Likelihood Inference for Order Restricted Models

by

Afnan Almazmomi

A Thesis

Submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

In partial fulfillment of the requirements for the

Degree of Master of Science

in

Applied Statistics

by

May 2018

APPROVED:

Professor Thelge Buddika Peiris, Major Thesis Advisor

Abstract

As we know the most popular inference methods for order restricted model are likelihood inference. In such models, the Maximum Likelihood Estimation (MLE) and Likelihood Ratio Testing (LRT) appear some suspect behaviour and unsatisfactory. In this thesis, I review the articles that focused in the behaviour of the Likelihood methods on Order restricted models. For those situations, an alternative method is preferred. However, likelihood inference is satisfactory for simple order cone restriction. But it is unsatisfactory when the restrictions are of the tree order, umbrella order, star-shaped and stochastic order types.

Keywords: Preservation, Reversal, Convex cones, Projection, Likelihood Methodology, Cone order, Cone Order Monotonicity, Simple Order, Tree Order, Stochastic Order.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Professor Thelge Buddika Peiris for the continuous support of my study and related research and his great helped in research and writing of this thesis. Also, for his patience, motivation, and knowledge. I could not have imagined having a better advisor for my Master study.

Also, I would like to dedicate my thesis to my beloved parents Aaliah Almaghrabi and Ali Almazmomi, my husband Abdulaziz and my two little daughters Retal and Alin Almaghrabi for their infinitely love, and supporting me spiritually through my years of study and the process of researching and writing this thesis and my life in general. This accomplishment would not have been possible without them.

Contents

List of Figures

List of Tables

Chapter 1

Intorduction

Inference had been a significant hurdle in understanding the behaviour of a parameter or properties of a population, where analysing a drawing sample from a population to have an estimate or conclusions about the population properties or parameter. Generally, the likelihood methodology has been based inferences for order restricted models. The primary reference for this thesis is Cohen and Kemperman and Sackrowitz (2000). In most of the cases, the maximum likelihood estimation (MLE) is the primary approach used in parameter estimation and the likelihood ratio test (LRT) for hypotheses tests.

There are many research papers on Likelihood inference on Restricted models. Peiris and Bhattacharya (2016) discus the restricted inference of a regression model with two predictors when parameters β_1 and β_2 have sign constrains. Restricted inferences on regression with circular data have been discussed by Peiris and Kim (2016). In this paper they have used the technique invented by Mukerjee and Tu (1995) on two regression models called Circular-Linear regression and Linear-Circular regression. Moreover, Chaudhuri and Perlman (2003) investigates where the LRT is cone order monotonic (COM) or not, which is an important property for reverse the cone order or preserve it. Cohen et.al., (1995) considers the normal models where the alternative hypothesis test always on order restricted model, and provide conditions where the class of tests are completed and unbiased.

In this work likelihood methods have been tested on the cones for different order restricted models such as the simple order, umbrella, tree, star and stochastic orders. The purpose of the study is identifying situations where the likelihood estimation techniques can provide sufficient information and the advantages that could be received by it. Secondly, this study also focus on determining the drawback, or it is lack of effectiveness in other situations or modelling techniques. The implications of the practical application of likelihood testing method were, therefore, to be investigated in detail.

It is found through the overview and comprehensive understanding of the likelihood methodology in the detailed examples, that if the results of the method are not in favour of initially proposed or tested hypothesis, then it would be much better to go for an alternative option. The term reversal is therefore introduced in the paper which refers to going with the exact opposite after testing of any hypothetical situation. The results, however, revealed that there is need of some critical explanation.

By considering the constraints of the system the equation and inequalities are formed. The procedure is tested on two different kinds of violations observed. The first is the reverse of the order in which the inequalities completely fail under the given situation or parameters for testing. Secondly, the violation on estimator, which is less critical,is neither reversal but also neither preserves the order.

Such a behaviour is usually a consequence of restricting the parameter space. Also, unwanted behaviour of the likelihood inferences is possible in many order restricted models. In chapter 2, I present some basic and necessary mathematical and statistical knowledge. In addition, examples are given to shows what type of issue will arise. This thesis will cover a simple introduction to Order Restricted hypothesis tests in chapter 3, and the relation between projection and their properties with the notation of preservation and reserves at in chapter 4. The results of preservation and reserves at likelihood inferences are in chapter 5. Chapter 6 contains some applications of likelihood methods for order restricted models.

Chapter 2

Overview and Basics

Before embarking on developing formal discussion, an overview of some related basic and necessary mathematical and statistical knowledge is introduced here.

2.1 Basics

The following are some basic definitions of the terms in order restricted inference and Linear Algebra.

1. **Convex cone :**

   A convex cone is a subset of vectors $\zeta \in R^k$ that if $x, y \in \zeta$, than $\beta_1 x + \beta_2 y \in \zeta$ for all $\beta_1 \geq 0$, $\beta_2 \geq 0$. The closed of convex cone is induces pre-ordering $\geq_\zeta$ such that $x \leq_\zeta y$ if and only if $y - x \in \zeta$. A cone ζ pointed if $x \in \zeta$ and $-x \in \zeta$ implies $x = 0$.

2. **Dual of convex cone** :

   Consider the cone ζ is a closed convex cone $\zeta \subseteq R^k$, and $x, y \in \zeta$, then $\lambda_1 x + \lambda_2 y$ for all $\lambda_1 \geq 0n, \lambda_2 \geq 0$.

Then the positive dual of ζ is

$$\zeta^* = \{\theta \in R^k :\ \langle x, \theta \rangle \geq 0,\ \text{for all } x \in \zeta\}, \tag{2.1}$$

And the negative dual (polar dual) for the cone ζ is

$$\hat{\zeta} = \{\theta \in R^k :\ \langle x, \theta \rangle \leq 0,\ \text{for all } x \in \zeta\}, \tag{2.2}$$

where $\langle . , . \rangle$ is the inner product. For closed convex cone ζ the dual of the dual is $(\zeta^*)^* = \zeta$.

3. **Orthocomplement**:

   Let W be a subspace from R^k, then $W^\perp$ is the set of vectors which are orthogonal to all elements of W. That is, the inner product for all elements on W with $W^\perp$ equal zero.

$$< w, x >= 0,$$

   where $w \in W$ and $x \in W^\perp$.

   So

$$W^\perp = \{x \in R^k :< w, x >\}. \tag{2.3}$$

4. **Linear Span:**

   Leaner span is the set of all linear combinations of vectors in a vectors space. Let $\vec{S} \in R^k$ be a vector space, and $s_1, s_2, ...s_k \in \vec{S}$ then

$$span(\vec{S}) = \{\sum_{i=0}^{k} \lambda_i s_i \mid \lambda_i \in R\}. \tag{2.4}$$

5. **Polyhedral cone:**

A Polyhedron is an intersection of half-spaces in R^k it is generated by a set of vectors such as

$$\zeta = \{\theta \in R^k; \theta = \sum_{i=1} \lambda_i \theta_i \mid \lambda_i \geq 0\}, \tag{2.5}$$

Also a bounded polyhedron is a convex hull of a set of points.

