A PRELIMINARY TEST ESTIMATOR

FOR MULTIVARIATE RESPONSE FUNCTIONS

by

Paul West Blackmon, Jr.

Dissertation submitted to the Graduate Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Statistics

APPROVED:

R. H. Myers, Chairman

J. C. Arnold

B. Harshbarger

K. H. Hinkelmann

C. Y. Kramer

December 1974

Blacksburg, Virginia

ACKNOWLEDGMENTS

I wish to express my deep appreciation to Dr. Raymond H. Myers for his guidance throughout this research. Thanks are also due Dr. Klaus H. Hinkelmann for his careful co-reading of the dissertation. I am grateful to Dr. Boyd Harshbarger for his assistance in obtaining financial aid.

Finally, I am indebted to Mrs. for her excellent typing of the manuscript.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter I

INTRODUCTION

Response functions are the target variables which arise in all experimental design systems. An industrial researcher will be interested in particular responses such as process yield or operation cost, which occur as functions of k independent variables subject to the control of the experimenter. This relationship can be represented as

$$\eta = f(\xi_1, \xi_2, \ldots, \xi_k) \tag{1.1.1}$$

for some response η.

In the conduct of an experiment, the natural variables ξ_1, ξ_2, ..., ξ_k must be confined to a region of interest R limiting their range. As an example in dealing with two factors, if ξ_1 and ξ_2 represent reaction temperature and amount of reactant present respectively, R might be taken as the region 100°C $\leq \xi_1 \leq$ 200°C, 5 grams $\leq \xi_2 \leq$ 15 grams. For mathematical convenience it is frequently desirable to deal with coded or design variables x_1, x_2, ..., x_k obtained from the original variables by a simple linear transformation. Often this transformed region is taken to be the hypersphere defined by $\sum_{i=1}^{k} x_i^2 \leq 1$, or a hypercube such that $-1 \leq x_i \leq 1$, i = 1, 2, ..., k. In the two variable example, the transformed variables become

$$x_1 = \frac{\xi_1 - 150}{50}, \quad x_2 = \frac{\xi_2 - 10}{5}.$$

The standard representation of the response is then

$$\eta = f(x_1, x_2, \ldots, x_k) \; . \tag{1.1.2}$$

The exact form of the relationship in (1.1.2) will be unknown; the usual practice is to approximate it by a polynomial of low degree within R. A linear or first order approximation might be

$$\eta = \beta_{1(0)} + x_1\beta_{1(1)} + x_2\beta_{1(2)} \tag{1.1.3}$$

where the β's are unknown parameters and must be estimated. A corresponding possible second order model would be

$$\begin{aligned} \eta = \beta_{1(0)} + x_1\beta_{1(1)} + x_2\beta_{1(2)} + x_1^2\beta_{2(1)} \\ + x_2^2\beta_{2(2)} + x_1x_2\beta_{2(12)} \; . \end{aligned} \tag{1.1.4}$$

The extended notation is necessary to avoid confusion with subsequent models.

Although the design variables x_1, x_2, ..., x_k are fixed by the experimenter and assumed to be measured with negligible error, the response is also dependent on the constant coefficient parameters, the unknown β's. In order to estimate these parameters for a single response function, N observations of η are made, resulting in an estimator for the response itself. For a response y, a linear estimated response function is

$$\hat{y} = \hat{\beta}_{1(0)} + \sum_{i=1}^{k} x_i\hat{\beta}_{1(i)} \tag{1.1.5}$$

where the $\hat{\beta}$'s are estimators of the true β's. Similarly, a full second order response is estimated as

$$\hat{\hat{y}} = \hat{\hat{\beta}}_{1(0)} + \sum_{i=1}^{k} x_i \hat{\hat{\beta}}_{1(i)} + \sum_{i=1}^{k} x_i^2 \hat{\beta}_{2(i)} + \sum_{i<j}\sum x_i x_j \hat{\beta}_{2(ij)} , \quad (1.1.6)$$

the $\hat{\hat{\beta}}$'s denoting estimators of the β's obtained when using a higher order model.

Now since η as given by (1.1.4) for example, is merely (1.1.3) plus the addition of higher order terms which should or should not be included in the model, on what basis should an experimenter choose one model over the other? Clearly, it is of central interest to accurately specify an approximating polynomial for the response function. The essential problem is how best to estimate this response so that ultimately, efficient determinations of η can be made using (1.1.2) as a prediction equation. An experimenter might employ a low, perhaps first order model. Alternatively, he could use a model of greater order containing some degree of curvature. He might wish to effect a compromise between these two extremes.

Preliminary test estimation is a widely used tool in statistics. It occurs most frequently in analysis of variance pooling procedures based on tests of hypotheses that particular variance components are negligible. It is quite natural to apply this general technique as an aid in developing a preliminary test estimator for the response. The general procedure will be to select $\hat{y}$ or $\hat{\hat{y}}$ contingent upon the results of a test of hypothesis consistent with the objective of estimating η with some degree of precision. It should be noted that there is no restriction on the order or form of the polynomial estimators $\hat{y}$ and $\hat{\overset{*}{y}}$, only that approximations of low degree of the type given by (1.1.5) and (1.1.6) are common in practice. Past researchers have primarily concentrated their investigations on preliminary tests of significance of higher order coefficients, either sequentially or as a whole, e.g., testing the hypothesis that $\beta_{2(1)}$, $\beta_{2(2)}$, and $\beta_{2(12)}$ in (1.1.4) are equal to zero. Either $\hat{\overset{*}{y}}$ or $\hat{y}$ would be chosen according to whether this hypothesis is rejected or not. We propose to construct a preliminary test estimator around a more meaningful hypothesis centered on the quality of estimation of the response. This hypothesis and the criterion of estimation on which it is founded will be examined in great detail in Chapter III.

Although this estimation criterion is peculiar to the body of statistical techniques known as response surface methodology, our models are within the framework of regression analysis, with the restriction that design level combinations are confined to the factor space R. In particular, we have outlined a univariate regression approach since in taking the observations on η, a single N x 1 response vector $\underline{y}$ can be formed, all observations considered as being similar polynomial functions of the same set of design variables, coefficient parameters, and corresponding experimental error terms. The representations $\hat{y}$ and $\hat{\overset{*}{y}}$ simply designate estimators of a typical individual response in the vector $\underline{y}$.

Often it is desirable to simultaneously treat not one but several N x 1 observation vectors $\underline{y_1}$, $\underline{y_2}$, ..., $\underline{y_p}$. For a given j, j = 1, 2, ..., p, each $\underline{y_j}$ is a univariate regression with the additional stipulation that there exists a covariance structure among them. For this multivariate regression model, it is frequently of interest to consider appropriate linear combinations of estimators of a single response function in each of the p models. These estimators may or may not reduce to the estimators obtained by treating each $\underline{y}_j$ as a separate univariate problem, depending upon factors such as covariance structure assumed and type of multivariate model involved. As was the case with a single univariate regression, we wish to formulate a preliminary test estimator constructed around the control of certain properties of a linear combination of estimated response functions.

Although each response vector has a unique set of coefficient parameters and error terms associated with it, this need not be the case with the design variables corresponding to a particular $\underline{y}_j$; however, the standard multivariate regression model does in fact postulate the same design for all p observation vectors. This design, of course, consists not only of the N design level combinations, but also higher order terms as functions of the basic design variables. Returning to the two factor example, it is now practicable to deal with two separate responses, quantity of yield of products A and B, say. The experimenter may wish to employ the same design for these product yields, both being dependent upon ξ_1, reaction temperature, and ξ_2, amount of reactant, a situation implying the use of the standard model.

Alternatively, suppose that a researcher is investigating the process yield of a given product from data acquired from experiments conducted by two different companies. In all probability, the firms will have used different combinations of levels of the design variables. In addition, they might have projected dissimilar models, both in degree and number of design variables. This is termed a generalized multivariate regression model, the distinguishing feature being a different design for each $\underline{y}_j$, all other conditions being equivalent to those of the standard model.

The two multivariate regression models discussed enable one to accommodate any number of responses of interest. In order to generate preliminary test estimators in both of these instances, we will devote considerable space to the development of a statistic for testing the hypothesis on which the estimators are based. As a consequence of model assumptions and covariance structure, several important special cases will be dealt with in detail. Graphical comparisons will be presented on the performance of our estimators relative to that of the estimators obtained under a test of the standard hypothesis. These comparisons also enable one to select an operating range of type I error probabilities with which to conduct a preliminary test.

Chapter II

REVIEW OF LITERATURE

Due to the widespread use of preliminary test estimation techniques in many areas of statistics as pointed out in Chapter I, we shall confine ourselves to a discussion of these procedures as they relate to regression functions within the framework of response surface methodology. This leads quite naturally to consideration of appropriate criteria by which to compare these estimated response functions.

One of the first investigators to look at estimators of this sort was Bancroft (1944). Basing a preliminary test procedure on the hypothesis H: $\beta_2 = 0$ when $\eta = x_1\beta_1 + x_2\beta_2$, he suggested the estimator

$$\hat{\beta} = \begin{cases} \hat{\tilde{\beta}}_1 & \text{if H is rejected} \\ \hat{\beta}_1 & \text{otherwise ,} \end{cases}$$

where $\hat{\tilde{\beta}}_1$ is the unrestricted least squares estimator of β_1, and $\hat{\beta}_1$ is the least squares estimator of β_1 under H. Utilizing normality assumptions, he also obtained the bias of $\hat{\beta}$, and tabulated this as a function of selected parameter values. The estimation procedure was extended to k variables by Bancroft (1950) in the treatment of subsets of the coefficient parameters in the linear model.

An interesting variation of this technique although still applied to first order models, was presented by Larson and Bancroft (1963a). A sequential procedure was developed whereby variables are consecutively deleted from the model if one fails to reject the hypothesis that the corresponding regression coefficient is zero. An inverse approach involves the sequential addition of variables to the model, again based upon repeated tests of significance. In both instances, the bias and mean squared error of the resulting estimators of the response function were determined and tabled.

A second paper by Larson and Bancroft (1963b) dealt with the bias and mean squared error of the estimator obtained under the more traditional procedure, i.e., testing the joint hypothesis that all uncertain coefficients are simultaneously zero.

An important contribution to the somewhat more general problem was made by Toro-Vizcarrondo and Wallace (1968). Using the framework of the general linear model, they introduced the hypothesis that the mean squared error for any non-zero linear combination of the regression parameters in $\hat{\beta}$ is greater than or equal to the mean squared error of the same linear combination subject to linear restrictions on the coefficient space. Employing the standard test statistic used in testing general linear hypotheses on parameter coefficients, the mean squared error hypothesis was shown to be equivalent to a test on the noncentrality parameter of the noncentral F distribution arising from the standard statistic under error normality assumptions. It was further shown that this method is a uniformly most powerful test for their reduced hypothesis.

Kennedy and Bancroft (1971) conducted extensive numerical investigations into the ratios of mean squared errors of the two sequential procedures, concluding that "sequential deletion" is to be preferred over "forward selection." In relation to an optimum range of test parameters, they also studied the relative efficiencies of the two procedures to that of retaining all uncertain variables in the fitted equation.

Still within the context of a single response vector, Ellerton (1973) developed a family of test statistics for the hypothesis that the integrated mean squared error of $\hat{\overset{*}{y}}$ is greater than or equal to the integrated mean squared error of $\hat{y}$, the integration being carried out over the factor space R. Under the assumption that the true model may contain terms in addition to those of $\hat{\overset{*}{y}}$, he determined a general expression for the integrated mean squared error of the response function estimator based upon the above hypothesis.

CHAPTER III

STANDARD MULTIVARIATE REGRESSION MODEL

3.1 The Problem in Detail

We wish to determine the form of p multivariate response functions which depend on known design variables restricted to some region of interest R. Let $\underline{y_1}, \underline{y_2}, \ldots, \underline{y_p}$ represent N x 1 vectors of independent observations. Using the framework of the general linear model, we postulate a model (linear in the parameters $\underline{\beta_{1j}}$) of the form

$$\underline{y_j} = \underset{N x q_1}{\underset{\sim}{X}_1 \underline{\beta_{1j}}} + \underline{\varepsilon_j} \quad , j = 1, 2, \ldots, p, \tag{3.1.1}$$

where $\text{cov}(\underline{\varepsilon_i}, \underline{\varepsilon_j}) = \sigma_{ij}\underset{\sim}{I}_N$, and $\underset{\sim}{X}_1$ consists of a column of 1's along with the N experimental combinations of the design variables with their powers and cross-products if applicable. Thus, the p observation vectors are correlated, and for a particular j, the model (3.1.1) is a univariate regression. In order to deal with the problem in a multivariate context, the basic model may be expressed more compactly as

$$\underline{y} = \underset{\sim}{X}\, \underline{\beta_1} + \underline{\varepsilon} \tag{3.1.2}$$

where $\underline{y}' = [\underline{y_1'}, \underline{y_2'}, \ldots, \underline{y_p'}]$

$$\underset{Npxpq_1}{\underset{\sim}{X}} = \text{diag}\,[\underset{\sim}{X}_1, \underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_1]$$

$$\underline{\beta_1'} = [\underline{\beta_{11}'}, \underline{\beta_{12}'}, \ldots, \underline{\beta_{1p}'}]$$

$$\underline{\varepsilon'} = [\underline{\varepsilon_1'}, \underline{\varepsilon_2'}, \ldots, \underline{\varepsilon_p'}] .$$

The true model, however, insofar as can be determined, may contain terms not specified in (3.1.1). We denote this by

$$\underline{y_j} = \underset{\sim}{X}_1 \underline{\beta_{1j}} + \underset{\substack{\sim \\ Nxq_2}}{X_2} \underline{\beta_{2j}} + \underline{\varepsilon_j} \quad , j = 1, 2, \ldots, p, \tag{3.1.3}$$

where $\underset{\sim}{X}_2$ consists of the q_2 contributions to the response over and above those of the basic model. Hereafter, we shall refer to this and similar models as the true model, although we can rarely ascertain the exact form of the true relationship. Equation (3.1.3) can be further consolidated to

$$\underline{y_j} = \underset{\sim}{X}_o^* \underline{\beta_j^*} + \underline{\varepsilon_j} \tag{3.1.4}$$

for $\underset{\substack{\sim \\ Nxq_o}}{X_o^*} = [\underset{\sim}{X}_1 : \underset{\sim}{X}_2]$

$$q_o = q_1 + q_2$$

$$\underline{\beta_j^*}' = [\underline{\beta_{1j}'}, \underline{\beta_{2j}'}] \ .$$

Similar to (3.1.2) we finally write

$$\underline{y} = \underset{\sim}{X}^* \underline{\beta}^* + \underline{\varepsilon} \tag{3.1.5}$$

where $\underset{\substack{\sim \\ Npxpq_o}}{X^*} = \text{diag}\ [\underset{\sim}{X}_o^*, \underset{\sim}{X}_o^*, \ldots, \underset{\sim}{X}_o^*]$

$$\underline{\beta^*}' = [\underline{\beta_1^*}', \underline{\beta_2^*}', \ldots, \underline{\beta_p^*}'] \ .$$

We make the standard assumptions on a multivariate regression model, i.e.,

$$E(\underline{\varepsilon}) = \underline{0}, \quad var(\underline{\varepsilon}) = \underset{p\times p}{\underset{\sim}{\Sigma}} \otimes \underset{\sim}{I}_N, \quad \underset{\sim}{\Sigma} \text{ positive definite.} \tag{3.1.6}$$

For $\underset{\sim}{\Sigma} = (\sigma_{ij})$, the assumption on the covariance structure is equivalent to cov $(\underline{\varepsilon_i}, \underline{\varepsilon_j}) = \sigma_{ij}\, \underset{\sim}{I}_N$. We further assume that rank $(\underset{\sim}{X}_o^*) = q_o$, and that there are available sufficient observations to estimate all unknown parameters in $\underline{\beta}^*$ and $\underset{\sim}{\Sigma}$.

Since the errors may be correlated and heteroscedastic, we apply generalized least squares to obtain estimators of the parameter vectors as

$$\begin{aligned}
\underline{\hat{\beta}_1} &= [\underset{\sim}{X}'(\underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_N)^{-1}\, \underset{\sim}{X}]^{-1}\; \underset{\sim}{X}'(\underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_N)^{-1}\, \underline{y} \\
&= [\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{X}_1'\underset{\sim}{X}_1]^{-1}\; [\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{X}_1']\, \underline{y} \\
&= [\underset{\sim}{\Sigma} \otimes (\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-1}]\; [\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{X}_1']\, \underline{y} \\
&= [\underset{\sim}{I}_p \otimes (\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-1}\, \underset{\sim}{X}_1']\, \underline{y} \\
&= \begin{bmatrix} (\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-1}\, \underset{\sim}{X}_1'\, \underline{y_1} \\ (\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-1}\, \underset{\sim}{X}_1'\, \underline{y_2} \\ \vdots \\ (\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-1}\, \underset{\sim}{X}_1'\, \underline{y_p} \end{bmatrix}
\end{aligned} \tag{3.1.7}$$

Similarly,

$$\underline{\hat{\beta}^*} = [\underset{\sim}{X}^{*\prime}(\underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_N)^{-1}\underset{\sim}{X}^*]^{-1} \; \underset{\sim}{X}^{*\prime}(\underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_N)^{-1} \; \underline{y}$$

$$= \begin{bmatrix} (\underset{\sim}{X}_o^{*\prime}\underset{\sim}{X}_o^*)^{-1} \; \underset{\sim}{X}_o^{*\prime} \; \underline{y_1} \\ (\underset{\sim}{X}_o^{*\prime}\underset{\sim}{X}_o^*)^{-1} \; \underset{\sim}{X}_o^{*\prime} \; \underline{y_2} \\ \vdots \\ (\underset{\sim}{X}_o^{*\prime}\underset{\sim}{X}_o^*)^{-1} \; \underset{\sim}{X}_o^{*\prime} \; \underline{y_p} \end{bmatrix} . \tag{3.1.8}$$

Thus, for $\underline{\hat{\beta}_1'} = [\underline{\hat{\beta}_{11}'}, \underline{\hat{\beta}_{12}'}, \ldots, \underline{\hat{\beta}_{1p}'}]$ and $\underline{\hat{\beta}^{*\prime}} = [\underline{\hat{\beta}_1^{*\prime}}, \underline{\hat{\beta}_2^{*\prime}}, \ldots, \underline{\hat{\beta}_p^{*\prime}}]$, the multivariate estimators reduce to the standard univariate least squares estimators

$$\underline{\hat{\beta}_{1j}} = (\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-1} \; \underset{\sim}{X}_1' \; \underline{y_j} \quad , \; j = 1, 2, \ldots, p$$

$$\underline{\hat{\beta}_j^*} = (\underset{\sim}{X}_o^{*\prime}\underset{\sim}{X}_o^*)^{-1} \; \underset{\sim}{X}_o^{*\prime} \; \underline{y_j},$$

making use only of the N observations associated with a particular regression. We shall see that this is not the case under the generalized multivariate regression model in Chapter IV.

For a given j, we fit either the response function

$$\hat{y}_j = \underline{x_1'} \; \underline{\hat{\beta}_{1j}} \tag{3.1.9}$$

or

$$\overset{*}{y}_j = \underline{x_1'} \; \underline{\hat{\hat{\beta}}_{1j}} + \underline{x_2'} \; \underline{\hat{\beta}_{2j}}$$

$$= \underline{x_o^{*\prime}} \; \underline{\hat{\beta}_j^*} \tag{3.1.10}$$

where $\underline{x}_1'$, $\underline{x}_2'$, and $\underline{x}_o^{*\prime}$ represent typical row vectors in the matrices $\underset{\sim}{X}_1$, $\underset{\sim}{X}_2$, and $\underset{\sim}{X}_o^*$ respectively, $\underline{x}_o^{*\prime} = [\underline{x}_1', \underline{x}_2']$, and $\underline{\hat{\beta}}_j^{*\prime} = [\underline{\hat{\hat{\beta}}}_{1j}', \underline{\hat{\beta}}_{2j}']$. To illustrate, suppose we are dealing with two types of responses, each a function of two independent variables. For j = 1, 2, we consider

$$\hat{y}_j = \hat{\beta}_{1j(0)} + x_1\hat{\beta}_{1j(1)} + x_2\hat{\beta}_{1j(2)} ;$$

however we wish to afford ourselves a measure of protection against a situation where we should have fitted

$$\hat{y}_j^* = \hat{\hat{\beta}}_{1j(0)} + x_1\hat{\hat{\beta}}_{1j(1)} + x_2\hat{\hat{\beta}}_{1j(2)} + x_1^2\hat{\beta}_{2j(1)} + x_2^2\hat{\beta}_{2j(2)} + x_1x_2\hat{\beta}_{2j(12)} .$$

Here, p = 2, $q_1 = 3$, $q_2 = 3$, $q_0 = 6$, $\underline{x}_1' = [1, x_1, x_2]$, $\underline{x}_2' = [x_1^2, x_2^2, x_1x_2]$, $\underline{x}_o^{*\prime} = [1, x_1, x_2, x_1^2, x_2^2, x_1x_2]$, $\underline{\hat{\beta}}_{1j}' = [\hat{\beta}_{1j(0)}, \hat{\beta}_{1j(1)}, \hat{\beta}_{1j(2)}]$, $\underline{\hat{\beta}}_{2j}' = [\hat{\beta}_{2j(1)}, \hat{\beta}_{2j(2)}, \hat{\beta}_{2j(12)}]$, $\underline{\hat{\beta}}_j^{*\prime} = [\hat{\hat{\beta}}_{1j(0)}, \hat{\hat{\beta}}_{1j(1)}, \hat{\hat{\beta}}_{1j(2)}, \hat{\beta}_{2j(1)}, \hat{\beta}_{2j(2)}, \hat{\beta}_{2j(12)}]$.

3.2 A Test Procedure for the Integrated Mean Squared Error Criterion

As mentioned previously, the standard hypothesis on which to base preliminary test estimators for a univariate regression has been $\underline{\beta}_{2j} = \underline{0}$ given a particular j. Frequently however, an experimenter is interested not so much in what values are assumed by this parameter vector, as in how best to control certain properties of his estimator such as variance and bias. If one is comparing estimators according to some arbitrary criterion, then it seems reasonable to use this criterion in the development of the estimator itself. For this reason in the multivariate problem, rather than testing the hypothesis $\underline{\beta_2} = \underline{0}$ where $\underline{\beta_2'} = [\underline{\beta_{21}'}, \underline{\beta_{22}'}, \ldots, \underline{\beta_{2p}'}]$ and choosing a response function model as a result of whether or not this is rejected, we propose to construct a performance oriented preliminary test estimator around a more meaningful hypothesis.

