

Isotonic Regression: Theory and Empirical Applications  
for Psychological Studies using Intrinsic Religious Orientation

by

John Grivell

University of Manitoba

A thesis submitted to the University of Manitoba in partial  
fulfillment of the requirements for the degree of Master of Arts

**THE UNIVERSITY OF MANITOBA**  
**FACULTY OF GRADUATE STUDIES**  
\*\*\*\*\*  
**COPYRIGHT PERMISSION**

**Isotonic Regression: Theory and Empirical Applications**  
**for Psychological Studies using Intrinsic Religious Orientation**

**BY**

**John Grivell**

**A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of**  
**Manitoba in partial fulfillment of the requirement of the degree**  
**Of**  
**MASTER OF ARTS**

**John Grivell © 2004**

**Permission has been granted to the Library of the University of Manitoba to lend or sell copies of this thesis/practicum, to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film, and to University Microfilms Inc. to publish an abstract of this thesis/practicum.**

**This reproduction or copy of this thesis has been made available by authority of the copyright owner solely for the purpose of private study and research, and may only be reproduced and copied as permitted by copyright laws or with express written authorization from the copyright owner.**

## Table of Contents

	<u>Page</u>
Abstract	1
Chapter One: Background of the Study	
1.1 Introduction	2
1.2 What is isotonic regression?	5
1.3 Statement of Purpose	8
1.4 Overview	9
Chapter Two: Hypothesis Testing in Isotonic Regression	
2.1 The Null and Alternative Hypotheses	12
2.2 The Pool-Adjacent-Violators Algorithm	13
2.3 Test Statistics and Critical Values	16
2.4 Illustrative Examples	17
Chapter Three: Goodness of Fit of Isotonic Regression Models	
3.1 The Role of Goodness of Fit in Model Assessment	22
3.2 Methods for Goodness of Fit Assessment	25
3.3 Using Isotonic Regression for Linear and Non-Linear Trends	28
3.4 An Illustrative Example	31
Chapter Four: Isotonic Regression, ANOVA and Non-Linear Regression	
4.1 Purpose of the Study	37
4.2 Methodology	38

4.3 Procedure and Design	41
4.4 Results for the ANOVA and Isotonic Regression Analyses	45
4.5 Results for the ANOVA and Isotonic Regression Goodness of Fit Indices	53
Chapter Five: Advantages and Limitations of Isotonic Regression	
5.1 The Advantages of Isotonic Regression	56
5.2 The Disadvantages of Isotonic Regression	58
5.3 Summary	60
References	61
Appendix	69
Glossary	73

## List of Tables

	<u>Page</u>
Table 1.1: Warm-up Times for a Three Level Factor of Tube Type	2
Table 2.1: Sample Mean Estimates become Isotonic Regression Estimates	14
Table 2.2: Blocks of Sample Mean Estimates become Isotonic Regression Estimates	15
Table 2.3: Likelihood Ratio Test Analysis of Mean Monthly Temperatures for the Palmero, Italy	17
Table 2.4 Critical Values for the $\chi^2$ Test ( $H_1: \mu_1 \leq \mu_2 \leq \dots \leq \mu_k$ or $\mu_1 \geq \mu_2 \geq \dots \geq \mu_k$ )	18
Table 2.5 Time to Complete the Test is Measured for Length of Test	20
Table 3.1 Vigilance Scores are Measured for Hours without Sleep	33
Table 3.2 Point-biserial Correlation Matrix of Vigilance and Hours without Sleep	33
Table 3.3 The Isotonic Regression Models for Vigilance Regressed on Hours of Sleep Deprivation	34
Table 4.1 Descriptive Statistics for the Variables in the Study	40
Table 4.2 Descriptive Statistics for Groups of Intrinsic Religious Orientation	43
Table 4.3 Correlations of Intrinsic Religious Orientation and Criterion Variables	44

## List of Tables

	<u>Page</u>
Table 4.4 Seven ANOVAs for Intrinsic Religious Orientation	47
Table 4.5 The Goodness of Fit for Three Models (Linear, Exponential and Logistic) using the Square of the Correlation Coefficient ( $R^2$ )	57
Table 5.1 Number of References using Isotonic Regression in Psychology and other Disciplines	60

## List of Figures

	<u>Page</u>
Figure 1.1: Plots of isotonic trends between population means	6
Figure 3.1: Plots that demonstrate goodness of fit for two regression models	23
Figure 3.2: Different models that explain monotonic functions	30
Figure 4.1: For ANOVA, Attachment to Mother as a function of Intrinsic Religiosity	48
Figure 4.2: For ANOVA, Attachment to Father as a function of Intrinsic Religiosity	48
Figure 4.3: For ANOVA, Neuroticism as a function of Intrinsic Religiosity	48
Figure 4.4: For ANOVA: Conscientiousness as a function of Intrinsic Religiosity	48
Figure 4.5: For ANOVA, Agreeableness as a function of Intrinsic Religiosity	48
Figure 4.6: For ANOVA, Sense of Coherence as a function of Intrinsic Religiosity	48
Figure 4.7: For ANOVA, Satisfaction with Life as a function of Intrinsic Religiosity	48
Figure 4.8: For isotonic regression, Attachment to Mother as a function of Intrinsic Religiosity	51
Figure 4.9: For isotonic regression, Attachment to Father as a function of Intrinsic Religiosity	51
Figure 4.10: For isotonic regression, Neuroticism as a function of Intrinsic Religiosity	51
Figure 4.11: For isotonic regression, Conscientiousness as a function of Intrinsic Religiosity	51
Figure 4.12: For isotonic regression, Agreeableness as a function of Intrinsic Religiosity	51

## List of Figures

	<u>Page</u>
Figure 4.13: For isotonic regression, Sense of Coherence as a function of Intrinsic Religiosity	51
Figure 4.14: For isotonic regression, Satisfaction with Life as a function of Intrinsic Religiosity	51

## Acknowledgements

There are many people who I would like to thank for the work pertaining to this thesis. First, I would like to thank Dr. Huynh for the many revisions that he has made on the drafts that I have sent him. His comments have been very construct and helpful during the writing process. I would also like to thank Dr. Schludermann for his comments pertaining to the data analysis portion of the thesis. Thanks are also extended to Dr. Xikui Wang for his advice on statistical regression. Finally, I would like to thank my family and friends for helping make this possible, namely Joshua and Esther Blum. I would like to dedicate my degree to my Mother, who passed away on January 1, 2002.

## Abstract

Isotonic regression is a technique for detecting trends that have the property of being non-decreasing or non-increasing. The first purpose of this study is to present the theory and computational procedures to conduct isotonic regression analysis for a one-way analysis of variance experimental design. The second purpose is to illustrate the use of the isotonic regression technique in analyzing psychological data. Using a data set from Hossack (1997), the property of isotonicity is tested by using a technique that preserves the order of group means. In this study, groups of Intrinsic Religious Orientation (IRO) are hypothesized to be significantly different for Attachment to Mother (ATM), Attachment to Father (ATF), Sense of Coherence (SOC) and Satisfaction with Life (SWL) at the 0.05 level of significance. Moreover, there is no expected significant difference hypothesized between IRO groups for Agreeableness (AGR), Conscientiousness (CON) or Neuroticism (NEU) at the same level of significance. Results confirm these hypotheses under both ANOVA and isotonic regression analyses. However, there are three exceptions. For ANOVA, the IOR groups were significantly different for CON ( $p < 0.01$ ) when a non-significant result was expected. For isotonic regression, the IOR groups were significantly different for NEU and AGR ( $p < 0.05$ ), when a non-significant result was expected. In addition, the goodness of fit values using isotonic regression estimates were greater than goodness of fit values using sample mean estimates for each of the criterion variables. Besides the linear ANOVA-typed model, isotonic regression techniques are also used with some nonlinear models such as exponential and logistic functions. The implications are that isotonic parameter estimates should be used instead of sample mean estimates. In the conclusion, the advantages and disadvantages of isotonic regression are discussed.

## Chapter One: Background of the Study

## 1.1 Introduction

Tests for hypotheses that involve order restricted inferences, such as  $\mu_1 \leq \mu_2 \leq \dots \leq \mu_k$  or  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_k$ , have been introduced by Bartholomew (1959a, 1959b), and developed by Page (1963), Barlow, Bartholomew, Bremner and Brunk (1972), and Robertson, Wright and Dykstra (1988). An introductory example in Nelson (1976)

**Table 1.1****Warm-up Times for a Three Level Factor of Tube Type**

		Tube Type					
		A		B		C	
	19	20	20	40	16	19	
	23	20	20	24	15	17	
	26	18	32	22	18	19	
	18	35	27	18	26	18	

Note: Data from Hicks' (1973, p.49) is used to demonstrate that significance for hypotheses with directional alternatives depends on the sequence of the population means.

demonstrates that tests for directional hypotheses may be applied to detect increasing warm-up time in glass containers shaped as tubes. He concludes that there is a directional trend of warm-up time between glass containers. The data are presented in Table 1.1 and can be used to show that there is warming trend between containers depending on container sequence. There is a significant directional trend for  $\mu_C \leq \mu_A \leq \mu_B$  ( $p < 0.05$ ) and  $\mu_C \leq \mu_B \leq \mu_A$  ( $p < 0.05$ ), but not  $\mu_A \leq \mu_C \leq \mu_B$  ( $p > 0.05$ ) according to an isotonic

regression analysis where the dependent variable is warm-up time and the factor is type of tube (levels A, B or C).

The concept of testing for directional hypotheses with continuous criterion variables and categorical factors has been developed from both a parametric and non-parametric points of view. From a parametric perspective, the likelihood ratio test and  $\bar{E}^2$  statistic are both derived from the  $\chi^2$  statistic using frequency data (Bartholomew, 1959a; Barlow et al, 1972, p.119, 121). From a non-parametric perspective, Joncheere (1954) has introduced a distribution-free test for directional alternatives, and Page (1963) presented a ranking statistic,  $L$ , for a test of monotonic relationships in a two-way ANOVA to be used in place of Friedman's chi-square of ranks statistic. Recently, this technique has also been used in an applied setting. Leuraud and Benichou (2001) introduce a chi-square test to detect trends for a monotonic dose-response relationship between exposure and disease in epidemiological and clinical studies. Karabatsos and Ullrich (2002) also use a Monte Carlo procedure to illustrate the importance of order-restricted statistical inference for the selection and testing of mathematical (axiomatic) models. The testing of directional hypotheses using isotonic regression analysis has therefore been developed substantially over the years.

Standard ANOVA methods have not usually been used for the testing of directional hypotheses. The standard ANOVA does not constrain the function to vary monotonically, which can give rise to inferences that are quite contradictory to our intuition (Mukerjee & Tu, 1995). The standard ANOVA tests for a non-directional alternative hypothesis where there is no sequence of population means, say  $H_1: \mu_i \neq \mu_j$  for at least one  $i$  and  $j$ ;  $i, j = 1, 2, \dots, n$ . Moreover, ANOVA uses numerous types of multiple

comparison procedures upon failure to retain the null hypothesis. Hypotheses with ordered-restricted alternatives, on the other hand, will result in combining the sample means where necessary and plotting the data to detect a directional trend. Although there are multiple comparison procedures for directional alternatives, see Ruberg (1989), a plot of the trend is used more frequently.

Trend analysis can be an alternative to isotonic regression. However, there are also limitations to the standard methods for trend analysis. A trend may be non-monotonic such as a quadratic or cubic function, so the constraint of non-decreasing (or non-increasing) parameter estimates must be imposed. A typical method of trend analysis is the use of planned contrasts, which will detect if a trend is not significantly different from a linear or quadratic function. A limitation to this type of analysis is that it is possible that there is not a planned contrast available to constrain for monotonicity. Several contrasts would be necessary to detect a monotonic function, which inflates the type I error rate. There are additional limitations to trend analysis such that sometimes only the significance of the means for the first and last treatment groups is used to detect a monotonic trend (Page, 1963). The two limitations of this approach are that, first, such a test discards large numbers of observations between the extremes and, second, the order of the middle treatment groups is not taken into consideration.

The importance of testing for directional hypotheses using isotonic regression techniques is demonstrated by overcoming the limitations of the standard methods for ANOVA and trend analysis. As a parametric test, isotonic regression is generally more powerful than non-parametric tests (Robertson et. al, 1988, p.86). In addition, isotonic regression techniques use the sums of squares (according to the least squares principle) to

maintain sensitivity to outlying observations and make sample size a factor for significance tests. Each of the treatment groups is used in the isotonic regression techniques and the shape of the trend is always tested for linear monotonicity (i.e. not quadratic or cubic). There is also only one test for isotonic regression as the order of the population means is usually determined by empirical evidence from related experiments, prior reasoning or expert opinion. The single test approach yields lower type I and experimental error than multiple comparison procedures of the standard ANOVA omnibus F-test (Page, 1963).

### 1.2 What is isotonic regression?

Isotonic regression is the statistical theory for applying the regression techniques “to deal with problems in which conditional expectations are subject to order restrictions” (Barlow et al, 1972, p.5). In general, the term ‘isotonic’ is a synonym for ‘order-preserving’ as the order of the population means,  $\mu_1 \leq \mu_2 \leq \dots \leq \mu_k$  or  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_k$  is tested when there are several possible reversals from their sample mean estimates. Without loss of generalizability, we only consider the case of monotonically increasing (non-decreasing) estimates (Barlow et al., 1972, p. 7). The process of preserving the direction of the population means requires the formation of isotonic regression estimates  $\mu_i^*$  where adjacent sample means that violate the order of the trend are combined into one estimate. The factor levels that underlie the amalgamation process demonstrate the properties of reflexivity, transitivity and asymmetry (Robertson et al., 1988, p.12). A typical hypothesis for isotonic regression would be that a criterion variable  $\mu_i$  (e.g.  $y =$  test scores in a course) will be increased when the strength of a stimulus  $x_i$  (e.g.  $x =$  number of hours studying) is gradually increased over several populations (age groups,

subject areas, etc.). A trend of score increments can be examined with test scores as the dependent measure and hours of study as the between-subjects factor. The levels of the factor  $x_1, x_2, \dots, x_k$  (or the number of hours for studying) can be nominal, categorical, or ordinal and do not need to be placed in sequence from largest to smallest or vice versa. As is the case for the usual one-way analysis of variance, the  $\mu_i$  are from a ratio or interval scale of measurement.

