WHEN THE COEFFICIENT OF A COVARIATE CHANGES FROM BLOCK TO BLOCK

by

Elinor S. Pape

Technical Report No. 108
Department of Statistics ONR Contract

November 18, 1971

Research sponsored by the Office of Naval Research
Contract N00014-68-A-0515
Project NR 042-260

DEPARTMENT OF STATISTICS
Southern Methodist University

CHAPTER I

INTRODUCTION

The one and two-way classification models with a covariate have been treated from many points of view. However the question of the covariate regression parameter changing from block to block has not been treated so extensively. That such changes might occur is certainly logical. Granting that "yield" might increase linearly with, say, temperature, it would not be at all surprising to find certain blocks, perhaps materials, to be more responsive to temperature than other blocks.

In the analysis of variance a single regression effect "costs" a single degree of freedom. If the regression effects change from block to block it will be necessary to estimate, or adjust for, as many regression effects as there are blocks, say r. Thus the additional "cost" of block regression effect differences is r-1 degrees of freedom.

As a tool in evaluating the adequacy of a one-way classification model with a covariate, Robson and Atkinson [7] propose an individual degree of freedom test for the homogeniety of regression coefficients. The same thing can be done in the two-way model. And in a two way model with individual regression coefficients in the blocks a single degree of freedom test can be constructed to test for regression coefficient differences among the treatments. An easy way to construct these tests is to mimic Scheffe's motivation [8] of Tukey's single degree of freedom for non-additivity [9]. If the model under consideration is $y_{ij} = \mu + \theta_i + \nu_j + \psi x_{ij} + \varepsilon_{ij}$ and one is concerned that the coefficient ψ might be changing from block to block, the procedure is to insert a non-linear term of the sort $\phi\nu_j x_{ij}$ into the model. Then this expanded model is approximated by the linear model $y_{ij} = \mu + \theta_i + \nu_j + \psi x_{ij} + \phi\hat{\nu}_j x_{ij} + \varepsilon_{ij}$ where $\hat{\nu}_j$ is the standard least squares estimate of ν_j under the original model. It is then the sum of squares for the dummy coefficient ϕ adjusted for μ, θ, ν, and ψ which is used to test the adequacy of the original model. The fact that the test is neither exact nor powerful does not keep it from being very useful in the field of model building.

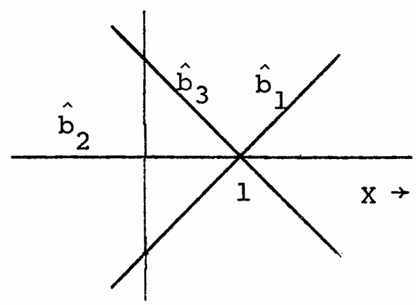
Granting that one has adopted a model with individual regression coefficients for the blocks and fixed effects for the treatments, one then finds the literature preoccupied with eliminating the effects of blocks and covariates in order to focus attention on the treatments. Much of this work has been done by C. P. Cox [1]. Noteworthy also is Zelen's adaptation of this problem to incomplete block designs [10].

Traditionally a covariate has been considered to be a nuisance factor to be eliminated. Perhaps the covariate has been the victim of an unfriendly press. If the covariate takes on a limited number of discrete values as in Cox's first, second, and third units of time [1] the covariate model is actually a shortcut to a factorial analysis in which all effects of the covariate factor other than linear effects are confounded. These effects may be of great interest. Even when the covariate is uncontrollable, it is quite possible that it is known. Of course if all experimental units show the same response to the covariate, that is if $y_{ij} = \mu + \theta_i + \nu_j + \psi x_{ij} + \varepsilon_{ij}$ then if high yields can be considered to be good, it takes no great mathematical calculations to say "the higher the X the better" or "the lower the X the better," and likewise, "the higher (or lower) the ν_j the better." But when the regression coefficient depends on the block, the value of the covariate can become quite important.

Consider three blocks with fixed effects (taken to sum to 0) of -1, 0, and +1. If their corresponding block regression effects (also taken to sum to 0) are +1, 0, and -1 (that is, $b_1 = -1 + X$, $b_2 = 0 + 0X = 0$, $b_3 = 1 - X$), then the question "Which block gives the highest yield?" can be answered only if we know the value of X. Clearly, if the covariate is less than 1, the third block will be preferred and if the covariate is more than 1, the first block will be preferable. By taking the block regression effects to sum to 0 the possibility of a large general regression effect with the covariate has not been eliminated. That is, in the model $y_{ij} = \mu + \theta_i + \nu_j + \psi x_{ij} + \phi_j x_{ij} + \varepsilon_{ij}$ the value of ψ may be very much larger than the values of the ϕ_j's. Then we would say "the higher the X the better--but if X happens to be less than 1, try to use block three and if X happens to be greater than 1, try to use block one." It is not difficult to see the usefulness of this sort of information which is based here on some insight into the "true" parameters of the model. In our real world, of course, we will have to estimate these parameters and we shall be interested, not only in which estimated block effect is larger at any X, but in whether these block effects are significantly different at any particular value of X in light of the normality assumption made for the ε_{ij}'s.

Our interest in a significance level arises from an assumption of a loss function associated with the choice of block. If blocks represent the type of material it is logical to assume that some types are more expensive than others. Even if they are all of the same price, keeping the materials separated might be expensive and should not be undertaken unless we are reasonably sure that the blocks are really different at some, if not all, values of the covariate X (perhaps the temperature or humidity). This thesis does not direct itself to any specific loss function; hence, it is left to the practitioner to bridge the gap between the loss function inherent in the application and the significance level of the test.

Returning to the previous example, consider the block and regression effects given above to be estimates rather than known parameters. That is, let $\hat{b}_1 = -1 + X$, $\hat{b}_2 = 0$, and $\hat{b}_3 = 1 - X$.



Now the question to be answered is "for what values of X are blocks significantly different?" Surely the point X = +1 will not be one of these points, for at that point the block estimates are exactly equal. We shall expect, intuitively, that the answer to the above question will be the real line with an interval about the point X = 1 deleted. And if that interval included the whole real line the answer would be that the blocks are nowhere significantly different. The higher the significance level, implying that there is a greater penalty for wrongly reporting block differences, the larger we shall expect that interval to be.

The fact that the most conventionally constructed test for block differences is very capable of answering the above question, not with one interval, but with two disjoint intervals (by saying that blocks are different everywhere along the real line except within those two disjoint intervals) is perhaps the most interesting aspect of this thesis.

Knowing the values for X at which blocks are significantly different enables one to know when it is worthwhile to be particular about blocks. The estimates of the block effects tell which block to favor at any given value of X. No consideration is given in this work to the problem of which blocks cause the difference when difference results.

A test is sought, then, for a null hypothesis of "no difference between blocks when $X < x_0$" under the assumption of the model

$$y_{ij} = \mu + \theta_i + \psi x_{ij} + \phi_j x_{ij} + \varepsilon_{ij}$$

$$i = 1, 2, \cdots , t; \; j = 1, 2, \cdots , r; \; \varepsilon_{ij} \sim \mathrm{NID}(0, \sigma^2) \quad .$$

The μ, θ_i, ν_j, ψ, and ϕ_j are unknown constants. μ represents a fixed mean effect, θ_i represents a fixed treatment effect, ν_j a fixed block effect, ψ a mean regression coefficient, and ϕ_j a block regression coefficient. In the manner of other two-way models these effects are not completely estimable and we will need to place three conditions on them. The usual condition $\sum_i \theta_i = 0$ is made to make the estimates of the θ_i unique. But two more conditions must be taken to discern the $\nu_j + \phi_j X$ effects from the fixed mean effect μ and the regression mean effect ψX. As usual we have a great deal of choice regarding these conditions. The criterion is generally mathematical neatness, but in this problem the conditions which result in mathematical neatness complicate the parametric statement of the null hypothesis. The conditions we shall use are $\sum_j (x_0 - x_{\cdot j})\phi_j = 0$ and $\sum_j (\nu_j + \phi_j x_{\cdot j} + \psi x_{\cdot j}) = 0$ where $x_{\cdot j} = \frac{1}{t}\sum_i x_{ij}$. Note that when $x_{\cdot j} = x_{\cdot j'}$ for all j and j' the first of these conditions simplifies to $\sum_j \phi_j = 0$ and consequently the second simplifies to $\sum_j (\nu_j + \psi x_{\cdot j}) = 0$. By defining $\kappa_j = \nu_j + \psi x_{\cdot j}$ the second condition further simplifies to $\sum_j \kappa_j = 0$.

This special case, $x_{\cdot j} = x_{\cdot j'}$ for all j and j', shall be presented fully, not with the thought that this case will occur frequently in practice, but rather to illuminate various theoretical aspects of the problem. After full consideration of this special case it will be shown that the general case follows by appropriate introduction of matrix multipliers.

It should not cause concern that in the general case the conditions used to solve the normal equations depend on x_0 . The null hypothesis to be tested also depends on x_0, i.e., H_o: $\nu_j + x_0\phi_j = 0$ for all j versus H_a: $\nu_j + x_0\phi_j \neq 0$.

When a statistic is found to test the hypothesis at $X = x_0$ that test will be inverted to yield the set of X's for which the test would be rejected. This test inversion will not yield a confidence set because the inversion does not extend to all alternate hypotheses. For example, it finds probable values of X if $\nu_j + X\beta_j = 0$ but not if $\nu_j + X\beta_j = 1$. A confidence set is of value when a parameter is unknown. In this problem the covariate, X, is known and we are interested instead in which X's will cause differences in the blocks. Hence, if x_p is in the set, then we have reason to believe that assignment of block will affect yield when $X = x_p$.

CHAPTER II

THE SPECIAL CASE OF EQUAL COVARIATE BLOCK MEANS

The least squares estimates of the block parameters are sought as a basis for a test of their difference. The experimental model under consideration can be written

$$y_{ij} = \mu + \theta_i + \kappa_j + \beta_j r_j a_{ij} + \psi r_j a_{ij} + \varepsilon_{ij}$$

$$i = 1, 2, \cdots, t;\ j = 1, 2, \cdots, r;\ \sum \kappa_j = 0;\ \sum \theta_i = 0;\ \sum \beta_j = 0$$

$$r_j = \sqrt{\sum_i (x_{ij} - x_{\cdot j})^2}\ ;\quad \sum_i a_{ij} = 0;\ \sum_i a_{ij}^2 = 1$$

$$a_{ij} = \frac{x_{ij} - x_{\cdot j}}{r_j}$$

where x_{ij} is the value of the covariate X in its original units and assuming, for this chapter, the condition

$$x_{\cdot j} = x_{\cdot j'}, \quad \text{for all } j, j'$$

$$x_{\cdot j} = \frac{1}{t} \sum_i x_{ij} \quad .$$

In referring to Chapter I, $\kappa_j = \nu_j + \psi x_{\cdot j}$ and $\beta_j = \phi_j$. The purpose of this reparameterization is to make $\hat{\kappa}_j$ and $\hat{\beta}_j$ independent. The purpose of factoring $(x_{ij} - x_{\cdot j})$ into r_j and a_{ij} is only to aid in analyzing sources of variation after a test is derived.

