ANALYSIS OF REMOVABLE INTERACTION

by

SHAROON HANOOK

Submitted in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

Dissertation Advisor: Robert C. Elston, Ph.D.

Department of Epidemiology and Biostatistics

CASE WESTERN RESERVE UNIVERSITY

January 2015
CASE WESTERN RESERVE UNIVERSITY
SCHOOL OF GRADUATE STUDIES

We hereby approve the thesis/dissertation of

__________________________
SHAROON HANOOK

candidate for the ______Ph.D.______ degree

(signed) ___________ Robert C. Elston, Ph.D. ____________

(Chair of committee)

__________________________
Mark Schluchter, Ph.D. ____________

__________________________
Jiayang Sun, Ph.D. ____________

__________________________
Jill Barnholtz-Sloan, Ph.D. ____________

(Date) __ August 27, 2014 ___
I dedicate this dissertation to:

My parents, Mr. Hanook and Mrs. Charlotte Hanook,

My sisters Rubeca, Naumi, Milka, Tamar and my fiancée Shonam
# Table of Contents

List of Tables vii  
List of Figures ix  
Acknowledgements x  
Abstract xii  

## Chapter 1  Introduction  
1.1 Definitions and Interpretation of Interaction  
1.1.1 Type A interaction  
1.1.2 Type B interaction  
1.1.3 Type C interaction  
1.1.4 Type D interaction  
1.1.5 Interpretation of interaction  
1.2 Detecting interaction  
1.3 Essential and Removable Interaction  
1.3.1 Testing of removable interaction  
1.4 Objective of the study  

## Chapter 2  Weighted Tukey’s test for removable interaction  
2.1 Non-additivity and tests for non-additivity  
2.2 A new test to detect removable interaction  
2.2.1 Two-way classification model  
2.2.2 Interaction structure  
2.2.3 Estimation of parameters  
2.2.4 Tests of hypothesis
2.2.5 Missing data 31
2.2.6 Numerical Examples 32
  2.2.6.1 Interaction structure I 32
  2.2.6.2 Interaction structure II 36
2.3 Extension to more than one observation per cell (Equal n) 42
  2.3.1 Two way classification model (n>1) 43
  2.3.2 Estimation of parameters 44
  2.3.3 Tests of hypothesis 46
2.4 Strategy for the analysis of a factorial design 52
2.5 Numerical example 53
2.6 Conclusion 58

Chapter 3 Modification of the weighted Tukey's test for removable interaction 59
3.1 Introduction 59
  3.1.1 Modification of the test 60
  3.1.2 Estimation of parameters 60
  3.1.3 Testing removable interaction 63
  3.1.4 Small sample approximation 64
3.2 Numerical example 65
3.3 Simulation study 70
3.4 Conclusion 72
Chapter 4  Transformation to remove Non-Additivity  73

4.1  Removing statistical interaction  73

4.2  General discretization algorithm  78

4.3  Empirical transformation  79

4.4  Transformation for one observation per cell  83

4.4.1  Transformation for different structures of interaction  83

4.4.1.1  Interaction structure I  83

4.4.1.2  Interaction structure II  88

4.4.1.3  Interaction structure III  92

4.5  Transformation for more than one observation per cell  96

4.5.1  Equal number of observations  96

4.5.2  Unequal number of observations  100

Chapter 5  Conclusion and future work  104

5.1  Conclusion  105

5.2  Future work  106

References  107
List of Tables

Table 2.2  ANOVA table for the weighted Tukey's test for one observation per cell  30

Table 2.2.1 percent change of LVP measure  32

Table 2.2.2  ANOVA table for Tukey's test for removable interaction  34

Table 2.2.3  ANOVA table for the weighted Tukey's test for removable interaction  35

Table 2.2.4  Yields of corn in Bushels per acre  36

Table 2.2.5  ANOVA table for Tukey's test for removable interaction  38

Table 2.2.6  ANOVA table for the weighted Tukey's test for removable interaction  40

Table 2.3.1  ANOVA table for testing removable interaction (n>1 & equal subclass size)  51

Table 2.5.1  dataset for two-way classification with more than one observation per cell  54

Table 2.5.2  ANOVA table for Tukey's test for removable interaction for more than one observation per cell  56

Table 2.5.3  ANOVA table for weighted Tukey's test for removable interaction for more than one observation per cell  57

Table 3.2.1  Dataset for two-way classification data with interaction  66

Table 3.2.2  ANOVA table for weighted Tukey's test for removable interaction  68

Table 3.2.3  Modified Tukey's test for removable interaction  68

Table 3.2.4  ANOVA table for weighted Tukey's test for removable interaction  69

Table 3.2.5  Modified weighted Tukey's test for removable interaction  70

Table 3.2.6  Power comparison of tests for removable interaction  71
| Table 4.4.1 | ANOVA table for Tukey’s test for removable interaction for one observation per cell, before transformation | 85 |
| Table 4.4.2 | Monotonic scores after transformation | 86 |
| Table 4.4.3 | ANOVA table for Tukey’s test for removable interaction for one observation per cell, after transformation | 87 |
| Table 4.4.4 | ANOVA table for Tukey’s test for removable interaction for one observation per cell, before transformation | 88 |
| Table 4.4.5 | Monotonic scores after transformation | 90 |
| Table 4.4.6 | ANOVA table for Tukey’s test for removable interaction for one observation per cell, after transformation | 91 |
| Table 4.4.7 | Volumes in milliliters of loaves | 92 |
| Table 4.4.8 | ANOVA table for Tukey’s test for removable interaction for one observation per cell, before transformation | 93 |
| Table 4.4.9 | Monotonic scores after transformation | 95 |
| Table 4.4.10 | ANOVA table for Tukey’s test for removable interaction for one observation per cell, after transformation | 95 |
| Table 4.5.1 | ANOVA table for Tukey’s test for removable interaction for more than one observation per cell, before transformation | 97 |
| Table 4.5.2 | ANOVA table for Unbalanced data, before transformation | 101 |
| Table 4.5.3 | Monotonic scores after transformation | 103 |
| Table 4.5.4 | ANOVA table for unbalanced data, after transformation | 104 |
List of Figures

Figure 1.1 Different types of interactions 7
Figure 2.2.1 Interaction plot for percentage change of LVP versus Propofol 33
Figure 2.2.2 Interaction plot of soil type versus liming treatment 37
Figure 2.2.3 Interaction plot of soil type versus liming treatment after weights 40
Figure 2.4.1 Strategy for analysis of a factorial design 52
Figure 2.5.1 Interaction plot of pentosans versus laboratory 55
Figure 3.2.1 Interaction plot of columns versus rows 67
Figure 4.1 Discretization process 78
Figure 4.4.1(a) Interaction plot of Pentosans versus Lab before transformation 85
Figure 4.4.1(b) Interaction plot of Pentosans versus Lab after transformation 85
Figure 4.4.2(a) Interaction plot of columns versus rows before transformation 90
Figure 4.4.2(b) Interaction plot of columns versus rows after transformation 90
Figure 4.4.3(a) Interaction plot of variety versus Potassium Bromate before transformation 94
Figure 4.4.3(b) Interaction plot of variety versus Potassium Bromate after transformation 94
Figure 4.5.1 Interaction plot of columns versus rows before transformation 98
Figure 4.5.2(a) Interaction plot of columns versus rows before transformation 102
Figure 4.5.2(b) Interaction plot of columns versus rows after transformation 102
Acknowledgements

The complete PhD process has been one of the most significant challenges I have faced. Without the support, patience and guidance of the following people and organizations, this would not have been completed. It is to them I owe my deepest gratitude.

First and foremost, I would like to give all praise and glory to heavenly father for giving me wisdom and strength to accomplish my doctorate. I would like to acknowledge my earnest gratitude to my advisor, Dr. Robert Elston, who undertook to be my research mentor despite his many other academic and professional commitments. I am so thankful for his wisdom, knowledge and candid guidance that always inspired and motivated me. It is him who always welcomed me to his office with a smile and enlightened me with other fields of research and research ethics. I have been very fortunate to work with such a learned and outstanding individual for past three years.

I gratefully acknowledge the financial support of Fulbright Program for my graduate studies at CWRU.

I would like to thank faculty and staff of Epidemiology and Biostatistics department for their time and guidance; especially Dr. Paul Jones, Dr. Jaffery Albert, Dr. Pingfu Fu, Dr. Mark Schluchter, Dr. Jill Barnholtz-Sloan, Dr. Abdus Sattar, Dr. Jiayang Sun, Dr. Daniel Tisch, Dr. Omar De la Cruz, Alicia Boscarello, Cynthia Moore, Victor Courtney and Alberto Santana.
I would also like to thank Dr. Xiaowei Guan, Ashley Lowry, Seunghee Margevicius, Apoorva Chander and all my classmates, friends and fellows for their support throughout my PhD.

I greatly acknowledge the support of my family in Cleveland, Arlene and Sonny, who welcomed me to their home, the very first day I arrived in Cleveland. I have no words to express my heartiest gratitude to Bill and Joan Wood, they have always been so welcoming, caring and loving that I never felt alone, their presence has always been a source of pleasure. I acknowledge Bok and Heidi Chew for trusting me and providing me an opportunity to lead ISF, which played a significant role in my professional development. I would like to thank Rev. Eileen Vizcaino and Rev. Robert Campbell for their presence and support. I am so thankful to all brothers and sisters, Krufinta Bun, Lingzi Hong, Megan Stonebraker, Leslie Cofie, Ivaylo Ivanov, Van Anh Tran, Dr. En Cheng, Adriel Jebin, Dr. Reena Bapputty, Wayne Gibson, Rosie Varela, Thomas Atta Fosu, Richard Nii Larte, Dawit Mesfin, and last but not least every individual who walked with me on this road to the completion of PhD, for helping, supporting and guiding me.

Last but not least, I would like to pay my highest regards to my parents, Mr. Hanook and Mrs. Charlotte Hanook, my sisters Rubeca, Naumi, Milka, Tamar and my brothers Aroon, Abdoon, Harmoon and my fiancée Shonam George, who have always encouraged and believed in me. Their constant love and support have been most important for me to endeavor for the best in my life.
Analysis of Removable Interaction

Abstract

By

SHAROON HANOOK

Researchers have always aspired to estimate the most useful and parsimonious model, which helps to obtain the most precise and efficient estimates. Once we have a large number of rows and columns in a factorial design, we have many parameters; so the question that arises is: can we obtain better estimates of the parameters of interest by modeling, or by transformation, or by both? In the presence of non-additivity in the model, Interpretation of the model is not very straightforward and it is not easy to describe the model in a most meaningful manner. The scale of measurement can cause interaction and it can be efficiently detected by using any auxiliary information in terms of weights (if at all available) regarding the row and column factor levels. We propose a test that uses available appropriate weights for Tukey’s one degree of freedom test known as weighted Tukey’s test for removable interaction to isolate removable interaction effect from the essential interaction effect. This test uses row and column level weights to effectively detect removable interaction effects in the presence of negligible essential interaction effect. We also proposed a modification of the weighted Tukey’s test due to Rasch et al (2009), using non-linear regression and testing overall model against the sub-model (i.e. no interaction model) by a likelihood ratio
test, and showed with simulation that proposed modification is better than Tukey’s, weighted Tukey’s and un-weighted modified Tukey’s test. Better estimates can also be obtained by a transformation, in this dissertation we are proposing an empirical monotonic transformation motivated by Fisher (1950) to remove non-additivity from a two-way analysis of variance model, given the condition that in the presence of negligible essential interaction, there exists some non-additivity in the model that is either completely or partially removable. This transformation provides satisfactory results in the case of different interaction structure types by eliminating removable interaction and making the model more parsimonious.
Chapter 1

Introduction

Box famously wrote, "All models are wrong, but some are useful". Researchers have always aspired to estimate the most useful and parsimonious model, which helps to obtain the most precise and efficient estimates. Better understanding of phenomena leads to the inclusion of more variables in models, resulting in larger models. Improvement in the model building tools and statistical software makes it possible to study complicated models.

Once we have a large number of rows and columns in a factorial design, we have many parameters; so the question that arises is: can we obtain better estimates of the parameters of interest by modeling, or by transformation, or by both? In practice, data are measured on any convenient scale, and such measurements obtained on a convenient scale are used to estimate the population parameters. Because of using an arbitrary scale of measurement, we may find many extra parameters due to the multiple factors and interaction terms involved. Mather (1949) and Elston (1959) have pointed out we have no a priori reason to believe that one scale is more appropriate than any other for measuring a character of a living organism. The scale we use for analysis should therefore be arrived at by empirical means. It should be one that facilitates both the analysis of the data and interpretation and use of the resulting statistics. For these reasons a scale on which we have additivity is desirable. In many cases, if we can assume that the model shows additivity, so that interaction terms are not necessary, then the estimates we
obtain are of far greater value to us (Elston, 1961). Changing a scale in such a way that interaction terms are negligible can help make the model more parsimonious. If we believe two factors are involved in the production of an outcome, we must believe that there is some physical interaction involved; but we should want, if at all possible, to model the outcome without the need for statistical interaction, so that the joint action is more parsimoniously described.

Statistically, the term interaction is defined as a departure from additivity in a linear model on a selected scale of measurement (Scheffé, 1959). Statistical interaction corresponds to higher order effects, and including interaction terms in a model is equivalent to including higher degree polynomial terms; large interactions may induce a curvature effect and the resulting model is non-additive in its effects (Box, 1991). On a selected scale of observations, interaction terms may be required to obtain the best fitting model that can explain the curvature effects. Certain types of interactions may be eliminated by a transformation of the outcome so that the resulting model is additive in the effects (Box, 1991; Elston, 1961). Such interactions are referred to as removable interactions.

Tukey (1949) introduced a test of non-additivity for one observation per cell in a factorial design, to help detect the non-additivity that may arise due to the joint action of variables in the model. Let us consider a model with two factors, one factor represented by rows called R, and another represented by columns called C, such that in this two-way array model a response in the $i - th$ row and $j - th$ column is modeled as:

$$y_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ij} \quad \left\{ i = 1, \ldots, r, \quad j = 1, \ldots, c \right\}$$  \hspace{1cm} (1.1.1)
where
\[ \sum_{i=1}^{r} \alpha_i = 0, \sum_{j=1}^{c} \beta_j = 0, \sum_{i=1}^{r} \gamma_{ij} = 0 \ \forall \ j \ \text{and} \ \sum_{j=1}^{c} \gamma_{ij} = 0 \ \forall \ i \]

and the \( \epsilon_{ij} \) are normally distributed independent random variables with mean zero and variance \( \sigma^2 \), i.e. \( \epsilon_{ij} \sim N(0, \sigma^2) \). In model (1.1.1), \( \alpha_i \) represents the \( i-th \) row effect, \( \beta_j \) represents the \( j-th \) column effect, and \( \gamma_{ij} \) denotes an interaction effect of the \( i-th \) row and \( j-th \) column in the \( ij-th \) cell.

When we have a situation of only one observation per cell, the conventional linear model theory (i.e. using ANOVA) may not be feasible, and one must use an interaction sum of squares to estimate (conservatively) the error variance (Johnson & Graybill, 1972). Discussion of interaction falls under three broad headings, namely its definition, detection and interpretation. In the next sections we will concisely discuss these three headings, but the first and third must necessarily be discussed together.

### 1.1 Definitions and Interpretation of Interaction

The concept of interaction, and certainly the very word itself, is extensively used in scientific discussion. Widespread use of this term “interaction” has created confusion among researchers across different fields. Owing to the advancement in interdisciplinary research, it is very important to have some consistency in the usage of terminology. Wang et al (2010) tried to clarify this confusion very eloquently, especially in genetics, owing to the use of some alternative terms like “epistasis”. Bateson, in 1909, used this term to describe
instances in which an effect of a particular genetic variant at one locus was masked by a variant at another locus, so that the variation of a phenotype due to a genotype at one locus was only apparent among those with certain genotypes at the second locus (Cordell, 2002; Vanderweele, 2010). Cordell et al (2005) mentioned that Bateson employed the term, “epistasis”, to explain a masking effect, but Fisher (1918) used a similar term, “epistacy”, to explain statistical interaction. Fisher used this term to define the departure from additivity such that $\gamma_{ij} \neq 0$ in equation (1.1.1). Cordell et al (2005) further argued that Fisher’s “epistacy” quickly evolved into “epistasis”, so that in the modern genetics literature the coexistence of two different meanings creates ambiguity (Phillips, 2008; Vanderweele, 2010; Wang et al, 2010). To lessen such confusion in terminology, Satagopan et al (2013) further described different types of epistasis by mentioning alternative terms used in the genetics literature: physical interaction as functional epistasis, masking effect as compositional epistasis, and departure from additivity (i.e. statistical interaction) as statistical epistasis.

Haldane (1946), Satagopan et al. (2013) and Wang et al. (2010) explained different types of gene-environment interactions, describing types of interactions in the case of a $2 \times 2$ contingency table.

Now we will discuss types of interaction in the case of the two-way classification model given in equation (1.1.1) where a row factor A has 2 levels and a column factor B has 3 levels. Let $\alpha_i$ represent row factor effects, $\beta_j$ represent column factor effects and the outcome summaries, such as the cell means, are $\zeta_{ij}$. The ordering of the $\zeta_{ij}$ induces different types of interactions (Satagopan et al, 2013;
Tukey, 1949). There are six possible ordering of $\zeta_{ij}$, providing four major types of interaction, denoted by A, B, C and D; types A and D each contain two subtypes. Using the proposition put forward by Mandel (1961), $b_i$ represents the slope of the $i$-th level of row factor A across different levels of column factor B. Hence four types of interactions can be defined as:

1.1.1 Type A interaction

This occurs when the outcome summary $\zeta_{ij}$ at $B_j \geq B_{j'}$ for all levels of $A_i$ when $j > j'$, and similarly for summary measures under all possible $A_i$, such that the following conditions hold true (i) $\zeta_{11} \leq \zeta_{12} \leq \zeta_{13}$ (ii) $\zeta_{21} \leq \zeta_{22} \leq \zeta_{23}$ (iii) $\zeta_{11} \leq \zeta_{21}, \zeta_{12} \leq \zeta_{22}, \zeta_{13} \leq \zeta_{23}$ and (iv) $\zeta_{13} \leq \zeta_{21}$. This type of interaction can be removed, and is shown in Fig (1.1a).

1.1.2 Type B Interaction

An interaction will be called a Type B interaction when the outcome summary $\zeta_{ij}$ at $B_j \geq B_{j'}$ for some levels of $A_i$ when $j > j'$, but their order is reversed (i.e. $B_j \leq B_{j'}$) for some levels of $A_i$ when $j > j'$. But in a given range of observations, any of the lines representing different row (column) effects at each level of column (row) do not intersect each other, i.e. (i) $\zeta_{11} \leq \zeta_{12} \leq \zeta_{13}$ (ii) $\zeta_{21} \geq \zeta_{22} \geq \zeta_{23}$ (iii) $\zeta_{11} \leq \zeta_{21}, \zeta_{12} \leq \zeta_{22}$ and $\zeta_{13} \leq \zeta_{23}$. These interactions are difficult to detect because, within a given range, they will not be very visible. Such interactions may be removed, but notice that the slopes have different signs and in this situation an interaction is not completely removable. This can be observed in Fig (1.1b).
1.1.3 Type C Interaction

An interaction will be called Type C interaction when the outcome summary $\zeta_{ij}$ at $B_j \geq B_{j'}$ for all levels of $A_i$ when $j > j'$, and similarly for the summary measures under all possible $A_i$; but in a given range of observations some of the lines representing different row (column) effects at each level of column (row) intersect each other i.e. (i) $\zeta_{11} \leq \zeta_{12} \leq \zeta_{13}$ (ii) $\zeta_{21} \leq \zeta_{22} \leq \zeta_{23}$ (iii) $\zeta_{11} \leq \zeta_{21}, \zeta_{12} \leq \zeta_{22} or \zeta_{12} \leq \zeta_{22}, \zeta_{13} \geq \zeta_{23}$. Type C interaction is similar to type B interaction, with the roles of factor A and factor B exchanged.