The figure below illustrates three monotonic trends which will be useful for explaining the properties of isotonic regression models. In each figure,  $y$  represents test scores (from 0 to 50) and  $x$  represents five categories of hours spent studying. In the

**Figure 1.1**

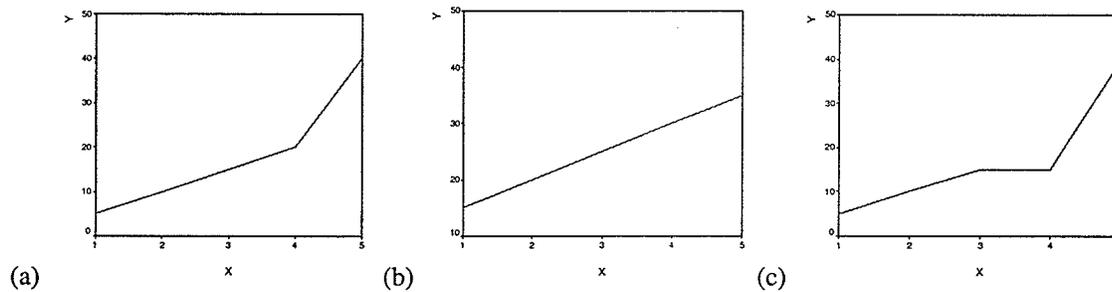


Figure 1.1 The plots are examples of trends between population means  $\mu_1 \leq \mu_2 \leq \dots \leq \mu_5$  that may be tested using isotonic regression. Each of these trends do not have to be linear and may be compared against the hypothesis of no trend.

first plot, (a), test scores in a course  $y_i$  increases without interruption across each group.

The curve has three linear, non-decreasing segments which are joined together to demonstrate that the level of the criterion variable is steadily increasing with study time. The trend could be non-linear, but since it increases without decreasing there is no need to combine the sample means. There are therefore no isotonic regression estimates,  $\mu_i^*$ , that are required for this sequence of population means. An isotonic regression analysis

may now be conducted directly on the sample means of test scores to determine if the population means demonstrate a monotonic trend. The second plot, (b), shows a linear, increasing trend. The test scores ( $y_i$ ) increase proportionately to the level of the regressor variable. No amalgamations of sample means are needed since the trend is monotonic. Therefore, the isotonic regression analysis can be conducted on the sample means immediately across each of the categories of hours of study. The third plot, (c) shows that  $y_i$  increases but at different rates, and the trend is nevertheless monotonic. An amalgamation of group means is required at the location where the trend is close to being constant (in the middle section of the curve). The two adjacent sample means are first combined and then the isotonic regression technique produces the arithmetic average of the two sample means. The result is a trend that would have potentially demonstrated a decrease prior to the amalgamation, but is now monotonic and can be analyzed by the isotonic regression procedure for hours of study. To summarize, isotonic regression is used primarily to detect if there is a non-decreasing trend between population means. The sample means must be order-preserving (i.e. monotonically trended) prior to deriving the test statistic, adhering to the direction of the population means. If this condition is violated, some of the group means must be combined and their averages are computed. This is called "the amalgamation process". The amalgamation process of producing isotonic regression estimates reassures that the sample means are in a monotonic sequence. Both the properties and procedures for the amalgamation process and the direction of sample means are discussed in chapter two.

For the results of isotonic regression analysis to be properly interpreted, several assumptions must be met. First, the criterion and regressor variables are normally

distributed. Second, the homogeneity of variance assumption is satisfied; that is, the error variance of the criterion variable is not significantly different between groups. Thirdly, the observations are independent. The final assumption is that the order of the population means is preserved, with at least one inequality (say,  $\mu_1 \leq \mu_2 \leq \dots \leq \mu_k$ ). The first assumption is verified by using a histogram or a plot of the normal quantiles (P-P plot). The second assumption is verified by using a residual plot and the third assumption is maintained by experimental design. The final assumption of monotonicity applies an algorithm which verifies that the predicted order of the sample means is preserved. Empirical studies have shown that test statistics in standard regression analysis are robust to departures of the normality and homogeneity of variance assumptions (Maxwell & Delaney, 1990, p.109).

### 1.3 Statement of Purpose

The purpose of this thesis is threefold. Firstly, the theory, development and application of isotonic regression are presented in such a way that is relevant for psychologists. The theory component involves presenting hypotheses that can be tested in isotonic regression analyses and the background for the use of isotonic regression estimates from some techniques such as the pool-adjacent-violators algorithm. In particular, the development of the two most prominent test statistics for isotonic regression is also illustrated, with an emphasis on their critical value tables and the appropriate use of the scientific method for psychological research (stating the hypothesis, statistic, decision and conclusion). The applied component is a study on the isotonic trend of groups varying in relation to the other variables. The regressor variable is represented by 10 balanced groups of 520 undergraduate psychology students, an

almost equal number of males and females, who have completed Allport's (1960) Intrinsic Religiosity Scale. Separate ANOVAs and isotonic regressions are computed on each of the seven criterion variables to be elaborated on in Section 1.4 and Chapter 4.

Secondly, this thesis discusses the issues and procedures for testing the role of goodness of fit for isotonic regression models. The roles as well as the methods for goodness of fit assessment in the context of linear and non-linear trends are studied. Illustrative examples are used to show the differences in the goodness of fit values between the standard ANOVA and isotonic regression models. The intention is to demonstrate the strengths of isotonic regression in estimation and statistical inference.

Thirdly, this thesis examines the advantages and limitations of isotonic regression and explores further avenues for psychological research. The advantages include a lower standard error of the sample mean estimate, better type I error control and increased power. The main limitation of isotonic regression is that the shape of the trend is compromised to satisfy the monotonicity assumption. Methods for improving this shortcoming are discussed.

#### 1.4 Overview

The main purpose of this thesis is to present the isotonic regression technique to researchers in the field of psychology. The focus will be on the significance testing and goodness of fit models for isotonic regression with Intrinsic Religious Orientation (IRO) as the independent variable. Chapter two introduces the null and alternative hypotheses for several order-restricted population means, as well as their test statistics and critical values. Chapter three presents goodness of fit in model assessment as well as methods for assessing goodness of fit when trends are monotonically increasing (or decreasing).

The ordinary least squares simple regression and multiple regression methods are compared to the isotonic regression technique with an illustrative example involving the effects of drugs on driving behavior. Chapter four involves an application of isotonic regression with psychology variables<sup>1</sup>. The IRO variable will be used as the regressor variable for seven psychology criterion variables. Seven ANOVAs and isotonic regressions are computed separately with the use of multiple comparison procedures to detect differences between 10 IRO groups. The results of the ANOVA models and isotonic regressions are compared by means of plots to detect any trends that may be attributed to the relation between the regressor and criterion variables. The hypothesis that isotonic regression estimates produce higher goodness of fit values than sample mean estimates for linear, exponential and logistic models is also tested. Each model has the property of being monotonically increasing. Chapter five presents the advantages and disadvantages of isotonic regression with the most recent literature for directional decisions. The advantages include the sophistication of the isotonic regression technique compared to standard ANOVA procedures and the trend analysis method, as well as the capacity to make decisions pertaining to non-normal data, statistical power and Bayesian inference. The disadvantages include the inability to test for a specific function or shape of a trend, the lack of programming software for directional alternatives, and the few empirical studies in the social sciences that apply isotonic regression to their psychology variables. The new contributions of isotonic regression to psychologists include a statistical technique to detect trends that are non-increasing or non-decreasing. This

---

1. Permission was given by Dr. Schludermann to use Hossack's (1997) data set.

technique also introduces isotonic regression estimates and overcomes several limitations of other trend analysis techniques. An application of isotonic regression demonstrates that this technique provides useful information to psychologists.

## Chapter Two: Hypothesis Testing in Isotonic Regression

## 2.1 The Null and Alternative Hypotheses

To briefly outline this chapter, the isotonic regression  $\chi^2$  and  $\bar{E}^2$  statistics are presented with research hypotheses, algorithms for isotonic regression estimates, critical values and illustrative examples. More specifically, the isotonic regression hypotheses are compared against the hypotheses for ANOVA, and a step-by-step approach is used to demonstrate the pool-adjacent-violators algorithm. The distributions for the  $\chi^2$  and  $\bar{E}^2$  statistics are also defined along with explanations pertaining to critical values. The chapter concludes by using average monthly temperatures in Palermo, Italy to illustrate the  $\chi^2$  test and time to complete a test as measured by length of test using the  $\bar{E}^2$  statistic.

In the standard one-way analysis of variance, the null hypothesis for isotonic regression postulates that all group means are the same, namely  $H_0: \mu_1 = \mu_2 = \dots = \mu_k$  for all  $j = 1, 2, \dots, k$ . However, the alternative hypotheses are different, in the ANOVA and isotonic regression techniques. In the isotonic regression context, Bartholomew (1959a) and Barlow et al (1972, p.2) postulate a sequence of population means, say  $H_1: \mu_1 \leq \mu_2 \leq \dots \leq \mu_k$  or  $H_1: \mu_1 \geq \mu_2 \geq \dots \geq \mu_k$ , with at least one inequality between population means. The population means demonstrate a two-sided alternative in either the non-decreasing or non-increasing direction. On the other hand, the alternative hypothesis for ANOVA states that the population means are not significantly different from one another, say  $\mu_i \neq \mu_j$  for all  $i \neq j$ , when  $i, j = 1, 2, \dots, k$  (Maxwell and Delaney, 1990, p.78).

Barlow et al (1972, p.2) mention that the simplest hypothesis for isotonic regression is similar to that of the one-tail t-test for two groups, in which the null

hypothesis is postulated as  $H_0: \mu_1 = \mu_2$  and the alternative hypothesis, as  $H_1: \mu_1 \leq \mu_2$  or  $H_1: \mu_1 \geq \mu_2$ . In fact, when there are only two groups, one should use a one-tailed (directional) t-test to evaluate the above hypothesis. Hence, given the homogeneity of variances ( $\sigma_1^2 = \sigma_2^2$ ), if the tests statistic  $t = (\bar{x}_1 - \bar{x}_2) / S_p \sqrt{(1/n_1 + 1/n_2)}$  where  $S_p = ((n_1 S_1^2 + n_2 S_2^2) / (n_1 + n_2 - 2))^{1/2}$  is greater than the corresponding critical value,  $t_{(n_1 + n_2 - 2), \alpha/2}$ , then the null hypothesis is not retained. The procedure and test statistics under the isotonic regression approach for two or more populations are discussed in Section 2.3.

## 2.2 The Pool-Adjacent-Violators Algorithm

In the isotonic regression testing procedure, when the sample means  $\bar{y}_1, \dots, \bar{y}_k$  do not preserve the directional sequence of the population means  $\mu_1 \leq \mu_2 \leq \dots \leq \mu_k$ , say, then one must obtain isotonic regression estimates of group means prior to computing the test statistics in the hypothesis testing procedure. Consider the case of three group means where the alternative hypothesis is postulated to be  $H_1: \mu_1 \leq \mu_2 \leq \mu_3$ . Suppose the sample mean estimates are observed to be  $\bar{y}_1 \leq \bar{y}_2 \geq \bar{y}_3$ . The sample mean  $\bar{y}_3$  has violated the direction of the hypothesized trend. Therefore new estimates of sample means must be obtained such that the non-decreasing sequence of the population means is preserved. For this purpose, the sample mean estimates are replaced by their isotonic regression estimates. In the example just presented, the sample mean estimate  $\bar{y}_3$  is replaced by the isotonic regression estimate  $\bar{y}_3^* = (\bar{y}_2 + \bar{y}_3)/2$ . The technique for arriving at isotonic regression estimates is referred to as the "pool-adjacent-violators algorithm" (Barlow et al., 1972, p.13). The algorithm consists of the following three steps.

Step 1: Align the sample mean estimates in the order of the population means under the alternative hypothesis, say  $H_1: \mu_1 \leq \mu_2 \leq \dots \leq \mu_k$ . Then  $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_k$  are placed adjacent to one another, in that order.

Step 2: Determine if the order of the population means has been violated. If  $\bar{y}_j > \bar{y}_{(j+1)}$  for any  $j = 1, 2, \dots, k$ , then both  $\bar{y}_j$  and  $\bar{y}_{(j+1)}$  are replaced by their arithmetic average, namely  $\bar{y}_j^* = \bar{y}_{(j+1)}^* = (\bar{y}_j + \bar{y}_{(j+1)})/2$ . For example, let the sample means be  $\bar{y}_1 = 1, \bar{y}_2 = 4, \bar{y}_3 = 2$  and  $\bar{y}_4 = 5$ . Suppose the alternative hypothesis in step1 is specified as  $H_1: \mu_1 \leq \dots \leq \mu_5$ , then  $\bar{y}_2$  and  $\bar{y}_3$  are replaced by their isotonic regression estimates of  $\bar{y}_2^* = \bar{y}_3^* = (4+2)/2 = 3$ , according to the pool-adjacent-violators algorithm. The outcome of this procedure is presented below in Table 2.1.

**Table 2.1**

**Sample Mean Estimates become Isotonic Regression Estimates**

<u>Sample Mean Estimates</u>			
$\bar{y}_1$	$\bar{y}_2$	$\bar{y}_3$	$\bar{y}_4$
1	<u>4</u>	<u>2</u>	5

---

<u>Isotonic Regression Estimates</u>			
$\bar{y}_1^*$	$\bar{y}_2^*$	$\bar{y}_3^*$	$\bar{y}_4^*$
1	3	3	5

Note: In the above table, a straight line connects group means that do not conform to the postulated direction in the alternative hypothesis. The group means are combined to become isotonic regression estimates and preserve the direction of the population means.

Step 3: There may be more than one violator among the sample means. Treat each violator as a block, pool it with its preceding sample mean, and obtain the isotonic regression estimate as the arithmetic average of these adjacent sample means. If there is more than one block, treat each one separately. When the algorithm is finished, the blocks will be combined into one segment which contains isotonic regression estimates that preserve the order of the population means. For the data presented below in Table 2.2,  $\bar{y}_3$  and  $\bar{y}_5$  are violators that must be amalgamated.

If the sequence of sample means  $\mu_1 \leq \mu_2 \leq \dots \leq \mu_k$  is not violated then this partition of sample means is also the final partition and  $\mu_j^* = \mu_j, j = 1, 2, \dots, k$ . If not,

**Table 2.2**

**Blocks of Sample Mean Estimates become Isotonic Regression Estimates**

---

							<u>Blocks</u>
$\bar{y}_1$	$\bar{y}_2$	$\bar{y}_3$	$\bar{y}_4$	$\bar{y}_5$	$\bar{y}_6$	$\bar{y}_7$	
1	<u>4</u>	<u>2</u>	<u>7</u>	<u>5</u>	6	7	

---

<u>Isotonic Regression Estimates</u>						
$\bar{y}_1^*$	$\bar{y}_2^*$	$\bar{y}_3^*$	$\bar{y}_4^*$	$\bar{y}_5^*$	$\bar{y}_6^*$	$\bar{y}_7^*$
1	3	3	6	6	6	7

---

Note: When there is more than one violator, the sample mean estimates are put into blocks and the isotonic regression estimates are ready to be calculated. As mentioned previously, a straight line connects means that do not conform to the postulated direction in the alternative hypothesis.

select any of the pairs of violators of the ordering, such that  $\mu_j > \mu_{j+1}$ , and ‘pool’ these two values placing them in the same block. When the ordering of  $\mu_1 \leq \mu_2 \leq \dots \leq \mu_k$  is observed, then the algorithm is finished.

