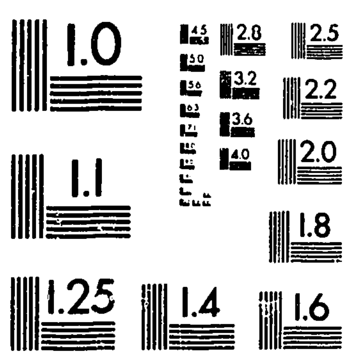
1

OF/DE

1

PM-1 3½"x4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT



PRECISION℠ RESOLUTION TARGETS

UNIVERSITY OF ALBERTA

Circumventing Analysis of Covariance's Homogeneity Assumption With the Johnson-Neyman Method and *Mathematica*

BY

JACQUELINE P. LEIGHTON ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree MASTER OF EDUCATION.

in

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

Edmonton, Alberta

Fall, 1995

ISBN 0-612-06405-0

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Jacqueline P. Leighton

TITLE OF THESIS: Circumventing Analysis of Covariance's Homogeneity Assumption With the Johnson-Neyman Method and *Mathematica*

DEGREE: Master of Education

YEAR THIS DEGREE GRANTED: 1995

8510 - 111 Street

Edmonton, Alberta

Canada T6G - 1H7

August 18, 1995

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommended to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled CIRCUMVENTING ANALYSIS OF COVARIANCE'S HOMOGENEITY ASSUMPTION WITH THE JOHNSON-NEYMAN METHOD AND *MATHEMATICA* submitted by JACQUELINE P. LEIGHTON in partial fulfillment of the requirements for the degree of MASTER OF EDUCATION.

S. Hunka

Steve Hunka

Jeffrey Bisanz

Jeffrey Bisanz

Maguire

Thomas Maguire

July 26, 1995

Abstract

This study describes two statistical methods, Analysis of Covariance (ANCOVA) and the Johnson-Neyman method. The Johnson-Neyman method is shown to be the better method to use when an important assumption in ANCOVA is not met--the homogeneity of within group regression slopes.

The Johnson-Neyman method is not as frequently applied as ANCOVA. Therefore, this study also demonstrates how researchers can apply the Johnson-Neyman method using *Mathematica*. *Mathematica* is the computer software package of choice when doing the Johnson-Neyman method because it can be used to manipulate interactively mathematical expressions in symbolic and numeric form. Three examples are presented that demonstrate how researchers would proceed in performing the Johnson-Neyman method using *Mathematica*.

Furthermore, a popular statistical package, SPSS, is used to perform a technique similar to the Johnson-Neyman method, an ANCOVA with heterogeneous within group regression slopes. The advantages and disadvantages of both packages are discussed.

Acknowledgement

I would like to thank my thesis and academic supervisor, Steve Hunka, for his knowledgeable and careful guidance throughout this project, yet unrestricted supervision.

I extend my gratitude to the Natural Science and Engineering Research Council of Canada (N.S.E.R.C.) for supporting this project and research throughout my Masters of Education training.

I also appreciate the support and patience provided by Greg Anderson and my family, once again.

Table of Contents

Chapter I 1

Introduction 1

What is Analysis of Covariance (ANCOVA)? 3

Assumptions of ANCOVA 6

History and Development of ANCOVA 11

Advantages and Limitations of ANCOVA 12

Why Is It Important to Understand ANCOVA
A Summary 13

Chapter II 15

An Introduction to the Johnson-Neyman Method 15

Assumptions of the Johnson-Neyman Method 16

History and Development of the Johnson-Neyman Method 17

Non-Simultaneous and Simultaneous Confidence Limits in
the Johnson-Neyman Method 18

Advantages and Limitations of the Johnson-Neyman Method 20

Importance of Understanding The Johnson-Neyman Method 22

Past Application of the Johnson-Neyman Method:
Non-Computer Based 23

Chapter III 24

Mathematica and Application of the Johnson-Neyman Method 24

The General Linear Model 25

Example 1:
Two Groups and One Covariate 31

Example 2:
Three Groups and Two Covariates 36

Example 3:
Two Way Analysis of Covariance and the Johnson-Neyman Method 43

Chapter IV 61

SPSS and Application of the Johnson-Neyman Method 61

Example 1:
Two Groups and One Covariate 62

Example 2:
Three Groups and Two Covariates 65

Comparison of Mathematica and SPSS
Advantages and Limitations 69

Chapter V 72

Summary and Discussion 72

References 76

Appendices 79

List of Tables

Table 4.1.
Analysis of Covariance. Customized Model....63

Table 4.2.
Analysis of Covariance. Regular Model....64

Table 4.3.
Analysis of Covariance. Customized Model....66

Table 4.4.
Analysis of Covariance. Regular Model....68

List of Figures

Figure 1.1.
Homogeneous Within Group Regression Lines 7

Figure 1.2.
Disordinal Within Group Regression Lines 8

Figure 1.3.
Ordinal Within Group Regression Lines...9

Figure 1.4.
Ordinal Within Group Regression Lines...10

Figure 3.1.
Region of Significance--Two Groups
and One Covariate 35

Figure 3.2.
Region of Significance--Three Groups and
Two Covariates 41

Figure 3.3.
Contour Plot--Alternative Representation
of Region of Significance...43

Figure 3.4.
Region of Significance--Factor B...58

Figure 3.5.
Region of Significance--Interaction AB.....................................60

Chapter I
Introduction

Research designs that make use of covariates are commonly employed by behavioural researchers (Maxwell, O'Callahan, & Delaney, 1993). Cochran (1957) suggests that some of the reasons for using covariates in a research design are (a) to increase precision in randomized experiments, (b) to remove effects of nuisance variables in observational or nonrandomized experiments, (c) to examine the true nature of a treatment effect, and (d) to analyze data, when observations are missing. The analysis of covariance (ANCOVA) is one statistical technique that is used to analyze data that contains covariates. Another statistical technique that can be used is the Johnson-Neyman method (Johnson & Fay, 1950). These two methods share two assumptions:

$$y_{ij} = \mu + \alpha_j + \beta\,(x_{ij} - \overline{X}..) + \varepsilon_{ij} \qquad (1)$$

$$\varepsilon_{ij}\text{s are NID } (0, \sigma^2) \qquad (2)$$

where: y_{ij} is observation i in group j,

μ is a constant or group effect,

α_j is a treatment effect,

$\beta\,(x_{ij} - \overline{X}..)$ is a covariate effect, and

ε_{ij} is the error.

The first of these assumptions is additivity or that an observation can be thought of as a sum of independent parameters, and the second assumption is that ε_{ij}s are normally distributed with a population mean of zero, variance σ^2, and are independent of each other (Bliss, 1970; Evans & Anastasio, 1968; Glass, Peckham & Sanders, 1972). Nevertheless, ANCOVA and the Johnson-Neyman method differ on a third assumption--the homogeneity of within group regression slopes. The use of

ANCOVA requires that the within group regression slopes be homogeneous i.e., the regression lines are parallel:

$$\beta_1 = \beta_2 = \beta_3 .... = \beta \qquad (3).$$

Whereas with the Johnson-Neyman method this assumption is not necessary and the within group regression slopes can be homogeneous or heterogeneous.

Researchers' ability to differentiate between these two methods will enable proper use of each method in the appropriate research situation. Ultimately, this will aid and improve the quality of behavioural research. In addition, before researchers can properly apply a statistical method, they must first understand the method and also have the resources, such as the computer software, to carry out the analysis. Without the resources, some statistical methods can be almost impossible to apply.

This study has three main purposes: (a) to briefly provide some background into ANCOVA and the Johnson-Neyman method by describing for each technique, its data organization, assumptions, history and development, advantages and limitations, and importance of its understanding; (b) to demonstrate how the Johnson-Neyman method can be performed with available computer software, such as *Mathematica* (Wolfram, 1988), because this statistical method is not as frequently or as easily performed as is ANCOVA; and (c) to show how a technique similar to the Johnson-Neyman method is performed with a widely used statistical software package, SPSS (SPSS, 1994), and compare SPSS and *Mathematica* in how well they handle the Johnson-Neyman method.

Furthermore, this study is geared towards those researchers who would like to use a statistical technique such as ANCOVA but whose data does not meet the assumption of homogeneity of within group regression slopes and, therefore need to use a less restrictive statistical method.

What is Analysis of Covariance (ANCOVA)?

Analysis of Covariance has been defined as:

> "[A] way to adjust the sample average [by] first [adjusting] the individual scores and then [computing] the average of the adjusted scores, [obtaining] an adjusted average. The method used for this type of adjustment employs the correlation between the stratifying variable and the dependent variable directly, through the corresponding linear regression equation. The resulting procedure is commonly referred to as the *analysis of covariance,* and the variable used for stratification...is called a *covariate*." (Marascuilo & Serlin, 1988, p. 598).

Although this definition is true, it is also not very helpful because of its generality. For example, the correlation can be used for adjusting the scores, but frequently it is the regression coefficient that is used. Moreover, one does not know for sure how the adjustment is made, and whether only one regression coefficient is used for all groups (because the within group regression slopes are all homogeneous) or a separate coefficient for each group, for example, $Y_{ij} - bX_{ij}$ or $Y_{ij} - b_jX_{ij}$. Nevertheless, this definition is presented here despite its limitations because it provides a general idea of the ANCOVA.

Before discussing the above definition by explaining each idea that is contained within the overall definition, it may be useful to first present an example of a situation where analysis of covariance is applied. The following example is taken from Marascuilo & Serlin (1988). All first-year students entering a university are required to take a course in composition. An instructor who teaches one of the courses in composition wants to find out if the course is equally effective for students in six different programs: (a) business, (b) administration, (c) engineering, (d) chemistry, (e) English, (f) history, and (g) mathematics. One of the problems with simply using the final composition exam scores to compare the groups of students is that the students enter the composition course with varying verbal abilities. The instructor has, however, access to each student's verbal SAT score and plans to use these scores as a covariate. The instructor will compare the groups' composition means after adjusting the composition scores for incoming verbal ability.

From the above example, it is clear that the instructor is interested in making comparisons among groups of students on a single dependent variable. Normally, when the comparison involves more than two groups, analysis of variance is used to analyze the data. In this case, however, it is also known that the students differ on an important variable that is very likely related to the dependent variable. The students differ in verbal ability and this, it is assumed, may be related to how effective the composition course is with the different groups of students as measured by the final exam. Normally, when there is reason to believe that there is a variate or covariate that, if included in the analysis of the dependent variable will provide "a method of achieving increased precision over the corresponding analysis of variance by employing statistical control of the sources of variation not directly controlled by the experimenter" (Glass et al., 1972, p. 272), analysis of covariance can be used to analyze the data.

Analysis of variance and analysis of covariance are similar statistical procedures, the only difference between them being that the model for the former is extended to accommodate the use of covariates in the form of linear regressions. Incorporating relevant variables, such as covariates, into the model serves to improve the estimates of the group parameters or group means (Pearce, 1983). By regressing the dependent variable on the covariate, it is possible to determine the linear association or correlation between these two variables, and to adjust the individual scores of the dependent variable so that they do not reflect that variation which can be attributed to the covariate (Cochran, 1957; Smith, 1957). The adjusted individual scores permit the calculation of adjusted group averages which results in improved estimates of the groups' dependent variable values. A statistically significant difference between or among the groups' adjusted values suggests that the groups differ on the dependent variable in spite of taking into account the covariate.

Because the adjustment in ANCOVA is based on the linear regression of the dependent variable on the covariate, it is important that the relation between the dependent variable and covariate be linear (Elashoff, 1969; Maxwell et al, 1993; Pigache, Graham, & Freedman, 1976). Although, Rutherford (1992) claims ANCOVA can be based on non-linear models, researchers prefer to work with linear models due to their simplicity of calculation and interpretation.

The covariate in ANCOVA does not need to have any specific distribution, but the dependent variable must meet certain assumptions (Bliss, 1970). Recall these assumptions are:

$$y_{ij} = \mu + \alpha_j + \beta\,(x_{ij} - \overline{X}_{..}) + \varepsilon_{ij} \qquad (1)$$

$$\varepsilon_{ij}\text{s are NID } (0, \sigma^2)\,. \qquad (2)$$

Moreover, ANCOVA can be used with data from individuals or groups of individuals that are initially different on an important variable related to the response measure. This important variable or covariate is considered a nuisance because it makes the comparison of individuals or groups of individuals a problem since the individuals differ on the covariate measure and any observed differences on the dependent variable may be due not to any treatment effect, but rather, to the covariate. For example, the composition instructor can only obtain accurate results regarding the effectiveness of the composition course with the different groups of students if a parameter is included within the analysis model which controls for the initial differences among the students on the covariate. By incorporating the covariate into the model, the instructor is controlling what would otherwise be a source of unaccounted variability in the results.

Another way of looking at the function of the covariate within the model of analysis is as a predictor. Because the use of the covariate within the model is partly dependent on its correlation with the dependent variable, the covariate can be used as a predictor of the response measure. The effectiveness of the covariate in predicting the response variable is seen by the decrease in the residual variation in the model (Cox & McCullagh, 1982; Smith, 1957). Thus, the covariance adjustment has two beneficial effects: (a) it reduces the residual variation that would have been obtained had a model that did not incorporate a covariate been used; in ANCOVA the covariate is no longer interpreted as an unaccounted source of variation or error in the model, but interpreted as an accounted source of variance (Cook & Campbell, 1979); (b) the adjusted parameter estimates are improved because they take into account the covariate's contribution.

Assumptions of ANCOVA

ANCOVA has three main assumptions: additivity, ε_{ij}s should be distributed normally and independently around the regression line, with a mean of zero and an equal variance, and homogeneity of within group regression slopes. The last assumption, homogeneity of within group regression slopes, will be explored further because it is from violation of this assumption that the need for the Johnson-Neyman method becomes evident.

Huitema (1980) claims that one of the most important assumptions related to the use of ANCOVA is the homogeneity of the within group regression slopes. Indeed, he states that a test to ensure homogeneity of regression slopes should be part of any regular ANCOVA procedure. The reason for concern regarding the violation of this assumption arises from the misleading results of the ANCOVA if homogeneity of slopes is not confirmed, and the likelihood that this assumption is the one most often violated (Maxwell et al., 1993; Rutherford, 1992).

Four examples are presented below. Each example serves to illustrate a possible position of the within group regression slopes and the consequences of this position on the interpretation of a treatment effect from the application of an ANCOVA.

Figure 1.1 illustrates the classic case in ANCOVA where the treatment effect, the vertical displacement between estimated regression lines, is constant regardless of the value of the covariate.

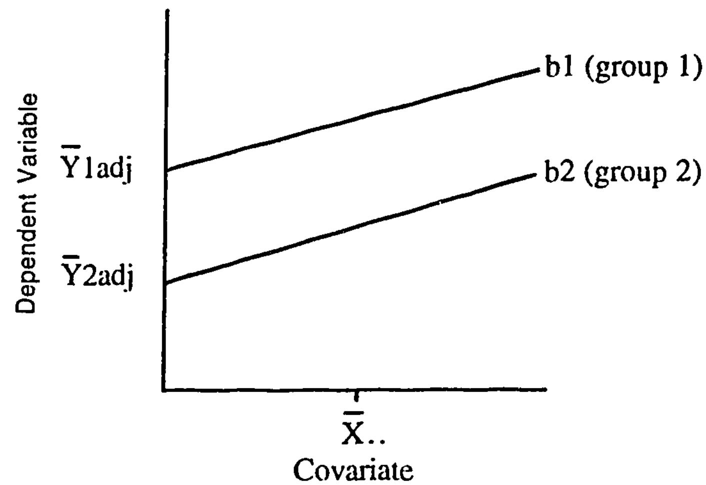


Figure 1.1 Homogeneous Within Group Regression Lines

This vertical displacement is usually, in ANCOVA, assessed at the grand covariate mean ($\bar{X}_{..}$) (Huitema, 1980; Marascuilo & Serlin, 1988). However, the treatment effect could be assessed at any point on the covariate distribution because the within group regression lines are homogeneous and, thus, the vertical displacement between the within group regression lines is constant regardless of the covariate value. Without homogeneity of regression slopes, interpretation of the ANCOVA treatment effect is problematic because the vertical displacement between the within group regression lines is not constant, but a function of the covariate value.

In Figure 1.2, the true within group regression lines are heterogeneous and they intersect at the grand covariate mean, but the estimated regression lines that would be fitted to the data from a traditional ANCOVA are represented by the broken lines, of equal slope, that fall on top of each other.