6. **Types of Order Restricted Cones:**

- **Simple Order cone**: The simple order cone is one of the most common cone which had been used in different practices. Order restriction on the parameter space μ, where we test $\mu_1 \geq \mu_2 .... \geq \mu_k$ as the cone $\zeta = \{\mu \in R^k : \mu_1 \geq \mu_2 .... \geq \mu_k\}$. This can be expressed as matrix form

$$\zeta = \{\mu \in R^k : B\mu \geq 0\}, \tag{2.6}$$

where

$$B_{m \times k} = \begin{bmatrix} 1 & -1 & 0 & 0 & .... & 0 \\ 0 & 1 & -1 & 0 & .... & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & .... & 1 & -1 \end{bmatrix}.$$

- **Tree Order cone:** The tree order cone is another important cone for comparing difference treatments group to an control or stander group. Th cone for it is $\zeta = \{\mu : \mu_1 \geq \mu_k, \mu_2 \geq \mu_k .... \mu_{k-1} \geq \mu_k\}$,and as matrix form

$$\zeta = \{\mu \in R^k : B\mu \geq 0\}, \tag{2.7}$$

where

$$B_{m\times k} = \begin{bmatrix} 1 & 0 & 0 & ... & 0 & -1 \\ 0 & 1 & 0 & ... & 0 & -1 \\ \vdots \\ 0 & 0 & 0 & ... & 1 & -1 \end{bmatrix}.$$

- **Umbrella order Cone:** The umbrella order cone restricted the parameter space μ , where the test take the form $\mu_1 \le \mu_2 \le ... \le \mu_m \ge \mu_{m+1} \ge ... \ge \mu_k\}$. Then the cone$\zeta$ can be express as

$$\zeta = \{\mu \in R^k : B\mu \ge 0\}, \tag{2.8}$$

where

$$B_{m\times k} = \begin{bmatrix} -1 & 1 & 0 & 0 & ... & ... & 0 \\ 0 & -1 & 1 & ... & ... & 0 & 0 \\ \vdots \\ 0 & ... & -1 & 1 & ... & 0 & 0 \\ 0 & ... & 0 & 1 & -1 & ... & 0 \\ \vdots \\ 0 & 0 & 0 & ... & ... & 1 & -1 \end{bmatrix}.$$

- **Stochastic Order cone :** This type of cone consider a two independent random vectors which follow the multinomial distribution with (n_i, k, p_i) $\mathbf{X}_1$= $(X_{11}, X_{12}, ..., X_{1k})$ and $\mathbf{X}_2$= $(X_{21}, X_{22}, ..., X_{2k})$. Now our interest to test if the distribution of X_2 is stochastically greater than or equal to X_1

such as $\sum_{i=1}^{k-1} p_{1i} \leq \sum_{i=1}^{k-1} p_{2i}$. This can be expressed in matrix form

$$\zeta = \{\theta \in R^k : B\theta \geq 0\}, \tag{2.9}$$

where

$$\theta = \begin{pmatrix} p_{11}\ p_{12}\ \cdots\ p_{1(k-1)} \\ p_{21}\ p_{22}\ \cdots\ p_{2(k-1)} \end{pmatrix}, \tag{2.10}$$

and

$$B_{(k-1)\times 2(k-1)} = \begin{bmatrix} 1 & 0 & ... & 0 & -1 & ... & 0 \\ 1 & 1 & ... & 0 & -1 & -1... & 0 \\ \vdots & & & & & & \\ 1 & 1 & ... & 1 & -1 & ... & -1 \end{bmatrix}.$$

It is clear that ζ the parameter space is not a cone because the p_{ji}'s are bounded. See Cohen et.al.,(1998)

- **Star-Shaped Cone:** there are two type of cone under it, the lower star-shaped cone take the form $\mu : \mu_1 \geq \frac{(\mu_1+\mu_2)}{2} \geq .... \geq \frac{(\mu_1+\mu_2+....+\mu_k)}{k}$ and the upper star-shaped cone will be $\mu : \mu_1 \leq \frac{(\mu_1+\mu_2)}{2} \leq .... \leq \frac{(\mu_1+\mu_2+....+\mu_k)}{k}$. The polyhedral cone for the lower star-shaped will be

$$\zeta = \{\mu : B\mu \geq 0\}, \tag{2.11}$$

where

$$B_{(k-1)\times 2(k-1)} = \begin{bmatrix} 1 & -1 & ... & ... & 0 & 0 \\ 1 & 1 & -2 & ... & ... & 0 \\ \vdots & & & & & \\ 1 & 1 & ... & 1 & ... & -(k-1) \end{bmatrix}.$$

2.2 Likelihood Methodology

The maximum likelihood method is used in wide range of statistical analyses. Likelihood functions has a fundamental role in frequentest inference, specificity methods of parameter estimating.

2.2.1 Maximum likelihood estimator (MLE)

The most traditional and fundamental way of parameter estimations in statistics is maximum likelihood estimation. Let $X_1, X_2, ..., X_n$ be random sample from a population with density function $f_{X_i}(x|\theta)$. The likelihood function is the joint density function regarded as a function of the parameter θ. That is

$$L(\theta|X) = \prod_{i=1}^{n} f(X_i|\theta). \tag{2.12}$$

Then the maximum likelihood estimator (MLE) of θ is the value of θ which maximize the likelihood function . So

$$\hat{\theta}(X) \quad = \quad arg\max_{\theta}(L(\theta|X)), \tag{2.13}$$

Maximizing the log likelihood function instead of the likelihood function is a most common technique in maximum likelihood estimation.

2.2.2 Likelihood ratio test (LRT)

likelihood ratio test is widely applicable in hypothesis testing. Let Θ denote the entire parameter space and consider the hypotheses

$$H_0 : \theta \in \Theta_0 \qquad \text{vs.} \qquad H_1 : \theta \in {\Theta_0}^c. \tag{2.14}$$

Then the likelihood ratio test is

$$\lambda(x) = \frac{\sup_{\Theta_0} L(\theta|X)}{\sup_{\Theta} L(\theta|X)}, \tag{2.15}$$

with critical region $C = \{X : \lambda(x) \geq c\}$. The p-value of the test is

$$\text{p-value} = \sup_{\theta \in \Theta_0} P(\lambda(x) \geq c). \tag{2.16}$$

In most of the applications the asymptotic distribution of $-2\ln(\lambda(x))$ is used and

$$-2\ln\lambda(x) \sim \chi^2_d, \tag{2.17}$$

where χ^2_d the chi-square distribution with degree of freedom d and $d = dim(\Theta) - dim(\Theta_0)$.

2.3 Suspect Behaviour of Likelihood Methodology

In order restricted inference, most often, the likelihood methodology is used. That is, most often, the parameter estimation done by the maximum likelihood estimation and the hypothesis tests done by likelihood ratio test. In some situations, the likelihood methodology is less satisfactory or have suspect behaviours. Here I give two examples, one for maximum likelihood estimation and the other one for likelihood ratio test when they are unsatisfactory.

2.3.1 Example for hypothesis test:

This example has been illustrated in Cohen and Sackrowitz (1998). Considering an experiment to comparing between a treatment and a control. Conduct 2×3 table where the columns are the responses which are ordered and categorical. 2-sample points were considering which presented in table 2.1 and 2.2, where the second table (table 2.2) has been created from table 2.1 keeping the marginal totals fixed.