Suppose we define vectors of the p estimated responses, $\hat{\underline{y}}' = [\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_p]$ and $\hat{\underline{y}}^{*\prime} = [\hat{y}_1^*, \hat{y}_2^*, \ldots, \hat{y}_p^*]$. It is often of interest to study appropriate linear combinations of these responses such as their sum. The criterion used will be that of mean squared error (MSE), averaged or integrated over the region of interest R of the independent variables $x_1, x_2, \ldots, x_k$. We shall test the hypothesis

$$H_0: J_1 \leq J_2 \quad \text{against} \quad H_1: J_1 > J_2 ,$$

where
$$J_1 = NK \int_R \text{MSE}\,(\underline{d}'\hat{\underline{y}})\, d\underline{x}$$

$$J_2 = NK \int_R \text{MSE}\,(\underline{d}'\hat{\underline{y}}^*)\, d\underline{x}$$

$$\underline{x}' = [x_1, x_2, \ldots, x_k]$$

$$K^{-1} = \int_R d\underline{x}$$

$\underline{d}'$ is a 1xp vector chosen by the experimenter to reflect weighting of the responses.

The integrated mean squared error criterion allows us to consider the performance of an estimator not just at a single point $x_1, x_2, \ldots, x_k$, but over the entire region R. It further enables us to examine both variance and bias of the estimators, similarly averaged over R. We denote these components as

$$V_1 = NK \int_R \mathrm{var}(\underline{d}'\hat{\underline{y}})\, d\underline{x}$$

$$B_1 = NK \int_R \mathrm{bias}^2(\underline{d}'\hat{\underline{y}})\, d\underline{x}$$

$$V_2 = NK \int_R \mathrm{var}(\underline{d}'\hat{\underline{y}}^*)\, d\underline{x}$$

where $\hat{\underline{y}}$ and $\hat{\underline{y}}^*$ are understood to be functions of $\underline{x}$. Since the mean squared error of an estimator is the sum of its variance and the square of its bias, it is immediate that $J_1 = V_1 + B_1$. It is also clear that there will be no integrated bias contribution to J_2 since $\hat{\underline{\beta}}^*$ is an unbiased estimator of $\underline{\beta}^*$ (Press (1972) page 199). We assume that the vector of true responses is best represented by

$$\begin{aligned} \underline{\eta'} &= [\eta_1, \eta_2, \ldots, \eta_p] \\ &= [\underline{x}_o^{*\prime}\underline{\beta}_1^*,\ \underline{x}_o^{*\prime}\underline{\beta}_2^*,\ \ldots,\ \underline{x}_o^{*\prime}\underline{\beta}_p^*] \,. \end{aligned} \tag{3.2.1}$$

In testing H_o: $J_1 \leq J_2$, we are essentially attempting to determine whether the bias component B_1 incurred by the addition of the terms in $\underset{\sim}{X}_2$ to the basic model (3.1.1), increases V_1 to the extent that $J_1 = V_1 + B_1$ is larger than $J_2 = V_2$, the variance arising from these same supplementary terms. The additional terms can only increase the variance as shown in the following:

<u>Lemma 3.2.1</u>: $V_1 < V_2$.

<u>Proof</u>: $\text{var}(\underline{d}'\hat{\underline{y}}) = \underline{d}'[\text{var}(\hat{\underline{y}})]\underline{d}$

$$= \underline{d}'[\underset{pxpq_1}{\text{diag}}\ (\underline{x_1'},\underline{x_1'},\ldots,\underline{x_1'})]\ \text{var}(\hat{\underline{\beta_1}})[\underset{pq_1xp}{\text{diag}}\ (\underline{x_1},\underline{x_1},\ldots,\underline{x_1})]\underline{d}$$

$$= \underline{d}'[\text{diag}(\underline{x_1'},\underline{x_1'},\ldots,\underline{x_1'})][\underset{\sim}{\Sigma} \otimes (\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-1}]$$

$$[\text{diag}(\underline{x_1},\underline{x_1},\ldots,\underline{x_1})]\underline{d}$$

(Press (1972) page 214)

$$= \underline{d}'[(\underline{x_1'}(\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-1}\underline{x_1})\ \underset{\sim}{\Sigma}]\underline{d}$$

$$= \underline{x_1'}(\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-1}\underline{x_1}\ (\underline{d}'\underset{\sim}{\Sigma}\underline{d})\ . \qquad (3.2.2)$$

Therefore, $V_1 = NK \int_R \text{var}(\underline{d}'\hat{\underline{y}})d\underline{x}$

$$= (\underline{d}'\underset{\sim}{\Sigma}\underline{d})\ NK \int_R \text{tr}(\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-1}\underline{x_1}\ \underline{x_1'}\ d\underline{x}$$

$$= (\underline{d}'\underset{\sim}{\Sigma}\underline{d})\ \text{tr}(\underset{\sim}{M}_{11}^{-1}\underset{\sim}{\mu}_{11}) \qquad (3.2.3)$$

where $$\underset{\sim}{M}_{ij} = N^{-1}(\underset{\sim}{X}_i'\underset{\sim}{X}_j) \qquad (3.2.4)$$

and $$\underset{\sim}{\mu}_{ij} = K \int_R \underline{x_i}\ \underline{x_j'}\ d\underline{x} \qquad (3.2.5)$$

are referred to as design and region moment matrices respectively.

Similarly,

$$\text{var}(\underline{d}'\hat{\underline{y}}^*) = \underline{x_o^*}'(\underset{\sim}{X}_o^{*\prime}\underset{\sim}{X}_o^*)^{-1}\underline{x_o^*}\ (\underline{d}'\underset{\sim}{\Sigma}\ \underline{d})\ . \qquad (3.2.6)$$

Now

$$(\underset{\sim}{X}_o^{*\prime}\underset{\sim}{X}_o^*)^{-1} = \begin{bmatrix} \underset{\sim}{X}_1'\underset{\sim}{X}_1 & \underset{\sim}{X}_1'\underset{\sim}{X}_2 \\ \underset{\sim}{X}_2'\underset{\sim}{X}_1 & \underset{\sim}{X}_2'\underset{\sim}{X}_2 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} (N\underset{\sim}{M}_{11})^{-1} + \underset{\sim}{A}\underset{\sim}{M}\underset{\sim}{A}' & -\underset{\sim}{A}\underset{\sim}{M} \\ -\underset{\sim}{M}\underset{\sim}{A}' & \underset{\sim}{M} \end{bmatrix} \tag{3.2.7}$$

where $\underset{\sim}{A} = \underset{\sim}{M}_{11}^{-1}\underset{\sim}{M}_{12}$ (3.2.8)

$$\underset{\sim}{M} = N^{-1}(\underset{\sim}{M}_{22} - \underset{\sim}{M}_{21}\underset{\sim}{M}_{11}^{-1}\underset{\sim}{M}_{12})^{-1} \tag{3.2.9}$$

(Press (1972)(2.6.4) and (2.6.5) and Graybill (1961) Theorem 1.49).

Thus,

$$\frac{\text{var}(\underline{d}'\hat{\underline{\overset{*}{y}}})}{\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}} = [\underline{x}_1', \underline{x}_2'] \begin{bmatrix} (N\underset{\sim}{M}_{11})^{-1} + \underset{\sim}{A}\underset{\sim}{M}\underset{\sim}{A}' & -\underset{\sim}{A}\underset{\sim}{M} \\ -\underset{\sim}{M}\underset{\sim}{A}' & \underset{\sim}{M} \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix}$$

$$= N^{-1}\underline{x}_1'\underset{\sim}{M}_{11}^{-1}\underline{x}_1 + \underline{x}_1'\,\underset{\sim}{A}\underset{\sim}{M}\underset{\sim}{A}'\,\underline{x}_1 - \underline{x}_2'\,\underset{\sim}{M}\underset{\sim}{A}'\underline{x}_1$$

$$- \underline{x}_1'\,\underset{\sim}{A}\underset{\sim}{M}\,\underline{x}_2 + \underline{x}_2'\,\underset{\sim}{M}\,\underline{x}_2$$

$$= N^{-1}\underline{x}_1'\,\underset{\sim}{M}_{11}^{-1}\underline{x}_1 + (\underline{x}_1'\,\underset{\sim}{A}-\underline{x}_2')\underset{\sim}{M}(\underset{\sim}{A}'\underline{x}_1-\underline{x}_2)\ . \tag{3.2.10}$$

Therefore, $V_2 = NK\int_R \text{var}(\underline{d}'\hat{\underline{\overset{*}{y}}})d\underline{x}$

$$= (\underline{d}'\underset{\sim}{\Sigma}\,\underline{d})NK\int_R [N^{-1}\text{tr}(\underset{\sim}{M}_{11}^{-1}\underline{x}_1\underline{x}_1') + (\underline{x}_1'\underset{\sim}{A}-\underline{x}_2')\underset{\sim}{M}(\underset{\sim}{A}'\underline{x}_1-\underline{x}_2)]d\underline{x}$$

$$= (\underline{d}'\underset{\sim}{\Sigma}\ \underline{d})[\mathrm{tr}(\underset{\sim}{M}_{11}^{-1}\underset{\sim}{\mu}_{11}) + NK \int_R (\underline{x}_1'\underset{\sim}{A}-\underline{x}_2')\ \underset{\sim}{M}(\underset{\sim}{A}'\underline{x}_1-\underline{x}_2)d\underline{x}] \ . \quad (3.2.11)$$

Comparing (3.2.11) and (3.2.3) gives

$$V_2 - V_1 = (\underline{d}'\underset{\sim}{\Sigma}\ \underline{d})\ NK \int_R (\underline{x}_1'\underset{\sim}{A}-\underline{x}_2')\ \underset{\sim}{M}(\underset{\sim}{A}'\underline{x}_1-\underline{x}_2)d\underline{x} \ . \quad (3.2.12)$$

Assuming that R is such that $(\underline{x}_1'\underset{\sim}{A}-\underline{x}_2') \neq \underline{0}$ for at least one $\underline{x} \in R$, it remains to show that $\underset{\sim}{M}$ is positive definite. But since $(\underset{\sim}{X}_o^{*}{}'\underset{\sim}{X}_o^{*})^{-1}$ is positive definite from Theorem 1.24 of Graybill (1961), it is immediate from (3.2.7) that $\underset{\sim}{M}$ is also positive definite using Theorem 1.23 of Graybill (1961). Hence, $V_2 - V_1 > 0$.

At this point we note that

$$\mathrm{var}(\hat{\underline{\beta}}^{*}) = \underset{\sim}{\Sigma} \otimes (\underset{\sim}{X}_o^{*}{}'\underset{\sim}{X}_o^{*})^{-1}$$

(Press (1972), page 214). This, along with (3.2.7), implies

$$\mathrm{var}(\underline{\hat{\beta}_2}) = \underset{\sim}{\Sigma} \otimes \underset{\sim}{M} \quad (3.2.13)$$

for $\underline{\hat{\hat{\beta}}_1'} = [\underline{\hat{\hat{\beta}}_{11}'}, \underline{\hat{\hat{\beta}}_{12}'}, \ldots, \underline{\hat{\hat{\beta}}_{1p}'}]$ and $\underline{\hat{\beta}_2'} = [\underline{\hat{\beta}_{21}'}, \underline{\hat{\beta}_{22}'}, \ldots, \underline{\hat{\beta}_{2p}'}]$.

For what follows, it will be convenient to rewrite H_o. Using (3.2.12),

$$\begin{aligned} V_2 - V_1 &= (\underline{d}'\underset{\sim}{\Sigma}\ \underline{d})\ NK \int_R \mathrm{tr}[\underset{\sim}{M}(\underset{\sim}{A}'\underline{x}_1-\underline{x}_2)(\underline{x}_1'\underset{\sim}{A}-\underline{x}_2')]\ d\underline{x} \\ &= (\underline{d}'\underset{\sim}{\Sigma}\ \underline{d})\ N\ \mathrm{tr}(\underset{\sim}{M}\ \underset{\sim}{M}_{212}) \end{aligned}$$

where $$\underset{\sim}{M}_{212} = K \int_R (\underset{\sim}{A}'\underline{x}_1-\underline{x}_2)(\underline{x}_1'\underset{\sim}{A}-\underline{x}_2')d\underline{x} \ . \quad (3.2.14)$$

Let $\quad a_1 = N\ tr(\underset{\sim}{M}\ \underset{\sim}{M}_{212})$ (3.2.15)

so that $\quad V_2 - V_1 = a_1(\underline{d}'\underset{\sim}{\Sigma}\ \underline{d})$. (3.2.16)

This enables us to express our hypothesis in the form

$$H_o: \ \frac{B_1}{\underline{d}'\underset{\sim}{\Sigma}\ \underline{d}} \leq a_1 \ . \tag{3.2.17}$$

In order to obtain B_1, we first require

$$E(\underline{\hat{\beta}_1}) = \underset{pq_1 xNp}{\text{diag}}\ [(\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1',\ (\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1',\ \ldots,\ (\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1']\ E(\underset{\sim}{X}^*\underline{\beta}^*+\underline{\varepsilon})$$

(using (3.1.7) and (3.1.5))

$$= \underset{pq_1 xpq_o}{\text{diag}}\ [(\underset{\sim}{I}_{q_1} : \underset{\sim}{A}),\ (\underset{\sim}{I}_{q_1} : \underset{\sim}{A}),\ \ldots,\ (\underset{\sim}{I}_{q_1} : \underset{\sim}{A})]\ \underline{\beta}^*$$

$$= \underline{\beta_1} + \underset{pq_1 xpq_2}{\text{diag}}\ [\underset{\sim}{A},\ \underset{\sim}{A},\ \ldots,\ \underset{\sim}{A}]\ \underline{\beta_2} \ . \tag{3.2.18}$$

Therefore,

$$E(\underline{d}'\underline{\hat{y}}) = \underline{d}'[\text{diag}(\underline{x_1'},\ \underline{x_1'},\ \ldots,\ \underline{x_1'})]\ E(\underline{\hat{\beta}_1})$$

$$= \underline{d}'[\text{diag}(\underline{x_1'},\ \underline{x_1'},\ \ldots,\ \underline{x_1'})][\underline{\beta_1} + \text{diag}(\underset{\sim}{A},\ \underset{\sim}{A},\ \ldots,\ \underset{\sim}{A})\underline{\beta_2}] \ . \tag{3.2.19}$$

From (3.2.1),

$$\underline{d}'\underline{\eta} = \underline{d}'[\text{diag}(\underline{x_o^{*}}',\underline{x_o^{*}}',\ldots,\underline{x_o^{*}}')]\ \underline{\beta}^*$$

$$= \underline{d}'[\text{diag}(\underline{x_1'},\underline{x_1'},\ldots,\underline{x_1'})\underline{\beta_1} + \underset{pxpq_2}{\text{diag}}(\underline{x_2'},\underline{x_2'},\ldots,\underline{x_2'})\underline{\beta_2}] \ . \tag{3.2.20}$$

As a result of (3.2.19) and (3.2.20), let

$$b_1 = \text{bias}(\underline{d}'\hat{\underline{y}})$$

$$= E(\underline{d}'\hat{\underline{y}}) - \underline{d}'\underline{\eta}$$

$$= \underline{d}'[\underset{pxpq_2}{\text{diag}}\ (\underline{x}_1'\underset{\sim}{A}-\underline{x}_2',\ \underline{x}_1'\underset{\sim}{A}-\underline{x}_2',\ \ldots,\ \underline{x}_1'\underset{\sim}{A}-\underline{x}_2')]\underline{\beta}_2\ . \qquad (3.2.21)$$

Then, $b_1{}^2 = \underline{\beta}_2'[\underline{d}\ \underline{d}' \otimes (\underset{\sim}{A}'\underline{x}_1-\underline{x}_2)(\underline{x}_1'\underset{\sim}{A}-\underline{x}_2')]\underline{\beta}_2$, and since b_1 is a function of $\underline{x}$, we can write

$$B_1 = NK \int_R b_1{}^2 d\underline{x}$$

$$= N\ \underline{\beta}_2'[\underline{d}\ \underline{d}' \otimes \underset{\sim}{M}_{212}]\ \underline{\beta}_2\ . \qquad (3.2.22)$$

This suggests as the numerator of a test statistic for (3.2.17), the quantity

$$\hat{B}_1 = N\ \hat{\underline{\beta}}_2'\ [\underline{d}\ \underline{d}' \otimes \underset{\sim}{M}_{212}]\ \hat{\underline{\beta}}_2\ . \qquad (3.2.23)$$

We obtain the denominator of our statistic by using the standard covariance estimator for a multivariate regression model, i.e.,

$$\hat{\underset{\sim}{\Sigma}} = \frac{1}{N-q_o}\ (\underset{\sim}{Y}-\underset{\sim}{X}_o^*\hat{\underset{\sim}{B}}^*)'\ (\underset{\sim}{Y}-\underset{\sim}{X}_o^*\hat{\underset{\sim}{B}}^*) \qquad (3.2.24)$$

where $\underset{\sim}{Y} = [\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_p]$

$$\hat{\underset{\sim}{B}}^* = [\hat{\underline{\beta}}_1^*, \hat{\underline{\beta}}_2^*, \ldots, \hat{\underline{\beta}}_p^*]\ .$$

Our test statistic is

$$F_o = \frac{\hat{B}_1}{\underline{d}'\hat{\underset{\sim}{\Sigma}}\ \underline{d}}\ . \qquad (3.2.25)$$

For the univariate case defined by $\underline{d}' = 1$, it is interesting to note that if $q_2 = 1$, then (3.2.25) is equivalent (except for a constant multiplier) to the usual statistic used in testing the hypothesis $\underline{\beta_{2j}} = \underline{0}$ for a given j. This statistic is

$$F_c = \frac{\underline{\hat{\beta}'_{2j}}\ \underset{\sim}{M}^{-1}\ \underline{\hat{\beta}_{2j}}/q_2}{\hat{\sigma}_{jj}} \tag{3.2.26}$$

where $\hat{\sigma}_{jj} = \frac{1}{N-q_o}(\underline{y_j} - \underset{\sim}{X}^*_o\underline{\hat{\beta}^*_j})'\ (\underline{y_j} - \underset{\sim}{X}^*_o\underline{\hat{\beta}^*_j})$.

In particular if $q_2 = 1$, then $\underline{\hat{\beta}_{2j}} = \hat{\beta}_{2j}$, $\underset{\sim}{M} = m$, and $\underset{\sim}{M}_{212} = m_{212}$ are scalars so that

$$F_c = \frac{1}{N\ m\ m_{212}}\left(\frac{Nm_{212}\hat{\beta}^2_{2j}}{\hat{\sigma}_{jj}}\right)$$

$$= F_o/a_1 \tag{3.2.27}$$

for $\underline{d}' = 1$. This relationship does not hold in general, however. We shall see in Chapter V that once distributional assumptions are made, the two procedures have different critical regions even for $q_2 = 1$, owing to the different hypotheses on which they are based.

We now obtain numerator and denominator expected values in F_o. Utilizing Press (1961) (3.2.11),

$$E(\hat{B}_1) = NE[\underline{\hat{\beta}'_2}\ (\underline{d}\ \underline{d}' \otimes \underset{\sim}{M}_{212})\underline{\hat{\beta}_2}]$$

$$= N\ \underline{\beta'_2}\ [\underline{d}\ \underline{d}' \otimes \underset{\sim}{M}_{212}]\underline{\beta_2} + N\ \mathrm{tr}[(\underset{\sim}{\Sigma} \otimes \underset{\sim}{M})(\underline{d}\ \underline{d}' \otimes \underset{\sim}{M}_{212})]$$

$$= B_1 + N[\mathrm{tr}(\underset{\sim}{\Sigma}\, \underline{d}\, \underline{d}')\ \mathrm{tr}(\underset{\sim}{M}\, \underset{\sim}{M}_{212})]$$

$$= B_1 + a_1(\underline{d}'\underset{\sim}{\Sigma}\, \underline{d})\ . \tag{3.2.28}$$

Since $\hat{\underset{\sim}{\Sigma}}$ is unbiased (Press (1961) page 212),

$$E(\underline{d}'\hat{\underset{\sim}{\Sigma}}\, \underline{d}) = \underline{d}'\underset{\sim}{\Sigma}\, \underline{d}\ . \tag{3.2.29}$$

The ratio of expected values in (3.2.25) yields

$$a_1 + B_1/\underline{d}'\underset{\sim}{\Sigma}\, \underline{d} = a_1 + a_3\ , \tag{3.2.30}$$

letting $a_3 = B_1/\underline{d}'\underset{\sim}{\Sigma}\, \underline{d}\ .$ (3.2.31)

The hypothesis $J_1 \leq J_2$ can now be written

$$H_0:\ a_3 \leq a_1\ . \tag{3.2.32}$$

If we are unwilling to make distributional assumptions on the errors, then a reasonable test procedure (and thus an estimation procedure) based on (3.2.30) is

reject H_0 if $F_0 > 2a_1$ and fit $\hat{\underline{y}}^*$

accept H_0 otherwise and fit $\hat{\underline{y}}$.

We remark that the standard statistic F_c is unsuitable for testing H_0 from the standpoint of the ratio of expected values since

$$\frac{E[\hat{\underline{\beta}}_{2j}'\underset{\sim}{M}^{-1}\hat{\underline{\beta}}_{2j}/q_2]}{E(\hat{\sigma}_{jj})} = \frac{\underline{\beta}_{2j}'\underset{\sim}{M}^{-1}\underline{\beta}_{2j}}{q_2\sigma_{jj}} + \frac{\mathrm{tr}(\underset{\sim}{M}\,\underset{\sim}{M}^{-1})}{q_2}$$

$$= \frac{\underline{\beta}'_{2j} \underset{\sim}{M}^{-1} \underline{\beta}_{2j}}{q_2 \sigma_{jj}} + 1 \ . \tag{3.2.33}$$

The explanation, of course, is that F_c is designed for testing hypotheses on the parameter vector, e.g., $\underline{\beta}_{2j} = \underline{0}$. In the multivariate problem, while it is true that $\underline{\beta}_2 = \underline{0}$ implies $B_1 = 0$ and thus that $J_1 < J_2$ by virtue of Lemma 3.2.1, the equivalence is only one-way, i.e., it may be the case that $J_1 < J_2$ although $\underline{\beta}_2 \neq \underline{0}$. Thus, we could find ourselves in the position of rejecting one of the two hypotheses while failing to reject the other.