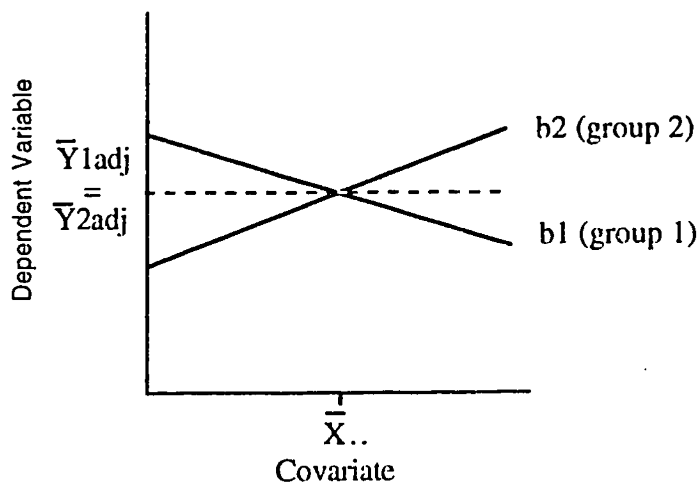


Figure 1.2. Disordinal Within Group Regression Lines

In Figure 1.2, the results from a traditional ANCOVA would indicate no treatment effect because the vertical displacement between the estimated regression lines (broken lines) is zero at the grand covariate mean and throughout the covariate distribution. This conclusion, however, is misleading. Obviously, if the researcher knew of the true nature of the within group regression slopes a treatment effect could be uncovered if the vertical displacement between the true lines was assessed below or above the grand covariate mean. Furthermore, the vertical displacement between true lines would have to be assessed at many different covariate values both below and above the grand covariate mean to obtain an accurate representation of the results.

Similar to the situation illustrated in Figure 1.2, Figure 1.3 and Figure 1.4 show other examples where a traditional ANCOVA does not provide an accurate representation of the data because it fits estimated regression lines of equal slope even when these lines do not represent the true nature of the within group regression lines.

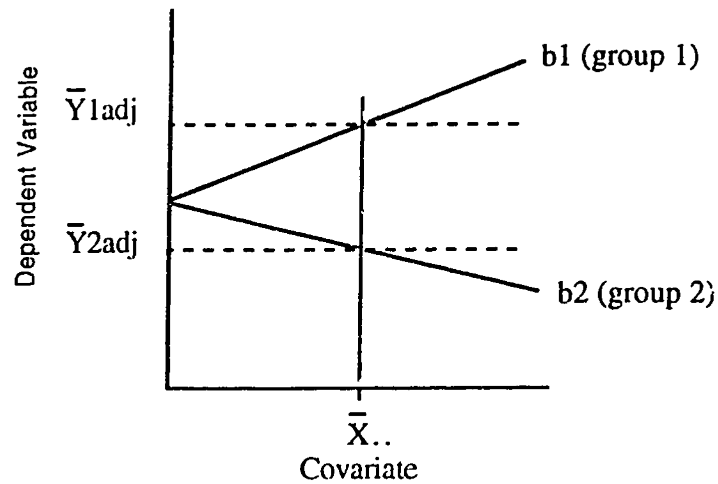


Figure 1.3. Ordinal Within Group Regression Lines

Figure 1.3, for example, illustrates a situation where the true within group regression slopes are heterogeneous and, therefore, the treatment effect is not constant but a function of the covariate. The results from a traditional ANCOVA, however, would likely indicate a large and constant treatment effect because the vertical displacement between the estimated regression lines (broken lines) is substantial at the grand covariate mean and throughout the covariate distribution. The result from a traditional ANCOVA would be inaccurate because the vertical displacement between the true within group regression lines does not remain constant for different values of the covariate. A significant treatment effect may be present for a large covariate value but not for a low covariate value. As in Figure 1.2, the true nature of the treatment effect in Figure 1.3 is a function of the covariate and in order to obtain an accurate picture of the data, the estimated regression lines should be calculated to allow for the identification of their unequal slopes.

Figure 1.4 illustrates a similar situation as the preceding examples. In this case, however, the true within group regression slopes are heterogeneous but group 1 is always superior to group 2. In all these examples, except for the first example, in order for the researcher to get a complete and accurate picture of the data, the estimated regression lines should be calculated so that unequal slopes can be identified. The ANCOVA does not allow for lines of unequal slope to be identified because it fits estimated regression lines of equal slope to the data even when this is inappropriate.

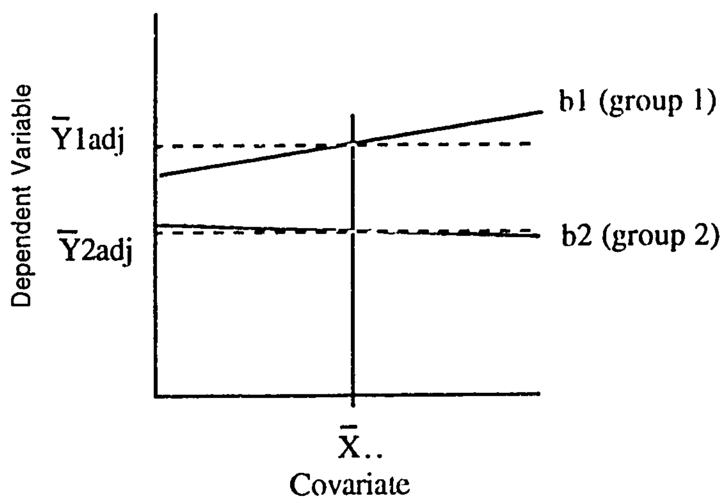


Figure 1.4. Ordinal Within Group Regression Lines

Hence, homogeneity of within group regression slopes is a key assumption in ANCOVA. Satisfying this assumption confirms that the group treatment effect is consistent over the range of the covariate. In other words, the treatment effect can be evaluated at the grand covariate mean without worry that the treatment effect will change with different values of the covariate (Huitema, 1980; Marascuilo & Serlin, 1988). However, it is important not to confuse the need for independence between the treatment and the covariate and the relationship between the dependent variable and the covariate. The covariate is and must be related to the dependent variable for it to be useful in the ANCOVA design, and this relationship is readily seen from the slopes of the within group regression lines. On the other hand, the covariate and the treatment must be independent in order for the treatment effect to not become a function of the covariate.

History and Development of ANCOVA

The idea of using covariates in the analysis of data originated with R.A. Fisher in 1932 in his book entitled, *Statistical Methods for Research Workers*. In his work, Fisher demonstrated how ANCOVA combined the advantages of two widely applicable procedures--regression and ANOVA (Cochran, 1957; Fisher, 1932). In applying ANCOVA he recommended examining the data before deciding whether or not to make an adjustment with a covariate. Fisher suggested that if the residual error decreased when the covariate was added then it should be included, but if the residual error stayed the same then adding a covariate was not necessary. Furthermore, Fisher claimed that the chief advantage of ANCOVA was in the guidance it gave in the design of an observational program because it increased the value of the design by accounting for relevant variables that would, if left unaccounted, likely bias a study. It is interesting that in his original formulation of ANCOVA, Fisher did not mention the important assumption of homogeneous regression slopes for the proper interpretation of ANCOVA results.

Three years later, in his book, *The Design of Experiments,* Fisher (1935) elaborated many of his original ideas. For example, he wrote that to accurately assess any difference between means, the magnitude of the variation from uncontrolled causes should be calculated, and the uncontrolled causes should be made into covariates. He also wrote that the precision of experiments can be increased by eliminating causes of variation which are commonly uncontrolled in observational studies. Fisher also discussed how ANCOVA can be generalized for the use of 2 or 3 covariates, and how ANCOVA is the process of calculating the average treatment effect based on one value of the covariate. Once again, however, there is no mention that within group regression lines ought to be tested for homogeneity or that treatment effects can become a function of the covariate if the within group regression slopes are heterogeneous.

Advantages and Limitations of ANCOVA

Many of the advantages in using ANCOVA have already been mentioned. In general, ANCOVA increases the precision of estimated group effects in randomized experiments , and adjusts for preexsisting group differences in nonrandomized studies (Maxwell et al., 1993). However, these advantages need to be qualified because there are limitations to them that should be considered. For example, Maxwell et al. (1993) state that ANCOVA, as originally conceived by Fisher, is a method for increasing the precision of estimated group effects in randomized studies. Unfortunately over the years, ANCOVA has been used primarily as a method for adjusting preexisting group differences and the assumptions required in this case are more stringent than those required when assignment is randomized (Maxwell et al., 1993). Furthermore, according to Maguire and Haig (1976), applying statistical techniques that supposedly control nuisance variables in Not Truly Experimental (NTE) research may add a more serious problem than it solves. For example, if a construct is conceptualized partly from the relationships it shares with other contructs and variables, then when a "nuisance" variable is statistically removed from the construct under study, in effect, the construct changes. Therefore, it may be preferable to accept the complexity of an experiment or study and describe it in its entirety in order to arrive at any conclusions than to try and statistically remove important variables that, in their absence, may change the contruct of interest.

Why Is It Important to Understand ANCOVA: A Summary

It is important to understand the statistical procedures that are employed to analyze data primarily because the results and conclusions obtained from a study are contingent on the specific procedures that were used with the data. This dilemma is compounded with the use of computer software packages.

Statistical software packages are often treated as magic boxes that will make sense of and manage data. It is important to remember that statistical software packages should be used to reduce or even eliminate the need for researchers to calculate messy equations, and efficiently and accurately handle large data sets. Software packages should not replace conceptual understanding on the part of the researcher. Just like a calculator is used for convenience and ease by children after they have mastered mathematical procedures on their own, statistical software packages should be used for their convenience and ease by researchers after they have understood the statistical methods they wish to employ. Although, understanding ANCOVA may not be considered essential precisely because of the many statistical software packages that provide simple ANCOVA analyses without requiring much knowledge from the user, proper interpretation of output from these software packages is contingent on a knowledgeable user.

To summarize, a key assumption in the use of ANCOVA and the one that is likely most often violated is the homogeneity of within group regression slopes (Maxwell et al., 1993; Rutherford, 1992). In those cases where homogeneity is not present and within group regression slopes are heterogeneous, the group treatment effect becomes a function of the covariate. This can produce seriously misleading results. For example, a claim may be made that two groups do not differ on a certain variable, but this lack of treatment effect may be observed only for a finite range in the distribution of the covariate and the groups may differ markedly at other points along the covariate distribution (see for example Figure 1.2).

Given that this ANCOVA assumption should be met, how can its violation be avoided? The violation of the homogeneity assumption can be avoided by always testing for homogeneity of slopes before proceeding with an ANCOVA. If the within group regression slopes are homogeneous, the test simply confirms the homogeneity and a traditional ANCOVA can be used to analyze the data. On the other hand, if the within group regression slopes are not homogeneous, but heterogeneous, ANCOVA is no longer the statistical procedure of choice. The procedure that should be employed when the within group regression slopes are heterogeneous is the Johnson-Neyman method (Johnson & Fay, 1950). The Johnson-Neyman method has been described by some researchers as simply a more general form of the ANCOVA since it provides not only similar information, but additional information as well (Rogosa, 1980, 1981).

Chapter II
An Introduction to the Johnson-Neyman Method

The Johnson-Neyman method can be used instead of ANCOVA if researchers want an analysis similar to ANCOVA (both methods incorporate a linear regression of the dependent variable on the covariate in order to increase the accuracy of the estimated group parameters), but without having to meet the restriction of homogeneous within group regression slopes.

The Johnson-Neyman method can be a very useful tool in behavioural research for the following reasons: (a) social scientific data does not always meet the required ANCOVA assumptions. For example, there are situations where interactions surface between the independent variable and the covariate, and consequently, the assumption of homogeneity of within group regression slopes is violated (Cronbach & Snow, 1977); (b) educators and theorists may be interested in more than just the treatment differences between groups at a single point along the covariate distribution, and they may be interested in knowing the effects of different covariate values on a treatment effect; and (c) researchers may expect, from the basis of their research, an interaction between the independent variable and covariate, and may want to know a priori the statistical technique to use with the data. For instance, in the area of Aptitude Treatment Interaction (ATI), Cronbach and Snow (1977), claim that the heterogeneity of within group regression slopes is part of what is expected in the results of their studies.

The Johnson-Neyman method can be simply defined as a statistical procedure that makes possible the identification of covariate values which yield statistically significant group or treatment differences on a dependent variable (Rutherford, 1992). The Johnson-Neyman method makes this possible by producing a "region of significance" where significant treatment differences are assessed for a range of covariate values (Johnson & Fay, 1950). For example, the region of significance, in the case of two groups and one covariate, is a coordinate space where the dependent variable and covariate are represented by coordinates on the Y and X axes, respectively. The region of significance encompasses all those group differences that are identified from the vertical displacement between the within group regression lines, drawn at specific values of the covariate.

In summary, the Johnson-Neyman method can circumvent the assumption of homogeneous within group regression slopes by producing a region of significance. In addition, some researchers recommend using the Johnson-Neyman method even in cases where one would normally use ANCOVA (Rogosa, 1980, 1981).

Assumptions of the Johnson-Neyman Method

The Johnson-Neyman method and ANCOVA share two main assumptions which are (a) additivity and (b) ε_{ij}s are normally and independently distributed with a mean of zero and variance, σ^2. However, these two methods differ on a third assumption: ANCOVA requires that within group regression slopes be parallel while the Johnson-Neyman method does not require that this assumption be met.

Pigache et al. (1976) suggest that some additional conditions be considered before the Johnson-Neyman method is applied. These conditions are complex issues and are mentioned here for the interested reader who may wish to investigate them in greater depth. Briefly, these conditions, which fall under research design considerations, are: (a) pre-treatment levels not subject to random error, (b) independence from measurement carry-over effects, (c) regression toward the mean in prediction, and (d) problems of extrapolation.

History and Development of the Johnson-Neyman Method

Originally, the Johnson-Neyman method was developed in an attempt to provide a solution to the problem of matching subjects (Johnson & Neyman, 1936). Johnson and Neyman claimed that the method of matching subjects possessed two weaknesses: First, matching individuals on one trait was difficult, and matching subjects on a number of traits was even more difficult. Second, treatment effects found between groups may be observed not because of any genuine effect, but may originate from an underlying variable that was systematically "unmatched" among the subjects when the subjects were matched on a specific trait. For example, if two groups are created so that individuals in one group match individuals in the other group on I.Q., a danger arises from this systematic matching. The danger lies in that the groups may have been systematically "unmatched" on another important variable, such as motivation. Although, the groups may be matched on I.Q., they are "unmatched" on motivation, and an observed effect between the groups may not be due to a treatment, but their different motivational levels. In an effort to relieve the researcher from the process of matching subjects, Johnson and Neyman demonstrated that both these weaknesses could be overcome by regressing the dependent variable on the trait used for matching the groups. Through regression, Johnson and Neyman argued, researchers would not need to match subjects, but could instead make use of all subjects irrespective of their trait values. Moreover, the use of a larger number of subjects in an experiment had the added benefit of increasing the power of the experiment.

Although, Johnson and Neyman (1936) developed their technique as a solution to the problem of matching subjects, it actually may not change the problem. By matching the subjects on one variable (the covariate) in order to compare them on the dependent variable, can cause an "unmatching" to occur on other, potentially important variables relevant to the dependent variable (Maguire & Haig, 1976). Thus, this may be an important consideration to keep in mind when applying any statistical procedure that will "match" subjects.

Non-Simultaneous and Simultaneous Confidence Limits in the Johnson-Neyman Method

The region of significance in the Johnson-Neyman method can also be formulated in terms of a confidence interval (Potthoff, 1964, 1983). First, recall that the region of significance contains those covariate values that are associated with statistically significant group differences. That is, the vertical displacement between the within group regression lines is large enough to be statistically significant. Outside of this region of significance, no differences will be found between the groups for any covariate value. Now, if a 95 percent confidence interval were calculated for each of the treatment differences that is found within the region of significance, the number zero will not fall within any one of the confidence intervals constructed. Thus, for any point that falls within the region of significance, its corresponding confidence interval will also show the point's significance. A problem that arises, however, is the distinction between two forms of confidence intervals that can be constructed from the region of significance--simple confidence intervals and simultaneous confidence intervals. Simple confidence intervals are associated with the region of significance labeled, R, and simultaneous confidence intervals are associated with a smaller subset of the region of significance labeled, R' (Potthoff, 1964).