1.1.4 Type D Interaction

Type D interaction can be called a non-removable or essential interaction. This type of interaction occurs when there is a cross-over in the lines representing an individual row effect across different columns. Type D interaction does not preserve any type of ordering of outcome summaries but rather depends on a cross-over effect, i.e. (i) $\zeta_{11} \leq \zeta_{12} \leq \zeta_{13}$ (ii) $\zeta_{21} \geq \zeta_{22} \geq \zeta_{23}$ (iii) $\zeta_{11} \leq \zeta_{21}, \zeta_{12} \leq \zeta_{22} or \zeta_{12} \leq \zeta_{22} and \zeta_{13} \geq \zeta_{23}$. It is also known as perfectly antagonistic interaction (Wang et al., 2010), and such a type of interaction cannot be removed by any transformation (Box, 1991).
Figure 1.1: The different types of interactions as classified by Haldane (1946) and further explained by Satagopan et al (2013). Factor A has two levels, $A_1$ and $A_2$; factor B has three levels $B_1$, $B_2$ and $B_3$; and $\zeta_{11}, \zeta_{12}, \zeta_{13}, \zeta_{21}, \zeta_{22}, \zeta_{23}$ are the outcomes in the 6 subclasses of a $2 \times 3$ classification table. In each of the four plots the horizontal axis shows the three values of factor B, and the vertical axis shows the cell values $\zeta_{ij}$. In all four figures, under the different types of interactions the cell values $\zeta_{ij}$ are ordered differently. The black and red lines in the figures above show the outcome as a function of factor A and factor B.

A test statistic that assumes only removable alternatives hypotheses, $A_{RI}$, is likely to be a powerful approach for identifying interactions of type A and certain types of interactions of types B and C but not interactions of type D. A test
statistic under general alternatives, $\Lambda_{ij}$, is likely to be beneficial mainly for identifying interactions of type D.

### 1.1.5 Interpretation of Interaction

As defined in section 1.1, an interaction is said to occur if the separate effects of the factors do not combine additively. González et al (2007) elaborated several different types of statistical interaction and the complexities in interpreting these different types of interaction. Gonzalez et al (2007) and Satagopan et al (2013) both emphasized that any transformation for removing interaction from the model should be one that can be transformed back for interpretation, though transformation may help to improve the presentation of the conclusions and its resulting interpretation. Removable interactions are inconsistent with average effect reversal, as might occur in the case of gender and treatment effect. Whatever the transformation of the measurement scale used, “absence of interaction on a transformed scale eliminates the possibility that a treatment is on an average beneficial for men and on an average harmful for women” (González et al (2007)).

Transformations to remove an interaction from the model have been proposed by Anscombe & Tukey (1963), Brown (1975), Cressie (1978), Kruskal (1965), McNeil & Tukey (1975), Moore & Tukey (1954) and Schlesselman (1973) and will be discussed in section 4.1.

When interaction effects are present in the data it is not easy to interpret the main effects; as Wang et al (2010) say, “in this case the interpretation of main effects is either misleading or incomplete”. We can say that the
interpretation of interaction effects plays a very significant role in determining a transformation such that, on the transformed scale, the results may be better interpreted and one can make sense out of it.

1.2 Detecting interaction

In any biological process one may know that there is biological or physical interaction; the existence of the main effects in a model automatically implies an existence of biological interaction, but for appropriate modeling the presence or absence of a multiplicative or other joint effects also needs to be determined. Non-additivity often arises because of the scale of measurement; hence one should carefully detect these non-additivities in the model and remove them by changing to an appropriate scale. Our major focus is to propose a transformation to eliminate removable interaction from the model; but this is a secondary step in the process - first we have to make sure that interaction is present in the model we are using and that it is removable. Various methods to detect interaction have been discussed in the case of one observation per cell (Johnson & Graybill, 1972; Mandel, 1959, 1961, 1969, 1971; Tukey, 1949, 1955; Tusell, 1990) using different criteria and techniques. Milliken & Graybill, (1971) also discussed one of the extensions of Tukey’s test.

Two common things should be noted about the tests discussed by Mandel (1961) and Tukey (1949): firstly, all these tests are based on models where an interaction is a known function of the row and column effects, and hence they probably have reasonably good power for these models; secondly, these papers
mentioned that the main problem is to devise a test to determine whether there is an interaction in the model. If a test indicates a presence of interaction, the recommendation has generally been to try to find a suitable transformation of the data so that the resulting model is additive. Then a conventional analysis can be performed (Johnson & Graybill, 1972). Elston (1961) and Cox (1984) added to this notion by emphasizing that this transformation should be monotonic. These tests will be discussed in section 2.1.

1.3 Essential and Removable Interaction

Equation (1.1.1) explains an overall model, considering that there is interaction in the model. Now interaction is usually defined as departure from additivity of effects on a specific outcome scale. If a monotonic transformation exists that induces additivity, then the interaction is called a removable interaction; otherwise the interaction is essential (Wu et al., 2009).

The problem of interest here is, if there is interaction, does there exist a suitable transformation of a scale of measurement so that, on a new scale, there will be additivity? The practical problem is complicated by the fact that we do not know the interaction parameters but have only estimates, which we know are in error. Given there is negligible essential interaction, if an interaction in equation (1.1.1) is removable, then there exists a monotonic transformation of a response variable \( y_{ij} \) say \( f(y_{ij}) \) such that equation (1.1.1) can be re-defined as an additive function:

\[
E(f(y_{ij})) = \mu + \alpha_i + \beta_j \tag{1.1.2}
\]
Equation (1.1.2) is an additive model, which is obtained after eliminating removable interaction from the model. If \( f\{y_{ij}\} \) is linear in \( y_{ij} \), then the model in equation (1.1.2) can be fitted to attain parsimony and a precise fit. If \( y_{ij} \) is a non-linear function of \( f\{y_{ij}\} \), then a quadratic approximation may provide a better fit than an additive linear model, i.e. we may have

\[
y_{ij} \approx \eta_0 + \eta_1 f\{y_{ij}\} + \eta_2 [f\{y_{ij}\}]^2,
\]  

(1.1.3)

where \( \eta_0, \eta_1 \) and \( \eta_2 \) are unknown parameters. Further, suppose the right hand side is a strictly monotonic function of \( f\{y_{ij}\} \) in the range \([\text{Min} f\{y_{ij}\}, \text{Max} f\{y_{ij}\}]\), assuming that \( f\{y_{ij}\} \) lies within a definite range, determined by the extent of the available data (Elston, 1961; Satagopan & Elston, 2013). Provided those data lie within the range where this quadratic polynomial is strictly monotonic, there exists a single-valued function such that, when applied to the data, the expectations of the transformed data satisfy the condition for additivity.

Assuming that the monotonicity condition holds, we may further approximate equation (1.1.3) as follows (Elston, 1961):

\[
E(y_{ij}) \approx \mu + \alpha_i + \beta_j + \lambda \times \alpha_i \times \beta_j.
\]  

(1.1.4)

The last term on the right hand side of equation (1.1.4) is a quadratic term \( y_{ij} = \lambda \times \alpha_i \times \beta_j \) that characterizes the interaction effects. The parameter \( \lambda \) is a scalar that quantifies the non-additivity, and the score test for \( \lambda = 0 \) is essentially Tukey’s (1949) one degree of freedom test for nonadditivity (Cox, 1984). When there are main effects (i.e., \( \alpha_i \neq 0 \) and \( \beta_j \neq 0 \) for at least one \( i \) and \( j \)), there is no interaction if and only if \( \lambda = 0 \), and then equation (1.1.4) becomes an additive
When $\lambda \neq 0$, then the model is non-additive (Elston, 1961; Satagopan et al. 2013). It may still not be possible to eliminate the interaction completely. So we can call that portion of the interaction that can be removed by a monotonic transformation a “removable interaction”, also known as ordinal interaction, and the remaining portion that cannot be removed, even after a transformation, an “essential interaction”.

When the null hypothesis for an additive model is not accepted, i.e. $\lambda \neq 0$, under model (1.1.4) “because monotonicity holds, the model is intrinsically linear, and we may empirically look for a transformation of the outcome and then fit an additive model on the transformed scale” (Finney, 1948; Satagopan et al., 2013). While choosing a transformation, one should keep in mind that a transformation should be strictly monotonic, so that the results can be back-transformed for interpretation on the original scale; this removed interaction will reappear in the model on the original scale (Satagopan et al., 2013; Wang et al., 2010).

1.3.1 Testing for removable interaction

Elston & Bush (1964) discussed a class of hypotheses that can be tested when there are interactions in an analysis of variance model. They suggested that even if we assume that interactions are small, we still have to test the hypothesis of no interaction. If the number of degrees of freedom for interaction is not much smaller than that for the residual sum of squares, it may well be advantageous to assume an additive model. Satagopan et al. (2013) further elaborated a case of testing interactions in the model, emphasizing the strict
monotonicity condition. Monotonicity induces a useful class of alternatives for evaluating interactions, namely removable interactions (Satagopan et al, 2013). Under the assumption of monotonicity, we can test the null hypothesis of no interaction against an alternative that an interaction exists and is removable, instead of considering a general family of alternatives; that is, we can test $\gamma_{ij} = 0$ for all $i$ and $j$ against the alternative $\gamma_{ij} \neq 0$ for some $(i, j)$ and that it takes the form $\gamma_{ij} = \lambda \times \alpha_i \times \beta_j$. When we assume that the main effects are not all zero, this is equivalent to testing $H_0: \lambda = 0$ against the alternative $H_A: \lambda \neq 0$. When in the presence of negligible essential interaction the $H_0$ is rejected, this would be an indication that there is removable interaction, suggesting that it may be sufficient to fit an additive model to the data. However, when $H_0$ is not rejected, this would mean that there is no removable interaction; either there is no interaction at all or there is only non-removable interaction, i.e. essential interaction (possibly because the monotonicity assumption fails).

Tukey’s one degree of freedom test for non-additivity is a test of removable non-additivity (interaction). Later in section 2.1, we will discuss different modifications of Tukey’s test in detail. We will use the terms interaction and non-additivity interchangeably.
1.4 Objective of the study

Our ultimate objective is an efficient estimation of the average responses using a parsimonious additive model in a two-way classification model. In the presence of negligible essential interaction, if there is a removable interaction in the model, then how it is to be detected; and upon detection of interaction, how to remove it? Our main focus is in evaluating and removing from the model (if possible) non-additivity, known as statistical interaction, to make the model more parsimonious and hence obtain better parameter estimates. When there is evidence of removable interaction, it is pragmatic to eliminate it via a transformation of the outcome to attain model parsimony and a precise fit, especially if all the interactions are removable.
Chapter 2

Weighted Tukey's test for removable interaction

2.1 Non-additivity and tests for non-additivity

Non-Additivity can be described as a joint effect of the row and column effects that brings a curvature effect into the two-way classification model, which makes the model complicated and less parsimonious, making it hard to interpret for practical applications. In this section we will discuss methods to detect non-additivity in the two-way classification model, as given in equation (1.1.1).

Methods to detect non-additivity in the two-way classification model are mainly categorized into three broad categories based on the structure of the interaction present in the data. These methods consider three different interaction structures, which are defined as:

I. Interaction depends on both or either of the main factor effects.

II. Interaction does not depend on any of the main effects.

III. Structure of interaction is unknown (unstructured).

Firstly we will discuss methods developed to test non-additivity in the case of interaction structure I and then we will discuss methods for interaction structures II and III.

Tukey (1949) introduced a test for non-additivity in a two-way classification model with one observation per cell. Though Tukey did not propose any specific type of interaction structure, later research showed that Tukey's test
has relatively good power when an interaction is proportional to the product of the row and column effects i.e. \( \gamma_{ij} = \lambda \times \alpha_i \times \beta_j \), which is represented in equation (1.1.4), but the power of Tukey’s test in a case of more general interaction, i.e. structure II and III, is very poor. Later, Ghosh & Sharma (1963) showed that in a situation where an interaction is a function of the product of row and column factor effects, Tukey’s test has better overall performance. Ward and Dick (1952) used the model (1.1.4) and considered an iterative solution to the normal equations resulting from the likelihood function; they showed that the sum of squares due to the hypothesis of no interaction (i.e. \( \lambda = 0 \)) after one iteration is the same as the sum of squares for non-additivity in the test proposed by Tukey (1949). A multivariate extension of Tukey’s (1949) test has been obtained by McDonald (1972), and Fujikoshi (1993) discussed it in the case of unbalanced data.

Mandel (1961) developed an extended model in the two-way analysis of variance with one observation per cell assuming an interaction structure in (1.1.4) that is a function of either or both of the row and column effects. He assumed an interaction term to be \( Q_i \beta_j \), defining \( Q_i \) to be a parameter of the \( i \) – \( th \) row, and \( \beta_j \) to be the main effect for column \( j \). He developed a method to test the interaction effect in a two-way classification by modeling and separating degrees of freedom for the slopes \( \beta_j \) and then testing their concurrence. Tukey’s (1949) test is a special case of Mandel (1961). One advantage of Mandel’s (1961) method is that it can also detect non-additivity that occurs in a few rows or columns (sparse interaction). Later, Mandel (1969) discussed the partition of interaction effects into a sum of the terms obtained as the product of the row factor and column factor separately.
Mandel’s (1969) method is derived from principal components analysis, where he used eigenvalues to estimate the partitioned sum of squares for the interaction. Recently, Rasch et al., (2009) modified Tukey’s (1949) test to improve the performance of Tukey’s (1949) and Mandel’s (1961) test in the case of a more general form of interaction. The idea behind Rasch’s modification was to fit a model as in equation (1.1.4) by a non-linear regression and test it against an additive model, as given in equation (1.1.2), by a likelihood ratio test. We will discuss Rasch’s modified Tukey’s test in Chapter 3.

Starting from here we will discuss the non-additive model with interaction structure II, for which the non-additive effect does not depend on any of the main effects and is given as:

\[ y_{ij} = \mu + \alpha_i + \beta_j + k \varphi_i \psi_j + \varepsilon_{ij} , \]

\[
\sum_{i=1}^{r} \alpha_i = 0, \sum_{j=1}^{c} \beta_j = 0, \sum_{i=1}^{r} \varphi_i = 0, \sum_{j=1}^{c} \psi_j = 0 \text{ and } \sum_{i=1}^{r} \varphi_i^2 = \sum_{j=1}^{c} \psi_j^2 = 1 .
\]

Mandel (1971) proposed a new analysis of variance model for non-additive data that is based on the model in (2.1.1), and partitioned any interaction effect into the product of the levels of row and column factors. He considered the interaction structure II where an interaction effect \( y_{ij} \) does not depend on any of the main effects, and is given as:

\[ y_{ij} = \sum_{k=1}^{p} \lambda_k \varphi_{kj} v_{kj} , \]

where “\( k \)” represents the rank of the interaction matrix.

The model discussed by Mandel (1971) is of the form given as follows:

\[ y_{ij} = \mu + \alpha_i + \beta_j + \sum_{k=1}^{p} \lambda_k \varphi_{kj} v_{kj} + \varepsilon_{ij} . \]
Model (2.1.3) is based on a singular value decomposition of the interaction matrix, when an interaction effect can be represented by (2.1.2), \( \Gamma = \{ Y_{ij} \} \) is an interaction matrix of order \( (r \times c) \), and \( \boldsymbol{\varphi} = \{ \varphi_{ki} \} \) and \( \boldsymbol{v} = \{ v_{kj} \} \) are orthogonal matrices of orders \( (r \times p) \) and \( (c \times p) \) respectively. He proposed a test statistic based on the characteristic roots \( \lambda_k \) of the interaction matrix \( \Gamma \). Snee (1982) has indicated that often only one of the \( \lambda \varphi \psi \) terms is needed to describe the interaction in Mandel's model. Mandel's model as given in (2.1.1) contains special cases such as Mandel (1961) and Tukey (1949). Snee (1982) used the model in (2.1.3) to study interaction effects and non-homogeneous variances in a two-way table with one observation per cell. He developed a method to isolate an effect of non-additivity and a non-homogeneous variance in the model.

Johnson & Graybill (1972) derived maximum likelihood estimates of the parameters of the product model given in equation (2.1.1). They also discussed an interaction structure II, when the interaction may not be a function of row and column effects. In situations where interaction is present in only one, two, or a few scattered cells (interaction outliers), which is only found for one or two treatments, then a transformation may do more harm than good. Also, in these cases there may be an unbiased estimate of \( \sigma^2 \) in terms of the original units and, if so, it is desirable to determine it. Moreover, they defined a likelihood ratio test to test the significance of the departure from no interaction, i.e. \( H_0: \lambda = 0 \). They also derived a hypothesis of no treatment effect in the presence of interaction and developed distributional properties for the test of no interaction by deriving maximum likelihood estimates of the parameters in the model (2.1.1). Tusell (1990) suggested that Johnson and
Graybill’s (1972) test performs the best in a case when the interaction matrix has rank of order one.

From this point forward we will discuss tests derived for interaction structure III. Tusell (1987) proposed a test for additivity restricted to a few rows or columns of a rectangular un-replicated two-way ANOVA table, testing for parallelism of complete row and columns in a manner reminiscent of standard profile analysis (Tusell, 1990). Later, Tusell (1990) proposed a test of additivity in two-way ANOVA with one observation per cell. Tusell used a likelihood ratio test of sphericity to test the non-additivity in the model. The notion motivating this test is
to carry out successive linear operations on the rows and columns of data in two-way ANOVA. Normally we assume additivity when a test of sphericity is carried out, but Tusell reasoned that if we assume sphericity, then a test for additivity could be applied using the LR criterion. Tusell’s test statistics for \( c \leq r \) is given as:

\[
\psi = \frac{|Z'Z|}{\text{trace} \left( \frac{Z'Z}{(c-1)} \right)^{c-1}} = \frac{\lambda_1 \lambda_2 \ldots \lambda_{c-1}}{[(c-1)^{-1} \Sigma_k l_k]^{c-1}},
\]

where the \( l_k \) denote non-zero eigenvalues of \( Z'Z \), \( Z = X'U_b \) and \( X = Y'U_a \), where \( Z \) is a matrix of order \( (r-1) \times (c-1) \); moreover, \( U_a \) and \( U_b \) can be any \( r \times (r-1) \) and \( c \times (c-1) \) orthogonal matrices of contrast coefficients.

Boik (1993b) proposed a test named “Locally Best Invariant (LBI) test of additivity” which does not assume any particular structure of interaction, making it less subject to model misspecification. He gives a limiting distribution of the LR sphericity criterion under non-additivity. Boik (1993a) carried out a simulation study to compare three invariant tests of additivity in a case of two-way
classification with one observation per cell. He compared Tusell's (1990) test with Johnson & Graybill's (1972) LR test against a rank-1 alternative and with Boik's (1993a) LBI test. He deduced that Tusell's LR Sphericity test is the most powerful test among the three when an interaction matrix has multiple non-zero eigenvalues that are comparable in magnitude. Boik (1993a) suggested that the LBI test is the most powerful invariant test when the magnitude of the non-additivity is small. He proposed a simple index of non-additivity magnitude, which is the squared scaled (by σ) norm of the interaction given as:

\[ \Lambda = \frac{1}{\sigma^2} \sum_{k=1}^{p} \lambda_k . \]  
(2.1.5)

For (2.1.5), even if a value of Λ is large, a poor fit translates into low power. The test statistic (2.1.6) proposed by Boik (1993a) is invariant under scalar multiplication and orthogonal rotation of the residual matrix. Consequently, invariant tests depend on the data through the eigenvalues of \( Z'Z / \text{tr}(Z'Z) \), which are defined as \( L_1 > L_2 > \ldots > L_p > 0 \), hence

\[ L_k = \frac{i_k}{l_1+l_2+\ldots+l_p} . \]  
(2.1.6)

Boik (1993b) showed that if the coefficient of variation among the eigenvalues of \( \text{E}(Z)'\text{E}(Z) \) is not too small, then the locally best invariant test of additivity is to reject \( H_0 \) for small \( \hat{\epsilon}_{p,q} \).

\[ \hat{\epsilon}_{p,q} = \frac{[\text{tr}(Z'Z)]^2}{p \ 	ext{tr}([Z'Z]^2)} = \left[ p \ \sum_{k=1}^{p}(L_k - p^{-1})^2 + 1 \right]^{-1} , \]  
(2.1.7)

where \( Z = \{ e_{ij} \} \) in the additive model \( y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij} \) and \( p = \min(a - 1, b - 1) \), with the restriction of being greater than or equal to 2, and \( q = \max(a - 1, b - 1) \).
For $p = 2$ and $q = 3$ John (1972) gave the exact null distribution of

$$T_{p,q} = \frac{1}{p-1}\left(\frac{1}{\hat{c}_{p,q}} - 1\right)$$

and, for $p = 2$, $T_{2,q} \sim beta \left(1, \frac{1}{2}(q - 1)\right)$.