Let

$$A = [a_{ij}]_{t\times r} = (\underline{a}_1, \underline{a}_2, \underline{a}_3, \cdots, \underline{a}_r) \quad .$$

That is, the vector $\underline{a}_j$ has elements a_{ij} . Then let

$$A_d = \begin{bmatrix} \underline{a}_1 & & & \psi \\ & \underline{a}_2 & & \\ & & \ddots & \\ \phi & & & \underline{a}_r \end{bmatrix}_{rt\times r}$$

Let

$$R_{r\times r} = \text{diag}\,(r_j) \qquad \underline{\beta}_{r\times 1} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_r \end{bmatrix} \quad \underline{\kappa}_{r\times 1} = \begin{bmatrix} \kappa_1 \\ \kappa_2 \\ \cdot \\ \cdot \\ \cdot \\ \kappa_r \end{bmatrix} \quad \underline{\theta}_{t\times 1} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \cdot \\ \cdot \\ \cdot \\ \theta_r \end{bmatrix}$$

$$\sum_i \sum_j y_{ij} = rt\mu + \sum_i \sum_j \varepsilon_{ij}$$

$$\sum_j y_{ij} = r\mu + r\theta_i + \sum_j \beta_j a_{ij} r_j + \psi \sum_j a_{ij} r_j + \sum_j \varepsilon_{ij}$$

$$\sum_i y_{ij} = t\mu + t\kappa_j + \sum_i \varepsilon_{ij}$$

$$r_j \sum_i a_{ij} y_{ij} = r_j \sum_i a_{ij}\theta_i + r_j^2\beta_j + r_j^2\psi + r_j \sum_i a_{ij}\varepsilon_{ij}$$

Note that the conditions $\sum \theta_i = 0$ and $\sum \kappa_j = 0$ were used in forming the above relationships, but the third condition $\sum \beta_j = 0$ has not yet been used.

Let $\hat{\underline{\beta}}$, $\hat{\underline{\theta}}$, $\hat{\underline{\kappa}}$, $\hat{\psi}$, and $\hat{\mu}$ be the unbiased estimates of the parameters of the model retaining the conditions restricting the parameters. That is $\underline{1}'\hat{\underline{\beta}} = 0$, $\underline{1}'\hat{\underline{\theta}} = 0$, $\underline{1}'\hat{\underline{\kappa}} = 0$. Then the normal equations become

$$\frac{1}{rt} J\underline{y} = \hat{\mu}\underline{1}$$

$$\frac{1}{r}(I_t \otimes \underline{1}_r')\underline{y} = \hat{\mu}\underline{1} + \hat{\underline{\theta}} + \frac{1}{r} AR\hat{\underline{\beta}} + \hat{\psi}\frac{1}{r} AR\underline{1}$$

$$\frac{1}{t}(\underline{1}_t' \otimes I_r)\underline{y} = \hat{\mu}\underline{1} + \hat{\underline{\kappa}}$$

(1) $$RA_d'\underline{y} = RA'\hat{\underline{\theta}} + R^2\hat{\underline{\beta}} + R^2\hat{\psi}\underline{1} = RA'\hat{\underline{\theta}} + R^2[\hat{\underline{\beta}} + \hat{\psi}\underline{1}]$$

Solving:

(2) $$RA'[\frac{1}{r}(I_t \otimes \underline{1}_r') - \frac{1}{rt} J]\underline{y} = RA'\hat{\underline{\theta}} + \frac{1}{r} RA'AR[\hat{\underline{\beta}} + \hat{\psi}\underline{1}] \ .$$

Subtracting (2) from (1)

$$RA_d'\underline{y} - RA'[\frac{1}{t}(\underline{1}_t' \otimes I_r) - \frac{1}{rt} J]\underline{y} = [R^2 - \frac{1}{r} RA'AR](\hat{\underline{\beta}} + \hat{\psi}\underline{1}) \ .$$

Assuming the matrix $[R^2 - \frac{1}{r} RA'AR]$ to be nonsingular*

$$\hat{\underline{\beta}} + \hat{\psi}\underline{1} = [R^{-1}(I - \frac{1}{r} A'A)^{-1}R^{-1}][RA_d'\underline{y} - RA'\frac{1}{r}(I_t \otimes \underline{1}_r')\underline{y}]$$

and

$$(I - \frac{1}{r} J)[\hat{\underline{\beta}} + \hat{\psi}\underline{1}] = \hat{\underline{\beta}} = (I - \frac{1}{r} J)[R^{-1}(I - \frac{1}{r} A'A)^{-1}R^{-1}](RA_d'[I_{rt} - \frac{1}{r}(I_t \otimes J_r)]) \ .$$

This is the first time the condition $\sum \beta_j = 0$ or $\sum \hat{\beta}_j = 0$ has been utilized, and it is used here to eliminate $\hat{\psi}$.

$$\hat{\underline{\kappa}} = [\frac{1}{t}(\underline{1}_t' \otimes I_r) - \frac{1}{rt} J_{r \times rt}]\underline{y} \ .$$

---

*See Appendix.

Note that $RA_d'[I_{rt} - \frac{1}{r}(I_t \otimes J_r)]A_dR = R[I - \frac{1}{r}A'A]R$. We assume $\varepsilon_{ij} \sim NID(0, \sigma^2)$. Hence,

$$\hat{\underline{\beta}} \sim N[\underline{\beta}, \sigma^2(I - \frac{1}{r}J)R^{-1}(I - \frac{1}{r}A'A)^{-1}R^{-1}(I - \frac{1}{r}J)] \quad .$$

Likewise,

$$\hat{\underline{\kappa}} \sim N[\underline{\kappa}, \sigma^2 \frac{1}{t}(I - \frac{1}{r}J)] \quad .$$

We shall be interested in block effects at x_0, so we seek the distribution of $\hat{\kappa}_j + \hat{\beta}_j(x_0 - x_{\cdot j})$, and since $x_{\cdot j} = x_{\cdot j'}$, for all j and j', the vector of block effects at x_0 is $\hat{\underline{\kappa}} + (x_0 - x_{\cdot j})\hat{\underline{\beta}}$. Letting $z_0 = x_0 - x_{\cdot j}$, we seek the distribution of $\hat{\underline{\kappa}} + z_0\hat{\underline{\beta}}$.

If the independence of $\hat{\underline{\kappa}}$ and $\hat{\underline{\beta}}$ is not apparent from the manner in which they have been constructed, it can be easily verified by matrix multiplication.

$$A_d'[I_{rt} - \frac{1}{r}(I_t \otimes J_r)][\frac{1}{t}(\underline{1}_t \otimes I_r) - \frac{1}{rt}J_{rt\times r}] = 0 \quad .$$

Hence,

$$\hat{\underline{\kappa}} + z_0\hat{\underline{\beta}} \sim N\left(\underline{\kappa} + z_0\underline{\beta}, \sigma^2(I - \frac{1}{r}J)[\frac{1}{t}I + z_0^2R^{-1}(I - \frac{1}{r}A'A)^{-1}R^{-1}](I - \frac{1}{r}J)\right)$$

The matrix $(I - \frac{1}{r}J)R^{-1}(I - \frac{1}{r}A'A)^{-1}R^{-1}(I - \frac{1}{r}J)$ will, of course, figure heavily in the solution of any problem involving block differences. Some properties of this matrix are explored in the appendix.

CHAPTER III

AN EXACT TEST FOR THE HYPOTHESIS IN THE SPECIAL CASE

In terms of the parameters of Chapter I where $y_{ij} = \mu + \theta_i + \nu_j + \psi x_{ij} + \phi_j x_{ij} + \varepsilon_{ij}$, a statistic is desired to test the hypothesis that $\nu_j + x_0\phi_j = 0$ for all j . In terms of the parameters of Chapter II, this hypothesis becomes $\kappa_j + (x_0 - x_{.j})\beta_j = 0$.

In the special case of equal covariate block means where $x_0 - x_{.j} = x_0 - x_{.j'} = z_0$, the null hypothesis of no difference between blocks can be expressed parametrically as

$$H_o\colon\ \underline{\kappa} + z_0\underline{\beta} = \underline{0}$$

$$\text{vs.}\quad H_a\colon\ \underline{\kappa} + z_0\underline{\beta} \neq \underline{0} \quad .$$

From the work of Chapter II we see that

$$\hat{\underline{\kappa}} + z_0\hat{\underline{\beta}} \sim N[\underline{0},\ \sigma^2(I - \frac{1}{r}J)(\frac{1}{t}I + z_0^2 M)(I - \frac{1}{r}J)]$$

under under H_o where $M = R^{-1}(I - \frac{1}{r}A'A)^{-1}R^{-1}$. Thus under the null hypothesis

$$\hat{\underline{\kappa}} + z_0\hat{\underline{\beta}} \sim N[\underline{0},\ \sigma^2(I - \frac{1}{r}J)(\frac{1}{t}I + z_0^2 M)(I - \frac{1}{r}J)] \quad .$$

To test the null hypothesis, then, we seek a statistic with a well known (i.e., tabled) distribution under the null hypothesis and with expectation under the alternate hypothesis dependent on $(\underline{\kappa} + z_0\underline{\beta})'(\underline{\kappa} + z_0\underline{\beta})$ or on $(\underline{\kappa} + z_0\underline{\beta})'\Lambda(\underline{\kappa} + z_0\underline{\beta})$ where Λ is a positive definite matrix.

Such a statistic is available since

$$\frac{1}{\sigma^2} T(z_0) = \frac{1}{\sigma^2} (\hat{\underline{\kappa}} + z_0\hat{\underline{\beta}})'[\frac{1}{t} I + z_0^2(I - \frac{1}{r} J)M(I - \frac{1}{r} J)]^{-1}(\hat{\underline{\kappa}} + z_0\hat{\underline{\beta}}) \sim \chi^2_{r-1,\lambda(z_0)}$$

where

$$\lambda(z_0) = \frac{1}{2\sigma^2} (\underline{\kappa} + z_0\underline{\beta})'[\frac{1}{t} I + z_0^2(I - \frac{1}{r} J)M(I - \frac{1}{r} J)]^{-1}(\underline{\kappa} + z_0\underline{\beta}) \quad .$$

The matrix of this quadratic form is a nonsingular generalized inverse of the variance matrix of the random vector $\hat{\underline{\kappa}} + z_0\hat{\underline{\beta}}$.