By way of illustration, let us return to the example of section 3.1. Suppose N = 9, $\underline{d}' = [1,1]$, $\underline{y}_1' = [2,1,-1,3,-4,0,-2,4,-1]$, $\underline{y}_2' = [-3,2,3,0,1,-1,4,-2,3]$, and

$$\underset{\sim}{X}_o^* = \begin{array}{cccccc} & x_1 & x_2 & x_1^2-\overline{x_1^2} & x_2^2-\overline{x_2^2} & x_1x_2 \end{array}$$
$$\begin{bmatrix} 1 & -1 & -1 & 1/3 & 1/3 & 1 \\ 1 & -1 & 0 & 1/3 & -2/3 & 0 \\ 1 & -1 & 1 & 1/3 & 1/3 & -1 \\ 1 & 0 & -1 & -2/3 & 1/3 & 0 \\ 1 & 0 & 0 & -2/3 & -2/3 & 0 \\ 1 & 0 & 1 & -2/3 & 1/3 & 0 \\ 1 & 1 & -1 & 1/3 & 1/3 & -1 \\ 1 & 1 & 0 & 1/3 & -2/3 & 0 \\ 1 & 1 & 1 & 1/3 & 1/3 & 1 \end{bmatrix}$$

where $\overline{x_1^2} = 2/3$ and $\overline{x_2^2} = 2/3$ represent means of x_1^2 and x_2^2. The matrix $\underset{\sim}{X}_o^*$ corresponds to the slightly rewritten model

$$\eta_j = \beta'_{1j(0)} + x_1\beta_{1j(1)} + x_2\beta_{1j(2)} + (x_1^2-\overline{x_1^2})\beta_{2j(1)}$$
$$+ (x_2^2-\overline{x_2^2})\beta_{2j(2)} + x_1x_2\beta_{2j(12)}$$

where $\beta'_{1j(0)} = \beta_{1j(0)} + \overline{x_1^2}\beta_{2j(1)} + \overline{x_2^2}\beta_{2j(2)}$, $j = 1, 2$.

The revised model is used merely for computational ease in obtaining estimates of the parameters since $(\underset{\sim}{X}_o^{*\prime}\underset{\sim}{X}_o^*)^{-1} = \text{diag}[1/9,1/6,1/6,1/2,1/2,1/4]$. Also,

$$\underset{\sim}{X}_1' = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 \end{bmatrix}$$

$$\underset{\sim}{X}_2' = \begin{bmatrix} 1/3 & 1/3 & 1/3 & -2/3 & -2/3 & -2/3 & 1/3 & 1/3 & 1/3 \\ 1/2 & -2/3 & 1/3 & 1/3 & -2/3 & 1/3 & 1/3 & -2/3 & 1/3 \\ 1 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 1 \end{bmatrix}$$

so that $\underset{\sim}{X}_1'\underset{\sim}{X}_2 = \underset{\sim}{0}$ implies $\underset{\sim}{A} = \underset{\sim}{0}$. If the region of interest R is $-1 \leq x_1 \leq 1$, $-1 \leq x_2 \leq 1$, then

$$\underset{\sim}{M}_{212} = K \int_R \underline{x_2}\, \underline{x_2'}\, d\underline{x}$$
$$= K \int_{-1}^{1}\int_{-1}^{1} [x_1^2-2/3, x_2^2-2/3, x_1x_2]'[x_1^2-2/3, x_2^2-2/3, x_1x_2]$$
$$dx_1dx_2$$

$$= \begin{bmatrix} 1/5 & 1/9 & 0 \\ 1/9 & 1/5 & 0 \\ 0 & 0 & 1/9 \end{bmatrix}$$

where $K^{-1} = \int_{-1}^{1} \int_{-1}^{1} dx_1 dx_2 = 4.$

From (3.2.7), $\underset{\sim}{M} = \text{diag}[1/2, 1/2, 1/4]$ and

$$\underset{\sim}{M}\,\underset{\sim}{M}_{212} = \begin{bmatrix} 1/10 & 1/18 & 0 \\ 1/18 & 1/10 & 0 \\ 0 & 0 & 1/36 \end{bmatrix} .$$

Therefore, $a_1 = N \text{ tr } (\underset{\sim}{M}\,\underset{\sim}{M}_{212}) = 9(41/180)$
$= 2.05$.

From (3.2.7) and using the fact that $\underset{\sim}{X}_o^* = [\underset{\sim}{X}_1 : \underset{\sim}{X}_2]$ with $\underset{\sim}{X}_1' \underset{\sim}{X}_2 = \underset{\sim}{0}$, we obtain

$$\underline{\hat{\beta}_{21}'} = (\underset{\sim}{X}_2' \underset{\sim}{X}_2)^{-1} \underset{\sim}{X}_2' \, \underline{y_1} = [5/6, -1/6, 1]$$

$$\underline{\hat{\beta}_{22}'} = (\underset{\sim}{X}_2' \underset{\sim}{X}_2)^{-1} \underset{\sim}{X}_2' \, \underline{y_2} = [7/6, 2/3, -7/4] .$$

Employing (3.2.23), we can write

$$\hat{B}_1 = N(\underline{\hat{\beta}_{21}} + \underline{\hat{\beta}_{22}})' \, \underset{\sim}{M}_{212} \, (\underline{\hat{\beta}_{21}} + \underline{\hat{\beta}_{22}}) \qquad (3.2.34)$$

for $\underline{d}' = [1, 1]$. This yields

$$\hat{B}_1 = 9[2,1/2,-3/4]\begin{bmatrix}1/5 & 1/9 & 0\\ 1/9 & 1/5 & 0\\ 0 & 0 & 1/9\end{bmatrix}\begin{bmatrix}2\\ 1/2\\ -3/4\end{bmatrix}$$

$$= 10.2125\ .$$

Applying (3.2.24) gives

$$\hat{\underset{\sim}{\Sigma}} = \begin{bmatrix}13.93 & -7.48\\ -7.48 & 9.17\end{bmatrix}$$

so that $\underline{d}'\hat{\underset{\sim}{\Sigma}}\ \underline{d} = 8.14$. Therefore,

$$F_o = \frac{\hat{B}_1}{\underline{d}'\hat{\underset{\sim}{\Sigma}}\ \underline{d}} = 1.25 < 4.10 = 2a_1.$$

We are unable to reject H_o and as a result, we fit $\hat{\underline{y}}' = [\hat{y}_1, \hat{y}_2]$ where $\hat{y}_j = \underline{x}_1'\ \underline{\hat{\beta}_{1j}} = \hat{\beta}_{1j(0)} + x_1\hat{\beta}_{1j(1)} + x_2\hat{\beta}_{1j(2)}$, $j = 1, 2$.

3.3 An Approximation to the Distribution of F_o

Thus far, we have made no distributional assumptions, and consequently, have been unable to determine type I and type II error probabilities. In order to investigate the power function for a test procedure based on F_o, we now invoke error normality and assume

$$\underline{\varepsilon} \sim N(\underline{0}, \underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_N)\ . \tag{3.3.1}$$

The distribution of the denominator of F_o can be obtained with little difficulty and is, in fact, a special case of the multivariate Wishart distribution. From Press (1961) (8.4.13),

$$(N-q_o)\,\hat{\underset{\sim}{\Sigma}} \sim W(\underset{\sim}{\Sigma},\ p,\ N-q_o)\ .$$

Therefore, $(N-q_o)\underline{d}'\hat{\underset{\sim}{\Sigma}}\,\underline{d} \sim W(\underline{d}'\underset{\sim}{\Sigma}\,\underline{d},\ 1,\ N-q_o)$

using Press (1961) Theorem (5.1.6). The density function of $(N-q_o)\underline{d}'\hat{\underset{\sim}{\Sigma}}\,\underline{d}$ is

$$f(v) = \frac{1}{\Gamma(\frac{N-q_o}{2})(2\underline{d}'\underset{\sim}{\Sigma}\,\underline{d})^{\frac{N-q_o}{2}}}\, v^{\frac{N-q_o}{2}-1}\, e^{-v/2\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}}\ ,\ v > 0.$$

If we make the transformation $u = v/\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}$, then the density function of $(N-q_o)\underline{d}'\hat{\underset{\sim}{\Sigma}}\,\underline{d}/\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}$ is

$$f(u) = \frac{1}{\Gamma(\frac{N-q_o}{2})\ 2^{\frac{N-q_o}{2}}}\, u^{\frac{N-q_o}{2}-1}\, e^{-\frac{u}{2}}\ ,\ u > 0 \tag{3.3.2}$$

or

$$(N-q_o)\,\frac{\underline{d}'\hat{\underset{\sim}{\Sigma}}\,\underline{d}}{\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}} \sim \chi^2_{N-q_o}\ . \tag{3.3.3}$$

We now turn our attention to the distribution of $\hat{B}_1$. From Press (1961) page 214,

$$\underline{\hat{\beta}_2} \sim N(\underline{\beta_2},\ \underset{\sim}{\Sigma} \otimes \underset{\sim}{M})\ . \tag{3.3.4}$$

Equation (3.2.21) gives us

$$\hat{b}_1 = \underline{d}'[\mathrm{diag}(\underline{x}_1'\underset{\sim}{A}-\underline{x}_2',\ \underline{x}_1'\underset{\sim}{A}-\underline{x}_2',\ \ldots,\ \underline{x}_1'\underset{\sim}{A}-\underline{x}_2')]\ \underline{\hat{\beta}_2} \tag{3.3.5}$$

and $\hat{b}_1 \sim N(b_1,\ \mathrm{var}(\hat{b}_1))$ where

$$\mathrm{var}(\hat{b}_1) = \underline{d}'[\mathrm{diag}(\underline{x}_1'\underset{\sim}{A}-\underline{x}_2',\ \underline{x}_1'\underset{\sim}{A}-\underline{x}_2',\ \ldots,\ \underline{x}_1'\underset{\sim}{A}-\underline{x}_2')][\underset{\sim}{\Sigma} \otimes \underset{\sim}{M}]$$

$$[\mathrm{diag}(\underset{\sim}{A}'\underline{x}_1-\underline{x}_2,\ \underset{\sim}{A}'\underline{x}_1-\underline{x}_2, \ldots,\ \underset{\sim}{A}'\underline{x}_1-\underline{x}_2)]\ \underline{d}$$

$$= (\underline{x}_1'\underset{\sim}{A}-\underline{x}_2')\ \underset{\sim}{M}(\underset{\sim}{A}'\underline{x}_1-\underline{x}_2)\ \underline{d}'\underset{\sim}{\Sigma}\ \underline{d}$$

$$= b(\underline{x})\ \underline{d}'\underset{\sim}{\Sigma}\ \underline{d} \tag{3.3.6}$$

for $b(\underline{x}) = (\underline{x}_1'\underset{\sim}{A}-\underline{x}_2')\ \underset{\sim}{M}(\underset{\sim}{A}'\underline{x}_1-\underline{x}_2)$. (3.3.7)

Thus, $\hat{b}_1 \sim N(b_1,\ b(\underline{x})\underline{d}'\underset{\sim}{\Sigma}\ \underline{d})$ (3.3.8)

and $(\hat{b}_1)^2/\underline{d}'\underset{\sim}{\Sigma}\ \underline{d} \sim b(\underline{x})\ \chi'^{2}_{1,\lambda(\underline{x})}$

where $\lambda(\underline{x}) = b_1/[b(\underline{x})\underline{d}'\underset{\sim}{\Sigma}\ \underline{d}]^{1/2}$.

Since $\hat{B}_1 = NK \int_R (\hat{b}_1)^2\ d\underline{x}$, (3.3.9)

$$\hat{B}_1/\underline{d}'\underset{\sim}{\Sigma}\ \underline{d} \sim NK \int_R b(\underline{x})[w+\lambda(\underline{x})]^2\ d\underline{x} \tag{3.3.10}$$

where $w \sim N(0,1)$.

Expanding the right-hand side of (3.3.10) and applying (3.2.31) along with $a_1 = NK \int_R b(\underline{x})d\underline{x}$, ultimately yields

$$\hat{B}_1/\underline{d}'\underset{\sim}{\Sigma}\ \underline{d} \sim a_1 w^2 + 2a_2 w + a_3 \tag{3.3.11}$$

where $$a_2 = \frac{NK}{(\underline{d}'\underset{\sim}{\Sigma}\,\underline{d})^{1/2}} \int_R [b(\underline{x})]^{1/2}\, b_1 d\underline{x}\ . \tag{3.3.12}$$

The integration to be conducted in (3.3.12) does not lend itself to an explicit expression save for special cases to be discussed in section 3.5. The form of a_2, however, suggests a means by which we can approximate the distribution of $\hat{B}_1/\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}$. The integral version of the Cauchy-Schwarz inequality implies

$$\begin{aligned} a_2 &\leq [(NK \int_R b(\underline{x})d\underline{x})(NK \int_R (b_1{}^2/\underline{d}'\underset{\sim}{\Sigma}\,\underline{d})d\underline{x})]^{1/2} \\ &\leq (a_1 a_3)^{1/2}\ . \end{aligned}$$

Therefore, $$\frac{a_2}{a_1} \leq \left(\frac{a_3}{a_1}\right)^{1/2}\ . \tag{3.3.13}$$

Using this bound,

$$\begin{aligned} \frac{\hat{B}_1}{\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}} &\sim a_1[w^2 + 2(a_2/a_1)w + a_3/a_1] \\ &\approx a_1[w^2 + 2(a_3/a_1)^{1/2}w + a_3/a_1] \\ &\approx a_1[w + (a_3/a_1)^{1/2}]^2 && (3.3.14) \\ &\approx a_1\ \chi'^2_{1,(a_3/a_1)^{1/2}}\ . && (3.3.15) \end{aligned}$$

Under normality, the numerator and denominator of F_0 are independent quadratic forms (using Graybill (1961) Theorem 4.21), so that our statistic can be approximated by the ratio of independent chi-square variates, i.e.,

$$F_o = \frac{\hat{B}_1/\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}}{\underline{d}'\underset{\sim}{\hat{\Sigma}}\,\underline{d}/\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}}$$

$$\approx \frac{a_1\,\chi'^2_{1,(a_3/a_1)^{1/2}}}{\chi^2_{N-q_o}/(N-q_o)}$$

$$\approx a_1\,F'_{1,N-q_o,(a_3/a_1)^{1/2}} \,. \tag{3.3.16}$$

In order to obtain an explicit expression for the power function P of a test procedure structured around (3.3.16), we shall use (3.3.14). If D_α is a positive constant,

$$1 - P = Pr(F_o \leq D_\alpha)$$

$$\doteq Pr[a_1(w+(a_3/a_1)^{1/2})^2 \leq D_\alpha\, U/(N-q_o)] \tag{3.3.17}$$

where $$U = (N-q_o)\underline{d}'\underset{\sim}{\hat{\Sigma}}\,\underline{d}/\underline{d}'\underset{\sim}{\Sigma}\,\underline{d} \,. \tag{3.3.18}$$

Equation (3.3.17) can be written

$$1 - P \doteq Pr[-\left(\frac{D_\alpha U}{a_1(N-q_o)}\right)^{1/2} - \left(\frac{a_3}{a_1}\right)^{1/2} \leq w \leq \left(\frac{D_\alpha U}{a_1(N-q_o)}\right)^{1/2} - \left(\frac{a_3}{a_1}\right)^{1/2}]$$

$$\doteq \int_0^\infty [\frac{1}{\sqrt{2\pi}} \int_{-\left(\frac{D_\alpha u}{a_1(N-q_o)}\right)^{1/2} - \left(\frac{a_3}{a_1}\right)^{1/2}}^{\left(\frac{D_\alpha u}{a_1(N-q_o)}\right)^{1/2} - \left(\frac{a_3}{a_1}\right)^{1/2}} e^{-z^2/2}\,dz]$$

$$\frac{1}{2^{\frac{N-q_o}{2}}\,\Gamma\left(\frac{N-q_o}{2}\right)}\, u^{\frac{N-q_o}{2}-1}\, e^{-u/2}\,du \tag{3.3.19}$$

utilizing (3.3.2) and $w \sim N(0,1)$. This enables us to determine critical points D_α corresponding to designated α levels or probabilities of type I error by making the substitution $a_3/a_1 = 1$ under H_0. If the approximation of the distribution of F_0 by noncentral F with noncentrality parameter $(a_3/a_1)^{1/2}$ were exact, the resulting test procedure,

reject H_0 if $F_0 > D_\alpha$

accept H_0 otherwise

where D_α is such that $P = \alpha$, would constitute a uniformly most powerful test of H_0 (Lehmann (1959) page 68). This occurs in a special case to be outlined in section 3.6. Excepting this special case, the application of (3.3.19) will require numerical integration for determination of critical points and type II error probabilities. Alternatively, one may employ (3.3.16) in conjunction with existing approximations or tables of the noncentral F distribution.

Under H_0: $a_3/a_1 \leq 1$, the maximum difference between a_2/a_1 and $(a_3/a_1)^{1/2}$, incurred using (3.3.13), ensues when $a_2/a_1 = 0$, $(a_3/a_1)^{1/2} = 1$. This particular situation is, of course, impossible since $(a_3/a_1) \neq 0$ implies $b_1 \neq 0$ so that $a_2/a_1 \neq 0$. Nevertheless, we will use this as an indication of the most pessimistic comparison arising from the use of our approximation procedure versus the "true" distribution, recognizing that the actual disparity may be considerably less. The following table represents the differences between the nominal α levels obtained using (3.3.15) and the "true" α levels obtained using (3.3.11), i.e., nominal α - "true" α. As will be demonstrated in

Table 3.3.1 Effect of Bound Substitution (3.3.13)

Nominal α	α Difference
.01	.0086
.05	.0358
.10	.0596
.18	.0790
.25	.0749
.30	.0580
.35	.0278
.37	.0124
.38	.0012
.39	-.0054
.40	-.0238
.45	-.1052
.50	-.2490

Chapter V, our principal interest is in a suitable range of α values, not in a precise α per se. Thus, although the magnitudes of the true discrepancies will be smaller than those of Table 3.3.1, even the tabular differences shown are well within our tolerances.

3.4 Integrated Mean Squared Error of the Preliminary Test Estimator

Our preliminary test estimator is

$$\underline{\hat{y}_o} = \begin{cases} \hat{\hat{y}} & \text{if } H_0 \text{ is rejected} \\ \hat{\underline{y}} & \text{otherwise .} \end{cases} \tag{3.4.1}$$

For subsequent work, it will facilitate matters to be able to represent $\hat{\hat{y}}$ in terms of $\hat{\underline{y}}$. Using (3.2.7), consider

$$\begin{aligned} \underline{\hat{\beta}_j^*} &= (\underset{\sim}{X}_o^{*\prime}\underset{\sim}{X}_o^*)^{-1}\underset{\sim}{X}_o^{*\prime}\ \underline{y_j}\ , \quad j = 1, 2, \ldots, p \\ &= \begin{bmatrix} (N\underset{\sim}{M}_{11})^{-1}\underset{\sim}{X}_1'\ \underline{y_j} + \underset{\sim}{A}\underset{\sim}{M}\underset{\sim}{A}'\underset{\sim}{X}_1'\ \underline{y_j} - \underset{\sim}{A}\underset{\sim}{M}\underset{\sim}{X}_2'\ \underline{y_j} \\ -\underset{\sim}{M}\underset{\sim}{A}'\underset{\sim}{X}_1'\ \underline{y_j} + \underset{\sim}{M}\underset{\sim}{X}_2'\ \underline{y_j} \end{bmatrix} \\ &= \begin{bmatrix} (N\underset{\sim}{M}_{11})^{-1}\underset{\sim}{X}_1'\ \underline{y_j} - \underset{\sim}{A}(-\underset{\sim}{M}\underset{\sim}{A}'\underset{\sim}{X}_1'+\underset{\sim}{M}\underset{\sim}{X}_2')\ \underline{y_j} \\ -\underset{\sim}{M}\underset{\sim}{A}'\underset{\sim}{X}_1'+\underset{\sim}{M}\underset{\sim}{X}_2')\ \underline{y_j} \end{bmatrix} \\ &= \begin{bmatrix} \underline{\hat{\beta}_{1j}} - \underset{\sim}{A}\ \underline{\hat{\beta}_{2j}} \\ \underline{\hat{\beta}_{2j}} \end{bmatrix} . \end{aligned} \tag{3.4.2}$$

Thus, $\underline{\hat{\beta}_1} = [\underline{\hat{\beta}_1} - \text{diag}(\underset{\sim}{A},\underset{\sim}{A},\ldots,\underset{\sim}{A})\ \underline{\hat{\beta}_2}]$ and recalling (3.1.10),

$$\hat{\underline{y}}^* = \text{diag}[\underline{x}_1',\underline{x}_1',\ldots,\underline{x}_1']\ \underline{\hat{\beta}_1} + \text{diag}[\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\ \underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\ \ldots,\ \underline{x}_2'-\underline{x}_1'\underset{\sim}{A}]\ \underline{\hat{\beta}_2}\ .$$

We write

$$\underline{d}'\underline{\hat{y}_o} = \underline{d}'[\text{diag}(\underline{x}_1',\underline{x}_1',\ldots,\underline{x}_1')]\underline{\hat{\beta}_1} + \delta\underline{d}'[\text{diag}(\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\ldots,\underline{x}_2'-\underline{x}_1'\underset{\sim}{A})]\underline{\hat{\beta}_2} \tag{3.4.3}$$

$$\text{where } \delta = \begin{cases} 1 \text{ if } H_o \text{ is rejected} \\ 0 \text{ otherwise .} \end{cases}$$

Since the estimation criterion being studied is that of integrated mean squared error, it is only natural to investigate

$$\begin{aligned} J_o &= NK \int_R \text{MSE}(\underline{d}'\underline{\hat{y}_o})\ d\underline{x} \\ &= NK \int_R E(\underline{d}'\underline{\hat{y}_o}-\underline{d}'\underline{\eta})^2\ d\underline{x}\ . \end{aligned} \tag{3.4.4}$$

Comparing (3.4.3) and (3.2.20) gives

$$\begin{aligned} E(\underline{d}'\underline{\hat{y}_o}-\underline{d}'\underline{\eta})^2 = E\{&\underline{d}'[\text{diag}(\underline{x}_1',\underline{x}_1',\ldots,\underline{x}_1')(\underline{\hat{\beta}_1}-E(\underline{\hat{\beta}_1})) \\ &+ \delta\ \text{diag}(\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\ldots,\underline{x}_2'-\underline{x}_1'\underset{\sim}{A})(\underline{\hat{\beta}_2}-\underline{\beta_2}) \\ &+ (\delta-1)\ \text{diag}(\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\ldots,\underline{x}_2'-\underline{x}_1'\underset{\sim}{A})\underline{\beta_2}]\}^2 \end{aligned}$$

from (3.2.18). Continuing,

$$\begin{aligned} E(\underline{d}'\underline{\hat{y}_o}-\underline{d}'\underline{\eta})^2 &= E\{\underline{d}'[\text{diag}(\underline{x}_1',\underline{x}_1',\ldots,\underline{x}_1')](\underline{\hat{\beta}_1}-E(\underline{\hat{\beta}_1}))\}^2 \\ &+ E\{\delta\underline{d}'[\text{diag}(\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\ldots,\underline{x}_2'-\underline{x}_1'\underset{\sim}{A})](\underline{\hat{\beta}_2}-\underline{\beta_2})\}^2 \\ &+ E\{(\delta-1)\underline{d}'[\text{diag}(\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\ldots,\underline{x}_2'-\underline{x}_1'\underset{\sim}{A})]\underline{\beta_2}\}^2 \end{aligned}$$

$$+\ 2E\{\delta\underline{d}'[\mathrm{diag}(\underline{x}_1',\underline{x}_1',\ldots,\underline{x}_1')](\underline{\hat{\beta}_1}-E(\underline{\hat{\beta}_1}))(\underline{\hat{\beta}_2}-\underline{\beta_2})'$$

$$[\mathrm{diag}(\underline{x_2}-\underset{\sim}{A}'\underline{x_1},\underline{x_2}-\underset{\sim}{A}'\underline{x_1},\ldots,\underline{x_2}-\underset{\sim}{A}'\underline{x_1})]\underline{d}\}$$

$$+\ 2E\{(\delta-1)\underline{d}'[\mathrm{diag}(\underline{x}_1',\underline{x}_1',\ldots,\underline{x}_1')](\underline{\hat{\beta}_1}-E(\underline{\hat{\beta}_1}))$$

$$\underline{\beta}_2'[\mathrm{diag}(\underline{x_2}-\underset{\sim}{A}'\underline{x_1},\underline{x_2}-\underset{\sim}{A}'\underline{x_1},\ldots,\underline{x_2}-\underset{\sim}{A}'\underline{x_1})]\underline{d}\}$$

employing $\delta(\delta-1) = 0$. Simplifying gives

$$\begin{aligned} E(\underline{d}'\underline{\hat{y}_0}-\underline{d}'\underline{n})^2 &= \underline{d}'[\mathrm{diag}(\underline{x}_1',\underline{x}_1',\ldots,\underline{x}_1')]\mathrm{var}(\underline{\hat{\beta}_1})[\mathrm{diag}(\underline{x_1},\underline{x_1},\ldots,\underline{x_1})]\underline{d} \\ &+ E\{\delta\underline{d}'[\mathrm{diag}(\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\ldots,\underline{x}_2'-\underline{x}_1'\underset{\sim}{A})](\underline{\hat{\beta}_2}-\underline{\beta_2})\}^2 \\ &+ (1-E(\delta))\{\underline{d}'[\mathrm{diag}(\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\ldots,\underline{x}_2'-\underline{x}_1'\underset{\sim}{A})]\underline{\beta_2}\}^2 \\ &+ 2E\{\underline{d}'[\mathrm{diag}(\underline{x}_1',\underline{x}_1',\ldots,\underline{x}_1')](\underline{\hat{\beta}_1}-E(\underline{\hat{\beta}_1})) \\ &(\delta\underline{\hat{\beta}_2})'[\mathrm{diag}(\underline{x_2}-\underset{\sim}{A}'\underline{x_1},\underline{x_2}-\underset{\sim}{A}'\underline{x_1},\ldots,\underline{x_2}-\underset{\sim}{A}'\underline{x_1})]\underline{d}\} \qquad (3.4.5) \end{aligned}$$

where $(\delta-1)^2 = (1-\delta)$. Under normality assumptions, the fourth term in (3.4.5) will vanish by virtue of the following:

<u>Lemma 3.4.1</u>: If $\underline{\varepsilon} \sim N(\underline{0}, \underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_N)$, then $E[(\underline{\hat{\beta}_1}-E(\underline{\hat{\beta}_1}))(\delta\underline{\hat{\beta}_2})'] = \underset{\sim}{0}$.