## 2.3 Test Statistics and Critical Values

There are two approaches in calculating the test statistics for isotonic regression. The first is the likelihood ratio test derived by Bartholomew (1959a, 1959b). This statistic uses the weighted sums of squares after amalgamation as the numerator and  $\sigma^2$  as the denominator,  $\chi^2 = \sum n_j (\mu_j^* - \mu)^2 / \sigma^2$  where  $n_j$  is the sample size of the  $j$ th population,  $\mu_j^*$  is the isotonic regression parameter of the  $j$ th population mean, and  $\mu$  is the grand population mean. This approach may not be practical since the population variance  $\sigma^2$  is assumed to be known a priori (Barlow et al, 1972, p.116). The critical values for the  $\chi^2$  test are presented in Table 2.4, the null distribution being  $\Pr\{\chi_k^2 \geq C\} = \sum P(l,k) \Pr\{\chi_{l-1}^2 \geq C\}$  when  $C > 0$ , and  $\Pr\{\chi_k^2 = 0\} = P(l,k)$  when  $C = 0$  (Barlow et al, 1972, p. 126). The notations used in the above distribution are: 'k', to represent the number of levels in the factor, 'l', the number of levels after amalgamation and  $P(l,k)$ , the probability that the isotonic regression function  $\mu_i^*$  takes exactly  $l$  distinct values from  $k$  population means. The sum of the probabilities  $\sum P(l,k)$  is the probability of all possible  $l$  and  $k$  combinations. The second statistic, denoted as  $\bar{E}^2$ , is more practical for applied researchers. It is computed as  $\bar{E}_k^2 = \sum n_i (y_j^* - \bar{y})^2 / \sum \sum (y_{ij} - \bar{y})^2$ , for which the numerator is an estimate of the numerator in the  $\chi^2$  test, but the denominator is the total sums of squares (Barlow et al., 1972, p.120). This statistic represents the amount of criterion variance accounted for by the group means of  $y_{ij}$ . The null distribution is  $\Pr\{\bar{E}_k^2 \geq C\} = \sum P(l,k) \Pr\{B_{1/2(l-1), 1/2(N-l)} \geq C\}$  when  $C > 0$  and  $\Pr\{\bar{E}_k^2 = 0\} = P(l,k)$  when  $C = 0$  (Barlow et al., 1972, p.127). The notation  $B_{p,q}$  denotes the beta distribution with parameters  $p$  and  $q$  as the shape and location respectively. In general, as the  $\chi^2$  and  $\bar{E}^2$  values increase, the

more likely it is that the null hypothesis will be rejected. Some selected critical values for  $\bar{F}$  are found in Nelson (1977) and are reproduced in the Appendix. The tables are for applying isotonic regression to one-way ANOVA data which consists of  $j = 3, 4, \dots, 10$  factor levels.

#### 2.4 Illustrative Examples

There are two examples presented in this section to illustrate the isotonic regression techniques discussed above. In the first example, Cameron (1970) studied the

**Table 2.3**

**Likelihood Ratio Test Analysis of Mean Monthly Temperatures of Palermo, Italy**

	Month											
	Jan	Feb	Mar	April	May	June	July	Aug	Sept	Oct	Nov	Dec
$\bar{y}_i$	51	52	55	58	64	71	76	77	<u>73</u>	<u>67</u>	<u>59</u>	<u>53</u>
$n_i$	31	28	31	30	31	30	31	31	30	31	30	31

$$\mu_1 = \Sigma(n_i \bar{y}_i) / \Sigma n_i = 17\,600 / 274 = 64.23$$

$$\mu_2 = \Sigma(n_i \bar{y}_i) / \Sigma n_i = 92 / 179 = 59.67$$

$$\chi_1^2 = \Sigma n_i (\mu_i - \mu_1)^2 / \sigma^2 = 1059.962$$

$$\chi_2^2 = \Sigma n_j (\mu_j - \mu_2)^2 / \sigma^2 = 122.3287$$

**Note:** In the first row, there is an increasing trend until August, followed by a decreasing trend (underlined) until December. The  $\bar{y}_i$  represents mean monthly temperature and the  $n_i$  represents the number of days per month.

average monthly temperature in Palermo, Italy. The criterion variable is average monthly temperature and the regressor is the month of the year. The monthly temperatures tend to increase from January to September and then decrease from October to December as seen in Table 2.3. There are therefore two trends to be tested, one that is non-decreasing and another that is non-increasing. The null hypothesis  $H_0: \mu_{\text{Jan}} = \mu_{\text{Feb}} = \dots = \mu_{\text{Dec}}$  is tested against  $H_1: \mu_{\text{Jan}} \leq \mu_{\text{Feb}} \leq \dots \leq \mu_{\text{Sept}}$  (warming trend) and  $H_2: \mu_{\text{Oct}} \geq \mu_{\text{Nov}} \geq \mu_{\text{Dec}}$  (cooling

Table 2.4

**Critical Values for the  $\chi^2$  Test ( $H_1: \mu_1 \leq \mu_2 \leq \dots \leq \mu_k$  or  $H_1: \mu_1 \geq \mu_2 \geq \dots \geq \mu_k$ )**

K						
$\alpha$	3	4	5	6	7	
0.100	2.580	3.187	3.636	3.994	4.289	
0.050	3.820	4.528	5.049	5.460	5.800	
0.025	5.098	5.891	6.471	6.928	7.304	
0.010	6.823	7.709	8.356	8.865	9.284	
0.005	8.144	9.092	9.784	10.327	10.774	
0.001	11.249	12.318	13.098	13.711	14.214	
$\alpha$	8	9	10	11	12	
0.100	4.542	4.761	4.956	5.130	5.288	
0.050	6.088	6.339	6.560	6.758	6.937	
0.025	7.624	7.901	8.145	8.363	8.561	
0.010	9.638	9.945	10.216	10.458	10.676	
0.005	11.152	11.480	11.768	12.025	12.257	
0.001	14.640	15.009	15.333	15.622	15.883	

Note: The  $\chi^2$  critical values for a factor with up to 12 levels are presented above. The levels of significance ( $\alpha$ ) range from 0.100 to 0.001 (see Robertson et al., 1988, p. 416).

trend) using the likelihood ratio test. The trends are expected to be significant at  $\alpha = 0.05$  as no amalgamations or isotonic regression estimates are required. The critical values for  $H_1$  and  $H_2$  are set at  $k = 9$  and  $k = 3$  factor levels respectively. The variance,  $\sigma^2$ , is unknown, however, to illustrate the use of the  $\chi^2$  test, assume  $\sigma^2 = 25$ . The design is unbalanced as there are an unequal number of days per month and the sample mean estimates are used in place of isotonic regression estimates. The numerator for the likelihood ratio test is therefore the between sums of squares,  $\sum n_i(\mu_i - \mu)^2$ , and not the sums of squares after amalgamation,  $\sum n_i(\mu_i^* - \mu)^2$  where  $\mu_i^*$  is the  $i$ th isotonic regression estimate, for  $i = 1, 2, \dots, k$ . It can be concluded from the likelihood ratio test using the critical values found in Table 2.5 that there is a significant warming trend in Palermo, Italy ( $\chi_1^2 = 1059.96 \geq 6.937$ ) followed by a significant cooling trend ( $\chi_2^2 = 122.33 \geq 3.820$ ).

The Palermo, Italy example reveals some strengths and weaknesses of the isotonic regression technique. The strengths include a single, one test approach to determine if there is a monotonic trend among several population means that is robust to departures of the normality and homogeneity of variance assumptions (Chacko, 1963 & Shorack, 1967). The technique is more flexible than the standard ANOVA and trend analysis method for directional hypotheses and provides clear decisions. The weakness is that the variance  $\sigma^2$  must be assumed a priori. Hence, the conclusions are theoretical and tentative for the likelihood ratio test.

As an example of the  $\bar{E}$  calculation, we use a data set from Lehman (1995) for testing  $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$  against  $H_1: \mu_1 \leq \mu_2 \leq \mu_3 \leq \mu_4$ . The criterion variable is the

time in minutes to complete a multiple-choice test. The regressed variable is the length of the test (10, 15, 20 or 25 items). The assumption of monotonicity is that the sample means will have the ordering of  $\bar{y}_1 \leq \bar{y}_2 \leq \bar{y}_3 \leq \bar{y}_4$  so that the trend is non-decreasing. When there are violations to the order of the sample means, then isotonic regression estimates are used according to the pool-adjacent violators algorithm. For the example presented in Table 2.6, isotonic regression estimates are unnecessary since as length

**Table 2.6**

**Time to Complete the Test is Measured for Length of Test**

	Length of Test (Items)			
	10	15	20	25
	13	17	18	16
	10	5	17	21
	13	13	10	28
	5	14	23	22
	6	11	24	24
	6	12	17	12
	2	3	10	9
	12	16	27	18
	17	18	21	18
	12	8	24	9
	12	14	23	15
	2	7	8	15
	17	13	14	10
	3	6	1	7
	7	14	8	30
	1	18	10	14
Sum:	138	199	255	268
Mean:	8.63	12.43	15.94	16.75

$$\bar{E}^2 = \frac{\sum n_i (y_i^* - \bar{y})^2}{\sum \sum (y_{ij} - \bar{y})^2} = 662.125 / 11556.25 = 0.225232$$

Note: An independent groups design shows that as the length of the test increases so does the time to complete the test. The  $\bar{E}^2$  statistic is used since raw data has been collected.

of the test increases, so does length of time to complete the test. The sample means are placed in order from smallest to largest according to the order of the population means. Given that  $\sigma^2$  can be estimated from the data, the  $\bar{E}^2$  statistic is then used in this example. It is expected that  $H_0$  will not be retained at the 0.05 level of significance as the sample mean estimates increase at a steady rate.

Since  $\bar{E}^2 = 0.2252 \geq \bar{E}_{0.05}^2 = .0571$  (where  $\alpha = 0.05$ ,  $n = 16$  and  $k = 4$ ), the critical value obtained from the Appendix, there is sufficient evidence to reject  $H_0$  and conclude that as the length of the test increases, so does the time to complete the test. The illustrative examples of the data sets by Carmen (1970) and Lehman (1995) show the uses of the  $\chi^2$  test for theoretical purposes and the  $\bar{E}^2$  statistic for applied situations.

## Chapter Three: Goodness of Fit of Isotonic Regression Models

## 3.1 The Role of Goodness of Fit Model Assessment

There are several roles goodness of fit plays when assessing statistical models. In this section, we will discuss three of them. First, goodness of fit assessment refers to how well a model, or combination of models, explain the patterns of variation of exploratory criterion variables within a particular data set (King and Minium, 2003, p.187). The goodness of fit tests decide the best fit for a single regression line, and obtain the best regression line among several possible regression lines (Neter, Kutner, Nachtsheim, Wasserman, 1996, p.81). The better the fit, the more accurate the models are for making predictions about future outcomes. However, two models are often compared against one another to determine which exploratory variables, or combination of these variables, can be used to explain the criterion variation more accurately (Neter et al, 1996, p. 287). Secondly, goodness of fit assesses the strengths and weaknesses of each model, such as how well each model fits the data and how different combinations of variables in the model affect the fit more profoundly than others. In the context of regression models, the simplest measures of goodness of fit are the standard error,  $S_e$ , and the square of the correlation coefficient,  $R^2$  (King and Minium, 2003, p.157, 162). The standard error is the average standard distance of the observations from many samples for the line of best fit. The square of the correlation coefficient is the amount of criterion variance explained by one or many of the regressors in the model. The partial correlation is used to identify how much variation a group of variables explains when one variable is excluded from the group. Thirdly, the results from goodness of fit indices can help to identify (or eliminate) a set of regressors that reduce error variation for prediction or description purposes (Neter

et al, 1996, p.346). The concept involves minimizing error variance, according to the least squares principle, to provide an increasingly accurate representation of the real world. In regression analysis, some popular techniques for variable and model selection involve the forward, backward and stepwise procedures.

Let us compare the regression lines of two plots in Figure 3.1 to better understand goodness of fit. From a statistical perspective, a regression line will have a good fit if it minimizes the residuals,  $e_i$ , between the observed points ( $y$ ) and estimated points ( $\hat{y}$ ). The sum of the residuals,  $\sum e_i = \sum (y_i - \hat{y}_i)$  for  $i = 1, 2, \dots, n$ , can be used to determine the best fit

**Figure 3.1**

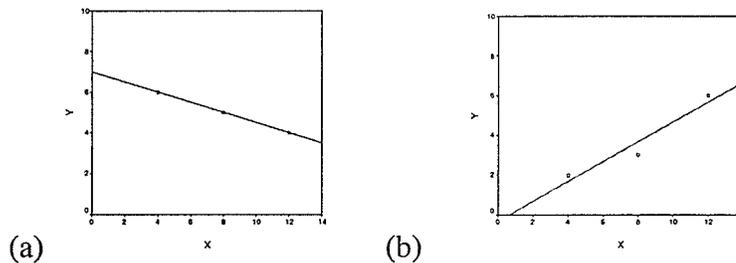


Figure 3.1 Two plots from Wayne (1986, p.378) demonstrate the estimated lines for the fit of two regression models. For (a), the estimated line fits the observed points exactly, whereas for (b), the observed points deviate more from the estimated line.

of the regression models. The difficulty with this approach is that this sum will always equal zero, providing no new information about which line has the best fit. On the other hand, taking the sum of the absolute value of the residuals between the observed and predicted points,  $\sum |y_i - \hat{y}_i|$  provides useful information as the line in plot (a) is a better measure of goodness of fit than the line in plot (b). The problem with absolute values is that they do not indicate the direction of the error for each of the observations. In effect, we want to find a way to penalize large errors so that we can have a model that is more

sensitive to individual differences (or outlying observations). This can be accomplished by taking the sum of the squared residuals,  $\sum e_i^2 = \sum (y_i - \hat{y}_i)^2$ , for  $i = 1, 2, \dots, n$ . Squaring the difference magnifies large errors and cancels the effect of the positive and negative signs in front of the values. Since we are looking for the line that minimizes the error sum of the squares, we call this the least squares principle. With this technique, plot (a) minimizes error (provides a better fit) more effectively than plot (b). When one line has more points than another line then the estimated standard error  $S_e$  is often used to compare the goodness of fit of the two lines. This is a measure of goodness of fit that is used as frequently as the square of the correlation coefficient,  $R^2$ . As goodness of fit improves,  $R^2$  values become larger and  $S_e$  values become smaller.

There is a trade-off between goodness of fit (minimizing the error sums of squares) and the principle of variable parsimony. Regressors that are added to a model explain more of the error variance that is not previously accounted for by other regressors already in the model. With enough regressors, a model could have a nearly perfect fit of the data. Further, a regression model with enough regressors will always yield a good fit. The drawback to too many regressors is the loss of degrees of freedom as there is no longer any potential to generalize a conclusion beyond the sample. There is also no applicability for interpolation within the sample since the error variance will be so small. The principle of parsimony was therefore necessary to prevent any unwanted or unimportant regressors from entering the model. The principle of parsimony states that we need to choose the model with the fewest regressors that still yield a good fit. This model would therefore also have the fewest degrees of freedom, the most generalizability

from the sample to the population, and be the simplest in interpolation. The principle of parsimony ensures that the conclusion has strength and applicability to the population.