“The test statistic given in (2.1.7) is a monotonic decreasing function of the coefficient of variation among the eigenvalues of $Z'Z$. A large coefficient of variation among the sample eigenvalues is evidence against additivity” (Alin et al (2006)).

2.2 A new test to detect removable interaction

In this section we are proposing a modification of Tukey's (1949) one degree of freedom test for non-additivity. This provides expressions for sums of squares of removable and essential interactions for a two-way classification model. It tests the hypotheses of main effect and removable interaction in the presence of non-existent or negligible essential interaction, when each level of the row factor and each level of the column factor is assigned some weight based on some auxiliary information such as: the population proportions or some known experimental circumstance. In practice we often collect data keeping convenience of the sampling in mind, and also perform analyses completely ignoring the fact, if there are any differences in the existing population proportions or if there is any other auxiliary information available for the levels of rows and columns. This raises concerns about the generalizability of the results to the whole population. Therefore, in the following sections we are proposing the use of auxiliary information as weights (if at all available) to test the non-additivity in the model and to draw inferences for the row and column effects in a two-way classification model. We are discussing the
traditional approach to decomposing the cell means into an overall mean, main and interaction effects, in which these effects satisfy certain restrictions.

2.2.1 Two-way classification model

A fixed effects model for a two-way classification design may be given as:

\[ y_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ij} \]

where the \( \epsilon_{ij} \) are independently distributed as normal with mean zero and a common variance \( \sigma^2 \), i.e.: \( \epsilon_{ij} \sim N(0, \sigma^2) \). \( \mu \) is a constant known as a grand mean, \( \alpha_i \) is considered to be a row main effect defined as an effect due to the \( i - th \) level of the row factor. Similarly, \( \beta_j \) is considered to be a column main effect defined as an effect due to the \( j - th \) level of the column factor, and \( \gamma_{ij} \) is an interaction effect of the \( i - th \) row level and \( j - th \) column level. Moreover, we assume that the observations in the \( ij - th \) cell are a random sample from a population corresponding to the cell.

2.2.2 Interaction structure

The proposed method considers an interaction structure I, i.e. \( \gamma_{ij} = \lambda \alpha_i \beta_j \), where an interaction is a function of either of the row and column factor effects. Considering this two-way classification with interaction structure I, the model can be given as:

\[ y_{ij} = \mu + \alpha_i + \beta_j + \lambda \alpha_i \beta_j + \epsilon_{ij} \]

\( i = 1, 2, ..., r \)

\( j = 1, 2, ..., c. \)  

(2.2.1)
An interaction term in (2.2.1) is a function of a real constant $\lambda$, row effect $\alpha_i$ and column effect $\beta_j$. Johnson & Graybill (1972) explained that in the case of no or non-significant removable interaction the value of $\lambda^2$ will be small resulting in the sum of squares for removable interaction being small.

### 2.2.3 Estimation of parameters

There are two approaches to performing an analysis of variance:

1. Decompose the cell means into an overall mean, main effects and interaction effects in which the effects satisfy certain restrictions.

2. Describe some common hypothesis in terms of $\hat{y}_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$ and to make statistical inference on the hypothesis.

Here we will use the first approach where the overall model with the usual assumptions of normality of $\varepsilon_{ij}$ can be given as in (2.2.1):

$$y_{ij} = \mu + \alpha_i + \beta_j + \lambda \alpha_i \beta_j + \varepsilon_{ij} \quad \{i = 1, 2, ..., r\} \quad \{j = 1, 2, ..., c\}.$$

This model is over-parameterized and constraints are needed to obtain unique estimates. It is customary to impose the following restrictions on the parameters to obtain unique solutions.

$$\sum_{i=1}^{r} \alpha_i = 0, \sum_{j=1}^{c} \beta_j = 0, \sum_{i=1}^{r} \gamma_{ij} = 0 \quad \forall j \quad \text{and} \quad \sum_{j=1}^{c} \gamma_{ij} = 0 \quad \forall i$$

This is equivalent to defining the estimated effects as follows:

$$\mu = \bar{y}. \quad \alpha_i = \bar{y}_i - \bar{y}. \quad \beta_j = \bar{y}_j - \bar{y}. \quad \gamma_{ij} = y_{ij} - \bar{y}_i - \bar{y}_j + \bar{y}. \quad (2.2.2)$$
where
\[
\bar{y}_r = \frac{\sum_{i=1}^r \sum_{j=1}^c y_{ij}}{rc} \quad (2.2.3)
\]
\[
\bar{y}_c = \frac{\sum_{j=1}^c y_{ij}}{c} \quad (2.2.4)
\]
\[
\bar{y}_r = \frac{\sum_{i=1}^r y_{ij}}{r} \quad (2.2.5)
\]

Scheffé (1956) introduced restrictions with non-negative weights \(u_i\) and \(v_j\) such that \(\sum_i u_i > 0\) and \(\sum_j v_j > 0\). In the case where weights are introduced, the corresponding restrictions in equation (1.1.1) turn out to be as follows:
\[
\sum_{i=1}^r u_i \alpha_i = 0, \quad \sum_{j=1}^c v_j \beta_j = 0, \quad \sum_{i=1}^r u_i y_{ij} = 0 \quad \forall \ j \text{ and } \sum_{j=1}^c v_j y_{ij} = 0 \quad \forall \ i \quad (2.2.6)
\]

“In order to get exact tests and confidence intervals concerning the main effects it is generally necessary with the fixed effects model to assume that there are no non-additivities” (Scheffé, 1959). The overall mean \(\mu\) can be defined and estimated using weighted least squares estimation by minimizing \(Q_1\) given as:
\[
Q_1 = \sum_{i=1}^r \sum_{j=1}^c u_i v_j (y_{ij} - \mu)^2 . \quad (2.2.7)
\]

Minimizing \(Q_1\) with respect to \(\mu\) provides a unique estimate for the weighted grand mean, where \(u_i\) represents weights for the \(i\) – \(th\) row and similarly \(v_j\) represents weights for the \(j\) – \(th\) column. An estimate of the grand mean can be given as:
\[
\hat{\mu} = \frac{\sum_{i=1}^r \sum_{j=1}^c u_i v_j y_{ij}}{\sum_{i=1}^r \sum_{j=1}^c u_i v_j} . \quad (2.2.8)
\]

An estimate of the row effects \(\alpha_i\) can be obtained by minimizing \(Q_2\) given below:
Equating the partial derivative of (2.2.9) with respect to $\alpha_i$ to zero and using the constraints given in (2.2.6) will provide a unique estimate of the row effect as:

$$\hat{\alpha}_i = \frac{\sum_{j=1}^c v_j y_{ij}}{\sum_{j=1}^c v_j} - \hat{\mu} \tag{2.2.10}$$

An estimate of the column effects $\beta_j$ can be obtained by minimizing $Q_2$ with respect to $\beta_j$. Equating partial derivative of (2.2.9) to zero and using the constraints given in (2.2.6) will give a unique estimate of the column effects as:

$$\hat{\beta}_j = \frac{\sum_{i=1}^r u_i y_{ij}}{\sum_{i=1}^r u_i} - \hat{\mu} \tag{2.2.11}$$

Estimates of row effects in (2.2.10) and column effects in (2.2.11) are the deviations of weighted row and column means from the grand weighted mean.

In the case of a single observation per cell, the usual F-test for interaction under the null hypothesis cannot be applied because there are no degrees of freedom left for a within-cell sum of squares. So, to obtain a removable interaction sum of squares we use Tukey’s method to isolate a single degree of freedom, where each row and each column has certain weights attached to it. Considering an interaction structure defined in model (2.2.1), the $E(y_{ij})$ are not linear functions of the parameters, but let us pretend that $\hat{\mu}, \hat{\alpha}_i$ and $\hat{\beta}_j$ are known. Under this situation we can easily obtain a least squares estimate of the coefficient $\lambda$ by minimizing

$$Q_3 = \sum_{i=1}^r \sum_{j=1}^c u_i v_j (y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \lambda \hat{\alpha}_i \hat{\beta}_j)^2, \tag{2.2.12}$$
taking partial derivative with respect to $\lambda$ and then equating them to zero. Using the restriction given in (2.2.6) an estimate of $\lambda$ can be given as:

$$
\hat{\lambda} = \frac{\sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j \bar{a}_i \bar{b} \bar{y}_{ij}}{\sum_{i=1}^{r} u_i \bar{a}_i^2 \sum_{j=1}^{c} v_j \bar{b}_j^2}.
$$

(2.2.13)

### 2.2.4 Tests of hypothesis

In this section we will discuss partitioning the total sum of squares for testing main effects. We will also present an expression for the removable interaction using a weighted Tukey’s one degree of freedom test for non-additvity, and will discuss an analysis of variance table to perform tests for the hypothesis of main effects and removable interaction, with the restriction of having the number of either rows or columns to be greater than 2.

Clearly, given a set of appropriate weights, one should only seek a transformation of the data to obtain additivity if any essential interaction is either non-existent or negligible in magnitude. In addition, one would not want to consider a transformation unless there is also significant removable interaction.

An expression for the total sum of squares in the case of weights attached to each row and column can be given as:

$$
Total SS = TSS = \sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j (y_{ij} - \bar{y}_j)^2.
$$

(2.2.14)

Partitioning the total sum of squares given in (2.2.14) into among and within sums of squares gives the following partition of the sums of squares:

$$
\sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j (y_{ij} - \bar{y}_j)^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j (y_{ij} - \bar{y}_j)^2 + \sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j (\bar{y}_j - \bar{y})^2 + \sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j (\bar{y}_i - \bar{y}_i + \bar{y})^2
$$

(2.2.15)
Equation (2.2.15) provides row, column and residual sums of squares with row and column level weights attached. Final expressions for the sums of squares can be given as follows:

Row Sum of Squares \( SS_R = \sum_{j=1}^{r} v_j \sum_{i=1}^{c} u_i (\bar{y}_i - \bar{y})^2 \). (2.2.16)

Column Sum of Squares \( SS_C = \sum_{i=1}^{r} u_i \sum_{j=1}^{c} v_j (\bar{y}_j - \bar{y})^2 \). (2.2.17)

Residual Sum of Squares \( SS_{Re} = \sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j (\bar{y}_{ij} - \bar{y}_i - \bar{y}_j + \bar{y})^2 \). (2.2.18)

Under the null hypothesis of no main effects \( H_R: u_i \alpha_i = 0 \) for all \( I \) and \( H_C: v_j \beta_j = 0 \) for all \( j \), the sum of squares for rows and columns can be given by (2.2.16) and (2.2.17), respectively. Since the null hypothesis for row effects, i.e.: \( H_R: u_i \alpha_i = 0 \), states that \((r - 1)\) linearly estimable functions are zero, the number of degrees of freedom for \( SS_R \) is \((r - 1)\). Similarly, the null hypothesis for column effects, i.e.: \( H_C: v_j \beta_j = 0 \), states that \((c - 1)\) linearly estimable functions are zero, and the number of degrees of freedom for \( SS_C \) is \((c - 1)\). Hence mean squares for row and column effects can be given as:

\[
MS_R = \frac{SS_R}{r - 1},
\]

(2.2.19)

\[
MS_C = \frac{SS_C}{c - 1}.
\]

(2.2.20)

\( MS_R \) in (2.2.19) is distributed as \( \sigma^2 \chi^2_{(r-1)} \) and \( MS_C \) given in (2.2.20) is distributed as \( \sigma^2 \chi^2_{(c-1)} \).
To establish tests of significance for testing the main effects and removable interaction, we first need to find out the exact expression for the residual sum of squares (which we call essential non-additivity) after isolating the sum of squares for removable interaction. The sum of squares for removable interaction when an interaction structure is given by $\gamma_{ij} = \lambda \alpha_i \beta_j$ and the estimate of $\lambda$ is given as in (2.2.13) can be obtained by using the following expression:

$$\sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j \gamma_{ij}^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j (\lambda \hat{\alpha}_i \hat{\beta}_j)^2.$$  \hspace{1cm} (2.2.21)

Substituting the value for $\lambda^2$ in (2.2.21) and solving we obtain an expression for the sum of squares for removable interaction:

$$SS_{RI} = \frac{\left(\sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j \hat{\alpha}_i \hat{\beta}_j \gamma_{ij}\right)^2}{\sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j \hat{\alpha}_i^2 \hat{\beta}_j^2}.$$ \hspace{1cm} (2.2.22)

The expression given in (2.2.22) is distributed as $\sigma^2 \chi^2_{(1)}$. This leaves us with the reduction in the sum of squares of residuals, which we name essential interaction/non-additivity. The sum of squares for essential interaction is distributed as $\sigma^2 \chi^2_{(r-1)(c-1)-1}$. An expression for essential interaction (in this case the remaining residual sum of squares) can be given as:

$$SS_{EI} = SS_{Res} - SS_{RI}.$$ \hspace{1cm} (2.2.23)

The three sets of linear form $\{\hat{\alpha}_i\}$, $\{\hat{\beta}_j\}$ and $\{\hat{\gamma}_{ij}\}$ span three mutually orthogonal spaces, and the three sets are then statistically independent. Under the respective null hypothesis, we have $E(\hat{\alpha}_i) = E(\hat{\beta}_j) = E(\hat{\gamma}_{ij}) = 0$, and hence all the chi-square distributions are central. Hence the $F - test$ of $H_R$ at significance level $\alpha$ consists in rejecting $H_R$ if
\[
\frac{MS_R}{MS_{EI}} > F_{\alpha;(r-1),(r-1)(c-1)-1}.
\] (2.2.24)

In a similar way one would find that the \textit{F-test} of \(H_c\) consists in rejecting \(H_c\) if

\[
\frac{MS_c}{MS_{EI}} > F_{\alpha;(c-1),(r-1)(c-1)-1}.
\] (2.2.25)

The hypothesis of removable interaction given as: \(H_\lambda: \lambda = 0\) can be rejected if

\[
\frac{MS_{RI}}{MS_{EI}} > F_{\alpha;(1),(r-1)(c-1)-1}.
\] (2.2.26)

A complete layout of an analysis of variance table for the weighted Tukey’s test with one observation per cell is given in Table 2.2.
Table 2.2: ANOVA table for the weighted Tukey's test for one observation per cell.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor R</td>
<td>$SS_R = \sum_{j=1}^{c} v_j \sum_{i=1}^{r} u_i (\bar{y}_{i.} - \bar{y}.)^2$</td>
<td>$r - 1$</td>
<td>$SS_R \over r - 1$</td>
<td>$MS_R \over MSE_{EI}$</td>
</tr>
<tr>
<td>Factor C</td>
<td>$SS_C = \sum_{i=1}^{r} u_i \sum_{j=1}^{c} v_j (\bar{y}_{.j} - \bar{y}.)^2$</td>
<td>$c - 1$</td>
<td>$SS_C \over c - 1$</td>
<td>$MS_C \over MSE_{EI}$</td>
</tr>
<tr>
<td>Residual</td>
<td>$SS_{Res} = \sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j (y_{ij} - \bar{y}<em>{i.} - \bar{y}</em>{.j} + \bar{y}.)^2$</td>
<td>$df_{Res} = (r - 1)(c - 1)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Removable Interaction</td>
<td>$SS_{RI} = \frac{\sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j \beta_i \beta_j y_{ij}^2}{\sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j \beta_i^2 \beta_j^2}$</td>
<td>1</td>
<td>$SS_{RI} \over df_{RI}$</td>
<td>$MS_{RI} \over MSE_{EI}$</td>
</tr>
<tr>
<td>Essential Interaction</td>
<td>$SS_{EI} = SS_{Res}^* SS_{RI}$</td>
<td></td>
<td>$df_{EI} = df_{Res} - 1$</td>
<td>$SS_{EI} \over df_{EI}$</td>
</tr>
<tr>
<td>Total</td>
<td>$SS_T = SS_R + SS_C + SS_{RI} + SS_{EI}$</td>
<td>$rc - 1$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.2 provides a basis for testing hypotheses related to row, column and removable interaction effects. The two-way classification table with one observation per cell does not provide any degrees of freedom to perform a test of hypothesis for essential interaction effects. In such a situation an interaction plot provides a tool to decide further action. If the interaction plot shows a sufficient evidence for the existence of interaction of type D, then one would not want to perform a transformation but perform the analysis using a non-additive model.
2.2.5 Missing data

Occasionally in a two or more factor study, the number of factors or even levels increases; a consequence of this proliferation is that some combinations of factor categories may not be observed. Especially in the case of one observation per cell, a single missing data point leads to an empty cell in a two-way table. Here we want to make it very clear that if any cell contains a value “zero” this doesn’t mean an empty cell, rather this should be treated as the normal value of a response variable. An empty cell only occurs when we don’t have any information about the response variable for that specific cell. When one or several cells are empty, an analysis of variance for a two-way classification cannot be performed efficiently. But this doesn’t mean that one should not perform any analysis because of missingness of a few data points.

Several solutions have been presented in the literature, like data imputation and partial analysis, etc. We suggest that, in the case of empty cells, one can preferably perform a partial analysis, because even if information for one cell is missing one can still obtain partial information about interactions by restricting attention to the available data. To perform a partial analysis, we can ignore a complete row or a column that contains an empty cell (the choice completely depends upon the researchers’ decision). Moreover, if we have any auxiliary information regarding the row and column factor then it makes the estimators of effect sizes much more accurate even in the presence of empty cells. For a further discussion about the empty subclass case we refer to Elston and Bush (1964); they discussed the use of appropriate hypotheses in the case of empty subclasses.
2.2.6 Numerical Examples

In this section we will discuss the effect of using weights in different interaction structure types with numerical examples.

2.2.6.1 Interaction structure I

Interaction structure I is the one where the interaction effect is a function of either of the row and column effects. Alin et al (2006) described a dataset, given in table 2.2.1, which is obtained from a study performed by Oztekin et al (2000). “Propofol is an intravenous anesthetic agent with favorable pharmacokinetic properties and some results indicate that it induces a depression in myocardial contractility resulting in a decrease in arterial blood pressure and cardiac output” (Alin et al (2006)). The data in table 2.2.1 are from a study by Oztekin et al (2000) carried out to compare the effect of propofol on cardiac contractile force in normal and hypercholesterolemic hearts. Oztekin et al (2000) compared the effect of different concentrations of drug on left ventricular pressure (LVP). For further details regarding this experiment we recommend Oztekin et al (2000) and Alin et al (2006).

Table 2.2.1: percent change of LVP measure

<table>
<thead>
<tr>
<th>Heart type</th>
<th>Propofol (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25</td>
</tr>
<tr>
<td>Normal Heart</td>
<td>-28.50</td>
</tr>
<tr>
<td>Hypercholesterolemic</td>
<td>-14.99</td>
</tr>
</tbody>
</table>
The data in table 2.2.1 are the percentage change in the LVP measure as a response variable, heart type as a row factor with two levels and propofol as a column factor with three levels. As suggested by Alin et al (2006), and can be seen in figure 2.2.1 that, the interaction structure in this data seems to be of structure I, because all the lines intercept near the baseline and mainly diverge on only one side of the interception. Hence any essential interaction is probably negligible.

![Interaction plot for percentage change of LVP versus Propofol](image)

**Figure 2.2.1: Interaction plot for percentage change of LVP versus Propofol**

Figure 2.2.1 is similar to figure 1 in the Alin et al (2006) paper. We will first perform Tukey's test, and then a weighted Tukey's test to see how including weights in the analysis may change the results. We will observe the effect of appropriate available weights on the testing of three hypotheses $H_R: \alpha_i = 0$, $H_C: \beta_j = 0$ and $H_A: \lambda = 0$. 
The single degree of freedom test for removable interaction proposed by Tukey (1949) was, performed on the data in table 2.2.1. Its results are given in table 2.2.2.