To verify this distribution note that

$$(I - \frac{1}{r} J)[\frac{1}{t} I + z_0^2M](I - \frac{1}{r} J) = (I - \frac{1}{r} J)[\frac{1}{t} I + z_0^2(I - \frac{1}{r} J)M(I - \frac{1}{r} J)](I - \frac{1}{r} J)$$

and recall that when a vector

$$w^* = (I - \frac{1}{r} J)\underline{w} \sim N[\underline{\mu}, (I - \frac{1}{r} J)V(I - \frac{1}{r} J)]$$

then

$$\frac{1}{\sigma^2} \underline{w}^{*\prime}V^{-1}w^* = \frac{1}{\sigma^2} \underline{w}'(I - \frac{1}{r} J)V^{-1}(I - \frac{1}{r} J)\underline{w} \sim \chi^2_{r-1,\frac{1}{\sigma^2}(\underline{\mu}'V^{-1}\underline{\mu})} \quad .$$

This statistic is independent of the mean square error so a true F-statistic can be formed which will eliminate the necessity to know the error variance σ^2 or to estimate it from other information. The independence follows from the fact that the sum of squares for error is constructed orthogonal to those for $\underline{\kappa}$ and those for $\underline{\beta}$. Hence, MSE is independent of any function of the $\hat{\underline{\kappa}}$ and $\hat{\underline{\beta}}$ vectors.

Having produced a statistic which will test the hypothesis, we would be satisfied if we were interested in only one value of X(or Z). But, of course, we wish to produce a set of x's for which the null hypothesis can be rejected. It is in this regard that the statistic $T(z_0)$ becomes unwieldy.

Note that if C_{r-1} is a critical point for χ^2 statistic with r-1 degrees of freedom, then we would be interested in solving the equation $T(z) = \sigma^2 C_{r-1}$ for the z points at which T(z) equals the critical point. With those values of Z we could deduce the set of points for which $T(z) > \sigma^2 C_{r-1}$ from the continuity of T(z).

Decompose $(I - \frac{1}{r} J)M(I - \frac{1}{r} J)$ into eigenroots and vectors by letting

$$(I - \frac{1}{r} J)M(I - \frac{1}{r} J) = \sum_{i=1}^{r-1} \gamma_i \underline{\mu}_i \underline{\mu}_i' \qquad \begin{array}{ll} \underline{\mu}_i' \underline{1} = 0 & \\ \underline{\mu}_i' \underline{\mu}_{i'} = 1 & i = i' \\ \underline{\mu}_i' \underline{\mu}_{i'} = 0 & i \neq i' \end{array}$$

Note that there is one zero eigenroot which corresponds to the vector $\frac{1}{\sqrt{r}} \underline{1}$.

Then T(z) can be expanded in terms of these same eigenvectors.

$$\begin{aligned} T(z) &= (\hat{\underline{\kappa}} + z\hat{\underline{\beta}})' \left(\frac{1}{t} I + z^2 \sum_{i=1}^{r-1} \gamma_i \underline{\mu}_i \underline{\mu}_i' \right)^{-1} (\hat{\underline{\kappa}} + z\hat{\underline{\beta}}) \\ &= (\hat{\underline{\kappa}} + z\hat{\underline{\beta}})' \left(\sum_{i=1}^{r-1} (\frac{1}{t} + z^2\gamma_i) \underline{\mu}_i \underline{\mu}_i' + \frac{1}{t}\frac{1}{r} \underline{1}\underline{1}' \right)^{-1} (\hat{\underline{\kappa}} + z\hat{\underline{\beta}}) \\ &= (\hat{\underline{\kappa}} + z\hat{\underline{\beta}})' \left(\sum_{i=1}^{r-1} \frac{\underline{\mu}_i \underline{\mu}_i'}{z^2\gamma_i + \frac{1}{t}} \right) (\hat{\underline{\kappa}} + z\hat{\underline{\beta}}) \quad \text{since} \quad (\hat{\underline{\kappa}} + z\hat{\underline{\beta}})'(\underline{1}\underline{1}')(\hat{\underline{\kappa}} + z\hat{\underline{\beta}}) = 0 \\ &= \sum_{i=1}^{r-1} \frac{\hat{\underline{\kappa}}' \underline{\mu}_i \underline{\mu}_i' \hat{\underline{\kappa}} + 2z\hat{\underline{\kappa}}' \underline{\mu}_i \underline{\mu}_i' \hat{\underline{\beta}} + z^2 \hat{\underline{\beta}}' \underline{\mu}_i \underline{\mu}_i' \hat{\underline{\beta}}}{z^2\gamma_i + \frac{1}{t}} \quad . \end{aligned}$$

Each of the r-1 terms is a ratio of two quadratic functions in z. Hence, if r=2 there would be only one term and we could solve the equation $T(z) = \sigma^2 C_{r-1}$ getting at most two solutions z_{c1} and z_{c2}, $z_{c1} < z_{c2}$. We would need to check the magnitude of T(z) and any one point, probably T(0). Without loss of generality assume that $T(0) < \sigma^2 C_{r-1}$. Then if $0 \varepsilon (z_{c1}, z_{c2})$

we would state that blocks are significantly different only outside the interval (z_{c1}, z_{c2}). Or, if $0 \notin (z_{c1}, z_{c2})$ we would state that blocks are significantly different only inside the interval (z_{c1}, z_{c2}).

This same easy solution is available in the event that all the eigenroots are identical, that is, $\gamma_i = \gamma_{i'} = \gamma$ for all i and i'. Then

$$T(z) = \frac{\hat{\underset{\sim}{\kappa}}'\hat{\underset{\sim}{\kappa}} + 2z\hat{\underset{\sim}{\kappa}}'\hat{\underset{\sim}{\beta}} + z^2\hat{\underset{\sim}{\beta}}'\hat{\underset{\sim}{\beta}}}{z^2\gamma + \frac{1}{t}} \quad .$$

And again equating T(z) to a critical value, $\sigma^2 C_{r-1}$, will yield at most two values for z. Again by continuity we know that blocks will be significantly different only when z is between those two values or outside that interval.

These special cases, $\gamma_i = \gamma_{i'}$ for all i and i', can be considered geometrically by examining the equation

$$T(z) = \frac{\hat{\underset{\sim}{\kappa}}'\hat{\underset{\sim}{\kappa}} + 2z\hat{\underset{\sim}{\kappa}}'\hat{\underset{\sim}{\beta}} + z^2\hat{\underset{\sim}{\beta}}'\hat{\underset{\sim}{\beta}}}{z^2\gamma + \frac{1}{t}} \quad .$$

First notice

$$T(0) = t\hat{\underset{\sim}{\kappa}}'\hat{\underset{\sim}{\kappa}} \qquad T(z) \geq 0 \text{ for all } z$$

$$T(\infty) = \hat{\underset{\sim}{\beta}}'\hat{\underset{\sim}{\beta}} \quad .$$

The extreme points of the function T(z) are obtained by

$$\frac{\partial T}{\partial z} = \frac{(z^2\gamma + \frac{1}{t})(2\hat{\underset{\sim}{\kappa}}'\hat{\underset{\sim}{\beta}} + 2z\hat{\underset{\sim}{\beta}}'\hat{\underset{\sim}{\beta}}) - 2z\gamma(\hat{\underset{\sim}{\kappa}}'\hat{\underset{\sim}{\kappa}} + 2z\hat{\underset{\sim}{\kappa}}'\hat{\underset{\sim}{\beta}} + z^2\hat{\underset{\sim}{\beta}}'\hat{\underset{\sim}{\beta}})}{(z^2\gamma + \frac{1}{t})^2}$$

$$= \frac{\frac{1}{t}\, 2\hat{\underset{\sim}{\kappa}}'\hat{\underset{\sim}{\beta}} + z\left(\frac{2\hat{\underset{\sim}{\beta}}'\hat{\underset{\sim}{\beta}}}{t} - 2\gamma\hat{\underset{\sim}{\kappa}}'\hat{\underset{\sim}{\kappa}}\right) - z^2 2\hat{\underset{\sim}{\kappa}}'\hat{\underset{\sim}{\beta}}\gamma}{(z^2\gamma + \frac{1}{t})^2} \quad .$$

Then

$$0 = z^2\underline{\hat{\kappa}}'\underline{\hat{\beta}}\gamma - z\,\frac{\underline{\hat{\beta}}'\underline{\hat{\beta}}}{t} - 2\gamma\underline{\hat{\kappa}}'\underline{\hat{\kappa}} \;-\; \frac{\underline{\hat{\kappa}}'\underline{\hat{\beta}}}{t}$$

where z is a maximum or minimun of the function T(z).

$$z_m = \frac{1}{2\underline{\hat{\kappa}}'\underline{\hat{\beta}}\gamma}\left[\frac{\underline{\hat{\beta}}'\underline{\hat{\beta}}}{t} - \gamma\underline{\hat{\kappa}}'\underline{\hat{\kappa}} \pm \sqrt{\left(\frac{\underline{\hat{\beta}}'\underline{\hat{\beta}}}{t} - \gamma\underline{\hat{\kappa}}'\underline{\hat{\kappa}}\right)^2 + \frac{\gamma 4(\underline{\hat{\kappa}}'\underline{\hat{\beta}})^2}{t}}\,\right] \quad \text{if } \underline{\hat{\kappa}}'\underline{\hat{\beta}} \neq 0 \quad .$$

$$\text{The discriminant} = \left[\left(\frac{\underline{\hat{\beta}}'\underline{\hat{\beta}}}{t} - \gamma\underline{\hat{\kappa}}'\underline{\hat{\kappa}}\right)^2 + \frac{\gamma 4(\underline{\hat{\kappa}}'\underline{\hat{\beta}})^2}{t}\right] > 0$$

and

$$z_c = 0 \qquad \text{if } \underline{\hat{\kappa}}'\underline{\hat{\beta}} = 0 \quad .$$

Hence, if $\underline{\hat{\kappa}}'\underline{\hat{\beta}} \neq 0$, T(z) has two extreme points $z_{c1} < 0 < z_{c2}$ and the single extreme point 0 if $\underline{\hat{\kappa}}'\underline{\hat{\beta}} = 0$.

If r=2, one of the extreme points yield a minimum of T(z) = 0 . If r > 2 it is doubtful that T(z) will actually attain 0 .

Then, if $\underline{\hat{\kappa}}'\underline{\hat{\beta}} > 0$, T(z) has two real extreme points and we can sketch the illustrative curves of Figure a.

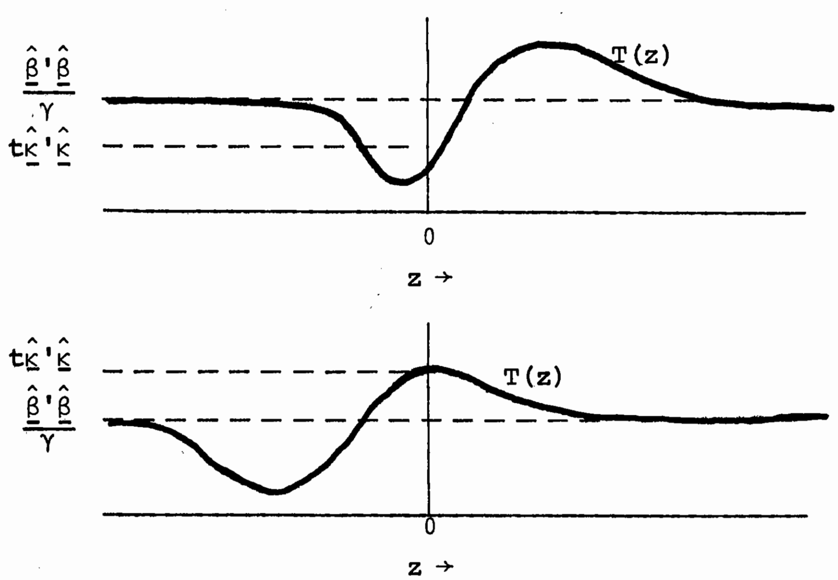


Figure a.