### 3.2 Methods of Goodness of Fit Assessment

The simple regression model  $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$  and its estimate  $\hat{y}_i = b_0 + b_1 x_i$  may be obtained by means of the method of ordinary least squares (OLS). The regression coefficients  $\beta_0$  and  $\beta_1$  are estimated in the sense of least squares as  $b_0 = \bar{y} - b_1 \bar{x}$  and  $b_1 = (\sum xy - n \bar{x} \bar{y}) / (\sum x^2 - n \bar{x}^2)$ . There are several methods of assessing goodness of fit for simple regression as listed below (Wayne, 1986, p. 381):

- (1) Standard error of the mean estimate of y:  $S_e = \sqrt{(y^2 - b_0 \sum y - b_1 \sum xy) / (n-2)}$
- (2) Standard error of the prediction:  $S_p = S_e \sqrt{1 + 1/n + (\bar{x} - x_0)^2 / (\sum x^2 - n \bar{x}^2)}$  where  $x_0$  is the specific value of x at which we want to predict the value of y.
- (3) Standard error of the regression coefficient:  $S_b = S_e / \sqrt{(\sum x^2 - n \bar{x}^2)}$
- (4) Squared correlation coefficient:  $R^2 = (b_0 \sum y - b_1 \sum xy - n \bar{y}) / (\sum y^2 - n \bar{y}^2)$

Any standard error is a measure of how well the parameters have been estimated from the data. Smaller standard error values lead to a model with better goodness of fit. Confidence intervals for parameter estimates of the mean of y, predicted value of y given a value of x, and the regression coefficient (slope) can be computed using the standard error. These confidence intervals tend to be narrower when standard error estimates are smaller. Even though confidence intervals are not measures of goodness of fit, a narrow confidence interval is a strong indication that the parameter is properly measured. There are several types of confidence intervals (Wayne, 1986, p. 384):

- (5) Confidence interval for estimate of the regression coefficient:  $b_1 \pm (t_{\alpha/2, n-2}) S_b$
- (6) Confidence interval for prediction of an observation:  $\hat{y} \pm (t_{\alpha/2, n-2}) S_e$

(7) Confidence interval for estimate of the mean:  $\bar{y} \pm (t_{\alpha/2, n-2})S_e$

(8) Prediction interval for an exact observation  $x_0$ :  $\hat{y} \pm (t_{\alpha/2, n-2})S_p$

When a multiple regression analysis is computed, there is at least one regressor in the model. The regression equation is  $y_{ij} = \beta_0 + \beta_j x_{ij} + \varepsilon_{ij}$  for  $i = 1, 2, \dots, n$  and  $j = 1, 2, \dots, k$  with the ordinary least squares estimation of the hyperplane being  $\hat{y}_{ij} = b_0 + b_j x_{ij}$  (Neter et al, 1996, p. 218). The objective in fitting a multiple regression model involves finding the best hyperplane that will minimize the sum of the squared error. According to the least squares principle, the standard errors presented in the standard (simple) linear regression model have direct counterparts under the multiple regression model. For the latter model, however, the formulas are often expressed in matrix notation. Properties of the confidence intervals discussed above are also applied to the multiple regression models. The best fit for several different models is determined by the highest  $R^2$  value. Because adding an additional variable into the equation always increases  $R^2$ , the  $R^2_{adjusted}$  statistic is used to account for degrees of freedom (Neter et al., p. 231).

For multiple regression, the most common methods for model selection are the forward, backward and stepwise procedures. Cook and Weisberg (1999, p. 275-283), Chatterjee and Price (1991, p. 236-237) and Chatterjee and Hadi (1988, p. 49-51) provide an excellent description of the three procedures, as well as indices for model fitting criteria and partial residual plots for model diagnostics. In forward regression, a statistical package such as SPSS or SAS uses a correlation matrix to find the regressor  $x_1, x_2, \dots, x_j$  with the highest correlation to the criterion variable. The regressor is then placed in the model and a t-ratio significance test is conducted to see if the corresponding regression coefficient  $\beta_j$  accounts for a significant amount of the criterion variance. If the

t-ratio is significantly different from zero, then the regressor is selected to be retained in the model. A partial correlation matrix between the criterion and the remaining regressors is then computed to determine if a second regressor is to be selected for the model. The regressor with the highest partial-correlation becomes the second regressor in the model and a second significance test is conducted to determine if a third regressor is required. When the addition of another variable no longer produces a significant t-ratio, the procedure comes to an end. Generally, a variable is added to the model as long as it contributes a significant increase in  $R^2$  of the model. For backward selection, the model is estimated with all the regressors and then the procedure successively drops one regressor at a time. When each of the regression coefficients in the model is significant, the procedure comes to an end. Those regressors with the smallest  $R^2$  value (or lowest residual sums of squares) are first deleted from the model. The stepwise procedure is a combination of both the forward and backward procedures since regressors are added or dropped from the model in successive steps. A forward selection is conducted, but regressors are deleted when they are no longer significant after the introduction of new regressors to the model. Chatterjee and Hadi (1988, p 84-87) provide an excellent discussion of residual plots used to determine which regressors should be added or deleted in a regression equation.

There is also a stepwise method for isotonic regression. The objective, from the principle of parsimony, is to determine the minimum number of regressors that yield a model with a significant  $F$  statistic. The regressors in this case would be population means  $\mu_j$  and not regression coefficients  $\beta_j$  as in OLS multiple regression. A full model with all of the regressors is first obtained, and then models with fewer regressors are

computed. For example, when  $k = 4$ , the  $\bar{E}^2$  statistic would be computed on all possible orderings of a model with four regressors. If significant, the  $\bar{E}^2$  statistic would be computed on all possible orderings of three regressors. The algorithm for isotonic regression then comes to an end as orderings of two regressors require a different statistic (namely the directional t-test). The isotonic regression stepwise method therefore deletes one regressor at a time to determine the model with the highest  $\bar{E}^2$  statistic. An equivalent alternative method to determine which regressor should be dropped from the model is the use of the p-value. Those models with p-values below the pre-determined significance level ( $\alpha = 0.05$  or  $0.01$ ) are retained. Models with p-values above the significance level are removed. The p-value may be obtained by linear interpolation when using the tables of the  $\bar{E}^2$  critical values in the appendix.

### 3.3 Using Isotonic Regression for Assessing Monotonic Linear and Non-Linear Trends

We can now extend the discussion pertaining to the goodness of fit tests to isotonic regression. The main differences are in the form of the regression equations between simple, multiple and isotonic regression models. The first main difference is that whereas the predictor can be continuous or categorical in simple and multiple linear regression, the regressor must be categorical in the isotonic regression context (Barlow et al, 1972, p.2). Most commonly, the regressors in simple and multiple regression models are continuous. The most pertinent implication of continuous versus categorical variables is that the unit of measurement is different. With the former, the unit of observation is the subject or participant in the study. With the latter, it is the group mean estimates of the populations. Further, the main parameters to be estimated in simple and multiple

regression are the regression coefficients,  $\beta_j$ , whereas the main parameter for isotonic regression is the isotonic regression estimate,  $\mu_j^*$ .

The second main difference between the three types of regressions is the objective of the study. A typical hypothesis for simple linear regression is  $H_0: \beta_1 = 0$  against  $H_1: \beta_1 \neq 0$  where the objective is to determine if the slope is significantly different from a pre-specified cut-off value (Neter et al. 1996, p. 51). For multiple regression, the objective is to determine the model with the fewest significant regression coefficients without compromising the fit of the model. For example, consider a model with one regressor, and the hypothesis to be tested is  $H_0: y_i = \beta_0 + \beta_1 x_{i1} + \varepsilon_i$  versus  $H_1: y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$ . The objective is to determine if  $x_2$  will be a significant regressor for predicting the criterion variable once  $x_1$  has already been added to the model (Neter et al, 1996, p. 261). The regressors  $x_1$  and  $x_2$  may have been selected using the forward, backward or stepwise procedures. On the other hand, the isotonic regression hypotheses, as stated in Section 2.1, are  $H_0: \mu_1 = \mu_2 = \dots = \mu_k$  and  $H_1: \mu_1 \leq \mu_2 \leq \dots \leq \mu_k$  or  $H_1: \mu_1 \geq \mu_2 \geq \dots \geq \mu_k$  (Barlow et al, 1972, p. 2). The alternative hypothesis tests whether there is a non-decreasing or non-increasing trend in the data. Examples in psychology include trends of stress and anxiety reduction, enhanced learning, progressive long-term memory development, intelligence across the life span and academic improvement at an educational institution. Isotonic regression has the additional assumption of preserving monotonicity, and multiple regression and isotonic regression also require the assumption of multivariate normality.

As the third main difference, isotonic regression may be used to test for non-linear patterns of group means. Isotonic regression analyzes several linear segments tip-to-tail in order to produce a piecewise non-linear trend. In other words, isotonic regression does

not require the continuity of the regression lines as in simple and multiple linear regression models. The additional flexibility enables isotonic regression to test a variety of different data sets.

Besides comparing simple and multiple regression models, properties of isotonic regression models can also be illustrated by comparing them against other non-linear regression models. There are several different types of non-linear models that may be used in place of isotonic regression. The two and three parameter exponential models as well as logistic model are examples of such non-linear models (Neter et al, 1996, p. 532, 533). All of these models can be used for testing trends. The difference between the non-linear models and isotonic regression is that, whereas exponential and logistic

**Figure 3.2**

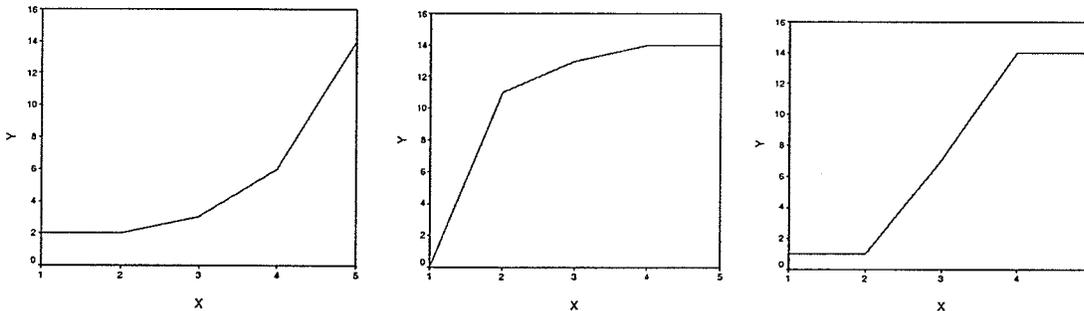


Figure 3.2. The (a) two-parameter exponential model, (b) three-parameter exponential model and (c) logistic model are monotonic functions that may explain a trend between population means (Neter et al, 1996, p.533).

models test for a particular type of monotonic trend, isotonic regression tests if a monotonic trend is present between categorical population means. Further, isotonic regression techniques involve differences between sample means that are increasing (or decreasing) whereas the other models see how well observations fit specified functions of a trend. In other words, the application of an isotonic regression technique does not

require the specification of a function, as illustrated in Figure 3.2.

The figure above demonstrates that monotonic trends may exist for both exponential and logistic models. In the first plot, (a), the trend is an exponential function that does not increase initially, but then gradually increases to display a concave form. The second plot, (b), has a trend with an immediate increase in slope but then gradually tapers to a uniform level. The final plot, (c), demonstrates a trend with no slope initially, but then gradually increases until the end where it levels out once more. In all of these cases, isotonic regression can be used to evaluate the monotonicity of the trend. In summary, the simple, multiple and isotonic regression models test different hypotheses, require different types of data and have different objectives and assumptions.

### 3.4 Illustrative Examples

This section presents an example from Keppel (1973, p. 70) for comparing the results by fitting simple, multiple and isotonic regression models to the same data set. Each of these three techniques has different hypotheses and objectives. The example consists of determining the effect of sleep deprivation on the ability of subjects to locate objects moving on a radar screen. Four different hours of sleep deprivation ( $x_1 = 4 \text{ hrs}$ ,  $x_2 = 12 \text{ hrs}$ ,  $x_3 = 20 \text{ hrs}$  and  $x_4 = 28 \text{ hrs}$ ) are tested for their effects on vigilance ( $y$ ) with subjects taking a simulated object detection test. The simple regression component involves estimating the linear equation  $y_{ij} = \beta_0 + \beta_1 x_{ij} + \varepsilon_{ij}$  for  $i = 1, 2, \dots, n_j$  and  $j = 1, 2, \dots, 4$  using the least squares principle. If there is a strong linear association between  $x$  and  $y$  then the omnibus  $F$  will be significant and then the goodness of fit measures  $R^2$  and  $S_e$  will be large and small respectively. For multiple regression, the model  $y_{ij} = \beta_0 + \sum \beta_j x_{ij} + \varepsilon_{ij}$  for  $i = 1, 2, \dots, n$ ,  $j = 1, \dots, 3$  is also estimated using the least squares principle

where  $x_{ij}$  is a dummy variable representing hours of sleep deprivation. The stepwise procedures are used to determine which of the three regressors  $x_1, x_2$  and  $x_3$  should be included in the model. The difference between this model and the simple regression model is the number of parameters to be estimated. Whereas the simple regression model estimates one regression coefficient,  $\beta_1$ , the multiple regression model must determine the best combination of regression coefficients,  $B_j$  for  $j = 1, 2, \dots, 4$ , to detect objects moving on a radar screen. For isotonic regression, the null hypothesis of no trend,  $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ , is tested against the alternative hypothesis of a non-decreasing trend,  $H_1: \mu_1 \leq \mu_2 \leq \mu_3 \leq \mu_4$  using the  $\bar{F}$  statistic. A stepwise procedure is used to determine both the smallest number of population means and their order. The process begins by determining if there is a non-decreasing trend among  $\mu_1, \dots, \mu_4$ , then a population mean is dropped and the isotonic regression analysis is recomputed. The process continues until there are too few population means or each ordering of the population means is non-significant. Since there are 24 possible combinations of the population means  $\mu_1, \dots, \mu_4$  and 18 possible combinations of the population means  $\mu_1, \dots, \mu_3$ , the ordering of the population means is predetermined. The criterion variable data ( $y$ ) is presented below in Table 3.1 along with a description of the coding of the regressors ( $x$ ) for simple and multiple regression.