### Table 2.2.2: ANOVA table for Tukey’s test for removable interaction

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F – value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart Type</td>
<td>236.2538</td>
<td>1</td>
<td>236.2538</td>
<td>61.3681</td>
<td>0.0808</td>
</tr>
<tr>
<td>Propofol</td>
<td>1977.2961</td>
<td>2</td>
<td>988.6481</td>
<td>256.8063</td>
<td>0.0441</td>
</tr>
<tr>
<td>Residual</td>
<td>2827.9564</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Removable Interaction</td>
<td>2824.1066</td>
<td>1</td>
<td>2824.1066</td>
<td>733.5760</td>
<td>0.0235</td>
</tr>
<tr>
<td>Essential Interaction</td>
<td>3.8498</td>
<td>2 – 1 = 1</td>
<td>3.8498</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5041.5063</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.2.2 provides the results to conduct tests of hypotheses for row, column and interaction effects, their hypotheses can be given as; $H_R: \alpha_i = 0$, $H_C: \beta_j = 0$ and the hypothesis of removable interaction $H_\lambda: \lambda = 0$. From the results we can notice that the heart type effect is non significant but the propofol effect is significant, and at the same time the joint effects (interaction effects) of heart type and propofol are also significant with a p-value of 0.0235. As we know that the hypothesis $H_\lambda: \lambda = 0$ is a hypothesis for removable interaction, and rejecting this hypothesis means there is removable interaction, this implies that we can eliminate the interaction by transformation. We will discuss the transformation approach later in chapter 4, but here we will apply the method proposed earlier in this chapter called weighted Tukey’s test for removable non-additivity, assuming that
we have some prior information regarding weights available associated with rows and columns. Suppose we know that the treatment, if used at, will be used on a population in which 75% have normal hearts, and 25% hearts are hypercholesterolemic. Similarly, suppose we know from previous experiments that propofol has some side effects and higher levels of propofol will have higher side effects, hence it should be recommended for less people. Keeping this scenario in mind we assign row weights $u_i = (0.75, 0.25)$ and the column weights by $v_j = (0.50, 0.25, 0.25)$. We should keep in mind that weights do not necessarily have sum to one, as Scheffé (1959) noted; all weights have to be positive.

Hence we used weighted Tukey’s test for removable interaction and obtained the following results as given in table 2.2.3.

**Table 2.2.3: ANOVA table for the weighted Tukey’s test for removable interaction**

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F – value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart Type</td>
<td>30.8401</td>
<td>1</td>
<td>30.8401</td>
<td>9.3096</td>
<td>0.2016</td>
</tr>
<tr>
<td>Propofol</td>
<td>40.3921</td>
<td>2</td>
<td>20.1961</td>
<td>6.0965</td>
<td>0.2753</td>
</tr>
<tr>
<td>Residual</td>
<td>265.1634</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Removable Interaction</td>
<td>261.8507</td>
<td>1</td>
<td>261.8507</td>
<td>79.0443</td>
<td>0.0713</td>
</tr>
<tr>
<td>Essential Interaction</td>
<td>3.3127</td>
<td>$2 - 1 = 1$</td>
<td>3.3127</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>336.3957</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From the results in table 2.2.3 we can clearly notice that none of the effects are significant, hence we can conclude that adding these particular weights to
each row and column changes the results. It also provides precise and unbiased estimates of row and column means for these particular weights.

2.2.6.2 Interaction structure II

Interactions of structure II are the ones where interaction does not depend on any of the main effects. Johnson & Graybill (1972) discussed a method to detect removable interaction in the case of interaction structure II. They discussed and considered a two-way classification design and data from Carter et al (1951). Johnson et al (1972) used a part of data from a larger experiment to determine “the effectiveness of blast furnace slags as agricultural liming materials on three types of soils”. We refer to Carter et al (1951) for further explanation of the experiment and dataset. Given that these data are known to have an interaction structure II, Johnson’s Likelihood ratio test performs better than Tukey’s test for removable interaction.

Table 2.2.4: Yields of corn in Bushels per acre

<table>
<thead>
<tr>
<th>Liming treatment</th>
<th>Soil Type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hartsells Very fine sandy loam</td>
</tr>
<tr>
<td>No lime + None</td>
<td>11.1</td>
</tr>
<tr>
<td>Course slag + None</td>
<td>15.4</td>
</tr>
<tr>
<td>Medium slag + None</td>
<td>22.7</td>
</tr>
<tr>
<td>Agricultural slag + None</td>
<td>23.8</td>
</tr>
<tr>
<td>Agricultural limestone + None</td>
<td>25.6</td>
</tr>
<tr>
<td>Agricultural slag + B, Z&lt;sub&gt;n&lt;/sub&gt;, M&lt;sub&gt;n&lt;/sub&gt;</td>
<td>31.2</td>
</tr>
<tr>
<td>Agricultural limestone + B, Z&lt;sub&gt;n&lt;/sub&gt;, M&lt;sub&gt;n&lt;/sub&gt;</td>
<td>25.8</td>
</tr>
</tbody>
</table>
The observations in table 2.2.4 are yields in bushels of shelled corn per acre. Liming treatment represents the row factor with 4000 lb per acre application of limestone or slag except for “no lime”, when no liming material was applied.

Any test for removable interaction assumes the presence of a negligible magnitude of essential interaction. In the case of a single observation per cell, we cannot perform the test of significance of essential interaction. Therefore, we will graphically check this assumption. Figure 2.2.2 is an interaction plot for soil type versus liming treatment. Note that, in the case of no interaction we would expect that all the lines in figure 2.2.2 to be parallel, but this is not actually the case. So, we can say that interaction exists between soil type and liming treatment and most probably there are both essential and removable effects.

Figure 2.2.2: Interaction plot of soil type versus liming treatment
Assuming that any essential interaction effect is negligible, we first will perform Tukey's test, and then weighted Tukey's test, to see how including weights in this analysis could help get much more precise results and obtain different results than originally can be seen by Tukey's method. We will observe the effect of weights on the testing of three hypotheses $H_R: u_i \alpha_i = 0$, $H_C: v_j \beta_j = 0$ and $H_A: \lambda = 0$ using these weights. The analysis of variance for the single degree of freedom for removable interaction was performed on the data in table 2.2.4. The results are given in table 2.2.5.

**Table 2.2.5: ANOVA table for Tukey's test for removable interaction**

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liming treatments</td>
<td>730.1980</td>
<td>6</td>
<td>121.6996</td>
<td>1.4317</td>
<td>0.28644</td>
</tr>
<tr>
<td>Soil type</td>
<td>5706.1800</td>
<td>2</td>
<td>2853.09</td>
<td>33.5654</td>
<td>0.00002</td>
</tr>
<tr>
<td>Residual</td>
<td>936.3533</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Removable Interaction</td>
<td>1.3457</td>
<td>1</td>
<td>1.3457</td>
<td>0.0158</td>
<td>0.90214</td>
</tr>
<tr>
<td>Essential Interaction</td>
<td>935.0076</td>
<td>$12 - 1 = 11$</td>
<td>85.0007</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>7372.7314</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From table 2.2.5 we can conclude that there is no significant removable interaction in the model. Johnson & Graybill (1972) showed that in this situation the Johnson's Likelihood Ratio test i.e., (LR-1) performs better to detect removable interaction from the data. The LR-1 test shows that there is significant removable interaction in this model.
We argue that if we have available weights or any prior auxiliary information for the soil types and/or liming treatment in table 2.2.4, we can improve the estimation and inference related to the study questions. For the sake of clarification, we assume that Lloyd sandy clay loam is not available very often, Hartsells very fine sandy loam is also very difficult to obtain, Norfolk loamy sand is the soil type that is most abundant in the area of application. On the other side, liming treatment can be easily obtained in almost equal proportion. Hence keeping this scenario in view, we can assign equal weights to the row factor given by $u_i = (1,1,1,1,1,1)$ but the column factor, i.e. soil type is given the following weights $v_j = (0.10,0.07,0.83)$. These weights are assumed to show how weights affect the results of Tukey’s one degree of freedom test in the presence of interaction structure II in the case of two-way classification table.

In case of a single observation per cell, we cannot test the hypothesis of essential interaction, so firstly we want to know the effect of these available weights on the interaction by plotting an interaction plot. Given these available weights are appropriate, we can see from figure 2.2.3 that lines for soil types Hartsells Very fine sandy loam and Lloyd Sandy clay loam are almost parallel and there is no essential interaction effect present in the model. Hence, we can proceed with the test of removable interaction using weighted Tukey’s test for removable interaction.
The analysis of variance of the data in table 2.2.4 is performed again and given in table 2.2.6, where the single degree of freedom for removable interaction is the weighted Tukey’s test for removable interaction.

**Table 2.2.6: ANOVA table for the weighted Tukey’s test for removable interaction**

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liming treatments</td>
<td>480.5997</td>
<td>6</td>
<td>80.1000</td>
<td>7.2189</td>
<td>0.0026</td>
</tr>
<tr>
<td>Soil type</td>
<td>965.0676</td>
<td>2</td>
<td>482.5338</td>
<td>43.4876</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Residual</td>
<td>178.9370</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Removable Interaction</td>
<td>56.8821</td>
<td>1</td>
<td>56.4350</td>
<td>5.1264</td>
<td>0.0448</td>
</tr>
<tr>
<td>Essential Interaction</td>
<td>122.0549</td>
<td>12 – 1 = 11</td>
<td>11.0959</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1624.6043</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
From the results in table 2.2.6 we can observe that, just by applying auxiliary information in the form of weights in Tukey’s one degree of freedom test for non-additivity, there is a change in the results and conclusion. We can clearly notice that not only is the removable interaction effect significant, but also the row and column effects are significant. Since the essential interaction effect is non-existent and hypothesis for removable interaction effect is significant, we recommend using transformation to eliminate this interaction effect.
2.3 Extension to more than one observation per cell (Equal n)

In factorial experiments with more than one observation per cell we usually observe and analyze row, column and interaction effects with the usual formulae and interpret the results based on the inferences drawn from the existing model. There has been a lot of discussion regarding the interpretation of the multifactor factorial design model in the presence of interaction. It is suggested that if a significant interaction is present in the model, then main effects shouldn’t be interpreted but if the interaction in the model is non-significant then one can interpret the main effects. Unlike the single observation per cell, in the case of more than one observation per cell we can test a hypothesis for the presence of essential interaction in the model. After partitioning overall interaction into essential and removable interaction effects, if there is negligible amount of essential interaction present in the model then we recommend testing a hypothesis of removable interaction. Given that there is significant removable interaction, we propose finding a transformation to eliminate the removable interaction and then using an additive model for estimation, this will be the topic in chapter 4. If there is a significant essential interaction present in the model, then it is recommended to use a non-additive model.

Due to arbitrary measurement scales our inferences can be wrong, which may lead to a wrong model formulation and interpretation. In this section we will discuss the use of weighted Tukey’s one degree of freedom test, assuming relevant weights are available, to test essential and removable interaction effects, separating the effect of essential and removable interaction from an overall
interaction effect in the model. Provided we can ignore essential interaction, and there is significant removable interaction, we can find a suitable transformation to remove this removable interaction, which will be discussed more fully in chapter 4. We will also see, with the help of a numerical example, the effect of using weights on the results.

\section{2.3.1 Two way classification model (n>1)}

The fixed effects model for a two-way classification with more than one observation per subclass may be given as:

\begin{equation}
\gamma_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk} \quad \left\{ \begin{array}{c} i = 1, \ldots, r \\
  j = 1, \ldots, c \\
  k = 1, 2, \ldots, n \end{array} \right. \tag{2.3.1}
\end{equation}

where the $\epsilon_{ijk}$ are independently distributed as normal with mean zero and a common variance $\sigma^2$, i.e.: $\epsilon_{ijk} \sim N(0, \sigma^2)$. $\mu$ is a constant known as a grand mean, $\alpha_i$ is considered to be a row main effect defined as an effect due to the $i$th level of the row factor. Similarly, $\beta_j$ is considered to be a column main effect defined as an effect due to the $j$th level of the column factor, and $\gamma_{ij}$ is an interaction effect of the $i$th row level and $j$th column level. Moreover, we assume that the observations in the $ij$th cell are a random sample from a population corresponding to the cell and each cell has an equal number of observations, i.e.; $n_{ij} = n$.

Considering an interaction structure I for interaction effect, model (2.3.1) can be defined as:
where \( \lambda \) is a real constant that provides the basis for a test of significance of removable non-additivity in the model (2.3.1).

### 2.3.2 Estimation of parameters

To begin with estimation, we shall assume that all the subclass sizes are equal, with each row and column having some weights available, and there are no empty cells in the two-way classification table. Model (2.3.2) is over-parameterized and, to obtain a unique solution for these parameters, we shall impose some constraints on (2.3.2). Scheffé (1956) introduced restrictions with non-negative weights \( u_i \) and \( v_j \) such that \( \sum_i u_i > 0 \) and \( \sum_j v_j > 0 \). In the case when weights are introduced, the corresponding restrictions turn out to be:

\[
\sum_{i=1}^{r} u_i \alpha_i = 0, \quad \sum_{j=1}^{c} v_j \beta_j = 0, \quad \sum_{i=1}^{r} u_i \gamma_{ij} = 0 \quad \forall \ j \quad \text{and} \quad \sum_{j=1}^{c} v_j \gamma_{ij} = 0 \quad \forall \ i. \tag{2.3.3}
\]

The effect \( \mu \) can be defined and estimated using weighted least squares estimation by minimizing \( Q_4 \) given as:

\[
Q_4 = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} u_i v_j (y_{ijk} - \mu)^2. \tag{2.3.4}
\]

Minimizing \( Q_4 \) with respect to \( \mu \) under the constraints given in (2.3.3) we can get a unique estimate of the weighted grand mean for the two-way classification table, where \( u_i \) represents a weight for the \( i \) – \( th \) row and similarly \( v_j \) represents a weight for the \( j \) – \( th \) column. An estimate of the weighted grand mean can be given as
\[ \hat{\mu} = \frac{\sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j \overline{y}_{ij}}{n \sum_{i=1}^{r} u_i \sum_{j=1}^{c} v_j}. \]  

(2.3.5)

An estimate of the row effects \( \alpha_i \) can be obtained by minimizing \( Q_5 \) given below:

\[ Q_5 = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} u_i v_j (y_{ijk} - \mu - \alpha_i - \beta_j)^2. \]  

(2.3.6)

Minimizing \( Q_5 \) with respect to \( \alpha_i \) under the constraints given in (2.3.3) will provide an estimator for the row effect:

\[ \hat{\alpha}_i = \frac{\sum_{j=1}^{c} \sum_{k=1}^{n} v_j \overline{y}_{ij}}{\sum_{j=1}^{c} v_j} - \hat{\mu}. \]  

(2.3.7)

An estimate of the column effect \( \beta_j \) can be obtained by minimizing \( Q_2 \) with respect to \( \beta_j \) under the constraints (2.3.3):

\[ \hat{\beta}_i = \frac{\sum_{i=1}^{r} \sum_{k=1}^{n} u_i \overline{y}_{ij}}{\sum_{i=1}^{r} u_i} - \hat{\mu}. \]  

(2.3.8)

Estimates of the row effect in (2.2.10) and column effect in (2.2.11) are deviations of weighted row and column means from the grand weighted mean.

We are considering structure I interaction in the model, where any interaction term is a function of the row and column effects, given as \( \lambda \alpha_i \beta_j \). The hypothesis of removable interaction tests the null hypothesis that the coefficient \( \lambda \) is zero against an alternative hypothesis that it is non-zero. The sum of squares that forms the basis for testing the hypothesis of no removable interaction can be obtained by estimating \( \lambda \) by the method of least squares. The estimate of \( \lambda \) can be obtained by minimizing \( Q_6 \):
\[ Q_6 = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} u_i v_j (y_{ijk} - \hat{\mu} - \alpha_i - \beta_j - \lambda \alpha_i \beta_j)^2. \] (2.3.9)

Assuming that \( \mu, \alpha_i \) and \( \beta_j \) are known we minimize \( Q_6 \) in (2.3.9) with respect to \( \lambda \). Using the constraints in equation (2.3.3), we get an estimate of the coefficient of the interaction term, i.e. \( \lambda \), which can be given as:

\[ \hat{\lambda} = \frac{\sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j \bar{y}_{ij} \bar{\alpha}_i \bar{\beta}_j}{\sum_{i=1}^{r} u_i \bar{\alpha}_i^2 \sum_{j=1}^{c} v_j \bar{\beta}_j^2}. \] (2.3.10)

### 2.3.3 Tests of hypothesis

In this section we will first discuss the partitioning of the total sum of squares into different components of variation and then later we will discuss hypothesis testing for the main effects and interaction effects in the model. Finally, we will present a complete weighted two-way classification analysis of variance table to summarize all the testing procedures. The total sum of squares for the full model (2.3.2) is given as:

\[
\text{Total Sum of Squares} = TSS = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} u_i v_j (y_{ijk} - \bar{y}_{..})^2. \] (2.3.11)

Now that there is more than one observation in each cell of the two-way table, we shall call the observations in a cell a group. Partitioning the total sum of squares given in (2.3.11) into among and within group variation can be obtained as shown below.
\[
\sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} u_i v_j (y_{ijk} - \bar{y}_{..})^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} u_i v_j (y_{ijk} - \bar{y}_{ij})^2 + \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} u_i v_j (\bar{y}_{ij} - \bar{y}_{..})^2.
\]  

(2.3.12)

We see that the row and column weights have been attached to all components of variation. From (2.3.12) the following results can be derived.

\[
Residual \ sum \ of \ squares = SS_{Res} = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} u_i v_j (y_{ijk} - \bar{y}_{ij})^2.
\]  

(2.3.13)

The within group variation (Residual Sum of squares) is given in (2.2.13) and similarly the among groups sum of squares (Effects sum of squares) can be given as follows.

\[
Effects \ sum \ of \ squares = SS_{between} = n \sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j (\bar{y}_{ij} - \bar{y}_{..})^2.
\]  

(2.3.14)

The effects sum of squares given in (2.3.14) can further be partitioned into other components of variation:

\[
\sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} u_i v_j (\bar{y}_{ij} - \bar{y}_{..})^2
\]

\[
= \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} u_i v_j (\bar{y}_{ij} - \bar{y}_{..})^2 + \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} u_i v_j (\bar{y}_{ij} - \bar{y}_{..})^2
\]

(2.3.15)

\[
+ \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} u_i v_j (\bar{y}_{ij} - \bar{y}_{ij} - \bar{y}_{..} + \bar{y}_{..})^2.
\]

Equation (2.3.15) provides row, column and overall interaction sums of squares with row and column level weights attached to each component of variation. Final expressions for row, column and interaction variation can be given as:
Row sum of squares $SS_R = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} u_{ij} (\bar{y}_{i.} - \bar{y}_{..})^2 = V n \sum_{i=1}^{r} (\bar{y}_{i.} - \bar{y}_{..})^2$. \(2.3.16\)

Column sum of squares $SS_C = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} u_{ij} (\bar{y}_{j.} - \bar{y}_{..})^2 = U n \sum_{j=1}^{c} (\bar{y}_{j.} - \bar{y}_{..})^2$. \(2.3.17\)

Overall interaction sum of square $SS_{o-int} = n \sum_{i=1}^{r} \sum_{j=1}^{c} u_{ij} (\bar{y}_{ij} - \bar{y}_{i..} - \bar{y}_{.j} + \bar{y}_{..})^2$. \(2.3.18\)

where $V = \sum_{j=1}^{c} v_j$ and $U = \sum_{i=1}^{r} u_i$. The null hypotheses of main effects $H_R: u_i \alpha_i = 0$ for all $i$ and $H_C: v_j \beta_j = 0$ for all $j$ and overall interaction $H_I: \gamma_{ij} = 0$ for all $ij$ in \(2.3.2\), These hypotheses are the same as saying that all the $\alpha_i = 0$ and all the $\beta_j = 0$. Estimates of sum of squares of row, column and interaction effects are given by \(2.3.16\), \(2.3.17\) and \(2.3.18\), respectively. Since the null hypothesis for the row effects, i.e.: $H_R: u_i \alpha_i = 0$ for all $i$, states that $(r - 1)$ linearly independent estimable conditions are imposed; the number of degrees of freedom for $SS_R$ is $(r - 1)$. The null hypothesis for column effects, i.e.: $H_C: v_j \beta_j = 0$ for all $j$, states that $(c - 1)$ linearly independent estimable conditions are imposed; the number of degrees of freedom for $SS_C$ is $(c - 1)$. Similarly the number of independent estimable conditions imposed by the null hypothesis of no interaction effect i.e. $H_I: \gamma_{ij} = 0$ is $rc$ “However, if we think of arranging the $\{\gamma_{ij}\}$ in an $r \times c$ table, we notice that, if all $\gamma_{ij} = 0$ in the sub-table obtained by deleting the last row and last column, then $\gamma_{ij} = 0$ in the whole table, since the rows and column sums must be zero. This suggests that the number of linearly independent restrictions imposed by the null hypothesis of no interaction effect, i.e.; $H_I$ is $(r - 1)(c - 1)$” (Scheffé, 1959). Mean squares for row, column and interaction effects can be given as:
\[ MS_R = \frac{SS_R}{r-1} \]  

(2.3.19)

\[ MS_C = \frac{SS_c}{c-1} \]  

(2.3.20)

\[ MS_{o-int} = \frac{SS_{o-int}}{(r-1)(c-1)}. \]  

(2.3.21)

\( MS_R \) in (2.3.19) is distributed as \( \sigma^2 \chi^2_{(r-1)} \), \( MS_C \) given in (2.3.20) is distributed as \( \sigma^2 \chi^2_{(c-1)} \), and \( MS_{o-int} \) is distributed as \( \sigma^2 \chi^2_{(r-1)(c-1)} \).

