INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the original text directly from the copy submitted. Thus, some dissertation copies are in typewriter face, while others may be from a computer printer.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyrighted material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is available as one exposure on a standard 35 mm slide or as a 17" × 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. 35 mm slides or 6" × 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Statistical inference based on ranks for some repeated measurement designs with exchangeable errors within blocks

Rashid, Md. Mushfiqur, Ph.D.

The Ohio State University, 1988
STATISTICAL INFERENCE BASED ON RANKS FOR SOME
REPEATED MEASUREMENT DESIGNS WITH EXCHANGEABLE
ERRORS WITHIN BLOCKS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Md. Mushfiqur Rashid, M.A. (Rajshahi), M.Sc. (Carleton)

The Ohio State University
1988

Reading Committee
Professor D. A. Wolfe
Dr. James C. Aubuchon
Dr. Jason C. Hsu

Approved By

Douglas A. Wolfe
Adviser
Department of Statistics
To My Mother Abeda and Father AHM Abdur Rashid
ACKNOWLEDGEMENTS

I would like to thank Professor D. A. Wolfe for stimulating my interest in Nonparametric Statistical Inference as well as for serving as Chairman of my dissertation committee. My special thanks go to Dr. James C. Aubuchon for his constant guidance throughout the preparation of this dissertation. I would also like to thank Dr. Jason Hsu for serving as a member of the dissertation committee. Finally, I would like to thank my wife Nazrin for her continual support throughout my graduate study and to my sons Naim and Shafique for accepting my absence during my busy days.
VITA

January 1, 1953  
Born - Bogra, Bangladesh

July 1975  
B.A. (Hons), Statistics, University of Rajshahi, Bangladesh

March 1977  
M.A., Statistics, University of Rajshahi Bangladesh

March 1977-August 1981  
Lecturer/Assistant Professor, Department of Statistics, University of Rajshahi, Rajshahi, Bangladesh

September 1981-December 1983  
Teaching Assistant, Department of Mathematics and Statistics, Carleton University, Ottawa, Canada

June 1984  
M.Sc., Mathematics, Carleton University, Ottawa, Canada

1984-1988  
Graduate Teaching Associate, Department of Statistics, The Ohio State University

PUBLICATIONS

FIELDS OF STUDY

Major Field: Statistics

Studies in Nonparametric Statistical Inference. Professor D. A. Wolfe
Studies in Statistical Inference. Dr. Jason C. Hsu.
Studies in Statistical Decision Theory. Dr. Mark Berliner.
Studies in Anova, and Design of Experiments. Dr. W. I. Notz.
Studies in Multivariate Statistical Analysis. Dr. J. Verducci.
Studies in Statistical Prediction Analysis. Dr. M. Berliner/Professor P. K. Goel
Studies in Likelihood Inference. Dr. M. Berliner/Professor P. K. Goel
TABLE OF CONTENTS

DEDICATIONS ........................................................................................................... ii
ACKNOWLEDGEMENTS ........................................................................................... iii
VITA ........................................................................................................................ iv
LIST OF TABLES ..................................................................................................... viii

CHAPTER PAGE

I. INTRODUCTION ............................................................................................ 1
    Statement of the Problem and Overview ......................................................... 1
    Background Ideas of Dispersion Function ..................................................... 4
    Inference Based on Dispersion Function ....................................................... 8
    Rank Based Inference for Some Repeated Measures ...................................... 9
    Results in this Dissertation ........................................................................... 14
    Assumptions for this Thesis ........................................................................... 16

II. REPEATED MEASURES RCBD .................................................................. 19
    Introduction .................................................................................................. 19
    Rank Estimates of Repeated Measures RCBD ............................................. 19
    Rank Tests in Repeated Measures RCBD ................................................... 27
    Multiple Comparison Based on Full Model Estimates ................................... 32
    Asymptotic Relative Efficiency .................................................................... 33
    ARE's for Different Distributions ................................................................... 36

III. ESTIMATION OF SCALE PARAMETER ................................................. 39
    Introduction ................................................................................................... 39
    Interpretation of $\tau$ ...................................................................................... 39
    Density Estimate of $\tau$ .................................................................................. 40
    Kernel Estimate of $\tau$ .................................................................................... 41
### LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Data for RCBD</td>
<td>10</td>
</tr>
<tr>
<td>2. Residual Table for RCBD</td>
<td>11</td>
</tr>
<tr>
<td>3. ARE (Rank, Leastsquare) for RCBD</td>
<td>38</td>
</tr>
<tr>
<td>4. Illusion Strengths</td>
<td>46</td>
</tr>
<tr>
<td>5. Anova Table for Example 1</td>
<td>47</td>
</tr>
<tr>
<td>6. Estimates of Contrasts</td>
<td>47</td>
</tr>
<tr>
<td>7. Amount of Iron Absorbed By Liver Cells</td>
<td>48</td>
</tr>
<tr>
<td>8. Anova Table for Example 2</td>
<td>48</td>
</tr>
<tr>
<td>9. Estimates of Independent Contrasts</td>
<td>49</td>
</tr>
<tr>
<td>10. Amount of Radioactive Iron In Liver Cells</td>
<td>50</td>
</tr>
<tr>
<td>11. Anova Table for Example 3</td>
<td>50</td>
</tr>
<tr>
<td>12. Estimates of Independent Contrasts</td>
<td>50</td>
</tr>
<tr>
<td>13. Average Time in Implementing Injection Systems</td>
<td>51</td>
</tr>
<tr>
<td>14. Anova Table for Example 4</td>
<td>52</td>
</tr>
<tr>
<td>15. Estimates of Independent Contrasts</td>
<td>52</td>
</tr>
<tr>
<td>16. Empirical Power for n=8, p=5, ρ=.5 and σ²=2 for Normal</td>
<td>53</td>
</tr>
<tr>
<td>17. Empirical Power for n=10, p=5, ρ=.5 and σ²=2 for Normal</td>
<td>53</td>
</tr>
<tr>
<td>18. Empirical Power for n=8, p=5, ρ=.5 and σ²=2 for Multivariate 't' with 3 d.f.</td>
<td>54</td>
</tr>
</tbody>
</table>
19. **Empirical Power for n=10, p=5, \( \rho=0.5 \) and \( \sigma^2=2 \) for**

   Multivariate 't' with 3 d.f. ................................................................. 54

20. **Empirical ARE** .................................................................................. 55

21. **Bias When Samples Come From Normal with \( \rho=0.5, \sigma^2=2, p=5 \) . . . 55

22. **Bias When Samples Come From 't' with 3 d.f., \( \rho=0.5, \sigma^2=6, p=5 \) . . . 56

23. **Empirical \( \tau \) and True \( \tau \). ................................................................. 56
CHAPTER I
INTRODUCTION

1.1 Statement of the Problems and Overview

In many data analysis techniques it is commonly assumed that the error variables are normally distributed. However, there are many instances in which this assumption of normality may not be valid. In such cases statisticians would prefer to use distribution-free or even asymptotically distribution-free tests. In this dissertation rank based inference for some experimental designs (e.g., randomized complete block designs, balanced incomplete block designs and two way models (one repeated measures factor and one grouping factor)) with exchangeable errors within blocks will be developed. Unlike available distribution-free techniques for the above mentioned designs, estimation and testing are found to be tied together, which is desirable for good procedures. Two methods are discussed for estimating a scale parameter which appears in the variance-covariance matrices of rank estimates, as well as in certain ARE expressions. Confidence intervals and multiple comparisons procedures can be developed using full model estimates. The ARE's of the rank based statistics with respect to traditional least square statistics are developed. Finally, the results of a Monte Carlo study are reported for repeated measures randomized complete block designs.

In order to provide some background to rank based inference and place the above mentioned problems in perspective, we discuss previous work for the i.i.d. and exchangeable error cases. The first paper which brought a tie between testing and estimation was by Hodges and Lehmann (1963). They used Wilcoxon's (1945)
statistic (equivalently, the Mann and Whitney (1947) statistic) for the two-sample problem to estimate the shift parameter. Estimation and testing have the same efficiency expression. Adichie (1967) extended Hodges and Lehmann's idea to estimate the parameter in a simple regression. His procedures did not give any closed form answer for the slope parameter. P. K. Sen (1968) used Kendall's tau to estimate the slope parameter in a linear regression. Unlike Adichie's procedure, his procedure gave closed form answers for the rank estimates. For the general linear model with i.i.d. errors, Jureckova (1971) developed estimates of the regression parameters which are based on the minimization of certain rank statistics and gave the conditions for the asymptotic normality of these estimates. The conditions given by Jureckova contain some restrictions, relative to the design matrix, which are difficult to check for the practitioners. Jaeckel (1972) developed a rank based inference for linear models with i.i.d. errors. First, he defined a dispersion function which has some desirable properties like convexity, continuity, nonnegativity, evenness and scale invariantness. His procedures were based on obtaining estimate of parameters by minimizing the dispersion function. It turned out that the derivative of his dispersion function is equivalent to the vector of test statistics proposed by Jureckova (1971). He showed that under Jureckova's assumptions, his estimates are asymptotically equivalent to Jureckova's; that is, $\sqrt{n}$ times the difference converges to 0 in probability.

McKean and Hettmansperger (1976) developed a statistic for testing linear hypotheses. This test statistic is based on the drop in the dispersion function due to fitting a reduced model as opposed to a full model. This method uses rank estimates of the regression parameters. They also discussed the estimation of the scale parameter which appears in the denominator of the test statistics. Heiller and Willers (1979) proved asymptotic normality of both Jaeckel's and Jureckova's estimates without using Jureckova's assumptions on the design matrix. They used only the necessary and
sufficient conditions for the asymptotic normality of least square estimates. Aubuchon (1982) also independently proved the asymptotic normality of Jaeckel's estimates for Wilcoxon scores without using Jureckova's assumptions.

For testing purposes, Hodges and Lehmann (1962) proposed aligned rank tests in a two way layout. Their approach was to remove the effect of the nuisance parameters and then base a test of hypothesis on the residuals. Aligned rank tests have received much attention in the literature since 1962. Koul (1970) discussed rank tests for regression parameters with two independent variables and some improvements were suggested by Puri and Sen (1973). Adichie (1978) and Sen and Puri (1977) described aligned rank tests for univariate and multivariate models. An approach similar to the aligned rank test was also developed in Hettmansperger (1984) by plugging the reduced model estimates into the gradient vector and then constructing a quadratic form out of the gradient vector. Friedman (1937) proposed a distribution-free test for randomized complete block designs. He introduced the idea of partial ranking. His test is based on within block rankings and is applicable to both randomized complete block designs and to randomized complete block designs with exchangeable errors within blocks. Friedman's statistic is not useful for testing general linear hypotheses. When Friedman's statistic rejects the null hypothesis, averages of the ranks are used to make multiple comparisons, but those ranks depends on other treatments. Nobody has tried to use his statistic to get estimates of the treatment effects. Durbin (1951) proposed a Friedman-like test for BIB designs which is also applicable to BIB designs with exchangeable errors within blocks. His statistic has the same disadvantages as Friedman's statistic. Bernard and Val Elteren (1953) proposed a test which is applicable to designs that have an arbitrary missing structure. Mack and Skillings (1980) developed a Friedman-type statistic for the main-effects in a two factor ANOVA with more than one observation per cell. Skillings and Mack (1981) developed distribution-free tests and
multiple comparisons procedures for randomized block designs. They showed that both
Friedman's statistic and Durbin's statistic are special cases of their statistic. Koch and
Sen (1968) developed some statistics for testing and estimation of the parameters in
some repeated measures designs with general covariance structure. However, in their
setting testing and estimation do not have the same efficiency factor. Koch (1970)
developed some results for complex split plot experiments. Koch (1972) also
considered nonparametric procedures for changeover designs. Koch, Amara, Stokes
and Gillings (1972) gave some views about parametric and nonparametric repeated
measurements. Jensen (1982) discussed efficiency and robustness in repeated
measurements.

In this thesis, Jaeckel's approach will be used to make inferences about the
parameters of RCBD, BIBD and two way models with exchangeable errors within
blocks. In the following section we discuss the background ideas behind the dispersion
function and statistical inference based on it.

1.2 Background Ideas about the Dispersion Function

In this section we discuss the evolution of the dispersion function as a basis for
making inference for linear models.

A. Hodges and Lehmann (1963) considered the Wilcoxon statistic for estimating the
shift parameter in the two-sample problem. The idea is as follows.

Let $Y_1, Y_2, \ldots, Y_m$ be i.i.d. from continuous $F_1 = F(t)$ and $Y_{m+1}, Y_{m+2}, \ldots, Y_{m+n}$
be i.i.d. from continuous $F_2 = F(t-\Delta)$. Also all $m+n$ of the $Y$'s are mutually
independent. We would like to test $H_0: \Delta = 0$ versus $H_1: \Delta > 0$. Let
$R_1, R_2, \ldots, R_m, R_{m+1}, \ldots, R_{m+n}$ be the ranks of $Y_1, Y_2, \ldots, Y_m, Y_{m+1},
Y_{m+2}, \ldots, Y_{m+n}$ in the combined samples. Wilcoxon (1945) proposed the statistic
\[ W(Y_1, Y_2, ..., Y_m, Y_{m+1}, Y_{m+2}, ..., Y_{m+n}) = \sum_{i=m+1}^{m+n} R_i. \quad (1.2.1) \]

This has the following properties:

1. Reject \( H_0 : \Delta = 0 \) for large values of \( W \).
2. \( W(Y_1, Y_2, ..., Y_m, Y_{m+1} + h, Y_{m+2} + h, ..., Y_{m+n} + h) = \sum_{i=m+1}^{m+n} R(Y_i + h) \) is a nondecreasing function of \( h \) for each \((Y_1, Y_2, ..., Y_m, Y_{m+1}, Y_{m+2}, ..., Y_{m+n})\).
3. When \( \Delta = 0 \), \( W \) is symmetrically distributed about \( \frac{(m+n+1)n}{2} \) for every continuous distribution \( F(\cdot) \).

\( Y_1, Y_2, ..., Y_m, Y_{m+1} - \Delta, Y_{m+2} - \Delta, ..., Y_{m+n} - \Delta \) are i.i.d. A reasonable estimate of \( \Delta \) is that statistic \( \hat{\Delta} = \hat{\Delta}(Y_1,...,Y_{m+n}) \) such that \( Y_1, Y_2, ..., Y_m, Y_{m+1} - \hat{\Delta}, Y_{m+2} - \hat{\Delta}, ..., Y_{m+n} - \hat{\Delta} \) look like i.i.d. variables. However, there must be some criterion to judge what it means to look like i.i.d. That is where the test statistic enters into the picture. If we apply \( W \) to the aligned samples \( Y_1, Y_2, ..., Y_m, Y_{m+1} - \Delta, Y_{m+2} - \Delta, ..., Y_{m+n} - \hat{\Delta}, \) its value will be approximately equal to the null distribution median

\[ \frac{(m+n+1)n}{2}. \quad (1.2.2) \]

In other words, solving \( \sum_{i=m+1}^{m+n} R(Y_i - \hat{\Delta}) = \frac{(m+n+1)n}{2} \) yields

\[ \hat{\Delta} = \text{Median} \{ Y_j - Y_i, j = m+1, ..., m+n; i = 1, ..., m \}. \quad (1.2.3) \]

It can be shown that \( \sqrt{n}[\hat{\Delta} - \Delta] \) is normally distributed with zero mean and variance

\[ \left( \frac{1}{12 \lambda (1-\lambda) \left[ \int f^2(x) \, dx \right]^2} \right) \text{ where } \lambda = \lim_{m,n \to \infty} \frac{m}{m+n} \neq 0,1. \quad (1.2.4) \]
B. Adichie (1967) considered a simple linear regression model. In fact he extended the idea of Hodges and Lehmann to this problem. To see the motivation, let us go back to the \( W \) statistic \( W(Y_1, Y_2, \ldots, Y_m, Y_{m+1}, Y_{m+2}, \ldots, Y_{m+n}) = \sum_{i=m+1}^{m+n} R_i \) and define some regression constants \( c_i = 1 \) if \( i > m \), 
\[ = 0 \text{ if } i \leq m. \]
Then \( W(\cdot) = \sum_{i=1}^{m+n} c_i R_i \). Let \( W^*(\cdot) = \sqrt{12} \sum_{i=1}^{m+n} (c_i - c^*) \left( \frac{R(Y_i)}{m+n+1} - \frac{1}{2} \right) \), where 
\[ c^* = \sum_{i=1}^{m+n} \frac{c_i}{m+n} = \frac{n}{m+n} \] 
\( W^* \) is a linear function of \( W \) whose distribution is symmetric about 0 under the null hypothesis. Hence solving 
\[ W^*(Y_1, Y_2, \ldots, Y_m, Y_{m+1} - \Delta, Y_{m+2} - \Delta, \ldots, Y_{m+n} - \Delta) = \sqrt{12} \sum_{i=1}^{m+n} (c_i - c^*) \left( \frac{R(Y_i - c_i \Delta)}{m+n+1} - \frac{1}{2} \right) = 0 \] 
we get the same result as before.

For the simple linear regression \( Y_i = \beta c_i + \epsilon_i, i = 1, 2, \ldots, n \) with \( F_i(t) = F(t - \beta c_i) \), Adichie used the statistic 
\[ T = \frac{1}{\sqrt{n}} \sqrt{12} \sum_{i=1}^{n} (c_i - c^*) \left( \frac{R(Y_i - \beta c_i)}{n+1} - \frac{1}{2} \right) \] 
to get an estimate of \( \beta \), where \( c^* = \sum_{i=1}^{n} \frac{c_i}{n} \). Under the null hypothesis, the median of \( T \) is 0.
He solved 
\[ \frac{1}{\sqrt{n}} \sqrt{12} \sum_{i=1}^{n} (c_i - c^*) \left( \frac{R(Y_i - \beta c_i)}{n+1} - \frac{1}{2} \right) = 0 \] 
(1.2.8) 
for \( \hat{\beta} \) and showed that \( \sqrt{n} (\hat{\beta} - \beta) \) is asymptotically \( (n \to \infty) \) normally distributed with mean 0 and variance 
\[ \frac{\delta^2}{\infty}, \quad \text{where } \delta^2 = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \frac{(c_i - c^*)^2}{\int_{-\infty}^{\infty} t^2 dx}. \]
C. Let \( Y_1, Y_2, \ldots, Y_n \) be \( n \) independent random variables with c.d.f.'s of the form 
\[
F_i(t) = F(t - X' i \beta), \quad i = 1, 2, \ldots, n,
\]
where \( X' i \) is the \( i \)th row of the design matrix and \( \beta \) is a \( p \)-component vector of regression parameters. Jureckova (1971) proposed 
using 
\[
S(\beta) = [s_1(\beta), s_2(\beta), \ldots, s_p(\beta)]'
\]
for estimating \( \beta \), with 
\[
s_j(\beta) = \sum_{i=1}^{n} \sqrt{12}(x_{ij} - x_j^*) \left[ \frac{R(Y_i - X' i \beta)}{n + 1} - \frac{1}{2} \right],
\]
(1.2.9)

where \( R(Y_i - X' i \beta) \) stands for the rank of the \( i \)th residual, \( x_j^* \) is the average of the 
values of the \( j \)th explanatory variable and \( s_j(\beta) \) is just \( \frac{\sqrt{12}}{n+1} \) times the dot product of 
the residual vector and the \( j \)th column of the design matrix in deviation form.

Under \( H_0: \beta = 0 \), \( ES(\beta) = 0 \). Jureckova's idea was to minimize 
\[
\frac{R(Y_i - X' i \beta)}{n + 1} \left[ \sum_{j=1}^{p} s_j(\beta) \right].
\]
Although she could also have used other norms. Under 
some concordance conditions on the design matrix, she showed that \( \sqrt{n}(\hat{\beta} - \beta) \)
converges (as \( n \to \infty \)) to a multinormal distribution with zero mean vector and 
covariance matrix \( \tau^2 \Sigma^{-1} \), where 
\[
\tau^2 = \frac{1}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sqrt{f^2} dx^2} \quad \text{and} \quad \Sigma = \lim_{n \to \infty} \frac{X'X}{n}
\]
(1.2.10)

Jaeckel (1972) introduced the idea of a dispersion function. Let \( D(Z) \) be a 
translation invariant measure of dispersion for \( (z_1, z_2, \ldots, z_n) \) (i.e. \( D(Z+b) = D(Z) \), 
for all \( b = (b_1, b_2, \ldots, b_n) \). His estimate of \( \beta \) is that vector which minimizes 
\( D(Y - X\beta) \), where \( X \) is the \( nxp \) design matrix. If \( D(Z) \) is the variance of \( Z \), then 
the minima of \( D(Y - X\beta) \) will be the least square estimates. Suppose \( a_1 \leq a_2 \leq \ldots \leq a_n \) is a nonconstant sequence of scores such that \( a_k + a_{n-k+1} = 0 \). Jaeckel (1972) 
defined the following measure of dispersion of the vector \( Z \):
where the $Z(i)$ are the ordered values of the $Z_i$'s. Further, if $R_1, R_2, ..., R_n$ are the ranks of the $Z_i$'s, then we can write $D(Z) = \sum_{i=1}^{n} a(R_i) Z_i$, where we assign average ranks to the tied $Z$ values. A rank estimate of $\beta$ then minimizes

$$D(\beta) = \sqrt{\frac{12}{n}} \sum_{i=1}^{n} R(Y_i - X_i'\beta) \left( \frac{Y_i - X_i'\beta}{n+1} \right)$$  

where $R(Y_i - X_i'\beta)$ is the rank of the $i$th residuals. This even, location-free measure is a linear, rather than quadratic function of the residuals. It is hoped that the estimates generated by this dispersion function will be more robust than least square estimates since the influence of outliers enters in a linear rather than a quadratic fashion. McKean and Schrader (1980) showed that this dispersion function is a pseudo-norm.

1.3 Inference based on the Dispersion Function

Jaeckel showed that the dispersion function defined in (1.2.11) is nonnegative, convex and continuous as a function of $\beta$, so that it is always possible to minimize it to obtain estimates of the $\beta$'s. Even though these estimates need not be uniquely defined, Jaeckel showed that $\sqrt{n}$ times the diameter of the set of minimizing points converges to zero in probability. It turns out that Jureckova's proposal to estimate $\beta$ is equivalent to making the gradient of the dispersion function approximately zero. Jaeckel (1972) showed that $\sqrt{n}$ times the difference between his estimates and Jureckova's estimates converges to zero in probability. Thus the two estimates have the same asymptotic distribution.
Suppose we wish to test the hypothesis \( H_0: A\beta = c \) versus \( H_1: A\beta \neq c \), where \( A \) is a full rank \( q \times p \) matrix with \( q < p \) and \( c \) is a specified \( q \times 1 \) vector. Often we take \( c = 0 \). It can be shown that, under the null hypothesis, \( Q(x) = \tau^2 (A\hat{\beta} - c)' [A(X'X)^{-1}A]^{-1} (A\hat{\beta} - c) \) converges to a chi-square distribution with \( q \) degrees of freedom. If we replace \( \tau^2 \) by a consistent estimate we can use \( Q \) as a test statistic.

Another approach to testing the hypothesis \( H_0: A\beta = c \) versus \( H_1: A\beta \neq c \) depends on fitting both the full model and the reduced model induced by the null hypothesis. Let \( \hat{\beta}_F \) and \( \hat{\beta}_H \) be the full model and reduced model estimates of \( \beta \), respectively. McKean and Hettmansperger (1975) showed that, under the null hypothesis \( D^* = 2[ D(\hat{\beta}_H) - D(\hat{\beta}_F)]/\tau \) is also asymptotically chi-square with \( q \) d.f. In addition \( D^*(\tau) \) follows an approximate chi-square distribution with \( q \) d.f., provided \( \hat{\tau} \) is a consistent estimate of \( \tau \).

McKean and Hettmansperger (1976,1978) developed a procedure for estimating \( \tau \). Aubuchon (1982) extended their approach to estimate \( \tau \) when the error distribution is not symmetric and the scores are Wilcoxon scores. Also Koul, Sievers and McKean (1980) developed a procedure to estimate \( \tau \).

A third approach is based on the gradient vector \( S(\beta) \) of \( D(\beta) \). If the reduced model estimate \( \hat{\beta}_H \) of \( \beta \) is plugged into \( S(\beta) \), then a test can be based on a quadratic form in \( S(\hat{\beta}_H) \). It can be shown that this quadratic form converges to a chi-square distribution with \( q \) d.f. under \( H_0 \). This is like Rao's score statistic in the maximum likelihood approach.

1.4 Rank Based Inference for Some Repeated Measures Designs

In social, biological, behavioral and medical sciences experiments the experimental units are often human subjects. Because of differences in background, experience and
training, the responses of different subjects will show relatively large variability. This variability between subjects becomes a part of the experimental error. Unless we control it, in some situations it will inflate the error mean-square. As a result we might be unable to find real differences between treatments. This problem can be overcome by designs in which all the treatments are applied to each subject. Such designs are called Repeated Measures Designs. By using subjects themselves as blocks we control the variability between subjects. In fact, each subject acts as its own control. In this thesis we will consider Repeated Measures Randomized Complete Block Designs (RCBD), Repeated Measures Balanced Incomplete Designs (BIBD) and the Two Way Model with one repeated measures factor and one grouping factor. Our aim will be to make inferences by ranking residuals within subjects.

1.4.1 Repeated Measures RCBD

Suppose there are p treatments and all p treatments are applied to each subject. Let \( Y_{ij} \) denote the observation corresponding to the ith treatment and the jth subject.

The data for this experiment can be presented as follows.

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data for RCBD</td>
</tr>
<tr>
<td>Subjects</td>
</tr>
<tr>
<td>Treatments</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>i</td>
</tr>
<tr>
<td>p</td>
</tr>
</tbody>
</table>
A model for this experiment is $Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$, $i = 1, 2, ..., p; j = 1, 2, ..., n$. For the jth subject, the model becomes:

$$Y_j = \mu 1 + I\alpha + \epsilon_j,$$  
(1.4.1.1)

where $Y_j = (Y_{1j}, ..., Y_{ij}, ..., Y_{pj})'$, $\mu$ = overall mean, $\alpha = (\alpha_1, \alpha_2, ..., \alpha_p)'$, and $\epsilon_j = (\epsilon_{1j}, \epsilon_{2j}, ..., \epsilon_{pj})'$, $j = 1, ..., n$. The $\epsilon_j$'s are i.i.d. random vectors such that the elements of the $\epsilon_j$'s are exchangeable random variables.