Table 2.1: sample point 1

	Same	Some improvement	Cured	Total
Control	5	11	1	17
Tretment	3	8	4	15
Total	8	19	5	

Table 2.2: sample point 2

	Same	Some improvement	Cured	Total
Control	0	16	1	17
Tretment	8	3	4	15
Total	8	19	5	

Assuming that the observation following an independent multinomial distribution with probabilities $P_c = (p_{1c}, p_{2c}, p_{3c})$ and $P_t = (p_{1t}, p_{2t}, p_{3t})$. The hypothesis test is

$$H_0 : \text{the two groups are the same} \quad \text{vs.} \quad H_1 : \text{ the treatment is batter.} \tag{2.18}$$

That is, the distribution of the treatment is stochastically larger than that of the control $\sum_{k=1}^{m} p_{kt} \geq \sum_{k=1}^{m} p_{kc}$. From these two sample point, we expect that the p-value for the sample point 1 from table 2.1 should be lower than that for the other sample point . But LRT statistic from table 2.1 is 2.777 (see chapter 3 section 3.1.2 for more details) and the condition p-value for fixed margins (for more details see Agresti and Coull (1998))is 0.169, and for the table 2.2 LRT statistic is 22.65 and the p-value is 0.019. This contradict what we expect and suggest that LRT may be less satisfactory for this problem. In this example we can say that The LRT give a reversal results.

2.3.2 Example for parameter estimation:

A company is interested to increase SAT scores of student and developed two programs for that. They want to know are the two programs effected and increase the SAT score. Three samples were taken with equal sample size for two treatment groups and a control group and determined the mean for each group $(\bar{X}_{T_1}, \bar{X}_{T_2}, \bar{X}_C)$. The hypothesis test

$$H_0 : \mu_{T_1} = \mu_{T_2} = \mu_C \quad \text{vs.} \quad H_1 : \mu_{T_1} \geq \mu_c \text{ and } \mu_{T_2} \geq \mu_C. \tag{2.19}$$

Sample means are assumed to have normal distributed with equal variance $\frac{\sigma^2}{n} = 144$. From these three samples, it is natural to expect that the p-value for S_1, p-value for

S_2 and p-value for S_3 are in increasing order. But the p-value of the LRT corresponding to S_1, S_2 and S_3 are 0.134,0.083 and 0.44 respectively, so only S_3 is significant. This is also contradict the naturally expect and show that likelihood ratio test is less satisfactory for this problem two. Furthermore, The MLE's for S_1, S_2, S_3 are present in table 2.4, are suggests that only S_3 has the stronger indication about the treatments than S_1 and S_2. This is also contradict the naturally expect and show that likelihood methodology is less satisfactory for this problem two.

Table 2.3: The mean average for each group

	$\bar{X}_{T_1}$	$\bar{X}_{T_2}$	$\bar{X}_C$
S_1	1124	1110	1096
S_2	1124	1096	1096
S_3	1124	1089	1096

Table 2.4: The MLE's for each sample

	$\hat{\mu}_{T_1}$	$\hat{\mu}_{T_2}$	$\hat{\mu}$	$\hat{\mu}_{T_1} - \hat{\mu}_C$	$\hat{\mu}_{T_2} - \hat{\mu}_C$
S_1	1124	1110	1096	28	14
S_2	1124	1096	1096	28	0
S_3	1124	1092.5	1092.5	31.5	0

Chapter 3

Order Restricted Hypothesis Test

Suppose our interest is testing means (or any location parameters) for different groups or treatments or populations. Generally, we have n independent samples from n different populations with means $\mu_1, \mu_2, ..., \mu_n$. Topically, the test

$$H_0 : \mu_1 = \mu_2 = ... = \mu_n \quad \text{vs.} \quad H_a : \text{not all equal ,} \tag{3.1}$$

The $F-$test in analysis of variance (ANOVA) and the nonparametric methods based on ranks are the stander methods. On the other hand, some times we may interest to test which treatment is better or worst than others. Consider the alternative hypotheses

$$H_a : \mu_1 \leq \mu_2 \leq \mu_3 \quad \text{or} \quad H_a : \mu_1 \leq \mu_2 \text{ and } \mu_1 \geq \mu_3, \tag{3.2}$$

as examples. Here we need to consider an alternative method for the analysis. In this chapter, I provide some general and basics approaches for these statistical inference problems. There are two main types of test problems. For the convenience they are called type A and type B problems. **Type A** problem has ordered alternative hypothesis. That is, this type has the order restricted parameter space for the alternative hypothesis. $H_0 : \theta = 0$ vs. $H_a : \theta \geq 0$, and $H_0 : B\theta = 0$ vs. $H_a : B_1\theta \geq 0$ where B_1 and B are matrices, are some examples for type A problems. **Type B** problems has ordered null hypothesis. Here the restriction are on the null hypothesis. For example , $H_0 : \theta \geq 0$ vs. $H_a : \theta \ngeq 0$, and $H_0 : \theta \geq 0$ and $\theta_b = 0$ vs. $H_a : \theta \in R^k$, where $\theta = (\theta_a, \theta_b)^T$. Also, combination of type A and type B problems exist. In this section, I focus on type A problems and their results.

3.1 Ordered Alternative Hypothesis

Here let's consider comparing k treatments and assume that we have k sample from those population with sample size n_i. Suppose observation $x_{ij} \sim N(\mu, \sigma^2)$ where $i = 1, ..., k$ and $j = 1, .., n_i$. Suppose we want to test

$$H_0 : \mu_1 = \mu_2 = ...\mu_k \qquad \text{vs.} \qquad H_1 : \mu_i - \mu_j \geq 0, \tag{3.3}$$

where $i, j = 1, 2, ..., k$. The standard F test statistics is

$$F = (k-1)^{-1}\{RSS(H_0) - RSS(H_a)\}/S^2, \tag{3.4}$$

where S^2 is the mean square error, and the Residual Sum of Squares under H_0 is

$$RSS(H_0) = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y})^2,$$

The MLE for μ_i is $\bar{y}_i$. Under 3.3 the standard F-test do not have a good power because it does not take the additional information with the restriction $\mu_i \geq \mu_j$ into account. An alternative test can be obtained from F-statistic

$$\bar{F} = \{RSS(H_0) - RSS(H_1)\}/S^2, \tag{3.5}$$

where $RSS(H_1) = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \tilde{\mu}_i)^2$, and the MLE's for μ under restriction is $\tilde{\mu}$, $S^2 = v^{-1}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2$, where $v = n_1 + n_2 + ... + n_k - k$. Here the p-value of $\bar{F}$ can be estimated by simulation. For more details See M. Silvapulle and P. Sen (2005).

On the other way, Testing the formula in 3.3 allows only pairwise contrasts, and can be rewriting the hypothesis in matrix form $H_0 : B\mu = 0$ Vs. $H_1 : B\mu \geq 0$ where $B_{m\times k}$ matrix such as

$$\begin{bmatrix} 1 & -1 & 0 & 0 & .... & 0 & \\ 0 & 1 & -1 & 0 & .... & 0 & \\ \vdots & & & & & & \\ 0 & 0 & 0 & 0 & .... & 1 & -1 \end{bmatrix}.$$

This case called simple order and it can be expressed as $\zeta = \{\mu : B\mu \geq 0\}$.

Let's $y_1, y_2, ...y_k$ follow density function

$$f(y_i, \theta_i, \phi_i) = \exp\{\frac{y_i\theta_i - b(\theta_i)}{a(\phi_i)} + c(y_i; \phi_i)\}, \tag{3.6}$$

where θ, ϕ are scalars, $a(.) \geq 0$ and $b(.), c(.)$ are some functions.