Simple Confidence Intervals

Rogosa (1980) labels simple confidence intervals, pick-a-point intervals, because they should be used only when assessing treatment differences at a single point along the covariate distribution. To this extent, simple confidence intervals are analogous to the way ANCOVA assesses treatment differences--treatment differences are assessed at only one point along the covariate distribution because it is assumed that the regression lines are parallel. The reason that simple confidence intervals should be used only when making group comparisons at a single covariate value is to control the probability of making a Type I error. For example, when a 95 percent simple confidence interval is constructed for a specific group comparison, there exists a 5 percent chance that the confidence interval will show a significant group difference when, in fact, there is no difference at all between the groups. Hence, if simple confidence intervals were constructed to assess a multitude of treatment differences, and there exists a 5 percent chance of making a type I error with every confidence interval constructed, the chances of committing a type I error with multiple comparisons are much higher than 5 percent.

Simultaneous Confidence Intervals

Simultaneous confidence intervals are recommended when more than one comparison of group treatment differences is desired. That is, when group differences are assessed at many points along the covariate distribution. The reason that simultaneous confidence intervals are suggested for this purpose is because they are calculated so that a 5 percent error rate is maintained over a set of comparisons (the probability of making a type I error in a single 95 percent simultaneous confidence interval is much smaller than 5 percent) and, therefore, even when making many comparisons the probability of making a Type I error is held constant at 5 percent.

Aitkin (1973) argues, however, that the simultaneous confidence interval is too conservative for practical purposes. For example, he claims that, in most cases, researchers are interested in knowing about group differences for a finite range of covariate values and not necessarily for all covariate values within the region of significance, R'. To this end, he recommends the use of Gafarian simultaneous bands, which allow researchers to choose a finite set of data points in which to examine group differences (see Gafarian, 1964).

Advantages and Limitations of the Johnson-Neyman Method

The Johnson-Neyman method has definite advantages. The most salient advantage in using this method is that it can be applied to data that contain heterogeneous within group regression slopes. In social scientific research this is an attractive advantage because not all data with covariates will have homogeneous within group regression slopes and, thus, fit within the framework of ANCOVA.

A second advantage is that when dealing with heterogeneous within group regression slopes, the Johnson-Neyman method assesses treatment effects for a range of covariate values. By producing a region of significance, the Johnson-Neyman method indicates which covariate values yield significant treatment effects, and which do not. This is accomplished first by showing, within the boundaries of the region of significance, all values of the covariate that are associated with significant group differences (in the case of 2 covariates, the combination of covariate values are identified), and second by the exclusion of those values of the covariate that are not associated with significant group differences from the region of significance. The boundaries that form the region of significance act as gatekeepers in the process of determining which covariate values are associated with significant group differences and which are not. Knowledge that a covariate value resides within the region of significance is just as important as the knowledge that a covariate value resides outside the region of significance because this is relevant information concerning the magnitude of the covariate value under which a treatment effect is observed or not.

It is important to note that the boundaries of the region of significance, which demarcate at which point significant group differences begin and end, can be established with a contrast matrix. For example, (a) one could use a "pick-a-point procedure," a covariate point of interest, or a series of covariate points to define a region. In the latter case, it is through trial and error that values are entered into the contrast matrix; or (b) one could use a symbolic value to find the values of the covariate at which significant treatment differences appear (it is important to consider that some programs do not allow entry of contrast matrices or they allow only contrasts with 1 degree of freedom). Essentially, the attractiveness of the Johnson-Neyman method is well stated by Rogosa (1980, 1981), who confirms that the Johnson-Neyman method provides not only the same information as ANCOVA, but more information with fewer restrictions.

The Johnson-Neyman method does have its limitations, however. First, there are no statistical computer software packages that perform the Johnson-Neyman method. (*Mathematica* can be used to perform the Johnson-Neyman method, but *Mathematica* is not considered a statistical software package). Moreover, this is not a statistical procedure that can be easily done by hand with a pocket calculator. Therefore, the availability of computer software dictates whether this method is employed or not. There have been attempts to circumvent the complex calculations involved with the Johnson-Neyman method. For example, Carroll and Wilson (1970) developed a program designed to perform all the calculations necessary for the Johnson-Neyman method, but it was slow (one half hour per problem) and it did not provide a graphical output of the region of significance.

Currently, there are statistical computer packages that will perform an ANCOVA with heterogeneous within group regression slopes, for example SPSS (1994) and SAS (1982), but the procedures do not provide the region of significance. Thus far, *Mathematica* may be the only computer software that can successfully perform the calculations necessary to the Johnson-Neyman method and provide the region of significance in both numerical and graphical form (*Maples* (Abell & Braselton, 1994), a computer software package, may also be used to perform the Johnson-Neyman method since it has been shown to be similar to *Mathematica*).

A second limitation to the Johnson-Neyman method is advanced by Abelson (1953), who cautions against using the Johnson-Neyman method when homogeneity of within group regression slopes is present. Abelson claims that if one performs the Johnson-Neyman method with homogeneous regression slopes that yield significant treatment effects, the corresponding region of significance may turn out to be so large that it includes all covariate values, it does not add any valuable information to the analysis, and is performed at the expense of a great deal of computation. Today, researchers such as Rogosa (1980, 1981) argue to the contrary, claiming that the Johnson-Neyman method can be used even when homogeneity of within group regression slopes is present. Obtaining a large region of significance that potentially includes all covariate values may not contribute new information to the analysis but it also does not rob the analysis from any valuable information. Furthermore, if the calculations are done by computer, the computational labour is not a problem.

Importance of Understanding The Johnson-Neyman Method

The importance of knowing and understanding the Johnson-Neyman method for social scientific researchers lies in that this statistical procedure can be very useful in their research. There are many instances in educational research where groups or treatments differ on important concomitant variables (Cronbach & Snow, 1977). For example, groups of students frequently differ on verbal or mathematical ability measures, and these measures often serve as covariates in testing or evaluation of new programs or curriculum initiatives. Thus, in educational research it is often the case that within group regression lines are not homogeneous and traditional ANCOVA procedures are, therefore, not justified to be used. In these cases, the Johnson-Neyman method is the method of choice.

Past Application of the Johnson-Neyman Method: Non-Computer Based

Previous non-computer based applications of the Johnson-Neyman method are difficult to reproduce here because of the amount of space that would have to be dedicated to the mathematical calculations. It is not the intent of this study to concentrate on the mathematical details of the Johnson-Neyman method. For those who are interested in reading about the original mathematical application of the Johnson-Neyman method or its derivation, Johnson and Neyman (1936) and Johnson and Fay (1950), respectively, provide complete mathematical details.

Chapter III
Mathematica and Application of the Johnson-Neyman Method

To circumvent the manual computational labour of the Johnson-Neyman method, Carroll and Wilson (1970), using a programming language labeled TELECOMP, wrote a computer program capable of executing the Johnson-Neyman method with two covariates. Although, the program performed all necessary calculations, including identification of the parameters that form the region of significance, it required too much time to execute and did not test for homogeneity of within group regression slopes. Researchers had to confirm this assumption independently before using the program.

Today, *Mathematica* (Wolfram, 1988), a software package that can be used to manipulate interactively mathematical expressions in symbolic and numeric form, successfully executes all calculations necessary for the Johnson-Neyman method, including a graphical output of the region of significance. In the simplest case, users identify the mathematical problems or equations that need to be solved and *Mathematica*, in turn, provides the solutions. In addition, users can plot equations by employing specific *Mathematica* commands and operators.

The option to plot equations on *Mathematica* is a highly useful feature for the user wishing to perform more complex applications of the Johnson-Neyman method. Although *Mathematica* can always solve polynomials (finding a solution to the Johnson-Neyman problem usually involves solving a polynomial which leads to finding the region of significance) when (a) there is one unknown, and (b) the highest power is 4, if these two conditions are not met it may prove to be mathematically impossible to provide an exact algebraic solution and plotting the polynomial becomes the method of choice by which the region of significance is identified.

Before presenting three examples that illustrate the use of the Johnson-Neyman method with *Mathematica*, a brief description of the *general linear model* is necessary because it is used to define all variables, equations, and contrasts relevant to performing the Johnson-Neyman method in *Mathematica*. Furthermore, a description of the semantics used in *Mathematica* is given.

The General Linear Model

The general linear model, in its most elementary form, is defined as $Y=X\beta+\varepsilon$; where Y is a vector of order (N,1) whose elements represent values of the dependent variable for N observations; X is a design matrix of order (N.n) and of rank n (N>n) whose elements represent the independent variables; β is a vector of order (n,1) whose elements represent the population parameters to be estimated; ε is a vector of order (N,1) whose elements are the population error values.

The design matrix (X) may be constructed in a number of different ways depending upon the parameters researchers wish to estimate. For example, Searle, Speed, and Henderson (cited in Hunka, 1993) claim that the three most commonly used design matrices are the Σ-restricted model, μ–model, and set-to-zero model. Although, all of these design matrices can be used in solving for the Johnson-Neyman method, each estimates different parameters. Commonly, researchers use the Σ-restricted model and μ–model because the parameters are more easily interpreted than the set-to-zero model. It is important, however, that researchers be aware of the design matrix they select because interpretation of parameter estimates and construction of contrast matrices are contingent upon the design matrix used. Lastly, an estimate of β is solved through the ordinary least squares method $B=(X'X)^{-1}X'Y$.

Use of the Σ-restricted design matrix in a simple one-way analysis of variance (ANOVA) design with j groups produces estimates of an overall mean or a constant effect μ, and treatment effects α_j (j=1,2,...,g-1) such that $\Sigma\alpha_j=0$. The Σ-restricted design matrix is created for an ANOVA design as follows:

X_{i1}= 1 for all observations i

$X_{i,j+1}$= 1 for all observations i in group j, otherwise a zero, except when i is in group j+1=g; then

$X_{j+i,1}$= -1 for all i

Thus, for a three group ANOVA with two observations in each group, the design matrix would be as follows:

$$x_1 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \end{bmatrix}$$

The parameters estimated by the design matrix x_1 are μ, α_1, and α_2. In order for the design matrix above to be used for an ANCOVA, some slight changes are needed. For example, if one covariate measure is included in the ANOVA design, three additional columns must be appended to the design matrix, each one containing the covariate values for a specific group.

$$x_2 = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 2 & 0 & 0 \\ 1 & 0 & 1 & 0 & 4 & 0 \\ 1 & 0 & 1 & 0 & 7 & 0 \\ 1 & -1 & -1 & 0 & 0 & 8 \\ 1 & -1 & -1 & 0 & 0 & 9 \end{bmatrix}$$

Now, the parameters estimated by x_2 are μ, α_1, α_2, δ_1, δ_2, δ_3; the δ parameters are estimates of individual within group regression coefficients. With the δ parameters, the α parameters are interpreted as treatment effects adjusted at a covariate value of zero.

If a μ-model design matrix had been constructed instead of the Σ-restricted design matrix, each of the first three columns of the design matrix would contain a one if the observation is in group j, otherwise a zero. For example, the three group ANOVA design with two observations in each group would take the following form:

$$x_3 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

The parameters estimated by x_3 are μ_1, μ_2, and μ_3. To modify this design matrix, once again, for an ANCOVA with one covariate, the same procedure is used as was used with the Σ-restricted model--three additional columns are appended, each column containing the covariate values for the specific group.

$$x_4 = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 4 & 0 \\ 0 & 1 & 0 & 0 & 7 & 0 \\ 0 & 0 & 1 & 0 & 0 & 8 \\ 0 & 0 & 1 & 0 & 0 & 9 \end{bmatrix}$$

The parameters, thus, estimated by this modified design matrix x_4 would be μ_1, μ_2, μ_3, δ_1, δ_2, δ_3. Once again, the μ parameters are interpreted as group means adjusted at a covariate value of zero, and the δ parameters are estimates of the individual within group regression coefficients.

In order to test specific hypotheses, contrast matrices are formed and their associated sum of squares calculated. For example, the sum of squares of a specific contrast matrix K', of order (r,n) and rank (r), that is associated with the hypothesis $K'\beta - m = 0$ in which m is set to zero and β is estimated from $B=(X'X)^{-1}X'Y$, is calculated by

$$SS_k=(K'B)'[K'(X'X)^{-1}K]^{-1}K'B \qquad (4)$$

In addition, to test for the statistical significance of the hypothesis associated with contrast K' at an α level of 0.05, the following relationship should hold:

$$(K'B)'[K'(X'X)^{-1}K]^{-1}K'B - r \text{ Mse} \times F_{0.05;dfk;dfe} \geq 0 \qquad (5)$$

where Mse = appropriate Mean Square Error for the data

dfk = the degrees of freedom (k) associated with the specific contrast formed

dfe = the degrees of freedom for the error term

The contrast matrix K' and its associated hypothesis $K'\beta - m = 0$ are used to assess all potential effects, including the region of significance in the Johnson-Neyman method. For those readers interested in further information regarding the general linear model, refer to Searle (1971).

Mathematica Semantics

All variables, equations, and contrasts must be specified with a particular syntax in *Mathematica*. For ease of understanding the *Mathematica* commands used to do the Johnson-Neyman method, 12 syntax descriptions and examples are illustrated below:

(a) Input into *Mathematica* is portrayed in **bold** courier font.

e.g., **20 x + 5**

(b) Output from *Mathematica* is portrayed in plain Courier font.

e.g., `15`

(c) An operation is positioned to the left of an argument with the argument in square brackets.

e.g., Inverse[x] where $x = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

(d) A (.) is a matrix multiplier.

e.g., x = (5,5) matrix; y = (5,5) matrix

x.y

(e) An assignment is made with the equal symbol (=).

e.g., a = 5; x = 2

(f) A vector or list is defined by enclosing a list of number(s) within one set of curly brackets.

e.g., {4, 5, 5, 7,..., n}

(g) A matrix is defined by enclosing subsets of numbers belonging to the matrix within single curly brackets and the entire set of numbers within another set of curly brackets.

e.g., {{4, 5, 2},{5, 1, 3},{7, 8, 9}}

(h) A right slanted slash (/) performs a division.

e.g., 20/5 = 4

(i) A left slanted slash (\) signals a line continuation.

(j) A space between or among a set of numbers or symbols is a multiplication operator.

e.g., 5 4 returns the result 20

(k) Two right slanted slashes (//) followed by an N gives numerical values, expresses result in real form.

e.g., 4/5 //N returns the result 0.80

(l) All *Mathematica* operators have the first character in upper case.

e.g., Inverse[x]

Many of the operations that are input into *Mathematica* to solve for the Johnson-Neyman problem generate very long output. Such long output takes a lot of space and detracts from the primary purpose of showing how the Johnson-Neyman method is performed with *Mathematica*. Therefore, to save space but not to diminish from a complete and comprehensive demonstration of how the Johnson-Neyman method is performed with *Mathematica*, all output from *Mathematica* is shown for the first example only. For the second and third example, the output is selectively shown, and only when it is relevant towards the application of the Johnson-Neyman method, for example, the polynomial equation needed to solve the Johnson-Neyman method.

In addition, the following general steps are executed for each example:

(a) The design matrix is defined.

(b) The Y vector is defined.

(c) The transpose of X is multiplied by X to form X'X.

(d) The inverse of X'X is calculated.

(e) The transpose of X is post-multiplied by Y.

(f) Estimated parameters are calculated by $B = [X'X]^{-1}X'Y$.

(g) The contrasts (K') are defined and used to solve for SS_k (sum of squares for contrast K').

(h) The contrasts are tested for statistical significance.

(i) The region of significance is found by forming the polynomial from

$$SS_k = (K'B)'[K'(X'X)^{-1}K]^{-1}(K'B)$$

and then subtracting the constant r MSe $F_{0.05,\, dfe, dfk}$ and solving for the unknowns embedded in K' by setting

$$(K'B)'[K'(X'X)^{-1}K]^{-1}(K'B) - r\ MSe\ F_{0.05,\, dfe, dfk} = 0.$$

Example 1: Two Groups and One Covariate

In order to demonstrate the Johnson-Neyman method as it is applied with *Mathematica*, a data set taken from Huitema (1980) is used for purposes of the analysis. The data set is made up of 30 subjects that are divided into two different therapy conditions. The dependent variable is an aggressiveness score on a behavioural checklist; the covariate represents individual scores on a sociability scale. The Σ-restricted model will be employed in all examples.