Table 2

<table>
<thead>
<tr>
<th>TREAT</th>
<th>1</th>
<th>2</th>
<th>j</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$Y_{11} - \mu - \alpha_1$</td>
<td>$Y_{12} - \mu - \alpha_1$</td>
<td>$Y_{1j} - \mu - \alpha_1$</td>
<td>$Y_{1n} - \mu - \alpha_1$</td>
</tr>
<tr>
<td>2</td>
<td>$Y_{21} - \mu - \alpha_2$</td>
<td>$Y_{22} - \mu - \alpha_2$</td>
<td>$Y_{2j} - \mu - \alpha_2$</td>
<td>$Y_{2n} - \mu - \alpha_2$</td>
</tr>
<tr>
<td>i</td>
<td>$Y_{i1} - \mu - \alpha_i$</td>
<td>$Y_{i2} - \mu - \alpha_i$</td>
<td>$Y_{ij} - \mu - \alpha_i$</td>
<td>$Y_{in} - \mu - \alpha_i$</td>
</tr>
<tr>
<td>p</td>
<td>$Y_{p1} - \mu - \alpha_p$</td>
<td>$Y_{p2} - \mu - \alpha_p$</td>
<td>$Y_{pj} - \mu - \alpha_p$</td>
<td>$Y_{pn} - \mu - \alpha_p$</td>
</tr>
</tbody>
</table>

The dispersion for the jth subject is $D_j(\alpha)$

$$= \sqrt{12} \sum_{i=1}^{p} \left[ \frac{R(Y_{ij} - \alpha_i)}{p+1} - \frac{1}{2} \right] [Y_{ij} - \alpha_i]$$  
(1.4.1.2)

The combined dispersion function is

$$D(\alpha) = \sqrt{12} \sum_{j=1}^{n} \sum_{i=1}^{p} \left[ \frac{R(Y_{ij} - \alpha_i)}{p+1} - \frac{1}{2} \right] [Y_{ij} - \alpha_i]$$  
(1.4.1.3)
We minimize $D(\alpha)$ with respect to $\alpha$ to get the R-estimate of $\alpha$. This is equivalent to simultaneously solving $s_i(\alpha) = \sqrt{12} \sum_{j=1}^{n} \left[ \frac{R(Y_{ij}-\alpha_i)}{p+1} - \frac{1}{2} \right] [Y_{ij}-\alpha_i] = 0$, $i = 1,2,..., p$. Let $\hat{\alpha}$ be the solution to these equations. This design will be discussed further in Chapter 2.

### 1.4.2. Repeated Measures BIBD

In these designs there are $n$ blocks (subjects or judges) and $t$ treatments. There are $k \leq t$ residuals ranked within blocks. Every treatment appears in $rn$ blocks and every treatment appears with every other treatment an equal number of times ($\lambda$). A model for this type of design is given by $Y_{ij} = (\mu + \alpha_i + \beta_j + \epsilon_{ij})n_{ij}$, $i = 1,2,..., t$, and $j = 1,2,..., n$, where $n_{ij} = 1$ if the $i$th treatment appears in the $j$th block and $n_{ij} = 0$ otherwise, $\mu$ = (overall mean), $\alpha_i$ = (effect of the $i$th treatment), $\beta_j$ = (effect of the $j$th block), and $\epsilon_{ij}$ is the error term. In such a design we will assume that the error vectors are i.i.d. and that the elements within each error vector are exchangeable random variables.

The dispersion for the $j$th subject is then

$$D_j(\alpha) = \sqrt{12} \sum_{i=1}^{t} n_{ij} \left[ \frac{R(Y_{ij}-\alpha_i)}{k+1} - \frac{1}{2} \right] [Y_{ij}-\alpha_i]$$  \hspace{1cm} (1.4.2.1)

and the combined dispersion function is

$$D_j(\alpha) = \sqrt{12} \sum_{j=1}^{n} n_{ij} \left[ \frac{R(Y_{ij}-\alpha_i)}{k+1} - \frac{1}{2} \right] [Y_{ij}-\alpha_i]$$  \hspace{1cm} (1.4.2.2)

It can be shown that this dispersion function of $\alpha$ is nonnegative, continuous, even, location-free, convex and scale invariant. Hence, a rank estimate of $\alpha$ can be obtained
by minimizing the dispersion function. This model will be discussed further in Chapter four.

1.4.3 Two Way Models With One Repeated Measures Factor and One Grouping Factor

Let us consider a two way model with $p$ rows (treatments) and $q$ columns (grouping factor) in which each subject receives each row treatment but only one column treatment. (This would occur, for example, if the columns represented race, sex, degree of illness etc.) We assume that there are $n$ subjects in the study.

Let $Y_{jk} = (Y_{1jk}, Y_{2jk}, \ldots, Y_{pjk})$ be the vector of observations on the $k$th subject who receives the $j$th column treatment, $j = 1, 2, \ldots, q$ and $k = 1, 2, \ldots, n$. Our model is then $Y_{ijk} = \mu + \alpha_i + \delta_j + \gamma_{ij} + \epsilon_{ijk}$, where $\mu$ = overall mean, $\alpha_i$ = (effect of the $i$th treatment), $\delta_j$ = (effect of the $j$th group treatment), $\gamma_{ij}$ = (interaction between the $i$th row treatment and the $j$th group treatment), and $\epsilon_{ijk}$ is the error term. Let $\varepsilon_{jk}$ be the error vector corresponding to $Y_{jk}$. We assume the error vector $\varepsilon_{jk}$ are i.i.d. and that the elements of each $\varepsilon_{jk}$ are exchangeable random variables. A dispersion function for the $k$th subject on the $j$th group is

$$D_{jk}(\alpha, \delta, \gamma) = \sqrt{12} \sum_{i=1}^{p} \left( \frac{R(Y_{ijk} - \mu - \alpha_i - \delta_j - \gamma_{ij})}{p+1} \right)^2 \frac{1}{2} [ Y_{ijk} - \mu - \alpha_i - \delta_j - \gamma_{ij} ]$$

(1.4.3.1)

$$= \sqrt{12} \sum_{i=1}^{p} \left( \frac{R(Y_{ijk} - \alpha_i - \gamma_{ij})}{p+1} \right)^2 \frac{1}{2} [ Y_{ijk} - \alpha_i - \gamma_{ij} ] ,$$

where $R$ stands for the rank of the residuals within subjects. The combined dispersion function for the model is

$$D(\alpha, \delta, \gamma) = \sqrt{12} \sum_{k=1}^{n} \sum_{j=1}^{q} \sum_{i=1}^{p} \left( \frac{R(Y_{ijk} - \alpha_i - \gamma_{ij})}{p+1} \right)^2 \frac{1}{2} [ Y_{ijk} - \alpha_i - \gamma_{ij} ]$$

(1.4.3.2).
This dispersion function shares all the properties of Jaeckel's dispersion function. We will come back to this model in Chapter 6.

1.5 Thesis Results

Statistical inference for Repeated Measures RBCD, BIBD and Two way model designs based on ranks of the residuals within subjects have been developed in this thesis. In fact, inference for these designs depends on several important results. Jureckova (1971) proved the asymptotic linearity of the gradient vector for linear models with i.i.d. errors under some complicated assumptions about the design matrix. Heiller and Willers (1979) proved the asymptotic linearity of the gradient vector for Jaeckel's dispersion function by invoking the convexity of the function and without using Jureckova's conditions on the design matrix. In fact, they used Huber's condition which is a necessary and sufficient condition for the asymptotic normality of the least square estimates. Using their approach the approximate linearity of the gradient vectors for the dispersion functions associated with these designs have been established.

Three different statistics for testing linear hypotheses have been developed for these repeated measures designs. The first is a Wald type statistic based on the full model rank estimate. The second test statistic is based on the drop in dispersion, an idea similar to that employed in least square and maximum likelihood techniques. The third is a Wald type statistic based on the gradient vector evaluated at a reduced model estimate. It is similar in principle to Rao's score maximum likelihood statistic. It has been shown that, under the null hypothesis, all three of these statistics have approximate chi-square distributions for large n. All are also asymptotically distribution-free, once a scale parameter has been replaced by a consistent estimate. Only the third one is strictly distribution-free under an appropriate simple null hypothesis \( \alpha_1 = \alpha_2 = \ldots = \alpha_p \). The
results for repeated measures RCBD are developed in Chapter 2, those for Repeated Measures BIBD in Chapter 5, and the results for the Two Way Model are presented in Chapter 6. Two different procedures have been prepared for estimating the scale parameter

\[ \tau = \sqrt{\frac{1}{12} \int_{-\infty}^{\infty} (f(x,x)dx)^2}, \]  

which appears in the denominator of the first two test statistics, in the covariance matrices of the rank estimates and in the ARE's and plays a role similar to that of \( \sigma^2(1-\rho) \) in the least square procedures. These estimates are shown to be consistent. A computer program has been developed for obtaining the first estimate. The results will be discussed in Chapter 3.

The asymptotic relative efficiencies of these rank tests with respect to their least square counterparts have been developed. It is seen that for nonnormal distributions the rank tests do better than their least square counterparts.

To date, distribution-free methods based on rank estimates for making multiple comparisons for contrasts have not been available. On the basis of the full model estimates we will be able to make such multiple comparisons similar to the LSD, Bonferroni, Tukey and Scheffe methods.

Finally, the results of a Monte Carlo study are reported in Chapter 4, and some further research problems are discussed in Chapter 7.

1.6 Assumptions for This Thesis

To prove the asymptotic linearity of the gradient vector for all the repeated measures designs we need Huber's condition. As we assume the number of subjects is large,
Huber's condition is automatically satisfied. We also need the following conditions for the joint p.d.f. of the observations within subjects.

**Assumption 1:** The bivariate c.d.f. $F(\cdot,\cdot)$ is absolutely continuous and
\[
\int_{-\infty}^{\infty} f(y_1,y_1) dy_1 < \infty \text{ where } f(y_1,y_1) \text{ is a bivariate p.d.f. evaluated at } y_1.
\]

**Assumption 2:** $a(i) = -a(n-i+1)$ and $a(1) \leq a(2) \leq \ldots \leq a(n)$ but not all $a$'s are equal.

**Assumption 3:** The errors within subjects are exchangeable.

**Assumption 4:** The problem considered in this dissertation are not model free since we are assuming a linear model.

### 1.7 Basic Strategy and Lemmas

The most important thing in this thesis is to show that the gradient vector is approximately linear in the parameters. We need some lemmas to prove this result.

**Lemma 1 (Rockafeller (1970))**

Let the sequence $\{f_n(x)\}$ converge for all $x \in C'$, where $C'$ is dense in $C$. Then the function $f_n$ converges on the whole set $C$ to a finite and convex function $f$ and the convergence is uniform on each compact subset of $C$.

**Lemma 2 (Rockafeller (1970))**

If, under the assumptions of Lemma 1, the limit function $f$ is differentiable, then
\[
\lim_{n} \frac{\delta f_n(x)}{\delta x} = \frac{\delta f(x)}{\delta x} \text{ for all } x \in C \text{ and the convergence is uniform on each compact subset of } C.
\]

The following lemma is also used in proving the asymptotic linearity of the gradient vector.
Lemma 3 (Heiller and Willers (1979))

Let \( \{f_n(x)\}, f, C \) and \( C' \) be as given in Lemmas 1 and 2 and let \( f \) be differentiable. If \( \lim_n \frac{\delta f_n(x)}{\delta x} = \frac{\delta f(x)}{\delta x} \) for all \( x \in C' \) and \( \lim_n f_n(x_0) = f(x_0) \) for at least one \( x_0 \in C' \), then \( \lim_n \frac{\delta f_n(x)}{\delta x} = \frac{\delta f(x)}{\delta x} \) for all \( x \in C \) and the convergence is uniform on each compact subset of \( C \).

Lemma 4 (Chung (1974))

Let \( a_n \) be a sequence of numbers such that \( \lim_n a_n = 0 \). Then \( \sum_{i=1}^{n} \frac{a_i}{n} \) converges to zero as \( n \) tends to \( \infty \).

Lemma 5 (Lebesgue Convergence Theorem Royden (1968))

Let \( g \) be integrable over \( E \) and suppose that \( f_n \) is a sequence of measurable functions such that \( |f_n(x)| \leq g(x) \) on \( E \) and such that \( f_n(x) \) converges to \( f(x) \) almost everywhere on \( E \).

Then \( \int_{E} f = \lim_{n} \int_{E} f_n \).

Lemma 6 (Kolmogrov's Theorem 1 (SLLN) Rao (1973))

Let \( \{X_i\}, i = 1, 2, \ldots, \) be a sequence of independent random variables such that \( E(X_i) = \mu_i \) and \( V(X_i) = \sigma_i^2 \). Then \( \sum_{i=1}^{\infty} \frac{\sigma_i^2}{i^2} < \infty \) implies \( \frac{1}{n} \sum_{i=1}^{n} X_i - \frac{1}{n} \sum_{i=1}^{n} \mu_i \) converges to zero almost surely.

Lemma 7 (Multivariate Central Limit Theorem Rao (1973))

Let \( X_1, X_2, \ldots, X_n \) be a sequence of independent \( p \)-dimensional random variables such that \( E(X_i) = 0 \) and \( V(X_i) = \Sigma_i \). Suppose that \( \frac{1}{n} \sum \Sigma_i \) converges to \( \Sigma \neq 0 \) and for every \( \varepsilon > 0 \), \( \frac{1}{n} \sum_{i=1}^{n} \int_{|X_i| > \varepsilon \sqrt{n}} |X_i|^2 \, dF_i \) converges to zero where \( F_i \) is the c.d.f. of \( X_i \) and
\|X_i\| is the Euclidean norm of the vector \(X_i\). Then the random variable \(\sqrt{n} \sum_{i=1}^{n} \frac{X_i}{n}\) converges in distribution to the \(p\)-variate normal distribution with mean vector zero and dispersion matrix \(\Sigma\).

The strategy for the limiting distribution is as follows:

1) From the linear approximation to the gradient vector we construct a quadratic approximation to the original dispersion function. 2) We show that the vector which minimizes the quadratic approximation has an asymptotic multivariate normal distribution, and 3) we show that the minimas of the original quadratic function and the quadratic approximation coincide asymptotically.
CHAPTER II
REPEATED MEASURES RCBD

2.1 Introduction

Suppose we have a repeated measures factor and the p levels of this factor are
applied to n subjects. Let \( Y_{ij} \) be the observation corresponding to the ith treatment and
jth subject. An appropriate model would be

\[
Y_{ij} = \mu + \alpha_i + \epsilon_{ij}, \quad i = 1, \ldots, p; \quad j = 1, \ldots, n \tag{2.1.1}
\]

where \( \alpha_i \) = the effect of ith repeated measures treatment; \( \mu \) is overall mean, \( \epsilon_{ij} \) is the
error term and \( \sum_{i=1}^{p} \alpha_i = 0 \). For the jth subject the model would be
\( Y_j = \mu 1 + I\alpha + I\epsilon_j; \quad j = 1, 2, \ldots, n; \quad \epsilon_j's \) are i.i.d.random vectors. Further \( \epsilon_j's \) are
continuous and elements of \( \epsilon_j \) are exchangeable random variables.

2.2. Rank Estimates of One-Way Repeated Measures Model

Our main objective in this Section is to develop R-estimates of the \( \alpha_i \)'s. Using
Jaeckel's (1972) approach, we can define a dispersion function within the jth subject by

\[
D_j(\alpha) = \sqrt{12} \sum_{i=1}^{p} \left( \frac{R(Y_{ij}-\alpha)}{p+1} - \frac{1}{2} \right) \{ Y_{ij} - \alpha_i \} \tag{2.2.1}
\]
where \( R \) denotes the rank within the subject. By Theorem 1, Jaeckel (1972), \( D_j(\alpha) \) is a nonnegative, continuous and convex function of \( \alpha \). The overall dispersion function of the model can be written as \( D(\alpha) = \sum_{j=1}^{n} D_j(\alpha) \)

\[
= \sqrt{12} \sum_{j=1}^{n} \sum_{i=1}^{p} \left\{ \frac{R(Y_{ij} - \alpha_j)}{p+1} \cdot \frac{1}{2} \right\} (Y_{ij} - \alpha_i)
\] (2.2.2)

Since the sum of nonnegative, continuous and convex functions is again a nonnegative, continuous and convex function, \( D(\alpha) \) is a nonnegative, continuous and convex function of \( \alpha \).

A rank estimate of \( \alpha \) is the value \( \hat{\alpha} \) which minimizes \( D(\alpha) \). The domain (\( \alpha \) space) of \( D(\alpha) \) is divided into a finite number of convex polygonal subsets, on each of which \( D(\alpha) \) is linear function of \( \alpha \). The partial derivatives exist almost everywhere and the negative of the partial derivatives are given by the vector

\[
S(\alpha) = (s_1(\alpha), s_2(\alpha), ..., s_p(\alpha))'
\] (2.2.3)

where

\[
s_i(\alpha) = \sqrt{12} \sum_{j=1}^{n} \left\{ \frac{R(Y_{ij} - \alpha_j)}{p+1} \cdot \frac{1}{2} \right\}
\] (2.2.4)

i = 1,..., p. Hence minimizing \( D(\alpha) \) is equivalent to solving \( S(\alpha) = 0 \). But this is an extension of Hodges-Lehmann estimation to the linear model since \( E_0(S(\alpha_0)) = 0 \).

2.2.1. Asymptotic Distribution of \( S(\alpha) \) under True Value \((\alpha)\)

Without loss of generality, let us assume that \( \alpha_0 = 0 \). Therefore \( E_0(S(0)) = 0 \) and \( \text{Var}_0(S(0)) = \Sigma \) where \( \Sigma = \frac{1}{p+1} [pI - 11'] \) where I is an identity matrix of order pxp and 1 is a px1 column vector of 1's. The vector \( S(0) \) can be written as \( Z_1 + Z_2 + ... + Z_n \).
where the ith component of $Z_j$ is given by $\sqrt{n} \left\{ \sum_{j=1}^{p} \left( \frac{R(Y_{ij} - \alpha_i)}{p+1} - \frac{1}{2} \right) \right\}$. $Z_i$'s are i.i.d. random vectors since we are ranking within subjects. Also $E_0[Z_j] = 0$ and $Cov_0[Z_j] = \Sigma$. By a multivariate central limit theorem (a particular case of Lemma 7, Section 1.7)

$$\sqrt{n} \frac{S(O)}{n} \sim N(0, \Sigma) \text{ as } n \text{ tends to infinity.}$$ (2.2.1.1)

### 2.2.2 Linear Approximation to the Negative of the Gradient Vector

A linear approximation to the gradient vector is crucial to the development of the distribution theory of rank estimates and tests of repeated measures model. First we would like to find the expected value of the negative of the gradient vector under true value $\alpha^0$. Without loss of generality let us assume that $\alpha^0 = 0$.

We have

$$E_0(s_i(\alpha)) = E_0[\sqrt{12} \sum_{j=1}^{n} \left( \frac{R(Y_{ij} - \alpha_i)}{p+1} - \frac{1}{2} \right)], \ i = 1, ..., p$$ (2.2.2.1)

But

$$R(Y_{ij} - \alpha_i) = 1 + \sum_{i \neq i}^{p} 1 \left( y_{ij} - \alpha_i \leq y_{ij} - \alpha_i \right).$$

Therefore

$$E_0(R(Y_{ij} - \alpha_i)) = 1 + \sum_{i \neq i}^{p} P(0(y_{ij} - \alpha_i \leq y_{ij} - \alpha_i))$$

$$= 1 + \sum_{i \neq i}^{p} P(0(y_{ij} - y_{ij} \leq \alpha_i - \alpha_i))$$

$$= 1 + \sum_{i \neq i}^{p} \int_{-\infty}^{\infty} \int_{-\infty}^{y_{ij} - \alpha_i - \alpha_i} f(y_{ij}, y_{ij}) dy_{ij} dy_{ij}$$

$$= 1 + g(\alpha_i, \alpha_i)$$ (2.2.2.2)
Expanding \( g(\alpha_i, \alpha_i') \) as a function of \( \alpha_i \) and \( \alpha_i' \) around \((0,0)\) and retaining the first two terms, we get

\[
g(\alpha_i, \alpha_i') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y_{ij}, y_{ij}) dy_{ij} dy_{ij} + (\alpha_i' - \alpha_i) \int_{-\infty}^{\infty} f(y_{ij}, y_{ij}) dy_{ij} 
\]

\[
= \frac{1}{2} + (\alpha_i' - \alpha_i) \int_{-\infty}^{\infty} f(y_{ij}, y_{ij}) dy_{ij} = \frac{1}{2} + (\alpha_i' - \alpha_i) \int_{-\infty}^{\infty} f(y, y) dy.
\]

Therefore \( E_0(s_i(\alpha)) = E_0[\sqrt{12} \sum_{j=1}^{n} \left( \frac{R(Y_{ij} - \alpha_i)}{p+1} - \frac{1}{2} \right)] \)

\[
= -n\sqrt{12} \left( \frac{1}{p+1} \right) \int_{-\infty}^{\infty} f(y, y) dy \cdot (-1, ..., p - 1, ..., -1) \cdot (\alpha_1, \alpha_2, ..., \alpha_i, ..., \alpha_p)
\]

Therefore

\[
E_0\left( \frac{1}{n} S(\alpha) \right) = \frac{-1}{\tau} \Sigma \alpha
\]

(2.2.2.3)

where \( \tau = \frac{1}{\sqrt{12} \int_{-\infty}^{\infty} f(y, y) dy} \). Hence under true value \( \alpha^0 \),

\[
E_0\left( \frac{1}{n} S(\alpha) \right) = -\frac{1}{\tau} \Sigma (\alpha - \alpha^0)
\]

(2.2.2.4)

Now we would like to prove a Theorem.

**Theorem 1:** \( \frac{1}{n} S(\alpha) \) converges almost surely to \(-\frac{1}{\tau} \Sigma (\alpha - \alpha^0)\) as \( n \) tends to infinity.

**Proof:** We can write \( s_i(\alpha) = \frac{1}{n} \Sigma_{j=1}^{n} t_j \) where \( t_j = \left( \frac{R(Y_{ij} - \alpha_i)}{p+1} \right) \cdot \frac{1}{2} \). Hence \( Z_j \) are independently distributed. \( \text{Var}_0(t_j) \leq E_0(t_j^2) \leq 12p^2 \). Therefore \( \sum_{j=1}^{n} \left( \frac{\text{Var}_0(t_j)}{f^2} \right) \leq 12p^2 \)
\((\sum \frac{1}{j^2}) < \infty\). Hence by Kolmogrov's SLLN (Lemma 6, Section 1.7), the result follows. It also follows from Theorem 1, \(\frac{1}{n} S(\alpha)\) converges to \(\frac{1}{\tau} \Sigma(\alpha - \alpha^0)\) in probability as \(n\) tends to infinity.

**Theorem 2:** Suppose \(\int f(y,y)dy < \infty\) and true value is \(\alpha^0 = 0\). For \(\varepsilon > 0\) and \(\Delta > 0\), \(\lim P_0 \sqrt{n} \left[ \frac{1}{n} S(\frac{\Delta}{\sqrt{n}}) - \frac{1}{n} S(0) + \frac{1}{\tau} \Sigma \Delta || \geq \varepsilon \right] = 0\) as \(n\) tends to infinity.

**Proof:** Using (2.2.2.4), we can write

\[
E_0 \sqrt{n} \left[ \frac{1}{n} S(\frac{\Delta}{\sqrt{n}}) - \frac{1}{n} S(0) \right] = \frac{1}{\tau} S \Delta
\]

and

\[
\text{Var}_0 \left( \sqrt{n} \left[ \frac{1}{n} s_i(\frac{\Delta}{\sqrt{n}}) - \frac{1}{n} s_i(0) \right] \right) = \frac{1}{n} \sum_{j=1}^{n} \left( R(y_{ij} - \frac{\Delta_i}{\sqrt{n}}) - R(y_{ij}) \right)^2 = \frac{1}{n} \sum_{j=1}^{n} a_j
\]

where \(a_n = E_0 \left( R(y_{in} - \frac{\Delta_i}{\sqrt{n}}) - R(y_{in}) \right)^2 \) but \( \left( R(y_{in} - \frac{\Delta_i}{\sqrt{n}}) - R(y_{in}) \right)^2 \leq (p-1)^2 \) and

\[
\lim_n \left( R(y_{in} - \frac{\Delta_i}{\sqrt{n}}) - R(y_{in}) \right)^2 \leq \lim_n \left[ \sum_{k=1}^{p} \left( I(y_{kn} - \frac{\Delta_k}{\sqrt{n}} \leq y_{in} - \frac{\Delta_i}{\sqrt{n}}) - I(y_{kn} \leq y_{in}) \right) \right] = 0
\]

where \(I\) is an indicator function. Hence by Lebesgue's Convergence Theorem (Lemma 5, Section 1.6) \(\lim n a_n = \lim_n E_0 \left( R(y_{in} - \frac{\Delta_i}{\sqrt{n}}) - R(y_{ij}) \right)^2 = 0\). Since \(a_n\) converges to 0 as \(n\) tends to infinity then by Lemma 4, Section 1.6, \(\lim_n = \frac{1}{n} \sum_{j=1}^{n} a_j = 0\). Hence \(\lim_n \text{Var}_0 \sqrt{n} \left[ \frac{1}{n} s_i(\frac{\Delta}{\sqrt{n}}) - \frac{1}{n} s_i(0) \right] = 0\). By appealing to Markov's inequality we have \(P_0 \sqrt{n} \left[ || \frac{1}{n} S(\frac{\Delta}{\sqrt{n}}) - \frac{1}{n} S(0) + \frac{1}{\tau} \Sigma \Delta || \geq \varepsilon \right] = 0\)
Theorem 3: \( \lim_{n \to \infty} P_0 \{ \sup |A| \leq c \sqrt{n} \| A \| \leq c \sqrt{n} \left[ \frac{1}{n} S \left( \frac{\Delta}{\sqrt{n}} \right) - \frac{1}{n} S(0) + \frac{1}{\tau} \Sigma \Delta \| \geq \varepsilon \right] \} = 0 \)

where \( c > 0 \).

Theorem 4: \( \lim_{n \to \infty} P_0 \{ \sup |A| \leq c \| D_n(A) - Q_n(A) \| > \varepsilon \} = 0 \) where

\[
D_n(A) = \sqrt{12} \sum_{j=1}^{n} \sum_{i=1}^{p} \frac{R(Y_{ij} - \Delta_i)}{p + 1} - \frac{1}{\sqrt{n}} \left \{ Y_{ij} - \frac{\Delta_i}{\sqrt{n}} \right \}
\]

and

\[
Q_n(A) = D_n(0) + \frac{1}{2} \frac{1}{\tau} \Delta' \Sigma \Delta - \frac{1}{\sqrt{n}} \Delta' S(0)
\]  

(2.2.2.6)

Cor: Theorems 2, 3 and 4 are equivalent.

