40	60	80
50	12.3	19.6	25.7	30.4
60	13.7	16.9	27.0	31.5
70	17.8	20.0	26.3	35.9
80	12.1	17.4	36.9	43.4
90	6.9	18.8	35.0	53.0

A decomposition of the residual generates:

$$\hat{\lambda}_1 = 18.9950$$
$$\hat{\lambda}_2 = 5.3380$$
$$\hat{\lambda}_3 = 3.3382$$

$$\hat{\tau} = \begin{bmatrix} -.3691 & .0759 \\ -.3161 & -.2260 \\ -.3461 & .4612 \\ .2788 & -.7395 \\ .7525 & .4284 \end{bmatrix} \quad \text{and} \quad \hat{\gamma} = \begin{bmatrix} -.6157 & .0021 \\ -.3016 & .3813 \\ .2250 & -.8166 \\ .6923 & .4333 \end{bmatrix}.$$

I present the singular vectors now and refer to them during the discussion of influence on the singular vectors. The residuals from the additive effects are fit with three multiplicative components. However only the influence on the first two components is studied, since the last component contains redundant information.

Table 17 presents the observations that are and are not influential. The four most influential observations on the first singular value occur at the four corners of the two-way table, whereas the three least influential cells exist in the middle of the table. I have listed four influential observations, because a recognizable pattern can be observed that can be seen in the singular vectors as well. There does not seem to be any pattern to the three influential observations on $\hat{\lambda}_2$. The three least influential observations once again are in the middle of the table.

Table 18 illustrates the most and least influential observations on the first and second singular vectors. The observation in cell (5,4) is the most influential cell on both singular vectors. It also corresponds to the largest weights in $\hat{\tau}_1$ row 5, .7525, and $\hat{\gamma}_1$ column 4, .6923. The other two influential observations on $\hat{\lambda}_1$, (1,4) and (5,1), are also influential on $\hat{\tau}_1$ and $\hat{\gamma}_1$, respectively. Examining the three most influential observations on $\hat{\gamma}_1$ shows that they all occur in row 5. This is probably the result of the weight, .7525, in row 5 of $\hat{\tau}_1$. The last corner observation that is influential on $\hat{\lambda}_1$, (1,1), is not influential on either $\hat{\tau}_1$ or $\hat{\gamma}_1$. Turning to the second set of singular vectors, I would expect that the observation in cell (4,3) is influential, since it follows the pattern of having large weights in row 2 of $\hat{\tau}_2$, -.7395, and column 3 of $\hat{\gamma}_2$, -.8166. However, I am surprised that cell (1,1) is influential on both singular vectors, $\hat{\tau}_2$ and $\hat{\gamma}_2$. This observation has the lowest weights on the second component for both singular vectors. I have no explanation for this occurrence. I would expect cell (1,1) to be more influential on the singular vectors of the first component, since it is influential on $\hat{\lambda}_1$ and not on $\hat{\lambda}_2$.

Although influential observations are frequently associated with large weights on the corresponding singular vectors, it is not necessary to have a large weight for an observation to be influential. This reflects the complexity of the data's structure.

Table 17. Comparison of the change in the singular values for the most and the least influential cells of the sorghum height data.

		$\hat{\lambda}_1 = 18.9950$			$\hat{\lambda}_2 = 5.3380$		
		Actual Change	Percent Change	Cell	Actual Change	Percent Change	Cell
Most Influential	1	6.0825	32.02%	(5,4)	-2.4610	-46.10%	(5,1)
	2	4.8633	25.60%	(5,1)	-1.4424	-27.02%	(3,4)
	3	1.5013	7.90%	(1,4)	1.3238	24.80%	(4,3)
	4	.8967	4.72%	(1,1)	.	.	.
	.	.	.	.	.	.	.
	.	.	.	.	.	.	.
	.	.	.	.	.	.	.
Least Influential	3	.1809	.95%	(1,3)	-.1507	-2.82%	(2,3)
	2	.1664	.88%	(2,2)	-.1428	-2.68%	(2,4)
	1	.0490	.26%	(2,3)	.0197	.37%	(4,2)

Table 18. Comparison of the angle that the singular vectors rotate for the most and the least influential cells of the sorghum data.