Data Input

```
x={{1,1,1,0},{1,1,2,0},{1,1,2,0},{1,1,3,0},
{1,1,4,0},{1,1,5,0},{1,1,5,0},{1,1,6,0},
{1,1,6,0},{1,1,7,0},{1,1,8,0},{1,1,8,0},
{1,1,9,0},{1,1,10,0},{1,1,11,0},{1,-1,0,1},
{1,-1,0,1.5},{1,-1,0,2.5},{1,-1,0,3.5},
{1,-1,0,4.5},{1,-1,0,4.5},{1,-1,0,5},{1,-1,0,6},
{1,-1,0,6},{1,-1,0,7},{1,-1,0,7},{1,-1,0,7.5},
{1,-1,0,8},{1,-1,0,9},{1,-1,0,10}}
```

The Σ-restricted model, as defined above, has four columns. The first column estimates the constant effect μ, the second column estimates the treatment effect for group 1, the third column estimates the within group regression coefficient for group 1, and the fourth column estimates the within group regression coefficient for group 2.

```
y={10,10,11,10,11,11,10,11,11.5,12,12,11,11,12.5,12,
5,6,6,7,8,9,9,9,10.5,11,12.5,12.5,14,14.5,16}
```

The Y vector is ordered with respect to the dependent variable; the first 15 values correspond to group 1 and the last 15 values correspond to group 2.

Basic Calculations

```
xx=Transpose[x].x

{{30., 0., 87., 83.}, {0., 30., 87., -83.},
{87., 87., 635., 0.}, {83., -83., 0., 558.5}}
```

```
Inverse[xx]

{{0.174963, -0.012642, -0.0222393, -0.0278804},
{-0.012642, 0.174963, -0.0222393, 0.0278804},
                                                   -21
{-0.0222393, -0.0222393, 0.00766871, 2.5411 10      },

                                       -21
{-0.0278804, 0.0278804, 1.69407 10       , 0.0100773}}

xy=Transpose[x].y

{316., 16., 990., 952.75}
```

In order to estimate the parameters, μ, α, and δ contained in vector B, X'X is calculated along with its inverse, and X'Y.

```
b=Inverse[xx].xy

{6.5061, 3.35075, 0.208589, 1.23698}
```

The parameter estimates given by vector B are μ=6.5061; α_1=3.35075; δ_1=0.208589; and δ_2=1.23698.

Contrasts

```
ssk[k_,xx_,b_]:=
(k.b).(Inverse[k.Inverse[xx].Transpose[k]]).
Transpose[(k.b)]
```

The above expression corresponds to $SS_k=(K'B)'[K'(X'X)^{-1}K]^{-1}K'B$ and, thus, calculates the sum of squares associated with a specific contrast. This expression will be used to assess the sum of squares for all contrasts. Essentially, one substitutes a specific contrast matrix, K_a, into k_; X'X into xx_; and B into b_.

Contrast K_1, shown below, is used to assess the homogeneity or heterogeneity of the within group regression slopes. Contrast K_1 has one degree of freedom. Parenthetically, the contrast used to assess heterogeneity of within group regression slopes is principal to the Johnson-Neyman method. The outcome of this contrast will determine how treatment effects are interpreted. For example, if contrast K_1 is significant (indicating that the within group regression slopes are heterogeneous) the treatment effects cannot be assessed as they would be in a traditional ANCOVA because the homogeneity assumption has been violated. In this case, the treatment effects must be assessed as a function of the covariate. On the other hand, if K_1 is not significant, the treatment effects can be assessed as they would be normally with the ANCOVA.

```
k1={{0,0,1,-1}}
{{0,0,1,-1}}
```

The null hypothesis associated with K_1 is the following:

$$Ho: \; K_1\beta = [\,0 \;\; 0 \;\; 1 \;\; -1\,] \begin{bmatrix} \mu \\ \alpha_1 \\ \delta_1 \\ \delta_2 \end{bmatrix} = [\delta_1 - \delta_2] = [0]$$

The sum of squares associated with K_1 is 59.5964 and its mean square remains 59.5964 because K_1 has only one degree of freedom. In addition, the mean square error for this example is 0.42.

```
ssk[k1,xx,b]
59.5964
```

In order to compute the F statistic for K_1, 59.5964 is divided by 0.42 and the F value produced is 141.896 which is statistically significant at an α level of 0.05. Thus, homogeneity of within group regression slopes is rejected and heterogeneity is assumed of the regression coefficients. This result also means that the treatment effects must be interpreted with regard to the covariate and the Johnson-Neyman method is used to assess the treatment effects.

<u>The Johnson-Neyman Method: Finding the Region of Significance</u>

Contrast K2 is set up (see Appendix 1) to find the points along the covariate distribution where significant treatment effects are found.

```
k2={{0,2,p,-p}}
{{0, 2, p, -p}}
```

$$\text{Ho: } K2\beta = \begin{bmatrix} 0 & 2 & p & -p \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \delta_1 \\ \delta_2 \end{bmatrix} = \begin{bmatrix} 2\alpha_1 + & p\delta_1 & -p\delta_2 \end{bmatrix} = [0]$$

The sum of squares associated with K2 is expressed in *Mathematica* in the form of a polynomial because K2 contains the symbolic, unknown value, p. An indicator of *Mathematica's* power is illustrated by the expression ssk[k2,xx,b] which can produce numeric or symbolic results--something unique in a computer language.

```
d=ssk[k2,xx,b]
                                2
        (6.70149 - 1.02839 p)
-----------------------------------
                                   2
0.699852 - 0.200479 p + 0.017746 p
```

In order to solve for the unknown value p, $(K'B)'[K'(X'X)^{-1}K]^{-1}K'B$ - r Mse x $F_{0.05;dfk;dfe} = 0$ is used. First, $(K'B)'[K'(X'X)^{-1}K]^{-1}K'B$ is known in the form of the polynomial above (d). Second, r Mse x $F_{0.05;dfk;dfe}$ (r=1, dfk=1, dfe=26) is calculated,

```
1 0.42  4.26
1.7892
```

Now, *Mathematica's* Solve operator is used to find the values of p that yield significant treatment effects and bind the region of significance.

```
Solve[d-1.7892  ==0,p]
{{p -> 6.03625}, {p -> 7.0504}}
```

Location of point p at p=6.03625 and p=7.0504 bind the region of significance. That is, these two points demarcate the points along the covariate distribution where significant treatment differences are observed.

The region of significance is plotted using *Mathematica's* Plot operator.

```
Plot[d-1.7892,{p,0,11},AxesLabel->{"p","ss"}]
```

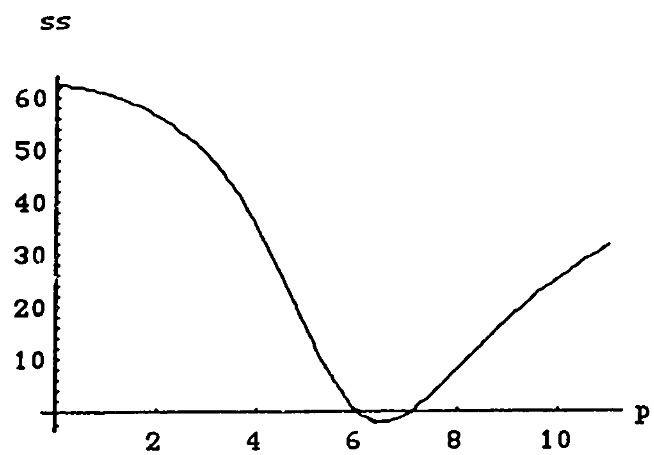


Figure 3.1.: Region of Significance--Two Groups and One Covariate

In Figure 3.1, the function intersects the p axis at two points, p=6.04 and p=7.05. These two points demarcate the origins of the region of significance; the region where statistically significant treatment effects are observed. For instance, all values of p for which the plot is above the horizontal axis yield statistically significant treatment effects at an α value of 0.05 or less. Hence, statistically significant effects between the two therapy conditions are observed when subjects have, as a group, average sociability scores equal to or less than 6.04 and equal to or greater than 7.05.

Example 2: Three Groups and Two Covariates

The preceding example involved only two groups and one covariate. The present example will address the application of the Johnson-Neyman method to a slightly more difficult data set. This hypothetical data set (N=15) is taken from Hunka (1995). In order to make the example less abstract, the three groups will be said to represent three instructional conditions: Seminar, lecture, and peer-learning. The two covariates represent motivation level and verbal ability, respectively. The dependent variable represents scores on a written test.

Data Input

```
x={{1,1,0,3,0,0,4,0,0},{1,1,0,3,0,0,5,0,0},
{1,1,0,4,0,0,6,0,0},{1,1,0,4,0,0,5,0,0},
{1,1,0,5,0,0,6,0,0},{1,0,1,0,2,0,0,4,0},
{1,0,1,0,4,0,0,5,0},{1,0,1,0,5,0,0,2,0},
{1,0,1,0,5,0,0,1,0},{1,0,1,0,7,0,0,6,0},
{1,-1,-1,0,0,5,0,0,7},{1,-1,-1,0,0,3,0,0,5},
{1,-1,-1,0,0,4,0,0,6},{1,-1,-1,0,0,5,0,0,5},
{1,-1,-1,0,0,6,0,0,7}}
```

Once again, the Σ-restricted design matrix is used to define the model. This design matrix has a total of nine columns and allows estimation of μ, α_1, α_2, δ_{11}, δ_{21}, δ_{31}, δ_{12}, δ_{22}, and δ_{32}.

```
y={10,9,9,7,7,3,3,3,4,4,6,4,6,4,7}
```

The Y vector is again ordered with respect to the dependent variable. For this example, the first 5 values correspond to group 1, the second 5 values correspond to group 2, and the last 5 values correspond to group 3.

Basic Calculations

```
xx=Transpose[x].x
```

```
Inverse[xx]
```

```
xy=Transpose[x].y
```

X'X, Inverse[X'X], and X'Y are calculated so that the B vector can be computed as shown below:

```
b=Inverse[xx].xy
{4.27825, 8.12175, -1.7418, -1.6, 0.218722, 0.0169492,
0.4,-0.039604, 1.23729}
```

The parameter estimates given by the vector B are μ=4.27825, α_1=8.12175, α_2=−1.7418, δ_{11}=−1.6, δ_{21}=0.218722, δ_{31}=0.0169492, δ_{12}=0.4, δ_{22}=−0.039604, and δ_{32}=1.23729.

<u>Contrasts</u>

```
k1={{0,0,0,1,-1,0,0,0,0},
{0,0,0,0,1,-1,0,0,0},
{0,0,0,0,0,0,1,-1,0},
{0,0,0,0,0,0,0,1,-1}}
```

$$\text{Ho:}\quad K_1\beta = \begin{bmatrix} 0&0&0&1&-1&0&0&0&0 \\ 0&0&0&0&1&-1&0&0&0 \\ 0&0&0&0&0&0&1&-1&0 \\ 0&0&0&0&0&0&0&1&-1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \delta_{11} \\ \delta_{21} \\ \delta_{31} \\ \delta_{12} \\ \delta_{22} \\ \delta_{32} \end{bmatrix} = \begin{bmatrix} \delta_{11}-\delta_{21} \\ \delta_{21}-\delta_{31} \\ \delta_{12}-\delta_{22} \\ \delta_{22}-\delta_{32} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Contrast K_1 in this example is used to determine the homogeneity or heterogeneity of the within group regression slopes corresponding to both covariates simultaneously. For example, the first two rows of K_1 assess the homogeneity of within group regression slopes for covariate 1, while the last two rows of K_1 assess the homogeneity of within group regression slopes for covariate 2.

The sum of squares associated with K_1 is 11.1612 and its mean square is 2.7903 because K_1 has four degrees of freedom.

```
ssk[k1,xx,b]
11.1612
```

```
msk1=11.1612/4
```

```
2.7903
```

The F statistic associated with K_1 is obtained by dividing 2.7903 by 0.654833, the mean square error for this data set. The F value for K_1 is 4.26109 ($F_{0.05,critical} = 4.53$) which just barely misses statistical significance at an α value of 0.05. However, because 4.26109 is so close to statistical significance, for purposes of illustrating the Johnson-Neyman method in this example the within group regression slopes for covariate 1 and for covariate 2 are taken to be heterogeneous. Due to heterogeneity, the treatment effects must be assessed as a function of the covariate.

The Johnson-Neyman Method: Finding the Region of Significance

Contrast K_2 is set up (see Apendix 2), as was explained previously in Example 1, to find the points along the covariate distribution where significant treatment differences are found.

```
k2={{0,1,-1,p,-p,0,q,-q,0},
{0,1,2,0,p,-p,0,q,-q}}
```

$$Ho:\ K_2\beta = \begin{bmatrix} 0 & 1 & -1 & p & -p & 0 & q & -q & 0 \\ 0 & 1 & 2 & 0 & p & -p & 0 & q & -q \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \delta_{11} \\ \delta_{21} \\ \delta_{31} \\ \delta_{12} \\ \delta_{22} \\ \delta_{32} \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_1-\alpha_2+p\delta_{11}-p\delta_{21}+q\delta_{12}-q\delta_{22} \\ \alpha_1+2\alpha_2+p\delta_{21}-p\delta_{31}+q\delta_{22}-q\delta_{32} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

As can be noticed from K_2, it is more complex than the contrast used to obtain the region of significance in Example 1. The reason for this is, of course, that in this example the region of significance is being defined for two covariates in three different treatment conditions, instead for only one covariate in two different treatment conditions. In this example, the values to which covariate adjustments are to be made are represented by p and q.

The sum of squares associated with K_2 is again in the form of a ratio of two polynomials, although this time it is of degree 6, instead of degree 2 as in Example 1.

```
d=ssk[k2,xx,b]
```

Simplify[d]

$$(1.33883\ 10^{6}\ (1166.63 - 402.764\ p + 102.544\ p^{2} - 11.6942\ p^{3} + 1.\ p^{4} - 396.463\ q + 17.3844\ p\ q - 4.60751\ p^{2}\ q - 1.91442\ p^{3}\ q + 92.9572\ q^{2} + 2.15573\ p\ q^{2} + 3.53078\ p^{2}\ q^{2} - 11.913\ q^{3} - 2.59113\ p\ q^{3} + 1.21071\ q^{4}))\ /$$

$$(132786565 - 6850808\ p + 15736360\ p^{2} - 650736\ p^{3} + 407916\ p^{4} - 88417620\ q - 20000708\ p\ q - 3977568\ p^{2}\ q - 1130560\ p^{3}\ q + 31277440\ q^{2} + 6441288\ p\ q^{2} + 1666092\ p^{2}\ q^{2} - 5365800\ q^{3} - 1244232\ p\ q^{3} + 484660\ q^{4})$$

The sum of squares associated with K2 is quickly calculated as a polynomial with *Mathematica*. This same calculation attempted by hand would not only take an inordinate amount of time, but would run the risk of calculation errors. In addition, even if this calculation could be programmed, as Carroll and Wilson (1970) attempted to do, it may not be possible to also program the solution set for p and q that is required to construct the region of significance. For example, it is essentially impossible to solve the polynomial equation in Fortran and to also solve for the p and q point sets (S. Hunka, personal communication, May, 1995).

This region of significance must be solved and displayed with the *Mathematica* operator Plot3D because the use of two covariates gives rise to a polynomial with two unknowns, thus, violating one of the requirements for an explicit algebraic solution.

```
Plot3D[(d-6.73168),{p,-10,15},{q,-5,20},
  PlotPoints->30,PlotLabel->
       "Polynomial Surface at ss=0",
  AxesLabel->{"x1","x2","ss"},
  PlotRange->{-3,0}]
```

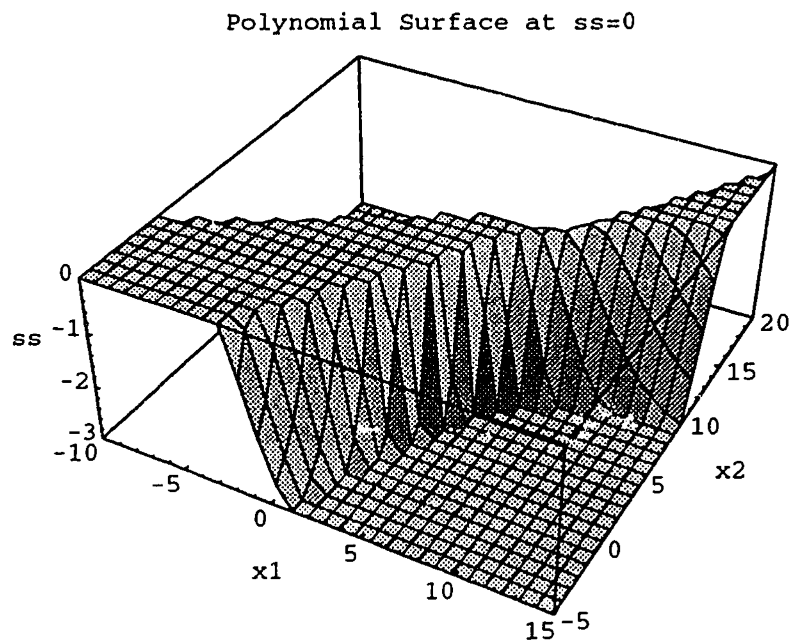


Figure 3.2. Region of Significance--Three Groups and Two Covariates

In Figure 3.2 displayed above, the first covariate, p (x1), ranges from a value of -10 to 15, and the second covariate, q (x2), ranges from a value of -5 to 20. These ranges are wide enough to encompass the covariate distributions of the data, as well as to provide a reasonable display of the results. The vertical axis represents values obtained from the equation:

$$ss=(K'B)'[K'(X'X)^{-1}K]^{-1}K'B - r\ Mse \times F_{0.05;dfk;dfe} \qquad (6)$$

Recall that these "ss" values need to be at least zero or larger for significant differences to be observed among the groups. Notice also that Figure 3.2 illustrates the results of equation 6 for values of ss in the range -3 to 0 only. The reason for this is that ss=0 is the minimum value at which significance appears for the point sets of p and q. Because the plot of equation 6 was truncated at ss=0, Figure 3.2 shows a flat upper surface.