Proof: We will be using the approach in Heiller and Willers (1979). Both \( D_n(A) \) and \( Q_n(A) \) are convex functions of \( A \) with

\[
D_n(0) = Q_n(0) \quad \text{and} \quad \frac{\delta(D_n(A) - Q_n(A))}{\delta \Delta} = -\frac{1}{\sqrt{n}} S \left( \frac{\Delta}{\sqrt{n}} \right) + \frac{1}{\sqrt{n}} S(0) - \frac{1}{\tau} \Sigma \]  

(2.2.2.7)

Since for any \( \Delta \in \mathbb{R}^p \), the right hand side of the above converges to zero in probability (by Theorem 2), we can select from each infinite index set \( n^* \in \mathbb{n} \) another infinite index set \( n^* \supset n^{**} \) such that \( \| \frac{\delta(D_n(A) - Q_n(A))}{\delta \Delta} \|_{n^{**}} \) converges to 0 almost surely. By using standard diagonal sequence of arguments it is therefore possible to find an index subset \( \mathfrak{f} \) such that \( \| \frac{\delta(D_n(A) - Q_n(A))}{\delta \Delta} \|_{n^{**}} \) converges to zero for all rational \( \Delta \in \mathbb{R}^p \).

Let us introduce the convex function
\[ H_n(\Delta) = D_n(\Delta) - D_n(0) + \frac{1}{\sqrt{n}} \Delta' S(0) \]  
(2.2.2.8)

Therefore

\[ D_n(\Delta) - Q_n(\Delta) = D_n(\Delta) - D_n(0) - \frac{1}{2} \Delta' \Sigma \Delta + \frac{1}{\sqrt{n}} \Delta' S(0) = H_n(\Delta) - \frac{1}{2} \frac{1}{\tau} \Delta' \Sigma \Delta \]  
(2.2.2.9)

and

\[ \frac{\delta(D_n(\Delta)Q_n(\Delta))}{\delta A} = \frac{\delta H_n(\Delta)}{\delta A} - \frac{1}{\tau} \Sigma A \]  
(2.2.2.10)

Since \( \|\frac{\delta(D_n(\Delta) - Q_n(\Delta))}{\delta A}\| \) converges to zero for all rational \( \Delta \in \mathbb{R}^p \) the first condition of Lemma 3 (Section 1.6) is satisfied. Since \( D_n(0) - Q_n(0) = 0 \), the second condition is satisfied. Hence, Theorem 2 implies Theorem 4. Theorem 4 implies Theorem 3 can be proved in the same way by using Lemma 2. Theorem 3 implies Theorem 4 automatically.

2.2.3. Asymptotic Distribution of Rank Estimates Under Full Model

Let \( \hat{\alpha} \) minimize \( Q_n(\alpha) \); then

\[ \hat{\alpha} = \frac{1}{n} \tau \Sigma^{-1} S(0) \]  
(2.2.3.1)

If we assume that \( \alpha^0 \) is the true value of \( \alpha \) then

\[ \hat{\alpha} = \alpha^0 + \frac{1}{n} (\tau \Sigma^{-1} S(\alpha^0)) \]  
(2.2.3.2)

From Theorem 4 \( \hat{\alpha} \) and \( \hat{\alpha} \) must coincide asymptotically. Since \( \sqrt{n} \frac{S(\alpha^0)}{n} \sim N(0, \Sigma) \) as \( n \) tends to infinity, therefore \( \sqrt{n} (\hat{\alpha} - \alpha^0) \) converges to the multinormal distribution with mean zero and variance \( \tau^2 \Sigma \).  
(2.2.3.3)
Hence $\sqrt{n} (\hat{\alpha} - \alpha)$ converges to the multinormal distribution with mean zero and variance $\tau^2 \Sigma$. (2.2.3.4)

2.2.4 Asymptotic Distribution of Rank Estimates under Reduced Model

We would like to minimize $D(\alpha) = \sum_{j=1}^{n} D_j(\alpha) = \sqrt{12} \sum_{j=1}^{n} \sum_{i=1}^{p} \left( \frac{R(Y_{ij} - \alpha_i)}{p + 1} - \frac{1}{2} \right) \{ Y_{ij} - \alpha_i \}$ subject to the restriction $A\alpha = 0$ where $A1 = 0$ and rows of $A$ are mutually orthogonal.

The vector $\hat{\alpha}_H$ which minimizes $D(\alpha)$ subject to the restriction $A\alpha = 0$ is called the reduced model estimate of $\alpha$. To minimize $D(\alpha)$ subject to $A\alpha = 0$, we will use the method of Lagrangian multipliers. Let us define $\phi = D(\alpha) + (\lambda)' A\alpha$. Therefore $\frac{\delta \phi}{\delta \alpha} = S(\alpha) + A' \lambda = 0$ and $\frac{\delta \phi}{\delta \lambda} = A\alpha = 0$. Now we would like to get an approximation to $\frac{\delta \phi}{\delta \alpha}$ using Theorem 3 Section 2.2.2. It can be shown that

$$\frac{1}{\sqrt{n}} \frac{\delta \phi}{\delta \alpha} = \frac{1}{\sqrt{n}} S(\alpha^0) - \sqrt{n} \frac{1}{\tau} \Sigma \alpha - \alpha^0)) + \frac{1}{\sqrt{n}} A' \lambda + o_p(1) \quad (2.2.4.1)$$

where $o_p(1)$ tends to zero in probability uniformly for all values of $\alpha$ such that $\sqrt{n} \| \alpha - \alpha^0 \| \leq c$ for any $c > 0$ and $\alpha^0$ is the true value under the null hypothesis. By Theorem 4, Section 2.2.2 we can construct a quadratic approximation to $\phi$ such as $\psi = Q(\alpha) + \lambda' A\alpha$. Hence the minima of $\phi$ and $\psi$ will coincide asymptotically. Let $\tilde{\alpha}_H$ be the minima of $Q(\alpha)$ subject to the restriction $A\alpha = 0$. Then $\tilde{\alpha}_H$ is a solution of the equation $\frac{\delta \psi}{\delta \alpha} = 0$ and $\frac{\delta \psi}{\delta \lambda} = 0$. By solving those two equations, we get $\tilde{\alpha}_H = (I - A'A) \bar{\alpha}$ where $\bar{\alpha}$ is the minima of $Q(\alpha)$ without any restriction. In Section 2.2.3 it has been shown that $\sqrt{n} (\tilde{\alpha} - \alpha^0)$ converges to the multinormal distribution with
mean zero and variance $\tau^2 \Sigma \cdot$ therefore $\sqrt{n} (\tilde{\alpha}_H - \alpha^0) = \sqrt{n} ((I - A'A) \tilde{\alpha} - \alpha^0)$ converges to a multivariate normal distribution with mean zero and variance

$$(I - A'A)\tau^2 \Sigma - (I - A'A) = \frac{\tau^2 p+1}{p} (I - \frac{1}{p} 1' - A'A)$$  \hspace{1cm} (2.2.4.2)$$

Hence $\sqrt{n} (\tilde{\alpha}_H - \alpha^0)$ converges to a multivariate normal distribution with mean zero and variance

$$\tau^2 \frac{p+1}{p} (I - \frac{1}{p} 1' - A'A)$$  \hspace{1cm} (2.2.4.3)$$

2.3. Rank Tests in Repeated Measures Randomized Complete Block Designs

In this Section we would like to develop three different statistics based on the rank estimate of $\alpha$ for testing the hypothesis of the form $A\alpha = 0$ where $A1 = 0$ and rank of $A = q < p$.

2.3.1. Test Based on Drop in Dispersion Function

In the last Section we used $D(\alpha) = \sqrt{12} \sum_{j=1}^{n} \sum_{i=1}^{p} \left( \frac{R(Y_{ij} - \alpha_i)}{p+1} - \frac{1}{2} \right) (Y_{ij} - \alpha_i)$ to get an estimate of $\alpha$. In fact $D(\alpha)$ is used as a criterion for fitting a linear model to data. $D(\alpha)$ represents the minimum distance, as measured by $D(\alpha)$, from the data vector to the subspace spanned by the linear model. Let $\tilde{\alpha}_H$ be the R-estimate of $\alpha$ under $A\alpha = 0$.

To test the hypothesis $A\alpha = 0$ versus $A\alpha \neq 0$, we will compare $D(\tilde{\alpha})$ to $D(\tilde{\alpha}_H)$. This is the same strategy as that used to develop F tests based on reduction in sum of squares due to fitting reduced and full models. To make the test operational we at least need the limiting distribution under the null hypothesis. The test is not distribution free for finite sample size. The following Theorem shows that the test is asymptotically distribution free under the null hypothesis.
Theorem 5: Given the Repeated Measures Randomized Complete Block Design and \[ 2[D(\hat{\alpha}) - D(\hat{\alpha}_H)] \]

\[ A\alpha = 0, D^* = \text{Drop} = \frac{\tau}{\tau} \]

has an asymptotic chi-square distribution with \( q \) d.f.

Proof: The argument proceeds by approximating \( D(\alpha) \) with \( Q(\alpha) \) and by approximating \( \hat{\alpha} \) by \( \tilde{\alpha} \) and \( \hat{\alpha}_H \) by \( \tilde{\alpha}_H \). We then show that the asymptotic distribution of \( D(\hat{\alpha}_H) - D(\hat{\alpha}) \) is determined by that of \( Q(\tilde{\alpha}_H) - Q(\tilde{\alpha}) \). Finally the argument is completed by showing that \( Q(\tilde{\alpha}_H) - Q(\tilde{\alpha}) \) when properly normalized has an asymptotic chi-square distribution.

We begin by writing

\[ D(\hat{\alpha}_H) - D(\hat{\alpha}) = [D(\hat{\alpha}_H) - Q(\tilde{\alpha}_H)] + [Q(\tilde{\alpha}_H) - Q(\tilde{\alpha})] + [Q(\tilde{\alpha}) - Q(\hat{\alpha})] + [Q(\hat{\alpha}) - D(\hat{\alpha})] \]  

(2.3.1.1)

From 2.2.4.2 under true value \( \alpha^0 \), \( \sqrt{n} (\hat{\alpha}_H - \alpha^0) \) and from (2.2.3.4), \( \sqrt{n} (\hat{\alpha} - \alpha^0) \) are asymptotically normally distributed and hence bounded in probability. Hence by Theorem 4 Section 2.2.2, the first and fifth differences on the right hand side of the above tend to zero in probability; i.e. they are \( o_p(1) \). Since \( \sqrt{n} (\hat{\alpha} - \tilde{\alpha}) \) is \( o_p(1) \) and

\[ [Q(\tilde{\alpha}) - Q(\alpha)] \]

can be written as \( \sqrt{n} (\hat{\alpha} - \alpha^0) \left[ \frac{1}{\sqrt{n}} S(\alpha^0) - \frac{\tau}{2} \sqrt{n} \Sigma (\hat{\alpha} - \tilde{\alpha}) \right] \), therefore the fourth difference \( [Q(\tilde{\alpha}) - Q(\hat{\alpha})] \) is also \( o_p(1) \). Since \( \sqrt{n} (\hat{\alpha}_H - \tilde{\alpha}_H) \) is \( o_p(1) \), it can be shown that the second difference \( Q(\tilde{\alpha}_H) - Q(\tilde{\alpha}) \) is also \( o_p(1) \). Therefore we can write

\[ D(\hat{\alpha}_H) - D(\hat{\alpha}) = Q(\tilde{\alpha}_H) - Q(\tilde{\alpha}) + o_p(1) \]. If we plug \( \tilde{\alpha} = \alpha^0 + \frac{1}{n} \tau \Sigma - S(\alpha^0) \) into \( Q(\alpha) \) we get, after simplification,

\[ Q(\tilde{\alpha}) = D(\alpha^0) - \frac{1}{2n} \tau [S(\alpha^0)]' \Sigma - [S(\alpha^0)]. \]

Also if we plug \( \tilde{\alpha}_H = (I - A'A) \tilde{\alpha} \) into \( Q(\alpha) \), we get, after simplification,
\[ Q(\tilde{\alpha}_H) = D(\alpha^0) - \frac{1}{2n} \tau [S(\alpha^0)]' \left( \frac{p+1}{p} (I - \frac{1}{p} A'A) \right) [S(\alpha^0)] \text{ since } I - A'A \text{ is idempotent and } \left[I' A'A\right] = 0. \]

Therefore the difference \( Q(\tilde{\alpha}_H) - Q(\alpha) \) is equal to

\[ D(\alpha^0) - \frac{1}{2n} \tau [S(\alpha^0)]' \left[ \frac{p+1}{p} (I - \frac{1}{p} A'A) \right] [S(\alpha^0)] - D(\alpha^0) \]

\[ - \frac{1}{2n} \tau [S(\alpha^0)]' [S(\alpha^0)] = \frac{1}{2n} \frac{p+1}{p} \tau [S(\alpha^0)]' (A'A) [S(\alpha^0)]. \]

Hence

\[ 2[Q(\tilde{\alpha}_H) - Q(\alpha)]/\tau = \frac{p+1}{p} \frac{1}{n} [S(\alpha^0)]' (A'A) [S(\alpha^0)] \] (2.3.1.2)

But under true value \( \alpha^0, \sqrt{n} \left[ \frac{1}{n} S(\alpha^0) \right] \) converges to the multinormal distribution with mean vector \( \theta \) and variance \( \Sigma \). Since \( A'A \) is idempotent, therefore \( A'A \sqrt{n} \left[ \frac{1}{n} S(\alpha^0) \right] \) converges to the multinormal distribution with variance \( \frac{n}{p} A'A \). Hence

\[ \frac{p+1}{p} \frac{1}{n} [S(\alpha^0)]' (A'A) [S(\alpha^0)] \] converges to a chi-square random variable with \( q \) d.f. under the null hypothesis as \( n \) tends to infinity. Since

\[ \text{Drop} = 2[D(\tilde{\alpha}_H) - D(\alpha)]/\tau = \frac{p+1}{p} \frac{1}{n} [S(\alpha^0)]' (A'A) [S(\alpha^0)] + o_p(1). \] (2.3.1.3)

therefore \( \text{Drop} = \frac{2[D(\tilde{\alpha}_H) - D(\alpha)]}{\tau} \) converges to a chi-square random variable with \( q \) d.f. under true value as \( n \) tends to infinity. If we replace \( \tau \) by a consistent estimate \( \hat{\tau} \) in \( \text{Drop} = \frac{2[D(\tilde{\alpha}_H) - D(\alpha)]}{\hat{\tau}} \), it will again follow a chi-square random variable for large \( n \) if we invoke Slutsky's Theorem. In fact \( \text{Drop} \) is analogous to \( -2\log \lambda \) in maximum likelihood technique.
2.3.2. Test based on the Gradient Vector

In this Section we would like to develop a test statistic based on the gradient vector evaluated at the reduced model rank estimate. From Theorem 3 Section 2.2.2, we can write

\[ \frac{1}{\sqrt{n}} S(\alpha) = \frac{1}{\sqrt{n}} S(\alpha^0) - \sqrt{n} \frac{1}{\tau} \Sigma(\alpha - \alpha^0) + o_p(1) \]  

(2.3.2.1)

where \( o_p(1) \) tends to zero in probability uniformly for all values of \( \alpha \) such that \( \sqrt{n} \|\alpha - \alpha^0\| \leq c \) for any \( c > 0 \) and \( \alpha^0 \) is under the null hypothesis.

Theorem 6: Given the Repeated Measures Randomized Complete Block Design and supposing that the null hypothesis \( A\alpha = 0 \) holds, then

\[ S^* = \frac{p+1}{p} \frac{1}{n} [S(\hat{\alpha}_H)]' (A'A) [S(\hat{\alpha}_H)] \]

follows a chi-square distribution with \( q \) d.f. for large \( n \).

Proof: Since \( \sqrt{n} (\hat{\alpha}_H - \alpha^0) \) is bounded in probability, hence we can replace \( \alpha \) by \( \hat{\alpha}_H \). Therefore plugging \( \hat{\alpha}_H \) into \( \frac{1}{\sqrt{n}} S(\alpha) \) we get

\[ \frac{1}{\sqrt{n}} S(\hat{\alpha}_H) = \frac{1}{\sqrt{n}} S(\alpha^0) - \sqrt{n} \frac{1}{\tau} \Sigma(\hat{\alpha}_H - \alpha^0) + o_p(1). \]

Therefore \( A'A \frac{1}{n} [S(\hat{\alpha}_H)] = A'A \frac{1}{\sqrt{n}} S(\alpha^0) - A'A \sqrt{n} \frac{1}{\tau} \Sigma(\hat{\alpha}_H - \alpha^0) + o_p(1). \)

But the second term of the above is 0 by hypothesis. Therefore

\[ A'A \frac{1}{n} [S(\hat{\alpha}_H)] = A'A \frac{1}{\sqrt{n}} S(\alpha^0) + o_p(1). \]

But \( \frac{1}{\sqrt{n}} S(\alpha^0) \) converges to a multinormal distribution with mean 0 and variance \( \Sigma \), hence \( A'A \frac{1}{n} [S(\hat{\alpha}_H)] \) converges to a multinormal random variable with mean 0 and covariance \( \frac{p+1}{p} A'A \). If we make a quadratic form out of \( A'A \frac{1}{n} [S(\hat{\alpha}_H)] \) that will follow a chi-square distribution with \( q \) d.f. But that quadratic form
{A'A \frac{1}{n} [S(\hat{\alpha}_H)]} \cdot \left\{ \frac{p+1}{p} A'A \right\}^{-1} \{A'A \frac{1}{n} [S(\hat{\alpha}_H)] \}

= \frac{p+1}{p} \frac{1}{n} [S(\hat{\alpha}_H)]' (A'A) [S(\hat{\alpha}_H)] \tag{2.3.2.2}

Hence the result. This result is analogous to Rao's score statistic in maximum likelihood techniques. When \( \alpha = 0 \), this test statistic becomes Friedman's statistic. This gives a justification of Friedman's statistic from a nonparametric Repeated Measures Model point of view.

2.3.3. Test Based on Full Model Estimate

A third approach to testing \( A\alpha = 0 \) is based directly on the full model R-estimate \( \hat{\alpha} \) determined by minimizing \( D(\alpha) \) or by solving \( S(\alpha) = 0 \). We know that \( \sqrt{n}(\hat{\alpha} - \alpha^0) \) converges to a multinormal distribution with mean zero and variance \( \tau^2 \Sigma' \). Therefore \( \sqrt{n} A \hat{\alpha} \) converges to a multinormal distribution with mean \( 0 \) and variance \( \tau^2 A \Sigma A' \) under \( H_0 \). Hence if we make a quadratic form out of \( A\hat{\alpha} \) we get

\[ W^* = n(A\hat{\alpha})' [\tau^2 A \Sigma A']^{-1} (A\hat{\alpha}) = n (\hat{\alpha})' A' [\tau^2 A \Sigma A']^{-1} A \hat{\alpha} \]

\[ = \frac{p}{p+1} \frac{n\hat{\alpha}' A' A \hat{\alpha}}{\tau^2} \tag{2.3.3.1} \]

and this follows a chi-square distribution for large \( n \) with \( q \) d.f. If we replace \( \tau \) by a consistent estimate \( \hat{\tau} \) in \( n \frac{p}{p+1} (\hat{\alpha})' A' [\tau^2 A \Sigma A']^{-1} A \hat{\alpha} \) it will again follow a chi-square random variable for large \( n \) if we invoke Slutsky's Theorem. This test statistic is like Wald's statistic.
2.4. Multiple Comparisons Based on Full Model Estimate

If the test for $A \hat{a} = 0$ is significant, the next step is to decide which of the $q$ contrasts are responsible for rejection. In this Section we will develop some multiple comparison procedures based on full model estimates. From Section 2.2.3 we have

$$\sqrt{n}(\hat{\alpha} - \alpha^0) \sim \text{MVN}[0, \tau^2 \Sigma].$$

Hence $a_i' \hat{\alpha}$ is approximately normally distributed with mean zero and variance $\frac{\tau^2(p+1)}{np}$, where $a_i'1 = 0$. We can have the following asymptotic results:

2.4.1 Least Significant Difference

A $100(1 - \alpha)\%$ confidence interval for a contrast $a_i' \alpha$ is given by

$$[a_i' \hat{\alpha} - z_{\alpha/2} \sqrt{\frac{\tau^2(p+1)}{np}}, a_i' \hat{\alpha} + z_{\alpha/2} \sqrt{\frac{\tau^2(p+1)}{np}}],$$

where $z_{\alpha/2}$ is the $(1 - \alpha/2)$th quantile of the standard normal distribution.

2.4.2 Bonferroni Intervals

$$[a_i' \hat{\alpha} - z_{\alpha/2k} \sqrt{\frac{\tau^2(p+1)}{np}}, a_i' \hat{\alpha} + z_{\alpha/2k} \sqrt{\frac{\tau^2(p+1)}{np}}, \ldots, a_i' \hat{\alpha} + z_{\alpha/2k} \sqrt{\frac{\tau^2(p+1)}{np}}]$$

$i = 1, 2, \ldots, k$. The overall probability for all $k$ intervals is at least $1 - \alpha$.

2.4.3. Tukey’s Procedure

$$[\hat{\alpha}_i - q_{\alpha, p, \infty} \sqrt{\frac{\tau^2(p+1)}{np}}, \hat{\alpha}_i + q_{\alpha, p, \infty} \sqrt{\frac{\tau^2(p+1)}{np}}]$$

where $q_{\alpha, p, \infty}$ is the upper $(1 - \alpha)$th percentile for the range of $p$ independent $N(0,1)$ random variables.

2.4.4. Sheffe’s Procedure

Let $f$ be a $q$-dimensional vector of contrasts in $\alpha_i$’s. A $100(1 - \alpha)\%$ confidence interval for any linear function of $h'f$ is given by
Asymptotic Relative Efficiency

The three statistics developed in Section 2.2 for testing $A\alpha = 0$ have the same asymptotic distribution under the null hypothesis. In this section we will develop Pitman efficiency by considering the limiting distribution when a sequence of alternatives converges to the null hypothesis. It is well known that when statistics have asymptotic chi-square distributions, the efficiency is developed by the ratio of noncentrality parameters. For the linear model with i.i.d. errors, McKean and Hettmansperger (1976),; Sen and Puri (1977), and Adichie (1978) showed that for the three tests, the efficiency of any one of these tests relative to the least squares F test is

$$E(\text{Rank, Least-square}) = 12\tau^2 \left( \int (f(x))^2 dx \right)$$

Thus the three tests are asymptotically equivalent in the sense of Pitman efficiency and they inherit the efficiency of the Wilcoxon signed rank test, the Mann-Whitney-Wilcoxon test and Kruskall-Wallis test. In this section we will show that for the Repeated Measures Randomized Block Design the three tests, i.e., Drop in dispersion, Wald's statistic and Aligned test, the efficiency of any one of these tests relative to the corresponding least squares F statistic is

$$E(\text{Rank, Least Square}) = 12\sigma^2 (1-p) \left( \int f(x,x)^2 dx \right)$$
Sen (1971) also obtained the above expression for the ARE of Friedman statistic with respect to the least square $F$ under local alternatives.

### 2.5.1. Efficacy of Wald Statistic

From Section 2.2, we know that $\sqrt{n} (\hat{\alpha} - \alpha)$ is of $o_p(1)$. Also $\tilde{\alpha} = \alpha^0 + \frac{1}{n} \tau \Sigma^{-1} S(\alpha^0)$. Without loss of generality let us assume that $\alpha^0 = 0$. We can write $\tilde{\alpha}$ as $\frac{1}{n} \frac{p+1}{p} \tau S(\theta)$ since $1'S(\theta) = 0$. Let $\tilde{\theta} = \frac{\alpha}{\sqrt{n}}$. From Section 2.2.2, $E_{\theta}[S(\theta)] = \frac{p+1}{p} \frac{1}{\tau} [\theta - \frac{1}{n}]$. Therefore $E_{0}[\tilde{\theta}^2] = \frac{1}{p} \frac{p+1}{p} \tau E_{\theta}[S(\theta)] = [\theta - \frac{1}{n}]$. Hence the noncentrality parameter of the Wald statistic $\frac{\tau^2}{\Sigma_{i=1}^{p+1} (\alpha_i - \alpha^*)^2}$ becomes, after simplification,

$$\frac{p}{(p+1)\tau^2} (\alpha - 1'\alpha)' A' A (\alpha - 1'\alpha). \tag{2.5.1.1}$$

If we would like to test $\alpha_1 = \alpha_2 = \ldots = \alpha_p$, then under local alternatives the noncentrality parameter becomes

$$\frac{p}{(p+1)\tau^2} \sum_{i=1}^{p} (\alpha_i - \alpha^*)^2. \tag{2.5.1.2}$$

### 2.5.2. Efficacy of the Test Statistic Based on the Gradient Vector

From 2.3.2.1 we have $A' A \frac{1}{n} [S(\hat{\alpha}_{H})] = A' A \frac{1}{\sqrt{n}} S(\alpha^0) + o_p(1)$. Therefore $E_{0}(A' A \frac{1}{n} [S(\hat{\alpha}_{H})] = E_{0}(A' A \frac{1}{\sqrt{n}} S(\alpha^0)) = \frac{p}{(p+1)\tau} A' A (\alpha - 1'\alpha)$ by using the results of the previous Section.

Hence the noncentrality parameter of $\frac{p+1}{p} \frac{1}{n} [S(\hat{\alpha}_{H})]' (A' A) [S(\hat{\alpha}_{H})]$ under local alternatives is
2.5.3. Efficacy of Test based on Drop in Dispersion

From (2.3.1.2) Drop = 2[D(\alpha_H) - D(\alpha)]/\tau = \frac{p+1}{p} \frac{1}{n} [S(\alpha^0)'(A'A)[S(\alpha^0)] + o_p(1)].

Like the previous two Sections, it can be shown that, the efficacy of Drop in dispersion statistic under local alternatives is \( \frac{p}{(p+1)\tau^2} (\alpha - 11'\alpha)' A'A(\alpha - 11'\alpha) \) (2.5.2.1).