If the null hypothesis of no interaction effects is rejected, then usually we interpret the results with interaction. Then the question arises: are these non-additive effects true? To answer this question we further investigate the interaction sum of squares, acknowledging the fact that we obtain most of our data on an available convenient scale of measurement. We are not certain if our measured scale is a “true” scale, which raises concerns over the validity of the results. Many interactions exist due to such ambiguity of measurement scale. Keeping the above scenario in mind, we suggest using Tukey’s one degree of freedom test to isolate the effect of essential and removable non-additivity from the overall interaction effect. We believe that Tukey’s test will help us detect if there is any non-additivity due to scale of measurement.

The sum of squares for removable non-additivity when an interaction structure is given by \( \gamma_{ij} = \lambda \alpha_i \beta_j \), and the estimate of \( \lambda \) is given as in (2.3.10), may be obtained by using the following expression:

\[ \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} u_i v_j \gamma_{ij}^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} u_i v_j (\lambda \alpha_i \beta_j)^2. \]  

(2.3.22)
Substituting the value of $\lambda$ in (2.3.22) and solving provides an expression for the sum of squares for removable non-additivity:

$$SS_{RI} = \frac{n\left(\sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j \bar{y}_{ij} \bar{a}_i \bar{b}_j\right)^2}{\sum_{i=1}^{r} u_i \bar{a}_i^2 \sum_{j=1}^{c} v_j \bar{b}_j^2}.$$  \hfill (2.3.23)

The sum of squares for removable interaction (2.3.23) isolated from the overall interaction in model (2.3.2) is distributed as $\sigma^2 \chi^2_{(1)}$. Moreover, $SS_{RI}$ is separated from overall interaction, which leaves us with a reduced amount of interaction, called essential interaction/non-additivity or non-removable interaction/non-additivity, which is distributed as $\sigma^2 \chi^2_{(r-1)(c-1)-1}:

$$SS_{EI} = SS_{o-int} - SS_{RI}.$$  \hfill (2.2.24)

The three sets of linear form $\{\bar{a}_i\}, \{\bar{b}_j\}$ and $\{\bar{y}_{ij}\}$ span three mutually orthogonal spaces, and the three sets are thus statistically independent. Under the respective null hypothesis, we have $E(\bar{a}_i) = E(\bar{b}_j) = E(\bar{y}_{ij}) = 0$, and hence all the Chi-Square distributions are central. The $F - test$ of $H_R$ at significance level $\alpha$ consists in rejecting $H_R$ if

$$\frac{MS_R}{MS_{Res}} > F_{\alpha; (r-1), rc(n-1)}.$$  \hfill (2.2.25)

In a similar way one would find that the $F - test$ of $H_C$ consists in rejecting $H_C$ if

$$\frac{MS_C}{MS_{Res}} > F_{\alpha; (c-1), rc(n-1)}.$$  \hfill (2.2.26)

The hypothesis of overall non-additivity given as $H_Y: \gamma_{ij} = 0$ can be rejected if

$$\frac{MS_{o-int}}{MS_{Res}} > F_{\alpha; (r-1)(c-1), rc(n-1)}.$$  \hfill (2.2.27)

The hypothesis of removable non-additivity given as $H_\lambda: \lambda = 0$ can be rejected if
\[ \frac{MS_{RI}}{MS_{Res}} > F_{\alpha; (1), rc(n-1)} \cdot \tag{2.2.28} \]

The hypothesis of essential non-additivity given as \( H_{EI} : \gamma_{EI} = 0 \) can be rejected if

\[ \frac{MS_{EI}}{MS_{Res}} > F_{\alpha; ((r-1)(c-1)-1), rc(n-1)} \cdot \tag{2.2.29} \]

A complete layout of an analysis of variance table for the weighted Tukey’s test with more than one observation per cell is given in Table 2.3.1.

### Table 2.3.1: ANOVA table for testing removable interaction (n>1 & equal subclass size)

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor R</td>
<td>( SS_R = \frac{1}{Vn} \sum_{i=1}^{r} u_i (\bar{y}_i - \bar{y})^2 )</td>
<td>( df_R = r - 1 )</td>
<td>( \frac{SS_R}{r-1} )</td>
<td>( \frac{MS_R}{MS_{Res}} )</td>
</tr>
<tr>
<td>Factor C</td>
<td>( SS_C = \frac{1}{Un} \sum_{j=1}^{c} v_j (\bar{y}_j - \bar{y})^2 )</td>
<td>( df_C = c - 1 )</td>
<td>( \frac{SS_C}{c-1} )</td>
<td>( \frac{MS_C}{MS_{Res}} )</td>
</tr>
<tr>
<td>Overall Interaction</td>
<td>( SS_{O-int} = \sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j (\bar{y}_{ij} - \bar{y}_i - \bar{y}_j + \bar{y})^2 )</td>
<td>( df_{O-int} = (r-1)(c-1) )</td>
<td>( \frac{SS_{O-int}}{df_{O-int}} )</td>
<td>( \frac{MS_{O-int}}{MS_{Res}} )</td>
</tr>
<tr>
<td>Removable Interaction</td>
<td>( SS_{RI} = \frac{n (\sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j (\bar{y}_{ij} - \bar{y}<em>i - \bar{y}<em>j + \bar{y})^2)}{\sum</em>{i=1}^{r} u_i \sum</em>{j=1}^{c} v_j \bar{y}_j^2} )</td>
<td>( df_{RI} = 1 )</td>
<td>( \frac{SS_{RI}}{1} )</td>
<td>( \frac{MS_{RI}}{MS_{Res}} )</td>
</tr>
<tr>
<td>Essential Interaction</td>
<td>( SS_{EI} = SS_{O-int} - SS_{RI} )</td>
<td>( df_{EI} = (r-1)(c-1) - 1 )</td>
<td>( \frac{SS_{EI}}{df_{EI}} )</td>
<td>( \frac{MS_{EI}}{MS_{Res}} )</td>
</tr>
<tr>
<td>Residual</td>
<td>( SS_{Res} = \sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j (\bar{y}<em>{ijk} - \bar{y}</em>{ij})^2 )</td>
<td>( df_{Res} = rc(n-1) )</td>
<td>( \frac{SS_{Res}}{df_{Res}} )</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>( SS_T = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} u_i v_j (y_{ijk} - \bar{y}_{ij})^2 )</td>
<td>( rc - 1 )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.4  **Strategy for the analysis of a factorial design**

In this section we present and extend the strategy flow chart given by Neter *et al* (1996) to incorporate the presented idea of removable non-additivity.

**Figure 2.4.1:** Strategy for analysis of a factorial design.
2.5 Numerical example

To evaluate the performance of the weighted Tukey’s test for more than one observation per cell over Tukey’s test for more than one observation per cell, we generated a data set using a model-based simulation method. The simulated dataset is based on the observations obtained from the dataset provided by Mandel (1991) and Alin et al (2006). The aim of this experiment is to determine pentosans in wood pulp by measuring absorbency. Each of 7 laboratories (representing rows in this case) received portions of each of the 9 materials (representing columns). A response variable is the absorbance values obtained by a method known as orcinol. The original data includes three replicates of each material per lab but Mandel (1991) and Alin et al (2006) used the mean of the replications of each material in each laboratory. We used the means obtained from this dataset to generate a dataset with three observations per cell. The simulated dataset is given in table 2.5.1.
Table 2.5.1: dataset for two-way classification with more than one observation per cell

<table>
<thead>
<tr>
<th>Lab Number</th>
<th>Materials</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>1</td>
<td>0.084</td>
</tr>
<tr>
<td></td>
<td>0.334</td>
</tr>
<tr>
<td></td>
<td>0.438</td>
</tr>
<tr>
<td>2</td>
<td>1.021</td>
</tr>
<tr>
<td></td>
<td>0.944</td>
</tr>
<tr>
<td></td>
<td>0.392</td>
</tr>
<tr>
<td>3</td>
<td>1.323</td>
</tr>
<tr>
<td></td>
<td>1.020</td>
</tr>
<tr>
<td></td>
<td>0.650</td>
</tr>
<tr>
<td>4</td>
<td>1.415</td>
</tr>
<tr>
<td></td>
<td>1.861</td>
</tr>
<tr>
<td></td>
<td>1.022</td>
</tr>
<tr>
<td></td>
<td>3.999</td>
</tr>
<tr>
<td></td>
<td>3.767</td>
</tr>
</tbody>
</table>

Figure 2.5.1 has been obtained from the data in table 2.5.1. The interaction plot in figure 2.5.1 shows the existence of removable and essential interaction in the model. We can see cross over effect as well as some nearly parallel lines.
Figure 2.5.1: Interaction plot of Pentosans versus laboratory

To test if there are essential and removable interaction effects present in the data we have performed Tukey's test for removable interaction for more than one observation per cell and the results are given in Table 2.5.2. It provides a basis for testing the hypothesis for the significance of row, column, overall interaction, essential interaction and removable interaction in the model. From table 2.5.2 we can infer that all the effects present in the model are significant. Moreover, we can also conclude that there is removable interaction present in the model.
Table 2.5.2: ANOVA table for Tukey’s test for removable interaction for more than one observation per cell

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laboratory</td>
<td>4625.5871</td>
<td>6</td>
<td>1052.8077</td>
<td>2546.5814</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Materials</td>
<td>118.4185</td>
<td>8</td>
<td>2042.2719</td>
<td>48.8957</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Overall Interaction</td>
<td>261.9554</td>
<td>48</td>
<td>39.7638</td>
<td>18.0271</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Removable Interaction</td>
<td>121.4738</td>
<td>1</td>
<td>111.1116</td>
<td>401.25867</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Essential Interaction</td>
<td>140.4816</td>
<td>47</td>
<td>33.2777</td>
<td>9.8733</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Residual</td>
<td>38.1442</td>
<td>126</td>
<td>6.2097</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5044.1052</td>
<td>188</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.5.2 provides the results to conduct tests of hypotheses for row, column, overall interaction, essential and removable interaction effects; their hypotheses can be given as; \( H_R: \alpha_i = 0, H_C: \beta_j = 0, H_{Y_{ij}}: \gamma_{ij} = 0, H_{EI}: \gamma_{EI} = 0 \) and the hypothesis of removable interaction \( H_{\lambda}: \lambda = 0 \). From the results we can notice that the Laboratory effect is significant as well as the among materials effect is significant, and at the same time the joint effects (overall interaction effects) of laboratory and material are also significant with a p-value of less than 0.00001. Since the overall interaction effect is significant, we can proceed by separating the effects of essential and removable interaction from the model. We can notice in table 2.5.2 that the p-value for essential interaction is significant; hence we can conclude that there is no need to test for the significance for removable interaction and one should use a non-additive model for estimation.
Let us assume that we already have information regarding the availability of materials, and based on this information, we assign weights \( v_j = (0.06, 0.28, 0.05, 0.05, 0.03, 0.48) \) to the different levels of the column variable, i.e., materials. Higher weight for a material represents availability in abundance and smaller weight represents less availability of a material. Since there is no information regarding weights for laboratories, so we give equal or no weights to the row factor levels \( u_i = (0.11, 0.11, 0.11, 0.11, 0.11, 0.11, 0.11, 0.11). \) After using these weights we applied the weighted Tukey's test for removable interaction. Given that the weights used for this test are available and appropriate, we can see the results in table 2.5.3.

### Table 2.5.3: ANOVA table for weighted Tukey's test for removable interaction for more than one observation per cell

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laboratory</td>
<td>138.1083</td>
<td>6</td>
<td>32.8478</td>
<td>2433.4529</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Materials</td>
<td>4.9016</td>
<td>8</td>
<td>41.9874</td>
<td>64.7736</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Overall Interaction</td>
<td>4.6701</td>
<td>48</td>
<td>5.7889</td>
<td>10.2858</td>
<td>&lt; 0.00001</td>
</tr>
<tr>
<td>Removable Interaction</td>
<td>4.02169</td>
<td>1</td>
<td>5.2225</td>
<td>425.1702</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Essential Interaction</td>
<td>0.6484</td>
<td>47</td>
<td>5.8698</td>
<td>1.4585</td>
<td>0.05101</td>
</tr>
<tr>
<td>Residual</td>
<td>1.1918</td>
<td>126</td>
<td>1.4329</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>148.8718</td>
<td>188</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Given the entire row weights and columns weights are appropriate and available, the results in table 2.5.3 shows that all effects are significant at the
5% level except the essential interaction effect, which is non-significant. This leads to the conclusion that if the given weights are true and appropriate and essential interaction is non-significant we can proceed to testing for the existence of the removable interaction effect in the model, which is significant in this case. In such a situation we recommend using a transformation, in order to eliminate removable interaction effect from the model. Moreover, doing this will help achieve precise estimates and a parsimonious model.

2.6 Conclusion

In conclusion, the proposed test is better to detect the effects traditional test are unable to detect properly. Improper detection of results may lead to invalid and misleading conclusions, which can be avoided by using proper weights. If available, using appropriate weights may also help eliminate interaction from the model, though this is not always the case. In the case of a single observation per cell, the decision regarding the existence of essential interaction must be made by observing the interaction plot and looking for interaction type D, but in the case of more than one observation per cell we can test the null hypothesis of essential interaction. Given there is non-significant essential interaction in the model, a test for the significance of the removable interaction should be performed. In situations where the test for removable interaction is significant and the test for essential interaction is non-significant, our recommendation is to perform a transformation of the response variable to eliminate removable interaction from the model and use an additive model.
Chapter 3

Modification of the weighted Tukey’s test for removable interaction

3.1 Introduction

In the previous chapter we have discussed different tests for removable non-additivity. Each individual test has its own limitations for application depending on the structure of the interaction described by the dataset, and hence the two-way classification model also changes according to the structure of the interaction. It is known from the literature that Tukey’s one degree of freedom test has good power for the interaction structure I. But unfortunately, in cases of interaction structures II and III (more general forms of interaction), it has less power, as compared to other alternative tests for non-additivity. Ghosh & Sharma (1963) discussed the power of Tukey’s test and later Johnson & Graybill (1972) also discussed the low power of Tukey’s one degree of freedom test.

To improve the power of this test for more general cases of interaction structure, in the case of fixed effects and normally distributed data, Rasch et al (2009) and Simecek & Simeckova (2012) discussed a modification of Tukey’s one degree of freedom test for non-additivity. In the next section we will discuss modification of the weighted Tukey’s one degree of freedom test for the case of one observation per cell, when we have some appropriate weights available.
3.1.1 Modification of the test

We propose a modification of the weighted Tukey’s test discussed in chapter 2 to overcome the problem of low power in more general interaction structures. Consider the two-way analysis of variance model with the number of either rows or column greater than 2, is of the form:

\[ y_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ij} \quad \left\{ \begin{array}{l} i = 1, \ldots, r \\ j = 1, \ldots, c \end{array} \right. \]  

(3.1.1)

where the interaction effect \( \gamma_{ij} \) has an interaction of structure I, given as; \( \gamma_{ij} = \lambda \alpha_i \beta_j \). The above model can be written as:

\[ y_{ij} = \mu + \alpha_i + \beta_j + \lambda \alpha_i \beta_j + \epsilon_{ij} \quad \left\{ \begin{array}{l} i = 1, \ldots, r \\ j = 1, \ldots, c. \end{array} \right. \]  

(3.1.2)

The main idea behind the modification is that a two-way ANOVA model given in (3.1.2) is fitted by non-linear regression and tested against the sub-model (i.e. no interaction model) by a likelihood ratio test. Two estimators of row and column effects are calculated in the same way in a saturated model and in a sub-model, but dependency of the response \( y_{ij} \) on the effect parameters is not linear for the saturated model, resulting in different estimates of the effect estimators.