		ϕ_1				ϕ_2			
		$\hat{\tau}_1$		$\hat{\gamma}_1$		$\hat{\tau}_2$		$\hat{\gamma}_2$	
		Angle	Cell	Angle	Cell	Angle	Cell	Angle	Cell
Most Influential	1	23.77	(5,4)	30.42	(5,4)	82.96	(4,3)	73.88	(5,4)
	2	16.85	(1,4)	27.20	(5,1)	71.37	(1,1)	72.31	(1,1)
	3	14.44	(3,1)	10.50	(5,2)	57.19	(5,4)	64.56	(4,3)
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Least Influential	3	1.29	(2,2)	1.22	(4,4)	8.27	(4,4)	2.79	(1,2)
	2	.97	(5,3)	.57	(2,3)	4.86	(2,3)	1.98	(1,4)
	1	.38	(2,3)	.24	(3,4)	4.05	(5,3)	.14	(2,3)

CHAPTER 6 - CONCLUSIONS

In this chapter, I summarize the findings of this study on outliers and interaction in an unreplicated two-way layout. I critique my procedure, the leverage investigation and the influence study. Throughout this chapter, I address some areas of further research.

The main contribution of this investigation is the development of a procedure to identify both outliers and rank 1 interaction in an nonreplicated two-way table. My technique appears to complement Stefansky's test and John and Draper's test for detecting one or two outliers, respectively. As long as the interaction effect is strong, my procedure performs well in locating the outlier even if the outlier effect is weaker than the interaction effect. The advantage of this procedure is that it does not rely on the largest residual like the current outlier tests, so it is able to detect an outlier that may be hiddened in the interaction.

In the presence of no or mild interaction, my technique is not as powerful as either of the above mentioned tests. The reason for this is largely do to the power of the likelihood ratio tests for testing the significance of λ_i in the additive-plus-multiplicative model. If all $\lambda_i = 0$, then the data would not be examined for nonadditivity. However, rotating the singular vectors frequently uncovers the outlier effect even if λ_i is nonsignificant. This suggests that the procedure can be used as a exploratory tool to look for potential outliers.

As mentioned in the simulation of two outliers and interaction, section 3.6.3, occasionally the first two singular values are not significant whereas the third singular value is significantly different from zero. Usually, this occurs because three nonadditive effects of similar magnitude are present in the data which in turn produce three singular values close in value. According to Milliken and Johnson (1989), these data should not be analyzed for nonadditivity because two consecutive nonsignificant test results appeared. The high power of the procedure in these circumstances indicates that many samples would have been overlooked had this recommendation been followed for detecting outliers. An appropriate way to handle this situation might be to plot the values of the singular values against the order of the components to reveal the relative closeness of the λ_i's. If the first three singular values appear to clump together compared to the remaining singular values, then my procedure can be implemented to investigate the nonadditivity.

I study the performance of my procedure to detect an outlier and interaction (3.6.1), two outliers (3.6.2) and two outliers and interaction (3.6.3). In all simulations the interaction effect is rank 1. I do not present any results for the case where the nonadditive term is composed of two interaction effects, i.e., rank 2. This is because the procedure does not work well in distinguishing two complex interaction factors. The Givens rotation is designed to zero a particular element of a singular vector and place all the weight on the opposing element of the same row, $[r \quad 0]$. This method functions well in identifying the row (or column) location of the outlier. It could also be used to separate simple interaction, i.e., interaction that is confined to either a row or column of a two-way table. Interaction that has a more complicated form, eg., linear, quadratic, etc., would not benefit from this method. Several researchers have tried using factor analysis in this situation, however they have been unable to produce any interpretative results (Johnson 1990, personal communication). This problem needs further investigation and solving it would make a tremendous contribution to the statistical community.

During this study of outliers and interaction in a two-way setup, I briefly investigate the topics of leverage and influence. I extend the leverage tables for an additive-plus-multiplicative model of rank 1 (Emerson *et al.*, 1984) to a rank k model. However in practice, these tables of rank k are not useful since they depend upon knowing the true value of the outlier effect, θ. A computational form for leverage was developed by Emerson *et al.* for the rank 1 model which did not rely on θ, but I could not devise a similar formula for the rank k model. Further research in this area might produce a form for the rank k model which does not depend on knowing θ.