<u>Proof</u>: For i, j = 1, 2, ..., p,

$$\mathrm{cov}(\underline{\hat{\beta}_{1i}},\underline{\hat{\beta}_{2j}}) = (\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-1}\underset{\sim}{X}_1'\ \mathrm{cov}(\underline{y_i},\underline{y_j})(-\underset{\sim}{X}_1\underset{\sim}{A}\underset{\sim}{M}+\underset{\sim}{X}_2\underset{\sim}{M})$$

from the proof of (3.4.2). Thus,

$$\mathrm{cov}(\underline{\hat{\beta}_{1i}},\underline{\hat{\beta}_{2j}}) = \sigma_{ij}(-\underset{\sim}{A}\underset{\sim}{M}+\underset{\sim}{A}\underset{\sim}{M}) = 0.$$

Since $\hat{\underline{\beta}}_1$ and $\hat{\underline{\beta}}_2$ are distributed normally, the two are independent so that

$$E[(\hat{\underline{\beta}}_1 - E(\hat{\underline{\beta}}_1))(\delta\hat{\underline{\beta}}_2)'] = E(\hat{\underline{\beta}}_1 - E(\hat{\underline{\beta}}_1))E(\delta\hat{\underline{\beta}}_2)' = \underset{\sim}{0}.$$

Utilizing (3.4.4), (3.2.2), (3.2.3), (3.2.21), (3.2.22), and Lemma 3.4.1 enables us to write (3.4.5) as

$$\begin{aligned} J_o &= V_1 + (1-P)B_1 \\ &\quad + NK \int_R E\{\delta\underline{d}'[\mathrm{diag}(\underline{x}_2'-\underline{x}_1'\underset{\sim}{A}, \underline{x}_2'-\underline{x}_1'\underset{\sim}{A}, \ldots, \underline{x}_2'-\underline{x}_1'\underset{\sim}{A})](\hat{\underline{\beta}}_2-\underline{\beta}_2)\}^2 \qquad (3.4.6) \end{aligned}$$

for $E(\delta) = P$.

The evaluation of the first two terms of (3.4.6) is straightforward for specified parameter values. The major problem is in the determination of

$$\begin{aligned} J_{03} &= NK \int_R E\{\delta\underline{d}'[\mathrm{diag}(\underline{x}_2'-\underline{x}_1'\underset{\sim}{A}, \underline{x}_2'-\underline{x}_1'\underset{\sim}{A}, \ldots, \underline{x}_2'-\underline{x}_1'\underset{\sim}{A})](\hat{\underline{\beta}}_2-\underline{\beta}_2)\}^2 \\ &= E\{\delta[NK \int_R (\hat{b}_1 - b_1)^2 d\underline{x}]\} \\ &= a_1 \underline{d}' \underset{\sim}{\Sigma}\, \underline{d}\, E(\delta Y) \qquad (3.4.7) \end{aligned}$$

where $Y = NK \int_R (\hat{b}_1 - b_1)^2 d\underline{x} / a_1 \underline{d}' \underset{\sim}{\Sigma}\, \underline{d}$. (3.4.8)

Conditions under which J_o can be evaluated exactly will be discussed in section 3.5. For the present, we shall confine ourselves to an estimation procedure for $E(\delta Y)$. Our preliminary test critical region

$$\hat{B}_1 / \underline{d}' \hat{\underset{\sim}{\Sigma}}\, \underline{d} > D_\alpha$$

is equivalent to (using (3.3.9))

$$\frac{NK \int_R (\hat{b}_1 - b_1)^2 d\underline{x}}{a_1 \underline{d}' \underset{\sim}{\Sigma}\, \underline{d}} > \frac{D_\alpha \underline{d}' \hat{\underset{\sim}{\Sigma}}\, \underline{d} - 2NK \int_R b_1 \hat{b}_1 d\underline{x} + B_1}{a_1 \underline{d}' \underset{\sim}{\Sigma}\, \underline{d}}$$

or $\quad Y > \hat{y}_\alpha$

$$\text{where} \quad \hat{y}_\alpha = \frac{D_\alpha U}{a_1 (N - q_o)} - \frac{2NK \int_R b_1 \hat{b}_1 d\underline{x}}{a_1 \underline{d}' \underset{\sim}{\Sigma}\, \underline{d}} + \frac{a_3}{a_1} . \tag{3.4.9}$$

Therefore, the random variable δY has a truncated distribution, suggesting as an estimator

$$\hat{J}_{03} = a_1 \underline{d}' \underset{\sim}{\Sigma}\, \underline{d} \int_{\hat{y}_\alpha}^{\infty} t\ f(t) dt \tag{3.4.10}$$

where $f(t)$ is the density function of Y. We remark that one could integrate with respect to the random variable U in $\hat{y}_\alpha$ and write (3.4.10) as a double integral; however, since b_1 in (3.4.9) must be estimated by $\hat{b}_1$ using the observation vector $\underline{y}$, it seems reasonable to use these same observations to estimate $\underset{\sim}{\Sigma}$. In order to determine $f(t)$, we proceed as in section 3.3. From (3.3.8),

$$\frac{\hat{b}_1 - b_1}{[b(\underline{x}) \underline{d}' \underset{\sim}{\Sigma}\, \underline{d}]^{1/2}} \sim N(0,1)$$

$$(\hat{b}_1 - b_1)^2 \sim b(\underline{x}) \underline{d}' \underset{\sim}{\Sigma}\, \underline{d}\ \chi_1^2$$

$$NK \int_R (\hat{b}_1 - b_1)^2 d\underline{x} \sim a_1 \underline{d}' \underset{\sim}{\Sigma}\, \underline{d}\ \chi_1^2$$

$$Y \sim \chi_1^2 \tag{3.4.11}$$

Thus, $\hat{J}_{03} = a_1\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}[1 - \frac{1}{\sqrt{2\pi}} \int_0^{\hat{y}_\alpha} t^{1/2}e^{-t/2}dt]$

$$= a_1\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}[1 - \frac{2}{\sqrt{2\pi}} \int_0^{\sqrt{\hat{y}_\alpha}} t_1^2 e^{-t_1^2/2} dt_1]$$

making the transformation $t_1 = t^{1/2}$.

Integrating by parts, i.e.,

$$\int r\, d\ell = r\ell - \int \ell\, dr \qquad (3.4.12)$$

where we equate $r = t_1$ and $\ell = -e^{-t_1^2/2}$, results in

$$\hat{J}_{03} = a_1\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}[(2\hat{y}_\alpha/\pi)^{1/2}\, e^{-\hat{y}_\alpha/2} + 2\Phi(-\sqrt{\hat{y}_\alpha})] \qquad (3.4.13)$$

and $$\hat{J}_0 = V_1 + (1-P)B_1 + \hat{J}_{03}\,. \qquad (3.4.14)$$

Of course the difficulty in this procedure is that in general we are unable to evaluate $E(\delta Y)$ exactly, owing to the fact that $\hat{y}_\alpha$ is not a true constant but a random variable. We shall now discuss a special case for which exact expressions for J_0 can be obtained.

3.5 The Single Independent Variable

If $q_2 = 1$, much of the preliminary test estimation problem is simplified. Without loss of generality, we shall restrict consideration to perhaps the most common example of this, the situation in which each of our p estimated responses is a function of a single independent variable x. As in the example of section 3.2, let

$$\begin{array}{c} \\ \underset{\sim}{X}_o^* = \end{array} \begin{array}{c} \begin{array}{ccc} \quad & x & x^2-\overline{x^2} \end{array} \\ \begin{bmatrix} 1 & x_{11} & x_{11}{}^2-\overline{x^2} \\ 1 & x_{12} & x_{12}{}^2-\overline{x^2} \\ \vdots & \vdots & \vdots \\ 1 & x_{1N} & x_{1N}{}^2-\overline{x^2} \end{bmatrix} \end{array}$$

and assume $\sum_{j=1}^{N} x_{1j} = 0$ so that

$$\underset{\sim}{X}_o^{*\prime}\underset{\sim}{X}_o^* = N \begin{bmatrix} 1 & 0 & 0 \\ 0 & \overline{x^2} & \overline{x^3} \\ 0 & \overline{x^3} & \overline{x^{4\prime}} \end{bmatrix}$$

where $\overline{x^2} = \sum_{j=1}^{N} x_{1j}{}^2/N$, $\overline{x^3} = \sum_{j=1}^{N} x_{1j}{}^3/N$, $\overline{x^{4\prime}} = \sum_{j=1}^{N} (x_{1j}{}^2-\overline{x^2})^2/N$.

Hence, $(\underset{\sim}{X}_o^{*\prime}\underset{\sim}{X}_o^*)^{-1} = \frac{1}{N} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \overline{x^{4\prime}}/D & -\overline{x^3}/D \\ 0 & -\overline{x^3}/D & \overline{x^2}/D \end{bmatrix}$

and $\underset{\sim}{M} = \overline{x^2}/ND$

where $D = (\overline{x^2})(\overline{x^{4\prime}}) - (\overline{x^3})^2$. (3.5.1)

Also, $\underset{\sim}{A} = \begin{bmatrix} N & 0 \\ 0 & N\overline{x^2} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ N\overline{x^3} \end{bmatrix} = \begin{bmatrix} 0 \\ \overline{x^3}/\overline{x^2} \end{bmatrix}$

$$\underline{x_1'}\underset{\sim}{A} - \underline{x_2'} = \overline{x^2} + (\overline{x^3}/\overline{x^2})x - x^2$$

$$\underset{\sim}{M}_{212} = K \int_{-1}^{1} [\overline{x^2} + (\overline{x^3}/\overline{x^2})x - x^2]^2 \, dx$$

$$= [(\overline{x^2}-1/3)^2 + 4/45 + (\overline{x^3})^2/3(\overline{x^2})^2]$$

for $K = 1/2$ and R such that $-1 \leq x \leq 1$.

Thus, $$a_1 = N \text{ tr } (\underset{\sim}{M}\, \underset{\sim}{M}_{212})$$

$$= \overline{x^2}\, [(\overline{x^2}-1/3)^2 + 4/45 + (\overline{x^3})^2/3(\overline{x^2})^2]/D \, . \tag{3.5.2}$$

Similar to (3.2.34), we can write

$$a_3 = N(\underline{d}'\underline{\beta_2})' \underset{\sim}{M}_{212}(\underline{d}'\underline{\beta_2})/\underline{d}'\underset{\sim}{\Sigma}\, \underline{d}$$

$$= N(\underline{d}'\underline{\beta_2})^2[(\overline{x^2}-1/3)^2 + 4/45 + (\overline{x^3})^2/3(\overline{x^2})^2]/\underline{d}'\underset{\sim}{\Sigma}\, \underline{d}$$

giving $$a_3/a_1 = ND(\underline{d}'\underline{\beta_2})^2/\overline{x^2}(\underline{d}'\underset{\sim}{\Sigma}\, \underline{d}). \tag{3.5.3}$$

Critical points D_α and probability of type II error are then obtained by employing (3.5.2) and (3.5.3) in (3.3.19) with $q_o = 3$. From (3.3.12),

$$a_2 = \frac{(N\overline{x^2})^{1/2}(\underline{d}'\underline{\beta_2})}{[D(\underline{d}'\underset{\sim}{\Sigma}\, \underline{d})]^{1/2}} \, K \int_{-1}^{1} [\overline{x^2} + (\overline{x^3}/\overline{x^2})x - x^2]^2 dx$$

$$= \frac{(N\overline{x^2})^{1/2}(\underline{d}'\underline{\beta_2})}{[D(\underline{d}'\underset{\sim}{\Sigma}\, \underline{d})]^{1/2}} \, [(\overline{x^2}-1/3)^2 + 4/45 + (\overline{x^3})^2/3(\overline{x^2})^2]$$

implies $$a_2/a_1 = (ND)^{1/2}(\underline{d}'\underline{\beta_2})/[\overline{x^2}(\underline{d}'\underset{\sim}{\Sigma}\, \underline{d})]^{1/2}$$

$$= (a_3/a_1)^{1/2} \, . \tag{3.5.4}$$

Thus, the distributional results obtained in section 3.3 are exact for the single independent variable case, and we do not have to make use of the bound substitution in (3.3.13).

To determine the integrated mean squared error of our preliminary test estimator for the case of the single independent variable, we shall examine its two components separately, i.e.,

$$J_o = NK \int_R MSE(\underline{d}'\underline{\hat{y}_o})d\underline{x}$$

$$= V_o + B_o$$

where $$V_o = NK \int_R var(\underline{d}'\underline{\hat{y}_o})d\underline{x} \tag{3.5.5}$$

$$B_o = NK \int_R bias^2(\underline{d}'\underline{\hat{y}_o})d\underline{x} \; . \tag{3.5.6}$$

From (3.4.3), (3.2.20), and (3.2.18),

$$bias(\underline{d}'\underline{\hat{y}_o}) = E(\underline{d}'\underline{\hat{y}_o}) - \underline{d}'\underline{n}$$

$$= \underline{d}'[diag(\underline{x_2'}-\underline{x_1'}\tilde{A},\underline{x_2'}-\underline{x_1'}\tilde{A},\ldots,\underline{x_2'}-\underline{x_1'}\tilde{A})][E(\delta\underline{\hat{\beta}_2})-\underline{\beta_2}]$$

$$= [x^2-(\overline{x^3}/\overline{x^2})x-\overline{x^2}]\{E[\delta(\underline{d}'\underline{\hat{\beta}_2})]-\underline{d}'\underline{\beta_2}\} \tag{3.5.7}$$

for the single independent variable case. Expanding (3.5.7) yields

$$bias(\underline{d}'\underline{\hat{y}_o}) = [x^2-(\overline{x^3}/\overline{x^2})x-\overline{x^2}]\{(var(\underline{d}'\underline{\hat{\beta}_2}))^{1/2}\; E[\delta(\frac{\underline{d}'\underline{\hat{\beta}_2}-\underline{d}'\underline{\beta_2}}{(var(\underline{d}'\underline{\hat{\beta}_2}))^{1/2}})]$$

$$+ E[\delta(\underline{d}'\underline{\beta_2})] - \underline{d}'\underline{\beta_2}\}$$

$$= [x^2-(\overline{x^3}/\overline{x^2})x-\overline{x^2}][(var(\underline{d}'\underline{\hat{\beta}_2}))^{1/2}E(\delta w)+(P-1)\underline{d}'\underline{\beta_2}] \tag{3.5.8}$$

where $w \sim N(0,1)$, $E(\delta) = P$, and

$$\mathrm{var}(\underline{d}'\hat{\underline{\beta}}_2) = \overline{x^2}\,(\underline{d}'\underset{\sim}{\Sigma}\,\underline{d})/ND\ . \tag{3.5.9}$$

$$\text{Therefore, } B_o = N[(\overline{x^2}(\underline{d}'\underset{\sim}{\Sigma}\,\underline{d})/ND)^{1/2}E(\delta w)+(P-1)\underline{d}'\underline{\beta}_2]^2$$
$$K\int_{-1}^{1}[x^2-(\overline{x^3}/\overline{x^2})x-\overline{x^2}]^2dx$$
$$= N[(\overline{x^2}-1/3)^2+4/45+(\overline{x^3})^2/3(\overline{x^2})^2][(\overline{x^2}(\underline{d}'\underset{\sim}{\Sigma}\,\underline{d})/ND)^{1/2}$$
$$E(\delta w)+(P-1)\underline{d}'\underline{\beta}_2]^2\ . \tag{3.5.10}$$

In order to evaluate $E(\delta w)$, we now reformulate our preliminary test critical region for the single independent variable case. The inequality $F_o > D_\alpha$ reduces to

$$N(\underline{d}'\hat{\underline{\beta}}_2)^2\,K\int_{-1}^{1}[x^2-(\overline{x^3}/\overline{x^2})x-\overline{x^2}]^2dx > D_\alpha\underline{d}'\hat{\underset{\sim}{\Sigma}}\,\underline{d}$$

and $\delta = 1$ only if

$$\begin{cases} \underline{d}'\hat{\underline{\beta}}_2 > \{D_\alpha\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}U/N(N-3)[(\overline{x^2}-1/3)^2+4/45+(\overline{x^3})^2/3(\overline{x^2})^2]\}^{1/2} \\ \text{or} \\ \underline{d}'\hat{\underline{\beta}}_2 < -\{D_\alpha\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}U/N(N-3)[(\overline{x^2}-1/3)^2+4/45+(\overline{x^3})^2/3(\overline{x^2})^2]\}^{1/2} \end{cases}$$

where $\quad U = (N-3)\dfrac{\underline{d}'\hat{\underset{\sim}{\Sigma}}\,\underline{d}}{\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}} \sim \chi^2_{N-3}\ .$

Normalizing $\underline{d}'\hat{\underline{\beta}}_2$ gives the equivalent condition

$$\begin{cases} w > R_H|U \\ \text{or} \\ w < R_L|U \end{cases} \qquad (3.5.11)$$

$$\text{for } R_H|U = \{DD_\alpha U/\overline{x^2}(N-3)[(\overline{x^2}-1/3)^2+4/45+(\overline{x^3})^2/3(\overline{x^2})^2]\}^{1/2}$$

$$- \underline{d}'\underline{\beta_2}/[\overline{x^2}(\underline{d}'\underset{\sim}{\Sigma}\ \underline{d})/ND]^{1/2}$$

$$= [D_\alpha U/a_1(N-3)]^{1/2} - (a_3/a_1)^{1/2} \qquad (3.5.12)$$

$$R_L|U = -[D_\alpha U/a_1(N-3)]^{1/2} - (a_3/a_1)^{1/2} \qquad (3.5.13)$$

where we agree to restrict $\underline{d}'\underline{\beta_2}/[\overline{x^2}(\underline{d}'\underset{\sim}{\Sigma}\ \underline{d})/ND]^{1/2}$ to the positive root $(a_3/a_1)^{1/2}$ due to the following:

Lemma 3.5.1: If $\underline{d}'\underline{\beta_2} \neq 0$, then the multiplication of $\underline{d}'\underline{\beta_2}$ by (-1) results in the multiplication of $E(\delta w)$ by (-1).

Proof: The random variable δw has a truncated distribution, allowing us to write for $\underline{d}'\underline{\beta_2} > 0$,

$$E(\delta w) = \int_0^\infty [\int_{-\infty}^{R_L|U} zf(z)dz + \int_{R_H|U}^{\infty} zf(z)dz]\ f(u)du$$

$$= \int_0^\infty [\int_{-R_H|U}^{-R_L|U} zf(z)dz]\ f(u)du \qquad (3.5.14)$$

where $f(u)$ is given by (3.3.2), and $f(z)$ being the $N(0,1)$ density function implies $zf(z)$ is an odd function. If $\underline{d}'\underline{\beta_2} < 0$, then using (3.5.14),

$$E(\delta w) = \int_0^\infty \Big[\int_{-[D_\alpha u/a_1(N-3)]^{1/2} - (a_3/a_1)^{1/2}}^{[D_\alpha u/a_1(N-3)]^{1/2} - (a_3/a_1)^{1/2}} zf(z)dz \Big] f(u)du$$

$$= -\int_0^\infty \Big[\int_{-R_H|U}^{-R_L|U} zf(z)dz \Big] f(u)du \;.$$

Clearly then, the sign of $\underline{d}'\beta_{\underline{2}}$ will have no effect on (3.5.10), and will similarly have no effect on the variance component of J_o as will be seen when we turn our attention to V_o. In Appendix I, it is shown that (3.5.14) can ultimately be expressed as

$$E(\delta w) = \frac{e^{-h^2/2(g+1)}/\sqrt{2\pi}}{2^{(\nu-2)/2}\Gamma(\nu/2)\sqrt{g+1}} \Bigg[\sum_{\theta=0}^{\nu-1} \binom{\nu-1}{\theta} \frac{(h\sqrt{g})^\theta}{(\sqrt{g+1})^{\nu+\theta-1}} \int_{-h\sqrt{\frac{g}{g+1}}}^{h\sqrt{\frac{g}{g+1}}} u_2^{\nu-1-\theta} e^{-u_2^2/2} du_2 + 2 \sum_{\phi=0}^{[\frac{\nu-2}{2}]} \binom{\nu-1}{2\phi+1} \frac{(h\sqrt{g})^{2\phi+1}}{(\sqrt{g+1})^{\nu+2\phi}} \int_{h\sqrt{\frac{g}{g+1}}}^{\infty} u_3^{\nu-2-2\phi} e^{-u_3^2/2} du_3 \Bigg] \;. \tag{3.5.15}$$

with g and h given by (A.2), $\nu = N - 3$, and $[\frac{\nu-2}{2}]$ denoting the largest integer less than or equal to $(\nu-2)/2$. The integrations in (3.5.15) can be evaluated by successive use of (3.4.12), identifying $u_i e^{-u_i^2/2}$ $(i = 2,3)$ with $d\ell$. Determination of $E(\delta w)$ is further facilitated by utilizing the fact that $u_2^{\nu-1-\theta} e^{-u_2^2/2}$ is either an odd or an even function being integrated over a finite symmetric range about zero. Thus, we obtain B_o from (3.5.10) and (3.5.15).