The analysis for simple regression consists of several components. The fitted regression equation is  $\hat{y} = 14.5 + 12.55x$ , the estimate of the slope is  $b = 12.55$  with a standard error of  $S_b = 2.65229$ . The t-ratio is significant for the intercept ( $p < 0.0001$ ) and the slope ( $p < 0.0003$ ). Consequently, the  $R^2$  value for the model is very high, 0.6153 ( $F = 22.39, p = 0.003$ ). For this data set, there is a significant difference between

Table 3.1

Vigilance Scores are Measured for Hours without Sleep

	<u>Hours without Sleep</u>			
	4	12	20	28
	37	36	43	76
	22	45	75	66
	22	47	66	43
	25	23	46	62
Sum:	106	151	230	247
Mean:	26.5	37.75	57.50	61.75

Coding for Simple Regression

<u>Y</u>	<u>X</u>
37	1
22	1
22	1
25	1
36	2
45	2
47	2
23	2
43	3
75	3
66	3
46	3
76	4
66	4
43	4
62	4

Coding for Multiple Regression

<u>Y</u>	<u>X<sub>1</sub></u>	<u>X<sub>2</sub></u>	<u>X<sub>3</sub></u>
37	1	0	0
22	1	0	0
22	1	0	0
25	1	0	0
36	0	1	0
45	0	1	0
47	0	1	0
23	0	1	0
43	0	0	1
75	0	0	1
66	0	0	1
46	0	0	1
76	0	0	0
66	0	0	0
43	0	0	0
62	0	0	0

Note: The data consists of vigilance scores for a simple, multiple and isotonic regression with one regressor for simple regression and four regressors for multiple regressions.

hours of sleep deprivation when the dependent variable is vigilance toward objects moving on a radar screen. As hours of sleep deprivation increase so do the number of objects that fail to be detected on a radar screen. Very similar results are found for the

multiple regression analysis. First, Table 3.2 provides a point-biserial correlation matrix of the criterion variable (vigilance) with the three regressors (hours of sleep),  $x_1$ ,  $x_2$  and  $x_3$ , each is a dummy variable that represents amount of sleep. The criterion variable correlates highly with  $x_1$  ( $r = -0.6321$ ),  $x_2$  ( $r = -0.2649$ ) and  $x_3$  ( $r = 0.4186$ ). Notice also that  $x_3$  correlates with  $x_1$  and  $x_2$  ( $r = -0.3636$ ). Similarly, the same correlations between  $x_1$  and  $x_2$  are observed for  $x_4$  ( $r = -0.3015$ ). In other words, there are significant differences among the effects of the four groups of sleep deprivation for  $x_3$  and  $x_4$ . This observation is reconfirmed from the multiple regression analysis of  $y$ , against  $x_1$ ,  $x_2$  and  $x_3$ . The

**Table 3.2**

**Point-biserial Correlation Matrix of Vigilance and Hours of Sleep Deprivation**

	Variables				
	VIGILANCE (y)	4 HRS ( $x_1$ )	12 HRS ( $x_2$ )	20 HRS ( $x_3$ )	28 HRS ( $x_4$ )
1. VIGILANCE	1.00	-0.6321	-0.2649	0.4186	0.5184
2. 4 HRS	-0.6321	1.00	-0.3636	-0.3015	-0.3636
3. 12 HRS	-0.2649	-0.3636	1.00	-0.3015	-0.3636
4. 20 HRS	-0.4186	-0.3015	-0.3015	1.00	-0.3015
5. 28 HRS	0.5184	-0.3636	-0.3636	-0.3015	1.00

Note: The criterion variable is vigilance and the regressors are the number of hours of sleep deprivation. The correlations are significant at the 0.05 level of significance.

predicted regression equation is  $\hat{y} = 61.75 - 35.25x_1 - 12x_2 - 1.42x_3$  with p-values of the regression coefficients at 0.0096, 0.0102 and 0.6630 respectively. Each procedure concludes that the regressors  $x_1$ ,  $x_2$  and  $x_3$  should be included in the model.

For the isotonic regression component, there are 24 possible orderings for the population means  $\mu_1, \dots, \mu_4$  so a strategy was developed to reduce the number of  $\bar{E}^2$  tests that would be computed on the population means. The strategy involved placing the population means in ascending order and then randomizing the order of the first three population means while maintaining that the last population mean is always largest. The number of possible orderings is then reduced to 6 models and there are fewer  $\bar{E}^2$  tests to

**Table 3.3****Isotonic Regression Models for Vigilance Regressed on Hours of Sleep Deprivation**

Model	$\bar{E}^2$
1. $\mu_1 \leq \mu_2 \leq \mu_3 \leq \mu_4$	0.7844
2. $\mu_1 \leq \mu_3 \leq \mu_2 \leq \mu_4$	0.7006
3. $\mu_2 \leq \mu_1 \leq \mu_3 \leq \mu_4$	0.7625
4. $\mu_2 \leq \mu_3 \leq \mu_1 \leq \mu_4$	0.5124
5. $\mu_3 \leq \mu_1 \leq \mu_2 \leq \mu_4$	0.5124
6. $\mu_3 \leq \mu_2 \leq \mu_1 \leq \mu_4$	0.5124

Note: The isotonic regression technique produces a significant monotonic trend for the order of the population means  $\mu_1, \dots, \mu_4$  at the 0.05 significance level.

be computed. Those models that are significant are then tested again while randomizing the order with one fewer population mean. The data is the same as the simple and multiple regression analyses with the exception that sample means are estimated and not regression coefficients. The analysis is presented in Table 3.3.

There is enough evidence to suggest that there is a non-decreasing trend of vigilance for different hours of sleep deprivation. The critical value taken from the appendix for this analysis with  $k = 4$  factor levels and  $n = 4$  at  $\alpha = 0.05$  is 0.2923. This value is much smaller than the smallest  $\bar{E}$  value in Table 3.4. Nelson (1976) reports  $\bar{E}$  values for a three level factor as " $B = T / SS_{Total}$ " where  $T$  is the between groups sums of squares taken after amalgamation,  $\sum n_i(y_i^* - \bar{y})^2$ , and  $SS_{Total}$  is  $\sum \sum (y_{ij} - \bar{y})^2$ . The analysis above is the same as Nelson's (1976) procedure, taking into account the most possible orderings for a factor with four levels. To summarize briefly, simple regression suggests that there is a linear association between hours of sleep deprivation and vigilance. As the hours of sleep deprivation increase so do the number of objects not detected on a radar screen. Multiple regression suggests that the model with  $x_1$ ,  $x_2$  and  $x_3$  are the best predictors of vigilance, and isotonic regression suggests that there is a non-decreasing trend of vigilance across hours of sleep deprivation. Each type of analysis provides information about the association between vigilance and hours of sleep deprivation.

## Chapter Four: Isotonic Regression, ANOVA and Non-Linear Regression

## in the study of Intrinsic Religious Orientation

## 4.1 Purpose of the Study

The purpose of this study is to apply and compare several statistical techniques in analyzing an empirical data set in psychological research. For these objectives, seven ANOVAs and isotonic regressions with Scheffe post hoc multiple comparison procedures are investigated. To illustrate the importance of the current study, Hossack (1996) proposed a model that would integrate “trait” and “process” approaches to personality development in a sample of undergraduate Psychology students at the University of Manitoba. In the primary analysis, a structural equation model demonstrated how the interaction of traits and interpersonal variables tend to influence parental attachment styles, providing adults with a sense of well-being (happiness). In the secondary analysis, the relationship between Intrinsic Religious Orientation (IRO), several personality traits and parental attachment types was explored. Specifically, IRO was correlated with Extraversion ( $r = 0.12, p < 0.05$ ), Attachment to Mother ( $r = 0.22, p < 0.01$ ), Attachment to Father ( $r = 0.16, p < 0.01$ ), and Satisfaction with Life ( $r = 0.19, p < 0.01$ ). This study examined how trends that have the property of being monotonic could be used to describe the relation between IRO and parental attachment types, three NEO subscales (Neuroticism, Agreeableness and Conscientiousness), as well as Satisfaction with Life. In particular, the methods of ANOVA and isotonic regression are illustrated with these variables. The hypotheses for this study are presented in section 4.3.

## 4.2 Methodology

Participants: There were 520 Introductory Psychology students, 258 males and 262 females from the University of Manitoba. A majority (85.6%) were within the age range of 18 to 22 years. These students were placed into one of ten different groups based on their level of Intrinsic Religious Orientation (Allport & Ross, 1967).

Materials: The undergraduate psychology students were measured on seven different measures: Attachment to Mother, Attachment to Father, Neuroticism, Agreeableness, Conscientiousness, Sense of Coherence, and Satisfaction with Life. A description of these variables and Intrinsic Religious Orientation is presented below.

*Intrinsic Religious Orientation (Allport & Ross, 1967):* IRO is a 20-item scale that measures an individual's motive for their behavior and way of life as it relates to their belief in God. Other needs and values are brought into harmony with this belief in God. A sample item would be: 'My religious beliefs are what really lie behind my whole approach to life.' This item would be responded to according to a 5-point Likert scale: (a) this is definitely not so, 5 (b) probably not so, 4, (c) neutral, 3, (d) probably so, 2, (e) definitely so, 1. A score of 1 indicates high intrinsic religiosity whereas a score of 5 indicates a score of low intrinsic religiosity (or high extrinsic religiosity). This subscale has good internal consistency reliability ( $\alpha = .83$ ) and correlates highly with variables such as mental health, altruism and religious commitment (Trimble, 1997). The intrinsically religious person considers religion to be a way of life.

*Attachment to Mother (Armsden & Greenberg, 1987):* The ATM subscale contains 25 items that ask about feelings toward their mother and about their tendency to approach their mother when they are upset or have to make a decision. Items include “My mother helps me understand myself better” and “I tell my mother about my problems and troubles.” The items are responded to using a five-point Likert scale from (1) seldom true to (5) almost always or always true. The ATM construct is measured from the Inventory of Parent and Peer Attachment (IPPA) which has three subscales (trust, communication & alienation).

*Attachment to Father (Armsden & Greenberg, 1987):* The ATF subscale contains 25 items pertaining to how the individual views, thinks and feels about their father. The word “father” replaces the word “mother” for each item in the IPPA subscales. The scoring methodology for ATF is the same for the ATM construct.

*Neuroticism (Costa & McCrae, 1989):* The NEO personality inventory measures the big-five dimensions of personality: Neuroticism, Extraversion, Openness, Agreeableness and Conscientiousness. The NEU component contains 12 items that measure anxiety, hostility, depression, self-consciousness, impulsiveness and vulnerability. Reliabilities for this subscale are in the high .80s to low .90s for both internal consistency and test-retest reliability (Kaplan & Saccuzzo, 1997).

*Agreeableness (Costa & McCrae, 1989):* The AGR component of the NEO contains 12 items which measure the degree to which a person is warm and cooperative as opposed to

unpleasant and disagreeable. Agreeable individuals are communicative, supportive and tend to be very kindhearted.

*Conscientiousness (Costa & McCrae, 1989):* The CON component of the NEO is a 12 item subscale which measures the degree to which a person is persevering, responsible, and organized as opposed to lazy, irresponsible and impulsive. Conscientious individuals are concerned about others and plan their activities ahead of the scheduled time.

*Sense of Coherence (Antonovsky, 1987):* The SOC subscale has 29 items with 7-point likert scale ratings, where 1 is “strongly disagree” and 7 is “strongly agree.” The three components of the subscale (comprehensibility, manageability and meaningfulness) allow an individual to see the world as predictable, understandable and well structured. Comprehensibility is the ability to make cognitive sense of information that is ordered and consistent rather than random or accidental. Manageability is the extent to which an individual perceives that resources are available to attend to the information that is presented to them. A person with high manageability will not feel victimized by events or that life is unfair. Meaningfulness is the extent that an individual shapes their own destiny and daily experience.

*Satisfaction with Life (Diener, Larsen & Griffin, 1985):* The SWL subscale contains five items and is measured on a seven point likert scale. Items from the scale would be “So far I have gotten the important things I want in life” and “The conditions of my life are

excellent.” The retest reliability and internal consistency reliability for this scale are 0.82 and 0.87 (Diener et al, 1985).

### 4.3 Hypotheses, Procedure and Design

To illustrate the use of isotonic regression methodology, seven one-way ANOVAs and seven isotonic regressions were calculated on IRO, a ten-level factor. The null hypothesis  $H_0: \mu_1 = \mu_2 = \dots = \mu_{10}$  was tested against  $H_1: \mu_i \neq \mu_j, i \neq j, i \text{ and } j = 1, 2, \dots, 10$  for ANOVA and  $\mu_1 \leq \mu_2 \leq \dots \leq \mu_{10}$  for the isotonic regression. The first hypothesis is that, for ANOVA, there will be a significant difference between the 10 groups of Intrinsic Religious Orientation (IRO) for Attachment to Mother (ATM), Attachment to Father (ATF), Sense of Coherence (SOC) and Satisfaction with Life (SWL) but not Neuroticism (NEU), Conscientiousness (CON) and Agreeableness (AGR) at the 0.05 level of significance. The second hypothesis is similar to the first, in that for isotonic regression, there will also be differences between the 10 groups for ATM, ATF, SOC and SWL but not NEU, CON and AGR at the same significance level. The third hypothesis is that isotonic regression estimates will produce higher goodness of fit values ( $R^2$  values) than the sample mean estimates for linear, exponential and logistic models. This chapter represents a reanalysis of the data collected by Hossack (1996) as an empirical example of isotonic regression.

The criterion variables of interest were the two measures of attachment, the three NEO subscales, SOC and SWL. In all the statistical tests below, the alpha level was set at 0.05. It must be pointed out at the outset that all the  $F_{\max}$  tests for variance homogeneity are significant. Hence, the assumption of variance homogeneity is not

tenable. Strictly speaking, the ANOVA results would be invalid. However, for the purpose of illustrating the uses of isotonic regression, ANOVA designs will be taken into consideration. The descriptive statistics for the variables in the study are presented below

**Table 4.1**

**Descriptive Statistics for the Variables in the Study**

Variable	Descriptive Statistics		
	Mean	S.D.	Var
ATM	80.025	7.59	57.61
ATF	76.06	8.62	74.30
NEU	20.89	4.03	16.24
CON	29.37	3.49	12.18
AGR	25.02	4.00	16
SOC	92.05	6.36	40.45
SWL	17.68	3.87	14.98
IOR	27.37	9.59	91.97

**Note:** The descriptive statistics for the variables in the study ( $N = 520$ ) suggest that observations for the variables are close to the mean except for ATM, ATF, SOC and IOR.

in Table 4.1. The table demonstrates that the mean for SOC is largest, followed by ATM and ATF. The standard deviations for these variables are between 3.49 for CON and 9.59 for IOR. The variances for each of the variables are quite large and should therefore be applicable for statistical analyses.