I look into the effect deleting an observation in an unreplicated two-way table has on the estimated parameters of the additive-plus-multiplicative model. Influential observations frequently occur in cells whose row and column location in the left and right singular vectors have large weights. However, sometimes an influential observation does not adhere to this pattern. This reminds me of how complex a data structure can be in a two-way table.

Lastly, my entire examination of outliers and interaction focuses on data in an unreplicated two-way table. I did not study any three-way tables or more sophisticated designs. The ability to detect both outliers and interaction needs to be investigated for nonreplicated designs in general.

BIBLIOGRAPHY

Barnett, V. and Lewis, T. (1984). *Outliers in Statistical Data, 2nd ed.* Wiley, New York, New York, 463p.

Black, A. L. (1970). "Adventitious Roots, Tillers, and Grain Yields of Spring Wheat as Influenced by N-P Fertilization." Agronomy Journal, 62, 32-36.

Bradu, D. and Hawkins, D. M. (1982). "Location of Multiple Outliers in Two-way Tables Using Tetrads." Technometrics, 24, 103-108.

Bross, I. D. J. (1961). "Outliers in Patterned Experiments: A Strategic Appraisal." Technometrics, 3, 91-102.

Brown, M. B. (1975). "Exploring Interaction Effects in the ANOVA." Applied Statistics, 24, 288-298.

Calder, P. (1986). *Influence Functions in Multivariate Analysis.* Unpublished Ph.D. Thesis, University of Kent.

Corsten, L. C. A. and van Eijnsbergen, A. C. (1972). "Multiplicative Effects in Two-way Analysis of Variance." Statistical Neerlandica, 26, 61-67.

Daniel, C. (1960). "Locating Outliers in Factorial Experiments." Technometrics, 2, 149-156.

Daniel, C. (1978). "Patterns in Residuals in the Two-Way Layout." Technometrics, 20, 385-395.

Draper, N. R. and John, J. A. (1980). "Testing for Three of Fewer Outliers in Two-Way Tables." Technometrics, 22, 9-15.

Emerson, J. D., Hoaglin, D. C. and Kempthorne, P. J. (1984). "Leverage in Least Squares Additive-Plus-Multiplicative Fits for Two-Way Tables." Journal of the American Statistical Association, 79, 329-335.

Gabriel, K. R. (1978). "Least Squares Approximation of Matrices by Additive and Multiplicative Models." Journal of the Royal Statistical Society, B, 40, 186-196.

Gabriel, K. R. and Zamir, S. (1979). "Lower Rank Approximations of Matrices by Least Squares with any Choice of Weights." Technometrics, 21(4), 489-499.

Galpin, J. S. and Hawkins, D. M. (1981). "Rejection of a Single Outlier in Two- of Three-Way Layouts." Technometrics, 23, 65-70.

Gauch, H. G. (1988). "Model Selection and Validation for Yield Trials with Interaction." Biometrics, 44, 705-715.

Gentleman, J. F. and Wilk, M. B. (1975a). "Detecting Outliers in a Two-Way Table: I. Statistical Behavior of Residuals." Technometrics, 17, 1-14.

Gentleman, J. F. and Wilk, M. B. (1975b). "Detecting Outliers. II. Supplementing the Direct Analysis of Residuals." Biometrics, 31, 387-410.

Givens, W. (1954). "Numerical Computation of the Characteristic Values of a Real Symmetric Matrix." Oak Ridge National Laboratory report ORNL-1574, Oak Ridge, TN.

Gnanadesikan, R. and Kettenring, J. R. (1972). "Robust Estimates, Residuals, and Outlier Detection with Multiresponse Data." Biometrics, 28,81-124.

Goldsmith, P. L. and Boddy, R. (1973). "Critical Analysis of Factorial Experiments and Orthogonal Fractions." Applied Statistics, 22, 141-160.

Gollob, H. F. (1968). "A Statistical Model which Combines Features of Factor Analytic and Analysis of Variance Techniques." Psychometrika, 33, 73-116.