2.5.4. Efficacy of Least Square F

Using equation (16) P.K. Sen (1971), it can be shown that the efficacy of the least squares F under exchangeable normal errors and local alternatives is

\[ \frac{p}{(p+1)\tau^2} (\alpha - 11'\alpha)' A'A(\alpha - 11'\alpha) \] since \( q_F \) converges to a noncentral chi-square distribution with parameter

\[ \frac{(\alpha - 11'\alpha)' A'A(\alpha - 11'\alpha)}{\sigma^2(1 - \rho)} \] (2.5.3.1)

2.5.5. ARE of Rank Tests with respect to Least square Test

Sen (1971) pointed out that that under local alternatives and equicorrelated errors, \( \frac{(p-1)}{\rho}F \) is approximately distributed like non-central chi-square with \( p-1 \) degrees of freedom as \( n \) tends to infinity. All the rank statistics developed in this chapter converge to non-central chi-square random variable with \( (p-1) \) d.f. under local alternatives as \( n \) tend to infinity. Consequently using equation (5) (which is the ratio of noncentrality parameters) in Hannan (1956) we get the ARE of all the rank tests with respect to least squares test as \( E(\text{Rank}, \text{Least square}) \)
2.5.6 ARE's of Rank Tests versus Least-Square F for Different Distributions

2.5.6.1 Errors are Exchangeable Normal

If errors are exchangeable normal, then the bivariate p.d.f. is given by

\[ f(y_1, y_2) = \frac{1}{2\pi\sigma^2(1-p)} \exp \left( \frac{-1}{2\sigma^2(1-p^2)} \left( (y_1 - \mu_1)^2 + (y_2 - \mu_2)^2 - 2\rho(y_1 - \mu_1)(y_2 - \mu_2) \right) \right) \]

Therefore

\[ \left( \int f(x,x) \, dx \right)^2 = \left[ \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma^2(1-p)} \exp \left( \frac{-1}{\sigma^2(1-p^2)} (x - \mu)^2 \right) \, dx \right]^2 = \frac{1}{4\pi\sigma^2(1-p)}. \]

Hence

\[ \text{ARE(Rank, Least-square)} = 12\sigma^2(1-p) \frac{1}{p+1} \frac{1}{4\pi\sigma^2(1-p)} = \frac{3p}{\pi(p+1)}. \]

2.5.6.2 Errors are Exchangeable Multivariate Logistic

A multivariate logistic distribution has been derived by Gumbell (1961). Its bivariate p.d.f. in standard form is

\[ f(y_1, y_2) = \frac{2\exp(-y_1-y_2)}{[1+\exp(-y_1)+\exp(-y_2)]^3} \]

Therefore

\[ \int_{-\infty}^{\infty} f(x,x) \, dx = \frac{1}{4}, \quad \text{Var}(y_1) = \frac{\pi}{3} \quad \text{and} \quad \text{Corr}(y_1, y_2) = \frac{1}{2}. \]

Therefore

\[ \text{ARE(Rank, Least-square)} = 12\sigma^2(1-p) \frac{1}{p+1} \frac{1}{4} \frac{\pi\sigma^2}{8(p+1)} = \frac{p\pi^2}{8(p+1)}. \]
2.5.6.3. Exchangeable Cauchy Distribution

A multivariate Cauchy distribution was derived by Ferguson (1962). The bivariate p.d.f. has the form \( f(y_1, y_2) = \frac{G(3/2)}{\pi^{3/2}} \left[1+(y_1-\mu)^2+(y_2-\mu)^2\right]^{-3/2} \). If \( y \) has an exchangeable normal distribution and \( \chi^2 \) is an independent chi-square random variable with 1 d.f., then \( \frac{y}{\sqrt{\chi^2}} \) has an exchangeable Cauchy distribution. Since \( \text{var}(y) = \infty \), the ARE (Rank, Least-square) is always infinity.

2.5.6.4. Exchangeable t Distribution

If \( y \) has an exchangeable normal distribution and \( \chi^2 \) is an independent chi-square random variable with \( k \) d.f., then \( \sqrt{k} \frac{y}{\sqrt{\chi^2}} \) has an exchangeable t distribution with \( k \) d.f. Its bivariate p.d.f. is given by

\[
f(y_1, y_2) = \frac{1}{2\pi\sigma^2 (1-\rho)} \left\{1+\frac{(y_1-\mu)^2+(y_2-\mu)^2-2\rho(y_1-\mu)(y_2-\mu)}{k(1-\rho^2)}\right\}^{-(k+1)/2}
\]

Therefore \( \int_{-\infty}^{\infty} (\int f(x,x)dx)^2 dx = \frac{(\sqrt{k\pi} G((k+1)/2))^2}{2\pi\sigma^2 (1-\rho) (\sqrt{2} G(k+2)/2)^2} \).

For \( k = 3 \), \( \text{Var}(y) = 3\sigma^2 \). Therefore ARE(Rank,Least-square)

\[
eq \frac{24p}{(p+1)^2} \frac{12p\sigma^2 (1-\rho)}{\pi^2} \frac{(\sqrt{k\pi} G((k+1)/2))^2}{p+1} = 12p \frac{1}{p+1} \frac{\sqrt{k} G((k+1)/2)^2}{2 \sqrt{2} G(k+2)/2}
\]

For \( k = 3 \), \( \text{Var}(y)=3\sigma^2 \) and therefore ARE(Rank,Least-square) = \( \frac{24p}{(p+1)^2} \).
2.5.6.5. Errors are Exchangeable Exponential

The p.d.f. of bivariate exponential is given by

\[ f(y_1, y_2) = \exp(-y_1 - y_2) \left[ 1 + a \{ 2 \exp(-y_1) - 1 \} \{ 2 \exp(-y_2) - 1 \} \right]; \quad |a| \leq 1, \quad 0 \leq y_1, y_2 \leq \infty. \quad \text{Also} \]

\[ \text{corr}(y_1, y_2) = \frac{a}{4}, \quad \text{Var}(y_1) = 1. \quad \text{Also} \quad \left( \int f(x, x) \, dx \right)^2 = \frac{(3 + \alpha)^2}{36}. \]

Hence

\[ \text{ARE} = \frac{12(4 - \alpha)p(3 + \alpha)^2}{(p+1)144} \cdot \sigma^2 (1 - \rho) \left[ \int f(x, x) \, dx \right]^2. \]

**TABLE 3**

<table>
<thead>
<tr>
<th>Distribution</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>10</th>
<th>20</th>
<th>( \infty )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ex Nor</td>
<td>.64</td>
<td>.72</td>
<td>.76</td>
<td>.8</td>
<td>.87</td>
<td>.91</td>
<td>.955</td>
</tr>
<tr>
<td>Ex Log.</td>
<td>2</td>
<td>.92</td>
<td>.98</td>
<td>1.03</td>
<td>1.12</td>
<td>1.17</td>
<td>1.232</td>
</tr>
<tr>
<td>Ex Cau</td>
<td>( \infty )</td>
<td>( \infty )</td>
<td>( \infty )</td>
<td>( \infty )</td>
<td>( \infty )</td>
<td>( \infty )</td>
<td>( \infty )</td>
</tr>
<tr>
<td>Ex t 3 d.f.</td>
<td>1.62</td>
<td>1.82</td>
<td>1.95</td>
<td>2.03</td>
<td>2.2</td>
<td>2.32</td>
<td>2.434</td>
</tr>
<tr>
<td>Ex t 4 d.f.</td>
<td>1.12</td>
<td>1.26</td>
<td>1.34</td>
<td></td>
<td>1.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex t 5 d.f.</td>
<td>.96</td>
<td>1.08</td>
<td>1.15</td>
<td></td>
<td>1.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex exp (Corr.25)</td>
<td>2.66</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Ex exp (corr.12)</td>
<td>1.21</td>
<td>1.36</td>
<td>1.45</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>
CHAPTER III
ESTIMATION OF SCALE PARAMETER

3.1. Introduction

The parameter $\tau^2 = \frac{1}{12 \int f(x,x)dx^2}$ plays a prominent role in rank
based inference for repeated measures designs with exchangeable errors within blocks.
For example $\tau^2$, appears in the variance and covariance matrix of the rank estimates of
$\alpha$. The $\tau$ appears both in the quadratic approximation of the dispersion function and in
the denominator for Drop in dispersion test as a standardizing parameter. It appears in
the confidence interval for the contrasts and in the different multiple comparison
procedures and also in the ARE expression. Hence it is important to have a consistent
estimate of $\tau$. We will consider two different approaches to estimating $\tau$.

3.2. Interpretation of $\tau$

Let $z_1^{(i)} = y_{i1} - y_{11}$, $z_2^{(i)} = y_{i2} - y_{12}$, ..., $z_n^{(i)} = y_{in} - y_{1n}$; $i = 2, 3, ..., p$; where $i$
refers to treatment. Then $z_1^{(i)}, z_2^{(i)}, ..., z_n^{(i)}$ are i.i.d. random variables and
symmetrically distributed about $\alpha_2 - \alpha_1$. Since

$$P_{\alpha 1, \alpha 2} (y_{21} - y_{11} < a) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y_{21} - \alpha_2, y_{11} - \alpha_1) \ dy_{11}.$$ (3.2.1)
Therefore the derivative of \( P_{\alpha_1,\alpha_2} (y_{21} - y_{11} < a) \) with respect to

\[
\int_{-\infty}^{\infty} f(y_{21} - \alpha_2 + \alpha_1 y_{11}) \, dy_{11}
\]

is

\[
\int_{-\infty}^{\infty} f(y_{21}, y_{11}) \, dy_{11}.
\]

Hence \( \frac{\delta(P_{\alpha_1,\alpha_2} (y_{21} - y_{11} < a))}{\delta a} \) when \( \alpha_1 = \alpha_2 \) and \( a = 0 \) is

which is the density of \( y_{21}, y_{11} \) at 0. Hence \( \sqrt{12} \frac{1}{\tau} \) is the density of \( y_{21} - y_{11} \) at 0. Thus estimating \( \tau \) just involves estimating the p.d.f. of \( y_{21} - y_{11} \) at 0.

### 3.3. Density Estimate of \( \tau \)

Let us first see how we can estimate a density at a particular point using a random sample. After that we will develop a consistent estimate of \( \tau \) using \( z_1, z_2, ..., z_n \).

Let \( x_1 \leq x_2 \leq ... \leq x_n \) be an ordered sample of size \( n \) from an absolutely continuous c.d.f. \( F(x) \) with positive density \( f(x) \) having a continuous derivative in a neighborhood of the \( p \)th population quantile \( v_p = \{F^{-1}(p)\} \). In order to convert the median or any other 'quick estimate' into a test, we need to estimate its variance, or for large samples its asymptotic variance which depends on \( 1/f(v_p) \). Siddiqui (1960) proposed the estimator

\[
s_{mn} = n \frac{1}{2m} \left[ x[np+m] - x[np-m+1] \right] \text{ for } \frac{1}{f(v_p)}
\]

and showed that it is asymptotically normally distributed and suggested that \( m \) be chosen to be of order \( \sqrt{n} \). Bloch and Gastwirth (1968) showed the value of \( m \)
minimizing the asymptotic mean-square error is of order $n^{4/5}$ and if $m = o(n)$ and $m$ tends to infinity, then the statistic $s_{mn}$ is a consistent estimator of $\frac{1}{f(\nu_p)}$. For our case $z_1^{(i)}, z_2^{(i)}, ..., z_n^{(i)}$ are i.i.d. Without loss of generality let us assume that $z_1^{(i)}, z_2^{(i)}, ..., z_n^{(i)}$ are ordered observations ($i = 2, 3, ..., p$).

Then $s^{(i)}_{mn} = n \frac{1}{2m} [ z_{[np]} + m - z_{[np]-m+1} ]$ is a consistent estimate of $\frac{1}{f(\nu_p)}$. Therefore

$$\frac{1}{p-1} \sum_{i=2}^{p} s^{(i)}_{mn}$$

is also a consistent estimate of $\frac{1}{f(\nu_p)}$. One can choose $m = cn^{4/5}$. For $\nu_p = 0$ the choice of $c$ as recommended by Bloch and Gastwirth (1968) is as follows:

1) for normal $c = .5$

2) for cauchy $c = .4$

3) for logistic $c = .58$.

### 3.4 Kernel Estimate of $\tau$

In this section we develop another estimate of $\tau$ which is different from the one we developed in the last section. We consider the i.i.d. case where $x_1, x_2, ..., x_n$ is a random sample from $F$ with median $0$ and p.d.f. $f(x)$. The estimate of $f(x)$, called a kernel or window estimate, was proposed by Rosenblatt (1956) and studied further by Parzen (1962). Wegman (1972) and Bean and Tsokos (1980) provide extensive surveys of the inference in the areas of density estimation.
Definition: Suppose $F$ has median 0 with p.d.f. $f(x)$. Suppose $w(x)$ is a square integrable density, symmetric about 0 and $h_n$ a sequence of constants such that $h_n$ tends to 0 and $nh_n$ tends to infinity.

Then

$$f_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} w\left(\frac{x-x_i}{h_n}\right)$$

is called a window estimate of $f(x)$. The function $w(x)$ is called the window or kernel and $h_n$ is called the window width or band width. Usually attention is given to a uniform window: $w(x) = 1$ if $-\frac{1}{2} < x \leq \frac{1}{2}$

= 0 otherwise

Using the indicator function $I$, we have

$$f_n(y) = \frac{1}{nh_n} \left[ \sum_{j=1}^{n} I(x_i \leq y + \frac{h_n}{2}) - \sum_{j=1}^{n} I(x_i \leq y - \frac{h_n}{2}) \right]$$

$$= \frac{1}{h_n} \left[ F_n(y + \frac{h_n}{2}) - F_n(y - \frac{h_n}{2}) \right]$$

(3.4.2)

where $F_n(y)$ is the empirical c.d.f. It can be shown that $f_n(y)$ converges in probability to $f(y)$. Let $z_{ij}' = y_{ij} - y_{ij}$ (i$\neq$I'=1,2,..., p; j = 1,2,..., n). Since $\int f(x,x)dx$ is the p.d.f. of $y_{ij}-y_{ij}'$ at 0 under $\alpha_1 = \alpha_2 = ... = \alpha_p$ the density estimate of $\gamma = \int f(x,x)dx$ under the null hypothesis is

$$\hat{\gamma} = \frac{1}{p(p-1)h_n} \sum_{j=1}^{n} \sum_{i=1}^{p} \sum_{i'=1, i' \neq i}^{p} \left[ F_{ij}(y + \frac{h_n}{2}) - F_{ij}(y - \frac{h_n}{2}) \right]$$

(3.4.3)

Therefore

$$E_{0}[\hat{\gamma}] = \int f(x,x)dx = \gamma.$$
Also

\[ \text{var}_0 \left[ \hat{\gamma} \right] = \text{Var} \left[ \frac{1}{p(p-1)h_n} \sum_{j=1}^{n} \sum_{i=1}^{p} \sum_{i'=1}^{p} \left[ F_{i_j} - F_{j_i} - F_{i'j} + F_{j'i} \right] \right] \]

\[ = \frac{1}{n^2p^2(p-1)^2h_n^2} \sum_{j=1}^{n} \text{Var} \left[ \sum_{i=1}^{p} \sum_{i'=1}^{p} I_{ii'} \right] \]

\[ = \frac{1}{n^2p^2(p-1)^2h_n^2} n \text{Var} \left[ \sum_{i=1}^{p} \sum_{i'=1}^{p} I_{ii'} \right] \quad (3.4.4) \]

where \( I_{ii'} = 1 \) if \( \frac{h_n}{2} < z_{ii'} < \frac{h_n}{2} \)

\[ = 0 \] otherwise.

\[ \text{var}_0 \left[ \hat{\gamma} \right] = \frac{1}{n^2p^2(p-1)^2h_n^2} n[p(p-1)\text{Var}_0(I_{12})] \]

\[ + E_0 \sum_{i=1}^{p} \sum_{i=1}^{p} \sum_{k=1}^{p} \sum_{k'=1}^{p} \left( I_{ii'} I_{kk'} - E_0(I_{ii'}) E_0(I_{kk'}) \right) \]

\[ \leq \frac{1}{n^2p^2(p-1)^2h_n^2} n[p(p-1)E_0(I_{12})] E_0(1-I_{12}) \]

\[ + \sum_{i=1}^{p} \sum_{i=1}^{p} \sum_{k=1}^{p} \sum_{k'=1}^{p} \left( E_0(I_{ii'}) + E_0^2(I_{ii'}) \right) \]

\[ \leq \frac{1}{n^2p^2(p-1)^2h_n^2} n[p(p-1)\left( F_2 - F_{-2} \right) + p^4 \left( E_0 I_{12} + (E_0 I_{12})^2 \right)] \]

\[ \leq \frac{1}{np(p-1)h_n} \left( \frac{F_{h_n} - F_{-h_n}}{h_n} \right) + \frac{1}{nh_n} \left( \frac{F_{h_n} - F_{-h_n}}{h_n} \right) / h_n E^2(I_{12}) \quad (3.4.5) \]

which tends to zero as nh tends to infinity. Therefore \( \hat{\gamma} \) converges to \( \gamma \). Further work would be needed to show that \( \hat{\gamma} \) based on residuals would also be consistent.
4.1 Computational Consideration

In this section we discuss computational procedures for estimating the scale parameter and for the full model R-estimates of the treatment effects.

4.1.1 Computation of the Estimate of the Scale Parameter $\tau$

In this section we outline an algorithm which calculates the estimate of $\tau$ from the data.

**Step 1:** Subtract all the observations corresponding to the first treatment from the rest of the treatments. This gives $Z_1^{(i)} = y_{11} - y_{11}$, $Z_2^{(i)} = y_{12} - y_{12}$, $Z_n^{(i)} = y_{1n} - y_{1n}$, Without loss of generality let us assume that $Z_1^{(i)}$, $Z_2^{(i)}$, ..., $Z_n^{(i)}$ are the ordered differences.

**Step 2:** Sort the resulting data within treatments. This is done by the procedure qsort-row in the main program.

**Step 3:** For each $i$, compute $S_m^{(i)} = \frac{1}{m} \left[ Z^{(i)}_{[np]+m} - Z^{(i)}_{[np]-m+1} \right]$

**Step 4:** Finally, we compute $\frac{1}{p-1} \sum_{i=2}^{p} S_m^{(i)}$, which is a consistent estimate of $\tau$. All these steps are done by the procedure entitled cal-scale-param.

4.1.2 Computation of the R-Estimate

In this section we discuss the computational procedure for rank based inference for Repeated Measures RCBD. In fact, we have developed a computer program in Pascal.
language that calculates the full model R-estimates, Drop in dispersion statistic (D*), S statistic (S*), and Walds statistic (W*). The following are the outlines which show how the program is structured.

**Step 1:** Procedure Readin reads the data from a file. The observations are arranged by treatment. For example, if there are p treatments and n blocks and $y_{ij}$ is the observation corresponding to the ith treatment and the jth block, then the data will be in the form $(Y_1, Y_2, ..., Y_p)'$ where $Y_i = (y_{i1}, y_{i2}, ..., y_{in})$

**Step 2:** Procedure Calguessvalue calculates the least square estimates of the contrasts $\alpha_i - \alpha_1$. In fact this least square estimate serves as a guess value for finding the rank estimate.

**Step 3:** Procedure observminus guess calculates the residuals.

**Step 4:** Procedure qsort sorts the residuals within blocks, using a quick sort method.

**Step 5:** Procedure find-direction computes the direction of the current Newton step. This procedure also gives the value of the S statistic (S*).

**Step 6:** Procedure Dispersion (t) calculates the value of the dispersion function at guessvalue [treatment] + t direction [treatment].

**Step 7:** Procedure findroot finds the values of t that bracket the minima of D(t).

**Step 8:** Procedure goldensection finds the minima of D(t). The guess values corresponding to the minima of D(t) are the rank estimates of $\alpha_i$'s.

**Step 9:** Once we know the rank estimates and a consistent estimate of $\tau$, we can compute the Wald statistic (W*). The procedure calwald does this.

**Step 10:** The Drop in dispersion statistic is computed by $\frac{[D(\hat{\alpha}) - D(0)]}{\hat{\tau}}$

**Step 11:** The procedure Anova calculates the least square F statistic.
4.2 Data Analysis

In this section we analyse some data which have been selected from selected journals.

Example 1: Illusion Data. Larsen and Marx (1986)

Hypnotic age regression is a procedure whereby an adult subject is instructed under hypnosis to return to an earlier chronological age. Although this process can recapture behavioral patterns of an earlier age, it is not known whether it can reinstate perceptual patterns. A recent study addressed itself to that point by using the Ponzo illusion.

The ponzo illusion was shown to eight college students. These same students were then regressed under hypnosis to age nine and then to age five. At each age, their perceptions of the illusion were measured. If hypnosis can, in fact, recapture perceptual patterns, the illusion strengths at the three ages should be significantly different.

<p>| Table 4 |
| Illusion Strengths  |
| Subjects  |</p>
<table>
<thead>
<tr>
<th>Treat</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>AW</td>
<td>0.81</td>
<td>0.44</td>
<td>0.44</td>
<td>0.56</td>
<td>0.19</td>
<td>0.94</td>
<td>0.44</td>
<td>0.06</td>
</tr>
<tr>
<td>NINE</td>
<td>0.69</td>
<td>0.31</td>
<td>0.44</td>
<td>0.44</td>
<td>0.19</td>
<td>0.44</td>
<td>0.44</td>
<td>0.19</td>
</tr>
<tr>
<td>FIVE</td>
<td>0.56</td>
<td>0.44</td>
<td>0.44</td>
<td>0.44</td>
<td>0.31</td>
<td>0.69</td>
<td>0.44</td>
<td>0.19</td>
</tr>
</tbody>
</table>
### Table 5

ANOVA Table for Example 1

<table>
<thead>
<tr>
<th>Test-Stat</th>
<th>Value</th>
<th>D.F.</th>
<th>Tab-value(.05)</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>1.462</td>
<td>2&amp;14</td>
<td>3.74</td>
<td>Accept H₀</td>
</tr>
<tr>
<td>W*</td>
<td>0</td>
<td>2</td>
<td>5.99</td>
<td>Accept H₀</td>
</tr>
<tr>
<td>D*</td>
<td>.0003</td>
<td>2</td>
<td>5.99</td>
<td>Accept H₀</td>
</tr>
<tr>
<td>S*</td>
<td>1.312</td>
<td>2</td>
<td>5.99</td>
<td>Accept H₀</td>
</tr>
</tbody>
</table>

### Table 6

Estimates of Contrasts

<table>
<thead>
<tr>
<th>Contrasts</th>
<th>Method</th>
<th>$\alpha_2-\alpha_1$</th>
<th>$\alpha_3-\alpha_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rank</td>
<td>.00000166</td>
<td>.000000815</td>
</tr>
<tr>
<td></td>
<td>LS</td>
<td>-.092</td>
<td>-.046</td>
</tr>
</tbody>
</table>


Sixteen animals were randomly placed into one of two groups - an experimental group receiving ethionine in their diets and a pair-fed control group. The liver of each animal was split into two parts one of which was treated with radioactive iron and oxyzen and the other with radioactive iron and nitrozen. The data consist of the amounts of iron absorbed by the variously treated liver halves. If matched pairs of animals are regarded as subjects and the combinations ethionine-oxyzen (EO), ethionine-nitrozen (EN), control-oxyzen(CO) and control-nitrozen(CN) are regarded as
treatments, then whether either diet or gas has any effect may be discussed by the results of this thesis.

Table 7
Amount of Iron Absorbed By Liver Halves

<table>
<thead>
<tr>
<th>Treat</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>EO</td>
<td>38.43</td>
<td>36.09</td>
<td>34.49</td>
<td>37.44</td>
<td>35.53</td>
<td>32.35</td>
<td>31.54</td>
<td>33.37</td>
</tr>
<tr>
<td>EN</td>
<td>31.47</td>
<td>29.89</td>
<td>34.50</td>
<td>38.86</td>
<td>32.69</td>
<td>32.69</td>
<td>31.89</td>
<td>33.26</td>
</tr>
<tr>
<td>CO</td>
<td>36.09</td>
<td>34.01</td>
<td>36.54</td>
<td>39.87</td>
<td>33.38</td>
<td>36.07</td>
<td>35.88</td>
<td>34.17</td>
</tr>
<tr>
<td>CN</td>
<td>32.53</td>
<td>27.73</td>
<td>29.51</td>
<td>33.03</td>
<td>29.88</td>
<td>29.29</td>
<td>31.53</td>
<td>30.16</td>
</tr>
</tbody>
</table>

We would like to test $H_0: \alpha_1=\alpha_2=\alpha_3=\alpha_4$

Table 8
Anova for Example 2

<table>
<thead>
<tr>
<th>Test stat.</th>
<th>Value</th>
<th>D.F.</th>
<th>TabValue</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>F*</td>
<td>15.40</td>
<td>3 &amp; 27</td>
<td>2.57</td>
<td>Reject $H_0$</td>
</tr>
<tr>
<td>D*</td>
<td>18.16</td>
<td>3</td>
<td>7.81</td>
<td>Reject $H_0$</td>
</tr>
<tr>
<td>S*</td>
<td>15.89</td>
<td>3</td>
<td>7.81</td>
<td>Reject $H_0$</td>
</tr>
<tr>
<td>W*</td>
<td>14.327</td>
<td>3</td>
<td>7.81</td>
<td>Reject $H_0$</td>
</tr>
</tbody>
</table>
Table 9
Estimates of Independent Contrasts

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Method</th>
<th>$\alpha_2-\alpha_1$</th>
<th>$\alpha_3-\alpha_1$</th>
<th>$\alpha_4-\alpha_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LS</td>
<td>-1.74</td>
<td>0.84</td>
<td>-4.44</td>
</tr>
<tr>
<td></td>
<td>Rank</td>
<td>-1.19</td>
<td>1.05</td>
<td>-4.37</td>
</tr>
</tbody>
</table>

Tukey's Procedure

Rank: Treatment 4 is different from treatments 1, 2, and 3.

L.S.: Same Conclusion.


Sixteen animals were randomly placed into one of two groups - an experimental group receiving ethionine in their diets and a pair-fed control group (i.e., a control animal was given the same amount of diet as the experimental animal with which it was paired). The data for each animal consisted of the measurements of amounts of radioactive iron among various subcellular fractions from liver cells. The fractions used were nuclei (n), mitochondria (Mit), Mitcrisomes (Mic) and supernatant (S). One question of interest to experimenters was whether the ratio of the measurements for the experimental group to those for the control group was the same for all cell fractions. If a matched pair of animals are regarded as subjects and cell fractions are regarded as treatments, then this problem may be attacked by the results of this thesis.
Table 10
Amount Of Radioactive Iron In Liver Cells
Subjects

<table>
<thead>
<tr>
<th>Treat</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1.73</td>
<td>2.50</td>
<td>1.17</td>
<td>1.54</td>
<td>1.53</td>
<td>2.61</td>
<td>1.86</td>
<td>2.21</td>
</tr>
<tr>
<td>Mit</td>
<td>1.08</td>
<td>2.55</td>
<td>1.47</td>
<td>1.75</td>
<td>2.71</td>
<td>1.37</td>
<td>2.13</td>
<td>1.06</td>
</tr>
<tr>
<td>Mic</td>
<td>2.60</td>
<td>2.51</td>
<td>1.49</td>
<td>1.55</td>
<td>2.51</td>
<td>1.15</td>
<td>2.47</td>
<td>0.95</td>
</tr>
<tr>
<td>S</td>
<td>1.67</td>
<td>1.80</td>
<td>1.47</td>
<td>1.72</td>
<td>2.25</td>
<td>1.67</td>
<td>2.50</td>
<td>0.98</td>
</tr>
</tbody>
</table>

We would like to test $H_0: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$.