And if $\hat{\underline{\kappa}}'\hat{\underline{\beta}} < 0$, T(z) will yield curves such as those of Figure b.

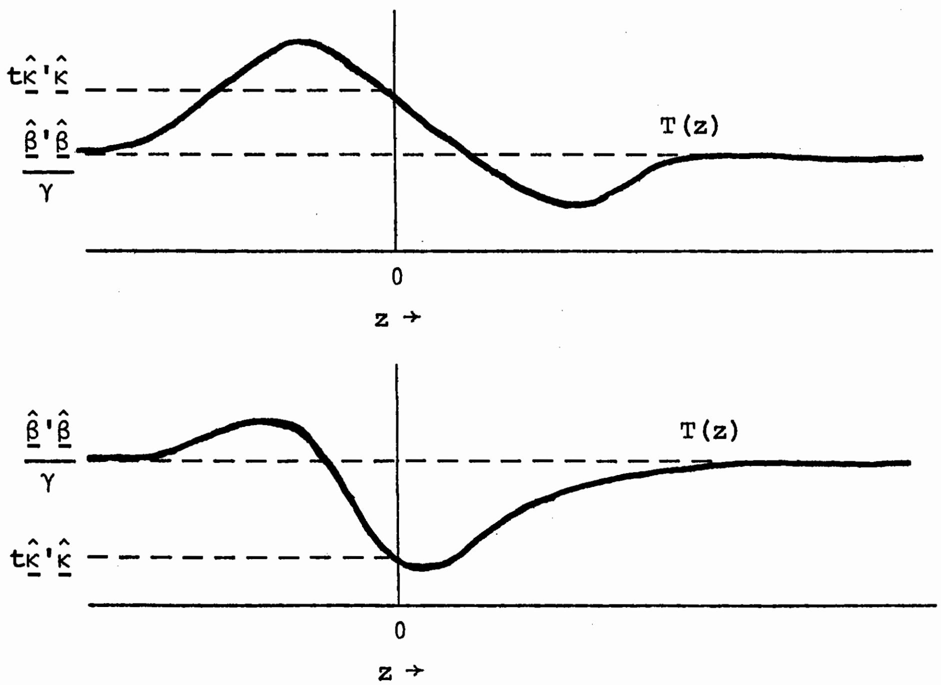


Figure b.

And if $\hat{\underline{\kappa}}'\hat{\underline{\beta}} = 0$, the shape is determined by whether $\frac{\hat{\underline{\beta}}'\hat{\underline{\beta}}}{\gamma} < t\hat{\underline{\kappa}}'\hat{\underline{\kappa}}$ or $t\hat{\underline{\kappa}}'\hat{\underline{\kappa}} < \frac{\hat{\underline{\beta}}'\hat{\underline{\beta}}}{\gamma}$. See Figure c.

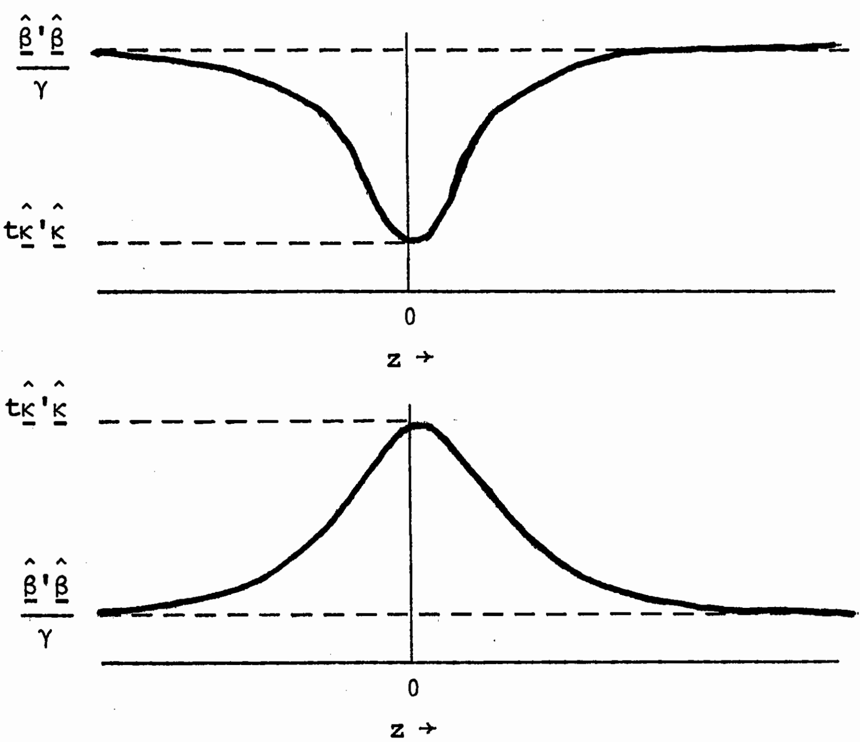


Figure c.

In all of these figures a critical point for the statistic at any z_0 is the point $(z_0, \sigma^2 C_{r-1})$. Thus a critical curve for the function $T(z)$ would be a straight line $\sigma^2 C_{r-1}$ units above the origin. Any horizontal line which intersects any of the curves intersects it twice giving rise to end points of a critical interval (z_{c1}, z_{c2}). [The only exception to this generality would be a line $\frac{\hat{\underline{\beta}}'\hat{\underline{\beta}}}{\gamma}$ units above the origin in Figures a. and b. which would produce a critical interval of the form $(-\infty, z_c)$ or $(z_c, +\infty)$.]

Let Ω_T be the set of z's such that $T(z) > \sigma^2 C_{r-1}$. That is, if $z_0 \in \Omega_T$, block effects are significantly different when $Z = z_0$. Let $\overline{U}$ be the totality of points on the real line.

Then when $\gamma_i = \gamma_{i'}$ for all i, i', the set $\Omega_T = (z_{c1}, z_{c2})$ or $\Omega_T = \overline{U} - (z_{c1}, z_{c2})$. The symbolism $\overline{U} - (z_{c1}, z_{c2})$ is used as shorthand notation for $(-\infty, z_{c1}) \cup (z_{c2}, +\infty)$, the union of two open-ended intervals. No distinction will be made here between open and closed intervals which shall be justified only by saying that $T(z) = \sigma^2 C_{r-1}$ only on a set of z's of measure zero.

The above reference to infinite values of the covariate occur only for completeness in consideration of the function $T(z)$. As in any regression type problem, we are not justified in extrapolating our results beyond the range of the covariate actually covered in the experiment. We are even more restricted in this particular model, and should not place much weight on covariate values beyond the range covered in each and every block. Hence, our interest in $T(\infty)$ is as an unattainable limit point rather than an asymptote.

Thus far we have considered only the situation when all the eigenroots are identical. When the roots are different the situation is much more complex. Let r=3 so that $(I - \frac{1}{r} J)M(I - \frac{1}{r} J)$, the covariance matrix of $\underline{\beta}$, has two roots γ_1 and γ_2 such that $\gamma_1 \neq \gamma_2$. Then

$$T(z) = \frac{(\hat{\underline{\kappa}}'\underline{\mu}_1)^2 + 2z(\hat{\underline{\kappa}}'\underline{\mu}_1)(\hat{\underline{\beta}}'\underline{\mu}_1) + z^2(\hat{\underline{\beta}}'\underline{\mu}_1)^2}{\gamma_1 z^2 + \frac{1}{t}} + \frac{(\hat{\underline{\kappa}}'\underline{\mu}_2)^2 + 2z(\hat{\underline{\kappa}}'\underline{\mu}_2)(\hat{\underline{\beta}}'\underline{\mu}_2) + z^2(\hat{\underline{\beta}}'\underline{\mu}_2)^2}{\gamma_2 z^2 + \frac{1}{t}}$$

If T(z) is set equal to a critical value the resulting equation is of 4^{th} degree in z. That is, we could have four values z_1, z_2, z_3, and z_4 such that $\Omega_T = (z_1, z_2) \cup (z_3, z_4)$ or $\overline{U} - [(z_1, z_2) \cup (z_3, z_4)]$.

When there are three different eigenroots equating T(z) to a critical value produces as many as 6 critical z's. Then the set Ω_T could have the form $(z_1, z_2) \cup (z_3, z_4) \cup (z_5, z_6)$ or $\overline{U} - [(z_1, z_2) \cup (z_3, z_4) \cup (z_5, z_6)]$. Likewise, when there are r-1 different eigenroots, there can be as many as 2(r-1) different critical z's.

Consider the following example.

<u>Example 1.</u> Let

$$\sigma^2 = 1 \ ; \ r = 3 \ ; \ t = 5 \ ; \ \gamma_1 = \frac{1}{5} \ ; \ \gamma_2 = 1 \ ;$$

$$\underline{\mu}_1' = \frac{1}{\sqrt{2}}(1, -1, 0) \ ; \ \underline{\mu}_2' = \frac{1}{\sqrt{6}}(1, 1, -2) \ ;$$

$$x_{\cdot j} = x_{\cdot j'} \ , \text{ for all } j, j' \quad .$$

That is, this is an example of the special case of equal covariate block means.

$$\hat{\underline{\kappa}}' = \left(\frac{2\sqrt{6} - 3\sqrt{2}}{12} , \frac{2\sqrt{6} + 3\sqrt{2}}{12} , -\frac{\sqrt{6}}{3}\right)$$

$$\hat{\underline{\beta}}' = \left(\frac{\sqrt{6} + 3\sqrt{2}}{6} , \frac{\sqrt{6} - 3\sqrt{2}}{6} , -\frac{\sqrt{6}}{3}\right)$$

Let

$$T_1(z) = \frac{(\hat{\underline{\kappa}}'\underline{\mu}_1)^2 + 2z(\hat{\underline{\kappa}}'\underline{\mu}_1)(\hat{\underline{\beta}}'\underline{\mu}_1) + (\hat{\underline{\beta}}'\underline{\mu}_1)^2 z^2}{.2z^2 + .2} = \frac{z^2 - z + .25}{.2z^2 + .2}$$

and let

$$T_2(z) = \frac{(\hat{\underline{\kappa}}'\underline{\mu}_2)^2 + 2z(\hat{\underline{\kappa}}'\underline{\mu}_2)(\hat{\underline{\beta}}'\underline{\mu}_2) + (\hat{\underline{\beta}}'\underline{\mu}_2)^2}{z^2 + .2} = \frac{z^2 + 2z + 1}{z^2 + .2}$$

Then

$$T(z) = T_1(z) + T_2(z) \sim \left(\chi^2_{2,\lambda}\right)\sigma^2$$

where $\lambda = 0$ when $\hat{\underline{\kappa}} + z\hat{\underline{\beta}} = 0$ since $\sigma^2 = 1$.