When the IOR variable observations are placed into ten groups, the means, standard deviations and variances are mentioned in Table 4.2. The sample sizes were

**Table 4.2**

**Descriptive Statistics for Groups of Intrinsic Religious Orientation**

Group	Descriptive Statistics			
	Sample Size	Mean	S.D.	Var
1	50	12.18	1.35	1.8225
2	50	16.62	1.00	1.0
3	57	19.84	0.86	0.7396
4	55	23.04	0.84	0.7056
5	51	25.80	0.80	0.64
6	52	28.36	0.86	0.7396
7	53	30.98	0.77	0.5929
8	49	33.94	0.83	0.6889
9	52	38.44	1.70	2.89
10	51	45.12	1.89	3.5369

**Note:** The sample means for the ten groups of Intrinsic Religious Orientation are used to detect trends for the seven dependent measures. Sample sizes are approximately equal.

selected to be as equal as possible. The  $F_{\max}$  statistic indicates that the homogeneity of variance assumption is not supported for this analysis as  $F_{\max} = 91.97 / 12.18 = 7.55$  is greater than  $F_{(N-1, = 519, k = 8)} = 2.22$ . Table 4.2 shows that upon parsing the IRO variable into ten groups, the means of these groups were between 12.18 and 45.12 with standard deviations between 0.80 and 1.89. The standard deviations indicate that although there is

Table 4.3

**Correlations of Intrinsic Religious Orientation and Criterion Variables**

		Variables							
		IRO	ATM	ATF	NEU	CON	AGR	SOC	SWL
		Students ( $N = 520$ )							
1.	IRO	1.00	0.19**	0.16**	-0.07	0.25**	-0.03	0.20**	0.17**
2.	ATM	0.19**	1.00	0.44**	0.01	0.22**	-0.16**	0.14	0.23**
3.	ATF	0.16**	0.44**	1.00	-0.07*	0.14*	-0.18**	0.16**	0.20**
4.	NEU	-0.07	0.01	-0.07	1.00	-0.11*	0.28**	-0.06	-0.31**
5.	CON	0.25**	0.22**	0.14*	-0.11	1.00	0.06	0.22**	0.37**
6.	AGR	-0.03	-0.16*	-0.18**	0.28**	0.06	1.00	0.14	-0.24**
7.	SOC	0.20**	0.14	0.16**	-0.06	0.22**	0.14**	1.00	0.18**
8.	SWL	0.17**	0.23**	0.20**	-0.31**	0.37**	-0.24**	0.18**	1.00

\*  $p < 0.001$ , \*\*  $p < 0.0001$

Note: Correlations between IRO and the seven criterion variables indicate that ATM, ATF, CON, SOC and SWL are expected to produce significant monotonic trends. variation due to individual differences (randomization and sampling error) between the not much error in the ten groups, they have been varied widely. Prior to computing the ANOVA and isotonic regression analyses, correlations of the criterion variables with Intrinsic Religious Orientation have been examined for possibly identifying the monotonic trends that are significant at  $\alpha = 0.05$  (as seen in Table 4.3). Specifically, those variables that have a high correlation with Intrinsic Religious Orientation are

expected to produce significant linear trends. Results in the correlation matrix demonstrate that ATM, ATF, CON, SOC and SWL are expected to produce significant monotonic trends with IRO ( $p < 0.0001$ ). Additional findings suggest that ATM is highly correlated with ATF, CON and SWL ( $p < 0.0001$ ), but not with NEU and SOC ( $p > 0.05$ ). The SWL variable was significantly correlated with each of the remaining variables ( $p < 0.0001$ ). Non-significant correlations between IRO and both NEU and AGR suggest that there may be no trend between these variables, or a trend in the other direction.

#### 4.4 Results for the ANOVA and Isotonic Regression Analysis

Results for the ANOVA and isotonic regression analyses were summarized in Table 4.4. The omnibus hypothesis in  $H_0: \mu_1 = \mu_2 = \dots = \mu_{10}$  in the ANOVA analysis for five of the seven variables were significantly different at  $\alpha = 0.05$ , ATM ( $p < 0.01$ ), ATF ( $p < 0.05$ ), CON ( $p < 0.01$ ), SOC ( $p < 0.01$ ) and SWL ( $p < 0.01$ ). The p-value for CON is smaller than expected as it was hypothesized that there would be no significant difference between the IRO groups for this variable. Scheffe multiple comparison procedures at the 0.05 significance level indicate that there is a significant difference between groups one and ten for ATM and ATF. Similar results can be found for CON, where group ten is

Table 4.4

**Statistics for ANOVA and Intrinsic Religious Orientation and the Seven Criterion Variables**

Variable	Group										F	F(p-value)	$\bar{E}^2$	$\bar{E}^2$ (p-value)
	1	2	3	4	5	6	7	8	9	10				
ATM	77.0	79.28	79.47	79.69	79.55	80.60	79.85	80.12	80.19	84.51	3.13	0.01**	0.176	0.03*
ATF	73.36	75.16	75.05	76.25	75.73	77.63	76.43	74.92	76.19	79.78	2.10	0.026*	0.134	0.02*
NEU	20.30	21.56	21.26	21.51	20.90	20.92	20.96	20.37	21.17	19.80	1.01	0.434	0.057	0.0023**
CON	28.84	28.40	28.00	29.0	29.51	28.94	29.87	29.45	29.98	31.82	5.15	0.001**	0.023	0.06
AGR	25.10	25.88	24.39	24.87	24.47	25.69	25.13	25.69	24.31	24.75	1.06	0.369	0.237	0.005**
SOC	88.46	91.82	90.56	92.18	92.18	92.69	92.57	94.18	91.33	94.73	4.12	0.001**	0.196	0.05*
SWL	17.50	16.42	17.53	17.36	17.24	17.77	17.61	17.61	17.58	20.20	3.26	0.007**	0.151	0.02*

(\*\*) -  $p < 0.01$ , (\*) -  $p < 0.05$

Note: The ANOVA indicates that Attachment to Mother, Conscientiousness, Sense of Coherence and Satisfaction with Life are statistically significantly at  $\alpha = 0.01$  level. Attachment to Father is significant at  $\alpha = 0.05$ . The isotonic regression suggests that Neuroticism and Agreeableness are significant at  $\alpha = 0.01$  and Attachment to mother, Attachment to Father, Sense of Coherence and Satisfaction with Life are significant at  $\alpha = 0.05$ .

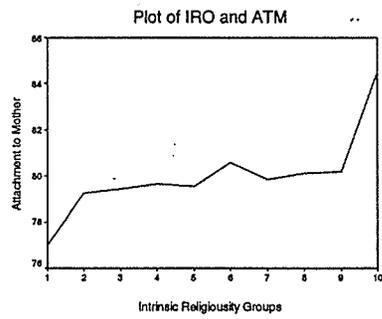


Figure 4.1

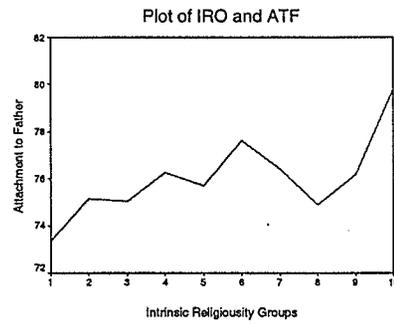


Figure 4.2

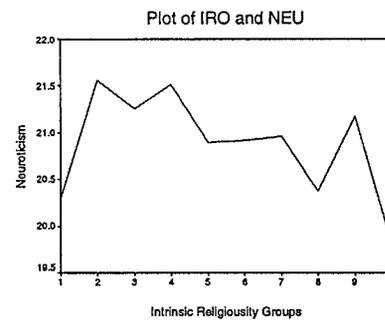


Figure 4.3

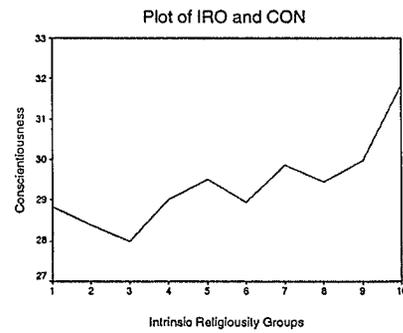


Figure 4.4

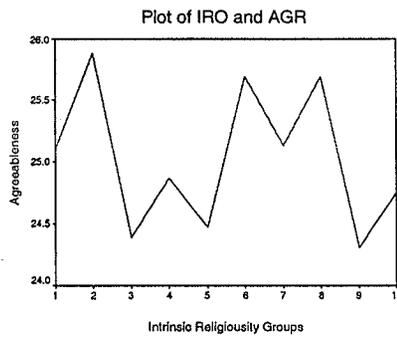


Figure 4.5

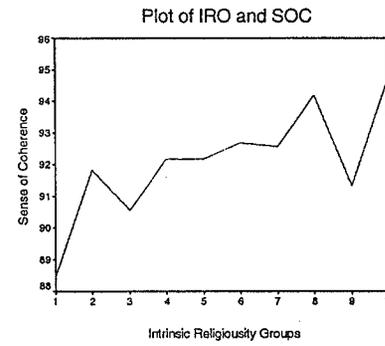


Figure 4.6

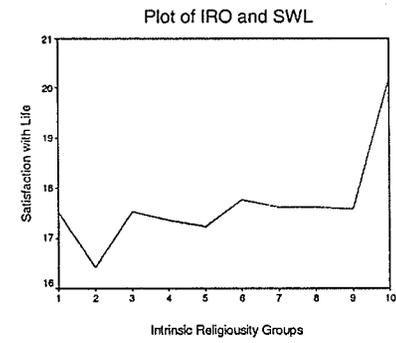


Figure 4.7

significantly different from groups one, two, three and six. The SOC variable also shows significant differences between groups one and six, as well as groups one and ten. The SWL variable demonstrates a significant difference between groups two and ten. Each of these comparisons maintain a mean square error of 14.39 and a critical value of  $F = 1.89$  using the SAS programming language. The Scheffe multiple comparison procedure has more power than most multiple comparison procedures so the results are useful for detecting group differences.

The plots of the variables in Figures 4.1 to 4.7 for the ANOVA analysis suggest that the monotonically increasing trend for three variables (ATM, CON and SWL) is quite likely, but nevertheless there is no clear trend for other variables. The purpose of these plots is to observe differences and trends between the sample means. The ATM variable increases from group 1 to group 2 and then remains steady until group 9, where it increases to group 10. The ATF variable increases from group 1 to group 6 and then decreases to from group 6 to group 8 before increasing to group 10. The NEU variable demonstrates a sharp incline from group 1 to group 2 and then gradually decreases until group 8, where it increases to group 9 and then diminishes to group 10. The CON variable decreases from groups 1 to 3, and then increases slowly until group 10. The AGR variable increases from group 1 to group 2, but then decreases sharply from group 2 to group 3, before increasing to group 6. The variable then decreases to group 7, before increasing to group 8 and finally decreasing to group 9. The SOC variable increases from group 1 to group 2 and then decreases from group 2 to group 3. The variable remains steady until group 8, where it decreases sharply to group 9, before increasing sharply to group 10. The SWL variable decreases initially from group 1 to group 2, and then increases for the remainder of the trend. The y-axis (level of the dependent variable) is not the same for each of these variables.

For the isotonic regression analysis, the planned multiple comparison procedures have the same significance level and a critical t-value of 3.053. The comparisons indicate that group 10 is significantly different from groups 1, 2 and 3 for the ATM variable and groups 1 and 3 for the ATF variable ( $p < 0.05$ ). The CON, SOC and SWL variables demonstrate that group 10 is significantly different from group 3 ( $p < 0.05$ ). The p-value for CON was smaller than anticipated as a non-significant result was expected. The comparisons were for those of the endpoints and tended to demonstrate that groups 8 and 9 were not significantly different from groups 1, 2 or 3 ( $p > 0.05$ ). Each variable had a potential nine multiple comparisons that could be performed (group 8, 9 and 10 compared to groups 1, 2 and 3 respectively) and no more than three were significant per variable. The alpha level was controlled to account for capitalization on chance by dividing 0.05 by the number of t-tests. The result,  $0.05/9 = 0.006$  suggest that none of the previous comparisons would be lower than this number. The possibility of rejecting a null hypothesis on the basis of chance and chance alone is therefore very small.

The planned comparisons have more power than post hoc comparisons since fewer tests are performed and the familywise error rate is lower. The tendency is for differences between groups to occur at the end points of the trends. Differences that occur in the middle of a curve can be identified from a plot and the number of amalgamations can be observed when the trend is constant. A critical evaluation of the plots allows the comparison of the different variables. There are trends the may be reproduced among several variables and there may be evidence to suggest that levels of IRO are different depending on the variable (this is a multivariate statistics problem).

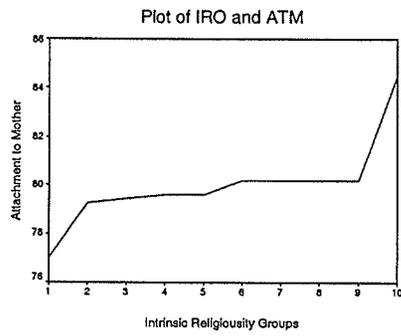


Figure 4.8

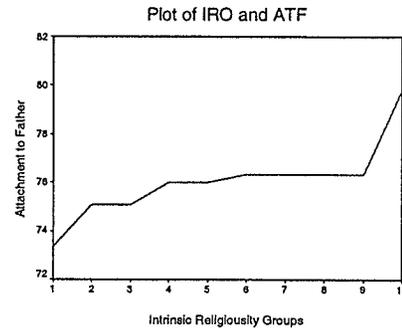


Figure 4.9

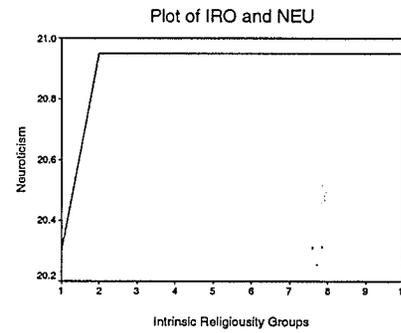


Figure 4.10

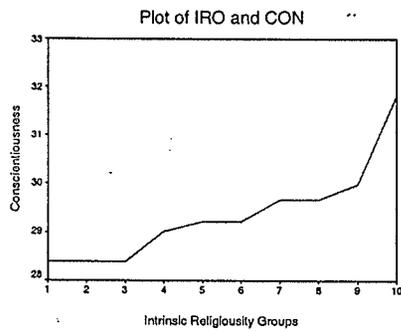


Figure 4.11

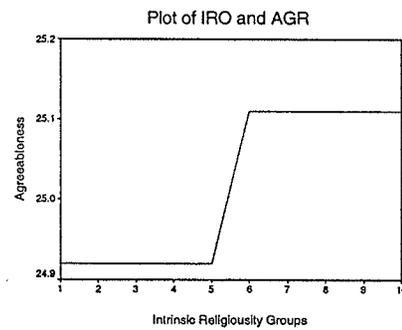


Figure 4.12

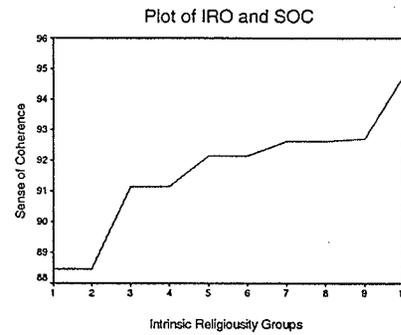


Figure 4.13

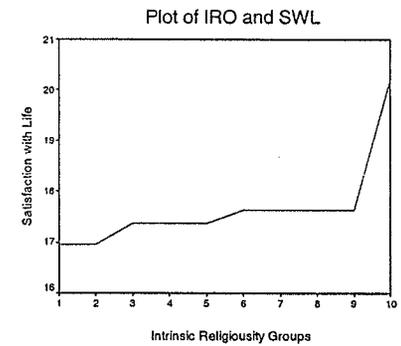


Figure 4.14

The group means after conducting the “amalgamation process” in the isotonic regression procedure were then estimated and plotted as seen in Figures 4.8 to 4.14. As the number of amalgamations increase, the slope of the trend becomes more level. The plots depict the number of amalgamations as well as the steepness of the slope. When the value of the slope is very low between the means and there are many amalgamations, then the null hypothesis of monotonic trend is not expected to be rejected. The ATM variable requires one amalgamation (groups 7 to 9 are combined) to make the trend monotonic. The trend tends to increase from group 1 to group 2 with not much change until group 9 where it increases to group 10. The ATF variable shows a very similar trend. There is also one amalgamation which occurs at the same location as the ATM variable. The variable tends to have slightly more of an incline from groups 1 to 6, and then remains steady until group 9 where it increases to group 10. The level of IRO tends to increase in a stepwise manner from group 2 to group 6. The NEU variable demonstrates several amalgamations (groups 2 to 10 are combined) after a sharp increase from group 1 to group 2. The trend remains constant until the 10th group. The CON variable does not have any amalgamations and nevertheless shows lower levels of the IRO variable than the NEU variable. The trend is uniform from group 1 to group 2, and then increases gradually from group 2 to group 10. The AGR variable demonstrates four amalgamations (groups 1 to 4 and 5 to 10 are combined) such that the trend remains constant until there is dramatic increase from group 5 to group 6. The trend then remains at the same level of IOR until group 10. The SOC variable does not show any amalgamations and increases gradually from the group 2 to group 10. The variable becomes greater as the trend progresses, with a gradual improvement from group 3 to group 9. The SWL variable has a very gradual increase from

group 1 to group 9, suggesting that most individuals are satisfied with their life regardless of their level of IRO. There is then a sharp increase from group 9 to group 10, which suggests that those individuals in the last group are also extremely satisfied with their life. These seven variables did not tend to demonstrate linear trends, however several variables such as the ATM, ATF, CON and SOC variables were increasing without amalgamations and have the potential to be fit by a linear trend. These variables demonstrate the property of monotonicity.