Equation 6 can be used to identify--based on a contrast's sum of squares and rank, mean square error, and selected F value--if the contrast is statistically significant. Hence, the region of significance is assessed by finding the values of covariate 1 and covariate 2 that, in conjunction, yield a value of ss=0 or greater on the vertical axis. Those covariate values, that taken simultaneously, produce a minimum value of zero on the vertical axis are associated with significant treatment differences. For example, if the subjects in the three instructional conditions had average covariate values of 15 and 19 for covariate 1 and 2, respectively, the three instructional conditions would show to be significantly different from one another (because a combination of 15 and 19 reach a value of zero on the vertical axis). Any combination of values for p and q which fall on the flat surface of Figure 3.2 will produce significant treatment differences at an alpha value of 0.05 exactly.

The region of significance can also be represented in another form by using *Mathematica's* Contour Plot operator. In Figure 3.3 covariate 1 (p), labelled x1, is shown on the x- or horizontal axis, while covariate 2 (q), labelled x2, is shown on the y- or vertical axis. The region of significance (at an α=0.05) is that second outermost layer (contour line) of the figure which represents the result of zero from equation 6. Those values of covariate 1 and covariate 2 that fall within the second outermost layer are associated with significant treatment differences. From Figure 3.3, a covariate point-set of (0,0) is significant (p=.016), but (10,5) is not (p=.529). The contour lines represent alpha levels of 0.1, 0.05, .025, and 0.01 which correspond to the values of

4.536, 6.736, 9.508, and 14.307 respectively on the y axis. These values are found from calculating r Mse x $F_{\alpha;dfk;dfe}$ for various values of α and represent the minimum value required for the sum of squares to produce significance.

```
ContourPlot[(d-6.73168),{p,-10,15},{q,-5,20},
   AxesLabel->{"x1","x2"},
   Contours->{4.536,6.736,9.508,14.307}
   FrameLabel->"Regions of Significance"]
```

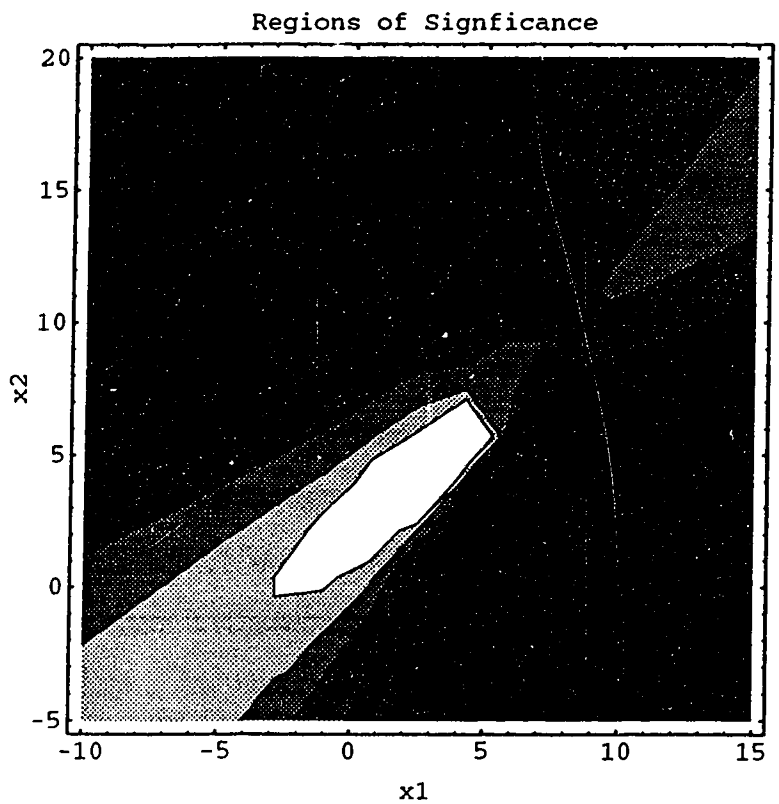


Figure 3.3. Contour Plot--Alternative Representation of Region of Significance

Example 3: Two Way Analysis of Covariance and the Johnson-Neyman Method

Thus far, the examples have involved comparing groups at only one factor. That is, the Johnson-Neyman method has been applied exclusively to one-way ANCOVA data sets. In this section, the Johnson-Neyman method is applied to a two-way ANCOVA data set, where groups are compared on two factors. In applying the Johnson-Neyman method to a two-way ANCOVA design, an important point must be noted. With the addition of another factor to the ANCOVA design, the assumption and verification of homogeneous within group regression slopes is further complicated. For example, in a one-way ANCOVA with two groups, heterogeneity of within group regression slopes is a concern among the two groups at factor A. On the other hand, in a 2 by 3, two-way ANCOVA with six groups, heterogeneity of within group regression slopes is a concern among all levels of factors A and B, as well as within cells (AxB).

Following the work of Hendrix, Carter, and Short (1982), a hierarchical approach is used in application of the Johnson-Neyman method to the analysis of a two-way ANCOVA . The hierarchical approach is employed only as an example for the Johnson-Neyman method and is presented without a discussion of its merits. In the hierarchical approach, the covariate is treated as a separate factor, thereby transforming the analysis into a sort of three-way ANCOVA with the covariate representing the third factor. For example, for a data set that comprises two factors, A (two reading programs) and B (three ability levels--low, medium, and high), and one covariate (score on a motivation scale), a Σ-restricted design matrix is constructed to estimate the following parameters:

$$\mu,\ \alpha_1,\ \beta_1,\ \beta_2,\ \alpha\beta_{11},\ \alpha\beta_{12},\ \delta,\ \alpha\delta,\ \beta_1\delta,\ \beta_2\delta,\ \alpha\beta_{11}\delta,\ \alpha\beta_{12}\delta$$

From the above parameters, it can be seen that the covariate is paired with all the treatment and interaction terms. The result is that for each treatment and interaction term (ANOVA term) there is a corresponding ANCOVA term (where the ANOVA term is paired with the covariate). In addition, as a consequence of the hierarchical approach, the complete Σ-restricted design matrix, whose parameters are represented above, is not always used but subsets of the complete matrix are used to assess the effects of different parameters.

The general procedure of this approach is as follows:

1. The effect of the covariate parameter is assessed in the context of a reduced model.

   If the covariate effect is significant at a predetermined alpha value, then:

2. The effect of all covariate terms are assessed together (pooled) in the context of a full model.

   If the pooled covariate terms' effect is significant at a predetermined alpha value, then:

3. The effect of the highest order ANCOVA term is assessed in the context of a reduced model.

   If the highest order ANCOVA term is significant at a predetermined alpha value, then interpret its corresponding ANOVA term as a function of the covariate. If not significant, pool this term with the error and interpret corresponding ANOVA term without respect for the covariate.

4. The effect of all second order ANCOVA terms are assessed together (pooled) in the context of a reduced model.

   If the pooled second order ANCOVA terms' effect is significant at a predetermined alpha value, then:

5. The effect of individual second order ANCOVA terms are assessed in the context of a reduced model.

   If any or all terms are significant at a predetermined alpha value, interpret the corresponding ANOVA terms with respect to the covariate. If not significant, interpret the corresponding ANOVA terms as usual without respect for the covariate.

The data set that is used to demonstrate the Johnson-Neyman method as applied to a two-way ANCOVA is taken from Hunka (1993). The data set comprises two factors, A (two reading programs) and B (three ability levels--low, medium, and high), and one covariate (score on a motivation scale). There are six groups with five observations in each group.

Data Input: Assessment of Covariate--Reduced Design Matrix 1

```
x1={{1,1,1,0,1,0,40},{1,1,1,0,1,0,35},
{1,1,1,0,1,0,40},{1,1,1,0,1,0,50},
{1,1,1,0,1,0,45},{1,1,0,1,0,1,30},
{1,1,0,1,0,1,40},{1,1,0,1,0,1,45},
{1,1,0,1,0,1,40},{1,1,0,1,0,1,40},
{1,1,-1,-1,-1,-1,50},{1,1,-1,-1,-1,-1,40},
{1,1,-1,-1,-1,-1,40},{1,1,-1,-1,-1,-1,30},
{1,-1,1,0,-1,0,50},{1,-1,1,0,-1,0,30},
{1,-1,1,0,-1,0,35},{1,-1,1,0,-1,0,45},
{1,-1,1,0,-1,0,30},{1,-1,0,1,0,-1,50},
{1,-1,0,1,0,-1,30},{1,-1,0,1,0,-1,40},
{1,-1,0,1,0,-1,45},{1,-1,0,1,0,-1,40},
{1,-1,-1,-1,1,1,45},{1,-1,-1,-1,1,1,30},
{1,-1,-1,-1,1,1,25},{1,-1,-1,-1,1,1,50},
{1,-1,-1,-1,1,1,35}}
```

Design matrix x_1 estimates parameters μ, α_1, β_1, β_2, $\alpha\beta_{11}$, $\alpha\beta_{12}$, δ. This is called a reduced model design matrix because, although, the main effect of the covariate term is assessed, the other covariate terms that arise from treating the covariate as a factor are not assessed.

Basic Calculations

```
x1x1=Transpose[x1].x1

Inverse[x1x1]

x1y=Transpose[x1].y

b1=Inverse[x1x1].x1y
{61.9354, -2.23287, 3.98695, -0.313054, -0.586014,
0.992308, 0.759441}
```

The parameter estimates given by b_1 are μ=61.9354, α_1=−2.23287, β_1=3.98695, β_2=−0.313054, $\alpha\beta_{11}$=−0.586014, $\alpha\beta_{12}$=0.992308, and δ=0.759441.

Contrasts

Next the significance of the covariate is determined.

```
k1={{0,0,0,0,0,0,1}}
```

```
ssk[k1,x1x1,b1]
824.752
```

The sum of squares associated with contrast K_1 is 824.752 and its mean square remains the same because K_1 has only one degree of freedom. The F statistic relevant to K_1 is 824.752 divided by 26.1064, which is the mean square error for this data set. The resulting F value is 31.5919 and it is statistically significant at the 0.05 alpha level.

Because the test of the covariate effect turned out to be significant, the next step is to test the pooled results of all the ANOVA sum of squares terms involving the δ term (ANCOVA terms) using a liberal α level of 0.25 (Hendrix et al., 1982). If this contrast is significant, the highest order ANCOVA term with a full model is tested.

Data Input: Assessment of Pooled ANCOVA terms and Highest Order ANCOVA terms--Complete Design Matrix 2

```
x2={{1,1,1,0,1,0,40,40,40,0,40,0},
{1,1,1,0,1,0,35,35,35,0,35,0},
{1,1,1,0,1,0,40,40,40,0,40,0},
{1,1,1,0,1,0,50,50,50,0,50,0},
{1,1,1,0,1,0,45,45,45,0,45,0},
{1,1,0,1,0,1,30,30,0,30,0,30},
{1,1,0,1,0,1,40,40,0,40,0,40},
{1,1,0,1,0,1,45,45,0,45,0,45},
{1,1,0,1,0,1,40,40,0,40,0,40},
{1,1,0,1,0,1,40,40,0,40,0,40},
{1,1,-1,-1,-1,-1,50,50,-50,-50,-50,-50},
{1,1,-1,-1,-1,-1,40,40,-40,-40,-40,-40},
{1,1,-1,-1,-1,-1,40,40,-40,-40,-40,-40},
{1,1,-1,-1,-1,-1,30,30,-30,-30,-30,-30},
{1,1,-1,-1,-1,-1,40,40,-40,-40,-40,-40},
{1,-1,1,0,-1,0,50,-50,50,0,-50,0},
{1,-1,1,0,-1,0,30,-30,30,0,-30,0},
{1,-1,1,0,-1,0,35,-35,35,0,-35,0},
{1,-1,1,0,-1,0,45,-45,45,0,-45,0},
{1,-1,1,0,-1,0,30,-30,30,0,-30,0},
{1,-1,0,1,0,-1,50,-50,0,50,0,-50},
{1,-1,0,1,0,-1,30,-30,0,30,0,-30},
{1,-1,0,1,0,-1,40,-40,0,40,0,-40},
{1,-1,0,1,0,-1,45,-45,0,45,0,-45},
{1,-1,0,1,0,-1,40,-40,0,40,0,-40},
{1,-1,-1,-1,1,1,45,-45,-45,-45,45,45},
{1,-1,-1,-1,1,1,30,-30,-30,-30,30,30},
{1,-1,-1,-1,1,1,25,-25,-25,-25,25,25},
{1,-1,-1,-1,1,1,50,-50,-50,-50,50,50},
{1,-1,-1,-1,1,1,35,-35,-35,-35,35,35}}
```

The design matrix x_2 estimates the parameters μ, α_1, β_1, β_2, $\alpha\beta_{11}$, $\alpha\beta_{12}$, δ, $\alpha\delta$, $\beta_1\delta$, $\beta_2\delta$, $\alpha\beta_{11}\delta$, $\alpha\beta_{12}\delta$.

Basic Calculations

```
x2x2=Transpose[x2].x2

Inverse[x2x2]

x2y=Transpose[x2].y

b2=Inverse[x2x2].x2y   //N
{63.1534, -3.73357, -11.8854, 17.3295, -17.1499,
```

```
5.12561, 0.724983, 0.023949, 0.379097, -0.438998,
0.410433, -0.101601}
```

The parameter estimates provided by b_2 are μ=63.1534, α=-3.73357, β_1=−11.8854, β_2=17.3295, $\alpha\beta_{11}$=−17.1499, $\alpha\beta_{12}$=5.12561, δ=0.724983, $\alpha\delta$=0.023949, $\beta_1\delta$=0.379097, $\beta_2\delta$=−0.438998, $\alpha\beta_{11}\delta$=0.410433, and $\alpha\beta_{12}\delta$=−0.101601.

Contrasts

The contrast matrix for the pooled results of all ANCOVA terms ($\alpha\delta$, $\beta_1\delta$, $\beta_2\delta$, $\alpha\beta_{11}\delta$, $\alpha\beta_{12}\delta$) excluding δ is:

```
k2={{0,0,0,0,0,0,0,1,0,0,0,0},
{0,0,0,0,0,0,0,0,1,0,0,0},
{0,0,0,0,0,0,0,0,0,1,0,0},
{0,0,0,0,0,0,0,0,0,0,1,0},
{0,0,0,0,0,0,0,0,0,0,0,1}}

ssk[k2,x2x2,b2]
207.335

207.335/5

41.467
```

The sum of squares associated with contrast K_2 is 207.335 and its mean square is 41.467. To obtain an F statistic for this contrast, 41.467 is divided by 21.84 which is the mean square error for this data set. The resulting F value is 1.89867 and it is significant due to the liberal α level of 0.25. Because contrast K_2 proved to be significant, the next step is to test the highest order ANCOVA terms ($\alpha\beta_{11}\delta$, $\alpha\beta_{12}\delta$).

```
k3={{0,0,0,0,0,0,0,0,0,0,1,0},
{0,0,0,0,0,0,0,0,0,0,0,1}}
```

```
ssk[k3,x2x2,b2]
116.324
```

```
116.324/2
```

```
58.162
```

The sum of squares associated with contrast K_3 is 116.324 and its mean square is 58.162. To evaluate the significance of contrast K_3, the F statistic is computed by dividing 58.162 by 21.84. The resulting F value is 2.6631 and it is significant at an α level of 0.25. Therefore, the ANOVA term corresponding to this highest order ANCOVA term must be interpreted with reference to specific values of the covariate.