Golub, G. H. and Reinsch, C. (1971). "Singular Value Decomposition and Least Squares Solutions." In *Linear Algebra, Handbook for Automatic Computation* Vol. II. Springer-Verlag, New York, New York. 439p.

Golub, G. H. and VanLoan, C. F. (1989). *Matrix Computations, 2nd.* The Johns Hopkins University Press, Baltimore, MD 642p.

Hawkins, D. M. (1984). *Identification of Outliers.* Monographs on Applied Probability and Statistics. Chapman and Hall, London, 188p.

Hegemann, V. J. and Johnson, D. E. (1976). "On Analyzing Two-Way AOV Data with Interaction." Technometrics, 18, 273-281.

John, J. A. and Draper, N. R. (1978). "On Testing for Two Outliers or One Outlier in Two-Way Tables." Technometrics, 20, 69-78.

Johnson, D. E. and Graybill, F. A. (1972). "An Analysis of a Two-Way Model with Interaction and No Replication." Journal of the American Statistical Association, 67, 862-868.

Jolliffe, I. T. (1986). *Principal Component Analysis.* Springer-Verlag, New York, 271p.

Kennedy, W. J. and Gentle, J. E. (1980). *Statistical Computing.* Statistics: Textbooks and Monographs. Vol 33. Marcel Dekker, Inc., New York, New York, 591p.

Kolman, B. (1976). *Introductory Linear Algebra with Applications.* Macmillan, New York, 426p.

Mandel, J. (1961). "Non-additivity in Two-Way Analysis of Variance." Journal of the American Statistical Association, 56, 878-888.

Mandel, J. (1969). "The Partitioning of Interaction in Analysis of Variance." Journal of Research of the National Bureau of Standards, 73B, 309-328.

Mandel, J. (1971). "A New Analysis of Variance Model for Non-additive Data." Technometrics, 13, 1-18.

Marasinghe, M. G. and Johnson, D. E. (1981). "Testing Subhypotheses in the Multiplicative Interaction Model." Technometrics, 23, 385-393.

Milliken, G. A. and Johnson, D. E. (1989). *Analysis of Messy Data. Vol 2: Nonreplicated Experiments*. Van Nostrand Reinhold, New York, New York. 199p.

Pack, P., Jolliffe, I. T. and Morgan, B. J. T. (1980). "Influential Observations in Principal Component Analysis: A Case Study." Journal of Applied Statistics, 15, 39-52.

Schuurmann, F. J., Krishnaiah, P. R. and Chattopadhyah, A. K. (1973). "On the Distribution of the Ratios of the Extreme Roots to the Trace of the Wishart Matrix." Journal of Multivariate Analysis, 3, 445-453.

Snee, R. D., Acuff, S. K. and Gibson, J. R. (1979). "A Useful Method for the Analysis of Growth Studies." Biometrics, 35, 835-848.

Stefansky, W. (1971). "Rejecting Outliers by Maximum Normed Residual." Annals of Mathematical Statistics, 42, 35-45.

Stefansky, W. (1972). "Rejecting Outliers in Factorial Designs." Technometrics, 14, 469-479.

Yochmowitz, M. G. and Cornell, R. G. (1978). "Stepwise Tests for Multiplicative Components of Interaction." Technometrics, 20, 79-84.

Appendix A. Givens Rotation - Case 3

A rotation when θ_1 (or θ_2) is not in row i or k, say row m.

Recall that the mth row of τ is:

$$\begin{bmatrix} 0 & \dfrac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \end{bmatrix}$$

where:

m $\neq$ i and m $\neq$ k,

therefore

$$\text{SIN } \theta = \frac{\dfrac{-2\theta_1}{\sqrt{2r(r-2)\theta_1^2}}}{\sqrt{0^2 + \dfrac{2^2\theta_1^2}{2r(r-2)\theta_1^2}}} \quad \text{and} \quad \text{COS } \theta = \frac{0}{\sqrt{0^2 + \dfrac{2^2\theta_1^2}{2r(r-2)\theta_1^2}}}$$