A more thorough investigation reveals that the ATM and SOC variables are the most strongly increasing of the variables. Both of these variables begin and end with very steep, increasing slopes with a tendency to be more flat and stable in the middle of the trend. The SWL variable is also similar, however the beginning of the trend has a steep declination. This declination, while being somewhat immediate, does not tend to damage the shape of a predominantly increasing trend in the middle and end. The same should be mentioned of the CON variable as its shape is the same as the SWL variable. The ATF variable is increasing at the beginning and end of the trend with a sharp declination in the middle. This variable is therefore similar to the ATM and SOC variables with a decline in the middle in place of being flat and stable.

#### 4.5 Results for ANOVA and Isotonic Regression Goodness of Fit Indices

The second purpose of this paper is to determine when goodness of fit values (measured using standard deviates) of isotonic regression estimates are greater than the goodness of fit values of sample mean estimates for three non-linear models. The results are summarized in Table 4.5. The goodness of fit values using isotonic regression

Table 4.5

**The Goodness of Fit for Three Models (Linear, Exponential and Logistic) using the Square of the Correlation Coefficient ( $R^2$ )**

	<u>Sample Mean Estimates</u>			<u>Isotonic Regression Estimates</u>		
	Linear	Exponential	Logistic	Linear	Exponential	Logistic
ATM	0.714	0.719	0.663	0.736	0.741	0.682
ATF	0.588	0.59	0.573	0.800	0.804	0.764
NEU	0.203	0.204	0.102	0.281	0.281	0.432
CON	0.732	0.734	0.59	0.871	0.882	0.747
AGR	0.048	0.048	0.042	0.685	0.685	0.663
SOC	0.591	0.589	0.654	0.872	0.868	0.907
SWL	0.51	0.514	0.374	0.629	0.647	0.501

Note: The  $R^2$  values are presented for ATM, ATF, NEU, AGR, CON, SOC and SWL. The regressor variable for each of the regressions is IRO. It is hypothesized that the  $R^2$  values will be greater for isotonic regression than for sample mean estimates.

estimates and sample mean estimates are not the same for each model. For the sample mean estimates, goodness of fit values for each of the models were high for ATM, ATF and SOC, moderate for CON and SWL, and low for NEU and AGR. The exponential model showed greater goodness of fit values than the linear and logistic models for each of the variables except AGR, where the linear model was slightly higher than the exponential model. Similar to the analysis using sample mean estimates, the goodness of fit isotonic regression values were high for ATM, ATF and SOC, moderate for CON and

SWL, and low for NEU and AGR. The exponential model had greater goodness of fit values than the linear and logistic models for each of the variables except AGR, where the logistic model was greater than the exponential model. The isotonic regression goodness of fit values were greater than the sample mean goodness of fit values for each of the variables except for NEU, CON and AGR (the three NEO subscales).

### 5.1 Advantages of Isotonic Regression

When considering the methods for testing directional hypotheses, isotonic regression is a sophisticated technique comparable to the standard ANOVA method and trend analysis method. The ANOVA method uses multiple comparison procedures after rejection of the omnibus F-test. The trend analysis method tests for specific functions, such as quadratic or cubic trends using contrasts. Isotonic regression, however, tests the equality of population means against order-restricted alternatives. The technique is different from other techniques in that it is a one-test procedure applicable in numerous mathematical models (Kabatsos & Ullrich, 2002). Isotonic regression can also be used for estimating confidence intervals of parameters, such as  $\mu$ , when the observations are subject to order restrictions (Hwang and Das Peddada, 1994). Further applications of isotonic regression can be found from Dykstra (1983), who presents an algorithm for restricted least squares regression when the data is constrained to an order restriction. This algorithm is useful for finding the closest quadratic function to a set of points while using the least squares principle. There is also a multivariate version of the  $\chi^2$  and  $F^2$  isotonic regression test statistics (Sasabuchi, Inutsuka and Kulatunga, 1983). Therefore, the isotonic regression technique has many uses for statistical inference.

Another advantage of isotonic regression is that the technique can provide a clear statistical decision on whether trends are non-decreasing or non-increasing (i.e. monotonic). The procedures that are available to make this decision are numerous. For example, Wang (1996) used a Monte Carlo simulation study to derive a likelihood ratio test for order restrictions when distributions are unknown under the null hypothesis. Also, there are

methods for estimating confidence bands of monotonic dose-response regression functions (Lee, 1996). The confidence bands provide an indication of how much error can be tolerated for making decisions about the curve. Kelly (1989) has shown that the absolute error of the isotonic regression estimate  $\mu_i^*$  is smaller than that of the sample mean estimate. Isotonic regression estimates are therefore preferable to sample mean estimates as a smaller error variance is expected in the  $\chi^2$  and  $\bar{E}$  tests. There have also been additional refinements to the amalgamation process for clearer decision making, e.g. Dykstra and Robertson (1982) present an algorithm for isotonic regression with two or more independent variables.

A third advantage to isotonic regression is that there are several studies pertaining to non-normal data, statistical power, and the Bayesian approach to order restricted statistical inference. When data are non-normal, Wang's (1996) approach of using a distribution free likelihood ratio test is appropriate for testing directional hypotheses. Other distribution free tests for non-normal data have been presented by Page (1963) for two independent variables and Joncheere (1954) for one independent variable. With respect to statistical power, Mancuso, Ahn and Chen (2000) have studied an empirical distribution of the test statistics and critical values for isotonic regression. They also discuss an adjustment to the likelihood ratio test to increase the power of isotonic regression for dose-response curves. Marcus (1976) concludes that the power of the likelihood ratio test is slightly greater than the William's and modified William's test (1971) for the minimum effective dose in the context of a one-way design. The Bayesian approach to order restricted statistical inference has been studied most recently by Dunson and Neelon (2003), who present non-parametric bayesian inference on order-constrained parameters in general linear models. An isotonic regression transformation of the data is used to derive a posterior distribution for the general

linear model. Deaton (1980) presents a parametric Bayesian approach to polynomial regression under order restrictions. A Monte Carlo procedure is used to demonstrate that the Bayes approach of estimating order-constrained parameters in a prior distribution has improved goodness of fit compared to other methods of polynomial regression. Isotonic regression is therefore useful for Bayesian statistical inference.

## 5.2 Disadvantages of Isotonic Regression

A potential disadvantage to isotonic regression is that the technique is used for testing the monotonicity of a trend, but not a specific function or shape of a trend. A researcher may use isotonic regression estimates to correct for reversals that may occur between population means. The potential difficulty is that amalgamations will change the shape of the trend despite preserving the order of the population means in the alternative hypothesis. If a particular shape of a trend is to be tested and an amalgamation is required then there is a compromise for the shape of the trend. Since researchers may be interested in testing for a particular shape (e.g. linear, quadratic and cubic), further research is required to preserve both the shape and sequence of the population means. Brownie, Boos and Hughes-Olivier (1990) present a technique for pooling the homogeneity of variance of two or more treatment groups with a control group. This procedure helps to increase power without compromising the shape of the trend. The modified t-test and ANOVA are the preferred techniques for this type of analysis.

Another disadvantage to isotonic regression is the lack of programming software for directional hypotheses. The prominent use of isotonic regression in psychology will rely on a simple procedure that can be implemented in statistical packages such as SAS and SPSS.

Nelson (1976) has developed a program using the Basic programming language, but a more practical program or module is necessary. A computer program to assess the monotonicity of a trend would allow users to analyse their data without too much complication.

As a third disadvantage, there are few empirical studies in the social sciences that apply isotonic regression to their data sets. There is nevertheless an application of isotonic regression by Deary, Caryl and Gibson (1993) for testing psychophysical responses (reaction times) in a visual inspection task. Some clinical studies that use isotonic regression consist of research pertaining to AIDS by Ancukiewicz, Finkelstein and Schoenfeld (2003), radiation dosage by Morton-Jones, Diggle, Parker, Dickinson and Binks (2000) and irreversible diseases such as cancer (Dinse and Lagakos, 1982). A survey of the research pertaining to isotonic regression reported in Table 5.1 reveals that the PsychInfo database contain only four references. The database with the most references is JSTOR (Journal Storage), followed by MathSciNet (the Math Science Network) and Medline. These three databases contain research papers from the disciplines of Statistics, Mathematics and Medicine. The Science database has the next most number of references pertaining to isotonic regression, predominantly in the discipline of Biology. From the results for PsychInfo, it is clear that social researchers have not exploited isotonic regression to detect monotonic trends. This technique should be made more available to social scientist, particularly in the discipline of Psychology.

**Table 5.1**

**Number of References using Isotonic Regression in Psychology and other Disciplines**

---

Isotonic Regression (Keyword)

1. PsychInfo	4
2. EconLit	5
3. ERIC	2
4. JSTOR	200
5. Medline	45
6. MathSciNet	123
7. Science	22

---

Note: The number of references in a key word search using the terms isotonic regression is recorded. The JSTOR, MathSciNet and Medline databases have the most references.

## 5.3 Summary

This thesis had three purposes. Firstly, it presented the theory, development and application of isotonic regression for the benefit of researchers in Psychology. The theory component involved discussing the alternative hypothesis for isotonic regression and the background for the use of isotonic regression estimates from the pool-adjacent-violators algorithm. The applied component was a study on the isotonic trend of groups varying in relation to the other variables. In this study, the criterion variables were the constructs of Attachment to Mother, Attachment to Father, Neuroticism, Conscientiousness, Agreeableness, Sense of Coherence and Satisfaction with Life. Separate ANOVA and isotonic regressions were computed on each of these seven criterion variables. Secondly, goodness of fit tests for isotonic regression models were discussed. The role and methods of goodness of fit procedures in assessing models using isotonic regression have been discussed and illustrated in the context of linear and non-linear models. Further, their goodness of fit was compared when isotonic regression estimates were used in place of

sample mean estimates. The intention was to demonstrate that isotonic regression estimates may be useful in situations other than hypothesis testing.

## References

Allport, G.W. & Ross, M.J. (1967). Personal Religious Orientation and Prejudice. Journal of Personality and Social Psychology, 5(4), 432-443.

Antonovsky, A. (1987). Unraveling the Mystery of Health. San Francisco: Jossey-Bass Publishers.

Ancukiewicz, M., Finkelstein, D.M., & Schoenfeld, D.A. (2003). Modelling the relationship between continuous covariates and clinical events using isotonic regression. Statistics in Medicine, 22, 3151 - 3159.

Armsden, G.C., & Greenberg, M.T. (1987). The inventory of parent and peer attachment: Individual differences and their relationship to psychological well-being. Journal of Youth and Adolescence, 16, 427-454.

Barlow, R.E., Bartholomew, D.J., Bremner, J.M. & Brunk, H.D. (1972). Statistical Inference under Order Restrictions. London: John Wiley & Sons.

Bartholomew, D.J. (1959a). A test of homogeneity for ordered alternatives. Biometrika, 46, 36-48.

Bartholomew, D.J. (1959b). A test of homogeneity for ordered alternatives II. Biometrika, 46, 328 - 335.

Brownie, C., Boos, D.D. & Hughes-Olivier, J. (1990). Modifying the t and ANOVA F tests when treatment is expected to increase variability relative to controls. Biometrics, 46, 259 - 266.

Cameron, A.J. (1970). A guide to graphs. New York: Pergamon Press.

Chacko, V.J. (1963). Testing homogeneity against ordered alternatives. Annals of Mathematical Statistics, 34, 945 - 956.

Chatterjee, S. & Hadi, A.S. (1988). Sensitivity analysis in linear regression. New York: John Wiley & Sons.

Chatterjee, S & Price, B. (1991). Regression analysis by example. New York: John Wiley & Sons.

Cook, R.D. & Weisberg, S. (1999). Applied regression including computing and graphics. New York: John Wiley & Sons.

Costa, P.T. & McCrae, R.R. (1989). The NEO PI/FFI manual supplement. Odessa, Florida: Psychological Assessment Resources.

- Deary, I.J., Caryl, P.G., & Gibson, G.J. (1993). Nonstationarity and the measurement of psychophysical response in a visual inspection-time task. Perception, 22(10), 1245 - 1256.
- Deaton, L.W. (1980). An empirical Bayes approach to polynomial regression under order restrictions. Biometrika, 67, 111 - 117.
- Diener, E, Emmons, R.A., Larsen, J. & Griffin, S. (1985). The satisfaction with life scale. Journal of Personality Assessment, 49, 71-75.
- Dinse, G.E. & Lagakos, S.W. (1982). Nonparametric estimation of lifetime and disease onset distributions from incomplete observations. Biometrics, 38, 921 - 932.
- Dunson, D.B. & Neelon, B. (2003). Bayesian inference on order-constrained parameters in generalized linear models. Biometrics, 59, 286 - 295.
- Dykstra, R.L. (1983). An algorithm for restricted least squares regression. Journal of the American Statistical Association, 78, 837 - 842.
- Dykstra, R.L. & Robertson, T. (1982). An algorithm for isotonic regression for two or more independent variables. The Annals of Statistics, 10(3), 708-716.

Hicks, C.R. (1973). Fundamental Concepts in the Design of Experiments.

New York: Holt, Rinehart and Winston, p.49.

Hossack, R.C. (1997). Salutogenic and pathogenic orientations to life: Attachment, personality, sense of coherence and well-being in late adolescence-- A structural equation model. Winnipeg, Manitoba: Unpublished Ph.D. Dissertation.

Hwang, J.T.G. & Das Peddada, S. (1994). Confidence interval estimation subject to order restrictions. The Annals of Statistics, 22, 67 - 93.

Jonckheere, A.R. (1954). A distribution-free k-sample test against ordered alternatives. Biometrika, 41, 133-45.

Karabatsos, G. and Ullrich, J.R. (2002). Enumerating and testing conjoint measurement models. Mathematical Social Sciences, 43, 485 - 504.