Turning our attention to the variance component of J_o, we recall (3.4.3) and write

$$\text{var}(\underline{d}'\hat{\underline{y}}_o) = \underline{d}'[\text{diag}(\underline{x}_1',\underline{x}_1',\ldots,\underline{x}_1')]\text{var}(\hat{\underline{\beta}}_1)[\text{diag}(\underline{x}_1,\underline{x}_1,\ldots,\underline{x}_1)]\underline{d}$$

$$+ E\{\delta\underline{d}'[\text{diag}(\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\ldots,\underline{x}_2'-\underline{x}_1'\underset{\sim}{A})]\ \hat{\underline{\beta}}_2\}^2$$

$$- E^2\{\delta\underline{d}'[\text{diag}(\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\ldots,\underline{x}_2'-\underline{x}_1'\underset{\sim}{A})]\ \hat{\underline{\beta}}_2\}$$

where $\text{cov}\{\underline{d}'[\text{diag}(\underline{x}_1',\underline{x}_1',\ldots,\underline{x}_1')]\hat{\underline{\beta}}_1,\ \delta\underline{d}'[\text{diag}(\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\underline{x}_2'-\underline{x}_1'\underset{\sim}{A},\ldots,\underline{x}_2'-\underline{x}_1'\underset{\sim}{A})]\ \hat{\underline{\beta}}_2\} = 0$, invoking Lemma 3.4.1. Employing (3.2.2), (3.2.3), (3.3.5), and (3.3.9) yields

$$V_o = V_1 + E(\delta\hat{B}_1) - N\{E[\delta(\underline{d}'\hat{\underline{\beta}}_2)]\}^2\ K\int_{-1}^{1}[x^2-(\overline{x^3}/\overline{x^2})x-\overline{x^2}]^2\,dx.$$

Manipulations similar to those leading to (3.5.8) and (3.5.10) result in

$$V_o = V_1 + a_1\underline{d}'\underset{\sim}{\Sigma}\ \underline{d}E(\delta Y_o) - N[(\overline{x^2}-1/3)^2+4/45+(\overline{x^3})^2/3(\overline{x^2})^2]$$
$$[(\overline{x^2}(\underline{d}'\underset{\sim}{\Sigma}\ \underline{d})/ND)^{1/2}E(\delta w)+P\ \underline{d}'\underline{\beta}_2]^2 \tag{3.5.16}$$

where $$Y_o = \hat{B}_1/a_1\underline{d}'\underset{\sim}{\Sigma}\ \underline{d}\ . \tag{3.5.17}$$

As was the case with (3.5.10), we note that the sign of $\underline{d}'\underline{\beta}_2$ does not affect V_o as a consequence of Lemma 3.5.1.

To evaluate $E(\delta Y_o)$, we use (3.3.18) and write $F_o > D_\alpha$ as

$$Y_o > R|U$$

where $$R|U = D_\alpha U/a_1(N-3)\ . \tag{3.5.18}$$

The random variable δY_0 has a truncated distribution for which

$$E(\delta Y_0) = \int_0^\infty \left[\int_{R|u}^\infty sf(s)ds\right] f(u)du \tag{3.5.19}$$

where f(u) is given by (3.3.2), and f(s) is the noncentral χ_1^2 density function of Y_0, i.e., using (3.3.15) and Rao (1965) (3b.1.15),

$$f(s) = e^{-a_3/2a_1} \sum_{i=0}^\infty \frac{(a_3/a_1)^i (1/2)^{(4i+1)/2}}{\Gamma(i+1)\Gamma(i+1/2)} s^{(2i-1)/2} e^{-s/2}. \tag{3.5.20}$$

Making the transformation $s_1 = s^{1/2}$

$$E(\delta Y_0) = 2e^{-a_3/2a_1} \int_0^\infty \left[\sum_{i=0}^\infty \frac{(a_3/a_1)^i (1/2)^{(4i+1)/2}}{\Gamma(i+1)\Gamma(i+1/2)} \int_{R|u}^\infty s_1^{2(i+1)} e^{-s_1^2/2} ds_1\right] \frac{1}{2^{\frac{N-3}{2}} \Gamma(\frac{N-3}{2})} u^{\frac{N-3}{2}-1} e^{-u/2} du . \tag{3.5.21}$$

Save for the special case of the next section, the evaluation of $E(\delta Y_0)$ will require numerical integration similar to (3.3.19), which for the case of the single independent variable is

$$1 - P = \int_0^\infty \left[\frac{1}{\sqrt{2\pi}} \int_{R_L|u}^{R_H|u} e^{-z^2/2} dz\right] \frac{1}{2^{\frac{N-3}{2}} \Gamma(\frac{N-3}{2})} u^{\frac{N-3}{2}-1} e^{-u/2} du . \tag{3.5.22}$$

Therefore, V_0 is obtained from (3.5.16) and (3.5.21).

3.6 A Special Case: $\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}$ Known

Knowledge of $\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}$ considerably reduces the magnitude of the problem. Our test statistic is now simply

$$C_o = \hat{B}_1/\underline{d}'\underset{\sim}{\Sigma}\,\underline{d} \ , \qquad (3.6.1)$$

and from (3.2.28) and (3.2.31) with no distributional assumptions, we

reject H_o: $a_3 \le a_1$ if $C_o > 2a_1$

accept H_o otherwise.

Under normality, we use (3.3.14) for

$$\begin{aligned} 1 - P &= Pr(C_o \le D_\alpha) \\ &\doteq Pr[a_1(w+(a_3/a_1)^{1/2})^2 \le D_\alpha] \\ &\doteq Pr[-(D_\alpha/a_1)^{1/2}-(a_3/a_1)^{1/2} \le w \le (D_\alpha/a_1)^{1/2}-(a_3/a_1)^{1/2}] \\ &\doteq \Phi[(D_\alpha/a_1)^{1/2}-(a_3/a_1)^{1/2}] - \Phi[-(D_\alpha/a_1)^{1/2}-(a_3/a_1)^{1/2}] \ . \end{aligned} \qquad (3.6.2)$$

The critical region $\hat{B}_1/\underline{d}'\underset{\sim}{\Sigma}\,\underline{d} > D_\alpha$ is equivalent to

$$Y > \hat{\hat{y}}_\alpha$$

where Y is given by (3.4.8) and analogous to (3.4.9),

$$\hat{\hat{y}}_\alpha = \frac{D_\alpha + a_3}{a_1} - \frac{2NK \int_R b_1 \hat{b}_1 d\underline{x}}{a_1 \underline{d}'\underset{\sim}{\Sigma}\,\underline{d}} \ . \qquad (3.6.3)$$

Replacing $\hat{y}_\alpha$ by $\hat{\tilde{y}}_\alpha$ in (3.4.13) determines $\hat{J}_{03}$ and thus $\hat{J}_0$.

For the single independent variable case, we adjust our critical region once more so that $\delta = 1$ only if

$$N(\underline{d}'\hat{\underline{\beta}}_2)^2 \, K \int_{-1}^{1} [x^2-(\overline{x^3}/\overline{x^2})x-\overline{x^2}]^2 dx > D_\alpha \underline{d}'\underset{\sim}{\Sigma}\,\underline{d} \, ,$$

and proceeding along lines similar to those resulting in (3.5.11), the preliminary test condition becomes

$$\begin{cases} w > R_H \\ \text{or} \\ w < R_L \end{cases} \tag{3.6.4}$$

with
$$R_H = (D_\alpha/a_1)^{1/2} - (a_3/a_1)^{1/2} \tag{3.6.5}$$

$$R_L = -(D_\alpha/a_1)^{1/2} - (a_3/a_1)^{1/2} \, . \tag{3.6.6}$$

From (A.1) with $\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}$ known,

$$E(\delta w) = (1/\sqrt{2\pi}) \, [e^{-R_H^2/2} - e^{-R_L^2/2}] \, . \tag{3.6.7}$$

Finally, we write $C_o > D_\alpha$ as

$$Y_o > D_\alpha/a_1$$

where Y_o is given by (3.5.17) so that from (3.5.21),

$$E(\delta Y_o) = 2e^{-a_3/2a_1} \sum_{i=0}^{\infty} \frac{(a_3/a_1)^i (1/2)^{(4i+1)/2}}{\Gamma(i+1)\Gamma(i+1/2)} \int_{D_\alpha/a_1}^{\infty} s_1^{2(i+1)} e^{-s_1^2/2} ds_1 . \tag{3.6.8}$$

Applying (3.4.12) once more, (3.6.8) converges quite rapidly for representative values of a_3/a_1. Again $J_o = V_o + B_o$ is evaluated by using (3.6.7) and (3.6.8).

We present a simple example illustrating the concepts of the last four sections. Suppose we are dealing with a single independent variable and

$N = 3$, $\underline{d}' = [1,1]$, $\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}$ (known) $= 1$,

$\underline{y_1'} = [2,-2,-1]$, $\underline{y_2'} = [1,0,2]$, and

$$\underset{\sim}{X}_o^* = \begin{array}{c} \begin{array}{ccc} & x & x^2-\overline{x^2} \end{array} \\ \begin{bmatrix} 1 & -1 & 1/3 \\ 1 & 0 & -2/3 \\ 1 & 1 & 1/3 \end{bmatrix} \end{array} .$$

Since $\overline{x^2} = 2/3$, $\overline{x^3} = 0$, and $\overline{x^{4}{}'} = 2/9$, we use the results of section 3.5 to find

$$D = (\overline{x^2})(\overline{x^{4}{}'}) - (\overline{x^3})^2 = 4/27$$

$$\underset{\sim}{M} = \overline{x^2}/ND = 3/2$$

$$\underset{\sim}{A} = \underline{0}$$

$$\underset{\sim}{M}_{212} = [(\overline{x^2}-1/3)^2 + 4/45 + (\overline{x^3})^2/3(\overline{x^2})^2] = 1/5$$

$$a_1 = N\,\mathrm{tr}(\underset{\sim}{M}\,\underset{\sim}{M}_{212}) = 9/10$$

$$a_3 = N(\underline{d}'\underline{\beta_2})'\,\underset{\sim}{M}_{212}(\underline{d}'\underline{\beta_2})/\underline{d}'\underset{\sim}{\Sigma}\,\underline{d} = (3/5)(\beta_{21}+\beta_{22})^2$$

$$(\underline{\beta_2'} = [\beta_{21},\beta_{22}])$$

$$a_3/a_1 = (2/3)(\beta_{21}+\beta_{22})^2 \ .$$

If $\alpha = .05$, then we substitute $a_3/a_1 = 1$ in (3.6.2) and obtain

$$.95 = \Phi[(D_\alpha/.9)^{1/2}-1] - \Phi[-(D_\alpha/.9)^{1/2}-1]$$

for which $D_\alpha = .9(2.65)^2 = 6.320$.

Since $\underset{\sim}{X}_1'\underset{\sim}{X}_2 = \underline{0}$,

$$\hat{\beta}_{21} = (\underset{\sim}{X}_2'\underset{\sim}{X}_2)^{-1}\underset{\sim}{X}_2' \ \underline{y_1} = 5/2$$

$$\hat{\beta}_{22} = (\underset{\sim}{X}_2'\underset{\sim}{X}_2)^{-1}\underset{\sim}{X}_2' \ \underline{y_2} = 3/2$$

$$C_o = \frac{N(\hat{\beta}_{21}+\hat{\beta}_{22})^2\underset{\sim}{M}_{212}}{\underline{d}'\underset{\sim}{\Sigma}\ \underline{d}} = 9.600 > 6.320 = D_\alpha \ .$$

We reject H_o and fit the quadratic model $\hat{\hat{y}}$. To evaluate J_o, we require

$$\underset{\sim}{M}_{11}^{-1} = N(\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 3/2 \end{bmatrix}$$

$$\underset{\sim}{\mu}_{11} = K\int_{-1}^{1} \underline{x_1}\ \underline{x_1'}\ d\underline{x} = \begin{bmatrix} 1 & 0 \\ 0 & 1/3 \end{bmatrix}$$

$$V_1 = (\underline{d}'\underset{\sim}{\Sigma}\ \underline{d})\mathrm{tr}(\underset{\sim}{M}_{11}^{-1}\underset{\sim}{\mu}_{11}) = 3/2 \ .$$

Suppose $(\beta_{21}+\beta_{22})^2 = 15$ so that $a_3/a_1 = 10$. Equation (3.6.2) gives

$$1 - P = \Phi[(D_\alpha/.9)^{1/2}-\sqrt{10}] - \Phi[-(D_\alpha/.9)^{1/2}-\sqrt{10}]$$

whence P = 0.695. Also,

$$R_H = (D_\alpha/.9)^{1/2} - \sqrt{10} = -0.512$$

$$R_L = -(D_\alpha/.9)^{1/2} - \sqrt{10} = -5.812$$

$$E(\delta w) = (1/\sqrt{2\pi})\,[e^{-R_H^2/2} - e^{-R_L^2/2}] = 0.350 .$$

Applying (3.6.8) results in $E(\delta Y_o) = 9.691$. Substituting in (3.5.10) and (3.5.16) yields $B_o = 0.340$ and $V_o = 4.625$ so that $J_o = 4.965$.

3.7 Design Considerations

Before leaving the standard model, we shall briefly touch upon the problem of "optimal" design, in particular that of choosing a design to maximize power for the case of the single independent variable, i.e., for $a_3/a_1 > 1$, we wish to maximize

$$\begin{aligned} P &= Pr(F_o > D_\alpha | a_3/a_1 > 1) \\ &= Pr(F'_{1,N-q_o,(a_3/a_1)^{1/2}} > D_\alpha/a_1 | a_3/a_1 > 1) . \end{aligned} \quad (3.7.1)$$

It is clear from (3.5.20) that P is an increasing function of a_3/a_1; hence for the single independent variable case, we seek to maximize (3.5.3) for fixed N.

Examining $D/\overline{x^2} = \overline{x^4}' - (\overline{x^3})^2/\overline{x^2}$

gives $\overline{x^3} = 0$ as an initial condition. We also wish to design such that

$$\overline{x^{4}}' = \sum_{j=1}^{N} (x_{1j}^{2}-\overline{x^{2}})^{2}/N$$

$$= \sum_{j=1}^{N} x_{1j}^{4}/N - (\sum_{j=1}^{N} x_{1j}^{2}/N)^{2} \tag{3.7.2}$$

is maximized subject to $|x_{1j}| \leq 1$, $j = 1, 2, \ldots, N$. Applying the inequality

$$\sum_{j=1}^{N} x_{1j}^{4}/N \leq \sum_{j=1}^{N} x_{1j}^{2}/N \quad , \; |x_{1j}| \leq 1$$

to (3.7.2) gives $\overline{x_4}' \leq 1/4$. If N is a multiple of four, then the design which maximizes P is N/4 points at -1, N/2 points at 0, and N/4 points at 1 since this configuration achieves $\overline{x_4}' = 1/4$. Designs with a concentration of center points and remaining points split equally at $\pm$ 1 also seem to be effective if N is not a multiple of four.

Maximizing a_3/a_1 in general proves much more difficult than for the single independent variable case. Also, in the search for design values which minimize B_o or V_o, much less J_o, even for the case of the single independent variable, one is led to trial and error with respect to different designs or empirical minimization as the only practical solution.

Chapter IV

GENERALIZED MULTIVARIATE REGRESSION MODEL

4.1 An Expanded Notation

In Chapter III, each of the p response vectors was dependent upon the same regression matrix, either $\underset{\sim}{X}_1$ or $\underset{\sim}{X}_o^*$. We now relax this requirement and postulate a model of the form

$$\underline{y_j} = \underset{N\times q_{1j}}{\underset{\sim}{X}_{1j}} \underline{\beta_{1j}} + \underline{\varepsilon_j} \quad , j = 1, 2, \ldots, p, \tag{4.1.1}$$

or

$$\underline{y} = \underset{\sim}{X} \underline{\beta_1} + \underline{\varepsilon} \tag{4.1.2}$$

where now, $\underset{Np\times q_1^*}{\underset{\sim}{X}} = \mathrm{diag}[\underset{\sim}{X}_{11}, \underset{\sim}{X}_{12}, \ldots, \underset{\sim}{X}_{1p}]$

$$q_1^* = \sum_{j=1}^{p} q_{1j} \ .$$

The true model becomes

$$\underline{y_j} = \underset{\sim}{X}_{1j} \underline{\beta_{1j}} + \underset{N\times q_{2j}}{\underset{\sim}{X}_{2j}} \underline{\beta_{2j}} + \underline{\varepsilon_j} \ , j = 1, 2, \ldots, p \tag{4.1.3}$$

$$= \underset{\sim}{X}_j^* \underline{\beta_j^*} + \underline{\varepsilon_j} \tag{4.1.4}$$

for $\underset{N\times q_{oj}}{\underset{\sim}{X}_j^*} = [\underset{\sim}{X}_{1j} : \underset{\sim}{X}_{2j}]$

$$q_{oj} = q_{1j} + q_{2j} \ .$$

Consolidating gives

$$\underline{y} = \underset{\sim}{X}^{*}\underline{\beta}^{*} + \underline{\varepsilon} \tag{4.1.5}$$

with $\underset{Npxq}{\underset{\sim}{X}^{*}} = \text{diag}[\underset{\sim}{X}_1^{*}, \underset{\sim}{X}_2^{*}, \ldots, \underset{\sim}{X}_p^{*}]$

$$q = \sum_{j=1}^{p} q_{oj} \ .$$

If rank $(\underset{\sim}{X}_j^{*}) = q_{oj}$, our model assumptions are identical to those of section 3.1.

Unlike the standard model, the multivariate generalized least squares estimators of the parameter vectors do not reduce to univariate least squares estimators so we write only

$$\hat{\underline{\beta_1}} = [\underset{\sim}{X}'(\underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_N)^{-1}\underset{\sim}{X}]^{-1} \underset{\sim}{X}' (\underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_N)^{-1}\underline{y} \tag{4.1.6}$$

$$\hat{\underline{\beta}}^{*} = [\underset{\sim}{X}^{*}{}'(\underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_N)^{-1}\underset{\sim}{X}^{*}]^{-1}\underset{\sim}{X}^{*}{}'(\underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_N)^{-1}\underline{y} \ . \tag{4.1.7}$$

Employing (4.1.6) and (4.1.7), we fit either

$$\hat{y}_j = \underline{x_{1j}'}\ \underline{\hat{\beta}_{1j}} \tag{4.1.8}$$

or

$$\hat{\hat{y}}_j = \underline{x_{1j}'}\ \underline{\hat{\tilde{\beta}}_{1j}} + \underline{x_{2j}'}\ \underline{\hat{\beta}_{2j}}$$

$$= \underline{x_j^{*}{}'}\ \underline{\hat{\beta}_j^{*}} \tag{4.1.9}$$

where $\underline{x_{1j}'}$, $\underline{x_{2j}'}$, and $\underline{x_j^{*}{}'}$ are typical row vectors in the matrices $\underset{\sim}{X}_{1j}$, $\underset{\sim}{X}_{2j}$, and $\underset{\sim}{X}_j^{*}$.

We again wish to estimate our response function based on the hypothesis comparing the integrated mean squared errors of linear combinations of the estimated responses for the two models, i.e., H_o: $J_1 \leq J_2$. If R_j denotes the region of interest associated with $\underset{\sim}{X}_{1j}$, we define

$$K_j^{-1} = \int_{R_j} d\, \underline{x_{1j}}$$

$$K^{-1} = \int_R d\underline{x} = \int_{R_1} \int_{R_2} \cdots \int_{R_p} d\, \underline{x_{11}}\, d\, \underline{x_{12}} \cdots d\, \underline{x_{1p}}$$

$$= \prod_{j=1}^{p} K_j^{-1} .$$

There is again no integrated bias contribution to J_2 since $E(\hat{\underline{\beta}}^*) = \underline{\beta}^*$. We write our hypothesis as

$$H_o\colon\ B_1/(V_2-V_1) \leq 1 . \tag{4.1.10}$$

To obtain the quantities in H_o, we require

$$E(\underline{\hat{\beta}_1}) = [\underset{\sim}{X}'(\underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_N)^{-1}\underset{\sim}{X}]^{-1}\underset{\sim}{X}'(\underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_N)^{-1}E(\underset{\sim}{X}^*\underline{\beta}^*+\underline{\varepsilon})$$

$$= [\underset{\sim}{X}'(\underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_N)^{-1}\underset{\sim}{X}]^{-1}\underset{\sim}{X}'(\underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_N)^{-1}[\underset{\sim}{X}\ \beta_1 + \underset{Npxq_2^*}{\mathrm{diag}}(\underset{\sim}{X}_{21},\underset{\sim}{X}_{22},\ldots,\underset{\sim}{X}_{2p})\underline{\beta_2}]$$

$$(q_2^* = \sum_{j=1}^{p} q_{2j})$$

$$= \underline{\beta_1} + \underset{\sim}{A}_o\ \underline{\beta_2} \tag{4.1.11}$$

where $\underset{\sim}{A}_o = [\underset{\sim}{X}'(\underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_N)^{-1}\underset{\sim}{X}]^{-1}\underset{\sim}{X}'(\underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_N)^{-1}[\mathrm{diag}(\underset{\sim}{X}_{21},\underset{\sim}{X}_{22},\ldots,\underset{\sim}{X}_{2p})]$. (4.1.12)

Therefore, $E(\underline{d}'\hat{\underline{y}}) = \underline{d}'[\mathrm{diag}(\underline{x_{11}'},\underline{x_{12}'},\ldots,\underline{x_{1p}'})]E(\hat{\underline{\beta_1}})$

$$= \underline{d}'[\mathrm{diag}(\underline{x_{11}'},\underline{x_{12}'},\ldots,\underline{x_{1p}'})](\underline{\beta_1}+\underset{\sim}{A}_o\ \underline{\beta_2}) \ .$$

Also, $\underline{d}'\underline{\eta} = \underline{d}'[\mathrm{diag}(\underline{x_1^{*}}',\underline{x_2^{*}}',\ldots,\underline{x_p^{*}}')]\underline{\beta^{*}}$

$$= \underline{d}'[\mathrm{diag}(\underline{x_{11}'},\underline{x_{12}'},\ldots,\underline{x_{1p}'})\underline{\beta_1} + \mathrm{diag}(\underline{x_{21}'},\underline{x_{22}'},\ldots,\underline{x_{2p}'})\underline{\beta_2}] \ .$$

Thus, $b_1 = E(\underline{d}'\hat{\underline{y}}) - \underline{d}'\underline{\eta}$

$$= \underline{d}'[\mathrm{diag}(\underline{x_{11}'},\underline{x_{12}'},\ldots,\underline{x_{1p}'})\underset{\sim}{A}_o - \mathrm{diag}(\underline{x_{21}'},\underline{x_{22}'},\ldots,\underline{x_{2p}'})]\underline{\beta_2} \tag{4.1.13}$$

and $B_1 = NK \int_R b_1^{\,2} d\underline{x}$.