The functions T_1 and T_2 each have the shape of a T(z) function with identical roots as previously discussed; but their sum does not (see Figure I). (The illustrations with Roman numerals appear in Appendix II.)

With a significance level of .10, i.e., $\alpha = .10$; $\Omega_T = \overline{\cup} - (-.53, -.31) - (.61, 2.4)$. The $\alpha = .10$ critical line for χ^2_1 is shown on Figure I also. The $T_1(z)$ and $T_2(z)$ cruves can be individually compared with that line.

CHAPTER IV

INTERPRETATION OF EXACT TEST RESULTS

Not only is it difficult to solve a problem for Ω_T in this special case of equal covariate block means when the eigenroots of the covariance matrix are different, it is also difficult to interpret the solution. (Such a solution is even more difficult to find in the general case, as will be seen later.)

In an attempt to interpret Ω_T , visualize r lines with intercepts κ_j and slopes β_j . Unless a set of r straight lines have exactly the same slope they will become "infinitely" far apart as $|z|$ approaches ∞. So in deciding how to answer the question "Are block effects different when $Z = z_0$?" or "Are these straight lines different at $Z = z_0$?", one must use the magnitude of z_0 as part of the judgment criteria. The larger $|z|$ becomes, the more difference between the lines it takes to be "surprising."

$T(0) = t\hat{\underline{\kappa}}'\hat{\underline{\kappa}}$ will be recognized as a test statistic for "block mean effects." Likewise $T(\infty) = \hat{\underline{\beta}}'[R(I - \frac{1}{r} A'A)R]\hat{\underline{\beta}}$ is the usual test for block regression effects. It is essentially this test to which Cox [1] refers. In this problem $T(\infty)$ is the value of the test statistic when $|z|$ is so large that the fixed effects, or intercept effects, κ_j, are completely "washed out." Of course, in practice, this point $(Z = \infty)$ never occurs because the upper and lower values of z at which T(z) is meaningful is limited by the range of the covariate within each block in the experiment.

If $T(\infty) > \sigma^2 C_{r-1}$, i.e., $\infty \in \Omega_T$, then the r lines in question have "different" or "significantly different" slopes. The r lines are restricted in the sense that the sum of their slopes is 0 and the sum of their intercepts is 0. For examples with 4 lines, see Figures d. and e.

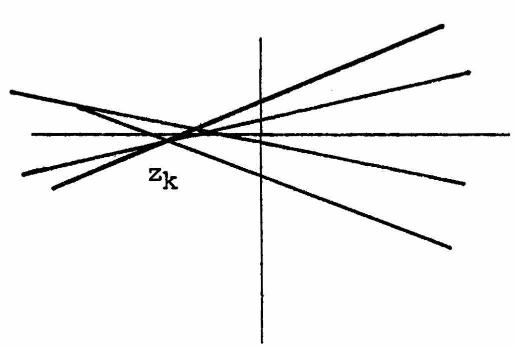


Figure d.

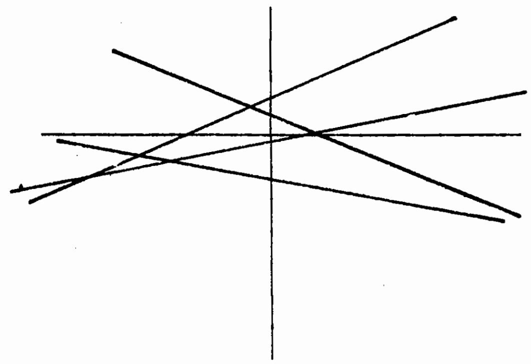


Figure e.

In Figure d, although the lines do not meet at a single point it would not be surprising to find an interval about z_k in which block differences are uncertain. Whether or not that interval would include the point z = 0 would depend on whether or not $T(0) < \sigma^2 C_{r-1}$. T(0) will be recognized as the usual test for the $\underset{\sim}{\kappa}$ effects. Figures d and e have been drawn with the same κ_j values so that if $T(0) < \sigma^2 C_{r-1}$ in Figure d, T(0) is also less than $\sigma^2 C_{r-1}$ in Figure e, and it would be logical to assume an interval about 0 in which the lines would not be considered significantly different.

Similarly if $T(\infty) < \sigma^2 C_{r-1}$ then the r lines will have approximately the same slopes; and since the slopes are restricted to add to zero, the lines will all have slopes very close to zero. See Figures f and g for examples with 4 lines.

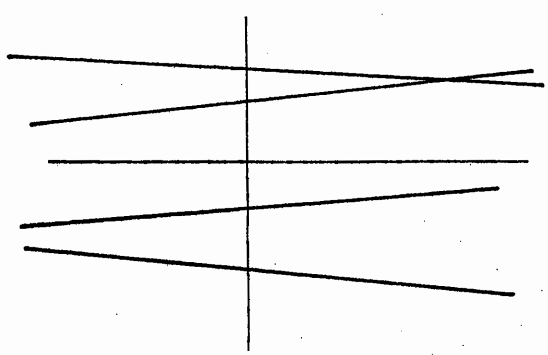


Figure f.

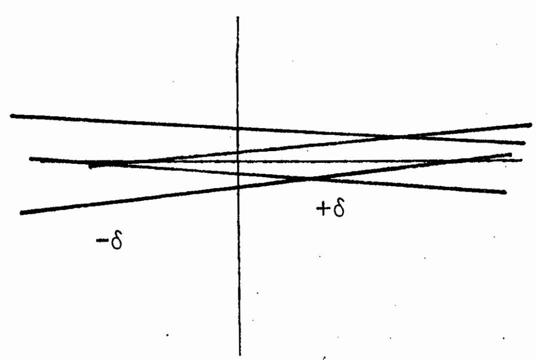


Figure g.

These cases are more difficult to evaluate subjectively because the effect of κ_j "washes out" as $|z|$ increases.

If $T(0) > \sigma^2 C_{r-1}$ (perhaps Figure f) then there would probably be an interval about 0 in which $T(z) > \sigma^2 C_{r-1}$. However if $T(\infty) < \sigma^2 C_{r-1}$ and $T(0) < \sigma^2 C_{r-1}$ (Figure g) it may be difficult to understand that there could still be an interval (in Figure g in the region $z < 0$) in which the lines could be considered significantly different. If this is so it is because the lines spread apart faster than we would expect with regard to $|z|$. Note that the lines are farther apart at $-\delta$ than at $+\delta$ in Figure g.

Scale has intentionally been omitted from these figures to force subjective thinking rather than mathematical manipulation. Certainly any decisions must take scale into account. The purpose here is to suggest that the logical systems for evaluating whether a set of lines are appreciably different at $Z = z_0$ relative to $|z_0|$ give rise to single intervals of difference (as in Figures f and g) and to complements of single intervals of difference (as in Figures d and e).

Stated differently, let Ω_k be the set of z's for which a set of lines is considered to be different by some criterion k. If $\infty \in \Omega_k$, $z' \notin \Omega_k$, and $z'' \notin \Omega_k$, then we would want $Z \notin \Omega_k$ if $z' < z < z''$. (Figures d and e.) And if $\infty \notin \Omega_k$, $z' \in \Omega_k$, and $z'' \in \Omega_k$, then we would want $z \in \Omega_k$ if $z' < z < z''$. (Figures f and g.)

This desire for Ω_k to be of the form (z_1, z_2) or $\mho - (z_1, z_2)$ arises from placing equal weight on each of the lines, or on each contrast of the lines since they are restricted to add to 0. And, indeed, Ω_T is of this form when $\gamma_i = \gamma_{i'}$ for all i and i'.

However, $T(z)$ does not place equal weight on each contrast of the β_j's unless $\gamma_i = \gamma_{i'}$ for all i and i'. Thus Ω_T is not necessarily of the form (z_1, z_2) or $\overline{U} - (z_1, z_2)$. Through the matrix $(I - \frac{1}{r} J)R^{-1}(I - \frac{1}{r} A'A)^{-1}R^{-1}(I - \frac{1}{r} J)$ the statistic places more weight on the contrasts about which there is the most information. For clarification, consider again Example 1 (page 18).

In Figure II (Appendix II), the lines $v_1 = \frac{2\sqrt{6} - 3\sqrt{2}}{12} + \frac{\sqrt{6} + 3\sqrt{2}}{6} z$, $v_2 = \frac{2\sqrt{6} + 3\sqrt{2}}{12} + \frac{\sqrt{6} - 3\sqrt{2}}{6} z$, and $v_3 = -\frac{\sqrt{6}}{3} - \frac{\sqrt{6}}{3} z$ are plotted. The shaded area represents $\overline{U} - \Omega_T$, the z values at which we cannot be certain of block differences testing with a .10 significance level. Because of the $\sum \beta_j = 0$ and $\sum \kappa_j = 0$ restrictions, the 3 lines of Figure II are redundant.

In Figure III, the two unrestricted lines $w_1 = \hat{\underline{\kappa}}'\underline{v}_1 + \hat{\underline{\beta}}'\underline{\mu}_1 z = .5 + z$ and $w_2 = \hat{\underline{\kappa}}'\underline{v}_2 + \hat{\underline{\beta}}'\underline{\mu}_2 z = 1 + z$ are plotted. The region of $\overline{U} - {}_T$ is again shaded.

In Figure III, where the two lines are perfectly parallel, the multiple Ω_T-zone is particularly enigmatic as an answer to the question "For what values of z, considering $|z|$, can w_1 and w_2 be considered estimates of zero?"

It is understandable that when 0 is in Ω_T points on either side of 0 might not be, due to the "washing out" of $\underline{\kappa}$ with increasing $|z|$. But that even larger values of $|z|$ should indicate block differences is difficult to accept.

The result is due, of course, to the difference in roots causing more emphasis on w_1 as $|z|$ increases.

The unequal weighting reflects the fact that we have more information about $\hat{\underline{\kappa}}'\underline{\mu}_1$ and $\hat{\underline{\beta}}'\underline{\mu}_1$ than $\hat{\underline{\kappa}}'\underline{\mu}_2$ and $\hat{\underline{\beta}}'\underline{\mu}_2$. Certainly this ability to concentrate on the qualities about which we have the most information can be a desirable property for a testing statistic. But when there is equal interest in each of the contrasts between effects this property may not be so desirable.

In terms of power, T(z) would have higher power against alternatives which make $(\underline{\kappa}'\underline{\mu}_1 + z\underline{\beta}'\underline{\mu}_1)^2$ large than against alternatives which make $(\underline{\kappa}'\underline{\mu}_2 + z\underline{\beta}'\underline{\mu}_2)^2$ large, whereas equal power against both of these types of alternative hypotheses would be desirable.

If we could control the covariate Z we would have equal information about each of the contrasts. When we don't have equal information, i.e., equal roots, we must consider whether we prefer a more exact answer to a question we are only approximately asking or a more approximate answer to the exact question we are asking.