In the next phase, the second highest order ANCOVA terms are assessed. First, they are assessed as a set and, if significant, the individual terms are tested.

Data Input: Assessment of Pooled Second Order ANCOVA terms--Reduced Design

Matrix 3

```
x3={{1,1,1,0,1,0,40,40,40,0},
{1,1,1,0,1,0,35,35,35,0},
{1,1,1,0,1,0,40,40,40,0},
{1,1,1,0,1,0,50,50,50,0},
{1,1,1,0,1,0,45,45,45,0},
{1,1,0,1,0,1,30,30,0,30},
{1,1,0,1,0,1,40,40,0,40},
{1,1,0,1,0,1,45,45,0,45},
{1,1,0,1,0,1,40,40,0,40},
{1,1,0,1,0,1,40,40,0,40},
{1,1,-1,-1,-1,-1,50,50,-50,-50},
{1,1,-1,-1,-1,-1,40,40,-40,-40},
{1,1,-1,-1,-1,-1,40,40,-40,-40},
{1,1,-1,-1,-1,-1,30,30,-30,-30},
{1,1,-1,-1,-1,-1,40,40,-40,-40},
{1,-1,1,0,-1,0,50,-50,50,0},
{1,-1,1,0,-1,0,30,-30,30,0},
{1,-1,1,0,-1,0,35,-35,35,0},
{1,-1,1,0,-1,0,45,-45,45,0},
{1,-1,1,0,-1,0,30,-30,30,0},
{1,-1,0,1,0,-1,50,-50,0,50},
{1,-1,0,1,0,-1,30,-30,0,30},
{1,-1,0,1,0,-1,40,-40,0,40},
{1,-1,0,1,0,-1,45,-45,0,45},
{1,-1,0,1,0,-1,40,-40,0,40},
{1,-1,-1,-1,1,1,45,-45,-45,-45},
{1,-1,-1,-1,1,1,30,-30,-30,-30},
{1,-1,-1,-1,1,1,25,-25,-25,-25},
{1,-1,-1,-1,1,1,50,-50,-50,-50},
{1,-1,-1,-1,1,1,35,-35,-35,-35}}
```

The design matrix x_3 estimates the parameters μ, α_1, β_1, β_2, $\alpha\beta_{11}$, $\alpha\beta_{12}$, δ, $\alpha\delta$, $\beta_1\delta$, $\beta_2\delta$.

Basic Calculations

```
x3x3=Transpose[x3].x3

Inverse[x3x3]

x3y=Transpose[x3].y
```

```
b3=Inverse[x3x3].x3y  //N
{64.4238, -1.97997, -4.45452, 15.3776, -0.564518,
0.858056, 0.699082, -0.014337, 0.209902, -0.394475}
```

The parameters as given by b_3 are μ=64.4238, α=–1.97997, β_1=–4.45452, β_2=15.3776, $\alpha\beta_{11}$=–0.564518, $\alpha\beta_{12}$=0.858056, δ=0.699082, $\alpha\delta$=–0.014337, $\beta_1\delta$=0.209902, and $\beta_2\delta$=–0.394475.

Contrasts

First all the second highest order ANCOVA terms are tested ($\alpha\delta$, $\beta_1\delta$, $\beta_2\delta$) as a set. If, from this contrast, there is any indication that individual ANCOVA terms may be significant, the individual ANCOVA terms are tested next.

```
k4={{0,0,0,0,0,0,0,1,0,0},
{0,0,0,0,0,0,0,0,1,0},
{0,0,0,0,0,0,0,0,0,1}}

ssk[k4,x3x3,b3]  //N
91.0111

91.0111/3

30.337
```

The sum of squares associated with contrast K_4 is 91.0111 and its mean square is 30.337. The F statistic associated with this contrast is obtained by dividing 30.337 by 25.4718, the mean square error corresponding to this data set. The resulting F value is 1.191 and its probability value is 0.338455. Although, this F value is not highly significant at an alpha value of 0.25, individual ANCOVA terms will be tested for the purpose of illustration.

Data Input: Assessment of Individual Second Order ANCOVA terms--Reduced Design

Matrix 4

```
x4={{1,1,1,0,1,0,40,40},
{1,1,1,0,1,0,35,35},
{1,1,1,0,1,0,40,40},
{1,1,1,0,1,0,50,50},
{1,1,1,0,1,0,45,45},
{1,1,0,1,0,1,30,30},
{1,1,0,1,0,1,40,40},
{1,1,0,1,0,1,45,45},
{1,1,0,1,0,1,40,40},
{1,1,0,1,0,1,40,40},
{1,1,-1,-1,-1,-1,50,50},
{1,1,-1,-1,-1,-1,40,40},
{1,1,-1,-1,-1,-1,40,40},
{1,1,-1,-1,-1,-1,30,30},
{1,1,-1,-1,-1,-1,40,40},
{1,-1,1,0,-1,0,50,-50},
{1,-1,1,0,-1,0,30,-30},
{1,-1,1,0,-1,0,35,-35},
{1,-1,1,0,-1,0,45,-45},
{1,-1,1,0,-1,0,30,-30},
{1,-1,0,1,0,-1,50,-50},
{1,-1,0,1,0,-1,30,-30},
{1,-1,0,1,0,-1,40,-40},
{1,-1,0,1,0,-1,45,-45},
{1,-1,0,1,0,-1,40,-40},
{1,-1,-1,-1,1,1,45,-45},
{1,-1,-1,-1,1,1,30,-30},
{1,-1,-1,-1,1,1,25,-25},
{1,-1,-1,-1,1,1,50,-50},
{1,-1,-1,-1,1,1,35,-35}}
```

The design matrix x_4 estimates the parameters μ, α_1, β_1, β_2, $\alpha\beta_{11}$, $\alpha\beta_{12}$, δ, $\alpha\delta$.

Basic Calculations

```
x4x4=Transpose[x4].x4

Inverse[x4x4]

x4y=Transpose[x4].y

b4=Inverse[x4x4].x4y
{62.3556, -1.15189, 4.02366, -0.357804, -0.560695,
0.987434,
0.749376, -0.0271542}
```

The parameters as given by b_4 are μ=62.3556, α_1=−1.15189, β_1=4.02366, β_2=−0.357804, $\alpha\beta_{11}$=−0.560695, $\alpha\beta_{12}$=0.987434, δ=0.749376, $\alpha\delta$=−0.0271542.

Contrasts

K_5 tests the second order ANCOVA term $\alpha\delta$.

```
k5={{0,0,0,0,0,0,0,1}}

ssk[k5,x4x4,b4]   //N
0.909571
```

The sum of squares associated with contrast K_5 is 0.909571 and it mean square remains the same because K_5 has only one degree of freedom. The corresponding F statistic for this contrast is obtained by dividing 0.909571 by 27.2517, the mean square error for this model. The resulting F value is 0.0333767 and its probability value is 0.856712. Hence, the ANOVA term α can be interpreted in the usual covariance manner given that its corresponding ANCOVA term is not significant at an alpha value of 0.25.

The $\beta\delta$ terms are tested next. For these contrasts, however, the design matrix above cannot be used because it does not estimate the $\beta\delta$ parameters. Thus, design matrix x_3 is used.

Basic Calculations

The basic calculations are the same as those computed for design matrix x_3.

Contrasts

K_6 tests the second order ANCOVA terms $\beta_1\delta$ and $\beta_2\delta$.

```
k6={{0,0,0,0,0,0,0,0,1,0},
{0,0,0,0,0,0,0,0,0,1}}

ssk[k6,x3x3,b3]   //N
90.1015

90.1015/2
```

45.0507

The sum of squares associated with contrast K_6 is 90.1015 and its mean square is 45.0507. The F statistic corresponding to this contrast is obtained by dividing 45.0507 by 25.4718, which is the mean square error for model x_3. The resulting value is 1.76865 and its probability value is 0.196216. This value is significant at an alpha value of 0.25.

<u>The Johnson-Neyman Method: Finding the Region of Significance--Reduced Design Matrix 5</u>

From the preceding an[illegible] ne non-covariate or ANOVA term had a corresponding ANCOVA t[illegible] as not significant. Thus, the α term can be interpreted without re[illegible] specific value of the covariate. On the other hand, β and $\alpha\beta$ had corresponding ANCOVA terms that were significant. Hence, β and $\alpha\beta$ must be interpreted with reference to specific covariate values. This suggests that the Johnson-Neyman method should be used.

Data Input

```
x5={{1,1,1,0,1,0,40,40,0,40,0},
{1,1,1,0,1,0,35,35,0,35,0},
{1,1,1,0,1,0,40,40,0,40,0},
{1,1,1,0,1,0,50,50,0,50,0},
{1,1,1,0,1,0,45,45,0,45,0},
{1,1,0,1,0,1,30,0,30,0,30},
{1,1,0,1,0,1,40,0,40,0,40},
{1,1,0,1,0,1,45,0,45,0,45},
{1,1,0,1,0,1,40,0,40,0,40},
{1,1,0,1,0,1,40,0,40,0,40},
{1,1,-1,-1,-1,-1,50,-50,-50,-50,-50},
{1,1,-1,-1,-1,-1,40,-40,-40,-40,-40},
{1,1,-1,-1,-1,-1,40,-40,-40,-40,-40},
{1,1,-1,-1,-1,-1,30,-30,-30,-30,-30},
{1,1,-1,-1,-1,-1,40,-40,-40,-40,-40},
{1,-1,1,0,-1,0,50,50,0,-50,0},
{1,-1,1,0,-1,0,30,30,0,-30,0},
{1,-1,1,0,-1,0,35,35,0,-35,0},
{1,-1,1,0,-1,0,45,45,0,-45,0},
{1,-1,1,0,-1,0,30,30,0,-30,0},
{1,-1,0,1,0,-1,50,0,50,0,-50},
{1,-1,0,1,0,-1,30,0,30,0,-30},
{1,-1,0,1,0,-1,40,0,40,0,-40},
{1,-1,0,1,0,-1,45,0,45,0,-45},
{1,-1,0,1,0,-1,40,0,40,0,-40},
{1,-1,-1,-1,1,1,45,-45,-45,45,45},
{1,-1,-1,-1,1,1,30,-30,-30,30,30},
{1,-1,-1,-1,1,1,25,-25,-25,25,25},
{1,-1,-1,-1,1,1,50,-50,-50,50,50},
{1,-1,-1,-1,1,1,35,-35,-35,35,35}}
```

The design matrix above represents a reduced model because $\alpha\delta$ has been removed since it was not significant; in other words, this term was pooled with the error term. The 11 columns of x_5 represent the same parameters as described before with the exception that $\alpha\delta$ has been excluded.

Basic Calculations

```
x5x5=Transpose[x5].x5

Inverse[x5x5]

x5y=Transpose[x5].y

b5=Inverse[x5x5].x5y   //N
{63.5101, -2.77672, -11.759, 17.295, -17.0887, 5.35673,
 0.716368, 0.376881, -0.439183, 0.409472, -0.10757}
```

The parameters as given by b_5 μ=63.5101, α_1=–2.77672, β_1=–11.759, β_2=17.295, $\alpha\beta_{11}$=–17.0887, $\alpha\beta_{12}$=5.35673, δ=0.716368, $\beta_1\delta$=0.376881, $\beta_2\delta$=–0.439183, $\alpha\beta_{11}\delta$=0.409472, $\alpha\beta_{12}\delta$=–0.10757.

Contrast: β

```
k7={{0,0,1,-1,0,0,0,p,-p,0,0},      (See Appendix 3)
{0,0,1,2,0,0,0,p,2p,0,0}}

d=ssk[k7,x5x5,b5]

Simplify[d]
              7                          2
(5.29619 10   (1547.81 - 78.0459 p + p )

                                2
    (1467.79 - 73.9983 p + p )) /

                                                      2
  (1241710687245 - 121302490290 p + 4482241741 p   -

                   3            4
    74231352 p   + 464996 p )
```

The contrast K_7 is a ratio of two polynomials (degree of seven and one unknown). In order to solve for this polynomial, $(K'B)'[K'(X'X)^{-1}K]^{-1}(K'B)$- r Mse $F_{0.05;dfk;dfe} = 0$, is again employed. First, the $(K'B)'[K'(X'X)^{-1}K]^{-1}(K'B)$ part of the equation is already known in the form of the polynomial d. Second, r Mse $F_{0.05;dfk;dfe}$ (r=2, dfk=2, dfe=19) is calculated as 2 x 21.84 x 3.57 = 155.94. Now, *Mathematica's* Solve operator is used to find the roots of the polynomial.

```
Solve[d-155.94  ==0,p]
{{p -> 38.2609 - 5.59376 I},
{p -> 38.2609 + 5.59376 I}, {p -> 38.4314},
{p -> 65.2595}}
```

Using the *Mathematica* Solve operator, the zero roots of equation d are solved and, thus, p is known. The p values are those specific covariate values for which significant treatment (β) differences appear at α=.05. Now, the region of significance may be plotted using *Mathematica's* Plot operator.

```
Plot[d-155.94,{p,30,80},AxesLabel->{"p","ss"}]
```

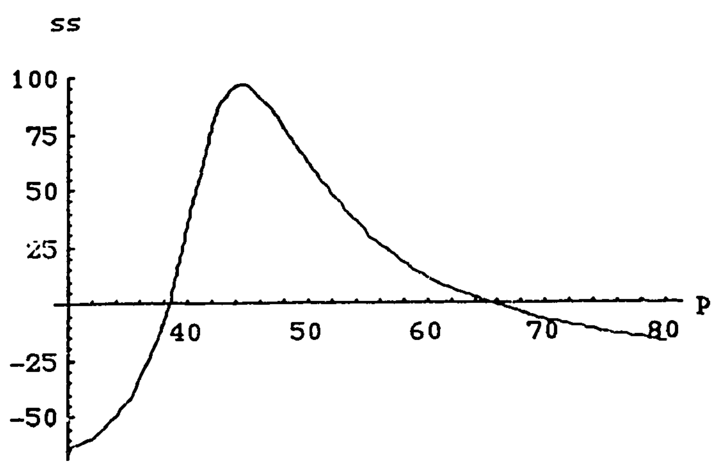


Figure 3.4. Region of Significance--Factor B

In Figure 3.4, the function intersects the horizontal axis at two points, p=38.43 and p=65.25. These two points demarcate the origins of the region of significance; the region where statistically significant treatment effects are observed. For example, all values of p for which the curve is above the horizontal axis produce statistically significant treatment effects at an α value of 0.05 or less. Hence, statistically significant effects among the three ability levels are observed when subjects have average covariate values equal to or greater than 38.43 and equal to or less than 65.25.

Contrast: $\alpha\beta$

The same procedure is followed to find the region of significance for the $\alpha\beta$ treatment effects as was followed to find the region of significance for the β treatment effects.

```
k8={{0,0,0,0,1,-1,0,0,0,p,-p},       (See Appendix 4)
{0,0,0,0,1,2,0,0,0,p,2p}}

v=ssk[k8,x5x5,b5]

Simplify[v]
           7                          2
(5.13934 10   (1711. - 82.6029 p + p )

                                 2
    (1562.22 - 78.1149 p + p )) /

                                                                   2
  (1182134225205 - 115470204070 p + 4267804355 p   -

                  3            4
    70720200 p    + 443394 p )
```

The sum of squares associated with contrast K_8 is again a polynomial of degree seven with an unknown value p. The polynomial is solved for its zero roots in the same way it was done for β.

```
Solve[v-110.00   ==0,p]
{{p -> 21.1638}, {p -> 39.3597 - 5.89808 I},
{p -> 39.3597 + 5.89808 I}, {p -> 83.5509}}
```

Now, the region of significance can be plotted. Notice that instead of using the value 155.94 in place of r x Mse x $F_{0.25;dfk;dfe}$ (r=2, dfk=2, dfe=19), the value 110.00 is used. The reason for this is that the alpha level has to be relaxed in this example (the F value changes too) because the region of significance is very small. Thus, if the 0.05 alpha value is used it is too conservative to detect any region of significance, consequently, when one attempts to solve for the zero roots of the polynomial v, *Mathematica* gives back only imaginary roots; that is, they are not mathematically real roots.