Table 11
Anova for Example 3

<table>
<thead>
<tr>
<th>Test stat.</th>
<th>Value</th>
<th>D.F.</th>
<th>TabValue</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>F*</td>
<td>.21486</td>
<td>3&amp;27</td>
<td>2.57</td>
<td>Reject $H_0$</td>
</tr>
<tr>
<td>D*</td>
<td>.18207</td>
<td>3</td>
<td>7.81</td>
<td>Reject $H_0$</td>
</tr>
<tr>
<td>S*</td>
<td>1.2375</td>
<td>3</td>
<td>7.81</td>
<td>Reject $H_0$</td>
</tr>
<tr>
<td>W*</td>
<td>.0172</td>
<td>3</td>
<td>7.81</td>
<td>Reject $H_0$</td>
</tr>
</tbody>
</table>

Table 12
Estimates of Independent Contrasts

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Method</th>
<th>$\alpha_2-\alpha_1$</th>
<th>$\alpha_3-\alpha_1$</th>
<th>$\alpha_4-\alpha_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>LS</td>
<td>-.1287</td>
<td>.0099</td>
<td>-.1362</td>
<td></td>
</tr>
<tr>
<td>Rank</td>
<td>.0592</td>
<td>.0166</td>
<td>.0029</td>
<td></td>
</tr>
</tbody>
</table>
Multiple Comparison Based on Tukey's Procedure

Rank: Treatment 4 is different from treatments 1, 2, 3.
Least Square: Same conclusion.

Example 4: Roulette, Amos. (1972) "An Assessment of Unit Dose Injectable Systems" - American Journal of Hospital Pharmacy.

A comparison was made of the efficiency of four different unit-dose injection systems. A group of pharmacists and nurses were the blocks. For each system, they were to remove the unit from its outer package, assemble it, and simulate an injection. In addition to the standard system using a disposable syringe and needle to draw medication from the vial, the other systems tested were Vari-Ject (CIBA Pharmaceutical), Unimatic (Squibb), and Tubex (Wyeth). Listed in the table are the average times (in seconds) needed to implement each of the systems.

Table 13

Average Time In Implementing Injection Systems

<table>
<thead>
<tr>
<th>Subject</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stand</td>
<td>35.6</td>
<td>31.3</td>
<td>36.2</td>
<td>31.1</td>
<td>39.4</td>
<td>34.7</td>
<td>34.1</td>
<td>36.5</td>
<td>32.2</td>
<td>40.7</td>
</tr>
<tr>
<td>Vari</td>
<td>17.3</td>
<td>16.4</td>
<td>18.1</td>
<td>17.8</td>
<td>18.8</td>
<td>17.0</td>
<td>14.5</td>
<td>17.9</td>
<td>14.6</td>
<td>16.4</td>
</tr>
<tr>
<td>Uni</td>
<td>24.4</td>
<td>22.4</td>
<td>22.8</td>
<td>21.0</td>
<td>23.3</td>
<td>21.8</td>
<td>23.0</td>
<td>24.1</td>
<td>23.5</td>
<td>31.3</td>
</tr>
<tr>
<td>Tub</td>
<td>25.0</td>
<td>26.0</td>
<td>25.3</td>
<td>24.0</td>
<td>24.2</td>
<td>26.2</td>
<td>24.0</td>
<td>20.9</td>
<td>23.5</td>
<td>36.9</td>
</tr>
</tbody>
</table>

We would like to test $H_0: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$. 
Table 14
Anova for Example 4

<table>
<thead>
<tr>
<th>Test stat.</th>
<th>Value</th>
<th>D.F.</th>
<th>TabValue</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>100.59</td>
<td>3 &amp; 27</td>
<td>2.57</td>
<td>Reject $H_0$</td>
</tr>
<tr>
<td>$D^*$</td>
<td>130.99</td>
<td>3</td>
<td>7.81</td>
<td>Reject $H_0$</td>
</tr>
<tr>
<td>$S^*$</td>
<td>28.49</td>
<td>3</td>
<td>7.81</td>
<td>Reject $H_0$</td>
</tr>
<tr>
<td>$W^*$</td>
<td>252.18</td>
<td>3</td>
<td>7.81</td>
<td>Reject $H_0$</td>
</tr>
</tbody>
</table>

Table 15
Estimates Of Independent Contrasts

<table>
<thead>
<tr>
<th>Contrasts</th>
<th>Method</th>
<th>$\alpha_2-\alpha_1$</th>
<th>$\alpha_3-\alpha_1$</th>
<th>$\alpha_4-\alpha_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LS</td>
<td>-18.29</td>
<td>-11.41</td>
<td>-9.57</td>
</tr>
<tr>
<td></td>
<td>Rank</td>
<td>-18.09</td>
<td>-11.31</td>
<td>-9.47</td>
</tr>
</tbody>
</table>

4.3 Monte Carlo Study

Empirical powers were obtained for the four statistics $F$, $D^*$, $S^*$, and $W^*$ when the data come from a multivariate normal distribution with common variance 2 and common correlation coefficient 0.5 or from a multivariate $t$ with three d.f., variance 6 and $\rho = .5$. We considered Repeated Measures RCBD's with $p = 5$, $n = 8$, $p = 5$, and $n = 10$. The level was set at 0.05. The empirical powers are computed on the basis of 10,000 simulations. In each simulation independent normal random variables are generated by
Box and Muller method. Then we make a transformation to get equicorrelated random variables with in a block. For getting equicorrelated t random variables, we divide equicorrelated normal random variables with in each block by $\sqrt{\frac{\chi^2}{3}}$, where $\chi^2$ is an independent chisquare random variable with 3 d.f.

Table 16

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>F</th>
<th>D*</th>
<th>S*</th>
<th>W*</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_1=\alpha_2=\alpha_3=\alpha_4=\alpha_5=0$</td>
<td>.0522</td>
<td>.0643</td>
<td>.0445</td>
<td>.0799</td>
</tr>
<tr>
<td>$\alpha_1=0,\alpha_2=\alpha_3=\alpha_4=\alpha_5=.5$</td>
<td>.1298</td>
<td>.3661</td>
<td>.3041</td>
<td>.1499</td>
</tr>
<tr>
<td>$\alpha_1=0,\alpha_2=\alpha_3=\alpha_4=\alpha_5=2$</td>
<td>.997</td>
<td>.9576</td>
<td>.9159</td>
<td>.9382</td>
</tr>
<tr>
<td>$\alpha_1=0,\alpha_2=1,\alpha_3=2,\alpha_4=3, \alpha_5=4$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>.9999</td>
</tr>
</tbody>
</table>

Table 17

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>F</th>
<th>D*</th>
<th>S*</th>
<th>W*</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_1=\alpha_2=\alpha_3=\alpha_4=\alpha_5=0$</td>
<td>.0522</td>
<td>.0586</td>
<td>.0453</td>
<td>.0634</td>
</tr>
<tr>
<td>$\alpha_1=0,\alpha_2=\alpha_3=\alpha_4=\alpha_5=.5$</td>
<td>.1506</td>
<td>.3684</td>
<td>.3261</td>
<td>.1555</td>
</tr>
<tr>
<td>$\alpha_1=0,\alpha_2=\alpha_3=\alpha_4=\alpha_5=2$</td>
<td>.9948</td>
<td>.9736</td>
<td>.9641</td>
<td>.9616</td>
</tr>
<tr>
<td>$\alpha_1=0,\alpha_2=1,\alpha_3=2,\alpha_4=3, \alpha_5=4$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 18
Empirical Powers for n=8, p=5, ρ=.5, σ^2=2, multivariate 't' with 3 d.f.

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>F</th>
<th>D*</th>
<th>S*</th>
<th>W*</th>
</tr>
</thead>
<tbody>
<tr>
<td>α1=α2=α3=α4=α5=0</td>
<td>.0356</td>
<td>.0540</td>
<td>.043</td>
<td>.0716</td>
</tr>
<tr>
<td>α1=0,α2=α3=α4=α5=.5</td>
<td>.0936</td>
<td>.1068</td>
<td>.0991</td>
<td>.1166</td>
</tr>
<tr>
<td>α1=0,α2=α3=α4=α5=2</td>
<td>.8994</td>
<td>.8042</td>
<td>.8864</td>
<td>.6571</td>
</tr>
<tr>
<td>α1=0,α2=1,α3=2,α4=3,</td>
<td>.9923</td>
<td>.9937</td>
<td>1.0</td>
<td>.9292</td>
</tr>
<tr>
<td>α5=4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 19
Empirical Powers for n=10, p=5, ρ=.5, σ^2=2, for multivariate 't' with 3 d.f.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>F</th>
<th>D*</th>
<th>S*</th>
<th>W*</th>
</tr>
</thead>
<tbody>
<tr>
<td>α1=α2=α3=α4=α5=0</td>
<td>.0428</td>
<td>.0477</td>
<td>.0430</td>
<td>.0590</td>
</tr>
<tr>
<td>α1=0,α2=α3=α4=α5=.5</td>
<td>.1209</td>
<td>.1257</td>
<td>.1207</td>
<td>.1189</td>
</tr>
<tr>
<td>α1=0,α2=α3=α4=α5=2</td>
<td>.9561</td>
<td>.9314</td>
<td>.9585</td>
<td>.8476</td>
</tr>
<tr>
<td>α1=0,α2=1,α3=2,α4=3,</td>
<td>.9955</td>
<td>1.0</td>
<td>1.0</td>
<td>.9905</td>
</tr>
<tr>
<td>α5=4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 20
Empirical ARE

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Normal n = 8</th>
<th>Normal n = 10</th>
<th>t with 3 d.f. n = 8</th>
<th>t with 3 d.f. n = 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_1=\alpha_2=\alpha_3=\alpha_4=\alpha_5=0$</td>
<td>0.829</td>
<td>0.823</td>
<td>1.832</td>
<td>1.85</td>
</tr>
<tr>
<td>$\alpha_1=0,\alpha_2=\alpha_3=\alpha_4=\alpha_5=.5$</td>
<td>0.814</td>
<td>0.831</td>
<td>1.75</td>
<td>1.82</td>
</tr>
</tbody>
</table>

Empirical ARE is defined in section 4.4

Table 21
Bias of rank estimates when samples come from normal with $\rho = .5, \sigma^2 = 2, p = 5$

<table>
<thead>
<tr>
<th>n</th>
<th>Hypotheses</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\alpha_3$</th>
<th>$\alpha_4$</th>
<th>$\alpha_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>$\alpha_1=\alpha_2=\alpha_3=\alpha_4=\alpha_5=0$</td>
<td>0</td>
<td>0.008</td>
<td>0.005</td>
<td>0.007</td>
<td>0.003</td>
</tr>
<tr>
<td>8</td>
<td>$\alpha_1=0,\alpha_2=\alpha_3=\alpha_4=\alpha_5=.5$</td>
<td>0</td>
<td>0.008</td>
<td>-0.009</td>
<td>0.004</td>
<td>0.002</td>
</tr>
<tr>
<td>8</td>
<td>$\alpha_1=0,\alpha_2=\alpha_3=\alpha_4=\alpha_5=2$</td>
<td>0</td>
<td>0.003</td>
<td>0.0001</td>
<td>0.0004</td>
<td>0.0004</td>
</tr>
<tr>
<td>8</td>
<td>$\alpha_1=0,\alpha_2=1,\alpha_3=2,\alpha_4=3,\alpha_5=4$</td>
<td>0</td>
<td>-0.00097</td>
<td>-0.0020</td>
<td>0.0038</td>
<td>0.004</td>
</tr>
<tr>
<td>10</td>
<td>$\alpha_1=\alpha_2=\alpha_3=\alpha_4=\alpha_5=0$</td>
<td>0</td>
<td>0.007</td>
<td>0.007</td>
<td>0.009</td>
<td>0.002</td>
</tr>
<tr>
<td>10</td>
<td>$\alpha_1=0,\alpha_2=\alpha_3=\alpha_4=\alpha_5=.5$</td>
<td>0</td>
<td>0.001</td>
<td>-0.0001</td>
<td>0.004</td>
<td>0.008</td>
</tr>
<tr>
<td>10</td>
<td>$\alpha_1=0,\alpha_2=\alpha_3=\alpha_4=\alpha_5=2$</td>
<td>0</td>
<td>0.006</td>
<td>0.002</td>
<td>-0.002</td>
<td>-0.007</td>
</tr>
<tr>
<td>10</td>
<td>$\alpha_1=0,\alpha_2=1,\alpha_3=2,\alpha_4=3,\alpha_5=4$</td>
<td>0</td>
<td>-0.0052</td>
<td>-0.0038</td>
<td>-0.0021</td>
<td>-0.0048</td>
</tr>
</tbody>
</table>

The entries in the above table are the bias of the rank estimates of $\alpha_i$'s when the samples come from normal distribution.
Table 22
Bias of rank estimates when samples come from 't' with 3 d.f., \( \rho = .5, \sigma^2 = 6, p = 5 \)

<table>
<thead>
<tr>
<th>n</th>
<th>Hypotheses</th>
<th>( \alpha_1 )</th>
<th>( \alpha_2 )</th>
<th>( \alpha_3 )</th>
<th>( \alpha_4 )</th>
<th>( \alpha_5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>( \alpha_1=\alpha_2=\alpha_3=\alpha_4=\alpha_5=0 )</td>
<td>0</td>
<td>-.0025</td>
<td>.0004</td>
<td>-.00035</td>
<td>.0071</td>
</tr>
<tr>
<td>8</td>
<td>( \alpha_1=0,\alpha_2=\alpha_3=\alpha_4=\alpha_5=.5 )</td>
<td>0</td>
<td>.057</td>
<td>.064</td>
<td>.057</td>
<td>.056</td>
</tr>
<tr>
<td>8</td>
<td>( \alpha_1=0,\alpha_2=\alpha_3=\alpha_4=\alpha_5=2 )</td>
<td>0</td>
<td>.267</td>
<td>.271</td>
<td>.271</td>
<td>.264</td>
</tr>
<tr>
<td>8</td>
<td>( \alpha_1=0,\alpha_2=1,\alpha_3=2,\alpha_4=3, \alpha_5=4 )</td>
<td>0</td>
<td>.42</td>
<td>.28</td>
<td>.423</td>
<td>.556</td>
</tr>
</tbody>
</table>

The entries in the above table are the bias of the rank estimates of \( \alpha_i \)'s when the samples come from \( t \) distribution.

Table 23
Estimated \( \tau \) and true \( \tau \)

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Normal</th>
<th>'t' with 3 d.f.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>estimated</td>
<td>true</td>
</tr>
<tr>
<td>( \alpha_1=\alpha_2=\alpha_3=\alpha_4=\alpha_5=0 )</td>
<td>1.052</td>
<td>1.075</td>
</tr>
<tr>
<td>( \alpha_1=0,\alpha_2=\alpha_3=\alpha_4=\alpha_5=.5 )</td>
<td>1.053</td>
<td>1.075</td>
</tr>
<tr>
<td>( \alpha_1=0,\alpha_2=\alpha_3=\alpha_4=\alpha_5=2 )</td>
<td>1.050</td>
<td>1.070</td>
</tr>
<tr>
<td>( \alpha_1=0,\alpha_2=1,\alpha_3=2,\alpha_4=3, \alpha_5=4 )</td>
<td>1.052</td>
<td>1.075</td>
</tr>
</tbody>
</table>
4.4 Summary

For the normal distribution, it is obvious that $F$ is doing better than any other test statistic. When the sample size is 10, $D^*$ is best among the rank tests.

For the $t$ distribution with 3 d.f., only $D^*$, among all tests, is taking the correct level. The power for $D^*$ is smaller because the estimator of $\tau$, which appears in the denominator, is getting larger. Some further research may be done for getting an unbiased estimate of $\tau$. We recommend $D^*$ for non-normal distributions provided an unbiased estimate is plugged into $D^*$.

Table 21 shows that when data come from a normal distribution the rank estimates appear to be asymptotically unbiased. Table 22 shows that when data come from a $t$ distribution with 3 d.f., the rank estimates are close to being asymptotically unbiased.

The empirical $\text{ARE} = \frac{n \| \hat{\tau} - \tau^0 \|^2}{10000 \times 2(p-1)\sigma^2(1-p)}$ is close to .82, whereas the theoretical $\text{ARE}$ is .8, when data come from normal distribution. The empirical $\text{ARE}$ is close to 1.82, whereas the theoretical $\text{ARE}$ is 2.03, when the data come from $t$ distribution with 3 d.f.

For a normal distribution the estimated $\tau$ and the true $\tau$ are pretty close. For the $t$ distribution, the estimated $\tau$ and true $\tau$ differ a bit. Further research needs to be done on the estimation of $\tau$. 
5.1 Repeated Measures BIBD

In this design there are n blocks (subjects or judges) and t treatments. There are
k ≤ t residuals ranked within blocks. Every treatment appears in r ≤ n blocks and every
treatment appears with every other treatment an equal number of times (λ). A model for
this design is given by

$$Y_{ij} = (\mu + \alpha_i + \beta_j + \epsilon_{ij})n_{ij}; i = 1, 2, ..., t; j = 1, 2, ..., n,$$

where $n_{ij} = 1$ if the ith treatment appears in the jth block and $n_{ij} = 0$ otherwise,
$\mu = [\text{overall mean}], \alpha_i = [\text{effect of the ith treatment}], \beta_j = [\text{effect of jth block}],$
$\sum_{i=1}^{t} \alpha_i = 0$ and $\epsilon_{ij}$ is the error term $\sum_{i=1}^{t} \alpha_i = 0$. In this design we assume that the error
vectors are i.i.d. and that the non-zero elements within each error vector are
exchangeable random variables.

5.2 Rank Estimates in Repeated Measures BIBD

In this section we develop rank estimates using a dispersion function similar to that
of Jaeckel's. The dispersion for the jth subject is

$$D_j(\alpha) = \sqrt{\frac{1}{2}} \sum_{i=1}^{t} n_{ij} \left[ \frac{R(Y_{ij} - \alpha)}{k+1} - \frac{1}{2} \right] [Y_{ij} - \alpha_i]$$

The combined dispersion function is

58
\[ D(\alpha) = \sqrt{12} \sum_{j=1}^{n} \sum_{i=1}^{t} n_{ij} \left[ \frac{R(Y_{ij} - \alpha)}{k+1} \right] \frac{1}{2} \] (5.2.2)

It can be shown that this dispersion function in \( \alpha \) is nonnegative, continuous, even, location-free, convex and scale invariant. Hence, a rank estimate of \( \alpha \) can be obtained by minimizing the dispersion function. The domain (\( \alpha \) space) of \( D(\alpha) \) is divided into a finite number of convex polygonal subsets, on each of which \( D(\alpha) \) is a linear function of \( \alpha \). The partial derivatives exist almost everywhere and the negatives of the partial derivatives are given by the vector

\[ S(\alpha) = (s_1(\alpha), s_2(\alpha), ..., s_p(\alpha)) \] (5.2.3)

where

\[ s_i(\alpha) = \sqrt{12} \sum_{j=1}^{n} n_{ij} \left[ \frac{R(Y_{ij} - \alpha)}{k+1} \right] \frac{1}{2}, \quad i = 1, 2, ..., t. \] (5.2.4)

Hence minimizing \( D(\alpha) \) is equivalent to solving \( S(\alpha) = 0 \). But this is just an extension of the Hodges-Lehmann estimate to the linear model since \( E_0(S(\alpha_0)) = 0 \).

5.2.1 Asymptotic Distribution of \( S(\alpha_0) \) Under the True Value \( \alpha_0 \)

Without loss of generality, let us assume that \( \alpha_0 = 0 \). Therefore

\[ E_0(s_i(0)) = \sqrt{12} \sum_{j=1}^{n} n_{ij} E_0\left( \frac{R(Y_{ij} - \alpha_0)}{k+1} \right) \frac{1}{2} \]

\[ = \sqrt{12} \sum_{j=1}^{n} n_{ij} \frac{k+1}{2(k+1)} \frac{1}{2} = 0, \quad i = 1, 2, ..., t. \] (5.2.1.1)

and
\[ \text{Var}_0(s_i(0)) = \frac{12}{(k+1)^2} \sum_{j=1}^{n} n_j^2 \text{Var}_0(R[Y_{ij}]) \]

\[ = \frac{12}{(k+1)^2} \sum_{j=1}^{n} n_j^2 \frac{k^2-1}{12} = \frac{r(k-1)}{k+1} \quad (5.2.1.2) \]

since \( \sum_{j=1}^{n} n_j^2 = r. \) Also the null covariance between \( s_i(0) \) and \( s_i(0) \) is

\[ \text{Cov}_0[(s_i(0), s_i(0))] \]

\[ = 12 \text{Cov}_0[ \frac{\sum_{j=1}^{n} R(Y_{ij})}{k+1}, \frac{\sum_{j=1}^{n} R(Y_{ij})}{k+1} ], i \neq 1 \]

\[ = \frac{12}{(k+1)^2} \sum_{j=1}^{n} n_j n_{ij} \text{Cov}_0[R(Y_{ij}), R(Y_{ij})], i \neq 1 \]

\[ = \frac{12}{(k+1)^2} \sum_{j=1}^{n} n_j n_{ij} (\frac{k+1}{12}) \]

\[ = (\frac{k+1}{12}) \frac{12}{(k+1)^2} \sum_{j=1}^{n} n_j n_{ij} = \frac{-\lambda}{k+1} \quad (5.2.1.3) \]

since \( \sum_{j=1}^{n} n_j n_{ij} = \lambda. \) Also, for a BIBD design \( \lambda(t-1) = r(k-1). \) Therefore,

\[ \text{Cov}_0[(s_i(0), s_i(0))] = \frac{-r(k-1)}{(k+1)(t-1)} \quad (5.2.1.4) \]

Hence

\[ \text{Var-Cov}_0[S(0)] = \frac{r(k-1)}{(k+1)(t-1)} [tI-11'] \]

\[ = \frac{r(k-1)}{(k+1)(t-1)} (t+1) [tI-11'] \quad \frac{r(k-1)}{(k+1)(t-1)} (t+1) \Sigma \quad (5.2.1.5) \]

where \( \Sigma = \frac{[tI-11']}{t+1}. \)
Now we note that $S(0)$ can be written as $\sum_{j=1}^{n} X_j$, where

$X_j = (n_{ij}(R(y_{ij}) - \frac{1}{2}), \ldots, n_{ij}(R(y_{ij}) - \frac{1}{2}))$ and the $X_j$ are independent random vectors.

Also $\text{Var}(\sum_{j=1}^{n} X_j) = \frac{r(k-1)}{(k+1)(t-1)} \Sigma$. Therefore, $\frac{1}{n} \text{Var}(\sum_{j=1}^{n} X_j) = \frac{r(k-1)}{(k+1)(t-1)} \Sigma$. As $n$ tends to infinity $\frac{1}{n} \text{Var}(\sum_{j=1}^{n} X_j)$ converges to $\frac{m(k-1)}{(k+1)(t-1)} \Sigma$, where $\lim_{n} \frac{r}{n} = m$. But

$$\frac{1}{n} \sum_{j=1}^{n} \int_{\|x\| > \sqrt{n}} \|x\|^2 \, dF_{j}$$

$$= \frac{1}{n} \frac{k(k+1)(2k+1)}{12} \sum_{j=1}^{n} \int_{\|x\| > \sqrt{n}} dF_{j}$$

$$= \frac{1}{n} \frac{k(k+1)(2k+1)}{12} \sum_{j=1}^{n} P_{j}[\|x\| > \sqrt{n}] ,$$

where $F_{j}$ is the c.d.f of $X_j$. Using Markov's inequality, we have $P_{j}[\|x\| > \sqrt{n}] < \frac{E\|x\|}{\varepsilon \sqrt{n}}$. Therefore, $\lim_{n} P_{j}[\|x\| > \sqrt{n}] = 0$. By Kronecker's Lemma (Lemma 4, Section 1.7) we have $\lim_{n} \sum_{j=1}^{n} \frac{P_{j}[\|x\| > \sqrt{n}]}{n} = 0$. By a multivariate central limit theorem (Lemma 7, Section 1.7) $\frac{1}{\sqrt{n}} S(0)$ converges to a multinormal distribution with mean vector $0$ and variance-covariance matrix

$$\frac{m(k-1)}{(k+1)(t-1)} \Sigma $$

Finally $\frac{1}{\sqrt{n}} S(0)$ converges to a multinormal distribution with mean vector $0$ and variance-covariance matrix

$$\frac{(k-1)}{(k+1)(t-1)} \Sigma $$
5.2.2 Linear Approximation to the Negative of the Gradient Vector

A linear approximation to the gradient vector is crucial to the development of the distribution theory of rank estimates and tests for our repeated measures model. First we would like to find the expected value of the gradient vector under the true value $\alpha^0$. Without loss of generality let us assume that $\alpha^0 = 0$. We have

$$E_0(s_i(\alpha)) = E_0[\sqrt{\frac{12}{n}} \sum_{j=1}^{n} \frac{R(Y_{ij} - \alpha_i)}{k+1} - \frac{1}{2}], \quad i = 1, \ldots, t \quad (5.2.2.1)$$

But

$$R(Y_{ij} - \alpha_i) = 1 + \sum_{i \neq i} n_{ij} I_{y_{ij} - \alpha_i \leq y_{ij} - \alpha_i}$$

where $I$ is an indicator function.