Now $V_1 = NK \int_R \mathrm{var}(\underline{d}'\hat{\underline{y}})\ d\underline{x}$

$$= NK \int_R \underline{d}'[\mathrm{var}(\hat{\underline{y}})]\underline{d}\ d\underline{x}$$

$$= NK \int_R \underline{d}'[\mathrm{diag}(\underline{x_{11}'},\underline{x_{12}'},\ldots,\underline{x_{1p}'})]\mathrm{var}(\hat{\underline{\beta_1}})[\mathrm{diag}(\underline{x_{11}},\underline{x_{12}},\ldots,\underline{x_{1p}})]\underline{d}\ d\underline{x}$$

$$= NK \int_R \underline{d}'[\mathrm{diag}(\underline{x_{11}'},\underline{x_{12}'},\ldots,\underline{x_{1p}'})][\underset{\sim}{X}'(\underset{\sim}{\Sigma}\otimes \underset{\sim}{I}_N)^{-1}\underset{\sim}{X}]^{-1}[\mathrm{diag}(\underline{x_{11}},\underline{x_{12}},\ldots,\underline{x_{1p}})]\underline{d}\ d\underline{x} \ . \tag{4.1.14}$$

(Press (1972) (8.5.12))

Similarly,

$$V_2 = NK \int_R \underline{d}'[\mathrm{diag}(\underline{x_1^{*}}',\underline{x_2^{*}}',\ldots,\underline{x_p^{*}}')][\underset{\sim}{X}^{*}{}'(\underset{\sim}{\Sigma}\otimes \underset{\sim}{I}_N)^{-1}\underset{\sim}{X}^{*}]^{-1}[\mathrm{diag}(\underline{x_1^{*}},\underline{x_2^{*}},\ldots,\underline{x_p^{*}})]\underline{d}\ d\underline{x} \tag{4.1.15}$$

where $\mathrm{var}(\hat{\underline{\beta}}^*) = [\underset{\sim}{X}^{*\prime}(\underset{\sim}{\Sigma} \otimes \underset{\sim}{I}_N)^{-1}\underset{\sim}{X}^*]^{-1}$.

We denote by $\underset{\sim}{M}_o$ the submatrix of $\mathrm{var}(\hat{\underline{\beta}}^*)$ associated with $\hat{\underline{\beta}}_2$ so that

$$\mathrm{var}(\hat{\underline{\beta}}_2) = \underset{\sim}{M}_o \,. \tag{4.1.16}$$

In order to develop a test statistic, we estimate $\underset{\sim}{\Sigma}$ by

$$\hat{\underset{\sim}{\Sigma}} = (\hat{\sigma}_{ij})$$

where $[N - q_{oi} - q_{oj} + \mathrm{tr}(\underset{\sim}{X}_i^*(\underset{\sim}{X}_i^{*\prime}\underset{\sim}{X}_i^*)^{-1}\underset{\sim}{X}_i^{*\prime}\underset{\sim}{X}_j^*(\underset{\sim}{X}_j^{*\prime}\underset{\sim}{X}_j^*)^{-1}\underset{\sim}{X}_j^{*\prime})]\hat{\sigma}_{ij} = \hat{\underline{u}}_i' \, \hat{\underline{u}}_j$

$$\begin{aligned}[\hat{\underline{u}}_1, \hat{\underline{u}}_2, \ldots, \hat{\underline{u}}_p] &= \hat{\underset{\sim}{U}} \\ &= \underset{\sim}{Y} - \underset{\sim}{Z}\,\hat{\underset{\sim}{\beta}}^* \end{aligned} \tag{4.1.17}$$

$$\underset{\sim}{Y} = [\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_p]$$

$$\underset{\sim}{Z} = [\underset{\sim}{X}_1^*, \underset{\sim}{X}_2^*, \ldots, \underset{\sim}{X}_p^*]$$

$$\hat{\underset{\sim}{\beta}}^* = \mathrm{diag}[\hat{\underline{\beta}}_1^*, \hat{\underline{\beta}}_2^*, \ldots, \hat{\underline{\beta}}_p^*] \,.$$

The natural test statistic for (4.1.10) is

$$F_1 = \hat{B}_1/(\hat{V}_2 - \hat{V}_1) \tag{4.1.18}$$

where $\hat{V}_2$ and $\hat{V}_1$ are given by (4.1.15) and (4.1.14) with $\hat{\underset{\sim}{\Sigma}}$ vice $\underset{\sim}{\Sigma}$, $\hat{B}_1$ is given by (3.3.9), and $\hat{b}_1$ is given by (4.1.13) with $\hat{\underline{\beta}}_2$ vice $\underline{\beta}_2$. There are several difficulties with (4.1.18) in general. If we are to justify a test procedure in terms of a ratio of expected values, the expectations of matrices of the form $[\underset{\sim}{X}'(\hat{\underset{\sim}{\Sigma}} \otimes \underset{\sim}{I}_N)^{-1}\underset{\sim}{X}]^{-1}$ do not lend themselves to explicit determination. Further, $\hat{\underset{\sim}{\beta}}^*$, $\underline{\hat{\beta}_2}$, and $\underset{\sim}{A}_o$ involve $\underset{\sim}{\Sigma}$ which is unknown. If our test statistic is altered to reflect estimation of $\underset{\sim}{\Sigma}$, then we are unable to obtain $E(\hat{B}_1)$, nor $E(\hat{V}_2-\hat{V}_1)$. We shall see that (4.1.18) becomes more useful in the next section.

4.2 A Special Case: $\underset{\sim}{\Sigma}$ Diagonal Unknown

If the error covariance matrix is diagonal, then (4.1.6) and (4.1.7) reduce to

$$\underline{\hat{\beta}_1} = \begin{bmatrix} (\underset{\sim}{X}_{11}'\underset{\sim}{X}_{11})^{-1}\underset{\sim}{X}_{11}' \ \underline{y_1} \\ (\underset{\sim}{X}_{12}'\underset{\sim}{X}_{12})^{-1}\underset{\sim}{X}_{12}' \ \underline{y_2} \\ \vdots \\ (\underset{\sim}{X}_{1p}'\underset{\sim}{X}_{1p})^{-1}\underset{\sim}{X}_{1p}' \ \underline{y_p} \end{bmatrix} \tag{4.2.1}$$

$$\underline{\hat{\beta}^*} = \begin{bmatrix} (\underset{\sim}{X}_1^{*\prime}\underset{\sim}{X}_1^*)^{-1}\underset{\sim}{X}_1^{*\prime} \ \underline{y_1} \\ (\underset{\sim}{X}_2^{*\prime}\underset{\sim}{X}_2^*)^{-1}\underset{\sim}{X}_2^{*\prime} \ \underline{y_2} \\ \vdots \\ (\underset{\sim}{X}_p^{*\prime}\underset{\sim}{X}_p^*)^{-1}\underset{\sim}{X}_p^{*\prime} \ \underline{y_p} \end{bmatrix} , \tag{4.2.2}$$

the univariate least squares estimators for (4.1.1) and (4.1.4). If we define

$$\underset{\sim}{A}_j = (\underset{\sim}{X}_{1j}'\underset{\sim}{X}_{1j})^{-1}\underset{\sim}{X}_{1j}'\underset{\sim}{X}_{2j} \ , \ j = 1, 2, \ldots, p, \tag{4.2.3}$$

then $$\underset{\sim}{A}_o = \text{diag}[\underset{\sim}{A}_1,\underset{\sim}{A}_2,\ldots,\underset{\sim}{A}_p] \ . \tag{4.2.4}$$

Let $\underset{\sim}{M}_j$ denote the submatrix of $(\underset{\sim}{X}_j^{*\prime}\underset{\sim}{X}_j^*)^{-1}$ corresponding to $\underset{\sim}{M}$ in (3.2.7). Then

$$\text{var}(\hat{\underline{\beta}}^*) = \text{diag}[\sigma_{11}(\underset{\sim}{X}_1^{*\prime}\underset{\sim}{X}_1^*)^{-1},\sigma_{22}(\underset{\sim}{X}_2^{*\prime}\underset{\sim}{X}_2^*)^{-1},\ldots,\sigma_{pp}(\underset{\sim}{X}_p^{*\prime}\underset{\sim}{X}_p^*)^{-1}]$$

$$\text{var}(\hat{\underline{\beta}}_2) = \underset{\sim}{M}_o = \text{diag}[\sigma_{11}\underset{\sim}{M}_1,\sigma_{22}\underset{\sim}{M}_2,\ldots,\sigma_{pp}\underset{\sim}{M}_p] \ . \qquad (4.2.5)$$

From (4.1.13) and (4.1.14),

$$b_1 = \underline{d}'[\text{diag}(\underline{x}_{11}'\underset{\sim}{A}_1-\underline{x}_{21}',\underline{x}_{12}'\underset{\sim}{A}_2-\underline{x}_{22}',\ldots,\underline{x}_{1p}'\underset{\sim}{A}_p-\underline{x}_{2p}')]\ \underline{\beta}_2 \qquad (4.2.6)$$

$$V_1 = NK \int_R \underline{d}'[\text{diag}(\sigma_{11}\underline{x}_{11}'(\underset{\sim}{X}_{11}'\underset{\sim}{X}_{11})^{-1}\underline{x}_{11},\sigma_{22}\underline{x}_{12}'(\underset{\sim}{X}_{12}'\underset{\sim}{X}_{12})^{-1}\underline{x}_{12},$$

$$\ldots,\sigma_{pp}\underline{x}_{1p}'(\underset{\sim}{X}_{1p}'\underset{\sim}{X}_{1p})^{-1}\underline{x}_{1p})]\ \underline{d}\ d\underline{x}$$

$$= \sum_{j=1}^{p} NK \int_R d_j^2[\underline{x}_{1j}'(\underset{\sim}{X}_{1j}'\underset{\sim}{X}_{1j})^{-1}\underline{x}_{1j}]\sigma_{jj}\ d\underline{x}$$

where $\underline{d}' = [d_1,d_2,\ldots,d_p]$.

Similarly, from (4.1.15),

$$V_2 = \sum_{j=1}^{p} NK \int_R d_j^2[\underline{x}_j^{*\prime}(\underset{\sim}{X}_j^{*\prime}\underset{\sim}{X}_j^*)^{-1}\underline{x}_j^*]\ \sigma_{jj}\ d\underline{x}$$

$$= \sum_{j=1}^{p} NK \int_R d_j^2(\underline{x}_{1j}'(\underset{\sim}{X}_{1j}'\underset{\sim}{X}_{1j})^{-1}\underline{x}_{1j} + (\underline{x}_{1j}'\underset{\sim}{A}_j-\underline{x}_{2j}')\underset{\sim}{M}_j(\underset{\sim}{A}_j'\underline{x}_{1j}-\underline{x}_{2j}))\sigma_{jj}d\underline{x},$$

adapting the development of (3.2.10) to each of the p terms $[\underline{x}_j^{*\prime}(\underset{\sim}{X}_j^{*\prime}\underset{\sim}{X}_j^*)^{-1}\underline{x}_j^*]$. Thus,

$$V_2 - V_1 = \sum_{j=1}^{p} NK \int_R d_j^{\,2}(\underline{x}'_{1j}\underset{\sim}{A}_j-\underline{x}'_{2j})\underset{\sim}{M}_j(\underset{\sim}{A}'_j\underline{x}_{1j}-\underline{x}_{2j})\ \sigma_{jj}\ d\underline{x}$$

$$= \sum_{j=1}^{p} N\ d_j^{\,2}\ \mathrm{tr}[\underset{\sim}{M}_j(K \int_R (\underset{\sim}{A}'_j\underline{x}_{1j}-\underline{x}_{2j})(\underline{x}'_{1j}\underset{\sim}{A}_j-\underline{x}'_{2j})d\underline{x})]\ \sigma_{jj}$$

$$= \sum_{j=1}^{p} a_{jj}d_j^{\,2}\sigma_{jj} \tag{4.2.7}$$

for $a_{jj} = N\ \mathrm{tr}[\underset{\sim}{M}_j(K \int_R (\underset{\sim}{A}'_j\underline{x}_{1j}-\underline{x}_{2j})(\underline{x}'_{1j}\underset{\sim}{A}_j-\underline{x}'_{2j})d\underline{x})]$. (4.2.8)

Equation (4.1.18) becomes

$$F_1 = \hat{B}_1 / \sum_{j=1}^{p} a_{jj}d_j^{\,2}\hat{\sigma}_{jj}\ , \tag{4.2.9}$$

and $\hat{b}_1 = \underline{d}'[\mathrm{diag}(\underline{x}'_{11}\underset{\sim}{A}_1-\underline{x}'_{21},\underline{x}'_{12}\underset{\sim}{A}_2-\underline{x}'_{22},\ldots,\underline{x}'_{1p}\underset{\sim}{A}_p-\underline{x}'_{2p}]\ \hat{\underline{\beta}}_2$. (4.2.10)

In general, we can write (4.1.17) as

$$\hat{\sigma}_{ij} = \frac{(\underline{y}_i-\underset{\sim}{X}_i^*\hat{\underline{\beta}}_i^*)'(\underline{y}_j-\underset{\sim}{X}_j^*\hat{\underline{\beta}}_j^*)}{[N-q_{oi}-q_{oj}+\mathrm{tr}(\underset{\sim}{X}_i^*(\underset{\sim}{X}_i^{*\prime}\underset{\sim}{X}_i^*)^{-1}\underset{\sim}{X}_i^{*\prime}\underset{\sim}{X}_j^*(\underset{\sim}{X}_j^{*\prime}\underset{\sim}{X}_j^*)^{-1}\underset{\sim}{X}_j^{*\prime})]}$$

$$= \frac{\underline{y}'_i(\underset{\sim}{I}_N-\underset{\sim}{X}_i^*(\underset{\sim}{X}_i^{*\prime}\underset{\sim}{X}_i^*)^{-1}\underset{\sim}{X}_i^{*\prime})(\underset{\sim}{I}_N-\underset{\sim}{X}_j^*(\underset{\sim}{X}_j^{*\prime}\underset{\sim}{X}_j^*)^{-1}\underset{\sim}{X}_j^{*\prime})\underline{y}_j}{[N-q_{oi}-q_{oj}+\mathrm{tr}(\underset{\sim}{X}_i^*(\underset{\sim}{X}_i^{*\prime}\underset{\sim}{X}_i^*)^{-1}\underset{\sim}{X}_i^{*\prime}\underset{\sim}{X}_j^*(\underset{\sim}{X}_j^{*\prime}\underset{\sim}{X}_j^*)^{-1}\underset{\sim}{X}_j^{*\prime})]} .$$

Using $E(\underline{y}_j\ \underline{y}'_i) = \sigma_{ij}\underset{\sim}{I}_N + \underset{\sim}{X}_j^*\underline{\beta}_j^*\ \underline{\beta}_i^{*\prime}\underset{\sim}{X}_i^{*\prime}$

and $\underset{\sim}{X}_i^{*\prime}(\underset{\sim}{I}_N-\underset{\sim}{X}_i^*(\underset{\sim}{X}_i^{*\prime}\underset{\sim}{X}_i^*)^{-1}\underset{\sim}{X}_i^{*\prime}) = \underset{\sim}{0}$,

$$E(\hat{\sigma}_{ij}) = \frac{tr[E(\underline{y}_j\ \underline{y}_i')(\underset{\sim}{I}_N-\underset{\sim}{X}_i^*(\underset{\sim}{X}_i^{*}{}'\underset{\sim}{X}_i^*)^{-1}\underset{\sim}{X}_i^{*}{}')(\underset{\sim}{I}_N-\underset{\sim}{X}_j^*(\underset{\sim}{X}_j^{*}{}'\underset{\sim}{X}_j^*)^{-1}\underset{\sim}{X}_j^{*}{}')]}{[N-q_{oi}-q_{oj}+tr(\underset{\sim}{X}_i^*(\underset{\sim}{X}_i^{*}{}'\underset{\sim}{X}_i^*)^{-1}\underset{\sim}{X}_i^{*}{}'\underset{\sim}{X}_j^*(\underset{\sim}{X}_j^{*}{}'\underset{\sim}{X}_j^*)^{-1}\underset{\sim}{X}_j^{*}{}')]}$$

$$= \sigma_{ij}\ .$$

Hence, $E(\sum_{j=1}^{p} a_{jj}d_j^2\hat{\sigma}_{jj}) = \sum_{j=1}^{p} a_{jj}d_j^2\sigma_{jj}\ .$ (4.2.11)

Now $E(\hat{B}_1) = NK \int_R E(\hat{b}_1)^2\, d\underline{x}$

$$= NK \int_R [var(\hat{b}_1)+E^2(\hat{b}_1)]\, d\underline{x}\ . \tag{4.2.12}$$

Comparing (4.2.10) and (4.2.6) gives $E(\hat{b}_1) = b_1$.

Also from (4.2.10) and (4.2.5),

$$var(\hat{b}_1) = \sum_{j=1}^{p} d_j^2(\underline{x}_{1j}'\underset{\sim}{A}_j-\underline{x}_{2j}')\underset{\sim}{M}_j(\underset{\sim}{A}_j'\underline{x}_{1j}-\underline{x}_{2j})\ \sigma_{jj}\ , \tag{4.2.13}$$

so that applying (4.2.8), we have

$$E(\hat{B}_1) = \sum_{j=1}^{p} a_{jj}d_j^2\sigma_{jj} + B_1\ . \tag{4.2.14}$$

The ratio of expected values in (4.2.9) is

$$1 + B_1/\sum_{j=1}^{p} a_{jj}d_j^2\sigma_{jj}\ . \tag{4.2.15}$$

If no distributional assumptions are made, we

reject H_o if $F_1 > 2$

accept H_o otherwise.

We now assume normality of the error vector $\underline{\varepsilon}$ as in (3.3.1). Since

$$\hat{\sigma}_{jj} = (\underline{y}_j - \underset{\sim}{X}_j^* \underline{\hat{\beta}}_j^*)'(\underline{y}_j - \underset{\sim}{X}_j^* \underline{\hat{\beta}}_j^*)/(N-q_{oj})$$

$$= \underline{y}_j'[\underset{\sim}{I}_N - \underset{\sim}{X}_j^*(\underset{\sim}{X}_j^{*\prime}\underset{\sim}{X}_j^*)^{-1}\underset{\sim}{X}_j^{*\prime}]\ \underline{y}_j/(N-q_{oj}) ,$$

$$(N-q_{oj})\hat{\sigma}_{jj}/\sigma_{jj} \sim \chi^2_{N-q_{oj}} \tag{4.2.16}$$

(Graybill (1961) Theorem 6.1).

The $\hat{\sigma}_{jj}$, $j = 1, 2, \ldots, p$, are independent, and using an approximation due to Satterthwaite (1946), we write

$$\sum_{j=1}^{p} a_{jj} d_j^{\,2} \hat{\sigma}_{jj} \approx g_o \chi^2_{h_o}/h_o$$

where $$g_o = \sum_{j=1}^{p} a_{jj} d_j^{\,2} \sigma_{jj} = V_2 - V_1$$

$$h_o = \frac{\left(\sum_{j=1}^{p} a_{jj} d_j^{\,2} \sigma_{jj}\right)^2}{\sum_{j=1}^{p} (a_{jj} d_j^{\,2} \sigma_{jj})^2/(N-q_{oj})} .$$

In order to use this result, we estimate the latter quantity by

$$\hat{h}_o = \frac{\left(\sum_{j=1}^{p} a_{jj} d_j^{\,2} \hat{\sigma}_{jj}\right)^2}{\sum_{j=1}^{p} (a_{jj} d_j^{\,2} \hat{\sigma}_{jj})^2/(N-q_{oj})} , \tag{4.2.17}$$

so that $$\frac{\sum_{j=1}^{p} a_{jj} d_j^{\,2} \hat{\sigma}_{jj}}{V_2 - V_1} \approx \frac{\chi^2_{\hat{h}_o}}{\hat{h}_o} . \qquad (4.2.18)$$

From Press (1961) page 222,

$$\underline{\hat{\beta}_2} \sim N(\underline{\beta_2}, \underset{\sim}{M}_o) \qquad (4.2.19)$$

where for $\underset{\sim}{\Sigma}$ diagonal, $\underset{\sim}{M}_o$ is given by (4.2.5). For $\hat{b}_1$ given by (4.2.10), we have $\hat{b}_1 \sim N(b_1, \mathrm{var}(\hat{b}_1))$ with $\mathrm{var}(\hat{b}_1)$ as in (4.2.13).