In Figure 3.5, the function intersects the horizontal axis at two points, p=21.16 and p=83.55. These two points demarcate the origins of the region of significance; the region where statistically significant interaction effects are observed. For example, all values of p for which the curve is above the horizontal axis produce statistically significant $\alpha\beta$ treatment effects at an α value of less than 0.05. Hence, statistically significant interaction effects are observed when subjects have an average covariate value equal to or less than 21.16 and equal to or greater than 83.55. However, notice that the covariate points, at which significant interaction effects are observed, lie outside the range of the covariate values in the data set used for this example.

```
Plot[v-110.96,{p,0,100},AxesLabel->{"p","ss"}]
```

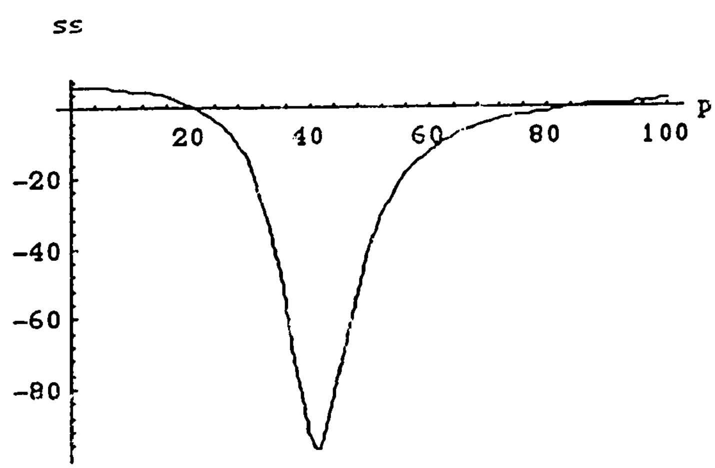


Figure 3.5. Region of Significance--Interaction AB

In this case, for all practical purposes, the groups do not differ on the dependent variable. Nevertheless, it is still of theoretical interest to know at which covariate values, if they had existed in the data set, the groups would have shown to be different.

Chapter IV
SPSS and Application of the Johnson-Neyman Method

SPSS does not specifically solve the Johnson-Neyman problem, i.e., the Johnson-Neyman method is not listed in SPSS as an optional analysis tool. However, it is possible to perform an ANCOVA with heterogeneous within group regression slopes on SPSS; that is, the researcher can calculate whether the within group regression slopes are heterogeneous on SPSS, but the researcher cannot obtain a region of significance anything of this sort on SPSS. Normally, a check for the homogeneity of within group regression slopes in an ANCOVA procedure is not done by SPSS and, consequently, the onus is on researchers to verify the homogeneity assumption before proceeding with a traditional ANCOVA design. If researchers do not know to check for homogeneity of within group regression slopes, and proceed with an ANCOVA, the results from the ANCOVA may be misleading.

A possible method by which to check for homogeneity of within group regression slopes in SPSS will be illustrated. This method is not explicitly described in SPSS manuals, and it is difficult to understand why it would not, considering that the homogeneity assumption in the application of an ANCOVA is critical.

For the two examples that will be used to show how one would go about testing for homogeneity of within group regression slopes in SPSS, the same data that was used to show how the Johnson-Neyman method was applied in *Mathematica*, is used. *Mathematica's* Example 3 will not be demonstrated because it becomes too complex to perform on SPSS. Data from Examples 1 and 2 are used.

Example 1: Two Groups and One Covariate

Data Input

The data are entered differently in SPSS from the way they are entered into *Mathematica*, for example, with two groups and one covariate, the following variables must be input into SPSS:

Group   Score   Covariate

From this input SPSS implicitly creates a column that reflects the covariate multiplied by group membership, which will assess the covariate by group interaction. If this interaction is significant it indicates that the within group regression slopes for the data set are heterogeneous. If the interaction is not significant, then there is no need to interpret treatment effects as a function of the covariate, and a traditional ANCOVA can be used.

Once users have set up their data matrix,

1. choose ANOVA MODEL option from the STATISTICS menu
2. choose GENERAL FACTORIAL within the ANOVA MODEL option
3. within GENERAL FACTORIAL specify the dependent variable, independent variable, and covariate(s); specify each variable by name -- dependent variable=score; independent variable=group; covariate=covariate
4. choose the MODEL option within GENERAL FACTORIAL and double click on CUSTOM (CUSTOM allows users to customize the model and this is needed because an additional effect--group*covariate--is required so that the homogeneity of within group regression slopes can be assessed.)
5. under the CUSTOM option, select the variables group, covariate, and request interaction
6. Execute the ANOVA analysis

The analysis proceeds and the results obtained from SPSS are given in Table 4.1.

Table 4.1.
Analysis of Covariance. Customized Model

* * * * * * A n a l y s i s   o f   V a r i a n c e -- design 1 * * * * * *

Tests of Significance for SCORE using UNIQUE sums of squares

Source of Variation	SS	DF	MS	F	Sig of F
WITHIN+RESIDUAL	10.92	26	.42		
GROUP	64.17	1	64.17	152.79	.000
COV	117.76	1	117.76	280.37	.000
COV * GROUP	59.60	1	59.60	141.90	.000
(Model)	166.05	3	55.35	131.78	.000
(Total)	176.97	29	6.10		

R-Squared = .938
Adjusted R-Squared = .931

The values produced by SPSS are similar to those obtained by *Mathematica*. In terms of the general linear model the model parameters that are being estimated (the effects shown in Table 4.1) are μ, α_1, δ, $\alpha_1\delta$. To better understand what these parameters and their values represent, each effect is explained in terms of the hypothesis tested. First, the group effect tests the null hypothesis $Ho{:}\alpha_1=0$ based on each groups' adjusted score on the dependent variable. This test is calculated at a covariate value of zero. The group effect in Table 4.1 indicates that the null hypothesis is rejected and the groups do differ in their adjusted scores at a covariate value of zero. Second, the covariate effect tests the null hypothesis that the pooled regression slope (for the two groups) is zero ($Ho{:}\delta=0$) in the context of the model parameters. The covariate effect in Table 4.1 indicates that the null hypothesis is rejected and the pooled regression slope is not likely equal to zero in the population; that is, it is likely relevant to include the covariate in the analysis because the relationship between the dependent variable and the covariate is different from zero. Finally, the COV*GROUP effect tests the null hypothesis that the within group regression slopes are homogeneous ($Ho{:}\alpha_1\delta=0$). The COV*GROUP effect in Table

4.1 indicates that the null hypothesis is rejected and the within group regression slopes are heterogeneous. In light of this effect, the treatment effects will not remain constant throughout the covariate distribution and should be examined at different covariate values.

Table 4.2 provides the results of the data analysis obtained from a non-customized model. That is, the additional effect--COVARIATE*GROUP--implicitly provided by SPSS was not chosen and no attempt was made to deviate from the regular, common SPSS ANCOVA procedure.

Table 4.2.

Analysis of Covariance. Regular Model

* * * * * A n a l y s i s o f V a r i a n c e -- design 1 * * * * * *

Tests of Significance for SCORE using UNIQUE sums of squares

Source of Variation	SS	DF	MS	F	Sig of F
WITHIN+RESIDUAL	70 52	27	2.61		
REGRESSION	97.92	1	97.92	37.49	.000
GROUP	5.96	1	5.96	2.28	.142
(Model)	106.45	2	53.23	20.38	.000
(Total)	176.97	29	6.10		

R-Squared = .602

Adjusted R-Squared = .572

In terms of the general linear model the model parameters that are being estimated (the effects shown in Table 4.2) are μ, α_1, δ. The REGRESSION effect shown on Table 4.2 tests the null hypothesis that the pooled regression slope is zero in the context of the model parameters. This effect is significant and so the null hypothesis is rejected and it can be concluded that the contribution of the regression is likely relevant in this analysis. The GROUP effect tests the null hypothesis (Ho: α_1=0) that the groups do not differ based on their adjusted dependent variable means at a covariate value of zero. From

Table 4.2, it is seen that the null hypothesis is not rejected and the groups do not differ at a covariate value of zero. What this analysis does not provide, however, is a test for the homogeneity of the within group regression slopes. In Table 4.1 this effect is tested by the null hypothesis, $\alpha_1\delta=0$, and it is significant. Thus, if users want to proceed with the type of analysis represented in Table 4.2, the onus is on them to check for homogeneity of the within group regression slopes. If users do not verify this, the mistake of concluding that the group effect is constant at all covariate values, could be made.

Although researchers can plot the within group regression slopes on SPSS and see how they intersect, there is no provision for plotting a region of significance and assessing at which points along the covariate distribution significant treatment effects are observed. Users interested in finding the points along the covariate distribution where significant treatment differences are found, must try to do so through trial and error by testing various covariate values. In addition, the group effect in Table 4.1 is different from the group effect in Table 4.2. The reason for this difference lies in the different models used to compute the effects. In the first analysis (Table 4.1) the model estimates parameters μ, α_1, δ, $\alpha_1\delta$, and in the second analysis (Table 4.2) the model estimates parameters μ, α_1, δ.

Example 2: Three Groups and Two Covariates

Data Input

With three groups and two covariates the model specification becomes more complex. For example, with two groups and two covariates, the following variables must be input into SPSS:

Group Score Cov1 Cov2

The user would follow the same procedure as in Example 1 to customize the model. The output SPSS generates following this analysis is given in Table 4.3.

Table 4.3.
Analysis of Covariance. Customized Model
* * * * * A n a l y s i s o f V a r i a n c e -- design 1 * * * * * *
Tests of Significance for SCORE using UNIQUE sums of squares

Source of Variation	SS	DF	MS	F	Sig of F
WITHIN+RESIDUAL	3.93	6	.65		
GROUP	11.76	2	5.88	8.98	.016
COV1	1.38	1	1.38	2.11	.197
COV2	1.78	1	1.78	2.72	.150
COV1 * GROUP	3.28	2	1.64	2.51	.162
COV2 * GROUP	3.34	2	1.67	2.55	.158
(Model)	75.00	8	9.38	14.32	.002
(Total)	78.93	14	5.64		

R-Squared = .950
Adjusted R-Squared = .884

The results from this analysis are similar to the results obtained from *Mathematica*. In terms of the general linear model the model parameters that are being estimated (the effects shown in Table 4.3) are $\mu, \alpha_1, \alpha_2, \delta_1, \delta_2, \alpha\delta_{11}, \alpha\delta_{12}, \alpha\delta_{21}, \alpha\delta_{22}$. Although the null hypotheses tested here are similar to those tested in Table 4.1, an explanation of each effect in Table 4.3 is still presented. First, the group effect tests the null hypothesis (α_1=0, α_2=0) that the three groups do not differ based on their adjusted values of the dependent variable at a covariate value of zero . From Table 4.3, it is clear that, at a covariate value of zero, the three groups do differ and the null hypothesis is rejected. Second, the covariate 1 effect tests the null hypothesis that the pooled regression slope for covariate 1 is zero (δ_1=0). From Table 4.3, and using a liberal alpha value of 0.25, the effect for covariate 1 can be considered significant and the null hypothesis is rejected. The covariate 2 effect tests essentially the same hypothesis (δ_2=0). The effect for covariate 2 is considered significant in light of the alpha value of 0.25. An important point to keep in mind when interpreting the effect of a pooled regression slope is that when within group regression slopes are very different (heterogeneous) and pooled (averaged), the act of pooling the slopes may mask or cancel their effects. For example, in a situation where two within group regression slopes are equal but opposite sign, taking the average of the slopes would result in a straight line. This result would misrepresent the true nature of the slopes. Third, the COV1*GROUP effect tests the null hypothesis that the within group regression slopes for covariate 1 are homogeneous (Ho:[$\alpha\delta_{11}$=0, $\alpha\delta_{12}$=0]). From Table 4.3, this effect is significant and indicates that the within group regression slopes for covariate 1 are heterogeneous. The COV2*GROUP effect tests essentially the same hypothesis (Ho:[$\alpha\delta_{21}$=0, $\alpha\delta_{22}$=0]). The effect for this test is also significant and it is concluded that the within group regression slopes for covariate 2 are heterogeneous. In light of the information that the within group regression slopes for both covariates are heterogeneous, treatment effects should be assessed at many different values along the covariate.

Table 4.4 provides the results produced by SPSS if the model had not been customized.

Table 4.4.
Analysis of Covariance. Regular Model

* * * * * A n a l y s i s   o f   V a r i a n c e -- design 1 * * * * * *

Tests of Significance for SCORE using UNIQUE sums of squares

Source of Variation	SS	DF	MS	F	Sig of F
WITHIN+RESIDUAL	15.09	10	1.51		
REGRESSION	.51	2	.25	.17	.847
GROUP	48.54	2	24.27	16.08	.001
(Model)	63.84	4	15.96	10.58	.001
(Total)	78.93	14	5.64		

R-Squared = .809
Adjusted R-Squared = .732

In terms of the general linear model the model parameters that are being estimated (the effects shown in Table 4.4) are μ, α_1, α_2, δ_1, δ_2. Once again, the regression effect tests the null hypothesis that the pooled regression slopes for covariate 1 and 2 *simultaneously* are zero, (Ho:[$\delta_1=0$, $\delta_2=0$]). The regression effect is not significant, but this result contradicts the pooled regression effects obtained in Table 4.3. From Table 4.3, the pooled regression slope for covariate 1 is considered significant at an alpha value of 0.25 as is the pooled regression slope for covariate 2. This regression effect in Table 4.4 may be masking the true nature of the pooled regression slopes for covariate 1 and 2. The group effect tests the null hypothesis that the three groups do not differ based on their adjusted dependent variable values at a covariate value of zero (Ho:[$\alpha_1=0$, $\alpha_2=0$]). From Table 4.4, the group effect is significant and the null hypothesis is not accepted. However, given that the within group regression slopes are heterogeneous, this group effect cannot be generalized to the entire range of the covariate distribution.

Comparison of *Mathematica* and SPSS: Advantages and Limitations

When performing the Johnson-Neyman method, *Mathematica* is preferable to SPSS because it gives the researcher a method by which to obtain the region of significance. Since the region of significance is likely the most important aspect of the Johnson-Neyman method that sets it apart from the ANCOVA, the computer package used to perform the Johnson-Neyman method should be capable of producing this important information.

It would be a mistake for researchers, however, to conclude that because *Mathematica* is the better package to use for the Johnson-Neyman method that it is the better package when it comes to data analysis in general. The effectiveness of a computer package depends on a number of factors. Some of these factors are: statistical knowledge of the user, speed, type of analysis desired, and the "user-friendliness" of the package. These factors, taken together, can determine whether a particular computer package is desirable to the user.

Both *Mathematica* and SPSS have strengths and weaknesses, albeit in different areas. In addition, many of the advantages of a particular statistical package can be disadvantages if they do not match the preferences of the user. For example, one of *Mathematica's* advantages is that researchers can easily input whatever specific statistical model they wish to use in the analysis of their data. The benefit of this lies not only in the flexibility that it provides the researcher, but also that it allows one to always know exactly what is being tested and how it is being tested. As the examples with *Mathematica* show, after the data are entered, the user must continue to specify every step thereafter. This control over the path the data analysis takes gives the user a great deal of flexibility. However, in using *Mathematica*, the user must be comfortable with many of the details of the statistics employed, as well as the semantics of *Mathematica*, and it is these prerequisites that may inhibit some researchers from using *Mathematica*.

On the other hand, SPSS does not make the same demands of users. In SPSS, after the data are entered, users do not have to input anything else. After the data are entered, users must make selections from a series of menus as to how they want to proceed next. So, if researchers want to perform an ANCOVA, they select this analysis tool from a specific menu (they may also choose particular options within the ANCOVA menu), indicate the variables to be included in the analysis, and execute the procedure. Of course, one has to have a certain amount of statistical knowledge because one will ultimately have to interpret what SPSS provides as output, but users do not need to construct any design matrices to test specific models or contrasts. Although, this is an advantageous feature for some SPSS users, it can also be a disadvantage for those users that would like more of an interactive, flexible software program in which to analyze their data. In addition, SPSS can run applications and produce output that can be misleading for users who lack the knowledge behind some of the statistical procedures they execute. For example, it was shown that an ANCOVA with heterogeneous within group regression slopes could be run on SPSS as a traditional ANCOVA. There were no cautionary messages from SPSS that the within group regression slopes were heterogeneous. The onus, then, is on the user to check this assumption beforehand.

In terms of speed, type of analysis desired, and "user-friendliness," *Mathematica* can be slow if the procedure is complex and requires a great amount of computer memory, but on the whole it is very fast in most analyses. Use of *Mathematica* is geared towards users who want more degrees of freedom in how they go about manipulating and analysing their data regardless of the type of analysis that is desired. The user-friendliness of *Mathematica* is contingent on the user. For example, for those users who may not enjoy inputting their own models and contrasts, *Mathematica* may not be for them. On the other hand, for researchers who cherish the flexibility it provides and enjoy inputting their own models, *Mathematica* can be very effective.