Recall that when we first sought a statistic to evaluate block differences (page 11) the first choice was one with expectation dependent on $(\underline{\kappa} + z_0\underline{\beta})'(\underline{\kappa} + z_0\underline{\beta})$, and T(z) with expectation dependent on $(\underline{\kappa} + z\underline{\beta})'[\frac{1}{t} I + z^2(I - \frac{1}{r} J)M(I - \frac{1}{r} J)]^{-1}(\underline{\kappa} + z\underline{\beta})$ was selected only because the exact distribution of T(z) is tabled.

Let us look at a statistic with expectation dependent on $(\underline{\kappa} + z\underline{\beta})'(\underline{\kappa} + z\underline{\beta})$.

CHAPTER V

AN APPROXIMATE TEST OF THE HYPOTHESIS IN THE SPECIAL CASE

As in Chapter III, a test is sought for the hypothesis which is expressed parametrically as

$$H_0: \ \underline{\kappa} + z_0\underline{\beta} = \underline{0}$$

$$\text{vs.} \quad H_a: \ \underline{\kappa} + z_0\underline{\beta} \neq \underline{\ }$$

$$\hat{\underline{\kappa}} + z_0\hat{\underline{\beta}} \sim N[\underline{\kappa} + z_0\underline{\beta} \ , \ \sigma^2(I - \frac{1}{r} J)(\frac{1}{t} I + z_0^2 M)(I - \frac{1}{r} J)]$$

where

$$M = R^{-1}(I - \frac{1}{r} A'A)^{-1}R^{-1}$$

and

$$\hat{\underline{\kappa}} + z_0\hat{\underline{\beta}} = (H + z_0 L)\underline{y}$$

where

$$H_{r\times rt} = [\frac{1}{t} \underline{1} \otimes I_r - \frac{1}{rt} \underline{1}_t' \otimes J_r]$$

and

$$L = (I - \frac{1}{r} J)MRA_d'(I_{rt} - [I_t \otimes \frac{1}{r} J_r]) \quad .$$

Let

$$S^*(z_0) = (\hat{\underline{\kappa}} + z_0\hat{\underline{\beta}})'(\hat{\underline{\kappa}} + z_0\hat{\underline{\beta}}) = \underline{y}'[H + z_0L]'[H + z_0L]\underline{y} \quad .$$

$T(z_0)$ is a central χ^2 under the null hypothesis. $S^*(z)$ is not a χ^2, but it is "central" in the sense that $\underline{\mu}'[H+z_0L]'[H+z_0L]\underline{\mu} = (\underline{\kappa}+z_0\underline{\beta})'(\underline{\kappa}+z_0\underline{\beta}) = 0$ only when $\underline{\kappa} + z_0\underline{\beta} = \underline{0}$, where $\underline{\mu} = E(\underline{y})$.

Although $S^*(z_0)$ is not recognizable as any single distribution, it can be decomposed as a sum of independent χ^2 statistics. Let

$$[H + z_0L]'[H + z_0L] = \sum_{i=1}^{r-1} \xi_i \underset{\sim}{\ell}_i \underset{\sim}{\ell}_i'$$

where

$$\underset{\sim}{\ell}_i' \underset{\sim}{\ell}_i = 1$$

$$\underset{\sim}{\ell}_i' \underset{\sim}{\ell}_{i'} = 0 \qquad i \neq i' ,$$

$\underset{\sim}{\ell}_i$ being a vector with rt elements. Note that the rank of $H + z_0L$ is r-1 so only r-1 of the rt eigenroots are non-zero.

We know that the non-zero roots of $(H + z_0L)'(H + z_0L)$ are identical to the non-zero roots of $(H + z_0L)(H + z_0L)'$. And $(H + z_0L)(H + z_0L)' = (I - \frac{1}{r} J)[\frac{1}{t} I + z_0^2(I - \frac{1}{r} J)M(I - \frac{1}{r} J)](I - \frac{1}{r} J)$, which we recall from $T(z_0)$. In the consideration of $T(z_0)$ we let $(H + z_0L)(H + z_0L)' = \sum_{i=1}^{r-1} (\frac{1}{t} + z_0^2\gamma_i)\underset{\sim}{\mu}_i\underset{\sim}{\mu}_i'$. Hence the set of $(\frac{1}{t} + z_0^2\gamma_i)$'s is identical to the set of ξ_i's and we shall use the notation $\frac{1}{t} + z_0^2\gamma_i$ since it shows the dependency of the roots on z_0 .

Therefore,

$$S^*(z_0) = \underset{\sim}{y}' \sum_{i}^{r-1} [(\frac{1}{t} + z_0^2\gamma_i)\underset{\sim}{\ell}_i\underset{\sim}{\ell}_i']\underset{\sim}{y} = \sum_{i=1}^{r-1} (\frac{1}{t} + z_0^2\gamma_i)\underset{\sim}{y}'\underset{\sim}{\ell}_i\underset{\sim}{\ell}_i'\underset{\sim}{y}$$

where

$$\frac{1}{\sigma^2} \underset{\sim}{y}'\underset{\sim}{\ell}_i\underset{\sim}{\ell}_i'\underset{\sim}{y} \sim \chi^2_{1,\lambda_i} \quad .$$

That is, each of these r-1 terms is proportional to a χ^2 statistic with one degree of freedom and all are independent. That $\lambda_i = 0$ under H_o, the hypothesis that $\underset{\sim}{\kappa} + z_0\underset{\sim}{\beta} = \underset{\sim}{0}$, follows from the "centrality" of $S^*(z_0)$. That is, $\underset{\sim}{\mu}'[\sum \xi_i\underset{\sim}{\ell}_i\underset{\sim}{\ell}_i']\underset{\sim}{\mu} = 0 \Rightarrow \underset{\sim}{\mu}'\underset{\sim}{\ell}_i\underset{\sim}{\ell}_i'\underset{\sim}{\mu} = 0$.

Being a weighted sum of independent χ^2 statistics under H_o, the distribution of $S^*(z_0)$, under H_o, can be approximated by fitting the first two moments to a multiple of a χ^2 statistic. Since

$$E[S^*(z_0)] = \sigma^2 \sum_{i}^{r-1} \left(\frac{1}{t} + z_0^2\gamma_i\right)$$

$$E\left[\frac{S^*(z_0)n(z_0)}{\sigma^2\sum\left(\frac{1}{t} + \gamma_i z_0^2\right)}\right] = n(z_0)$$

where $n(z_0)$ is the number of degrees of freedom of the χ^2 statistic which is being approximated. Then we wish

$$\mathrm{Var}\left[\frac{S^*(z_0)n(z_0)}{\sigma^2 \sum_{i=1}^{r-1} \left(\frac{1}{t} + \gamma_i z_0^2\right)}\right] = 2n(z_0) \quad ,$$

or

$$\frac{[n(z_0)]^2}{\sigma^4\left[\sum\left(\frac{1}{t} + \gamma_i z_0^2\right)\right]^2} \quad \mathrm{Var}\ S^*(z_0) = 2n(z_0) \ .$$

Now

$$\mathrm{Var}\ S^*(z_0) = \sum_{i}^{r-1} \mathrm{Var}\left(\gamma_i z_0^2 + \frac{1}{t}\right)\underline{y}'\underline{\ell}_i\underline{\ell}_i'\underline{y} = \sum_{i}^{r-1} \left(\gamma_i z_0^2 + \frac{1}{t}\right)^2 2\sigma^4$$

$$= 2\sigma^4\left[z_0^4 \sum_{i}^{r-1} \gamma_i^2 + \frac{2}{t} z_0^2 \sum_{i}^{r-1} \gamma_i + \frac{r-1}{t^2}\right]$$

$$= \frac{\sum_{i}^{r-1} \left(\gamma_i z_0^2 + \frac{1}{t}\right)^2}{\left[\sum_{i}^{r-1} \left(\gamma_i z_0^2 + \frac{1}{t}\right)\right]^2} \ .$$

Hence

$$n(z_0) = \frac{\left(z_0^2 \sum_i^{r-1} \gamma_i + \frac{r-1}{t}\right)^2}{\sum_i^{r-1} (\gamma_i z_0^2 + \frac{1}{t})^2} = \frac{\left(z_0^2 \sum_i^{r-1} \gamma_i + \frac{r-1}{t}\right)^2}{z_0^4 \sum_i^{r-1} \gamma_i^2 + 2\frac{z_0^2}{t} \sum_i^{r-1} \gamma_i + \frac{r-1}{t^2}}$$

$$= (r-1)\left[\frac{\left(z_0^2 \sum_i^{r-1} \gamma_i + \frac{r-1}{t}\right)^2}{\left(z_0^2 \sum_i^{r-1} \gamma_i + \frac{r-1}{t}\right)^2 + z_0^4\left[r-1 \sum_i^{r-1} \gamma_i^2 - \left(\sum_i^{r-1} \gamma_i\right)^2\right]}\right]$$

or

$$n(z_0) = r-1\left[\frac{(z_0^2\bar{\gamma} + \frac{1}{t})^2}{(z_0^2\bar{\gamma} + \frac{1}{t})^2 + z_0^4(\bar{\bar{\gamma}} - \bar{\gamma}^2)}\right] ,$$

where

$$\bar{\gamma} = \frac{1}{r-1}\left(\sum \gamma_i\right)$$

and

$$\bar{\bar{\gamma}} = \frac{1}{r-1}\left(\sum \gamma_i^2\right) .$$

Let

$$\delta = \bar{\bar{\gamma}} - \bar{\gamma}^2$$

and since there are r-1 γ_i's all greater than 0

$$\frac{\left(\sum \gamma_i\right)^2}{r-1} \le \sum \gamma_i^2 < \left(\sum \gamma_i\right)^2$$

$$\bar{\gamma}^2 = \frac{\left(\sum \gamma_i\right)^2}{r-1} \le \bar{\bar{\gamma}} < \frac{\left(\sum \gamma_i\right)^2}{r-1} = (r-1)\bar{\gamma}^2$$

$$0 \le \delta < (r-2)\bar{\gamma}^2 .$$

Hence,

$$n(z_0) \le r-1$$

and

$$n(z_0) > r-1 \left[\frac{(\bar{\gamma})^2}{\bar{\gamma}^2 + (r-2)\bar{\gamma}^2}\right] = 1 \quad .$$

$n(z_0) = r-1$ when $z_0 = 0$ or when $\delta = 0$. $\delta = 0$ corresponds to the case of equal eigenroots. Let

$$S(z_0) = \frac{S^*(z_0)}{z_0^2\bar{\gamma} + \frac{1}{t}} \quad .$$

Then

$$\frac{1}{\sigma^2} S(z_0) \frac{n(z_0)}{r-1} \mathrel{\dot{\sim}} \chi^2_{n(z_0)}$$

and

$$S(z_0) = T(z_0)$$

when $\gamma_i = \gamma_{i'} = \bar{\gamma}$ for all i and i' .