Therefore

$$E_0(R(Y_{ij} - \alpha_i)) = 1 + \sum_{i \neq i} n_{ij} P_0(y_{ij} - \alpha_i \leq y_{ij} - \alpha_i)$$

$$= 1 + \sum_{i \neq i} n_{ij} P_0(y_{ij} - \alpha_i \leq \alpha_i - \alpha_i)$$

$$= 1 + \sum_{i \neq i} n_{ij} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y_{ij}, y_{ij}) dy_{ij} dy_{ij}$$

$$= 1 + \sum_{i \neq i} n_{ij} g(\alpha_i, \alpha_i') \quad (5.2.2.2)$$

Expanding $g(\alpha_i, \alpha_i')$ as a function of $\alpha_i$ and $\alpha_i'$ around $(0,0)$, we get
\[
g(\alpha_i, \alpha_{i'}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y_{i'j}, y_{ij}) dy_{i'j} dy_{ij} + (\alpha_{i'} - \alpha_i) \int_{-\infty}^{\infty} f(y_{ij}, y_{ij}) dy_{ij}
\]

\[
= \frac{1}{2} + (\alpha_{i'} - \alpha_i) \int_{-\infty}^{\infty} f(y_{ij}, y_{ij}) dy_{ij} = \frac{1}{2} + (\alpha_{i'} - \alpha_i) \int_{-\infty}^{\infty} f(y, y) dy
\]

Therefore

\[
E_\theta(s_i(\alpha)) = E_\theta[\sqrt{12} \sum_{j=1}^{n} \frac{R(Y_{ij} - \alpha_i)}{k+1}]
\]

\[
= \left\{ \frac{1}{\sqrt{12} \sum_{j=1}^{n} \frac{\sum_{i \neq j} g(\alpha_i, \alpha_{i'})}{k+1}} \right\} \frac{1}{2}
\]

\[
= \left\{ \sqrt{12} \sum_{j=1}^{n} \frac{n_{ij}}{(k+1)} \right\} + \frac{\sqrt{12}}{2(k+1)} \sum_{j=1}^{n} \left\{ \sum_{i \neq j} n_{ij} n_{i'j} \frac{1}{2} + (\alpha_{i'} - \alpha_i) \int_{-\infty}^{\infty} f(y, y) dy \right\} - \sqrt{12} \frac{n}{2}
\]

\[
= \left\{ \sqrt{12} \sum_{j=1}^{n} \frac{r}{k+1} + \frac{\lambda(t-1)\sqrt{12}}{2(k+1)} + \frac{\sqrt{12}}{k+1} \sum_{j=1}^{n} \left\{ \sum_{i \neq j} n_{ij} n_{ij} (\alpha_{i'} - \alpha_i) \int_{-\infty}^{\infty} f(y, y) dy \right\} - \sqrt{12} \frac{r}{2}
\]

\[
= \lambda(t-1) - \frac{1}{k+1} \lambda(-1, -1, ..., 1, -1, ..., -1' \ (\alpha_1, \alpha_2, ..., \alpha_i, ..., \alpha_t)
\]

Therefore

\[
E_{\theta} S(\alpha^0) = - \frac{1}{\tau} \frac{(k-1)(t+1)}{(k+1)(t-1)} \sum_{i} (\alpha_i - \alpha_0), \quad (5.2.2.4)
\]

where \( \tau = \frac{1}{\sqrt{12} \int_{-\infty}^{\infty} f(y, y) dy} \)

Now we state and prove some useful theorems.
Theorem 1. \( \frac{1}{r} S(\alpha) \) converges almost surely to \(-\frac{1}{\tau} \frac{(k-1)(t+1)}{(k+1)(t-1)} \Sigma (\alpha - \alpha^0) \), as \( r \) tends to infinity such that \( \frac{r}{n} \) tends to \( m \), where \( 0 < m < 1 \).

Proof: We can write \( s_i(\alpha) = \frac{1}{n} \sum_{j=1}^{n} t_j \) where \( t_j = - \{ n_{ij} \frac{R(y_{ij}-\alpha)}{k+1} - \frac{1}{2} \} \) Hence, the \( t_j \) are independently distributed and \( \text{Var}_0(t_j) \leq E_0(t_j^2) \leq 12k^2 \). Therefore, \( \frac{1}{n} \sum_{j=1}^{n} (\text{Var}_0(t_j)) \leq 12k^2 \sum_{j=1}^{1} \frac{1}{2} < \infty \). Hence, by Kolmogrov's SLLN, \( \frac{1}{n} S_i(\alpha) \) converges almost surely to \(-\frac{1}{\tau} \frac{(k-1)(t+1)}{(k+1)(t-1)} \Sigma (\alpha - \alpha^0) \). Therefore, \( \frac{1}{r} S(\alpha) \) converges almost surely to \(-\frac{1}{\tau} \frac{(k-1)(t+1)}{(k+1)(t-1)} \Sigma (\alpha - \alpha^0) \), since \( \frac{r}{n} \) tends to \( m \), \( 0 < m < 1 \), as \( n \) tends to infinity.

Theorem 2: Suppose \( f(y,y)dy<\infty \) and that the true value of \( \Delta \) is \( \Delta^0 = 0 \). For \( \varepsilon > 0 \) and \( \alpha > 0 \), \( \lim P_0[\|\frac{1}{r} S(\sqrt{\frac{\Delta}{\sqrt{r}}} - S(0) + \frac{1}{\tau} \frac{(k-1)(t+1)}{(k+1)(t-1)} \Sigma (\alpha - \alpha^0)\| > \varepsilon] = 0 \) as \( n \) tends to infinity.

Proof: Using (5.2.2.4), we can write

\[
E_0 \sqrt{n} \left[ \frac{1}{n} S(\sqrt{\frac{\Delta}{\sqrt{r}}} - S(0)) - \frac{1}{n} S(0) \right] = -\frac{1}{\tau} \frac{m(k-1)(t+1)}{(k+1)(t-1)} \Sigma \Delta^0
\]

and

\[
\text{Var}_0(\sqrt{n}[\frac{1}{n} S_i(\sqrt{\frac{\Delta}{\sqrt{r}}}) - s_i(0)]) = \frac{1}{n} \text{Var}_0[\sum_{j=1}^{n} n_{ij} (R(y_{ij} - \frac{\Delta_i}{\sqrt{n}}) - R(y_{ij}))]
\]

\[
\leq \frac{1}{n} \sum_{j=1}^{n} n^2_{ij} E_0[ (R(y_{ij} - \frac{\Delta_i}{\sqrt{n}}) - R(y_{ij}))^2] = \frac{1}{n} \sum_{j=1}^{n} a_j.
\]
Therefore
\[ a_n = E_0 n^{2ij} \{ R(y_{in} - \frac{\Delta i}{\sqrt{n}}) - R(y_{in}) \}^2. \]

But
\[ n^{2ij} \{ R(y_{in} - \frac{\Delta i}{\sqrt{n}}) - R(y_{in}) \}^2 \leq (k - 1)^2 \]

and
\[ \lim_n n^{2ij} \{ R(y_{in} - \frac{\Delta i}{\sqrt{n}}) - R(y_{in}) \}^2 \]
\[ \leq \lim_n \left[ \sum_{k=1}^{1} \{ I(y_{kn} - \frac{D_k}{\sqrt{n}} \leq y_{in} - \frac{\Delta i}{\sqrt{n}}) - I(y_{kn} \leq y_{in}) n^{2ij} \} \right] = 0, \]

where I is an indicator function. Hence by Lebesgue's Convergence Theorem (Lemma 5, Section 1.6) \( \lim_n a_n = \lim_n E_0 \{ R(y_{in} - \frac{\Delta i}{\sqrt{n}}) - R(y_{in}) \}^2 = 0. \) Since \( a_n \) converges to 0 as \( n \) tends to infinity then by the Lemma 4, Section 1.6, \( \lim_n \frac{1}{n} \sum_{j=1}^{n} a_j = 0. \) Hence,
\[ \lim_n \text{Var}_0 \left[ \frac{1}{n} \sum_{i} \frac{\Delta i}{\sqrt{n}} - \frac{1}{n} s_i(0) \right] = 0. \]
By appealing to Markov's inequality, we have
\[ \lim_n P_0 \left[ \left\| \frac{1}{n} \sum_{i} \frac{\Delta i}{\sqrt{n}} - \frac{1}{r} s(0) + \frac{1}{r} \frac{m(k-1)(t+1)}{\tau (k+1)(t-1)} \Sigma \Delta \| \geq \varepsilon \right\| = 0. \]
By using Slutsky's theorem we have
\[ \lim_n P_0 \left[ \left\| \frac{1}{r} \sum_{i} \frac{\Delta i}{\sqrt{r}} - \frac{1}{r} s(0) + \frac{1}{r} \frac{m(k-1)(t+1)}{\tau (k+1)(t-1)} \Sigma \Delta \| \geq \varepsilon \right\| = 0. \]

**Theorem 3:**
\[ \lim_n P_0 \{ \text{Sup} \| \Delta \| \leq c \sqrt{n} \left[ \left\| \frac{1}{n} \sum_{i} \frac{\Delta i}{\sqrt{n}} - \frac{1}{n} s(0) + \frac{1}{r} \frac{m(k-1)(t+1)}{\tau (k+1)(t-1)} \Sigma \Delta \| \geq \varepsilon \right\| = 0, \]

where \( c > 0. \) In other words
Theorem 4: \( \lim P_0(\sup ||A|| \leq c \sqrt{r} \left\{ \frac{1}{r} S_n(A) - \frac{1}{\tau} S(0) + \frac{1}{\tau} \sum_{(k-1)(t+1)}^{(k+1)(t-1)} \Sigma A || \geq \varepsilon \} = 0 \). 

\[ D_n(A) = \sqrt{\frac{12}{\tau}} \sum_{j=1}^{n} \sum_{i=1}^{1} n_{ij} \left( \frac{R(Y_{ij} \frac{A_i}{\sqrt{n}})}{p+1} - \frac{1}{2} \right) \left( Y_{ij} - \frac{A_i}{\sqrt{n}} \right) \]

and

\[ Q_n(A) = D_n(0) + \frac{1}{\tau} \frac{1}{m(k-1)(t+1)} \frac{1}{(k+1)(t-1)} \Sigma A - \frac{1}{\sqrt{n}} \Sigma A S(0) \] \hspace{1cm} (5.2.2.5)

Cor: Theorem 2, 3 and 4 are equivalent.

Proof: Both \( D_n(A) \) and \( Q_n(A) \) are convex functions of \( A \) with \( D_n(0) = Q_n(0) \)

and

\[
\delta(D_n(A) - Q_n(A)) \over \delta A \n
= - \frac{1}{\sqrt{n}} S_n(A) + \frac{1}{\sqrt{n}} S(0) - \frac{1}{\tau} \frac{1}{m(k-1)(t+1)} \Sigma A \] \hspace{1cm} (5.2.2.6)

Since for any \( \Delta \in \mathbb{R}^p \), the right hand side of (5.2.2.6) converges to zero in probability (by Theorem 2), we can select from each infinite index set \( n^* \in n \) another infinite index set \( n^{**} \in n^* \) such that \( \frac{\delta(D_n(A) - Q_n(A))}{\delta A} \) converges to \( 0 \) almost surely. By using standard diagonal sequence arguments it is therefore possible to find an index subset \( \bar{n} \)
\[
\delta(D_n(\Delta) - Q_n(\Delta))
\]
such that \(\|\{\frac{\delta(D_n(\Delta) - Q_n(\Delta))}{\delta \Delta}\}\|_{nep} \) converges to zero for all rational \(\Delta \in \mathbb{R}^p\). Let us introduce the convex function \(H_n(\Delta)\)

\[
= D_n(\Delta) - D_n(0) + \frac{1}{\sqrt{n}} \Delta' \Sigma S(0)
\]  
(5.2.2.7)

Then

\[
D_n(\Delta) - Q_n(\Delta) = D_n(\Delta) - D_n(0) - \frac{1}{2} \tau m \frac{(k-1)(t+1)}{(k+1)(t-1)} \Delta' \Sigma \Delta + \frac{1}{\sqrt{n}} \Delta' S(0)
\]

\[
= H_n(\Delta) - \frac{1}{2} \tau m \frac{(k-1)(t+1)}{(k+1)(t-1)} \Sigma \Delta
\]  
(5.2.2.8)

and

\[
\left\{ \frac{\delta(D_n(\Delta) - Q_n(\Delta))}{\delta \Delta} \right\} = \frac{\delta H_n(\Delta)}{\delta \Delta} - \frac{1}{\tau m} \frac{(k-1)(t+1)}{(k+1)(t-1)} \Sigma \Delta
\]  
(5.2.2.9)

Since \(\|\{\frac{\delta(D_n(\Delta) - Q_n(\Delta))}{\delta \Delta}\}\|_{nep} \) converges to zero for all rational \(\Delta \in \mathbb{R}^p\), the first condition of Lemma 3 (Section 1.6) is satisfied. Since \(D_n(0) - Q_n(0) = 0\), the second condition is also satisfied. Hence Theorem 2 implies Theorem 4. Theorem 4 implies Theorem 3 can be proved in the same way by using Lemma 2. Theorem 3 implies Theorem 4 automatically. The above results are also valid if we replace \(n\) by \(r\).

5.2.3 Asymptotic Distribution of the Rank Estimates Under the Full Model

Let \(\tilde{\alpha}\) minimize \(Q_n(\alpha)\). Then

\[
\tilde{\alpha} = \frac{(k+1)(t-1)}{r(k-1)(t+1)} \Sigma S(0)
\]  
(5.2.3.1)
If we assume that $\alpha^0$ is the true value of $\alpha$, then

$$\tilde{\alpha} = \alpha^0 + \tau \frac{(k+1)(t-1)}{r(k-1)(t+1)} \Sigma^{-1} S(\alpha^0)$$

(5.2.3.2)

From Theorem 4, $\tilde{\alpha}$ and $\hat{\alpha}$ must coincide asymptotically. Since

$$\frac{S(\alpha^0)}{\sqrt{r}} \sim N(0, \frac{(t+1)(k-1)}{(k+1)(t-1)} \Sigma)$$

as $r$ tends to infinity, then $\sqrt{r} (\tilde{\alpha} - \alpha^0)$ converges in distribution to a multinormal distribution with mean vector zero and variance-covariance matrix $\tau^2 \frac{(k+1)(t-1)}{(k-1)(t+1)} \Sigma^{-1}$. Hence $\sqrt{r} (\hat{\alpha} - \alpha^0)$ converges in distribution to a multinormal distribution with mean vector zero and variance-covariance matrix.

$$\tau^2 \frac{(k+1)(t-1)}{(k-1)(t+1)} \Sigma^{-1}$$

(5.2.3.3)

5.2.4 Asymptotic Distribution of the Rank Estimates Under the Reduced Model

We would like to minimize $D(\alpha) = \sum_{j=1}^{n} D_j(\alpha) = \sqrt{n} \sum_{j=1}^{n} \sum_{i=1}^{t} \frac{R(\gamma_{ij} - \alpha_i)}{p+1} - \frac{1}{2} \{ Y_{ij} - \alpha_i \}$ subject to the restriction $A\alpha = 0$ where $A1 = 0$, and the rows of $A$ are mutually orthogonal. The vector $\hat{\alpha}_{H}$, which minimizes $D(\alpha)$ subject to the restriction $A\alpha = 0$, is called the reduced model estimate of $\alpha$. To minimize $D(\alpha)$ subject to $A\alpha = 0$, we will use the method of Lagrangian multiplier. Let us define $\phi = D(\lambda, \alpha)) + \lambda' A \alpha$. Therefore $\frac{\delta \phi}{\delta \alpha} = S(\alpha) + A' \lambda = 0$ and $\frac{\delta \phi}{\delta \lambda} = A \alpha = 0$. Now we would like to get an approximation to $\frac{\delta \phi}{\delta \alpha}$ using Theorem 3 in Section 5.2.2. It can be shown that

$$\frac{1}{\sqrt{n}} S(\alpha) + A' \lambda = \frac{1}{\sqrt{n}} S(\alpha^0) - \frac{1}{\tau} \sum_{j=1}^{n} \frac{m_0(k-1)(t+1)}{(k+1)(t-1)} \Sigma (\alpha - \alpha^0) + A' \lambda + o_p(1)$$

(5.2.4.1)
where $o_p(1)$ tends to zero in probability uniformly for all values of $\alpha$ such that 
\[ \sqrt{n} \|\alpha - \alpha^0\| \leq c \] for any $c > 0$ and where $\alpha^0$ is the true value under the hypothesis. By Theorem 4 in Section 2.2.2, we can construct a quadratic approximation to $\phi$ such as 
$\psi = Q(\alpha) + \lambda'A\alpha = Q(\alpha) + \lambda'A\alpha$. Hence the minima of $\phi$ and $\psi$ will coincide asymptotically. Let $\alpha_H$ be the minima of $Q(\alpha)$ subject to the restriction $A\alpha = 0$. Then $\alpha_H$ is a solution of the equations 
\[ \frac{\delta \psi}{\delta \alpha} = 0 \quad \text{and} \quad \frac{\delta \psi}{\delta \lambda} = 0. \]
By solving those two equations, we get $\alpha_H = (I - A'A) \bar{\alpha}$ where $\bar{\alpha}$ is the minima of $Q(\alpha)$ without any restriction. In Section 5.2.3 it was shown that $\sqrt{r} (\bar{\alpha} - \alpha^0)$ converges in distribution to a multinormal distribution with mean vector zero and variance-covariance matrix 
$\tau^2 \frac{(k+1)(t-1)}{(k-1)(t+1)} \Sigma$. Therefore, $\sqrt{r} (\bar{\alpha}_H - \alpha^0) = \sqrt{r} (I - A'A)(\bar{\alpha} - \alpha^0)$ converges in distribution to a multivariate normal distribution with mean vector zero and variance-covariance matrix 
\[ (I - A'A) \tau^2 \frac{(k+1)(t-1)}{(k-1)(t+1)} \Sigma^* (I - A'A)' = \tau^2 \frac{1}{t} \frac{(k+1)(t-1)}{(k-1)} (I - \frac{11'}{t} A'A). \]
Hence $\sqrt{r} (\bar{\alpha}_H - \alpha^0)$ converges in distribution to a multivariate normal distribution with mean vector zero and variance-covariance matrix 
$\tau^2 \frac{(k+1)(t-1)}{t(k-1)} (I - \frac{11'}{t} A'A)$ (5.2.4.2)

5.3 Rank tests in Repeated Measures Randomized Complete Block Designs

In this section we develop three different statistics based on rank estimates of $\alpha$ for testing hypotheses of the form $A\alpha = 0$, where $A1 = 0$, the rank of $A = q < t$ and the rows of $A$ are mutually orthogonal.
5.3.1 Test Based on Drop in the Dispersion Function

In the last section we used $D(\alpha) = \sqrt{12} \sum_{j=1}^{t} \sum_{i=1}^{n} \left( \frac{R(Y_{ij} - \alpha_i)}{k+1} - \frac{1}{2} \right) (Y_{ij} - \alpha_i)$ to get an estimate of $\alpha$. In fact $D(\alpha)$ is used as a criterion for fitting a linear model to the data. $D(\alpha)$ represents the minimum distance, as measured by $D(\alpha)$, from the data vector to the subspace spanned by the linear model. Let $\hat{\alpha}_H$ be the R-estimate of $\alpha$ under $A\alpha = 0$.

To test the hypothesis $A\alpha = 0$ versus $A\alpha \neq 0$, we compare $D(\alpha)$ to $D(\hat{\alpha}_H)$. This is the same strategy as that used to develop $F$ tests based on the reduction in the sum of squares due to fitting reduced and full models. To make the test operational we at least need the limiting distribution under the null hypothesis, since the test is not distribution-free for finite sample size. The following theorem shows that the test is asymptotically distribution-free.

Theorem 5:

Given the Repeated Measures Randomized Complete Block Designs and $A\alpha = 0$, $D^* = \text{Drop} = 2[D(\hat{\alpha}) - D(\hat{\alpha}_H)]/\tau$ has an asymptotic chi-square distribution with $q$ d.f.

Proof: The argument proceeds by approximating $D(\alpha)$ with $Q(\alpha)$ and by approximating $\hat{\alpha}$ by $\bar{\alpha}$ and $\hat{\alpha}_H$ by $\bar{\alpha}_H$. We then show that asymptotic distribution of $D(\hat{\alpha}_H) - D(\hat{\alpha})$ is determined by that of $Q(\hat{\alpha}_H) - Q(\hat{\alpha})$. Finally the argument is completed by showing that $Q(\hat{\alpha}_H) - Q(\hat{\alpha})$, when properly normalized has an asymptotic chisquare distribution. We begin by writing

$$D(\hat{\alpha}) - D(\hat{\alpha}_H) = [D(\hat{\alpha}_H) - (\bar{\alpha}_H)] + [Q(\hat{\alpha}_H) - Q(\bar{\alpha}_H)]$$
We have already shown that under the true value $\alpha^0$, $\sqrt{T} (\hat{\alpha}_H - \alpha^0)$ and $\sqrt{T} (\hat{\alpha} - \alpha^0)$ are asymptotically normally distributed and hence bounded in probability. Hence by Theorem 4, Section 2.2.2, the first and fifth difference on the right hand side of (5.3.1.1) converge to zero in probability (i.e. they are $o_p(1)$). Since $\sqrt{T} (\hat{\alpha} - \hat{\alpha})$ is $o_p(1)$ and $[Q(\hat{\alpha}) - Q(\hat{\alpha})]$ can be written as $\sqrt{T} (\hat{\alpha} - \hat{\alpha})$, then the fourth difference $[Q(\hat{\alpha}) - Q(\hat{\alpha})]$ is also $o_p(1)$. Since $\sqrt{T} (\hat{\alpha}_H - \hat{\alpha}_H)$ is $o_p(1)$, it can be shown that the second difference $Q(\hat{\alpha}_H) - Q(\hat{\alpha}_H)$ is also $o_p(1)$.

Therefore we can write $D(\hat{\alpha}_H) - D(\hat{\alpha}) = [Q(\hat{\alpha}_H) - Q(\hat{\alpha})] + o_p(1)$. If we plug $\bar{\alpha} = \alpha^0 + \frac{1}{n} \tau \frac{1}{m} \Sigma - S (\alpha^0)$ into $Q(\alpha)$ we get, after simplification,

$$Q(\bar{\alpha}) = D(\alpha^0) - \frac{1}{2n} \tau [S(\alpha^0)]' \frac{(k+1)(k-1)}{(k-1)} \frac{1}{m} \Sigma - [S(\alpha^0)].$$

Also if we plug $\bar{\alpha}_H = (I - A'A) \bar{\alpha}$ into $Q(\alpha)$, we get, after simplification,

$$Q(\bar{\alpha}_H) = D(\alpha^0) - \frac{1}{2n} \tau [S(\alpha^0)]' \frac{(k+1)(k-1)}{(k-1)} (I - A'A) [S(\alpha^0)].$$

Since $I - A'A$ is idempotent and $1'[I - A'] = 0$. Therefore, the difference $Q(\bar{\alpha}_H) - Q(\bar{\alpha})$ is equal to

$$D(\alpha^0) - \frac{1}{2n} \tau [S(\alpha^0)]' \frac{(k+1)(k-1)}{(k-1)} (I - A'A) [S(\alpha^0)] - D(\alpha^0)$$

$$- \frac{1}{2n} \tau [S(\alpha^0)]' \frac{(k+1)(k-1)}{(k-1)} \Sigma - [S(\alpha^0)].$$

$$= \frac{1}{2n} \tau [S(\alpha^0)]' \frac{(k+1)(k-1)}{(k-1)} \Sigma - [S(\alpha^0)].$$

Therefore

$$\frac{2(Q(\bar{\alpha}_H) - Q(\bar{\alpha}))}{\tau} = \frac{1}{t} \frac{(k+1)(k-1)}{(k-1)^r} [S(\alpha^0)]' (A'A) [S(\alpha^0)].$$

(5.3.1.2)
But under the true value $\alpha^0$, $\sqrt{t} \frac{1}{t} S(\alpha^0)$ converges in distribution to a multinormal distribution with mean vector $0$ and variance-covariance matrix $\frac{(t+1)(k-1)}{(k+1)(t-1)} \Sigma$. Since $A'A$ is idempotent, then $\sqrt{r} A'A \frac{1}{t} S(\alpha^0)$ converges in distribution to a multinormal distribution with variance-covariance matrix $\frac{(k+1)(t-1)}{(k-1)t} A'A$. Hence

$$\frac{1}{t} \frac{(k+1)(t-1)}{(k-1)r} [S(\alpha^0)]' (A'A) [S(\alpha^0)]$$

converges in distribution to a chi-square random variable with $q$ d.f. under the true value, as $n$ tends to infinity. Since

$$\text{Drop} = 2[D(\hat{\alpha}) - D(\hat{\alpha}_H)]/\tau$$

$$= \frac{1}{t} \frac{(k+1)(t-1)}{r(k-1)} [S(\alpha^0)]' (A'A) [S(\alpha^0)] + o_p(1)$$  \hspace{1cm} (5.3.1.3)

then $\text{Drop} = 2[D(\hat{\alpha}) - D(\hat{\alpha}_H)]/\tau$ converges to a chi-square random variable with $q$ d.f. under the true value, as $r$ tends to infinity. If we replace $\tau$ by a consistent estimate $\hat{\tau}$ in $\text{Drop} = 2[D(\hat{\alpha}) - D(\hat{\alpha}_H)]/\tau$, it will again follow (by Slusky's theorem) an asymptotic chi-square random distribution for large $r$. In fact, Drop is analogous to $-2\log \lambda$ in the maximum likelihood technique.

### 5.3.2 Test based on the Gradient Vector

In this section we develop a test statistic based on the gradient vector evaluated at the reduced model rank estimate. From Theorem 3, Section 2.2.2, we can write

$$\frac{1}{\sqrt{r}} S(\alpha) = \frac{1}{\sqrt{r}} S(\alpha^0) - \sqrt{r} \frac{1}{r} \frac{(t+1)(k-1)}{(k+1)(t-1)} \Sigma(\alpha - \alpha^0) + o_p(1)$$  \hspace{1cm} (5.3.2.1)

where $o_p(1)$ tends to zero in probability uniformly for all values of $\alpha$ such that $\sqrt{r} \|\alpha - \alpha^0\| \leq c$ for any $c > 0$ and $\alpha^0$ is the true value under the hypothesis.
Theorem 6: Consider the Repeated Measures BIBD and suppose that the null hypothesis $A\alpha = 0$ holds. Then $\frac{1}{r} \frac{(k+1)(t-1)}{r(k-1)} [S(\hat{\alpha}_H)]' (A'A) [S(\hat{\alpha}_H)]$ follows an approximate chi-square distribution with $q$ d.f. for large $r$.

Proof: Since $\sqrt{r} (\hat{\alpha}_H - \alpha^0)$ is bounded in probability, we can replace $\alpha$ by $\hat{\alpha}_H$. Therefore plugging $\hat{\alpha}_H$ into $\frac{1}{\sqrt{r}} S_\alpha$, we get $\frac{1}{r} [S(\hat{\alpha}_H)] = \frac{1}{\sqrt{r}} S(\alpha^0) - \sqrt{r} \frac{1}{r} \frac{(k+1)(t-1)}{(k-1)} (t+1) \Sigma (\hat{\alpha}_H - \alpha^0) + \text{op}(1)$. Therefore $A'A \frac{1}{r} [S(\hat{\alpha}_H)] = A'A \frac{1}{\sqrt{r}} S(\alpha^0) - A'A \sqrt{r} \frac{1}{r} \frac{(k+1)(t-1)}{(k-1)} (t+1) \Sigma (\hat{\alpha}_H - \alpha^0) + \text{op}(1)$. But the second term is 0 by hypothesis. Therefore $A'A \frac{1}{r} [S(\hat{\alpha}_H)] = A'A \frac{1}{\sqrt{r}} S(\alpha^0) + \text{op}(1)$. But $\frac{1}{\sqrt{r}} S(\alpha^0)$ converges to a multinormal distribution with mean 0 and variance $\frac{(t+1)(k-1)}{k+1} \Sigma$. Hence $A'A \frac{1}{r} [S(\hat{\alpha}_H)]$ converges to a multinormal random variable with mean 0 and variance $\frac{r(k-1)}{(k+1)(t-1)} A'A$. If we make a quadratic form out of $A'A \frac{1}{r} [S(\hat{\alpha}_H)]$ it will follow an approximate chi-square distribution with $q$ d.f. But that quadratic form is

$$\{A'A \frac{1}{r} [S(\hat{\alpha}_H)]\}' \frac{r(k-1)}{(k+1)(t-1)} A'A - 1 \{A'A \frac{1}{r} [S(\hat{\alpha}_H)]\}$$

$$= \frac{(k+1)(t-1)}{r(k-1)} \frac{1}{r} [S(\hat{\alpha}_H)]' (A'A) [S(\hat{\alpha}_H)] \quad (5.3.2.2)$$

Hence the result. This result is analogous to Rao’s score statistic in maximum likelihood techniques. When $\alpha = 0$, this test statistic becomes Friedman’s statistic. This gives justification for Friedman’s statistic from a nonparametric Repeated Measures Model point of view.
5.3.3 Test based on the Full Model Estimate

A third approach to testing $\alpha_0 = 0$ is based directly on the full model $R$-estimate $\hat{\alpha}$ determined by minimizing $D(\alpha)$ or by solving $S(\alpha) = 0$. We know that $\sqrt{r} (\hat{\alpha} - \alpha_0)$ converges in distribution to a multinormal distribution with a vector mean zero and variance-covariance matrix $\tau^2 (k+1)(t-1) \Sigma^{-1}$. Therefore $\sqrt{r} \hat{\alpha} \sim N(0, \tau^2 (k+1)(t-1) \Sigma^{-1})$. Hence, if we make a quadratic form out of $\hat{\alpha}$, we get

$$r(\hat{\alpha}^\prime A ') [\tau^2 (k+1)(t-1) \Sigma^{-1} A ' - A ']^{-1} A \hat{\alpha} = r(\hat{\alpha}^\prime A ') \tau^2 (k+1)(t-1) \Sigma^{-1} A \hat{\alpha}$$

(5.3.3.1)

and this follows a chi-square distribution with $q$ d.f. for large $n$. If we replace $\tau$ by a consistent estimate $\hat{\tau}$ in $\frac{\hat{\tau}^2}{\tau^2}$, it will again follow a chi-square random variable for large $n$ if we invoke Slusky's theorem. This test statistic is like Wald's statistic.