Then the loglikelihood function is

$$L(\theta, \phi) = \sum_{i=1}^{k} \frac{n_i\{\bar{y}_i\theta_i - b(\theta_i)\}}{a(\phi_i)} + H(y; \phi), \tag{3.7}$$

where H a function does not depend on θ. Let's $(\hat{\theta}, \hat{\phi})$ the unrestricted MLE and $(\tilde{\theta}, \tilde{\phi})$ denote the MLE under order restrict, then $\tilde{\theta} = \theta(\tilde{\mu})$. That is $\hat{\mu} = \bar{y}$ and $\tilde{\mu}$ is equal to the isotonic regression of $\hat{\mu}$, therefore $\tilde{\mu}$ can be obtained as follows

$$\tilde{\mu} = \min_{\mu B \geq 0} \sum (\bar{y_i} - \mu_i)^2 n_i. \tag{3.8}$$

3.1.1 Isotonic regression

Isotonic Regression is a way to fit a free-form line to observations under the condition of monotonic increasing. Further, the Isotonic Regression is illustrated the statistical inference on the means under order restricted. Let's $Y_1, Y_2, ...Y_k$ and consider the simple order cone in the alternative hypothesis $(\mu_1 \geq \mu_2 \geq ... \geq \mu_k)$, where $Y_i = \mu_i + \epsilon_i$, then it involves finding a weighted least-squares such as

$$\min_{(\mu_1 \geq \mu_2 \geq ... \geq \mu_k)} \sum_{i=1}^{k} w_i (Y_i - \hat{\mu_i})^2, \tag{3.9}$$

where $w_i \geq 0$ are giving weights. express in matrix form as

$$\min (Y - \hat{\mu})^T W (Y - \hat{\mu}), \tag{3.10}$$

where $Y = (Y_1, Y_2, .., Y_k)^T$, $\hat{\mu} = (\hat{\mu_1}, .., \hat{\mu_k})^T$ and $W = diag(w_1, w_2, .., w_k)$.
This problem can be solved by using the pool adjacent violators algorithm (PAVA). The figure 3.1 shows an example of comparing the isotonic regression to a linear regression, on the same dataset, where the isotonic regression is non-decreasing line and it is more flexible than the linear regression.

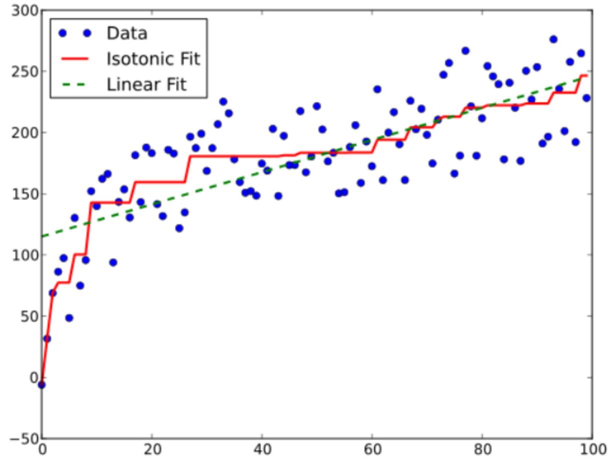

Figure 3.1: comparison of Isotonic regression and Linear regression

3.1.2 Examples in two dimension

Here I use an example illustrated in M. Silvapulle and P. Sen (2005). Let $X = (X_1, X_2)^T \sim N(\theta, I)$ and want to test of $H_0 : \theta = 0$ against $H_a : B\theta$ where $B_{2\times 2} = (1, 4|1, -2)$. Now let $\zeta = \{\theta : B\theta \geq 0\}$ is a closed convex cone , and $\hat{\zeta}$ be the polar cone of ζ as in (2.2). That is $\hat{\zeta} = \{\alpha : \langle\alpha, \theta\rangle \leq 0 \mid \text{where } \theta \in \zeta\}$.

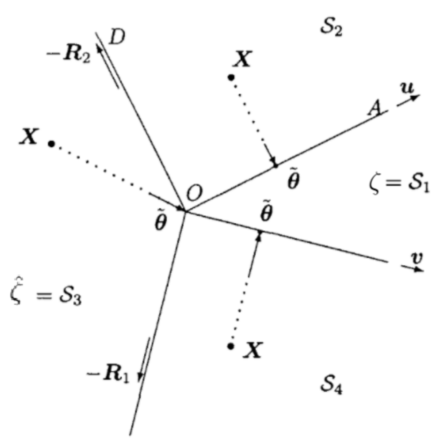

Figure 3.2: The MLE $\tilde{\theta}$ of θ where θ lie in ζ

From the figure 3.2 we can see that $\zeta = S_1$ and $\hat{\zeta} = S_3$. Now let u, v are the vectors parallel to the boundary of the cone ζ. Since the restricted MLE are the projections of X into cone ζ $[P(X|\zeta)]$,

$$\tilde{\theta} = \begin{cases} X & X \in S_1, \\ (u^T X)u & X \in S_2, \\ 0 & X \in S_3, \\ (v^T X)v & X \in S_4. \end{cases} \tag{3.11}$$

Then

$$LRT = ||X||^2 - ||X - \tilde{\theta}||^2 = ||\tilde{\theta}||^2, \tag{3.12}$$

under the null distribution of LRT

$$Pr(LRT \leq c) = \sum_{i=1}^{4} Pr(LRT \leq c|X \in S_i)Pr(X \in S_i). \tag{3.13}$$

Further

$$\begin{aligned} Pr(LRT \leq c \,|X \in S_1) &= Pr(X_1^2 + X_2^2 \leq c) \\ &= Pr(\chi_2^2 \leq c), \end{aligned}$$

where χ_2^2 is the chi-square distribution with degrees of freedom 2. Following the same argument, LRT can be obtained as

$$LRT = \begin{cases} X_1^2 + X_2^2 & X \in S_1 \sim \chi_2^2, \\ (u^T X)^2 & X \in S_2 \sim \chi_1^2, \\ 0 & X \in S_3 \sim \chi_0^2, \\ (v^T X)^2 & X \in S_4 \sim \chi_1^2. \end{cases} \tag{3.14}$$

$$Pr(LRT \leq c|H_0) = qPr(\chi_0^2 \leq c) + 0.5Pr(\chi_1^2 \leq c) + (0.5 - q)Pr(\chi_2^2 \leq c), \tag{3.15}$$

where

$$q = (2\pi)^{-1} \cos^{-1}[b_1^T b_2/\{(b_1^T b_1)(b_2^T b_2)\}^{1/2}],$$

and b_1, b_2 are the first tow rows of the matrix $B_{m\times k}$. The critical region$\{LRT \geq c\}$ is shows in figure 3.3.

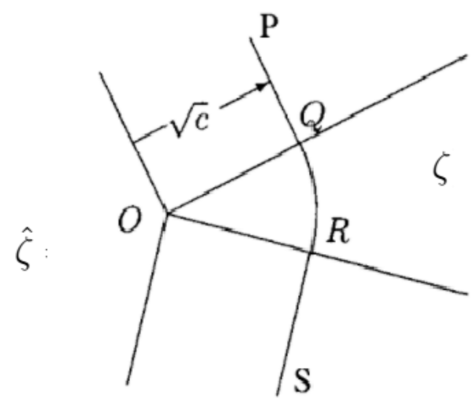

Figure 3.3: the critical region for LRT is to the upper-right region bounded by $PQRS$

As we can see from 3.14 the null distribution of LRT is a weighted chi-square distributions, which called chi-bar-square distribution $\bar{\chi}_v^2$ with degrres of freedom v . The main difficulty of using asymptotic tests based on the chi-bar squared is the computation of the weights associated with the various numbers of degrees of freedom. Analytical solutions are only available if the number of inequality restrictions is smaller than 5. However, several methods have been developed to approximate the weights of the chi-bar-squared distribution. Rather than using an asymptotic approach to obtain the p value associated with the LRT statistic, it can also be estimated using parametric bootstrapping. It is well-known that bootstrapping methods can be used to obtain an empirical approximation of the distribution of a test statistic when its asymptotic distribution is complicated or unknown.