3.1.2 Estimation of parameters

Estimators of parameters in the non-linear model (3.1.2) are computed by the standard iterative procedure discussed in Bates and Watts (1988), Rasch et al. (2009) and Simecek & Simeckova (2012). In the first step of the iterative estimation procedure two parameters of the model are assumed fixed and the third
parameter is estimated by linear regression. Under the additive model least squares estimators can be calculated simply as follows:

\[ \hat{\mu} = \bar{y}_- = \frac{\sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j y_{ij}}{\sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j} . \]  

(3.1.3)

The least squares estimator of the row effect can be obtained by minimizing:

\[ Q_2 = \sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j (y_{ij} - \mu - \alpha_i - \beta_j)^2 . \]  

(3.1.4)

An estimator of the row effect is a deviation of a weighted row mean from the overall weighted mean, where \( v_j \) denotes the \( j \)-th column weight. Initial row effects are given by \( \hat{\alpha}_i^0 \).

\[ \hat{\alpha}_i^0 = \bar{y}_i - \bar{y}_- = \frac{\sum_{j=1}^{c} v_j y_{ij}}{\sum_{j=1}^{c} v_j} - \bar{y}_- . \]  

(3.1.5)

Similarly, the least squares estimate of the column effect can also be obtained by minimizing \( Q_2 \) given in equation (3.1.4). An initial estimator of the column effect is also a deviation of a weighted column mean from the overall weighted mean, where \( u_i \) denotes the \( i \)-th row weight. Initial column effects are given as:

\[ \hat{\beta}_j^0 = \bar{y}_j - \bar{y}_- = \frac{\sum_{i=1}^{r} u_i y_{ij}}{\sum_{i=1}^{r} u_i} - \bar{y}_- . \]  

(3.1.6)

The residual sum of squares for initial estimators can be given as:

\[ Residual \ SS = RSS^0 = \sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j (y_{ij} - \hat{\mu} - \hat{\alpha}_i^0 - \hat{\beta}_j^0)^2 , \]
\[ RSS^0 = \sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j \left( y_{ij} - \bar{y}_i - \bar{y}_j - \bar{y} \right)^2. \]  

(3.1.7)

In the saturated model (3.1.2), the initial value of the coefficient \( \lambda \) of an interaction term can be obtained simply by using the weighted Tukey’s method:

\[
\hat{\lambda}^{(0)} = \frac{\sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j \left( \alpha_i^{(0)} \beta_j^{(0)} \right) \left( y_{ij} - \bar{y} - \alpha_i^{(0)} - \beta_j^{(0)} \right)}{\sum_{i=1}^{r} u_i \left( \alpha_i^{(0)} \right)^2 \sum_{j=1}^{c} v_j \left( \beta_j^{(0)} \right)^2}. \]  

(3.1.8)

Equation (3.1.8) provides the same estimator as the weighted Tukey’s method described earlier in chapter 2. In the second step of this iterative procedure, to obtain an estimator in a non-linear model, we continue updating the estimator based on an estimator obtained in the first step. To obtain an iterative solution to the estimators of row, column and interaction effects, we update the initial values of the estimators using the following equations, which are obtained by minimizing \( Q_7 \), given below:

\[ Q_7 = \sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j \left( y_{ij} - \mu - \alpha_i - \beta_j - \lambda \alpha_i \beta_j \right)^2. \]  

(3.1.9)

The least squares estimates of the row and column effects and \( \lambda \) thus obtained by minimizing (3.1.9) with respect to \( \alpha, \beta \) and \( \lambda \) respectively, give the following estimators of the row effects, column effects and \( \lambda \):

\[
\hat{\alpha}_i^{(n)} = \frac{\sum_{j=1}^{c} v_j \left( y_{ij} - \bar{y} - \hat{\beta}_j^{(n-1)} \left( 1 + \hat{\lambda}^{(n-1)} \hat{\beta}_j^{(n-1)} \right) \right)}{\sum_{j=1}^{c} v_j \left( 1 + \hat{\lambda}^{(n-1)} \hat{\beta}_j^{(n-1)} \right)^2}, \]  

(3.1.10)

\[
\hat{\beta}_j^{(n)} = \frac{\sum_{i=1}^{r} u_i \left( y_{ij} - \bar{y} - \hat{\alpha}_i^{(n-1)} \left( 1 + \hat{\lambda}^{(n-1)} \hat{\alpha}_i^{(n-1)} \right) \right)}{\sum_{i=1}^{r} u_i \left( 1 + \hat{\lambda}^{(n-1)} \hat{\alpha}_i^{(n-1)} \right)^2}, \]  

(3.1.11)
\[
\hat{\lambda}^{(n)} = \frac{\sum_{i=1}^{r} \sum_{j=1}^{c} u_i v_j \alpha_i^{(n-1)} \beta_j^{(n-1)} (y_{ij} - \mu - \alpha_i^{(n-1)} - \beta_j^{(n-1)})}{\sum_{i=1}^{r} u_i (\alpha_i^{(n-1)})^2 \sum_{j=1}^{c} v_j (\beta_j^{(n-1)})^2}.
\]  \tag{3.1.12}

In the above equations (3.1.8) and (3.1.12) as discussed by Rasch et al., (2009) and Simecek & Simeckova (2012), \(\lambda\) is a regression coefficient, and one can note that \(\gamma_{ij} = \lambda \alpha_i \beta_j\) is a more restricted interaction effect as compared to \(\gamma_{ij} = (\alpha \beta)_{ij}\). We are considering just a one-step iterative procedure of estimation of the parameters of model (3.1.2), since the iteration converges very quickly and the absolute difference between the residual sum of squares from the sub-model and the full model i.e., \(|\text{RSS}^{(n)} - \text{RSS}^{(n-1)}|\) is small.

### 3.1.3 Testing removable interaction

In this section the likelihood ratio test of \(H_0: \lambda = 0\) versus an alternative hypothesis \(H_A: \lambda \neq 0\) will be obtained for the model (3.1.2). This of course is a test for removable non-additivity. Let us denote \(\text{RSS} = \text{RSS}^{(n)}\). A likelihood ratio statistics is a difference of 2 log likelihoods and is given as:

\[
2 \text{ Log Likelihood} = \frac{|\text{RSS}^{(0)} - \text{RSS}^{(n)}|}{s^2}.
\]  \tag{3.1.13}

The test statistic given in (3.1.13) is asymptotically \(\chi^2_{(1)}\), and \(s^2\) is a consistent estimator of the residual variance \(\sigma^2\), which is given as:

\[
s^2 = \frac{\text{RSS}}{rc - r - c}.
\]  \tag{3.1.14}

\[
\text{and } \frac{\text{RSS}}{\sigma^2} \sim \chi^2_{(rc - r - c)}.
\]  \tag{3.1.15}

Thus, using a linear approximation of the given nonlinear model in (3.1.1),
\[
\frac{RSS^{(0)} - RSS}{RSS} \sim F_{(1, rc-r-c)}.
\]

Further manipulation of (3.1.16) gives the modified weighted Tukey's test, which rejects the additivity hypothesis if, and only if,

\[
RSS^{(0)} > RSS \left( 1 + \frac{1}{rc - r - c} F_{1, rc-r-c}(1 - \alpha) \right),
\]

where \( F_{1, rc-r-c}(1 - \alpha) \) stands for the \((1 - \alpha)\) quantile of the F-distribution.

“Likelihood ratio test statistics converge to a \( \chi^2 \) distribution rather slowly, resulting in the type-I risk to be around 6\% instead of 5\%. Thus a correction for small sample sizes is needed” (Rasch et al, 2009). In the following section we will discuss small sample approximation in the case of the modified weighted Tukey's one degree of freedom test for removable non-additivity.

### 3.1.4 Small sample approximation

Rasch et al (2009) and Simecek et al (2012) suggested small sample approximations. Simecek et al (2012) suggested two possibilities for when the number of rows or columns is less than 20; they based their empirical threshold on simulations. The reason for this approximation is the likelihood ratio test converges to \( \chi^2 \) rather slowly and a correction for small sample is needed. In this section we will use two approximations suggested by Simecek et al (2012) to improve the proposed test for significance of the removable non-additivity. One possibility to overcome this inflation of type-I risk is to estimate the residual variance given in
(3.1.14), i.e.: $s^2 = \frac{\text{RSS}}{r_c-r-c}$, and then use a parametric bootstrapping method to generate samples of the distribution using the following expression:

$$y_{ij}^{(\text{sample})}(z) = \hat{\mu} + \hat{\alpha}_i^{(0)} + \hat{\beta}_j^{(0)} + \epsilon_{ij}^{(\text{New})}(z), \quad z = 1,2,\ldots, N^{(\text{sample})}, \quad (3.1.18)$$

where $\epsilon_{ij}^{(\text{New})}(z)$ are identically independently distributed generated from $N(0,s^2)$, and $\hat{\alpha}_i^{(0)}$ and $\hat{\beta}_j^{(0)}$ are the initial parameter estimates of the row and column effects, respectively.

A second possibility to overcome this problem is to use a permutation test, i.e. generate data as follows:

$$y_{ij}^{(\text{Perm})}(z) = \hat{\mu} + \hat{\alpha}_i^{(0)} + \hat{\beta}_j^{(0)} + \gamma_{ij}(z), \quad z = 1,2,\ldots, N^{(\text{permutation})}, \quad (3.1.19)$$

where $\pi(z)$ is a random permutation of indexes of the interaction matrix. For each $z$ the statistic of interest $S^{(\text{perm})}(z) = \text{RSS}^{(0)}(z) - \text{RSS}(z)$ is computed. The critical value equals the $(1 - \bar{a})100\%$ quantile of $S^{(\text{perm})}(z)$.

Simecek et al (2012) discussed that “the proposed statistic of interest is $\text{abs}(\lambda_{(1)})$ mirroring the deviation from the null hypothesis $\lambda = 0$. As in the permutation test, the additivity hypothesis is rejected if more than $(1 - \bar{a})$ $100\%$ of the sampled statistics lie below the observed value of the statistic based on the real data.”

3.2 Numerical example:

Rasch et al (2009) and Simecek et al (2012) discussed modification of Tukey’s one degree of freedom test for removable interaction. In section 3.1 we have proposed a similar extension in the case of a weighted Tukey’s test for removable
interaction in the case of one observation per cell. For the sake of comparison in the estimation of the results and inferences drawn, a dataset has been generated from a normal distribution with mean zero and standard deviation equal to 6, with row effects $\alpha_i = (-10, -5, 5, 10)$, column effects $\beta_j = (5, 10)$ and interaction effect $\gamma_{ij} = (-2, 3, 0, 4, 4, 0, 3, -2)$. The data have been classified into the two-way classification model with one observation per cell and given in table 3.2.1.

Table 3.2.1: Dataset for two-way classification data with interaction

<table>
<thead>
<tr>
<th>Rows</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>41.03</td>
<td>45.39</td>
<td>48.80</td>
</tr>
<tr>
<td>R2</td>
<td>28.75</td>
<td>34.00</td>
<td>45.48</td>
</tr>
<tr>
<td>R3</td>
<td>41.68</td>
<td>56.48</td>
<td>48.10</td>
</tr>
<tr>
<td>R4</td>
<td>43.23</td>
<td>51.45</td>
<td>57.07</td>
</tr>
<tr>
<td>R5</td>
<td>50.61</td>
<td>68.30</td>
<td>51.79</td>
</tr>
</tbody>
</table>

For the sake of comparison, data have been generated with interaction, which can be seen in figure 3.2.1. We can see from figure 3.2.1 the existence of completely removable and partially removable interaction. Such a type of interaction effects can be removed by using transformation, given there is negligible essential interaction effect present in the data.

From figure 3.2.1, we can also notice the absence of type D interaction (i.e. essential interaction). Therefore our recommendation is to continue with the testing of removable interaction in the model followed by a transformation, given
the condition that in the absence of essential interaction, there is significant removable interaction present in the model.

![Figure 3.2.1: Interaction plot for columns versus rows.](image)

Firstly, we will perform an analysis using modified Tukey’s test and Tukey’s test by assuming equal weights for all rows and all columns. Later we will discuss and compare the results of the weighted Tukey’s test and the modified weighted Tukey’s test.

Table 3.2.2 provides the results of the weighted Tukey’s test with equal weights, when row weights are given as $u_i = (0.20,0.20,0.20,0.20,0.20)$ and column weights are given as $v_j = (0.33,0.33,0.33)$. We should keep in mind that the equal weights case will provide the same results from the weighted Tukey’s test as Tukey’s test. Since there is no essential interaction present in the data and can be seen in the figure 3.2.1, we will perform testing for removable interaction.
Table 3.2.2: ANOVA table for weighted Tukey’s test for removable interaction

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Row Factor</td>
<td>46.6916</td>
<td>4</td>
<td>11.6729</td>
<td>4.6212</td>
<td>0.0384</td>
</tr>
<tr>
<td>Column Factor</td>
<td>20.5117</td>
<td>2</td>
<td>10.2559</td>
<td>4.0602</td>
<td>0.0675</td>
</tr>
<tr>
<td>Removable Interaction</td>
<td>0.0503</td>
<td>1</td>
<td>0.0503</td>
<td>0.0199</td>
<td>0.8917</td>
</tr>
<tr>
<td>Residual</td>
<td>17.7317</td>
<td>8</td>
<td>2.2165</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Essential Interaction</td>
<td>17.6814</td>
<td>8 − 1 = 7</td>
<td>2.5259</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>84.9352</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2.2 shows failure to reject the hypothesis of no removable interaction i.e.: $H_\lambda = \lambda = 0$. Hence we can conclude that there is no removable interaction in the model; but in fact, if we look at the figure 3.2.1, we can notice the existence of A, B and C types of interactions in the model.

Keeping the same equal weights as described above we performed the modified Tukey’s test for removable interaction on the data given in table 3.2.1 and the results obtained are given in table 3.2.3.

Table 3.2.3: Modified Tukey’s test for removable interaction

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Statistic</td>
<td>0.1145</td>
</tr>
<tr>
<td>Critical Value</td>
<td>0.0202</td>
</tr>
</tbody>
</table>

From the table 3.2.3 we can obtain the results for the test of removable interaction proposed by Rasch et al (2009). Since the number of rows
and columns is less than 20, we used the small sample correction mentioned in equation (3.2.18). The likelihood ratio statistic is 0.1145 with a critical value calculated as the \((1 - \alpha)100\%\) of 1000 bootstrap results obtained from the real data, i.e.: 0.0202, which is smaller than the test statistic. Hence we can conclude that there is sufficient evidence for the existence of removable interaction in the model.

Now to observe the effects of weights on the dataset, we performed weighted Tukey’s test for removable interaction. Let us assume that due to previous research or some available relevant information, we know some population proportions regarding the levels of rows and column. Thus we utilize those available population proportions as weights to see the effect of weights on the results. Table 3.2.4 provides the results of the weighted Tukey’s test with unequal weights, when row weights are given as \(u_i = (0.20,0.20,0.15,0.25,0.20)\) and column weights are given as \(v_j = (0.30,0.35,0.35)\).

**Table 3.2.4: ANOVA table for weighted Tukey’s test for removable interaction**

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Row Factor</td>
<td>45.3243</td>
<td>4</td>
<td>11.3311</td>
<td>7.8029</td>
<td>0.01015</td>
</tr>
<tr>
<td>Column Factor</td>
<td>25.7956</td>
<td>2</td>
<td>12.8978</td>
<td>8.8818</td>
<td>0.01201</td>
</tr>
<tr>
<td>Removable Interaction</td>
<td>14.1705</td>
<td>1</td>
<td>14.1705</td>
<td>9.7582</td>
<td>0.01675</td>
</tr>
<tr>
<td>Residual</td>
<td>24.3356</td>
<td>8</td>
<td>3.0420</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Essential Interaction</td>
<td>10.1651</td>
<td>8 (-1 = 7)</td>
<td>1.4522</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>95.4557</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From the above table 3.2.4, we can notice that by adding weights to the analysis the results have changed. All the main effects and interaction effects are
significant, whereas for equal weights or no weights the column effect and interaction effects were non-significant. From table 3.2.4 we can conclude that there is significant removable interaction present in the model. We can see similar results from the modified weighted Tukey's test for removable interaction.

**Table 3.2.5: Modified weighted Tukey’s test for removable interaction**

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Statistic</td>
<td>0.1014</td>
</tr>
<tr>
<td>Critical Value</td>
<td>0.0861</td>
</tr>
</tbody>
</table>

Table 3.2.5 presents the results for the modified weighted Tukey's test for removable interaction. Since the number of rows and columns is less than 20, we used the small sample correction mentioned in equation (3.2.18). The likelihood ratio statistics is 0.1014 with a critical value calculated as the $(1 - \alpha)100\%$ of the 1000 bootstrap results obtained from the real data, i.e.: 0.0861, which is smaller than the test statistic. Hence we can conclude, given that there is no or negligible magnitude of essential interaction present, there is sufficient evidence to conclude the existence of removable interaction in the model. Moreover, we recommend eliminating this removable interaction by a transformation, which will help us to make the model more precise and parsimonious.

### 3.3 Simulation study

In this section we will discuss a simulation study to study the power of weighted Tukey and modified weighted Tukey’s test over Tukey’s one degree of freedom test for non-additivity and modified Tukey’s test.
For the sake of comparison, we have generated 10,000 datasets under the alternative hypothesis $H_A: \lambda \neq 0$, in the form of a two-way classification table with five rows and three columns, from a normal distribution with mean 0 and a standard deviation equal to 6. We fixed the parameters for the row effects to be (-10, -5, 5, 10), column effects to be (5, 10) and interaction effects (-2, 3, 0, 4, 4, 0, 3, -2). Assuming the availability of appropriate weights, such as population proportions, we assign each row and each column certain weights given as: row weights (1,1,1,1,1) and column weights (0.50,0.25,0.25). Given there is negligible magnitude of essential interaction present in the model, using these weights we will see if there is any effect of including weights to detect the removable interaction effect from the model.

**Table 3.2.6: power comparison of tests for removable interaction**

<table>
<thead>
<tr>
<th>Method</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tukey's one degree of freedom test</td>
<td>0.0426</td>
</tr>
<tr>
<td>Weighted Tukey's one degree of freedom test</td>
<td>0.1668</td>
</tr>
<tr>
<td>Modified Tukey's test for non-additivity</td>
<td>0.2642</td>
</tr>
<tr>
<td>Modified weighted Tukey's test for non-additivity</td>
<td>0.3161</td>
</tr>
</tbody>
</table>

Given the weights we have used are appropriate and there is negligible essential interaction in the model, table 3.2.6 shows the modified weighted Tukey's test performs better than other tests to detect removable interaction.
3.4 Conclusion:

In conclusion, we can say that the modified Tukey’s test indeed have more power to detect removable interaction from the model as compared with Tukey’s one degree of freedom test. However, we need to perform in depth simulation study to observe the changes in power of the proposed test with change in other parameters. If appropriate weights are available and essential interaction effect is negligible then the modified weighted Tukey’s test should provide better results with more power as compared with Tukey’s test, the un-weighted modified Tukey’s test and the unmodified Tukey’s test.
Chapter 4

Transformation to remove Non-Additivity

4.1 Removing statistical interaction

Schlesselman (1973) and Tukey (1957) discussed that transformations in statistical work have found use in two major problem areas—first, providing statistical approximations and, secondly, bending the data to conform to assumptions underlying conventional analyses.

The ANOVA model is considered to be an ideal model if it satisfies basic assumptions, where the observations $y_{ij}$ are realizations of normally distributed independent random variables with homoscedastic variance about the cell means $\mu_{ij}$ that are of the form of a row constant plus a column constant, thus $\mu_{ij} = \alpha_i + \beta_j$. If any of these three assumptions (i.e. additivity, normality and homoscedastic variances) are violated, then a model is considered to be misbehaving. Anscombe & Tukey (1963), Bartlett (1947) Kruskal (1965) and Schlesselman (1973) have mentioned that a transformation suitable for improving one approximation is often suitable for improving either or both of the others. Tukey (1949) mentioned that removal of non-additivity by transformation usually tends to stabilize the variance but, if it does not, then one should consider analyzing transformed data using weighted variances. In most cases we ignore the presence of non-additivity in the model and move on to fix non-normality and non-homoscedastic variances in the model. But one should know that a model is more
meaningful if we first remove from the model non-additive effects that are merely due to the scale of measurement.

Tukey (1957) discussed two possibilities for analyzing data that do not satisfy the ideal conditions for the model.

I. Bend the model to fit the assumptions by making a nonlinear transformation, or

II. Develop new methods of analysis, which better approximates the “original” form of the data.

Simplicity and ease commend transformation, but the question is, “what constitutes an adequate and effective transformation?” The best transformation is the one that can help achieve almost ideal conditions by making a model precise and parsimonious to provide better prediction and effective interpretation of model coefficients. Moreover, a transformed scale should be either transformed back to an original scale for interpretation or itself make practical sense.

Here our focus is on transformation to remove non-additivity from the two-way classification model. Tukey (1949) introduced a two-parameter family of transformations to eliminate removable non-additivity from the two-way classification model. Later Moore & Tukey (1954) and Elston (1961) discussed the same family of transformation to remove non-additivity from the model. They proposed to transform a response variable $y_{ij}$ into a new variable $z_{ij}$ such that:

$$ z_{ij} = (y_{ij} + k)^p. \tag{4.1.1} $$
Tukey (1949) used (4.1.1) for non-negative $y_{ij}$ and suggested different forms of (4.1.1) for different values of $p$. He also suggested a log-transformation of the response variable to obtain additivity. Moore & Tukey (1954) further discussed this class of transformation for a special case when $k \geq 0$ and $p \leq 1$. Moreover, they described a graphical approach to decide on the transformation. Elston (1961) suggested use of a similar transformation as (4.1.1) and suggested how to approximate the value of $k$. Moreover, he recommended using this approximation just as a rough guide.

Kruskal (1965) presented a non-parametric approach, which finds a monotone transformation that minimizes a residual sum of squares after removing effects explained by the model. He introduced a stress function:

$$s(f, \beta) = \sqrt{\frac{\sum_{i=1}^{r}(Z_i - z_i(\beta))^2}{\sum_{i=1}^{r}(z_i(\beta) - z(\beta))^2}}.$$  \hspace{1cm} (4.1.2)

The right-hand side of equation (4.1.2) is also considered to be an overall measure of goodness of fit; $Z_i$ is a transformed variable of a given variable $Y_i$; and $z_i(\beta) = \sum_i g_{ij} \beta_j$ is a linear model for the transformed variable. In equation (4.1.2) the $Z_i$ are subject to the constraint that they have the same rank order as the $Y_i$. Bradley et al (1962) presented a similar idea that minimizes the residual sum of squares; their basic approach can be described as that of monotonically transforming the data so as to maximize “a variance ratio $F$, for the test for treatment effects”.