$$= -1 \qquad\qquad = 0.$$

Now rotate τ by the appropriate Givens rotation:

$$= \tau \begin{bmatrix} \text{COS } \theta & \text{SIN } \theta \\ -\text{ SIN } \theta & \text{COS } \theta \end{bmatrix}$$

$$
= \begin{bmatrix}
0 & \frac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\
\vdots & \vdots \\
0 & \frac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\
\frac{-r\theta_1}{\sqrt{2r^2\theta_1^2}} & \frac{-(r-2)\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\
0 & \frac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\
\vdots & \vdots \\
0 & \frac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\
\frac{r\theta_1}{\sqrt{2r^2\theta_1^2}} & \frac{-(r-2)\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\
0 & \frac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\
\vdots & \vdots \\
0 & \frac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}} \\
\vdots & \vdots \\
0 & \frac{2\theta_1}{\sqrt{2r(r-2)\theta_1^2}}
\end{bmatrix}
\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}
$$

$$
= \begin{array}{c} \left[\begin{array}{cc}
\sqrt{\dfrac{2}{r(r-2)}} & 0 \\
\vdots & \vdots \\
\sqrt{\dfrac{2}{r(r-2)}} & 0 \\
-\sqrt{\dfrac{r-2}{2r}} & -\sqrt{\dfrac{1}{2}} \\
\sqrt{\dfrac{2}{r(r-2)}} & 0 \\
\vdots & \vdots \\
\sqrt{\dfrac{2}{r(r-2)}} & 0 \\
-\sqrt{\dfrac{r-2}{2r}} & \sqrt{\dfrac{1}{2}} \\
\sqrt{\dfrac{2}{r(r-2)}} & 0 \\
\vdots & \vdots \\
\sqrt{\dfrac{2}{r(r-2)}} & 0 \\
\vdots & \vdots \\
\sqrt{\dfrac{2}{r(r-2)}} & 0
\end{array} \right] \end{array}
\begin{array}{l}
\\ \\ \\ \leftarrow \text{ith row} \\ \\ \\ \\ \leftarrow \text{kth row} \\ \\ \\ \leftarrow \text{mth row} \\ \\ \\
\end{array}
$$

In general, rotating τ with a Givens rotation tailored to zero the second element in row m, generates a r x 2 matrix whose first column is a contrast between the two rows, i and k, containing outliers verses the remaining rows and whose second column is a contrast between row i and row k. In this case developed above, rotating $\hat{\tau}$ merely switched the order of the singular vectors since the necessary contrast already existed.

Recall Example 2-

$\theta_1 = 5$ is in row 1, $\theta_2 = 8$ is in row 5 and

$$\tau = \begin{array}{cl} \begin{bmatrix} -.4924 & .7467 \\ -.1179 & -.3456 \\ -.1179 & -.3456 \\ -.1179 & -.3456 \\ .8461 & .2901 \end{bmatrix} & \begin{array}{l} \\ \leftarrow \textbf{mth row} \\ \\ \\ \\ \end{array} \end{array}$$

therefore

$$\text{SIN } \theta = .9464 \quad \text{and} \quad \text{COS } \theta = -.3229.$$

Now, rotating τ by this rotation yields:

$$
= \begin{bmatrix} -.4924 & .7467 \\ -.1179 & -.3456 \\ -.1179 & -.3456 \\ -.1179 & -.3456 \\ .8461 & .2901 \end{bmatrix} \begin{bmatrix} -.3229 & .9464 \\ -.9464 & -.3229 \end{bmatrix}
$$

$$
= \begin{bmatrix} -.5477 & -.7071 \\ .3651 & 0 \\ .3651 & 0 \\ .3651 & 0 \\ -.5477 & .7071 \end{bmatrix}.
$$

When r = 5, this result is equal to the general result derived for $\theta_1 = \theta_2$:

$$
\begin{bmatrix} -\sqrt{\frac{r-2}{2r}} & \sqrt{\frac{1}{2}} \\ \sqrt{\frac{2}{r(r-2)}} & 0 \\ \sqrt{\frac{2}{r(r-2)}} & 0 \\ \sqrt{\frac{2}{r(r-2)}} & 0 \\ -\sqrt{\frac{r-2}{2r}} & -\sqrt{\frac{1}{2}} \end{bmatrix}.
$$

Appendix B. The Relationship between Interaction and Outlier Tests

The purpose of this section is to demonstrate the equivalence in expectation between Johnson and Graybill's (1972) statistic for inclusion of the interaction term, $\lambda_1\tau_{1i}\gamma_{1j}$, in an additive model and Stefansky's test statistic for rejection of a single outlier in an unreplicated two-way table.