When the roots are equal both $S(z_0)$ and $T(z_0)$ give the same result. When the roots are different, $T(z_0)$ becomes very cumbersome and can yield results which are difficult to interpret. When the roots are different $S(z_0)$ compensates by varying the critical point, the amount of compensation dependent on $\sum \gamma_i$ and $\sum \gamma_i^2$ which are, respectively the traces of the matrices $(I - \frac{1}{r} J)R^{-1}(I - \frac{1}{r} A'A)^{-1}R^{-1}(I - \frac{1}{r} J)$ and $[(I - \frac{1}{r} J)R^{-1}(I - \frac{1}{r} A'A)^{-1}R^{-1}(I - \frac{1}{r} J)]^2$.

In order to discover the nature and extent of the S(z) statistic's compensation for different roots, it will be necessary to examine the function $\frac{r-1}{n(z_0)} \chi^2_{n(z_0)}$.

CHAPTER VI

ANALYSIS OF APPROXIMATE TEST

The function S(z) has exactly the same shape characteristics as the function T(z) when the roots are identical. That function is discussed fully in Chapter III. However, the critical function for T(z) with identical roots is a straight line with ordinate $\sigma^2 C_{r-1,\alpha}$ where α gives the significance level of the test, whereas the critical function for S(z) is the function $\sigma^2 K_{n(z)} = \sigma^2 \frac{r-1}{n(z)} C_{n(z),\alpha}$. To analyze this critical function let us consider the variation of $C_{n,\alpha}$ with n . The functional notation showing that n is dependent on z will not be used here. From the previous chapter it should be clear that n is dependent not only on z but on r, t, and the values of the covariate X used in the experiment. At this point attention should be focused only on the variation of $C_{n,\alpha}$ with n regardless of how or why n might be varying. The notation "n(z)" will be resumed later.

$C_{n,\alpha}$ is defined by

$$\int_0^{C_{n,\alpha}} g(\chi_n^2)\, d\chi^2 = 1-\alpha \quad .$$

The curves $\frac{1}{n} C_n$ versus n are plotted in Figure IV for seven different significance levels. When α is small $\frac{1}{n} C_n$ decreases with n . When α is very large $\frac{1}{n} C_n$ increases with n . When $\alpha = .25$ the function $\frac{1}{n} C_n$ is very nearly a straight line. (Actually it is concave downward with its maximum at n = 2 .) Now if $\frac{1}{n} C_n$ could be well approximated by a constant, as is true for $\alpha = .25$ and for other significance levels in certain ranges of n, finding critical points for S(z) would be very easy. Note that the critical function for S(z) becomes $\sigma^2 K_n = \sigma^2 \frac{r-1}{n} C_{n-1} = \sigma^2 C_{r-1}$ because if $\frac{1}{n} C_n$ is constant for all n we may use n = (r-1) and it is equal to $\frac{1}{r-1} C_{r-1}$.

Hence if the desired significance level for the test is .25, as might be the case if the more serious testing error is to fail to recognize that the blocks are different, that is, Type I errors are tolerated to decrease the probability of Type II errors, then S(z) can be evaluated against the straight line $\sigma^2 C_{r-1}$. Where $S(z_0) > \sigma^2 C_{r-1}$, the point z_0 can be considered to belong to the set Ω_S , and where $S(z_0) < \sigma^2 C_{r-1}$, the point z_0 can be considered to belong to the complementary set $\overline{U} - \Omega_S$.

Unfortunately the function $\frac{1}{n} C_n$ is not so well approximated by a constant for small values of α, and when n is small the variation can be considerable. To evaluate the amount of error introduced by using $\frac{1}{r-1} C_{r-1}$ rather than the correct $\frac{1}{n} C_n$. Note that

$$\sigma^2 (r-1)\frac{1}{n} C_n = \sigma^2 (r-1)\frac{1}{r-1} C_{r-1} + \sigma^2 (r-1)\left[\frac{1}{n} C_n - \frac{1}{r-1} C_{r-1}\right] .$$

For small values of α the quantity in brackets is always positive since $n \leq r-1$, and it increases as n decreases. To find any bound on this error it is necessary to find out how small n can become.

From a theoretical standpoint it should be pointed out that α values of .50 or .75 will make the quantity in brackets negative, but the curves are still monotonic and the maximum error would occur when n is at its minimum value. From a practical point of view, it is difficult to imagine any situation with such a large α where one would not be satisfied to approximate $\frac{r-1}{n} C_n$ by C_{r-1} .

In Chapter V it was shown that n(z) is bounded below by 1 . This lower bound is not dependent upon the r, t, and x_{ij}'s of the experiment. When those factors are incorporated, a much higher lower-bound for n(z) can be obtained. Ignoring the practical limitation which the x_{ij}'s place on z, a lower limit for n can be achieved by letting z approach infinity since n(z) is a monotonic decreasing function of $|z|$. Note that

$$\lim_{|z| \to \infty} n(z) = \lim_{|z| \to \infty} (r-1) \left[\frac{(z^2\bar{\gamma} + \frac{1}{t})^2}{(z^2\bar{\gamma} + \frac{1}{t})^2 + z^4\delta} \right] = r-1 \frac{\bar{\gamma}^2}{\bar{\gamma}^2 + \delta} = (r-1)\frac{\bar{\gamma}^2}{\bar{\bar{\gamma}}}$$

That is

$$n(\infty) = \frac{\left(\sum \gamma_i\right)^2}{\sum \gamma_i^2}$$

where

$$\sum_i^{r-1} \gamma_i = tr(I - \frac{1}{r} J)R^{-1}(I - \frac{1}{r} A'A)^{-1}R^{-1}(I - \frac{1}{r} J)$$

and

$$\sum_i^{r-1} \gamma_i^2 = tr[(I - \frac{1}{r} J)R^{-1}(I - \frac{1}{r} A'A)^{-1}R^{-1}(I - \frac{1}{r} J)]^2 \quad .$$

Hence,

$$1 < n(\infty) < n(z_p) < n(z) \leq r-1 \quad ,$$

where z_p is a practical upper bound on $|z|$ imposed by the values of x_{ij} used in the experiment.

Referring again to Figure IV, with small α the function $C_{n(z)}$ is a monotonic decreasing function of n(z). So since n(z) is a monotonic decreasing function of $|z|$, the critical function for S(z) will be a monotonic <u>increasing</u> function of $|z|$.

The $S(z)$ curve will be shaped as in the equal eigenroot case of $T(z)$, however, the critical curve for $S(z)$ which shall be called $\sigma^2 K_{n(z)}$ is not a straight line, as is $\sigma^2 C_{r-1}$. The curve $\sigma^2 K_{n(z)}$ is symmetric and when α is not large, it is concave upward being asymptotic to the horizontal line represented by $\sigma^2 C_{n(\infty)}$. See Figure i.

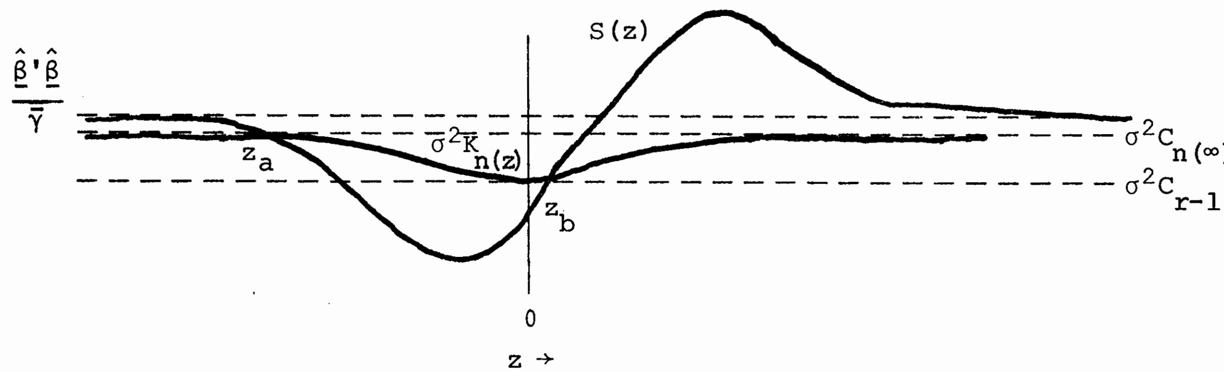


Figure i.

It should be pointed out that the critical curve is always symmetric, bounded and asymptotic to the line corresponding to $n(\infty)$, but its shape need not be that of Figure i. The only exception to its being concave upward (small α) or concave downward (very large α) is the case of $\alpha = .25$. In that case, the curve will be bounded above by C_b where b is the maximum of 2 and $n(\infty)$. In other words, when using $\alpha = .25$ if $n(\infty) < 2$ find an upper bound for the critical function with 2 rather than $n(\infty)$.

Since the critical curve based on a χ^2 statistic is virtually a straight line when $\alpha = .25$, the point may seem trivial. However, when σ^2 is not known and use is made of a F-statistic, the shapes show a little more variation as the degrees of freedom for error change. The χ^2 statistic discussed here is equivalent to the F-statistic situation with infinite degrees of freedom for error. Obviously, unless the degrees of freedom for error are very small, the situation will be similar to that illustrated in Figure i. The situation is discussed in Chapter VIII and is only brought up at this time to warn the reader that the critical function for an S(z) statistic, though bounded, may not always have the same shape.

Solving for the points z_a and z_b in Figure i, the intersections of S(z) and $\sigma^2 K_{n(z)}$ can be done iteratively, since the intersections of S(z) with any horizontal line are found easily. That there might be more than two intersections is discussed in Example 3 below.

Consider a few example problems. Reconsider Example 1, previously discussed in Chapters III and IV.