3.1.3 LRT for Type A Problems

In the previous section, for 2-dimensional cases were introduced and example were provided. Some higher dimensional example are introduced here. Suppose $\mathbf{X} \sim$ $\mathrm{N}(\theta, V)$ where V is known and the test is

$$H_0 : B\theta = 0 \qquad \text{vs.} \qquad H_a : B\theta \geq 0 \ \ , \tag{3.16}$$

where B is a matrix of the order $r \times p$, and the chi-bar-square distribution $\bar{\chi}^2$ statistic is giving by

$$LRT_A = X^T V^{-1} X - \min_{\theta \in \zeta} (X - \theta)^T V^{-1} (X - \theta). \tag{3.17}$$

Then the null distribution is give by

$$\mathrm{pr}(LRT_A \leq c | H_0) = \sum_{i=0}^{r} (r, BVB^T) \mathrm{pr}(\chi_i)^2 \leq c). \tag{3.18}$$

Then the p-value (critical value) of LRT_A with t is the observed value of LRT_A

$$\text{p-value} = \text{pr}(LRT_A \geq t \mid H_0) = \text{pr}(LRT_A \geq t \mid \theta = 0) \tag{3.19}$$

Chapter 4

Preservation Property of Projections

4.1 Properties of Projection

This section provide some important properties of projections which effect the reverse or preserver property.

Let ζ to be a close convex cone and a subset of R^k where $x, y \in \zeta$, then $\lambda_1 x + \lambda_2 y \in \zeta$ for all $\lambda_1, \lambda_2 \geq 0$. A preordering $\leq_\zeta$ in the cone ζ define as: $x \leq_\zeta y$ if and only if $x - y \in \zeta$. A cone ζ to be a pointed cone if $x \in \zeta$ and $-x \in \zeta$. Another important property is Cone Order Monotone with respect to the cone ζ (COM[ζ]), let function $F(x) : R^k \to R$ if whenever $x \leq_\zeta y$ then $F(x) \leq F(y)$, see Robertson and Wegman (1978).

The positive dual ζ^* of a convex closed cone as 2.1 , and let's Ω be the linear span of ζ^*. Where H is the orthocomplement of Ω in 2.3. As result, an $x \in R^k$ can be writing as $x = x_\Omega + x_H$, where that mean x_H is the projection of x into H and x_Ω is the projection of x into Ω.

Now define a cone ℓ as

$$\ell = \zeta \cap \Omega, \tag{4.1}$$

because Ω is liner subspace and ζ is a closed convex cone that imply ℓ is a closed convex cone. Additionally, where $H \subset \zeta$, $x \in H$ implies $x \in (\zeta^*)^* = \zeta$, and ℓ is pointed closed convex cone .

Lemma 4.1: $\zeta = \ell \oplus H$, every vector in ζ can be uniquely writing as a liner combinations of vectors from ℓ and H.

Lemma 4.2: $\zeta \subseteq \zeta^* \oplus H$ if and only if $\ell \subseteq \zeta^*$.

Next defined $P(x|\zeta)$ as the unique point of projection x in to ζ in R^k. Rewriting the theorem 8.2.7 of RWD(1988)as:

Lemma 4.3 : let $x, u \in R^k$, and ζ a closed convex cone in R^k. then $u = P(x|\zeta)$ if and only if $u \in \zeta$ and

$$\langle x - u, u \rangle = 0, \tag{4.2}$$

and

$$\langle x - u, f \rangle \leq 0, \text{ for all} f \in \zeta.$$

Using Lemma 4.1 we get

$$P(x|\zeta) = P(x_\Omega|\ell) + x_H, \tag{4.3}$$

where

$$x_\Omega \leq_\ell y_\Omega, \text{ if and only if } \; x \leq_\zeta y, \tag{4.4}$$

from 4.3 and 4.4 we get

$$P(x_\Omega|\ell) \leq_\ell P(y_\Omega|\ell), \text{ if and only if } P(x|\zeta) \leq_\zeta P(y|\zeta). \tag{4.5}$$

At this point we can defined the property of preservation and reversal for the projections.

Definition 4.1 : For any pair $x, y \in R^k$ such that $x \leq_\zeta y$, there is $P(x|\zeta) \leq_\zeta P(y|\zeta)$. Then $P(.|\zeta)$ have preservation property w.r.t cone ζ if and only if $P(.|\ell)$ has the preservation property w.r.t cone ℓ.

4.2 Result on Projections

Definition 4.2 : if exist $x, y \in R^k$ such that $x \leq_\zeta y$ and $P(x|\zeta) \geq_\zeta P(y|\zeta)$ the projection $P(.|\zeta)$ have a reverse property w.r.t cone ζ .

Theorem 4.1 : The projection P(.|ζ) has reversal property w.r.t cone ζ if and only if $\ell \not\subseteq \zeta^*$.

Considering the definition of a polyhedral cone in 2.5 can be as :

$$\zeta = \{\theta \in R^k : \langle b_i, \theta\rangle \geq 0, i = 1, 2..., m\} = \{\theta \in R^k : B\theta \geq 0\}, \tag{4.6}$$

where b_i are the rows of the matrix $B_{m\times k}$, And the b_i's are the generators of ζ^*, such as $\zeta^* = \{x \in R^k : x = \sum_{i=1}{}^m \lambda_i b_i | \lambda_i \geq 0\}$.

Theorem 4.2 : Let ζ be a closed convex cone in R^k , and suppose the projection $P(.|\zeta)$ has the preservation property w.r.t cone ζ, then ζ is a polyhedral cone.

This theorem shows that the projection has the preservation property only when ζ is a polyhedral cone. The following theorem shows when ζ a polyhedral cone should have a preservation property.

Theorem 4.3 : A projection $P(.|\zeta)$ has a preservation property w.r.t ζ , where ζ a polyhedral cone, if and only

$$\langle b_i, b_j \rangle \leq 0 \ ; \ i \neq j, \text{ where } i, j = 1, 2, 3, ..., m. \tag{4.7}$$

From definition 4.1 that projection on the cone ζ have preservation property if and only if $P(.|\ell)$ has the preservation property w.r.t cone ℓ. And following 4.1, since ζ a polyhedral cone as 4.6, also ℓ is a pointed polyhedral cone. Let a set of non-redundant generator of ℓ knows as $\{a_1, a_2, ..., a_p\}$ where($a_j \in \ell \subset \Omega$),

$$\langle a_i, a_j \rangle \geq 0 \ ; \ \text{where } i, j = 1, 2, ..., p. \tag{4.8}$$

Following theorem 4.1 and the property 4.8 hold that imply that $\ell \subseteq \zeta^*$, then the projections $P(.|\zeta)$ does not have a reverse property. Furthermore, from theorem 4.3 , if the inner product of all the generators of the cone ζ^* lower or equal zero 4.7 hold that implies the generator of the cone ℓ will be larger than zero 4.8, and we can say that the projections have a preservation property. However,the other way if 4.8 hold not implies 4.7. On other words, that the preservation 4.7 as strong preservation but 4.8 the shorting of reverse as weak preservation.