Brown (1975) suggested a stepwise procedure based on residual analysis to eliminate/replace outliers in the data that are responsible for the existence of non-additivity in certain rows and columns. Snee (1982) also pointed
out that non-additivity is often associated with a few rows or columns. An empirical method suggested by Brown (1975) allows a decision that a model is to be made additive after examining the data. Outliers being a cause of non-additivity have been discussed by Anscombe & Tukey (1963) and Schlesselman (1973). Cressie (1978) proposed a simple graphical procedure to calculate $\lambda$ in (4.1.3) for a special case of the Box & Cox (1964) family of power transformations with $k = 0$, hoping that there will always be one unique value of $\lambda$ that will make this model additive. Cressie (1978) utilized Tukey’s (1949) one degree of freedom test for non-additivity. The Box and Cox power family of transformations can be given as:

$$ y^{(\lambda,k)} \equiv \frac{(y+k)^{\lambda}-1}{\lambda} \text{ for } \lambda \neq 0 \text{ and } \equiv \log(y+k) \text{ for } \lambda = 0. \quad (4.1.3) $$

The transformation discussed above by Brown (1975), Cressie (1978), Elston (1961), Kruskal (1965), Moore & Tukey (1954) and Tukey (1949) can be used to eliminate removable interaction from a two-way ANOVA model, whereas the one proposed by Kruskal (1965) using the stress function in (4.1.2) can also be used when there is more than one observation per cell.

Anscombe & Tukey (1963) suggested a power transformation to remove non-additivity in a two-way analysis of variance with more than one observation per cell. If all $y$’s are positive, initially or after the addition of a common constant $k$, we can consider a power transformation. Replacing $y$ by $y^{p}$, for some fixed $p$, the required power may be roughly estimated as:

$$ p = 1 - 2k \bar{y}. \quad (4.1.4) $$
but if \( p = 0 \) in (4.1.4) then the transformation to be determined is a logarithmic transformation. In equation (4.1.4) the constant \( k = B/A \) where, \( A \) = residual sum of squares and

\[
B = \sum_{ij} z_{ij} y_{ij}^2 = \sum_{ij} z_{ij} (y_{ij} - \bar{y})^2, \tag{4.1.5}
\]

where \( z_{ij} \) in (4.1.5) are the residuals. Anscombe & Tukey (1963) also suggested a graphical procedure to detect non-additivity due to interaction.

Schlesselman (1973) also introduced a power transformation in the case of a two-way analysis of variance with more than one observation per cell:

\[
Y_{ijk}^{(\lambda)} = \begin{cases} 
  y_{ijk}^{\lambda} & \lambda \neq 0 \\
  \log y_{ijk} & \lambda = 0
\end{cases}
\tag{4.1.6}
\]

The only restriction on this family is that, for fixed \( \lambda \), \( y_{ijk}^{\lambda} \) be a monotonic function of \( y_{ijk} \) over some admissible range. Tukey (1957) has argued that power transformations are not only convenient to use but also work well in practice. Schlesselman (1973) suggested that by letting \( w \geq 0 \) denote a constant weight, a variable in the following equation:

\[
Z(\lambda) = \frac{z_1(\lambda) + wz_2(\lambda)}{\sqrt{(1 + w^2)}}. \tag{4.1.7}
\]

\( Z(\lambda) \) is a competitor of the Box and Cox likelihood procedure for getting point estimates and confidence intervals for \( \lambda \) in two-way ANOVA with more than one observation per cell. \( Z(\lambda) \sim N(0,1) \) and \( w \) are constant weights, chosen on the basis of the sample data. McNeil & Tukey (1975) introduced a transformation of the type

\[
p^{\lambda} - (1 - p)^{\lambda}, \tag{4.1.8}
\]
to remove non-additivity from the two-way analysis of variance model, where $p$
could be different functions of the response variable.

### 4.2 General discretization algorithm

A general algorithm to discretize a continuous response variable can be given as:

![Discretization process](image)

**Figure 4.1: Discretization process**

We will use the discretization process given in Figure 4.1 as a first step to empirically transform a continuous response variable in section 4.3.
4.3  Empirical transformation

In this section we propose an empirical monotonic transformation motivated by Fisher (1950) to remove non-additivity from a two-way analysis of variance model, given the condition that there exists some non-additivity in the model that is removable. Fisher’s method provides a general principle that may determine a set of scores in such a way as to maximize the ratio of the sum of the row and column sums of squares to the total sum of squares. Thus, given a two-way table of numerical observations, we may ask what values, or scores, should be assigned to them in order that the data be additive.

Because there are ties in the dataset due to the discretization, our proposed algorithm uses a secondary approach to ties as mentioned by Kruskal (1965).

In the problem under study we apply this empirical transformation to a two-way classification model when the response variable is quantitative. Hence, we are proposing to first discretize the response variable into categorical values, and then perform Fisher’s empirical transformation to obtain new monotonic scores that range in [0,1] and will help make the model additive and more parsimonious.

Before we perform this transformation algorithm, we have to make sure that there is removable interaction present in the data; for this we propose using Tukey’s one degree of freedom test for non-additivity or the weighted Tukey’s test (proposed in chapter 2 and 3), as appropriate, to see if the hypothesis for removable interaction is rejected or not. If with a negligible magnitude of essential interaction, there is sufficient evidence for the existence of removable non-additivity
in a model, then we recommend the following empirical transformation to eliminate removable non-additivity from the data.

The proposed empirical monotonic transformation is given as:

**Step 1:**

Sort the data into ascending order and then discretize the ordered response variable into $m$ disjoint groups using p-tiles, i.e. $p_l$ where $l = 1, 2, ..., m - 1$. P-tiles will discretize a response variable into equal-size categories (+/- one). We suggest the minimum number of percentiles should be 3, which will divide the data into 4~equal parts, and hence min $(m) = 4$. The $p - th$ p-tile is defined by:

$$Q(p) = (1 - \theta)y_j + \theta y_{j+1}$$

$$\frac{(j - s)}{N} \leq p < \frac{(j - s + 1)}{N},$$

where $y_j = j - th$ order statistic

$N = total\ number\ of\ observations, i.e.; n \times r \times c$

$\theta = Np + s - j$ and $s = 1 - p$

**Step 2:**

Once a continuous response variable is discretized into its categories, then replace the existing observations with $m$ new categorical labels, say $l_1, l_2, ..., l_m$. Here $l_m$ represents the $m - th$ interval. Assign the value zero to categorical label $l_1$ and the value 1 to $l_m$. If we have $m$ categories then create $(m - 1)$ new binary variables, say $x_1, x_2, ..., x_{m - 1}$, such that:

$$\begin{cases} 
  x_m = 1 & \text{for } l_{m+1} \\
  0 & \text{otherwise.} 
\end{cases}$$
Step 3:

Use the \( x_i \) obtained in step 2 as \((m - 1)\) new response variables and perform multivariate generalized regression on row and column factor variables as predictors, to obtain the following sum of squares and cross products:

\[
\mathbf{R} = c \sum_{i=1}^{r} (\overline{X}_{i.} - \overline{X}.)(\overline{X}_{i.} - \overline{X}.)' \tag{4.3.4}
\]

\[
\mathbf{C} = r \sum_{j=1}^{c} (\overline{X}_{.j} - \overline{X}.) (\overline{X}_{.j} - \overline{X}.)' \tag{4.3.5}
\]

\[
\mathbf{E} = \sum_{i=1}^{r} \sum_{j=1}^{c} (X_{ij} - \overline{X}_{i.} - \overline{X}_{.j} + \overline{X}.)(X_{ij} - \overline{X}_{i.} - \overline{X}_{.j} + \overline{X}.)' \tag{4.3.6}
\]

\[
\mathbf{T} = \sum_{i=1}^{r} \sum_{j=1}^{c} (X_{ij} - \overline{X}.) (X_{ij} - \overline{X}.)' \tag{4.3.7}
\]

where \( R, C, E \) and \( T \) respectively represent row factor, column factor, residual and total sum of squares and cross product matrices.

Step 4:

Post multiply the inverse of the total sum of squares and cross product matrix \( T^{-1} \) by a matrix of combined row and column effect factors \( R + C \), say \( \mathbf{F} \), and calculate the matrix

\[
\mathbf{G} = T^{-1} \mathbf{F}.
\]

where \( T \) and \( F \) are the matrices of the appropriate quadratic forms.

Step 5:

Calculate the largest eigenvalue of \( \mathbf{G} \), say \( \theta \); remove the last row from the matrix \( F - \theta T \) to give \( \mathbf{H} \), say. Then solve \( Hz = 0 \) to give \( m-2 \) new scores, say
$z_2, z_3, \cdots, z_{m-1}$. Add to these the minimum score $z_1 = 0$ and the maximum score $z_m = 1$. Hence we will have $m$ new scores.

**Step 6:**

Determine if a monotonicity condition holds true for these scores:

i.e., given that $y_{ij} \geq y_{i'j'}$, do we have $z_j \geq z_{j'}$

(where $z_j$ is the transformed score for data point $y_{ij}$ in the $i$–$th$ row and $j$–$th$ column)

**Step 7:**

If the monotonicity condition holds true, then continue from step 1, by adding an extra quintile point to discretize the complete data into one more ~ equal width intervals, and repeat in this manner until the monotonicity condition does not hold true. If the monotonicity condition is not satisfied at step 7, and this happens when there are $m - 1$ scores, then stop this algorithm and consider the $(m - 2)$ new scores as the new transformed scores, adding to it minimum score i.e. $z_1 = 0$ and maximum score $z_m = 1$ will make new $m$ transformed scores. Then perform an analysis of variance on these new scores as the transformed scale.

It should be noted that this procedure does not restrict the resulting scores, the $z_j$, to have the same rank order as the original scores. Occasionally two consecutive scores may have their order reversed, but when this occurs the data matrix has invariably contained only very small number of one of the scores involved, suggesting that small random errors can account for this reversal. The scores resulting from this transformation have always turned out to be positive. It is also required for $T$ to be non-singular, if a particular score does not exist in the data matrix; it causes $T$ to be singular.
4.4 Transformation for one observation per cell

This section includes numerical examples to evaluate the effect of the proposed transformation in different situations in the case of one observation per cell.

4.4.1 Transformation for different structures of interaction

As discussed in section 2.1, interactions in a two-way classification models follow three different structures. In the following sections we will discuss the presence of essential and removable interaction and will evaluate the performance and effects of the proposed empirical transformation in eliminating removable interaction in the presence of negligible essential interaction. In the case of a single observation per cell, it is not possible to perform a formal test of hypothesis for the significance of essential interaction; therefore we recommend using an interaction plot to observe the type of interaction. If an interaction plot shows the absence of type D interaction (i.e., essential interaction), and test of significance of removable interaction shows presence of removable interaction, then we recommend using the transformation to eliminate removable interaction, otherwise using a non-additive model for estimation.

4.4.1.1 Interaction structure I

Interaction structure I is the one where the interaction effect is a function of either of the row and column effects. In the following numerical example to demonstrate the performance of the proposed transformation, we will use the
dataset provided by Mandel (1991), which is an example of an inter-laboratory study. The aim of the experiment discussed by Mandel (1991) is to determine pentosans in wood pulp by measuring absorbency. Each of the 7 laboratories (representing rows in this case) received portions of each of the nine materials (representing columns). The response variable is the absorbance values obtained by a method known as orcinol. The original experiment included three replicates of each material per lab, Mandel (1991) and Alin et al (2006) used the mean of the three replicates of each material in each laboratory. Mandel (1991) and Alin et al (2006) both discussed that the data follows interaction structure I.

First of all we will observe an interaction plot to examine the existence or not of essential interaction in the data. We can see in Figure 4.4.1 (a) that all the lines run almost parallel but there is some overlapping, which shows the presence of essential interaction, but it is not very clear how much essential interaction is present. In the presence of interaction structure I, there is a gradual widening of the straight lines representing the 7 laboratories. Widening of the lines results in inflation of the standard deviation among the laboratories.

Table 4.4.1 provides the results for an analysis of variance table for Tukey’s one degree of freedom test for removable interaction.
Table 4.4.1: ANOVA table for Tukey’s test for removable interaction for one observation per cell, before transformation

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lab Number</td>
<td>1.8510</td>
<td>6</td>
<td>0.3085</td>
<td>7.5423</td>
<td>0.00001</td>
</tr>
<tr>
<td>Materials</td>
<td>1640.0241</td>
<td>8</td>
<td>205.0030</td>
<td>5012.0407</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Residual</td>
<td>8.0371</td>
<td>48</td>
<td>0.1674</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Removable Interaction</td>
<td>6.1147</td>
<td>1</td>
<td>6.1147</td>
<td>149.4951</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Essential Interaction</td>
<td>1.9224</td>
<td>12 – 1 = 47</td>
<td>0.0409</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1649.9122</td>
<td>62</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4.1 provides the basis for testing of hypotheses for the row, column and removable interaction effects, i.e.: $H_R: \alpha_i = 0$, $H_C: \beta_j = 0$ and $H_\lambda: \lambda = 0$ respectively. We can notice from table 4.4.1 that all the effects are significant, which leads to a conclusion that there exists removable interaction in the model.

![Interaction plots](image)

Figure 4.4.1: (a) Interaction plot of Pentosans versus Lab before transformation (b) Interaction plot of Pentosans versus Lab after transformation.

Tukey’s test of removable interaction shows significant existence of removable interaction in the model, hence we apply the proposed transformation to
the response variable to eliminate the discovered removable interaction effect. After iteratively dividing the data into 5 mutually disjoint intervals using 4 quantiles, we obtained monotonic scores by applying the generalized regression model. The monotonic scores obtained are given in table 4.4.2.

Table 4.4.2: Monotonic scores after transformation

<table>
<thead>
<tr>
<th></th>
<th>Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>0.00000</td>
</tr>
<tr>
<td>2.</td>
<td>0.08618</td>
</tr>
<tr>
<td>3.</td>
<td>0.31047</td>
</tr>
<tr>
<td>4.</td>
<td>0.57325</td>
</tr>
<tr>
<td>5.</td>
<td>1.00000</td>
</tr>
</tbody>
</table>

Table 4.4.2 provides monotonic scores obtained after applying the empirical transformation introduced in section 4.3, to eliminate removable interaction from the model. A new transformed variable, i.e. these transformed scores, is used to reanalyze the data. Table 4.4.3 provides the results obtained from the data after this monotonic transformation.
Table 4.4.3: ANOVA table for Tukey’s test for removable interaction for one observation per cell, after transformation

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lab Number</td>
<td>0.0606</td>
<td>6</td>
<td>0.0101</td>
<td>1.5646</td>
<td>0.1786</td>
</tr>
<tr>
<td>Materials</td>
<td>8.1099</td>
<td>8</td>
<td>1.0137</td>
<td>156.9203</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Residual</td>
<td>0.3225</td>
<td>48</td>
<td>0.0067</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Removable Interaction</td>
<td>0.0189</td>
<td>1</td>
<td>0.0189</td>
<td>2.9317</td>
<td>0.09345</td>
</tr>
<tr>
<td>Essential Interaction</td>
<td>0.3036</td>
<td>12 – 1 = 47</td>
<td>0.0065</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>8.4931</td>
<td>62</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4.3 provides the analysis of variance model for Tukey’s test for removable interaction with one observation per cell on the transformed response variable. In comparison with the results from table 4.4.1, the results in table 4.4.3 are different. The test of significance for row effect shows non-significant results, which represents that there is no effect among laboratories. However, the column effect is still significant, which leads to the conclusion that there is a significant difference among the material types. Figure 4.4.1 shows the interaction plot of the material versus laboratory. Finally, we can notice the hypothesis of removable interaction also shows non-significant results. We can conclude that there is no removable interaction effect in the model after the empirical transformation, which leads to the conclusion that we could use an additive model with row and column effects. But if we look at the interaction plot after the transformation, we can see that there still exists some essential interaction effect; therefore, in the absence of data to test for essential interaction, we may conclude
that a saturated model with interaction is an appropriate model to estimate row, column and interaction effects.

### 4.4.1.2 Interaction structure II

Interactions of structure II are the ones where interaction does not depend on any of the main effects. In the following example we will discuss an application of the proposed transformation of the response variable in the presence of interaction structure II, in the case of one observation per cell. For this purpose we generated a dataset with 7 rows and 4 columns from a normal distribution with mean zero and a standard deviation of 2, having row effects \( \alpha_i = (-15, -10, -5, 5, 10, 15) \), column effects \( \beta_j = (-10, 3, 7) \) and interaction effect \( \gamma_{ij} = (12, 4, 5, 8, 2, 13, 5, 2, 6, 8, 9, 1, 4, 6, 7, 12, 16, 2) \). Assuming there is negligible essential interaction present in the model, firstly we perform Tukey’s one degree of freedom test for removable interaction in the case of one observation per cell and the results of the test are given in table 4.4.4.

**Table 4.4.4: ANOVA table for Tukey’s test for removable interaction for one observation per cell, before transformation**

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor R</td>
<td>3207.0651</td>
<td>6</td>
<td>534.5109</td>
<td>31.1469</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Factor C</td>
<td>707.0059</td>
<td>3</td>
<td>235.6686</td>
<td>13.7328</td>
<td>0.00008</td>
</tr>
<tr>
<td>Residual</td>
<td>371.4406</td>
<td>18</td>
<td>20.6356</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Removable Interaction</td>
<td>79.7046</td>
<td>1</td>
<td>79.7046</td>
<td>4.6445</td>
<td>0.0458</td>
</tr>
<tr>
<td>Essential Interaction</td>
<td>291.7360</td>
<td>18 – 1 = 17</td>
<td>17.1609</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>4285.5117</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The results obtained in table 4.4.4 provide a complete analysis of variance table, which provides sums of squares and significance values for row, column and interaction effects. These results help evaluate the following three hypothesis $H_r: \alpha_i = 0$ for all $i$, $H_c: \beta_j = 0$ for all $j$ and the hypothesis of removable interaction $H_A: \lambda = 0$ for all $ij$. From table 4.4.4, we can conclude that both the main effects are significant. Removable interaction is also significant, with p-value 0.0458. Hence we can conclude that there exists removable interaction in the model, which is also evident from figure 4.4.2 (a): some lines are nearly parallel, showing type A interactions and some are showing type B and C interactions. In the case of a single observation per cell we cannot test the hypothesis for the presence of essential interaction in the model. It is recommended to observe the existence of essential interaction graphically by viewing an interaction plot. Figure 4.4.2(a) shows an interaction plot for columns versus rows. We can see from this figure that types A, B and C interactions are all present in the data; however, there is little crossover effect, suggesting absence of essential interaction effect. Therefore, we suggest proceeding with a transformation to eliminate removable interaction from the model.
Since the test of removable interaction shows significant results, in the presence of negligible essential interaction, we proceed with the monotonic transformation proposed in section 4.3. To eliminate the effect of the removable interaction we perform the proposed transformation of the data. After iteratively dividing the data into seven mutually disjoint intervals using six quantiles, we obtained monotonic scores by applying the generalized regression model. The monotonic scores obtained are presented in table 4.4.5.

**Table 4.4.5: Monotonic scores after transformation**

<table>
<thead>
<tr>
<th></th>
<th>Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00000</td>
</tr>
<tr>
<td>2</td>
<td>0.28027</td>
</tr>
<tr>
<td>3</td>
<td>0.46175</td>
</tr>
<tr>
<td>4</td>
<td>0.60287</td>
</tr>
<tr>
<td>5</td>
<td>0.74718</td>
</tr>
<tr>
<td>6</td>
<td>0.95009</td>
</tr>
<tr>
<td>7</td>
<td>1.00000</td>
</tr>
</tbody>
</table>
Table 4.4.5 provides monotonic scores obtained after applying the empirical transformation introduced in section 4.3, to eliminate removable interaction effects from the model. We created a new transformed variable based on these transformed scores, and used it to reanalyze the data. Table 4.4.6 provides the results obtained from the data after monotonic transformation.