Define $$b(\underline{x_j^*}) = (\underline{x_{1j}'}\underset{\sim}{A}_j - \underline{x_{2j}'})\underset{\sim}{M}_j(\underset{\sim}{A}_j'\underline{x_{1j}} - \underline{x_{2j}}) \qquad (4.2.20)$$

so that $$\mathrm{var}(\hat{b}_1) = \sum_{j=1}^{p} b(\underline{x_j^*}) d_j^{\,2} \sigma_{jj} .$$

Then, $$\hat{b}_1 / [\sum_{j=1}^{p} b(\underline{x_j^*}) d_j^{\,2} \sigma_{jj}]^{1/2} \sim N(b_1 / [\sum_{j=1}^{p} b(\underline{x_j^*}) d_j^{\,2} \sigma_{jj}]^{1/2}, 1)$$

$$(\hat{b}_1)^2 \sim [\sum_{j=1}^{p} b(\underline{x_j^*}) d_j^{\,2} \sigma_{jj}] \, \chi'^{\,2}_{1,\lambda(\underline{x}^*)}$$

where $$\lambda(\underline{x}^*) = b_1 / [\sum_{j=1}^{p} b(\underline{x_j^*}) d_j^{\,2} \sigma_{jj}]^{1/2} .$$

For $w \sim N(0,1)$,

$$\hat{B}_1 \sim NK \int_R [\sum_{j=1}^{p} b(\underline{x_j^*}) d_j^{\,2} \sigma_{jj}] (w + \lambda(\underline{x}^*))^2 d\underline{x}$$

$$\sim (V_2 - V_1) w^2 + 2wNK \int_R [\sum_{j=1}^{p} b(\underline{x_j^*}) d_j^{\,2} \sigma_{jj}]^{1/2} b_1 d\underline{x} + B_1 .$$

Analogous to (3.3.13), we base a bound approximation on

$$NK \int_R \left[\sum_{j=1}^{p} b(\underline{x}_j^*) d_j^2 \sigma_{jj}\right]^{1/2} b_1 d\underline{x} \le \left[\left(NK \int_R \sum_{j=1}^{p} b(\underline{x}_j^*) d_j^2 \sigma_{jj} d\underline{x}\right)\left(NK \int_R b_1^2 d\underline{x}\right)\right]^{1/2}$$

$$\le [(V_2 - V_1)B_1]^{1/2} .$$

Therefore, $$\frac{NK \int_R [b(\underline{x}_j^*) d_j^2 \sigma_{jj}]^{1/2} b_1 d\underline{x}}{V_2 - V_1} \le \left(\frac{B_1}{V_2 - V_1}\right)^{1/2} \qquad (4.2.21)$$

$$\hat{B}_1 \sim (V_2 - V_1)\left[w^2 + \frac{2wNK \int_R \left(\sum_{j=1}^{p} b(\underline{x}_j^*) d_j^2 \sigma_{jj}\right)^{1/2} b_1 d\underline{x}}{V_2 - V_1} + \frac{B_1}{V_2 - V_1}\right]$$

$$\approx (V_2 - V_1)[w^2 + 2(B_1/(V_2 - V_1))^{1/2} w + B_1/(V_2 - V_1)]$$

$$\approx (V_2 - V_1)[w + (B_1/(V_2 - V_1))^{1/2}]^2 \qquad (4.2.22)$$

$$\approx (V_2 - V_1)\, \chi'^2_{1,[B_1/(V_2 - V_1)]^{1/2}} . \qquad (4.2.23)$$

Using (4.2.23) and (4.2.18), the ratio of independent chi-square variates in (4.2.9) becomes

$$F_1 = \frac{\hat{B}_1/(V_2 - V_1)}{\sum_{j=1}^{p} a_{jj} d_j^2 \hat{\sigma}_{jj}/(V_2 - V_1)}$$

$$\approx \frac{\chi'^2_{1,[B_1/(V_2 - V_1)]^{1/2}}}{\chi^2_{\hat{h}_o}/\hat{h}_o}$$

$$\approx F'_{1,\hat{h}_o,[B_1/(V_2 - V_1)]^{1/2}} . \qquad (4.2.24)$$

The form of (4.2.22) is similar to that of (3.3.14) so that proceeding as in section 3.3, we have

$$1 - P \doteq \int_0^\infty \left[\frac{1}{\sqrt{2\pi}} \int_{-(D_\alpha u/\hat{h}_o)^{1/2} - (B_1/(V_2-V_1))^{1/2}}^{(D_\alpha u/\hat{h}_o)^{1/2} - (B_1/(V_2-V_1))^{1/2}} e^{-z^2/2}\, dz \right]$$

$$\frac{1}{2^{\hat{h}_o/2}\,\Gamma(\hat{h}_o/2)}\, u^{(\hat{h}_o/2)-1}\, e^{-u/2}\, du\ . \tag{4.2.25}$$

Under the hypothesis of (4.1.10), the substitution $B_1/(V_2-V_1) = 1$ enables us to determine D_α for specified $P = \alpha$. Also, for b_1 and $V_2 - V_1$ as given by (4.2.6) and (4.2.7), equation (4.2.25) can be employed to determine type II error probabilities for various values of the parameter $B_1/(V_2-V_1)$.

We now investigate the integrated mean squared error of our preliminary test estimator for $\underset{\sim}{\Sigma}$ diagonal unknown. The direct extensions of (3.4.2) and (3.4.3) are

$$\underline{\hat{\beta}_j^*} = \begin{bmatrix} \underline{\hat{\beta}_{1j}} - \underset{\sim}{A}_j \underline{\hat{\beta}_{2j}} \\ \underline{\hat{\beta}_{2j}} \end{bmatrix} \tag{4.2.26}$$

$$\underline{d}'\underline{\hat{y}_o} = \underline{d}'[\mathrm{diag}(\underline{x_{11}'}, \underline{x_{12}'}, \ldots, \underline{x_{1p}'})]\underline{\hat{\beta}_1} + \delta\underline{d}'[\mathrm{diag}(\underline{x_{21}'} - \underline{x_{11}'}\underset{\sim}{A}_1,$$

$$\underline{x_{22}'} - \underline{x_{12}'}\underset{\sim}{A}_2, \ldots, \underline{x_{2p}'} - \underline{x_{1p}'}\underset{\sim}{A}_p)]\underline{\hat{\beta}_2}\ . \tag{4.2.27}$$

Also, Lemma (3.4.1) holds for $\underline{\hat{\beta}_1}$ and $\underline{\hat{\beta}^*}$ of (4.2.1) and (4.2.2) with $\sigma_{ij} = 0$, $i \neq j$. Using the development (3.4.4) through (3.4.8), it is easy to show that

$$J_o = V_1 + (1-P)B_1 + (V_2-V_1)E(\delta Y)$$

where $\quad Y = NK \int_R (\hat{b}_1-b_1)^2 d\underline{x}/(V_2-V_1).$

Still utilizing section 3.4, if

$$J_{o3} = (V_2-V_1)E(\delta Y),$$

then $\quad \hat{J}_{o3} = (V_2-V_1)[(2\hat{y}_\alpha/\pi)^{1/2} e^{-\hat{y}_\alpha/2} + 2\Phi(-\sqrt{\hat{y}_\alpha})]$

where $\quad \hat{y}_\alpha = \dfrac{D_\alpha(\hat{V}_2-\hat{V}_1) - 2NK \int_R b_1\hat{b}_1 d\underline{x} + B_1}{V_2 - V_1}.$

Finally, $\quad \hat{J}_o = V_1 + (1-P)B_1 + \hat{J}_{o3} \qquad (3.4.14)$

where $V_2 - V_1$, P, and b_1 are obtained from (4.2.7), (4.2.25), and (4.2.6).

When the p response vectors are all functions of a single independent variable, we can generalize the results of section 3.5 so that

$$\underset{\sim}{X}_i^* = \begin{bmatrix} 1 & x_{i1} & x_{i1}^2-\overline{x_i^2} \\ 1 & x_{i2} & x_{i2}^2-\overline{x_i^2} \\ \vdots & \vdots & \vdots \\ 1 & x_{iN} & x_{iN}^2-\overline{x_i^2} \end{bmatrix}, \quad i = 1, 2, \ldots, p,$$

for $\overline{x_i^2} = \sum_{j=1}^{N} x_{ij}^2/N$, $\overline{x_i^3} = \sum_{j=1}^{N} x_{ij}^3/N$, $\overline{x_i^4}' = \sum_{j=1}^{N} (x_{ij}^2-\overline{x_i^2})^2/N$.

If we assume $\sum_{j=1}^{N} x_{ij} = 0$ for all i, then

$$\underset{\sim}{M}_j = \overline{x_j^2}/ND_j$$

where $D_j = (\overline{x_j^2})(\overline{x_j^4}') - (\overline{x_j^3})^2$,

and $\underset{\sim}{A}_j = \begin{bmatrix} 0 \\ \overline{x_j^3}/\overline{x_j^2} \end{bmatrix}$.

Scaling the R_j to the interval [-1,+1] enables us to write

$$K \int_R (\underset{\sim}{A}_j' \underline{x_{1j}} - \underline{x_{2j}})(\underline{x_{1j}'}\underset{\sim}{A}_j - \underline{x_{2j}'})d\underline{x} = [(\overline{x_j^2}-1/3)^2 + 4/45 + (\overline{x_j^3})^2/3(\overline{x_j^2})^2]$$

$$K \int_R (\underset{\sim}{A}_i' \underline{x_{1i}} - \underline{x_{2i}})(\underline{x_{1j}'}\underset{\sim}{A}_j - \underline{x_{2j}'})d\underline{x} = (\overline{x_i^2}-1/3)(\overline{x_j^2}-1/3) \quad , \; i \neq j$$

$$a_{jj} = \overline{x_j^2}\, [(\overline{x_j^2}-1/3)^2 + 4/45 + (\overline{x_j^3})^2/3(\overline{x_j^2})^2]/D_j \qquad (4.2.28)$$

$$B_1 = N[\sum_{j=1}^{p} d_j^2((\overline{x_j^2}-1/3)^2 + 4/45 + (\overline{x_j^3})^2/3(\overline{x_j^2})^2)\beta_{2j}^2$$

$$+ 2 \sum_{i<j}\sum d_i d_j (\overline{x_i^2}-1/3)(\overline{x_j^2}-1/3)\beta_{2i}\beta_{2j}] \qquad (4.2.29)$$

$$V_1 = \sum_{j=1}^{p} NK \int_R d_j^2[\underline{x_{1j}'}(\underset{\sim}{X}_{1j}'\underset{\sim}{X}_{1j})^{-1}\underline{x_{1j}}]\, \sigma_{jj}\, d\underline{x}$$

$$= \sum_{j=1}^{p} d_j^2(1 + 1/3\, \overline{x_j^2})\, \sigma_{jj} \qquad (4.2.30)$$

$$K \int_R b_1 \hat{b}_1 d\underline{x} = \sum_{j=1}^{p} d_j^2((\overline{x_j^2}-1/3)^2+4/45+(\overline{x_j^3})^2/3(\overline{x_j^2})^2)\beta_{2j}\hat{\beta}_{2j}$$

$$+ \sum_{i \neq j}\sum d_i d_j(\overline{x_i^2}-1/3)(\overline{x_j^2}-1/3)\beta_{2i}\hat{\beta}_{2j} \;. \qquad (4.2.31)$$

The noncentrality parameter $[B_1/(V_2-V_1)]^{1/2}$ for (4.2.25) is obtained using (4.2.28) and (4.2.29). To calculate $\hat{J}_{03}$, we substitute (4.2.31) in $\hat{y}_\alpha$; then $\hat{J}_o$ is given by (3.4.14) with (4.2.29) and (4.2.30).

Unlike the standard model, the distributional results obtained for $\hat{B}_1$ are not exact for the single independent variable case since the bound in (4.2.21) is not attained. Another dissimilarity from the case of the standard model single independent variable is that J_o does not lend itself to explicit evaluation, and we rely solely on $\hat{J}_o$.

4.3 A Special Case: $\underset{\sim}{\Sigma}$ Known

Knowledge of $\underset{\sim}{\Sigma}$ once more alleviates some of the difficulties inherent in our procedure. We write H_o as

$$B_1/NK \int_R \text{var}(\hat{b}_1)d\underline{x} \leq (V_2-V_1)/NK \int_R \text{var}(\hat{b}_1)d\underline{x} \qquad (4.3.1)$$

where, without assuming $\underset{\sim}{\Sigma}$ diagonal,

$$\hat{b}_1 = \underline{d}'[\text{diag}(\underline{x_{11}'},\underline{x_{12}'},\ldots,\underline{x_{1p}'})\underset{\sim}{A}_o-\text{diag}(\underline{x_{21}'},\underline{x_{22}'},\ldots,\underline{x_{2p}'})]\underline{\hat{\beta}_2} \qquad (4.3.2)$$

$$\text{var}(\hat{b}_1) = \underline{d}'[\text{diag}(\underline{x_{11}'},\underline{x_{12}'},\ldots,\underline{x_{1p}'})\underset{\sim}{A}_o-\text{diag}(\underline{x_{21}'},\underline{x_{22}'},\ldots,\underline{x_{2p}'})]\underset{\sim}{M}_o$$

$$[\underset{\sim}{A}_o'(\text{diag}(\underline{x_{11}},\underline{x_{12}},\ldots,\underline{x_{1p}}))-\text{diag}(\underline{x_{21}},\underline{x_{22}},\ldots,\underline{x_{2p}})]\underline{d} \qquad (4.3.3)$$

(from (4.1.16)).

Our test statistic is

$$C_1 = \hat{B}_1 / NK \int_R \mathrm{var}(\hat{b}_1) d\underline{x} \ . \tag{4.3.4}$$

Since $\hat{b}_1$ is unbiased, we have from (4.2.12) that

$$E(\hat{B}_1) = NK \int_R \mathrm{var}(\hat{b}_1) d\underline{x} + B_1,$$

and our procedure is

$$\text{reject } H_0 \text{ if } C_1 > 1 + (V_2 - V_1)/NK \int_R \mathrm{var}(\hat{b}_1) d\underline{x}$$

$$\text{accept } H_0 \text{ otherwise.}$$

For the error normality assumption, we first recall that $\mathrm{var}(\hat{b}_1) = \sum_{j=1}^{p} b(\underline{x}_j^*) d_j^{\,2} \sigma_{jj}$ if $\underset{\sim}{\Sigma}$ is diagonal. Thus, we can make use of the development leading to (4.2.21) to write

$$\hat{B}_1 \sim w^2 NK \int_R \mathrm{var}(\hat{b}_1) d\underline{x} + 2wNK \int_R [\mathrm{var}(\hat{b}_1)]^{1/2} b_1 d\underline{x} + B_1$$

$$\frac{NK \int_R [\mathrm{var}(\hat{b}_1)]^{1/2} b_1 d\underline{x}}{NK \int_R \mathrm{var}(\hat{b}_1) d\underline{x}} \leq \left[\frac{B_1}{NK \int_R \mathrm{var}(\hat{b}_1) d\underline{x}} \right]^{1/2} .$$

Generalizing (4.2.22) yields

$$\hat{B}_1 \approx (NK \int_R \mathrm{var}(\hat{b}_1) d\underline{x}) [w + (B_1 / NK \int_R \mathrm{var}(\hat{b}_1) d\underline{x})^{1/2}]^2$$

$$C_1 \approx \chi'^{\,2}_{1, [B_1/NK \int_R \mathrm{var}(\hat{b}_1) d\underline{x}]^{1/2}} \ .$$

Therefore, $1 - P = \Pr(C_1 \leq D_\alpha)$

$$\doteq \Phi[D_\alpha^{1/2} - (B_1/NK \int_R \mathrm{var}(\hat{b}_1)d\underline{x})^{1/2}]$$

$$- \Phi[-D_\alpha^{1/2} - (B_1/NK \int_R \mathrm{var}(\hat{b}_1)d\underline{x})^{1/2}] \qquad (4.3.5)$$

where B_1 and $\mathrm{var}(\hat{b}_1)$ are obtained from (4.1.13) and (4.3.3). Due to the complexity of our estimators (4.1.6) and (4.1.7), it is not feasible to develop a general expression for J_0 when $\underset{\sim}{\Sigma}$ is not diagonal even if it is known and we are dealing with single independent variables.

We shall briefly consider the simplest of all special cases, that of $\underset{\sim}{\Sigma}$ both diagonal and known. Now,

$$NK \int_R \mathrm{var}(\hat{b}_1)d\underline{x} = V_2 - V_1$$

where $V_2 - V_1$ is given by (4.2.7). Thus, our test statistic is simply

$$C_2 = \hat{B}_1/(V_2 - V_1) \;, \qquad (4.3.6)$$

for which we

reject H_0 if $C_2 > 2$

accept H_0 otherwise.

From (4.3.5),

$$1 - P \doteq \Phi[D_\alpha^{1/2} - (B_1/(V_2-V_1))^{1/2}] - \Phi[-D_\alpha^{1/2} - (B_1/(V_2-V_1))^{1/2}] \;. \qquad (4.3.7)$$

Generalizing (3.6.3) gives

$$\hat{\tilde{y}}_\alpha = \frac{D_\alpha - 2NK \int_R b_1 \hat{b}_1 d\underline{x} + B_1}{V_2 - V_1},$$

$$\text{and} \quad \hat{J}_{o3} = (V_2 - V_1)[(2\hat{\tilde{y}}_\alpha/\pi)^{1/2} e^{-\hat{\tilde{y}}_\alpha/2} + 2\Phi(-\sqrt{\hat{\tilde{y}}_\alpha})]$$

with $\hat{J}_o$ as in (3.4.14). Results for the case of the single independent variable are obtained using (4.2.28) through (4.2.31).

Design considerations are extremely difficult to treat for the generalized model even for the most restrictive assumptions on the error covariance. Combinations of design variables which increase power or decrease B_o, V_o, or J_o, seem best sought by empirical methods.

Chapter V

COMPARISON OF INTEGRATED MEAN SQUARED ERRORS

A variety of means by which to choose a model are available to the researcher. He may arbitrarily select $\hat{\underline{y}}$ or $\hat{\hat{\underline{y}}}$ having integrated mean squared errors J_1 and J_2 respectively. Another possibility is that of choosing a model by using a preliminary test estimation procedure based upon the usual statistic F_c given in (3.2.26). The resulting estimator for a multivariate model has integrated mean squared error J_3, say. We shall compare the performance of $\underline{\hat{y}_0}$ and the above estimators with respect to J_0, J_1, J_2, and J_3. We shall also discuss a reasonable range of α levels for the estimators structured around a preliminary test of hypothesis.

The subsequent graphs have been prepared utilizing the design in the example of section 3.6 for the case of the standard multivariate regression model, single independent variable, $\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}$ (known) = 1. Critical points for F_c were obtained using a similar procedure to (3.6.2) since

$$F_c = F_0/a_1 \tag{3.2.28}$$

for $q_2 = 1$. Due to the computational effort required, this is not intended as an exhaustive comparative study. Rather we are examining the special case of the single independent variable with $\underline{d}'\underset{\sim}{\Sigma}\,\underline{d}$ known as an indication of what is expected in more general cases. The symbol J in Figures 5.1.1 - 5.1.3 denotes integrated mean squared error with α values affecting only J_0 and J_3.

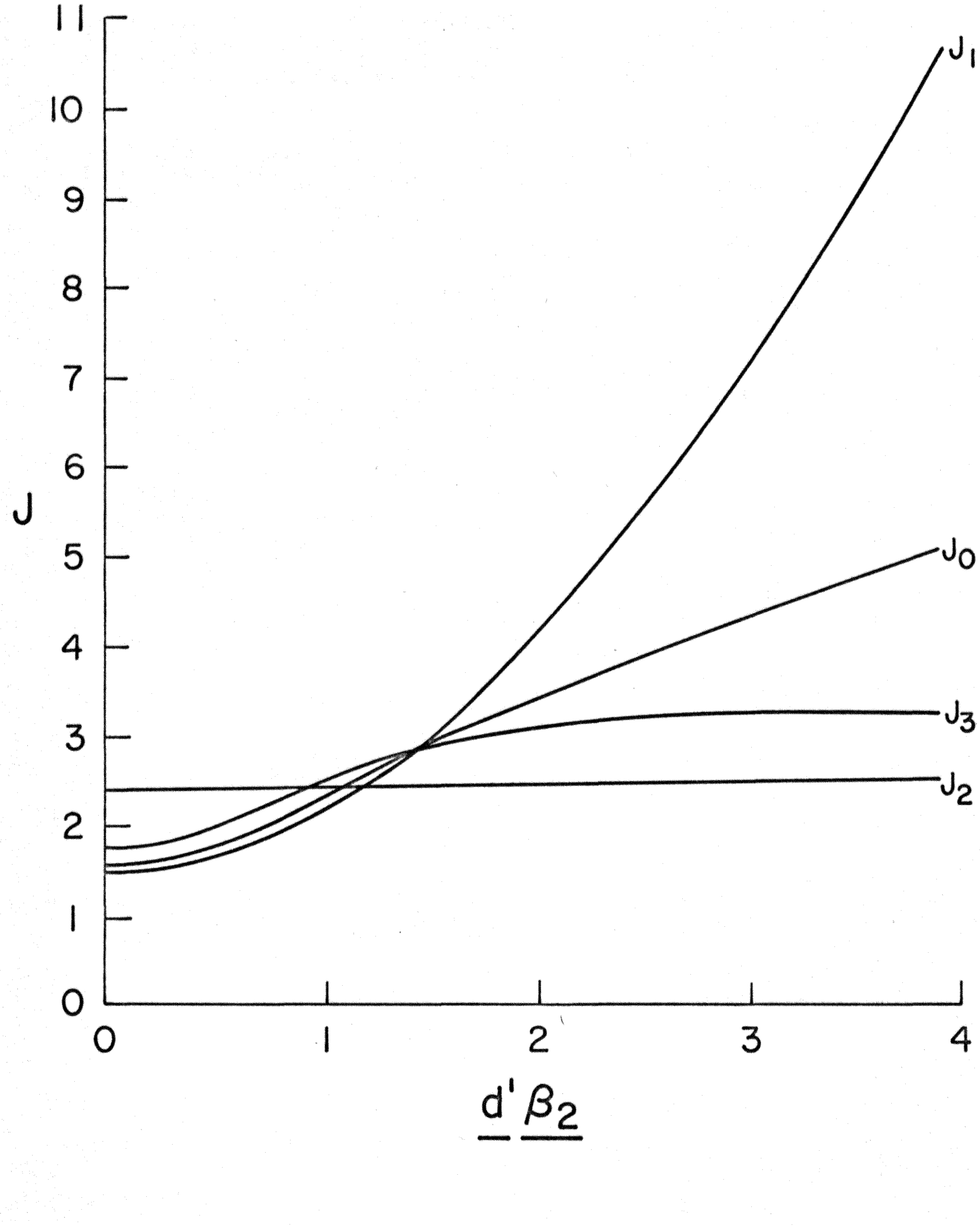

Figure 5.1.1 J Values (α = .05)