A very helpful concept related to preservation property is Linear independents of the rows of $B_{m\times k}$ which are the generators of the dual cone ζ^* .

Theorem 4.4 : let b_i's any set of vectors and $\langle b_i, b_j \rangle \leq 0$, for all $i \neq j$, where $i, j = 1, 2, ..., m$, in addition any of the condition below hold. if b_i's are the generators of pointed polyhedral cone. Or the exists of w such that $\langle b_i, w \rangle \geq 0$,all $i = 1, 2, ..., m$. If all b_i's in the same open half space. Then b_i's are linearly independent.

This theorem help to define which cones may have a preservation property such as contrast cones with $m = (k-1)$ row and linearly independent. Under order restricted model most comment cones used are pairwise contrast cone, which mean the cone in the forms 4.6 with $\left\langle \vec{1}, b_i \right\rangle = 0$ and only have two element non-zero in each rows and columns , and there is a few cones of this class which there projection $P(.|\zeta)$ have the preservation property. this led to the following corollary.

Corollary 4.1 : where $\zeta = \ell \oplus H$ is a pairwise contrast cone and H is one dimensional. if $P(.|\zeta)$ have preserve the order w.r.t cone ζ, then ζ is a simple order cone, which is the only cone who $P(.|\zeta)$ preserves order, $\zeta = \{\mu \in R^k : \mu_{j1} \geq \mu_{j2} \geq .... \geq \mu_{jk}\}$ for the arrangement $(j_1, j_2, ..., j_k)$.

That mean when the cone ζ is a pairwise contrast cone with H one dimensional, which is the orthocomplement of the span of the linear independent b_i's, the only cone with preserves order is the simple order cone.

Chapter 5

Likelihood Inferences and Preservations Property

The reverse and preserves property on projection are related to the statistical models special under order restricted model where the MLE's are the projection. Now let's $X_1, X_2, ..., X_k$ are independent variable with X_i have density function $f(x|\theta)$. Consider the exponential family, where the parameter space of μ is the common order restricted models.

$$f_{X_i}(x_i|\mu_i) = \exp[q_1(\mu_i)x_i + q_2(\mu_i)], \tag{5.1}$$

where μ_i is the most popular order restrict parameter space. The linear restriction on the mean μ take the form

$$\langle b_i, \mu\rangle \geq 0 \text{ , where } b_i \in R^k \text{ and } i = 1, 2, ..., m. \tag{5.2}$$

Let $y \geq x$ with the means $\bar{y}, \bar{x}$ such that $\langle b_i, \bar{y}\rangle \geq \langle b_i, \bar{x}\rangle$ for all $i = 1, ..., m$. Under this expression it can be said it is preserved order.

In terms of μ_i the cone ζ is a pairwise contrast cone then the projections $P(.|\zeta)$ are the MLE's for μ. See theorem (1.5.2) in Robertson et.al., (RWD) (1988) . That means the result on projections, if it having the requirements for preservation property or reversal will apply to the MLE's too.

Another concern is arise when the LRT for the hypothesis test $H_0 : B\mu$ vs. $H_1 : B\mu \geq 0$ where B as the forms in 4.6 is not $COM[\ell]$. In this case a lack of $COM[\ell]$ for the LRT may leed to unsatisfied or undesirable. Such an example present in the introduction. As a result, it is clear that there is an association between the reversal property and lack of $COM[\ell]$.

Now consider ζ as in 4.6 and the projection $P(X|\zeta)$ is the MLE $M(X)$ of μ, thus $P(X|\zeta) = M(X)$. And $B_{(k-1)\times k}$ is a pairwise contrast matrix, Where Tasting the hypothesis

$$H_0 : B\mu = 0 \qquad \textbf{vs.} \qquad H_1 - H_0 \qquad , \textit{where } H_1 : B\mu \geq 0. \tag{5.3}$$

Here note that, under H_0, the set of μ's satisfy is equal to the orthocomplement H, where H in this case is a linear combination of the vector $\mathbf{1}= (1, 1, ..., 1)$. Next assume that the LRT for testing $B\mu = 0$ vs. $B\mu \geq 0$ has $T = \sum X_i = t$ a fixed value which is a unique minimum sample points at $x = (\bar{x}\mathbf{1})$.

Remark: If the theorem 4.2 or $\ell \nsubseteq \zeta^*$ are not hold, that imply almost of the sample space will have a reverse property.

Which means there are a pair $x_i \leq_\zeta x_j$ when $i \neq j$, such as $M(x_i) >_\zeta M(x_j)$. Moreover, the LRT statistic at the point x_i and $M(x_i)$ will be equal or larger than the LRT statistic at the point x_j and $M(x_j)$, this is implies that the LRT may not be a COM[ℓ].

Theorem 5.1: Suppose the cone ζ as in 4.6 and $B_{(k-1)\times k}$ is pairwise contrast.

Consider that MLE has a reverse property. And the LRT for the test in 5.3 is not a COM[ℓ] when for each $T = \sum X_i = t$ the LRT for the test $B\mu = 0$ vs. $B\mu \neq 0$ has a unique minimum point through all sample points at $x = (\bar{x}\mathbf{1})$.

Theorem 5.2: Considering the exponential family density function in 5.1 , when $q_i(\mu_i)$ is increasing function of μ then the cone ζ in 4.6 is a pairwise contrast cone. If $\ell \subseteq \zeta^*$ then any test is COM $[\zeta^*]$ will be a COM [ℓ] also.

In this thesis I have provided a comprehensive understanding of the proper identification of situation where the likelihood ratio testing and maximum likelihood estimation techniques would be usable. Also the situations where the estimation tools can reveal some useful results can also provide the strong intuitions and reduce the ambiguities to reduce the uncertainty to maximum level to go with the most appropriate options.

Chapter 6

Applications

This chapter includes applications to the most frequently used models of order restricted inference. Here applying the result and the theorem of the last two chapter to different cones such as simple order cone, tree order cone, umbrella order cone and stochastic order cone.

Simple order cone:

Simple order cone is the most common cone us in many different practise. Now consider the mean parameters of the model 5.1 and the polyhedral cone $\zeta = \{\mu \in R^k : \mu_1 \geq \mu_2 \geq \mu_3 \geq \mu_4\}$, where the sample size are equal. The MLE $M(X)$ of μ are $P(X|\zeta)$. As an example, here let $k = 4$ so $\mu = (\mu_1\ \mu_2\ \mu_3\ \mu_4)'$. Then the condition 4.7 and 4.8 are applied for this cone model, where

$$B_{3\times 4} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

The inner product for the B matrix is $< b_i, b_j >= -2 \leq 0$ where $i \neq j = 1, 2, 3, 4$. Since 4.8 is hold that means that $\ell \subseteq \zeta^*$, which implies that theorem 5.2, the LRT is COM[ℓ]. As a result, the simple order cone have the preservation property for MLE and that implies that the LRT is COM[ℓ]. Generally, for any number of k the simple order cone as 2.6 always have a preservation property.