Kelly, R.E. (1989). Stochastic reduction of loss in estimating normal means by isotonic regression. The Annals of Statistics, 17(2), 937 - 940.

Keppel, G. (1973). Design and analysis: A researcher's Handbook. Englewood Cliffs, NJ: Prentice-Hall, Inc.

- King, B.M. & Minium, E.M. (2003). Statistical Reasoning in Psychology and Education. Hoboken, NJ: John Wiley and Sons, Inc.
- Lee, C.C. (1996). On estimation for monotone dose-response curves. Journal of the American Statistical Association, 91, 1110 -1119.
- Lehman, H.S. (1995). Statistics in the Behavioural Sciences: A conceptual Introduction. Pacific Groves: Brooks/Cole Publishing Company.
- Leuraud, K. & Benichou, J. (2001). A comparison of several methods to test for the existence of a monotonic dose-response relationship in clinical and epidemiological studies. Statistics in Medicine, 20, 3335 - 3351.
- Mancuso, J.Y., Ahn, H. & Chen, J.J. (2001). Order-restricted dose-related trend tests. Statistics in Medicine, 20, 2305 - 2318.
- Marcus, R. (1976). The power of some tests of the equality of normal means against an ordered alternative. Biometrika, 63, 177-183.
- Maxwell, S.E. & Delaney, H.D. (1990) Designing experiments and analyzing data: A model comparison perspective. Belmont, CA: Wadsworth Publishing Company.

Morton-Jones, T., Diggle, P., Parker, L., Dickinson, H.O. & Binks, K. (2000).

Additive isotonic regression models in epidemiology. Statistics in Medicine, 19, 849 - 859.

Mukerjee, H. & Tu, R. (1995). Order-restricted inferences in linear regression.

Journal of the American Statistical Association, 90, 717 - 722.

Nelson, L.S. (1976). Ordered tests for a three-level factor. Journal of Quality

Technology, 8(4), 241 - 243.

Nelson, L.S. (1977). Tables for testing ordered alternatives in an analysis of

variance. Biometrika, 64, 335-338.

Neter, J, Kutner, M.H., Nachtsheim, C.J., & Wasserman, W. (1996). Applied Linear

Statistical Models. Boston: WCB McGraw-Hill.

Page, E.B. (1963). Ordered hypotheses for multiple treatments: A significance test

for linear trends. Journal of the American Statistical Association, 216 - 230.

Robertson, T., Wright, F.T., & Dykstra, R.L. (1988). Order Restricted Statistical

Inference. New York: John Wiley and Sons.

Ruberg, S.J. (1989). Contrasts for identifying the minimum effective dose. Journal of

the American Statistical Association, 84, 816-822.

Sasabuchi, M. I., Inutsuka, M., & Kulatanga, D.D.S. (1983). A multivariate version of isotonic regression. Biometrika, 70(2), 465- 472.

Shorack, G.R. (1967). Testing against ordered alternatives in model I analysis of variance: Normal theory and non-parametric. Annals of Mathematical Statistics, 38, 1740 – 1752.

Trimble, D. E. (1997). The Religious Orientation Scale: Review and meta-analysis of social desirability effects. Educational and Psychological Measurement, 57(6), 970-986.

Wang, Y. (1996). A likelihood ratio test against stochastic ordering in several populations. Journal of the American Statistical Association, 91, 1676 - 1683.

Wayne, D.W. (1986). Business Statistics: Basic Concepts and Methodology. Boston: Houghton Mifflin Inc.

Williams, D.À. (1971). A test for differences between treatment means when several dose levels are compared with a zero dose control. Biometrics, 27, 103-117.

Zumbo, B.D. & Hubley, A.M. (1998). A note on misconceptions concerning prospective and retrospective power. The Statistician, 47(2), 385-388.

## Appendix

Critical values for  $\bar{E}^2$  statistic involving k groups each containing n observations

$\alpha$	k = 3					k = 4				
	0.10	0.05	0.025	0.01	0.001	0.10	0.05	0.025	0.01	0.001
n										
2	.5402	.6866	.7898	.8781	.9707	.4626	.5900	.6897	.7872	.9205
3	.3355	.4554	.5567	.6646	.8367	.2944	.3923	.4783	.5748	.7483
4	.2421	.3373	.4232	.5219	.7052	.2154	.2923	.3630	.4467	.6137
5	.1891	.2673	.3401	.4272	.6020	.1697	.2327	.2919	.3641	.5163
6	.1551	.2212	.2840	.3608	.5227	.1399	.1932	.2439	.3069	.4443
7	.1315	.1886	.2436	.3120	.4608	.1190	.1651	.2094	.2650	.3894
8	.1141	.1643	.2132	.2747	.4116	.1036	.1441	.1834	.2332	.3463
9	.1007	.1456	.1895	.2453	.3716	.0917	.1279	.1632	.2081	.3117
10	.0902	.1307	.1706	.2215	.3386	.0822	.1149	.1469	.1879	.2833
11	.0816	.1185	.1551	.2020	.3109	.0745	.1043	.1336	.1713	.2596
12	.0745	.1084	.1421	.1856	.2873	.0682	.0955	.1225	.1573	.2396
13	.0686	.0999	.1312	.1716	.2670	.0628	.0881	.1131	.1455	.2224
14	.0635	.0927	.1218	.1596	.2494	.0582	.0817	.1051	.1353	.2074
15	.0592	.0864	.1136	.1491	.2339	.0543	.0762	.0981	.1264	.1944
16	.0554	.0809	.1065	.1400	.2203	.0508	.0714	.0920	.1186	.1829
17	.0520	.0761	.1003	.1319	.2081	.0477	.0672	.0866	.1118	.1727
18	.0490	.0718	.0947	.1247	.1972	.0450	.0634	.0817	.1057	.1635
19	.0464	.0680	.0897	.1182	.1874	.0426	.0601	.0775	.1002	.1553
20	.0440	.0645	.0852	.1124	.1785	.0405	.0570	.0736	.0952	.1479
24	.0365	.0536	.0710	.0983	.1500	.0336	.0475	.0613	.0795	.1241
30	.0291	.0428	.0568	.0752	.1210	.0268	.0379	.0491	.0637	.0999
40	.0218	.0320	.0425	.0565	.0915	.0201	.0284	.0368	.0479	.0755
60	.0144	.0213	.0284	.0378	.0615	.0133	.0189	.0245	.0320	.0506

$\alpha$	k = 5					k = 6				
	0.10	0.05	0.025	0.01	0.001	0.10	0.05	0.025	0.01	0.001
n										
2	.4056	.5178	.6104	.7075	.8605	.3623	.4623	.5476	.6409	.8012
3	.2614	.3448	.4196	.5063	.6727	.2353	.3082	.3745	.4529	.6098
4	.1925	.2576	.3181	.3909	.5424	.1740	.2306	.2836	.3481	.4860
5	.1523	.2055	.2557	.3177	.4521	.1380	.1841	.2279	.2823	.4026
6	.1260	.1708	.2137	.2673	.3869	.1143	.1532	.1905	.2373	.3431
7	.1074	.1461	.1835	.2306	.3377	.0976	.1312	.1636	.2046	.2988
8	.0936	.1277	.1608	.2028	.2996	.0851	.1147	.1433	.1798	.2645
9	.0829	.1134	.1430	.1809	.2690	.0755	.1018	.1275	.1603	.2372
10	.0744	.1019	.1288	.1633	.2442	.0678	.0916	.1148	.1446	.2150
11	.0675	.0926	.1172	.1488	.2234	.0615	.0832	.1045	.1318	.1966
12	.0618	.0848	.1075	.1366	.2060	.0563	.0763	.0958	.1210	.1811
13	.0570	.0783	.0992	.1263	.1910	.0519	.0704	.0885	.1118	.1678
14	.0528	.0726	.0922	.1175	.1781	.0482	.0653	.0822	.1040	.1563
15	.0492	.0678	.0860	.1098	.1668	.0449	.0609	.0767	.0971	.1463
16	.0461	.0635	.0807	.1030	.1568	.0421	.0571	.0719	.0912	.1376
17	.0434	.0597	.0760	.0970	.1480	.0396	.0537	.0677	.0859	.1298
18	.0409	.0564	.0717	.0917	.1410	.0374	.0507	.0640	.0812	.1228
19	.0387	.0534	.0680	.0869	.1330	.0354	.0481	.0606	.0769	.1165
20	.0368	.0507	.0646	.0826	.1266	.0336	.0457	.0576	.0731	.1109
24	.0306	.0422	.0538	.0690	.1061	.0279	.0380	.0480	.0610	.0929
30	.0244	.0338	.0431	.0553	.0854	.0223	.0304	.0384	.0489	.0747
40	.0183	.0253	.0323	.0416	.0644	.0167	.0228	.0288	.0368	.0563
60	.0122	.0169	.0216	.0278	.0432	.0111	.0152	.0192	.0245	.0377

$\alpha$	k = 7					k = 8				
	0.10	0.05	0.025	0.01	0.001	0.10	0.05	0.025	0.01	0.001
n										
2	.3283	.4185	.4972	.5855	.7464	.3007	.3830	.4557	.5391	.6973
3	.2142	.2791	.3388	.4102	.5547	.1969	.2555	.3098	.3753	.5133
4	.1588	.2091	.2563	.3143	.4406	.1463	.1915	.2342	.2869	.4032
5	.1262	.1670	.2059	.2545	.3633	.1163	.1531	.1881	.2320	.3313
6	.1046	.1391	.1721	.2137	.3087	.0966	.1275	.1572	.1947	.2810
7	.0894	.1191	.1478	.1841	.2683	.0825	.1092	.1349	.1676	.2438
8	.0780	.1041	.1295	.1617	.2372	.0721	.0955	.1182	.1472	.2152
9	.0692	.0925	.1152	.1442	.2125	.0639	.0849	.1052	.1312	.1927
10	.0622	.0832	.1037	.1301	.1924	.0575	.0763	.0947	.1183	.1744
11	.0565	.0756	.0944	.1185	.1758	.0522	.0694	.0862	.1078	.1592
12	.0517	.0693	.0865	.1088	.1618	.0478	.0636	.0790	.0989	.1465
13	.0477	.0640	.0799	.1005	.1499	.0441	.0545	.0678	.0850	.1263
14	.0442	.0594	.0742	.0934	.1396	.0409	.0545	.0678	.0850	.1263
15	.0413	.0554	.0693	.0873	.1306	.0382	.0508	.0633	.0794	.1182
16	.0387	.0519	.0650	.0819	.1228	.0358	.0477	.0594	.0745	.1110
17	.0364	.0489	.0612	.0771	.1158	.0336	.0449	.0559	.0701	.1047
18	.0343	.0462	.0578	.0729	.1095	.0318	.0424	.0528	.0663	.0990
19	.0325	.0437	.0548	.0691	.1039	.0301	.0401	.0500	.0628	.0939
20	.0309	.0415	.0520	.0657	.0989	.0286	.0381	.0475	.0597	.0893
24	.0257	.0346	.0434	.0548	.0828	.0238	.0318	.0396	.0498	.0748
30	.0205	.0277	.0347	.0439	.0665	.0190	.0254	.0317	.0399	.0600
40	.0154	.0207	.0261	.0330	.0501	.0142	.0190	.0238	.0300	.0452
60	.0102	.0138	.0174	.0220	.0335	.0095	.0127	.0159	.0200	.0303

$\alpha$	k = 9					k = 10				
	0.10	0.05	0.025	0.01	0.001	0.10	0.05	0.025	0.01	0.001
n										
2	.2778	.3535	.4211	.4998	.6536	.2586	.3286	.3918	.4660	.6147
3	.1824	.2359	.2857	.3462	.4759	.1701	.2194	.2654	.3217	.4437
4	.1357	.1769	.2158	.2642	.3720	.1267	.1645	.2004	.2451	.3455
5	.1080	.1414	.1733	.2134	.3048	.1009	.1316	.1609	.1978	.2825
6	.0897	.1178	.1448	.1790	.2580	.0838	.1096	.1344	.1658	.2388
7	.0767	.1009	.1243	.1541	.2236	.0717	.0939	.1154	.1427	.2067
8	.0670	.0883	.1089	.1352	.1973	.0627	.0822	.1011	.1252	.1822
9	.0595	.0785	.0969	.1205	.1765	.0556	.0730	.0899	.1116	.1629
10	.0535	.0706	.0873	.1087	.1596	.0500	.0657	.0810	.1006	.1473
11	.0486	.0642	.0794	.0990	.1457	.0454	.0597	.0737	.0916	.1344
12	.0445	.0588	.0728	.0908	.1340	.0416	.0547	.0676	.0840	.1236
13	.0410	.0543	.0672	.0839	.1241	.0384	.0505	.0624	.0777	.1144
14	.0381	.0504	.0624	.0780	.1155	.0356	.0469	.0579	.0722	.1064
15	.0355	.0470	.0583	.0729	.1080	.0332	.0438	.0541	.0674	.0995
16	.0333	.0441	.0547	.0684	.1014	.0312	.0410	.0507	.0632	.0935
17	.0313	.0415	.0515	.0644	.0956	.0293	.0386	.0478	.0596	.0881
18	.0296	.0392	.0486	.0608	.0904	.0277	.0365	.0451	.0563	.0833
19	.0280	.0371	.0461	.0577	.0858	.0262	.0346	.0427	.0533	.0790
20	.0266	.0353	.0438	.0548	.0816	.0249	.0328	.0406	.0507	.0752
24	.0221	.0294	.0365	.0457	.0682	.0207	.0274	.0339	.0423	.0628
30	.0177	.0235	.0292	.0366	.0548	.0166	.0219	.0271	.0339	.0504
40	.0133	.0176	.0219	.0275	.0412	.0124	.0164	.0203	.0254	.0380
60	.0088	.0117	.0146	.0184	.0276	.0083	.0109	.0136	.0170	.0254

## Glossary of Statistical Terms

Between-Subjects Design: An experimental design where the differences between two or more treatment groups are compared. Each treatment group has different subjects and is uncorrelated with each other, usually being sampled from different populations.

$\bar{E}^2$ : The test statistic that is used to determine when there is a significant non-decreasing or non-increasing trend. The statistic is the ratio of the between sums of squares after amalgamation and the total sum of squares. High values of this statistic suggest that a significant amount of the criterion variable can be accounted for by the order of the levels of the independent variable.

Isotonic Regression Estimate: An isotonic regression estimate replaces a sample mean estimate to preserve a non-decreasing or non-increasing trend. The estimate is the arithmetic average of the sample mean that violates the trend and the sample mean that precedes this mean. When an isotonic regression estimate is required, a curvi-linear trend is obtained.

Likelihood Ratio Test: A statistic that measures the ratio of the between sums of squares after amalgamation and the population variance. This statistic is used for theoretical purposes. The likelihood ratio test indicates the amount of dependent variable variation accounted for by the trend.

Monotonic: The term used to describe the shape of a trend. A monotonic trend occurs when the dependent variable is either non-decreasing or non-increasing over the levels of the manipulated variable. The trend is usually curvi-linear but may be linear.

Within-Subjects Design: An experimental design where subjects are measured on each treatment group. The treatment groups are correlated and the order of the treatments must be counter-balanced.