Table 4.4.6: ANOVA table for Tukey’s test for removable interaction for one observation per cell, after transformation

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor R</td>
<td>2.3724</td>
<td>6</td>
<td>0.3954</td>
<td>40.0665</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Factor C</td>
<td>0.5800</td>
<td>3</td>
<td>0.1933</td>
<td>19.5894</td>
<td>0.00001</td>
</tr>
<tr>
<td>Residual</td>
<td>0.1758</td>
<td>18</td>
<td>0.0421</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Removable Interaction</td>
<td>0.0080</td>
<td>1</td>
<td>0.0080</td>
<td>0.8042</td>
<td>0.38237</td>
</tr>
<tr>
<td>Essential Interaction</td>
<td>0.1678</td>
<td>18 – 1 = 17</td>
<td>0.0099</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3.1281</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

After transformation, both row and column effects are highly significant but for the hypothesis of interaction effect, i.e.: \( H_0: \lambda = 0 \), the p-value 0.38237 does not show significant results. We can conclude that after applying the proposed transformation, the removable interaction effect has been removed from the model and the test of hypothesis for the significance of removable interaction is non-significant. Therefore, we can use an additive model to obtain much precise estimates.
4.4.1.3 Interaction structure III

Interactions of structure III are the ones where the interaction structure is unknown. In the following example we will discuss an application of the proposed transformation of the response variable in the presence of interaction structure III, in the case of one observation per cell. Table 4.4.7 “lists the volumes in milliliters of loaves of bread made under controlled conditions from 100-grams batches of dough made with 17 different varieties of wheat flour and containing x milligrams of potassium bromate”, Scheffé (1959). We refer to Scheffé(1959) for further details of the data.

Table 4.4.7: Volumes in milliliters of loaves

<table>
<thead>
<tr>
<th>Variety</th>
<th>Potassium bromate (mg)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>950</td>
<td>1075</td>
<td>1055</td>
<td>975</td>
<td>880</td>
</tr>
<tr>
<td>2</td>
<td>890</td>
<td>980</td>
<td>1000*</td>
<td>865</td>
<td>825</td>
</tr>
<tr>
<td>3</td>
<td>830</td>
<td>950</td>
<td>900*</td>
<td>770</td>
<td>735</td>
</tr>
<tr>
<td>4</td>
<td>770</td>
<td>815</td>
<td>765</td>
<td>725</td>
<td>700</td>
</tr>
<tr>
<td>5</td>
<td>860</td>
<td>1040</td>
<td>1065</td>
<td>975</td>
<td>945</td>
</tr>
<tr>
<td>6</td>
<td>835</td>
<td>960</td>
<td>985</td>
<td>915</td>
<td>845</td>
</tr>
<tr>
<td>7</td>
<td>795</td>
<td>900</td>
<td>905</td>
<td>880</td>
<td>785</td>
</tr>
<tr>
<td>8</td>
<td>800</td>
<td>860</td>
<td>870</td>
<td>850</td>
<td>850</td>
</tr>
<tr>
<td>9</td>
<td>750</td>
<td>940</td>
<td>1000</td>
<td>960</td>
<td>960</td>
</tr>
<tr>
<td>10</td>
<td>885</td>
<td>1000</td>
<td>1015</td>
<td>960</td>
<td>895</td>
</tr>
<tr>
<td>11</td>
<td>895</td>
<td>935</td>
<td>965</td>
<td>950</td>
<td>920</td>
</tr>
<tr>
<td>12</td>
<td>685</td>
<td>835</td>
<td>870</td>
<td>875</td>
<td>880</td>
</tr>
<tr>
<td>13</td>
<td>615</td>
<td>665</td>
<td>650</td>
<td>680</td>
<td>660</td>
</tr>
<tr>
<td>14</td>
<td>885</td>
<td>910</td>
<td>890</td>
<td>835</td>
<td>785</td>
</tr>
<tr>
<td>15</td>
<td>985</td>
<td>1075</td>
<td>1070</td>
<td>1015</td>
<td>1005</td>
</tr>
<tr>
<td>16</td>
<td>710</td>
<td>750</td>
<td>740</td>
<td>725</td>
<td>720</td>
</tr>
<tr>
<td>17</td>
<td>785</td>
<td>845</td>
<td>865</td>
<td>825</td>
<td>820</td>
</tr>
</tbody>
</table>

Assuming there is negligible essential interaction in the model, we perform Tukey’s one degree of freedom test for removable interaction in the case of
one observation per cell; the results of the test are given in table 4.4.8. This table provides the basis for testing hypotheses of main effects and interaction effects in a two-way classification model with one observation per cell. We will test the following hypothesis $H_R: \alpha_i = 0$, $H_C: \beta_j = 0$ and the hypothesis of removable interaction $H_A: \lambda = 0$.

Table 4.4.8: ANOVA table for Tukey’s test for removable interaction for one observation per cell, before transformation

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variety</td>
<td>770708.82</td>
<td>16</td>
<td>48169.30</td>
<td>30.8543</td>
<td>&lt;0.00001</td>
</tr>
<tr>
<td>Potassium Bromate</td>
<td>127831.18</td>
<td>4</td>
<td>31957.79</td>
<td>20.4702</td>
<td>&lt;0.00001</td>
</tr>
<tr>
<td>Residual</td>
<td>112804.82</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Removable Interaction</td>
<td>14454.15</td>
<td>1</td>
<td>14454.15</td>
<td>9.2584</td>
<td>0.00342</td>
</tr>
<tr>
<td>Essential Interaction</td>
<td>98354.67</td>
<td>8-1=63</td>
<td>1561.185</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1011348.82</td>
<td>84</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From table 4.4.8 we can conclude that row and column effects are significant. The hypothesis for removable interaction also shows significant results with p-value 0.00342. This leads to the conclusion that there exists removable interaction in the model. So, in this situation we could perform some monotonic transformation to eliminate this unnecessary effect. Figure 4.4.3 shows interaction plots of column versus rows and rows versus column before transformation.
Figure 4.4.3 (a) Interaction plot of variety versus Potassium Bromate before transformation (b) Interaction plot of variety versus Potassium Bromate after transformation.

Figure 4.4.3 (a) shows likely presence of essential interaction in the model. We can notice a lot of overlapping and cross over effect in the data. This situation suggests the existence of completely removable (i.e. Type A), partially removable (i.e. Type B and Type C) and essential interaction (Type D) in the model. Therefore, we can perhaps conclude that in this data set there is a sufficient amount of removable and partially removable interaction present, and moreover there is a negligible amount of essential interaction present in the data. In the light of this scenario, we can suggest the application of proposed transformation to eliminate the effect of removable interaction from the model.

Since the test of removable interaction also shows significant results for the presence of removable interaction, we proceed with the monotonic transformation proposed in section 4.3. To eliminate the effect of the removable and partially removable interaction, we perform the proposed transformation of the data. After iteratively dividing the data into seven mutually disjoint intervals using
six quantiles, we obtained monotonic scores by applying the generalized regression model. The monotonic scores obtained are given in table 4.4.9.

**Table 4.4.9: Monotonic scores after transformation**

<table>
<thead>
<tr>
<th>Scores</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>0.00000</td>
</tr>
<tr>
<td>2.</td>
<td>0.28656</td>
</tr>
<tr>
<td>3.</td>
<td>0.48630</td>
</tr>
<tr>
<td>4.</td>
<td>0.55638</td>
</tr>
<tr>
<td>5.</td>
<td>0.68330</td>
</tr>
<tr>
<td>6.</td>
<td>0.80450</td>
</tr>
<tr>
<td>7.</td>
<td>1.00000</td>
</tr>
</tbody>
</table>

Table 4.4.9 provides monotonic scores obtained after applying the empirical transformation introduced in section 4.3, to eliminate removable interaction effect from the model. We created a new transformed variable based on these transformed scores, and used it to reanalyze the data. Table 4.4.10 provides the results obtained from the data after monotonic transformation.

**Table 4.4.10: ANOVA table for Tukey’s test for removable interaction for one observation per cell, after transformation**

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variety</td>
<td>6.09081</td>
<td>16</td>
<td>0.3807</td>
<td>22.9795</td>
<td>&lt;0.00001</td>
</tr>
<tr>
<td>Potassium Bromate</td>
<td>1.00814</td>
<td>4</td>
<td>0.2520</td>
<td>15.2142</td>
<td>&lt;0.00001</td>
</tr>
<tr>
<td>Residual</td>
<td>1.10694</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Removable Interaction</td>
<td>0.06329</td>
<td>1</td>
<td>0.06329</td>
<td>3.8205</td>
<td>0.05507</td>
</tr>
<tr>
<td>Essential Interaction</td>
<td>1.04365</td>
<td>8-1=63</td>
<td>0.01657</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>8.2059</td>
<td>84</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

After transformation, both row and column effects are highly significant but for the hypothesis of interaction effect, i.e.: \( H_0: \lambda = 0 \), the p-value
0.05507 does not show significant results. We can conclude that after applying the proposed transformation some of the interaction effect has been removed from the model and we can use an additive model to obtain more precise estimates.

### 4.5 Transformation for more than one observation per cell

In this section we will evaluate the performance of the proposed transformation in the case of a two-way classification with more than one observation per cell.

#### 4.5.1 Equal numbers of observations

In the following numerical example we have used simulated data to see how the proposed transformation works in the case of a two-way classification table with more than one, but equal numbers of observations per cell. We used an R-program to generate a normal random sample for a two-way classification table with 5 rows and 3 columns. Each combination of rows and columns produced 4 repetitions of row and column factor in each cell. The response variable was generated from a normal distribution with zero mean and standard deviation 3, with row effects $a_i = (-10, -5, 5, 10)$, column effects $\beta_j = (5, 10)$ and interaction effect $\gamma_{ij} = (-2, 3, 4, 2, 2, 4, 3, -2)$. Then we performed the extended method of Tukey’s one degree of freedom test for non-additivity for the case of more than one observation per cell (equal cell size) to evaluate the significance of essential and removable interaction by testing $H_{EI} = \gamma_{EI} = 0$ and $H_\lambda = \lambda = 0$. If the hypothesis of essential interaction is non-significant, then we will proceed with the testing of
removable interaction followed by a transformation, otherwise we recommend using a non-additive model for estimation and modeling purposes. The results of Tukey’s test prior to the transformation is given below in table 4.5.1.

Table 4.5.1: ANOVA table for Tukey’s test for removable interaction for more than one observation per cell, before transformation

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor R</td>
<td>4211.2301</td>
<td>4</td>
<td>1052.8077</td>
<td>169.5413</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Factor C</td>
<td>6126.8158</td>
<td>3</td>
<td>2042.2719</td>
<td>328.8820</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Overall Interaction</td>
<td>477.1659</td>
<td>12</td>
<td>39.7638</td>
<td>6.4035</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Removable Interaction</td>
<td>111.1116</td>
<td>1</td>
<td>111.1116</td>
<td>17.8931</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Essential Interaction</td>
<td>366.0543</td>
<td>11</td>
<td>33.2777</td>
<td>5.3589</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Residual</td>
<td>372.5845</td>
<td>60</td>
<td>6.2097</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>11187.7964</td>
<td>79</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.5.1 provides an analysis of variance table to test the significance of row, column and removable interaction effects. We can notice that the hypothesis for row, column and interaction effects i.e.: $H_R: \alpha_i = 0$, $H_C: \beta_j = 0$ and overall interaction $H_y: \gamma_{ij} = 0$ all show significant results. Since the overall interaction effect is significant we can examine whether there is any essential and removable interaction effect present in the model or not. To answer this question we need to conduct tests of hypothesis for essential and removable interaction i.e.: $H_{EI} = \gamma_{EI} = 0$ and $H_\lambda = \lambda = 0$. Table 4.5.1 shows the sums of squares and p-values for the test of significance for essential and removable interaction. We can notice that the sums of squares for removable interaction is 111.1116 and the corresponding p-
value shows high significance, i.e.: p-value = <0.0001. Similarly hypothesis for no essential interaction can be rejected with the p-value = <0.0001; hence we can conclude the existence of essential interaction in the model. If we examine figure 4.5.1, we can see that almost all line runs parallel to each other but there is minor overlapping and cross-over in certain row and column means. From the table 4.5.1 and figure 4.5.1, it is very evident that there is significant overall interaction in the model and also there is significant removable interaction present in the data.

![Figure 4.5.1: Interaction plot of columns versus rows before transformation](image)

The essential interaction effect is significant; hence we can conclude that there is no need for transformation in this case. In such a situation our recommendation is to use a non-additive model to perform an analysis and estimate the overall model.

The results in table 4.5.1 show one aspect of the dataset but it could have happened that the hypothesis for essential interaction, i.e.: $H_{Ei}: \gamma_{ijEi} = 0$ is non-significant yet the hypothesis for removable interaction, i.e.: $H_{\lambda} = \lambda = 0$ is significant. In such a situation it is recommended to perform the proposed
transformation. If, after transformation, we are able to eliminate removable interaction from the model, then it is highly recommended to use a simple additive model.

It is very likely in some of the cases that after the rejection of the hypothesis for overall interaction, we are unable to reject both the hypotheses for removable interaction and essential interaction $H_{EI}: \gamma_{IJ EI} = 0$. In this situation we can interpret the results as having neither removable nor essential interaction effects present in the data. But if the hypothesis for essential interaction is significant and hypothesis for removable interaction, i.e.: $H_\lambda = \lambda = 0$, is non-significant then there is no need to proceed for transformation and one should consider the overall interaction as essential interaction.
4.5.2 Unequal numbers of observations

An unbalanced dataset is a very common situation that occurs in two-way classification data with more than one observation per cell. Lack of balance can be caused due to many reasons, such as missing observations for some row-column combination and/or wrong data entry. This lack of balance can also be caused when units of observation do not occur in the same proportions as in nature, or there is loss to follow-up, etc. The following numerical example will briefly illustrate an application of the proposed empirical transformation in the case of unbalanced data. We have simulated an unbalanced two-way classification data with 5 rows and 4 columns, by generating a random sample from a normal distribution with zero mean and standard deviation 2 with row effects $\alpha_i = (-10, -5, 5, 10)$, column effects $\beta_j = (5, 10, 15)$ and interaction effect $\gamma_{ij} = (-5, -2, 3, 0, 4, 2, 2, 4, 0, 3, -2, -5)$. Most of the cells in the two-way classification table have 4 observations each, but three of the cells contain 3 observations, which causes imbalance in the data; hence we cannot use methods for balanced data. Since we do not have any extensions of Tukey’s test for removable non-additivity for the case of an unbalanced analysis of variance, we will perform a simple two-way analysis for an unbalanced dataset and evaluate the performance of the proposed transformation. We will test if there is significant interaction. If the hypothesis for interaction, i.e.; $H_y: \gamma_{ij} = 0$ is significant, then we will proceed with an empirical transformation by maximizing the overall effects sums of squares. We are using type III sums of squares to calculate sums of squares for the following table. In the table 4.5.2 we get results for an unbalanced analysis of variance.
Table 4.5.2: ANOVA table for Unbalanced data, before transformation

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor R</td>
<td>1813.0</td>
<td>4</td>
<td>453.2500</td>
<td>73.2978</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Factor C</td>
<td>102.9</td>
<td>3</td>
<td>34.3000</td>
<td>5.5487</td>
<td>0.0021</td>
</tr>
<tr>
<td>Interaction</td>
<td>149.5</td>
<td>12</td>
<td>12.4583</td>
<td>2.0151</td>
<td>0.0394</td>
</tr>
<tr>
<td>Residual</td>
<td>352.5</td>
<td>57</td>
<td>6.1842</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2417.9</td>
<td>76</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.5.2 provides the basis for testing three different tests of hypothesis, i.e.; $H_R: \alpha_i = 0$, $H_C: \beta_j = 0$ and overall interaction $H_y: \gamma_{ij} = 0$. We can notice from the results in table 4.5.2 that all the three effects are significant. Figure 4.5.2 (a) shows that there is significant interaction present in the data as well. Hence we can also notice the likely presence of removable interaction in the data.
Figure 4.5.2: (a) Interaction plot of columns versus rows before transformation (b) Interaction plot of columns versus rows after transformation.
To eliminate the effect of the removable interaction given there is a negligible amount of essential interaction, we perform the proposed transformation on the data. After iteratively dividing the data into 7 mutually disjoint intervals using 6 quantiles, we obtained monotonic scores by applying the generalized regression model. The monotonic scores obtained are given in table 4.5.3.

Table 4.5.3: Monotonic scores after transformation

<table>
<thead>
<tr>
<th></th>
<th>Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0000000</td>
</tr>
<tr>
<td>2</td>
<td>0.3520859</td>
</tr>
<tr>
<td>3</td>
<td>0.5364991</td>
</tr>
<tr>
<td>4</td>
<td>0.6746017</td>
</tr>
<tr>
<td>5</td>
<td>0.7289989</td>
</tr>
<tr>
<td>6</td>
<td>0.8671841</td>
</tr>
<tr>
<td>7</td>
<td>1.0000000</td>
</tr>
</tbody>
</table>

We can see in table 4.5.3 that the minimum score is 0.0000 and the maximum score is 1.0000 and 5 new scores have been estimated. We used the scores given in table 4.5.3 to transform the original response variable into a transformed variable and performed a two-way analysis of variance for unbalanced data. The results obtained after the monotonic transformation are given in table 4.5.4.

We can notice that, after transformation, the hypothesis for the main effects are still showing significant row and column effects, but the hypothesis for interaction shows a non-significant result.
Table 4.5.4: ANOVA table for unbalanced data, after transformation

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square (MS)</th>
<th>F-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor R</td>
<td>1.16825</td>
<td>4</td>
<td>453.2500</td>
<td>23.6905</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Factor C</td>
<td>0.11570</td>
<td>3</td>
<td>34.3000</td>
<td>3.1283</td>
<td>0.03262</td>
</tr>
<tr>
<td>Interaction</td>
<td>0.14992</td>
<td>12</td>
<td>12.4583</td>
<td>1.0134</td>
<td>0.44922</td>
</tr>
<tr>
<td>Residual</td>
<td>0.70271</td>
<td>57</td>
<td>6.1842</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2417.9</td>
<td>76</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The hypothesis of interaction effect, i.e.: $H_0: \gamma_{ij} = 0$ is non-significant with p-value 0.44922. We can also see from figure 4.5.2 (b) that there is a reduction in the interaction effects and the lines are more parallel and linear which represents the presence of a partially removable interaction effect. Since the test of hypothesis for interaction effect is non-significant, we can conclude that any essential interaction effect in the transformed data is not important. Hence we can ignore this interaction effect and use a simpler main effect additive model for the estimation of the response.
Chapter 5

Conclusions and future work

5.1 Conclusion

In conclusion, the modified weighted Tukey’s test for removable interaction performs better than the corresponding weights Tukey’s one degree of freedom test. Including appropriate weights to the row and column factors levels provides more appropriate estimation of effects, and removing interaction leads to a more parsimonious model. Adding weights helps improves the estimation of parameters, but may or may not help eliminate removable interaction. The weighted Tukey’s and modified weighted Tukey’s methods require the number of either rows or columns to be greater than 2.

Assuming there is a negligible essential interaction effect present in the model, the transformation proposed in chapter 4, helps eliminate removable interaction from the model, in the case of one observation per cell and more than one observation per cell, with equal or unequal subclass sizes. We suggest the use of this transformation when the number of rows and columns is greater than 3. This transformation provides satisfactory results in the case of different interaction structure types by eliminating removable interaction and making the model more parsimonious. This transformation does not work for all types of datasets but, as Fisher mentioned, “this is not a consequence of the procedure by which transformed scores have been obtained but a property of the data examined”.

105
5.2 Future work

1. Comparison of the modified weighted Tukey’s test with other available tests for non-additivity in the case of one observation per cell.

2. Extension of the weighted Tukey’s test to factorial designs for unequal number of observations per cell.

3. Application and generalization of the proposed transformation to the case of general linear models, including fixed effects, random effects and mixed effects models.

4. Finding the properties of the data for which this transformation helps obtain a parsimonious model.

5. Finding the effect of this transformation in stabilizing the variance and normality of the data.

6. Comparison of the proposed transformation with the Box and Cox power transformations and other families of transformations.
References