5.4 Multiple Comparisons based on the Full Model Estimate

If the test for $\alpha_0 = 0$ is significant, the next step is to decide which of the $q$ contrasts are responsible for the rejection. In this section we develop some multiple comparison procedures based on the full model estimates. From Section 2.2.3 we have

$$\sqrt{n} (\hat{\alpha} - \alpha_0) \sim N[0, \tau^2 (k+1)(t-1) \Sigma^{-1}]$$. Hence $\sqrt{r} \hat{\alpha}_i \alpha \hat{\alpha}$ is approximately normally
distributed with mean zero and variance \( \tau^2 \frac{(k+1)(t-1)}{t(k-1)} \), where \( a'i1 = 0 \). We then have the following results:

5.4.1 Least Significant Difference

A 100(1-\( \alpha \))% confidence interval for a contrast \( a'i \alpha \) is given by

\[
[a'i \hat{\alpha} - z_{\alpha/2} \sqrt{\tau^2 \frac{(k+1)(t-1)}{t(k-1)r}}, a'i \hat{\alpha} + z_{\alpha/2} \sqrt{\tau^2 \frac{(k+1)(t-1)}{t(k-1)r}}],
\]

where \( z_{\alpha/2} \) is the \((1-\alpha/2)\)th quantile of the standard normal distribution.

5.4.2 Bonferroni Intervals

\[
[a'i \hat{\alpha} - z_{\alpha/2k} \sqrt{\tau^2 \frac{(k+1)(t-1)}{t(k-1)r}}, a'i \hat{\alpha} + z_{\alpha/2k} \sqrt{\tau^2 \frac{(k+1)(t-1)}{t(k-1)r}}],
\]

i = 1, 2, ..., k. The overall coverage probability for all k intervals is at least 1-\( \alpha \).

5.4.3 Tukey's Procedure

\[
[\hat{\alpha} - q_{\alpha,p,\infty} \sqrt{\tau^2 \frac{(k+1)(t-1)}{t(k-1)r}}, \hat{\alpha} + q_{\alpha,p,\infty} \sqrt{\tau^2 \frac{(k+1)(t-1)}{t(k-1)r}}],
\]

where \( q_{\alpha,p,\infty} \) is the upper \((1-\alpha)\)th percentile for the range of \( t \) independent \( N(0,1) \) random variables.

5.4.4 Sheffe's Procedure

Let \( f \) be a q-dimensional vector of contrasts in the \( \alpha_i \)'s. A 100(1-\( \alpha \))% confidence interval for any linear function \( h'f \) is given by
and the overall probability for the whole class of such intervals is exactly 1 - α.

5.5 Asymptotic Relative Efficiency

The three statistics developed in Section 2.2 for testing $A\alpha = 0$ have the same asymptotic distribution under the null hypothesis. In this section we develop the Pitman efficiency by considering the limiting distribution when a sequence of alternatives converges to the null hypothesis. It is well known that when statistics have limiting chi-square distribution, the efficiency is developed by the ratio of noncentrality parameters. For the linear model with i.i.d. errors, McKean and Hettmansperger (1976), Sen and Puri (1977), and Adichie (1978) show that for the three tests, the efficiency of any one of these tests relative to the least square F test is

$$e(\text{Rank, Least-square}) = 12\tau^2 \int_{-\infty}^{\infty} [f(x)]^2 dx$$

(5.5.1)

Thus the three tests are asymptotically equivalent in the sense of Pitman efficiency and they inherit the efficiency of the Wilcoxon signed rank test, the Mann-Whitney-Wilcoxon test and Kruskall-Wallis test.

Puri and Sen (1971) pointed out that for the Repeated Measures BIBD, the efficiency of Durbin's statistic relative to the corresponding least square F statistic, is

$$e(\text{Rank, Least Square}) = 12\sigma^2 (1-p)(\frac{k}{k+1})\int_{-\infty}^{\infty} [f(x,x)]^2$$

(5.5.2)

Our conjecture is that the same result will hold for other two statistics Wald and Drop in dispersion.
6.1 Two Way Models with One Repeated Measures Factor and one Grouping Factor

Let us consider a two way model with p rows (treatments) and q columns (grouping factor) in which each subject receives each row treatment but only one column treatment (this would occur, for example, if column treatments were race, sex, or degree of illness, etc.). We assume that there are n subjects who receive each column treatment. Let \( Y_{jk} = (y_{1jk}, y_{2jk}, ..., y_{pjk}) \) be the vector of observations on the kth subject who receives the jth column treatment, \( j = 1, 2, ..., q; k = 1, 2, ..., n \). A model would then be

\[
y_{ijk} = \mu + \alpha_i + \delta_j + \gamma_{ij} + \varepsilon_{ijk}
\]  

(6.1.1)

\( \mu \) = overall mean, \( \alpha_i \) = effect of ith treatment; \( \delta_j \) = effect of jth group treatment; \( \gamma_{ij} \) = interaction between ith row treatment and jth group treatment, \( \sum_{i=1}^{p} \alpha_i = 0, \sum_{j=1}^{q} \delta_j = 0, \sum_{j=1}^{q} \gamma_{ij} = 0 \) and \( \varepsilon_{ijk} \) = the error term. Let \( \varepsilon_{jk} \) be the error vectors corresponding to \( Y_{jk} \). We assume the error vectors \( \varepsilon_{jk} \) are i.i.d. and elements of each \( \varepsilon_{jk} \) are exchangeable random variables. A dispersion function for the kth subject in the jth group is

\[
D_{jk}(\alpha, \delta, \gamma) = \sqrt{\frac{p}{12}} \sum_{i=1}^{p} \left[ \frac{R(Y_{ijk} - \mu - \alpha_i - \delta_j - \gamma_{ij})}{p+1} \cdot \frac{1}{2} \right] [Y_{ijk} - \mu - \alpha_i - \delta_j - \gamma_{ij}] \]  

(6.1.2)

\[
= \sqrt{\frac{p}{12}} \sum_{i=1}^{p} \left[ \frac{R(Y_{ijk} - \alpha_i - \gamma_{ij})}{p+1} \cdot \frac{1}{2} \right] [Y_{ijk} - \alpha_i - \gamma_{ij}] \]  

(6.1.3)
where R stands for the rank of the residual within subjects.

The combined dispersion function for the model is

\[ D(\alpha_1, \beta_1, \gamma_1) = \sqrt{\frac{12}{n}} \sum_{k=1}^{n} \left( \sum_{j=1}^{q} \sum_{i=1}^{p} \frac{R(Y_{ijk} - \alpha_i - \gamma_{ij})}{p+1} - \frac{1}{2} \right) [Y_{ijk} - \alpha_i - \gamma_{ij}] \] (6.1.4)

This dispersion function shares all the properties of Jaeckel's dispersion function.

6.2 Rank Estimates for Two-Way Model

Our main objective in this section is to develop R-estimates of the \( \alpha_i \)'s. Rank estimates of \( \alpha \) and \( \gamma \) are the values \( \hat{\alpha} \) and \( \hat{\gamma} \) which minimize \( D(\alpha, \beta, \gamma) \). We would like to rewrite \( D(\alpha, \beta, \gamma) \) as \( D(\theta_1, \theta_2, ..., \theta_j, ..., \theta_q) \) where \( \theta_j = (\theta_{1j}, \theta_{2j}, ..., \theta_{ij}, ..., \theta_{pj}) \), \( j = 1, 2, ..., q \), \( \theta_{ij} = \alpha_i + \gamma_{ij} \) and \( \frac{1}{q} \sum_{j=1}^{q} \theta_{ij} = \alpha_i \). Therefore

\[ D(\theta_1, \theta_2, ..., \theta_j, ..., \theta_q) = \sqrt{\frac{12}{n}} \sum_{k=1}^{n} \left( \sum_{j=1}^{q} \sum_{i=1}^{p} \frac{R(Y_{ijk} - \theta_{ij})}{p+1} - \frac{1}{2} \right) [Y_{ijk} - \theta_{ij}] \] (6.2.1)

Also the dispersion function for the \( j \)th group is

\[ D(\theta_j) = \sqrt{\frac{12}{n}} \sum_{k=1}^{n} \left( \sum_{i=1}^{p} \frac{R(Y_{ijk} - \theta_{ij})}{p+1} - \frac{1}{2} \right) [Y_{ijk} - \theta_{ij}] \] (6.2.2)

Therefore
\[ D(\Theta) = D(\theta_1, \theta_2, \ldots, \theta_j, \ldots, \theta_q) \]
\[ = \sqrt{12} \sum_{k=1}^{n} \sum_{j=1}^{q} \sum_{i=1}^{p} \left( \frac{R(Y_{ijk} - \theta_{ij})}{p+1} - \frac{1}{2} \right) [Y_{ijk} - \theta_{ij}] \]
\[ = \sum_{j=1}^{q} D(\theta_j) = \sum_{j=1}^{q} \sqrt{12} \sum_{k=1}^{n} \sum_{i=1}^{p} \left( \frac{R(Y_{ijk} - \theta_{ij})}{p+1} - \frac{1}{2} \right) [Y_{ijk} - \theta_{ij}] \quad (6.2.3) \]

The domain (\( \Theta \) space) of \( D(\Theta) \) is divided into a finite number of convex polygonal subsets, on each of which \( D(\Theta) \) is linear function of \( \Theta \). The partial derivatives exist almost everywhere and the negative of the partial derivatives are given by the vectors

\[ S_j(\theta_j) = (s_1(\theta_j), s_2(\theta_j), \ldots, s_p(\theta_j))' \quad (6.2.4) \]

where

\[ s_i(\theta_j) = \sqrt{12} \sum_{k=1}^{n} \left( \frac{R(Y_{ijk} - \theta_{ij})}{p+1} - \frac{1}{2} \right), \quad i = 1, 2, \ldots, p \quad (6.2.5) \]

Hence minimizing \( D(\Theta) \) is equivalent to solving \( s_j(\theta_j) = 0 \). But this is an extension of Hodges-Lehmann estimation to the linear model since \( E_0 s_j(\theta_j) = 0 \).

### 6.2.1 Asymptotic Distribution of \( S_j(\theta_j^0) \) Under True Value \( \theta_j^0 \)

Without loss of generality, let us assume that \( \theta_j^0 = 0 \). Therefore \( E_0(S_j(0)) = 0 \) and \( \text{Var-Cov}_0(S_j(0)) = \Sigma \) where \( \Sigma = \frac{1}{p+1} [pI - 11'] \) where \( I \) is an identity matrix of order \( pxp \) and \( 1 \) is a \( px1 \) column vector of 1's. Using the result of section 2.2.1 it can be shown that

\[ \sqrt{n} \frac{S_j(0)}{n} \sim N(0, \Sigma) \text{ as } n \text{ tends to infinity; } j = 1, 2, \ldots, q. \quad (6.2.1.1) \]
6.2.2 Linear Approximation to the Negative of the Gradient Vector

A linear approximation to the gradient vector is crucial to the development of the distribution theory of rank estimates and tests for the repeated measures model. First we would like to find the expected value of the gradient vectors under true value $\theta_j^0$.

From section 2.2.2

$$E_0(\frac{1}{n} S_j(\theta_j)) = \frac{1}{\tau} \Sigma (\theta_j - \theta_j^0), j = 1, 2, ..., q \quad (6.2.2.1)$$

where $\tau = \frac{1}{\sqrt{12} \int_{-\infty}^{+\infty} f(y,y)dy}$. Now we would like to prove a theorem.

**Theorem 1:** $\frac{1}{n} S_j(\theta_j)$ converges almost surely to $-\frac{1}{\tau} \Sigma (\theta_j - \theta_j^0), j = 1, 2, ..., q$, as $n$ tends to infinity.

**Proof:** Replacing $\alpha$ by $\theta_j$, $\alpha^0$ by $\theta_j^0$, $S(\alpha)$ by $S_j(\theta_j)$, $s_i(\alpha)$ by $s_i(\theta_j)$, $Y_{ij}$ by $Y_{ijk}$ and $j$ by $j_k$ in the proof of Theorem 1 in Section 2.2.2, we get the desired result.

**Theorem 2:** Suppose $\int_{-\infty}^{+\infty} f(y,y)dy < \infty$ and the true value is $\theta_j^0 = 0$. For $\varepsilon > 0$ and $\Delta_j > 0$, $\lim P_0 \{ \| \sqrt{n} \frac{1}{n} S_j(\theta_j^0) - \frac{1}{n} S_j(0) + \frac{1}{\tau} \Sigma \Delta_j \| > = \varepsilon \} = 0$ as $n$ tends to infinity; $j = 1, 2, ..., q$.

**Proof:** Replacing $\Delta$ by $\theta_j$, $S(\Delta)$ by $S_j(\Delta_j)$, $s_i(\Delta)$ by $s_i(\Delta_j)$ in the proof of Theorem 2, Section 2.2.2, we get the desired result.
Theorem 3: \( \lim_n P_0 \{ \sup \{ ||A_j|| \leq c_j \} \left[ \| \sqrt{n} \frac{1}{n} S_j(\frac{A_j}{\sqrt{n}}) - \frac{1}{n} S_j(0) + \frac{1}{n} \sum \Lambda_j \| \geq \varepsilon \} \right] \} = 0 \)
where \( c_j > 0 \).

Proof: Replacing \( \Lambda \) by \( \Theta_j \), \( S(\Lambda) \) by \( S_j(\Lambda_j) \), \( s_i(\Lambda) \) by \( s_i(\Lambda_j) \) in Theorem 3, Section 2.2.2., we get the desired result.

Theorem 4: \( \lim P_0 \{ \sup ||A_j|| \leq c_j \| D_j(A_j) - Q_j(A_j) \| \geq \varepsilon \} = 0 \)
where

\[
D_j(A_j) = \sqrt{12} \sum_{k=1}^{n} \sum_{i=1}^{p} \frac{R(Y_{ijk} - \frac{A_{ij}}{\sqrt{n}})}{p+1} \frac{1}{2} \{ Y_{ijk} - \frac{A_{ij}}{\sqrt{n}} \}
\]

and

\[
Q_j(A_j) = D_j(0) + \frac{1}{2} \sum \frac{1}{n} \sum j \Lambda_j S_j(0)
\]  

(6.22.2)

Cor: Theorems 2, 3, and 4 are equivalent.

Proof: Replacing \( D_n(A) \) by \( D_j(A_j) \) and \( Q_n(A) \) by \( Q_j(A_j) \) in corollary of Section 2.2.2., we get the desired result.

Theorem 5: Suppose \( \int_{-\infty}^{\infty} f(y, y) \, dy < \infty \) and the true value is \( \Theta^0 = 0 \). For \( \varepsilon > 0 \) and \( \Lambda_{ij} > 0 \), \( \lim P_0 \left[ \| \sqrt{n} \sum_{j=1}^{q} \frac{1}{n} S_j(\frac{A_j}{\sqrt{n}}) - \frac{1}{n} S_j(0) + \frac{1}{n} \sum \Lambda_j \| \geq \varepsilon \right] = 0 \) as \( n \) tends to infinity.
Theorem 6: \( \lim_n P_0 \{ \sup \{ \| \Delta_j \| \leq c \} \left\| \sqrt{n} \sum_{j=1}^q \left( \frac{1}{n} S_j \left( \frac{\Delta_j}{\sqrt{n}} \right) - \frac{1}{n} S_j(0) \right) + \frac{1}{\tau} \Sigma \Delta_j \right\| \geq \varepsilon \} \leq \sum_{j=1}^q P_0 \left\{ \left\| \sqrt{n} \frac{1}{n} S_j \left( \frac{\Delta_j}{\sqrt{n}} \right) \right\| \geq \varepsilon \right\} = 0 \)

Proof: Similar to Theorem 5 of this section using theorem 3 of this section.

Theorem 7: \( \lim P_0 \{ \sup \{ \| \Delta_j \| \leq c \} \left\{ \left\| \sqrt{n} \sum_{j=1}^q \left( \frac{1}{n} S_j \left( \frac{\Delta_j}{\sqrt{n}} \right) - \frac{1}{n} S_j(0) \right) + \frac{1}{\tau} \Sigma \Delta_j \right\| > \varepsilon \} \right\} = 0 \).

Proof: Similar to theorem 5 of this section using theorem 4 of this section.

6.2.3 Asymptotic Distribution of Rank Estimates Under Full Model

Let \( \bar{\Theta}_1, \bar{\Theta}_2, \ldots, \bar{\Theta}_q \) minimize \( Q(\Theta) \); then \( \bar{\Theta}_j = \frac{1}{n} \Sigma S_j(0) \). If we assume that \( \Theta_j^0 \) is the true value of \( \Theta_j \) then \( \bar{\Theta}_j = \Theta_j^0 + \frac{1}{n} \Sigma S_j(\Theta_j^0) \). From theorem 4, \( \sqrt{n} (\bar{\Theta}_j - \Theta_j^0) \) converges to zero in probability. Since \( \sqrt{n} \frac{S_j(0)}{n} \sim N(0, \Sigma) \) as \( n \) tends to infinity, therefore \( \sqrt{n} (\bar{\Theta}_j - \Theta_j^0) \) converges to the multinormal distribution with mean zero and variance \( \tau^2 \Sigma \). Hence \( \sqrt{n} (\hat{\Theta}_j - \Theta_j^0) \) converges to the multinormal distribution with mean zero and variance \( \tau^2 \Sigma \).

6.2.4 Asymptotic Distribution of Rank Estimates Under Reduced Model

Case 1: We would like to minimize \( D(\Theta) = \sum_{j=1}^n D_j(\Theta_j) \) subject to the restriction \( A \alpha = 0 \) where \( A1 = 0 \) and \( A \) is an mxp matrix with rank \( m < p \) and rows of \( A \) are mutually orthogonal. The vector \( \hat{\Theta}^H \) which minimizes \( D(\alpha) \) subject to the restriction \( A \alpha = 0 \) is
called the reduced model estimate of $\alpha$. To minimize $D(\Theta)$ subject to $A\alpha = 0$, we will use the method of Lagrangian multipliers. Since $\frac{1}{q} \sum_{j=1}^{q} \Theta_j = \alpha$ we have $A\alpha = \frac{1}{q} A \sum_{j=1}^{q} \Theta_j = 0$; therefore $\sum_{j=1}^{q} A \Theta_j = 0$. Let $\theta_j^0$ be general solution to the above restrictions.

Let us define $\phi = D(\Theta) + \lambda^t \sum_{j=1}^{q} A \Theta_j$. Therefore $\frac{\partial \phi}{\partial \Theta_j} = -S_j(\Theta_j) + A^t \lambda = 0$, $j = 1, 2, \ldots, q$; and $\frac{\partial \phi}{\partial \lambda} = \sum_{j=1}^{q} A \Theta_j = 0$. Now we would like to get an approximation to $\phi$ using Theorem 3, Section 6.2.2. It can be shown that

$$\frac{1}{\sqrt{n}} \frac{\partial \phi}{\partial \Theta_j} = \frac{1}{\sqrt{n}} S(\Theta_j^0) - \sqrt{n} \frac{1}{\tau} \Sigma(\Theta_j - \Theta_j^0) + o_p(1)$$

where $o_p(1)$ tends to zero in probability uniformly for all values of $\Theta_j$ such that $\sqrt{n} \| \Theta_j - \Theta_j^0 \| \leq c_j$ for any $c_j > 0$ and $\Theta_j^0$ is the true value under the null hypothesis. By theorem 7, section 6.2.2, we can construct a quadratic approximation to $\phi$ such as $\psi = Q(\Theta) + \lambda^t A \alpha = Q(\Theta) + \lambda^t \sum_{j=1}^{q} A \Theta_j$. Hence $\sqrt{n} (\tilde{\Theta}_j^H - \tilde{\Theta}_j^0)$ converges to zero in probability. Let $\tilde{\Theta}_j^H$ be the minima of $Q(\Theta)$ subject to the restriction $\frac{1}{q} \sum_{j=1}^{q} A (\Theta_j) = 0$. Then $\tilde{\Theta}_j^H$ is a solution of the equation $\frac{\partial \psi}{\partial \Theta_j} = 0$, $j = 1, 2, \ldots, q$ and $\frac{\partial \psi}{\partial \lambda} = 0$. By solving those two equations, we get

$$\tilde{\Theta}_j^H = \theta_j^0 + \tilde{\Theta}_j - \frac{1}{q} A^t A \sum_{j=1}^{q} \tilde{\Theta}_j$$

(6.2.4.1)

where $\tilde{\Theta}_j$ are the minima of $Q(\Theta)$ without any restriction.

In section 6.2.3 it has been shown that $\sqrt{n} (\tilde{\Theta}_j - \theta_j^0)$ converges to a multinormal distribution with mean zero and variance $\tau^2 \Sigma$; therefore $\sqrt{n} (\tilde{\Theta}_j^H - \theta_j^0) = \sqrt{n} [\tilde{\Theta}_j - \frac{1}{q} A^t A \sum_{j=1}^{q} \tilde{\Theta}_j]$ converges to a multivariate normal distribution with mean zero and variance $(I - A^t A) \tau^2 \Sigma (I - A^t A)$. Hence $\sqrt{n} (\tilde{\Theta}_j^H - \theta_j^0)$ converges to a multivariate normal distribution with mean zero and variance.
Case 2: We would like to minimize $D(\Theta)$ subject to the restriction
$\theta_1 = \theta_2 = \ldots = \theta_q = \ldots = \beta$ (say). Therefore $D(\Theta)$ becomes

$$D(\beta) = \sqrt{12} \sum_{k=1}^{n} \sum_{j=1}^{q} \sum_{i=1}^{p} \left[ \frac{R(Y_{ijk} - \beta_i)}{p+1} \right] \frac{1}{2} [Y_{ijk} - \beta_i].$$

Let $\hat{\beta}$ minimize $D(\hat{\beta})$; then $\hat{\beta}$ is the reduced model rank estimate of $\beta$. Differentiating $D(\beta)$ with respect to $\beta$ we get the negative of the gradient as

$$s_i(\beta) = \sqrt{12} \sum_{k=1}^{n} \sum_{j=1}^{q} \sum_{i=1}^{p} \left[ \frac{R(Y_{ijk} - \beta_i)}{p+1} \right] \frac{1}{2} [Y_{ijk} - \beta_i].$$

An approximation to $S(\beta)$ is as follows: $\frac{1}{\sqrt{n}} S(\beta) = \frac{1}{\sqrt{n}} S(\beta^0) - q\sqrt{n} \frac{1}{\tau} \Sigma(\beta - \beta^0) + o_p(1)$.

Using the procedure in section 2.2.2 it can be shown that $\overline{\beta} = \beta^0 + \frac{1}{qn} \tau^2 \Sigma S(\beta^0)$ and $\hat{\beta}$ are asymptotically equivalent. But $\sqrt{n} \frac{1}{nq} S(\beta^0)$ converges to a multinormal distribution with mean $\theta$ and variance $\frac{1}{q^2} \Sigma$. Therefore $\sqrt{n} (\overline{\beta} - \beta^0)$ is asymptotically multinormally distributed with mean $0$ and variance $\frac{1}{q^2} \tau^2 \Sigma^-$. Hence $\sqrt{n} (\overline{\beta} - \beta^0)$ converges to a multinormal distribution with mean $0$ and variance $\frac{1}{q^2} \tau^2 \Sigma^-$. (6.2.4.3)

6.3.1 Rank Tests in Repeated Measures

In this section we would like to develop three different statistics based on the rank estimate of $\theta_j$ for testing hypothesis of the form $\frac{1}{q} \sum_{j=1}^{q} A \theta_j = 0$ where $A1 = 0$ and the rank of $A = q < p$. 
6.3.1.1 Test Based on Drop in Dispersion Function

In the last section we used \( D(\Theta) = \sqrt{12} \sum_{k=1}^{n} \sum_{j=1}^{q} \sum_{i=1}^{p} \left[ \frac{R(Y_{ijk} - \theta_{ij})}{p+1} - \frac{1}{2} \right] [Y_{ijk} - \theta_{ij}] \) to get an estimate of the \( \theta_j \)'s. In fact \( D(\Theta) \) is used as a criterion for fitting a linear model to data. \( D(\hat{\Theta}) \) represents the minimum distance, as measured by \( D(\Theta) \), from the data vector to the subspace spanned by the linear model. Let \( \hat{\theta}_j^H \) be the R-estimate of \( \theta_j \) under \( A\alpha = 0 \).

To test the hypothesis \( A\alpha = 0 \) versus \( A\alpha \neq 0 \), we will compare \( D(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_q) \) with \( D(\hat{\theta}_1^H, \hat{\theta}_2^H, ..., \hat{\theta}_q^H) \). This is the same strategy as that used to develop F tests based on the reduction in sum squares due to fitting reduced and full models. To make the test operational, we at least need the limiting distribution under the null hypothesis. The test is not distribution free for finite sample size. The following theorem shows that the test is asymptotically distribution free under the null hypothesis.