Theorem 1.

The test statistic, Λ^* (Johnson and Graybill), is related to the maximum normed residual, $|Z|^{(1)}$ (Stefansky), in expectation by the following equation:

$$\Lambda^* = \frac{rc}{(r-1)(c-1)}\left(|Z|^{(1)}\right)^2. \qquad [B.1]$$

Proof:

$$\Lambda^* = \frac{\hat{\lambda}_1^2}{\sum_{i=1}^{k} \hat{\lambda}_i^2} = \frac{\lambda_1^2}{\sum_{i=1}^{n} e_i^2}$$

$$= \frac{\left(\frac{\theta_1}{rc}\sqrt{r(r-1)c(c-1)}\right)^2}{\sum_{i=1}^{n} e_i^2} \qquad \text{Using equation [3.2.2]}$$

$$= \frac{\frac{\theta_1^2}{r^2c^2}\, r(r-1)c(c-1)}{\sum_{i=1}^{n} e_i^2}.$$

Now,

$$|Z|^{(1)} = \max\left|\frac{e_i}{\sqrt{\sum_{i=1}^{n} e_i^2}}\right|$$

$$= \frac{\frac{\theta_1(r-1)(c-1)}{rc}}{\sqrt{\sum_{i=1}^{n} e_i^2}}.$$

Therefore,

$$\frac{\frac{\theta_1^2}{r^2c^2}\, r(r-1)c(c-1)}{\sum_{i=1}^{n} e_i^2} = \frac{\frac{rc}{(r-1)(c-1)} \frac{\theta_1^2(r-1)^2(c-1)^2}{r^2c^2}}{\sum_{i=1}^{n} e_i^2}$$

and

$$\Lambda^{*} = \frac{rc}{(r-1)(c-1)} \left(|Z|^{(1)} \right)^{2}.$$

VITA

I was born December 29, 1955 in Little Falls, New York to Peter and Dorothy Kuzmak. I was educated in the Little Falls public school system. Upon graduation from high school, I gained entrance into Cornell University in Ithaca, New York.

While at Cornell, I majored in entomology. My main interest in entomology was integrated pest management, because it quantitatively studied an agricultural system before deciding how to manage a pest problem. I wanted to obtain some field experience in this discipline, so I enrolled in the Master's program at Kansas State University in Manhattan, Kansas. While working on my Master's thesis in entomology, I realized that I enjoyed analyzing data more than counting bugs.

Initially, I pursued my interest in statistics by taking some extra courses in this discipline, however my interest in statistics was not quenched by these additional courses. Eventually, I changed careers and enrolled in the Master's program in statistics at Kansas State University. My main interests in statistics were the topics of linear models and experimental design. I especially enjoyed applying statistical methodology to agricultural situations and teaching statistics. I wanted a Ph.D. to teach statistics at a university, and I also desired to relocate closer to my family and live by mountains and trees. So, I applied for and was granted admission to the doctorate program at Virginia Polytechnic Institute and State University.

I spent four years in the doctoral program at Virginia Tech. I have had the opportunity to take and teach many courses in statistics. I have had a very successful teaching career and have enjoyed being in the classroom. I worked under Eric Smith during my dissertation research and studied an area of messy data that dealt with nonreplicated designs. I felt privileged to experience Eric's enthusiasm for research while working on my dissertation. In my spare time, I have enjoyed visiting my brother's family in Baltimore, taking frequent hikes on the Appalachian Trail and relaxing moments doing cross stitch and needlepoint.

I have taken a position with The Procter & Gamble Company in Cincinnati, Ohio. My responsibility is to provide statistical support in product development. I plan to start work the end of December, 1990.

Barbara R. Kuzmak