Tree order cone:

Tree order cone is one of the basic models used in comparing treatments with control. Let Assume that $X's$ follows the exponential family as in 5.1 and with equal sample size, and the cone is $\zeta = \{\mu : \mu_1 \geq \mu_4, \mu_2 \geq \mu_4, \mu_3 \geq \mu_4\}$. Now let $k = 4$, and apply the condition 4.7 and 4.8

$$B_{3\times 4} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

The inner product for the all rows of B matrix is $< b_i, b_j >= 1 \geq 0$ where $i \neq j = 1, 2, 3, 4$, which imply that the condition 4.7 of preservation property not hold , and by theorem 4.1 that $\ell \not\subseteq \zeta^* \Longrightarrow \zeta \not\subseteq H \oplus \zeta^*$ that means that the MLE have reversal property. Also, LRT is not COM[ℓ] by theorem 5.1 where the statistic of LRT minimized when $t = \sum x_i$. Here the tree order cone always have reversal property because for all the rows of the matrix B, the inner products are large than zero and that violate the condition of preserving property.

Umbrella order cone:

As in the previous case I assumed the family model as in 5.1 with equal sample size. Let $k = 5$ where the cone $\zeta = \{\mu_1 \leq \mu_2 \leq \mu_3 \geq \mu_4 \geq \mu_5\}$.

$$B_{4\times 5} = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}.$$

now applying the condition of preservation property 4.7 The inner product of B matrix is $< b_2, b_3 >= 1 \geq 0$ where all other rows are less than zero which led to $\ell \subseteq \zeta^*$ does not hold. That mean reversal property for MLE's occur and LRT is not COM[ℓ]. See Cohen and Sackrowitz (1996b). For any number of k the umbrella order as 2.8 will reverse the MLE's and LRT not COM[ℓ], because there exists $< b_m - 1, b_m >= 1 \geq 0$ which contradict with proservation condition.

Stochastic order cone:

under the same assumption as previous cases, and the conditional parameter spas as 2.9. Here, as it clear, the parameter space is not a cone that means the MLE are not the projections, which imply that the condition of preservation and reversal can not be applied. Since we can not prove that MLE has preservation property we can focus on LRT is COM[ℓ]or not. First, find the generators of the cone ℓ. Assume that $\zeta = \{\mu : B\mu \geq 0\}$ where the matrix $B_{(k-1)\times k}$, and the row of the matrix B_H

are the basis of the Orthocomplement H of Ω. The matrix is

$$B_H^{(k-1)\times 2(k-1)} = \begin{bmatrix} 1 & 0 & 0 & ... & 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 & 0 & 1 & ... & 0 \\ \vdots & & & & & & & \\ 0 & 0 & ... & 1 & 0 & 0 & ... & 1 \end{bmatrix},$$

where the generator of the cone ℓ is the inverse of the matrix $\begin{pmatrix} B \\ B_H \end{pmatrix}$. See Cohen and Sackrowitz (1998), which COM[ℓ] called concordant monotone, and it is not easy to prove. As it shows previously in the example 2.3.1 the condition p-value for the LRT of the two sample point are not COM[ℓ] although they are COM[ζ^*], so a recommend a test with COM[ℓ] is the Wilcoxon Mann Whitney [WMW] test where is a conditionally with large value. A better test recommended is the COM[ℓ] Fisher test is concordant monotone, where it is sensitive to most stochastic order cone, similar power of the conditional p-value properties where WMW are satisfying and superior.

Unequal sample size:

Here we presenting the same result on preservation property under different samples size. Suppose we have four different random samples ($k = 4$) with samples size are $n = (20, 10, 25, 8)$, the MLE's of μ_i is $\bar{X}_i = \sum_{i=1}^{n_i} X_{ji}$ where $j = 1, 2, 3, 4$. Then the MLE's are the weighted projection of $\bar{X} = (\bar{X}_1, \bar{X}_2, \bar{X}_3, \bar{X}_4)$ on to ζ. Here treat $\bar{X}$ as normally distributed with mean vector μ and covariance matrix $(D^2)^{-1}$ to help find the projection. Now under the consider $Y = D\bar{X}$ where $Y \sim N(D\mu, I)$, also consider the polyhedral cone

$$\zeta_Y = \{\mu : BD^{-1}\mu \geq 0\}, \tag{6.1}$$

where D is the diagonal matrix of covariance matrix with $\sqrt{ni}$. Which implies the MLE's of $D\mu$ is $P(D\bar{X}|\zeta_Y)$

$$M(\bar{X}) = D^{-1}P(D\bar{X}|\zeta_Y). \tag{6.2}$$

Now let's apply the above result on the different cone order.

1. **simple order cone:** Under the assumption of the model 5.1 and the polyhedral cone in 6.1 . The matrix $B_y = BD^{-1}$

$$B_y^{3\times4} = \begin{bmatrix} \frac{1}{\sqrt{n_1}} & \frac{-1}{\sqrt{n_2}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{n_2}} & \frac{-1}{\sqrt{n_3}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{n_3}} & \frac{-1}{\sqrt{n_4}} \end{bmatrix} = \begin{bmatrix} 0.2236068 & -0.3162278 & 0.000 & 0.000 \\ 0.000 & 0.3162278 & -0.200 & 0.000 \\ 0.000 & 0.000 & 0.200 & -0.3535534 \end{bmatrix}.$$

   Then B_y is satisfy the conditions of Theorem 4.3, that preservation property of MLE hold and that led the LRT are COM[ℓ]. In fact having different sample sizes do not effect preservation property in simple order cone.

2. **Tree order cone:** under the same assumption, from 2.7 and 6.1 we get

$$B_y^{3\times4} = \begin{bmatrix} \frac{1}{\sqrt{n_1}} & 0 & 0\frac{-1}{\sqrt{n_2}} & \\ 0 & \frac{1}{\sqrt{n_2}} & 0 & \frac{-1}{\sqrt{n_3}} \\ 0 & 0 & \frac{1}{\sqrt{n_3}} & \frac{-1}{\sqrt{n_4}} \end{bmatrix} = \begin{bmatrix} 0.2236068 & 0.0000000 & 0.00 & -0.3535534 \\ 0.000 & 0.3162278 & 0.00 & -0.3535534 \\ 0.000 & 0.000 & 0.200 & -0.3535534 \end{bmatrix}.$$

   So in the light of theorem 4.3 the preservation property for MLE's not hold which imply that LRT are not COM[ℓ]

3. **Umbrella order cone:** under the same assumptions, and from 2.8 and 6.1. Where the alternative test is $\mu_1 \leq \mu_2 \leq \mu_3 \geq \mu_4$, then the matrix B_y given by

$$B_y^{3\times 4} = \begin{bmatrix} \frac{-1}{\sqrt{n_1}} & \frac{1}{\sqrt{n_2}} & 0 & 0 \\ 0 & \frac{-1}{\sqrt{n_2}} & \frac{1}{\sqrt{n_3}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{n_3}} & \frac{-1}{\sqrt{n_4}} \end{bmatrix}.$$

Under the condition 4.7 and 4,8 it is clear that preservation property not hold for any set of n_i's which imply that LRT not COM[ℓ].

The main conclusions that can be deduced from the study are follows. The maximum likelihood estimation technique would always preserve its property under the simple order cone (or simple order restrictions) and it is only for simple order cones. For the tree or umbrella order cones, the likelihood inference technique reverse the order. But for the other type of cones such as matrix and star-shaped the LRT might or might not preserve (or reverse), relying on the size of the samples and the corresponding distribution. When the likelihood methodology is less satisfactory, alternative methods can be used. Cone ordered monotonic Fisher test and Wilcoxon Mann Whitney (WMW) test (see Cohen and Sackrowitz, 1998) are some options. When the restriction is simple tree order, under some conditions, transforming polyhedral cone to a circular cone is also an option (see Pincus, 1975).