**Theorem 5:** Given the Repeated Measures Two way model and \( A\alpha = 0 \),
\[
D^* = \text{Drop} = \frac{2[D(\hat{\theta}_1^H, \hat{\theta}_2^H, ..., \hat{\theta}_q^H) - D(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_q)]}{\tau}
\]
has an asymptotic chi-square distribution with \( m \) d.f.

**Proof:** The argument proceeds by approximating \( D(\Theta) \) with \( Q(\Theta) \) and then showing that the asymptotic behavior of \( D(\hat{\theta}_1^H, \hat{\theta}_2^H, ..., \hat{\theta}_q^H) - D(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_q) \) can be determined by that of \( Q(\hat{\theta}_1^H, \hat{\theta}_2^H, ..., \hat{\theta}_q^H) - Q(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_q) \).

We begin by writing
\[ D(\theta^1_H, \theta^2_H, \ldots, \theta^q_H) - D(\tilde{\theta}^1_H, \tilde{\theta}^2_H, \ldots, \tilde{\theta}^q_H) \]

\[ = [D(\theta^1_H, \theta^2_H, \ldots, \theta^q_H) - Q(\theta^1_H, \theta^2_H, \ldots, \theta^q_H)] \]

\[ + [Q(\theta^1_H, \theta^2_H, \ldots, \theta^q_H) - Q(\tilde{\theta}^1_H, \tilde{\theta}^2_H, \ldots, \tilde{\theta}^q_H)] \]

\[ + [Q(\tilde{\theta}^1_H, \tilde{\theta}^2_H, \ldots, \tilde{\theta}^q_H) - Q(\theta^1, \theta^2, \ldots, \theta^q)] \]

\[ + [Q(\theta^1, \theta^2, \ldots, \theta^q)] - D(\theta^1, \theta^2, \ldots, \theta^q) \quad (6.3.1.1) \]

It can be shown that the first, second, third, and fifth terms are \( o_p(1) \) as in Theorem 5 of Chapter 2. Hence

\[ D(\theta^1_H, \theta^2_H, \ldots, \theta^q_H) - D(\theta^1_H, \theta^2_H, \ldots, \theta^q_H) \]

\[ = Q(\tilde{\theta}^1_H, \tilde{\theta}^2_H, \ldots, \tilde{\theta}^q_H) - Q(\theta^1, \theta^2, \ldots, \theta^q) + o_p(1) \]

Since \( \tilde{\theta}^j_H = \theta^0_j + \frac{1}{q} A'A \sum_{j=1}^{q} \tilde{\theta}^j \), therefore

\[ Q(\tilde{\theta}^1_H, \tilde{\theta}^2_H, \ldots, \tilde{\theta}^q_H) = D(\theta^0_1, \theta^0_2, \ldots, \theta^0_q) \]

\[ - \frac{1}{2} n \sum_{i=1}^{q} \left[ \tilde{\theta}^j_i - \frac{1}{q} A'A \tilde{\theta}^j_i \right] \frac{1}{q} \sum_{i=1}^{q} \left[ \tilde{\theta}^j_i - \frac{1}{q} A'A \tilde{\theta}^j_i \right] \]

\[ + \frac{1}{2} n \sum_{j=1}^{q} \left[ \tilde{\theta}^j - \frac{1}{q} A'A \tilde{\theta}^j \right] \frac{1}{q} \sum_{j=1}^{q} \left[ \tilde{\theta}^j - \frac{1}{q} A'A \tilde{\theta}^j \right] \]

Hence

\[ Q(\tilde{\theta}^1_H, \tilde{\theta}^2_H, \ldots, \tilde{\theta}^q_H) - Q(\tilde{\theta}^1, \tilde{\theta}^2, \ldots, \tilde{\theta}^q) \]

\[ = \frac{n}{2} \frac{1}{q} A'A \sum_{j=1}^{q} (\tilde{\theta}^j, \theta^0_j) \frac{1}{q} \sum_{j=1}^{q} A'A(\tilde{\theta}^j - \theta^0_j) \]
But $\sqrt{n} \frac{1}{q} \left[ \sum_{j=1}^{q} (\tilde{\theta}_j - \theta_j^0) \right]$ converges to a multivariate normal with variance $\frac{\tau^2}{q} \Sigma^{-}$. Therefore $\sqrt{n} \frac{1}{q} A' A \sum_{j=1}^{q} (\tilde{\theta}_j - \theta_j^0)$ is asymptotically multivariate normal with variance $\tau^2 A' A \Sigma^{-} A' A$. Therefore $\sqrt{n} \left( \frac{1}{q} A' A \left[ \sum_{j=1}^{q} (\tilde{\theta}_j - \theta_j^0) \right] \right) \Sigma^{-} A' A^{-1} \left[ \frac{1}{q} A' A \sum_{j=1}^{q} (\tilde{\theta}_j - \theta_j^0) \right] + o_p(1)$ (6.3.1.2)

is asymptotically chi-square with m d.f. under null hypothesis.

If we replace $\tau$ by a consistent estimate $\hat{\tau}$ in Drop, it will again follow a chi-square distribution for large $n$ if we invoke Slutsky's theorem. In fact Drop is analogous to $-2\log \lambda$ in maximum likelihood technique.

6.3.1.2 Test Based on the Gradient Vector

In this section we would like to develop a test statistic based on the gradient vector evaluated at the reduced model rank estimate. We have

$$\sum_{j=1}^{q} \frac{1}{\sqrt{n}} S_j(\theta_j)$$

$$= \sum_{j=1}^{q} \frac{1}{\sqrt{n}} S(\theta_j^0) - \sum_{j=1}^{q} \sqrt{n} \frac{1}{\tau} (\theta_j - \theta_j^0) + o_p(1) \quad (6.3.1.2.1)$$
Theorem 6: Given the Repeated Measures Two Way Design and supposing the null hypothesis \( A \alpha = 0 \) holds, then
\[
\frac{p+1}{pq} \frac{1}{n} \left[ \sum_{j=1}^{q} S(\hat{\theta}_j^H) \right]' (A'A) \left[ \sum_{j=1}^{q} S_j(\hat{\theta}_j^H) \right]
\]
follows a chi-square distribution with \( q \) d.f. for large \( n \).

Proof: Since \( \sqrt{n} (\hat{\theta}_j^H - \theta_j^0) \) is bounded in probability, hence we can replace \( \theta_j \) by \( \hat{\theta}_j^H \). Therefore plugging \( \hat{\theta}_j^H \) into
\[
\sum_{j=1}^{q} \frac{1}{\sqrt{n}} S_j(\theta_j) = \sum_{j=1}^{q} \frac{1}{\sqrt{n}} S(\theta_j^0) - \sum_{j=1}^{q} \frac{1}{\tau} \Sigma (\theta_j - \theta_j^0) + o_p(1)
\]
we get
\[
\sum_{j=1}^{q} \frac{1}{\sqrt{n}} S_j(\hat{\theta}_j^H) = \sum_{j=1}^{q} \frac{1}{\sqrt{n}} S(\theta_j^0) - \sum_{j=1}^{q} \frac{1}{\tau} \Sigma (\hat{\theta}_j^H - \theta_j^0) + o_p(1).
\]

Hence \( A'A \sum_{j=1}^{q} \frac{1}{\sqrt{n}} S_j(\hat{\theta}_j^H) = \sum_{j=1}^{q} A'A \frac{1}{\sqrt{n}} S(\theta_j^0) \) after using the hypothesis.

If we make a Wald type quadratic form out of above expression then
\[
\frac{p+1}{pq} \frac{1}{n} \left[ \sum_{j=1}^{q} S(\hat{\theta}_j^H) \right]' (A'A) \left[ \sum_{j=1}^{q} S_j(\hat{\theta}_j^H) \right] \quad (6.3.1.2.2)
\]
follows a chi-square distribution with \( m \) d.f. for large \( n \). Hence the result. This result is analogous to Rao's score statistic in maximum likelihood techniques. When \( A\alpha = 0 \), this test statistic becomes the generalized Friedman's statistic. This gives a justification of the generalized Friedman's statistic from a nonparametric Repeated Measures Model point of view.

6.3.1.3 Test Based on Full Model Estimate

Third approach to testing \( A\alpha = 0 \) is based directly on the full model R-estimate. We know that \( \sqrt{n} (\hat{\theta}_j - \theta_j^0) \) converges to the multinormal distribution with mean zero
and variance $\tau^2 \Sigma$. Therefore $\frac{1}{q} A \sum_{j=1}^{q} \sqrt{n} (\tilde{\theta}_j - \theta_j^0)$ follows the multinormal distribution with variance $\frac{\tau^2}{q} A \Sigma A'$. If we make a Wald type quadratic form out of

$$\frac{1}{q} A \sum_{j=1}^{q} \sqrt{n} (\tilde{\theta}_j - \theta_j^0)$$

we get

$$\frac{p q}{(p+1) \tau^2} n (\sum_{j=1}^{q} \tilde{\theta}_j)' A' A (\sum_{j=1}^{q} \tilde{\theta}_j)$$

(6.3.1.3.1)

converges to a chi-square random variable with $m$ d.f. when the null hypothesis is true. If we replace $\tau$ by $\hat{\tau}$ in

$$\frac{p q}{(p+1) \hat{\tau}^2} n (\sum_{j=1}^{q} \tilde{\theta}_j)' A' A (\sum_{j=1}^{q} \tilde{\theta}_j)$$

it will again follow a chi-square random variable for large $n$ if we invoke Slutsky's theorem. This test statistic is like Wald's statistic.

### 6.3.2 Tests for Interaction

In this section we would like to develop test statistics for testing the interaction between the repeated measures factor and the grouping factor.

#### 6.3.2.1 Drop in Dispersion Test

**Theorem 10:** Given the two way model with exchangeable errors within subjects and assuming $\theta_1 = \theta_2 = \ldots = \theta_q$ (i.e. all the interactions are zero),

$$\text{Drop} = \frac{2[D(\hat{\beta}, \hat{\beta}, \ldots, \hat{\beta}) - D(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_q)]}{\tau}$$

follows a chi-square distribution with $(p-1)(q-1)$ d.f. as $n$ tends to infinity.

**Proof:** Using the approach in theorem 8, it can be shown that

$$D(\hat{\beta}, \hat{\beta}, \ldots, \hat{\beta}) - D(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_q)$$

$$= Q(\tilde{\beta}, \tilde{\beta}, \ldots, \tilde{\beta}) - Q(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_q) + o_p(1).$$

(6.3.2.1.1)
From section 6.2.4 case 2, we have \( \tilde{\beta} = \beta^0 + \frac{1}{qn} \tau \Sigma^- S(\beta^0) \) where

\[
s_i(\beta^0) = (\sum_{k=1}^{n} \sum_{j=1}^{q} \frac{R(Y_{ijk} - \beta^0_j)}{p+1} - \frac{1}{2} [Y_{ijk} - \beta^0_j]; i = 1, 2, \ldots, p.
\]

Therefore

\[
Q(\tilde{\beta}, \tilde{\beta}_1, \ldots, \tilde{\beta}_q) = D(\beta^0) - \frac{\tau}{2qn} S'(\beta^0) \Sigma^- S(\beta^0).
\]

Also under the null hypothesis

\[
Q(\tilde{\beta}, \tilde{\beta}_1, \ldots, \tilde{\beta}_q) = D(\beta^0) - \frac{\tau}{2n} \sum_{j=1}^{q} S_j(\beta^0) \Sigma^- S_j(\beta^0).
\]

Hence

\[
Q(\tilde{\beta}, \tilde{\beta}_1, \ldots, \tilde{\beta}_q) - Q(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_q) = \frac{\tau}{2n} \sum_{j=1}^{q} [S_j(\beta^0) - S^*(\beta^0)] \Sigma^- [S_j(\beta^0) - S^*(\beta^0)]
\]

where

\[
S^*(\beta^0) = \frac{\sum_{j=1}^{q} S_j(\beta^0)}{q}.
\]

Hence

\[
D(\tilde{\beta}, \tilde{\beta}_1, \ldots, \tilde{\beta}_q) - D(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_q)
\]

\[
= \frac{\tau}{2n} \sum_{j=1}^{q} [S_j(\beta^0) - S^*(\beta^0)] \Sigma^- [S_j(\beta^0) - S^*(\beta^0)] + o_p(1)
\]

Therefore

\[
\frac{2[D(\tilde{\beta}, \tilde{\beta}_1, \ldots, \tilde{\beta}_q) - D(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_q)]}{\tau}
\]

\[
= \frac{1}{n} \sum_{j=1}^{q} [S_j(\beta^0) - S^*(\beta^0)] \Sigma^- [S_j(\beta^0) - S^*(\beta^0)] + o_p(1).
\]
But \( \frac{1}{n} \sum_{j=1}^{q} [S_{j}(\beta^0) - S^{*}(\beta^0)] \) follows the chi-square distribution with \((p-1)(q-1)\) d.f as \( n \) tends to infinity if the null hypothesis holds.

6.3.2.2 Test based on Gradient Vector

We have

\[
\frac{1}{\sqrt{n}} \sum_{j=1}^{q} S_j(\theta_j) = \frac{1}{\sqrt{n}} \sum_{j=1}^{q} S(\theta_j^0) - \frac{1}{\sqrt{n}} \sum_{j=1}^{q} \sqrt{n} \frac{1}{\tau} \Sigma(\theta_j - \theta^0) + o_p(1) \quad (6.3.2.2.1)
\]

Under \( \theta_1 = \theta_2 = \ldots = \theta_q = \beta^0, ..., \tilde{\beta} = \beta^0 + \frac{1}{qn} \tau \Sigma S(\beta^0) \). Replacing \( \theta_j \) by \( \hat{\theta} \) in

\[
\frac{1}{\sqrt{n}} \sum_{j=1}^{q} S_j(\theta_j)
\]

and setting \( \theta^0 = \beta^0 \) we get

\[
\frac{1}{\sqrt{n}} \sum_{j=1}^{q} S_j(\hat{\theta}) = \frac{1}{\sqrt{n}} \sum_{j=1}^{q} S(\beta^0) - \frac{1}{\sqrt{n}} \sum_{j=1}^{q} \sqrt{n} \frac{1}{\tau} \Sigma(\hat{\theta} - \beta^0) + o_p(1).
\]

Again replacing \( \hat{\theta} \) by \( \tilde{\beta} \) we get

\[
\frac{1}{\sqrt{n}} \sum_{j=1}^{q} S_j(\tilde{\beta})
\]

\[
= \frac{1}{\sqrt{n}} \sum_{j=1}^{q} S_j(\beta^0) - \frac{1}{\sqrt{n}} \sum_{j=1}^{q} \sqrt{n} \frac{1}{\tau} \Sigma(\tilde{\beta} - \beta^0) + o_p(1)
\]

\[
= \frac{1}{\sqrt{n}} \sum_{j=1}^{q} S_j(\beta^0) - \frac{1}{\sqrt{n}} \sum_{j=1}^{q} \frac{1}{\sqrt{n}} \sum_{j=1}^{q} S_j(\beta^0) + o_p(1)
\]

\[
= \frac{1}{\sqrt{n}} \sum_{j=1}^{q} S_j(\beta^0) - S^*(\beta^0) + o_p(1).
\]
Therefore
\[
\left(\sum_{j=1}^{q} \frac{1}{\sqrt{n}} S_j(\hat{\beta})\right)^\Sigma^- \left(\sum_{j=1}^{q} \frac{1}{\sqrt{n}} S_j(\hat{\beta})\right)
\]
\[
= \frac{1}{\sqrt{n}} \sum_{j=1}^{q} (S_j(\beta^0) - S^*(\beta^0))^\Sigma^- \frac{1}{\sqrt{n}} \sum_{j=1}^{q} (S_j(\beta^0) - S^*(\beta^0)) + o_p(1).
\]

But \( \frac{1}{\sqrt{n}} \sum_{j=1}^{q} (S_j(\beta^0) - S^*(\beta^0))^\Sigma^- \frac{1}{\sqrt{n}} \sum_{j=1}^{q} (S_j(\beta^0) - S^*(\beta^0)) \) follows a chi-square distribution with \((p-1)(q-1)\) d.f. as \( n \) tends to infinity, when the null hypothesis is true.

Hence
\[
\left(\sum_{j=1}^{q} \frac{1}{\sqrt{n}} S_j(\hat{\beta})\right)^\Sigma^- \left(\sum_{j=1}^{q} \frac{1}{\sqrt{n}} S_j(\hat{\beta})\right) \quad (6.3.2.2)
\]
follows the chi-square with \((p-1)(q-1)\) d.f. as \( n \) tends to infinity, and when the null hypothesis is true.

6.3.2.3 Test Based on Full Model Estimate

We know that \( \sqrt{n} (\hat{\theta}_j - \theta_j^0) \) converges to the multinormal distribution with mean zero and variance \( \tau^2 \Sigma^- \) when \( \theta_1 = \theta_2 = ... = \theta_q = \beta^0, \quad \hat{\theta} = \frac{1}{q} \sum_{j=1}^{q} \hat{\theta}_j \). Hence under the null hypothesis
\[
\left(\sum_{j=1}^{q} \frac{\hat{\theta}_j - \frac{1}{q} \sum_{j=1}^{q} \hat{\theta}_j}{\sum_{j=1}^{q} \hat{\theta}_j}\right)^\Sigma^- \left(\frac{\hat{\theta}_j - \frac{1}{q} \sum_{j=1}^{q} \hat{\theta}_j}{\sum_{j=1}^{q} \hat{\theta}_j}\right) \quad (6.3.2.3.1)
\]
follows the chi-square distribution with \((p-1)(q-1)\) d.f. as \( n \) tends to infinity.

6.4. Multiple Comparison Based on Full Model Estimate

If the test for \( \alpha = 0 \) is significant, the next step is to decide which of the \( m \) contrasts are responsible for rejection. In this section we will develop some multiple comparison procedures based on full model estimates.
From section 6.2.3 we have, asymptotically, $\sqrt{n} (\hat{\alpha} - \alpha^0) \sim \text{MVN}[0, \frac{\tau^2}{q} \Sigma^{-1}]$. Hence $a_1' \hat{\alpha}$ is approximately normally distributed with mean zero and variance $\frac{\tau^2(p+1)}{npq}$ where $a_1' = 0$. We have the following approximate results:

6.4.1 Least Significant Difference

A 100(1-\(\alpha\))% confidence interval for a contrast $a_1' \alpha$ is given by

$$[a_1' \hat{\alpha} - z_{\alpha/2} \sqrt{\frac{\tau^2(p+1)}{npq}}, a_1' \hat{\alpha} + z_{\alpha/2} \sqrt{\frac{\tau^2(p+1)}{npq}}]$$

where $z_{\alpha/2}$ is the $(1-\alpha/2)$th quantile of the standard normal distribution.

6.4.2 Bonferroni Intervals:

$$[a_1' \hat{\alpha} - z_{\alpha/k} \sqrt{\frac{\tau^2(p+1)}{npq}}, a_1' \hat{\alpha} + z_{\alpha/k} \sqrt{\frac{\tau^2(p+1)}{npq}}]$$
i = 1, 2, ..., k. The overall probability for all k intervals is at least 1-\(\alpha\).

6.4.3 Tukey's Procedure

$$[\alpha_i - q_{\alpha,p,\infty} \sqrt{\frac{\tau^2(p+1)}{npq}}, \alpha_i + q_{\alpha,p,\infty} \sqrt{\frac{\tau^2(p+1)}{npq}}]$$

where $q_{\alpha,p,\infty}$ is the upper (1-\(\alpha\))th percentile for the range of p independent N(0,1) random variables.

6.4.4 Sheffe's Procedure

Let $f$ be an m-dimensional vector of contrasts in $\alpha_i$'s. A 100(1-\(\alpha\))% confidence interval for any linear function $h'f$ is given by

$$[h'f - \sqrt{\frac{\tau^2 \chi^2(\alpha)}{npq}} \frac{(p+1)}{npq}, h'f + \sqrt{\frac{\tau^2 \chi^2(\alpha)}{npq}} \frac{(p+1)}{npq}]$$

and the overall probability for the whole class of such intervals is exactly 1 - \(\alpha\).
6.5 Asymptotic Relative Efficiency

The three statistics developed in section 6.2 for testing $\alpha = 0$ have the same asymptotic distribution under the null hypothesis. In this section we will develop Pitman efficiency by considering the limiting distribution when a sequence of alternatives converges to the null hypothesis. It is well known that when statistics have an asymptotic chi-square distribution, the efficiency is developed by the ratio of noncentrality parameters. For the linear model with i.i.d. errors, McKean and Hettmansperger (1976), Sen and Puri (1977), and Adichie (1978) show that for the three tests, the efficiency of any one of these tests relative to the least squares $F$ test is

$$E(\text{Rank,Least-square}) = 12\tau^2 \left( \int f^2(x)dx \right)^2$$

(6.5.1)

Thus the three tests are asymptotically equivalent in the sense of Pitman efficiency and they inherit the efficiency of the Wilcoxon signed rank test, the Mann-Whitney-Wilcoxon test and Kruskall-Wallis test. In this section we will show that for the Repeated Measures Randomized Block Design the three tests, e.g., Drop in dispersion, Wald's statistic and Aligned test, the efficiency of any one of these tests relative to the corresponding least squares $F$ statistic is $E(\text{Rank, Least Square})$

$$= 12\sigma^2 (1-p) \frac{p}{p+1} \left( \int f(x)dx \right)^2.$$  

(6.5.2)

6.5.1 Efficacy of Wald Statistic

From section 6.2, we know that $\sqrt{n} (\hat{\alpha} - \alpha)$ is $o_p(1)$. Proceeding like section 2.5.1, we get the noncentrality parameter of Wald statistic
\[
\frac{pq}{(p+1)\tau^2} n \left( \sum_{j=1}^{q} \hat{\theta}_j \right)' A'A \left( \sum_{j=1}^{q} \hat{\theta}_j \right), \text{ under local alternatives, which becomes, after simplification,}
\]
\[
\frac{pq}{(p+1)\tau^2} \left( \alpha - 11'\alpha \right)' A'A \left( \alpha - 11'\alpha \right)
\]
(6.5.11)

6.5.2 Efficacy of the Test statistic Based on the Gradient Vector

Proceeding like section 2.5.2, we get the noncentrality parameter of
\[
\frac{p+1}{pq} n \left[ \sum_{j=1}^{q} S_j(H) \right]' (A'A) \left[ \sum_{j=1}^{q} S_j(H) \right] \text{ under local alternatives is}
\]
\[
\frac{pq}{(p+1)\tau^2} \left( \alpha - 11'\alpha \right)' A'A \left( \alpha - 11'\alpha \right)
\]
(6.5.2.1)

6.5.3 Efficacy of Test Based on Drop in Dispersion

Proceeding like section 2.5.3, we get the noncentrality parameter of Drop in dispersion statistic under local alternatives is
\[
\frac{pq}{(p+1)\tau^2} \left( \alpha - 11'\alpha \right)' A'A \left( \alpha - 11'\alpha \right)
\]
(6.5.3.1)

6.5.4 Efficacy of Least Squares F:

The efficacy of least squares F under exchangeable errors and local alternatives is
\[
\frac{q(\alpha - 11'\alpha)' A'A(\alpha - 11'\alpha)}{\sigma^2(1-\rho)} \text{ since mF converges to a noncentral chi-square distribution with parameter}
\]
\[
\frac{q(\alpha - 11'\alpha)' A'A(\alpha - 11'\alpha)}{\sigma^2(1-\rho)}
\]
(6.5.4.1)
6.5.5 ARE of Rank Tests With Respect to Least Square Test

From the previous four sections we get the ARE of all the rank tests with respect to the least squares test as $E(\text{Rank,Least square})$

$$\frac{qp}{(p+1)\tau^2} \frac{(\alpha -1\alpha)' \ A' A(\alpha -1\alpha)}{q} = \frac{12p\sigma^2(1-p)}{p+1} \left( \int_{-\infty}^{\infty} f(x,x) \ dx \right)^2 \ (6.5.5.1).$$

6.5 Estimation of $\tau$

Two consistent estimates of $\tau$ can be obtained from each group as in 3.3 and in 3.4. If we take the average of those estimates in each case that will be again a consistent estimate of $\tau$. 
CHAPTER VII
FURTHER RESEARCH AND CONCLUSION

7.1 Further Reasearch

In this dissertation we have developed rank based inference for some Repeated Measures Designs using a suitable dispersion function based on Wilcoxon scores and Jaeckel's dispersion function. In fact, we have defined a dispersion function using residuals within each subject and then combined them to get an overall dispersion function. As with the general linear model with i.i.d. errors, one could use a general score function instead of the Wilcoxon score function.

We have used full model R-estimates for making multiple comparisons for contrasts. One could also use the Drop in dispersion statistic to make such multiple comparisons.

7.2 Conclusions

Unlike traditional nonparametric techniques for the designs considred in this thesis, the results of this thesis bring a tie between testing and estimation. Also the testing and estimation have the same efficiency factor which is not true for traditional nonparametric methods.

When the number of treatments is equal to two, the procedure proposed in this thesis yields an estimate which is identical to the Hodges-Lehmann estimate of shift in a paired sample based on sign statistic. In fact, when the number of treatments is greater than two, the developed procedure takes care of the compatibility of the estimates, as pointed out in Hodges-Lehmann (1962). Hence the results of this thesis are extensions
of the Hodges-Lehmann, Jureckova and Jaeckel ideas from the i.i.d. error case to the exchangeable error setting.

Since Friedman's statistic, Durbin's statistic and the Generalized Friedman-statistic are particular cases of one of the statistics proposed in this thesis, it can be concluded that our results give justification for these statistics from a nonparametric repeated measures model point of view.

For non-normal distributions, the results of this thesis show that rank based inferences are better than those based on least square theory. For example when the number of treatments is greater than four and the samples come from the exchangeable logistic distribution, the ARE of the rank tests with respect to their least square counterparts is greater than one. For the exchangeable Cauchy, t with 3 d.f., t with 4 d.f., and the exponential distributions, the ARE is also greater than one.

Our results for Two Way Models can be used for k sample repeated measures as well, for $k \geq 2$.

Like least square theory, multiple comparisons can be based on the full model estimates.

The conclusions obtained from the Monte Carlo study for Repeated Measures RCBD may be applicable to other designs as well.

Analyses of two data sets from Koch and Sen (1968) by our methods bring the same conclusions as they reached.
BIBLIOGRAPHY

Adichie, J. N. (1967a). Asymptotic efficiency of a class non-parametric tests for

of Mathematical Statistics 38, 894-904.


Indagationes Math 15, 358-369.

Bloch, D. A. and Gastwirth J. L. (1968). On a simple estimate of the reciprocal of the


Durbin J. (1951). Incomplete blocks in ranking experiments. British Journal of
Psychology 4, 85-90.


Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in

Association, 56, 335-349.

model. ABTEILUNG STATISTIK, UNIVERSITAT DORMUND, FORSCHUNGSBERICH
NR. 79/6.

Hannan, E. J. (1956). The asymptotic powers of certain tests based on multiple
correlations. Journal of the Royal Statistical Society (B)

99