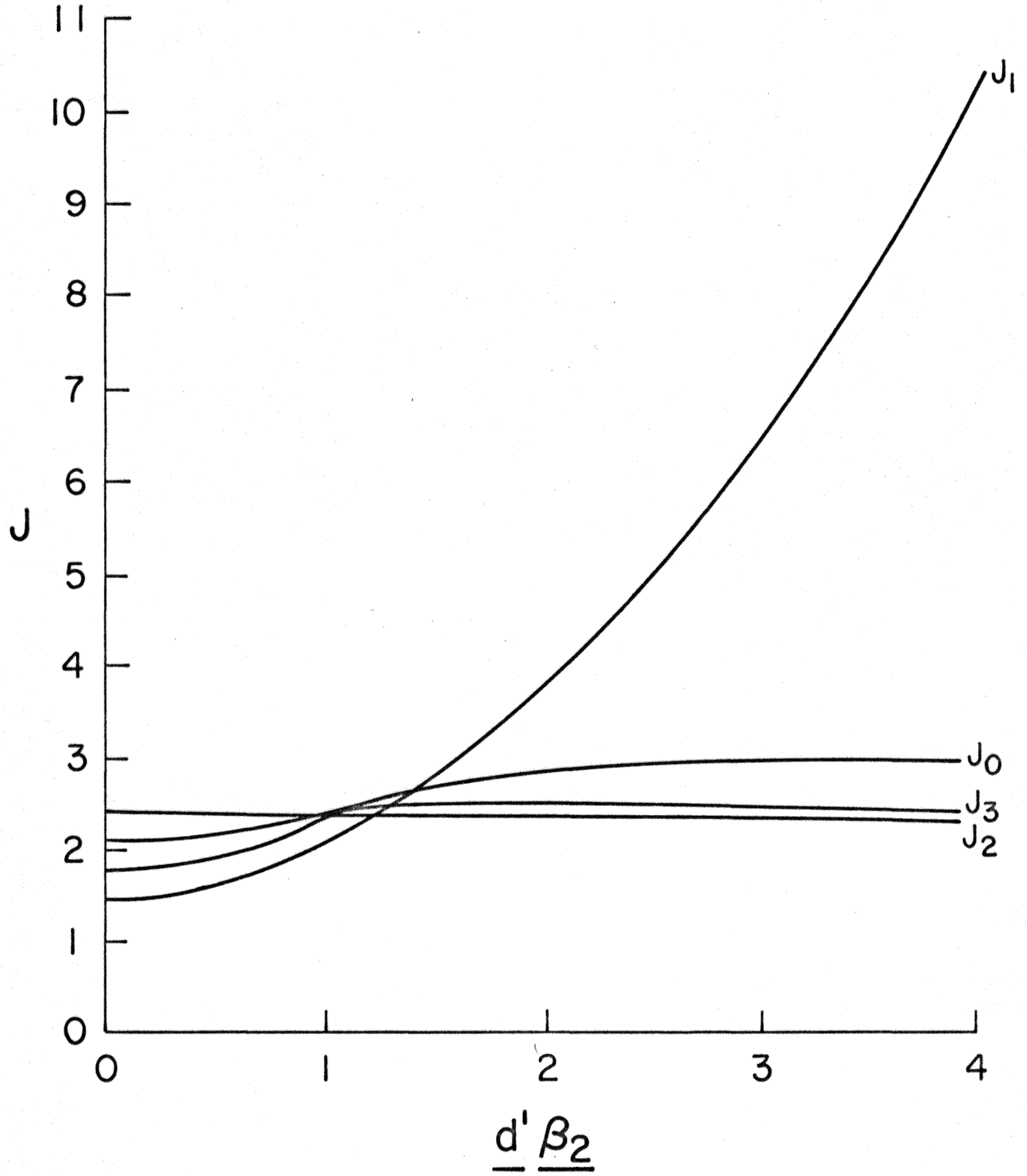

Figure 5.1.2 J Values (α = .18)

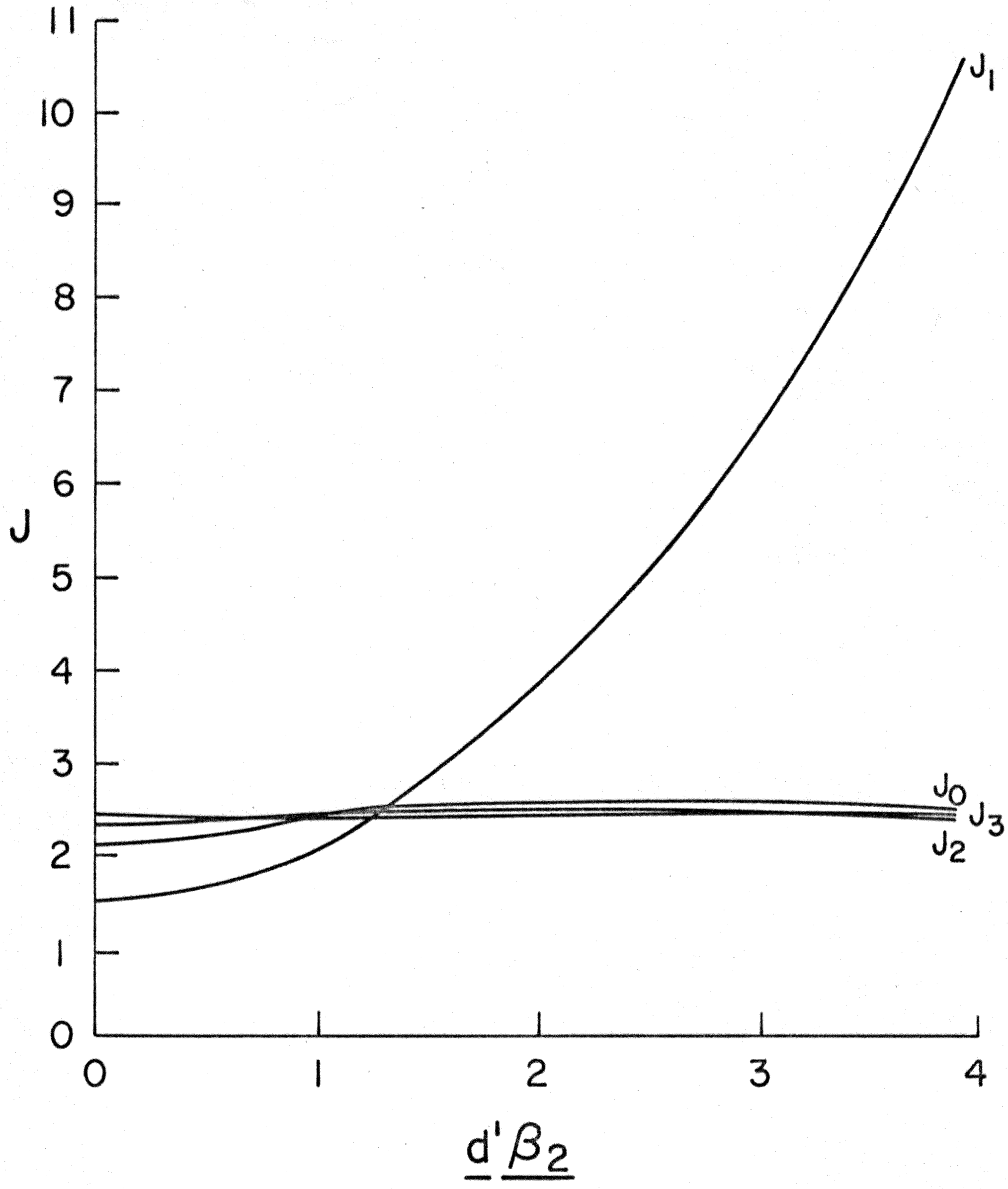

Figure 5.1.3 J Values (α = .50)

In general, it is to be expected that J_3 more closely resembles J_2 than does J_0 since the standard procedure is testing the hypothesis $a_3/a_1 = 0$ whereas J_0 is based on H_0: $a_3/a_1 \leq 1$. Since P' is an increasing function of the noncentrality parameter $(a_3/a_1)^{1/2}$, the classical procedure yields a lower critical value and rejects more often. However for $q_2 > 1$, we recall from (3.2.33) that F_c is unsuited for testing H_0 from the standpoint of a ratio of expected values.

The reason for the selection of $\alpha = .18$ as a tabular entry is illustrated by Figure 5.1.4. The graph of J_0 for $\alpha = .18$ seems to provide a reasonable compromise between the two extremes of Figure 5.1.4. While we may be unwilling to accept values of J_0 as great as those for $\alpha = .05$ and large $\underline{d}'\underline{\beta_2}$, we may also wish to discriminate more against J_2 than by the use of $\alpha = .50$. Of course, the range of α may be adjusted against the values of the parameter $\underline{d}'\underline{\beta_2}$ for which one wishes to obtain protection.

For the standard model with $q_2 = 1$, we can plot values of the integrated mean squared error for the preliminary test estimators exactly. If $q_2 > 1$ or we are dealing with the generalized model, then (3.4.14) can be employed for an estimate of J_0. Using Figures 5.1.1 through 5.1.4 as an indication, we conclude that ranges of α greater than the traditional testing values of .01, .05, and .10 seem best suited to preliminary test estimation in general. Although our α values for $q_2 > 1$ are not exact as indicated by Table 3.3.1, we are essentially interested in establishing a viable range of α's on which to base our estimators, not on the type I error probabilities themselves. Of

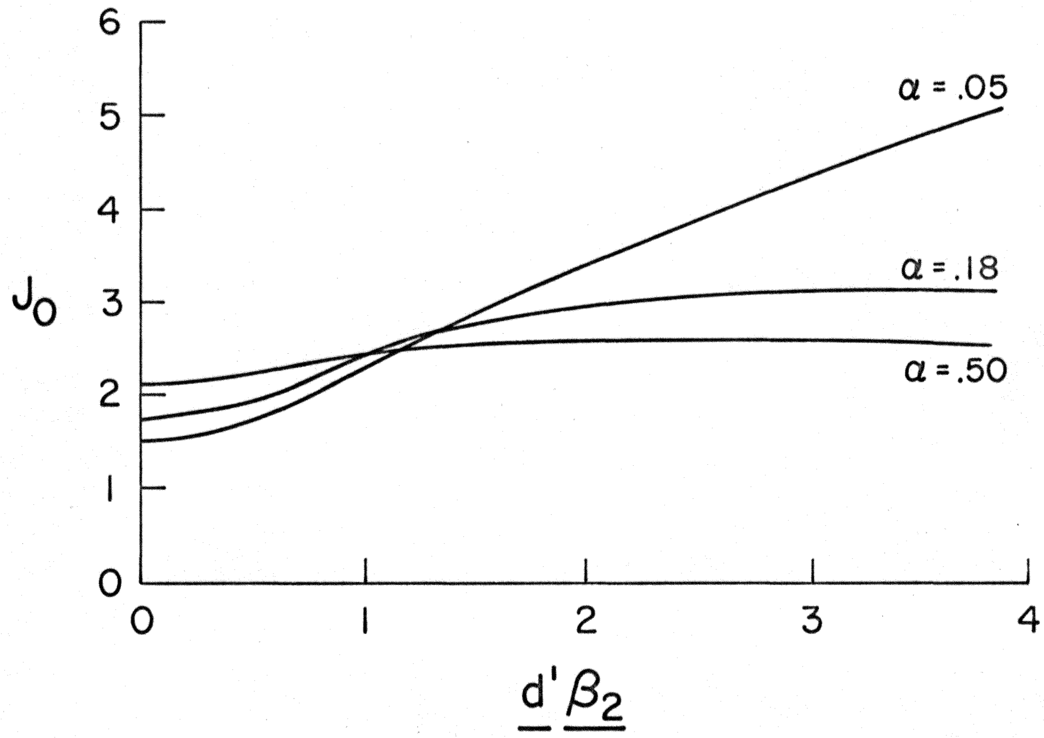

Figure 5.1.4 J_0 Values

further interest would be an extensive numerical investigation into the various models and special cases presented in the preceding chapters.

APPENDIX I

Proof of (3.5.15): From (3.5.14),

$$E(\delta w) = (1/\sqrt{2\pi}) \int_0^\infty [\int_{-R_H|u}^{-R_L|u} ze^{-z^2/2}\,dz]\,f(u)du$$

$$= (1/\sqrt{2\pi}) \int_0^\infty [\int_{(R_H|u)^2/2}^{(R_L|u)^2/2} e^{-z_1}\,dz_1]\,f(u)\,du$$

$(z_1 = z^2/2)$

$$= (1/\sqrt{2\pi}) \int_0^\infty [e^{-(R_H|u)^2/2} - e^{-(R_L|u)^2/2}]\,f(u)du\,. \tag{A.1}$$

$$\text{Let}\quad g = D_\alpha/a_1(N-3),\quad h = (a_3/a_1)^{1/2}, \tag{A.2}$$

so that $R_H|u = \sqrt{gu} - h$, $R_L|u = -\sqrt{gu} - h$,

$$e^{-(R_H|u)^2/2} - e^{-(R_L|u)^2/2} = e^{-h^2/2}(e^{-gu/2+h\sqrt{gu}} - e^{-gu/2-h\sqrt{gu}})\,.$$

Using (3.3.2) with $\nu = N - q_0$ gives

$$E(\delta w) = \frac{e^{-h^2/2}/\sqrt{2\pi}}{2^{\nu/2}\Gamma(\nu/2)} [\int_0^\infty u^{(\nu/2)-1}e^{-(g+1)u/2+h\sqrt{gu}}du - \int_0^\infty u^{(\nu/2)-1}$$

$$e^{-(g+1)u/2-h\sqrt{gu}}\,du]$$

$$= \frac{e^{-h^2/2}/\sqrt{2\pi}}{2^{(\nu-2)/2}\Gamma(\nu/2)} [\int_0^\infty u_1^{\nu-1}e^{-(g+1)u_1^2/2+h\sqrt{g}\,u_1}du_1 - \int_0^\infty u_1^{\nu-1}$$

$$e^{-(g+1)u_1{}^2/2-h\sqrt{g}\,u_1}\,du_1]$$

$(u_1 = u^{1/2})$

$$= \frac{e^{-h^2/2} e^{gh^2/2(g+1)}/\sqrt{2\pi}}{2^{(\nu-2)/2}\Gamma(\nu/2)} [\int_0^\infty u_1^{\nu-1} e^{-(\frac{g+1}{2})(u_1 - h\sqrt{g}/(g+1)^2} du_1$$

$$- \int_0^\infty u_1^{\nu-1} e^{-(\frac{g+1}{2})(u_1 + h\sqrt{g}/(g+1)^2} du_1]$$

$$= \frac{e^{-h^2/2(g+1)}/\sqrt{2\pi}}{2^{(\nu-2)/2}\Gamma(\nu/2)\sqrt{g+1}} [\int_{-h\sqrt{\frac{g}{g+1}}}^{\infty} (u_2/\sqrt{g+1} + h\sqrt{g}/(g+1)^{\nu-1}\ e^{-u_2^2/2} du_2$$

$$- \int_{h\sqrt{\frac{g}{g+1}}}^{\infty} (u_3/\sqrt{g+1} - h\sqrt{g}/(g+1))^{\nu-1}\ e^{-u_3^2/2} du_3]$$

$(u_2 = \sqrt{g+1}[u_1 - h\sqrt{g}/(g+1)]$ and $u_3 = \sqrt{g+1}[u_1 + h\sqrt{g}/(g+1)])$

$$= \frac{e^{-h^2/2(g+1)}/\sqrt{2\pi}}{2^{(\nu-2)/2}\Gamma(\nu/2)\sqrt{g+1}} [\int_{-h\sqrt{\frac{g}{g+1}}}^{h\sqrt{\frac{g}{g+1}}} (u_2/\sqrt{g+1} + h\sqrt{g}/(g+1))^{\nu-1} e^{-u_2^2/2} du_2$$

$$+ \int_{h\sqrt{\frac{g}{g+1}}}^{\infty} (u_2/\sqrt{g+1} + h\sqrt{g}/(g+1))^{\nu-1} e^{-u_2^2/2} du_2$$

$$- \int_{h\sqrt{\frac{g}{g+1}}}^{\infty} (u_3/\sqrt{g+1} - h\sqrt{g}/(g+1))^{\nu-1} e^{-u_3^2/2} du_3] . \qquad (A.3)$$

Employing the binomial expansion and cancelling terms in the last two integrals of (A.3) yields

$$E(\delta w) = \frac{e^{-h^2/2(g+1)}/\sqrt{2\pi}}{2^{(\nu-2)/2}\Gamma(\nu/2)\sqrt{g+1}} \Big[\int_{-h\sqrt{\frac{g}{g+1}}}^{h\sqrt{\frac{g}{g+1}}} \sum_{\theta=0}^{\nu-1} \binom{\nu-1}{\theta}(u_2/\sqrt{g+1})^{\nu-1-\theta}$$

$$(h\sqrt{g}/(g+1))^{\theta} e^{-u_2^2/2} du_2 + 2\int_{h\sqrt{\frac{g}{g+1}}}^{\infty} \sum_{\phi=0}^{[\frac{\nu-2}{2}]} \binom{\nu-1}{2\phi+1}(u_3/\sqrt{g+1})^{\nu-2-2\phi}$$

$$(h\sqrt{g}/(g+1))^{2\phi+1} e^{-u_3^2/2} du_3\Big] \tag{A.4}$$

where $[\frac{\nu-2}{2}] = [\frac{N-5}{2}]$ denotes the largest integer less than or equal to $(N-5)/2$. Similar to (3.3.2) for the single independent variable, (A.4) holds for $N \geq 4$ if we define the summation occurring under the second integral to be identically zero for $N = 4$. Simplifying gives

$$E(\delta w) = \frac{e^{-h^2/2(g+1)}/\sqrt{2\pi}}{2^{(\nu-2)/2}\Gamma(\nu/2)\sqrt{g+1}} \Big[\sum_{\theta=0}^{\nu-1} \binom{\nu-1}{\theta} \frac{(h\sqrt{g})^{\theta}}{(\sqrt{g+1})^{\nu+\theta-1}}$$

$$\int_{-h\sqrt{\frac{g}{g+1}}}^{h\sqrt{\frac{g}{g+1}}} u_2^{\nu-1-\theta} e^{-u_2^2/2} du_2 + 2\sum_{\phi=0}^{[\frac{\nu-2}{2}]} \binom{\nu-1}{2\phi+1} \frac{(h\sqrt{g})^{2\phi+1}}{(\sqrt{g+1})^{\nu+2\phi}}$$

$$\int_{h\sqrt{\frac{g}{g+1}}}^{\infty} u_3^{\nu-2-2\phi} e^{-u_3^2/2} du_3\Big] . \tag{3.5.15}$$

BIBLIOGRAPHY

Anderson, T. W. (1958), An Introduction to Multivariate Statistical Analysis. John Wiley and Sons, Inc., New York.

Bancroft, T. A. (1944), On Biases in Estimation Due to the Use of Preliminary Tests of Significance. Annals of Mathematical Statistics, 15: 190-204.

________ (1950), Bias Due to the Omission of Independent Variables in Ordinary Multiple Regression Analysis. Unpublished preliminary report, Iowa State University, Ames.

Ellerton, R. R. W. (1973), Preliminary Test Estimation in Response Surface Methodology. Unpublished Ph.D. thesis, Virginia Polytechnic Institute and State University, Blacksburg.

Folks, J. L. (1958), Comparison of Designs for Exploration of Response Relationships. Unpublished Ph.D. thesis, Iowa State University, Ames.

Graybill, F. A. (1961), An Introduction to Linear Statistical Models. McGraw-Hill Book Company, Inc., New York.

Kennedy, W. J., and Bancroft, T. A. (1971), Model Building for Prediction in Regression Based upon Repeated Significance Tests. Annals of Mathematical Statistics, 42: 1273-1284.

Larson, H. J., and Bancroft, T. A. (1963a), Sequential Model Building for Prediction in Regression Analysis, I. Annals of Mathematical Statistics, 34: 462-479.

________ (1963b), Biases in Prediction by Regression for Certain Incompletely Specified Models. Biometrika, 50: 391-402.

Lehmann, E. L. (1959), Testing Statistical Hypotheses. John Wiley and Sons, Inc., New York.

Myers, R. H. (1971), Response Surface Methodology. Allyn and Bacon, Inc., Boston.

Press, S. J. (1972), Applied Multivariate Analysis. Holt, Rinehart and Winston, Inc., New York.

Rao, C. R. (1965), Linear Statistical Inference and Its Applications. John Wiley and Sons, Inc., New York.

Satterthwaite, F. E. (1946), An Approximate Distribution of Estimates of Variance Components. Biometrics Bulletin, 2: 110-114.

Scheffé, H. (1959), The Analysis of Variance. John Wiley and Sons, Inc., New York.

Toro-Vizcarrondo, C., and Wallace, T. D. (1968), A Test of the Mean Square Error Criterion for Restrictions in Linear Regression. Journal of the American Statistical Association, 63: 558-572.

The vita has been removed from the scanned document

A PRELIMINARY TEST ESTIMATOR

FOR MULTIVARIATE RESPONSE FUNCTIONS

by

Paul West Blackmon, Jr.

(ABSTRACT)

If $\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_p$ represent vectors of independent observations, the generalized multivariate regression model is of the form

$$\underline{y}_j = X_{1j}\underline{\beta}_{1j} + X_{2j}\underline{\beta}_{2j} + \underline{\varepsilon}_j \quad , \; j = 1, 2, \ldots, p,$$

where X_{1j} and X_{2j} are general linear model regression matrices, $\underline{\beta}_{1j}$ and $\underline{\beta}_{2j}$ are vectors of unknown coefficients, and the $\underline{\varepsilon}_j$ are error vectors such that $\text{cov}(\underline{\varepsilon}_i, \underline{\varepsilon}_j) = \sigma_{ij}I$. If $X_{1j} = X_1$ and $X_{2j} = X_2$, $j = 1, 2, \ldots, p$, the above is a standard multivariate regression model.

Insofar as can be determined, the true relationship between the design variables and a response η_j is

$$\eta_j = \underline{x}_{1j}' \, \underline{\beta}_{1j} + \underline{x}_{2j}' \, \underline{\beta}_{2j}$$

where $\underline{x}_{1j}'$ and $\underline{x}_{2j}'$ are typical row vectors in the matrices X_{1j} and X_{2j}. For $\underline{x}_j^{*\prime} = [\underline{x}_{1j}', \underline{x}_{2j}']$ and $\underline{\beta}_j^{*\prime} = [\underline{\beta}_{1j}', \underline{\beta}_{2j}']$, the η_j are to be estimated either by $\hat{y}_j = \underline{x}_{1j}' \, \hat{\underline{\beta}}_{1j}$ or $\hat{\hat{y}}_j = \underline{x}_j^{*\prime}\hat{\underline{\beta}}_j^{*}$ where $\hat{\underline{\beta}}_{1j}$ and $\hat{\underline{\beta}}_j^{*}$ are the least squares estimators of $\underline{\beta}_{1j}$ and $\underline{\beta}_j^{*}$, obtained from the full multivariate regression model.

The estimators for the η_j are determined by a test of the hypothesis H_0: $J_1 \leq J_2$ where J_1 and J_2 denote the integrated mean

squared errors of a linear combination of the $\hat{y}_j$ and $\hat{\hat{y}}_j$ respectively. Rejection of H_0 results in selection of the $\hat{\hat{y}}_j$; otherwise the $\hat{y}_j$ are chosen.

A test statistic is developed to test H_0 with consideration extending to several important special cases. Distinctions are drawn between the preliminary test estimator constructed around H_0, and that based on the usual hypothesis $\underline{\beta_{2j}} = \underline{0}$, j = 1, 2, ..., p.

Under the assumption of error normality, an approximation to the distribution of the test statistic is developed in order to determine type I and type II error probabilities.

An explicit expression for J_0, the integrated mean squared error of the preliminary test estimator, is obtained, and difficulties in its evaluation are discussed. An estimator of J_0 is presented along with a special case in which J_0 can be evaluated exactly.

Graphical comparisons are made on the relative performance of the estimators based on H_0, and those constructed around the standard hypothesis. An operating range of type I error probabilities is also discussed.