Bibliography

[1] A. Agresti and B. Coull. Order restricted inference using odds ratios for monotone trend alternatives in contingency tables. *Comput. Statist. Data Anal. 28* , 139-155, (1998).

[2] A. Cohen, H. Kemperman and H. Sackrowitz. Properties of Likelihood inference for Order Restricted Models. *Journal of Multivariate Analysis* 72,50-77, (2000).

[3] A. Cohen, H. Sackrowitz, and E. Samuel-Cahn, Constructing tests for normal order restricted inference, *Bernoulli 1* , 321-333,(1995).

[4] A. Cohen and H. Sackrowitz. Tests for the umbrella alternative under normality *,Comm. Statist. 25* . 2807-2818, (1996b).

[5] A. Cohen and H. Sackrowitz, Directional tests for one sided alternatives in multivariate models, *Ann. Statist. 26*, 2321-2338, (1998).

[6] M. Silvapulle and P. Sen, Constrained Statistical Inference, *Ann. Statist. 26*, 2321-2338, (1998).

[7] T. Peiris and B. Bhattacharya, Constrained inference in linear regression, *Journal of Multivariate Analysis*, Vol. 151, (2016).

[8] T. Peiris and S. Kim, Restricted Inference in Circular-Linear and Linear-Circular Regression, *Sri Lankan Journal of Applied Statistics* 17 (1), (2016).

[9] S. Chaudhuri and M. Perlman, The Role of Reversals in Order Restricted Inference,*Technical report No.* 4334, (2003).

[10] T. Robertson, F. Wright and R. Dykstra, Order Restricted Statistical Inference, *Wiley, New York*, (1988).

[11] R. Pincus, Testing linear hypotheses under restricted alternatives, *Math. Operationsforsch. u. Statist.*, 6, 733-751, (1975)

Appendix A

R code for the property of preserving

```
k<- sample(x=c(1:20))[1]      #no. of column

#1- simple order cone
d <- matrix(0,(k-1),k)
for (i in 1:(k-1)){
  d[i,i]<-1
  d[i,i+1]=-1
}
d
# presrervation and revarsal property
Q<-rep(NA,time=k-1)
for(i in 1:(k-1)){
  for (j in 1:(k-2)) {
    if (i != j)
    v<-(d[i,] %*% d[j,])
  }
   Q[i]<-v
```

```
}
B<-matrix(Q,nrow=k-2,ncol=1,byrow=TRUE)
B
#other ways to proof preservation


dt=t(d)
t=d %*% dt
t
#-------------------------------------------------------------------------
#2 tree order cone


d <- matrix(0,(k-1),k)
for (i in 1:(k-1)){
  d[i,i]<-1
  d[i,k]=-1
}
d
# presrervation and revarsal property
Q<-rep(NA,time=k-1)
for(i in 1:(k-1)){
  for (j in 1:(k-1)) {
    if (i != j)
      v<-(d[i,] %*% d[j,])
  }
  Q[i]<-v
}
```

```
B<-matrix(Q,nrow=k-1,ncol=1,byrow=TRUE)
B
#other ways

dt=t(d)
t=d %*% dt
t

#-----------------------------------------------------------------------

#3 Umbrella cone:

m<-sample(x=c(1:k))[1]
d <- matrix(0,(k-1),k)
for (i in 1:(m)){
  d[i,i]<-1
  d[i,i+1]=-1
  }
for(i in (m+1):(k-1))
{
  d[i,i]=-1
  d[i,i+1]<-1
}
d
# presrervation and revarsal property
Q<-rep(NA,time=k-1)
```

```
for(i in 1:(k-1)){
  for (j in 1:(k-1)) {
    if (i != j)
      v<-(d[i,] %*% d[j,])
    }
  Q[i]<-v
}
v=(d[m,] %*% d[m+1,])
v
B<-matrix(Q,nrow=k-1,ncol=1,byrow=TRUE)
B
#other ways

dt=t(d)
t=d %*% dt
t
#-------------------------------------------------------------
# 4. Stochastic Order cone:

d <- matrix(0,(k-1),2*(k-1))

for(i in 1:(k-1)){
  for(r in 1:i){
    d[i,r]<-1
  }
```

```
}
d
for(i in 1:(k-1)){
  for(r in k:(i+(k-1))){

    d[i,r]=-1
  }
}
d
#-----------------------------------------------------------
#4. Star Shaped cone:

d <- matrix(0,(k-1),k)
for (i in 1:(k-1)){
  for (j in 1:(k)) {
    j=i+1
    d[i,j]=-(i)
  }
}
d
for(i in 1:(k-1)){
  for(r in 1:i){
    d[i,r]<-1
  }
}
d
```

```
###-------------------------------------##

## unequal sample sizes:

k<-4      # no. for row or column
n<-c(20,10,25,8) # different samples size

#1- simple order :

d <- matrix(0,(k-1),k)
for (i in 1:(k-1)){
  d[i,i]<-1/sqrt(n[i])
  d[i,i+1]=-1/sqrt(n[i+1])
}
d
Q<-rep(NA,time=k-1)
for(i in 1:(k-1)){
  for (j in 1:(k-1)) {
    if (i != j)
      v<-(d[i,] %*% d[j,])

        }
  Q[i]<-v
}
v=(d[m,] %*% d[m+1,])
v
```

```
B<-matrix(Q,nrow=k-1,ncol=1,byrow=TRUE)
B
#other ways

dt=t(d)
t=d %*% dt
t
##-------------------------------------------------------------------
#2 tree order cone

d <- matrix(0,(k-1),k)
for (i in 1:(k-1)){
  d[i,i]<-1/sqrt(n[i])
  d[i,k]=-1/sqrt(n[k])
}
d
Q<-rep(NA,time=k-1)
for(i in 1:(k-1)){
  for (j in 1:(k-1)) {
    if (i != j)
      v<-(d[i,] %*% d[j,])
  }
  Q[i]<-v
}
B<-matrix(Q,nrow=k-1,ncol=1,byrow=TRUE)
B
```

```
#other ways


dt=t(d)
t=d %*% dt
t
#--------------------------------------------------------------------
#3 Umbrella cone:
m<-sample(x=c(1:k))[1]
m=3
n<-rep(NA,time=k)


for (i in 1:k){
  prefs <- sample(x=c(1:100))
  n[i] <- prefs[1]
}
n=c(15,24,30,13,18)
k=4
m=3
d <- matrix(0,(k-1),k)
for (i in 1:(m)){
  d[i,i]<-1/sqrt(n[i])
  d[i,i+1]=-1/sqrt(n[i+1])
}
for(i in (m+1):(k-1))
{
  d[i,i]=-1/sqrt(n[i])
```

```
    d[i,i+1]=1/sqrt(n[i+1])
}
d

Q<-rep(NA,time=k-1)

for(i in 1:(k-1)){
    for (j in 1:(k-1)) {
        if (i != j)
            v <-(d[i,] %*% d[j,])
    }
    Q[i]<-v
}
B<-matrix(Q,nrow=k-1,ncol=1,byrow=TRUE)
B
v=(d[m,] %*% d[m+1,])
v

#other ways

dt=t(d)
t= d %*% dt
t
```