SPSS is also fast in its application of procedures and is geared towards persons who are satisfied with the options that SPSS provides. In addition, more sophisticated users can input SPSS commands themselves by using the syntax window (although, the user can only use commands that SPSS recognizes). Lastly, many persons find SPSS attractive because they are able to quickly execute statistical procedures.

The best conclusion is to be familiar with whatever statistical package is being used and with the statistical techniques being performed. When considerable effort is put into the background and research design of a study, the same effort should be extended to the data analysis.

Chapter V
Summary and Discussion

This study had three goals. First, to briefly describe ANCOVA and the Johnson-Neyman method, including each methods' organization of data, assumptions, history and development, advantages and disadvantages, and importance of its understanding. Second, to demonstrate how the Johnson-Neyman method is performed with available computer software--*Mathematica* --because this statistical method is not as frequently performed as is ANCOVA. Finally, to demonstrate how to perform a technique similar to the Johnson-Neyman method on a widely used statistical software package such as SPSS, and to discuss which computer package is preferable to use for applying the Johnson-Neyman method.

There are two conclusions that may be garnered from the present study. First, that it is important to understand and apply appropriate statistical methods to research data since the results obtained and the conclusions drawn from analyses are contingent upon the very statistical methods used to analyze data. For example, as Huitema (1980) asserts, one of the most important assumptions related to the use of ANCOVA is the homogeneity of within group regression slopes. The violation of this assumption is of great concern since the results obtained from an inappropriate application of ANCOVA can be very misleading, e.g., when there appears to be results from a traditional ANCOVA indicating no treatment effects between groups, and this result is not constant for all covariate points along the covariate distribution (see Figure 1.2).

In the cases where researchers know that within group regression slopes are heterogeneous, it is advantageous, if not critical, to know that the Johnson-Neyman method should be used instead. Although, the Johnson-Neyman method is not as frequently or as easily applied as is ANCOVA, it is to the researcher's advantage to know of the Johnson-Neyman method and how it can be used to accurately organize and analyze the data.

A second conclusion that can be drawn from this study is that the Johnson-Neyman method can be applied to data effectively when the computer software package *Mathematica* is used. Although, SPSS can be used to do a technique similar to the Johnson-Neyman method, the researcher cannot produce a region of significance. Moreover, the region of significance is the essential product of the Johnson-Neyman method and that which makes the Johnson-Neyman method so attractive. For example, on SPSS one can do an ANCOVA so that it also tests the validity of homogeneous within group regression slopes i.e., show whether the slopes are significantly different or not. However, the treatment effects are still calculated at a covariate value of zero only. So, the results of such an analysis should simply caution researchers from assuming that the treatment effect remains the same at all points along the covariate distribution. These results, however, help researchers only half-way. What researchers need to most accurately interpret the results (treatment effect(s)) of their studies is the region of significance; that is, where along the covariate distribution are there significant treatment effects. For this main reason--the inability of SPSS to produce a region of significance--*Mathematica* is the software of choice in performing the Johnson-Neyman method.

Finally, the advantages of using *Mathematica* for solving the Johnson-Neyman method should be emphasized. Because *Mathematica* is an interactive computer software program, it provides users with a great deal of flexibility and control over their analyses, especially important when the technique or problem that is being solved is complex. Although it can be argued that all software programs (including statistical software packages) are, to some extent, interactive because users must at least choose options from menus, *Mathematica* is at the high end of an interactive program because users are not choosing options, but actually constructing their own models and the path their analysis will take. In addition, *Mathematica* does supply users with a multitude of operators that can be used to construct, evaluate, and analyze a variety of expressions. For example, in Chapter three it was shown how *Mathematica's* Solve operator could be used to solve the polynomial needed to obtain the region of significance for two groups and one covariate. Moreover, *Mathematica's* graphing capabilities make possible obtaining the region of significance when an algebraic solution is not possible, such as in the three group and two covariate example.

Now that the procedures and advantages of using the Johnson-Neyman method, and *Mathematica* for solving the Johnson-Neyman method have been presented in this study, what is required to increase the use of the Johnson-Neyman method? The answer to this question is not an easy one because there are many reasons why particular statistical methods do not "catch on" in use. The first explanation for why the Johnson-Neyman method may not be used as much as it could is that *Mathematica* may demand too much statistical knowledge and "input" from the user. It is true that researchers may be very turned off by the idea that they have to construct their own models and contrasts and the time this could potentially take. The second reason is likely more true of human nature and it is the adherence to familiar ways of doing things, such as analyzing data with familiar methods and computer packages, and the resistance to new techniques.

Some ways that these two obstacles to the greater use of the Johnson-Neyman method and *Mathematica* could be overcome is to present and distribute the contents of this study in different scholarly journals. For example, it may be helpful to have a technical paper for users who already know about the Johnson-Neyman method and want to know how the method can be performed on *Mathematica*, and another paper for potential users who may not know of the Johnson-Neyman method, that presents the benefits of the Johnson-Neyman method to their research. In addition, researchers could start having students use *Mathematica* in class assignments which include the Johnson-Neyman procedure.

References

Abell, M.L. & Braselton, J.P. (1994). *The Maples V Handbook, AP Professional.* Boston: Academic Press.

Abelson, R. P. (1953). A note on the Neyman-Johnson technique. *Psychometrika, 18* (3), 213-2

Aitkin, M. A. (1973). Fixed-width confidence intervals in linear regression with applications to the Johnson-Neyman technique. *British Journal of Mathematical and Statistical Psychology, 26*, 261-269.

Bliss, C. I. (1970). *Statistics in biology.* Toronto: McGraw-Hill Book Company.

Carroll, J. B. & Wilson, G. (1970). An interactive-computer program for the Johnson-Neyman technique in the case of two groups, two predictor variables, and one criterion variable. *Educational and Psychological Measurement. 30*, 121-132.

Cochran, W.G. (1957). Analysis of covariance: Its nature and uses. *Biometrics, 13*, 261-281.

Cook, T. D. & Campbell D. T. (1979). *Quasi-experimentation. Design and analysis issues for field settings.* Boston: Houghton Mifflin Company.

Cox, D.R. & McCullagh, P. (1982). Some aspects of analysis of covariance. *Biometrics, 38*, 541-561.

Cronbach, L. J. & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbookfor research interactions.* Irvington Publishers, INC: Toronto.

Elashoff, J. D. (1969). Analysis of covariance: A delicate instrument. *American Educational Research Journal, 6*(3), 383-401.

Evans, S. H., & Anastasio, E. J. (1968). Misuse of analysis of covariance when treatment effect and covariate are confounded. *Psychological Bulletin, 69*, 225-234.

Fisher, R. A. (1932). *Statistical methods for research workers.* Edinburgh: Oliver and Boyd.

Fisher, R. A. (1935). *The design of experiments.* Edinburgh: Oliver and Boyd.

Gafarian, A. V. (1964). Confidence bands in straight line regression. *Journal of the American Statistical Association, 59,* 182-213.

Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of Educational Research, 42*(3), 237-288.

Huitema, B. E. (1980). *The analysis of covariance and alternatives.* New York: John Wiley and Sons.

Hendrix, L. J., Carter, M. W. & Short, D.T. (1982). Covariance analyses with heterogeneity of slopes in fixed models. *Biometrics, 38,* 641-650.

Hunka, S. (1993). *Univariate analysis student notes.* Unpublished notes, University of Alberta, Department of Educational Psychology, CRAME, Edmonton.

Hunka, S. (1995), Identifying regions of significance in ANCOVA. Problems having nonhomogeneous regressions. *British Journal of Mathematical and Statistical Psychology, 48,* 161-188.

Johnson, P. O. & Fay, L. C. (1950). The Johnson-Neyman technique, its theory and applications. *Psychometrika, 15* (4), 349-367.

Johnson, P. O. & Neyman, J. (1936). Tests of certain linear hypotheses and their application to some educational problems. *Statistical Research Memoirs, 1,* 57-93.

Maguire, T. O., & Haig, B. D. (1976). Problems of control in nonexperimental educational research. *The Alberta Journal of Educational Research,* 22(4), 289-296.

Marascuilo, L.A. & Serlin, R.C. (1988). *Statistical methods for the social and behavioral sciences.* New York: W.H. Freeman and Company.

Maxwell, S. E., O'Callahan, M. F., & Delaney, H. D. (1993). Analysis of covariance. In L. K. Edwards (Ed.), *Applied analysis of variance in behavioral science* (pp 63-104). New York: Marcel Dekker, Incorporated.

Pearce, S. C. (1983). Analysis of covariance. In *Encyclopedia of Statistical Sciences*, Eds. S. Kotz and N.L. Johnson, New York: John Wiley & Sons, *4*, 61-69.

Pigache, R. M., Graham, B. R. & Freedman, L. (1976). A modification of the Johnson-Neyman technique comparing two regressions, applied to treatment effects dependent on baseline levels. *Biological Psychology, 4*, 213-236.

Potthoff, R. F. (1964). On the Johnson-Neyman technique and some extensions thereof. *Psychometrika, 29* (3), 241-256.

Potthoff, R.F. (1983). Johnson-Neyman technique. In *Encyclopedia of Statistical Sciences*, Eds. S. Kotz and N.L. Johnson, New York: John Wiley & Sons, *4*, 299-303.

Rogosa, D. (1980). Comparing nonparallel regression lines. *Psychological Bulletin, 88* (2), 307-321.

Rogosa, D. (1981). On the relationship between the Johnson-Neyman region of significance and statistical tests of parallel within-group regressions. *Educational and Psychological Measurement, 41*, 73-84.

Rutherford, A. (1992). Alternatives to traditional analysis of covariance. *British Journal of Mathematical and Statistical Psychology, 45*, 197-223.

Searle, S.R. (1971). *Linear models*. New York: John Wiley & Sons.

Smith, H.F. (1957). Interpretation of adjusted treatment means and regressions in analysis of covariance. *Biometrics, 13*, 282-308.

SPSS Incorporated. (1989-1994). *SPSS 6.1 for windows*. Chicago: SPSS Incoporated.

Statistical Analysis System (SAS) Institute. (1982). *SAS user's guide: Satatistics*. North carolina: SAS Institute.

Wolfram, S. (1988). *Mathematica: A system for doing mathematics by computer*. Addison-Wesley Publishing Co.: Menlo Park, CA.

Appendix 1

Contrast K_2 In Example 1

In Example 1 contrast K_2 is given as

	μ	α_1	δ_1	δ_2
$k_2=$	0	2	p	- p

There are two regression lines (for the two groups) and the problem is to express SS_k as a function of the covariate values p following the general form of the null hypothesis given as $K'\beta=0$ in which β is estimated by B. The required hypothesis can be expressed as

$$(\alpha_1 + \delta_1 p) - (\alpha_2 + \delta_2 p) = 0$$

The left-hand side provides

$$\alpha_1 + \delta_1 p - (\alpha_2 + \delta_2 p) = \alpha_1 + \delta_1 p - \alpha_2 - \delta_2 p$$

because for the Σ-restricted model $\Sigma\alpha_j = 0$, then

$$\alpha_2 = -\alpha_1$$

$$\alpha_1 + \delta_1 p + \alpha_1 - \delta_2 p = 2\alpha_1 + \delta_1 p - \delta_2 p$$

The last statement provides the values in the contrast matrix to be 0 for μ, 2 for α_1 , p for δ_1, and -p for δ_2 .

Appendix 2

Contrast K_2 In Example 2

In Example 2 the contrast matrix K_2 is given as

	μ	α_1	α_2	δ_{11}	δ_{21}	δ_{31}	δ_{12}	δ_{22}	δ_{32}
$k_2=$	0	1	-1	p	-p	0	q	-q	0
	0	1	2	0	p	-p	0	q	-q

There are three regression lines (for the three groups) and the problem is to express SS_k as a function of the covariate values p for the the first covariate, and q for the second covariate, following the general form of the null hypothesis given as $K'\beta=0$ in which β is estimated by B. Given the 3 terms

(1) $\alpha_1 + \delta_{11}p + \delta_{12}q$

(2) $\alpha_2 + \delta_{21}p + \delta_{22}q$

(3) $\alpha_3 + \delta_{31}p + \delta_{32}q$

the required hypothesis can be expressed in two statements as

$$(\alpha_1 + \delta_{11}p + \delta_{12}q) - (\alpha_2 + \delta_{21}p + \delta_{22}q) = 0$$

$$(\alpha_2 + \delta_{21}p + \delta_{22}q) - (\alpha_3 + \delta_{31}p + \delta_{32}q) = 0$$

Expanding and collecting the terms,

$(1) - (2) = \alpha_1 + \delta_{11}p + \delta_{12}q - (\alpha_2 + \delta_{21}p + \delta_{22}q) = \alpha_1 + \delta_{11}p + \delta_{12}q - \alpha_2 - \delta_{21}p - \delta_{22}q$

$(2) - (3) = \alpha_2 + \delta_{21}p + \delta_{22}q - (\alpha_3 + \delta_{31}p + \delta_{32}q) = \alpha_2 + \delta_{21}p + \delta_{22}q - \alpha_3 - \delta_{31}p - \delta_{32}q$

Because $\Sigma\alpha_j = 0$, then

$$\alpha_3 = -\alpha_2 - \alpha_1$$

and substituting for α_3 in (2)-(3)

$$\alpha_2 + \delta_{21}p + \delta_{22}q + \alpha_2 + \alpha_1 - \delta_{31}p - \delta_{32}q = \alpha_1 + 2\alpha_2 + \delta_{21}p - \delta_{31}p + \delta_{22}q - \delta_{32}q$$

The first row of the contrast matrix is given by the coefficients of (1)-(2), and the second row by (2)-(3).

Appendix 3

Contrast K_7 In Example 3

In Example 3 the contrast matrix K_7 is given as

Appendix 4

Contrast K_8 In Example 3

The contrast matrix K_8 in Example 3 is given as

	μ	α_1	β_1	β_2	$\alpha\beta_{11}$	$\alpha\beta_{12}$	δ	$\beta_1\delta$	$\beta_2\delta$	$\alpha\beta_{11}\delta$	$\alpha\beta_{12}\delta$
$k_8 =$	0	0	0	0	1	-1	0	0	0	p	-p
	0	0	0	0	1	2	0	0	0	p	2p

Using the same general procedure as used in Appendix 3, the following results:

(1) $\alpha\beta_{11} + \alpha\beta_{11}\delta p$

(2) $\alpha\beta_{12} + \alpha\beta_{12}\delta p$

(3) $\alpha\beta_{13} + \alpha\beta_{13}\delta p$

Test of Ho: is that (1) - (2) = 0 and (2) - (3) = 0

$\Sigma\alpha\beta_{jk} = 0$ and $\Sigma\beta_{jk}\delta = 0$ (j=1;k=1,2,3) so,

$\alpha\beta_{13} = -\alpha\beta_{12} - \alpha\beta_{11}$ and

$\alpha\beta_{13}\delta = -\alpha\beta_{12}\delta - \alpha\beta_{11}\delta$

$$(1) - (2) = \alpha\beta_{11} + \alpha\beta_{11}\delta p - (\alpha\beta_{12} + \alpha\beta_{12}\delta p) = \alpha\beta_{11} + \alpha\beta_{11}\delta p - \alpha\beta_{12} - \alpha\beta_{12}\delta p$$

$$(2) - (3) = \alpha\beta_{12} + \alpha\beta_{12}\delta p - (\alpha\beta_{13} + \alpha\beta_{13}\delta p) = \alpha\beta_{12} + \alpha\beta_{12}\delta p - \alpha\beta_{13} - \alpha\beta_{13}\delta p$$

$$= \alpha\beta_{12} + \alpha\beta_{12}\delta p - (-\alpha\beta_{12} - \alpha\beta_{11}) - (-\alpha\beta_{12}\delta p - \alpha\beta_{11}\delta p)$$

$$= \alpha\beta_{12} + \alpha\beta_{12}\delta p + \alpha\beta_{12} + \alpha\beta_{11} + \alpha\beta_{12}\delta p + \alpha\beta_{11}\delta p$$

$$= \alpha\beta_{11} + 2\alpha\beta_{12} + \alpha\beta_{11}\delta p + 2\alpha\beta_{12}\delta p$$

The coefficients of (1)-(2) provide for row one, and (2)-(3) provide for row two of the contrast matrix.