Example 1, Figure V.

$$\sigma^2 = 1,\ \alpha = .10,\ r = 3,\ t = 5,\ \gamma_1 = \frac{1}{5},\ \gamma_2 = 1,\ \bar{\gamma} = \frac{3}{5},$$

$$\sum \gamma_i^2 = \frac{26}{25},\ \bar{\bar{\gamma}} = \frac{13}{25},\ \delta = \frac{13}{25} - \left(\frac{3}{5}\right)^2 = \frac{4}{25},$$

$$(\hat{\underline{\kappa}} + z\hat{\underline{\beta}})'(\hat{\underline{\kappa}} + z\hat{\underline{\beta}}) = (\hat{\underline{\kappa}}'\underline{v}_1 + z\hat{\underline{\beta}}'\underline{v}_1)^2 + (\hat{\underline{\kappa}}'\underline{v}_2 + z\hat{\underline{\beta}}'\underline{v}_2)^2 = 2z^2 + z + 1.25 \quad .$$

Let

$$S_1(z) = \frac{(\hat{\underline{\kappa}} + z\hat{\underline{\beta}})'\underline{v}_1\underline{v}_1'(\hat{\underline{\kappa}} + z\hat{\underline{\beta}})}{z^2\bar{\gamma} + \frac{1}{t}} = \frac{z^2 - z + .25}{.6z^2 + .2}$$

and let

$$S_2(z) = \frac{(\hat{\underline{\kappa}} + z\hat{\underline{\beta}})'\underline{v}_2\underline{v}_2'(\hat{\underline{\kappa}} + z\hat{\underline{\beta}})}{z^2\bar{\gamma} + \frac{1}{t}} = \frac{z^2 + 2z + 1}{.6z^2 + .2} \quad .$$

So that

$$S(z) = \frac{(\hat{\underline{\kappa}} + z\hat{\underline{\beta}})'(\hat{\underline{\kappa}} + z\hat{\underline{\beta}})}{z^2\bar{\gamma} + \frac{1}{t}} = \frac{2z^2 + z + 1.25}{.6z^2 + .2} = S_1(z) + S_2(z)$$

$$\frac{n(z)}{r-1} = \frac{(z^2\bar{\gamma} + \frac{1}{t})}{(z^2\bar{\gamma} + \frac{1}{t})^2 + z^4\delta} = \frac{(.6z^2 + .2)^2}{(.6z^2 + .2)^2 + .16z^4}$$

$$\frac{n(0)}{r-1} = 1$$

$$\frac{n(\infty)}{r-1} = \frac{(.6)^2}{(.6)^2 + .16} = \frac{.36}{.52} = .693$$

$$\frac{n(\pm 1)}{r-1} = \frac{.64}{.64 + .16} = .8$$

$$\frac{n(\pm .5)}{r-1} = \frac{.1215}{.1215 + .01} = .925$$

$$C_{n(\infty)} = C_{2(.693)} = C_{1.386} \doteq 3.4 \qquad K_{n(\infty)} = \frac{r-1}{n(\infty)} C_{n(\infty)} \doteq 4.9$$

$$C_{n(\pm 1)} = C_{2(.8)} = C_{1.6} \doteq 3.8 \qquad K_{n(\pm 1)} = \frac{r-1}{n(\pm 1)} C_{n(\pm 1)} \doteq 4.75$$

$$C_{n(\pm .5)} = C_{2(.925)} = C_{1.85} \doteq 4.3 \qquad K_{n(\pm .5)} = \frac{r-1}{n(\pm .5)} C_{n(\pm .5)} \doteq 4.65$$

$$C_{n(0)} = C_2 = 4.60 \qquad K_{n(0)} = C_{r-1} = 4.60$$

The component functions $S_1(z)$ and $S_2(z)$ are graphed along with $S(z)$ in Figure V only to allow comparison with $T_1(z)$ and $T_2(z)$. It is never necessary to break $S(z)$ into r-1 parts.

Although $K_{n(z)}$, the critical function for $S(z)$, is not a horizontal line through the point $(0, C_{r-1})$, the error incurred in such an assumption is small in this example. Let

$$S(z') = \sigma^2 K_{n(0)} = 4.6 = \frac{2z'^2 + z' + 1.25}{.6z'^2 + .2}$$

$$z' = \frac{1 \pm \sqrt{2.0032}}{1.52} = 1.59, -.27$$

Let

$$S(z'') = \sigma^2 K_{n(\infty)} = 4.9 = \frac{2z''^2 + z'' + 1.25}{.6z''^2 + .2}$$

$$z'' = \frac{1 \pm \sqrt{2.015}}{1.88} = 1.29, -.22 \quad .$$

With only two calculations we determine that

$$(-.22, 1.29) \subset \Omega_S \subseteq (-.27, 1.59)$$

where the symbol Ω_S is used to represent the set of z's such that $S(z) > \sigma^2 K_{n(z)}$. Whether or not more precision is justifiable considering the approximate nature of the distribution of S(z) may be subject to debate. If more precision is desired in the upper end point of the interval, one might solve for $K_{n(1.4)}$ and S(1.4).

$$\frac{n(1.4)}{r-1} = \frac{1.89}{1.89 + .615} = .753$$

$$C_{n(1.4)} = C_{1.506} = 3.67 \qquad K_{n(1.4)} = \frac{3.67}{.753} = 4.88$$

$$S(1.4) = \frac{2(1.96) + 1.4 + 1.25}{1.176 + .2} = 4.78 < K_{n(1.4)} = 4.88$$

and

$$S(1.29) > K_{n(1.29)} \quad .$$

Hence,

$$(-.22, 1.29) \subset \Omega_S \subset (-.27, 1.4) \quad .$$

In a similar manner any desired degree of precision can be attained with regard to either end point of the interval of the z's which constitutes Ω_S .

Example 1a.

Assume now that $\gamma_1 = \frac{2}{5}$ and $\gamma_2 = \frac{4}{5}$, maintaining the $\bar{\gamma} = \frac{3}{5}$ of Example 1, but having a ratio of 1:2 rather than the more extreme 1:5 ratio of Example 1.

Then letting

$$S(z) = \sigma^2 C_{r-1} = C_{r-1}$$

yields the same interval (-.27, 1.59). But

$$\delta = \frac{10}{25} - \frac{9}{25} = \frac{1}{25} ,$$

so that

$$\frac{n(\infty)}{r-1} = \frac{.36}{.40} = .9, \; C_{1.8} = 4.226, \; \sigma^2 K_{n(\infty)} = \frac{4.226}{.9} = 4.7 \quad .$$

Setting

$$S(z') = 4.7 \qquad z' = -.256, \; 1.47 \quad .$$

And if $z^2 < 1$ from theoretical or practical considerations, we would look, not to $n(\infty)$ for an inner limit, but to $n(\pm 1)$.

$$\sigma^2 K_{n(\pm 1)} = 4.66 = S(z'')$$

$$z'' = \frac{1 \pm \sqrt{1 + 1.0125}}{1.592} = -.262, \; 1.52 \quad .$$

And due to $z^2 < 1$, we would state only that z values greater than -.26 or -.27 result in significant block differences. Greater precision can be achieved, of course, from iteration.

Example 2. Figure VI.

$$\sigma^2 = 1,\ r = 5,\ t = 4,\ \gamma_1 = \frac{1}{4},\ \gamma_2 = \frac{1}{2},\ \gamma_3 = \frac{3}{4},\ \gamma_4 = 1,$$

$$\bar{\gamma} = \frac{5}{8},\ \delta = \frac{30}{64} - \frac{25}{64} = .078$$

$$\hat{\underline{\beta}}'\hat{\underline{\beta}} = 6,\ \hat{\underline{\kappa}}'\hat{\underline{\kappa}} = 1,\ \hat{\underline{\kappa}}'\hat{\underline{\beta}} = -\frac{3}{2} \quad .$$

Then if $\alpha = .25$,

$$S(z') = \frac{6z'^2 - 3z' + 1}{\frac{5}{8}z'^2 + \frac{1}{4}} = \sigma^2 C_{r-1,.25} = 5.385$$

$$z' = -.10,\ +1.24$$

$$S(z'') = \frac{6z''^2 - 3z'' + 1}{\frac{5}{8}z''^2 + \frac{1}{4}} = \sigma^2 K_{n(\infty)} = \frac{1}{.833} C_{3.33} = \frac{4.534}{.833} = 5.44$$

$$z'' = -.11,\ +1.26$$

$$\overline{\cup} - (-.10, +1.24) \subseteq \Omega_S \subset \overline{\cup} - (-.11, +1.26) \quad .$$

And if $\alpha = .05$,

$$S(z') = \frac{6z'^2 - 3z' + 1}{\frac{5}{8}z'^2 + \frac{1}{4}} = \sigma^2 C_{r-1,.05} = 9.488$$

$$z' = (-.50,\ 47.5)$$

$$S(z'') = (-10.1,\ -.53)$$

$$(-10.1, -.53) \subset \Omega_S \subseteq \overline{\cup} - (-.50, 47.5)$$

Then if $-z_p < z < z_p$ where z_p is a theoretical or practical limit on z and $-10.1 < -z_p < z_p < 47.5$

$$(-z_p, -.53) \subset \Omega_S \subseteq (-z_p, -.50) \quad .$$

It has been assumed, thus far, that $K_{n(z)}$ and $S(z)$ intersect at only two points. It should be clear that this will be the case unless the two curves closely parallel each other where $K_{n(z)}$ has the most slope. The following counter example is given.

Example 3. Figure VII.

$$r = 7,\ t = 3,\ \bar{\gamma} = \frac{1}{2},\ \delta = \frac{5}{6},\ \sigma^2 = 1,\ \alpha = .05,$$

$$\hat{\underline{\beta}}'\hat{\underline{\beta}} = 10,\ \hat{\underline{\kappa}}'\hat{\underline{\kappa}} = 4,\ \hat{\underline{\kappa}}'\hat{\underline{\beta}} = 0,$$

$$S(z) = \frac{10z^2 + 4}{\frac{1}{2}z^2 + \frac{1}{3}}$$

$$S(0) = 12 \quad < \quad K_{n(0)} = 12.6$$

$$S(\pm.5) = 14.2 > K_{n(\pm.5)} = 13.5$$

$$S(\infty) = 20 \quad < \quad K_{n(\infty)} = 20.2$$

This is a valid mathematical counter example to a contention that Ω_S must be of single interval form, however, two practical points should be made. First, the δ value for this example is very, very high. This means that the roots are very unequal, possibly 5 roots equal to .0925 and one root equal to 2.54 . With this much difference in the roots it is doubtful that the approximation of the distribution is very good. And, secondly, little real error would be introduced in this problem by reporting any Ω_S whatsoever.

Summary of the Special Case of Equal Covariate Block Means.

Letting $x_{\cdot j} = x_{\cdot j'}$ for all blocks, $j', j = 1, 2, \cdots, r$, to simplify the mathematics, we have found two tests, or criteria, for generating sets,

Ω_k, of z values for which blocks are significantly different. By adding $x_{\cdot j}$ to all z values we form the sets of all x values for which blocks are significantly different. That is, if $.67 \varepsilon \Omega_T$, then $(.67 + x_{\cdot j}) \varepsilon \Gamma_T$. And if $.67 \varepsilon \Omega_S$ then $(.67 + x_{\cdot j}) \varepsilon \Gamma_S$.

The sets Γ_T and Ω_T are found using the statistic $T(z)$, a sum of r-1 fractions whose numerators would add neatly but whose denominators contain the number γ_i which may differ from term to term; and $T(z) \sim \sigma^2 \chi^2_{r-1}$.

The sets Γ_S and Ω_S are found with the statistic $S(z)$ which averages the γ_i's, substitutes $\bar{\gamma}$ for each γ_i and adds the terms. $S(z)$ is approximately distributed as $\sigma^2 \frac{r-1}{n(z)} \chi^2_{n(z)}$, where $n(z)$ is a correction factor based on $|z|$, the differences between roots, and t.

Let us speak now of a third statistic, $R(z)$, which equals $S(z)$, but $R(z) \mathrel{\dot{\sim}} \sigma^2 \chi^2_{r-1}$. Naturally this approximation is rougher than the one for $S(z)$ since we know that the first two moments do not exactly fit, unless $n(z) = r-1$. Then the sets Ω_R and Γ_R will be generated by $R(z)$. The use of the $R(z)$ approximation derives as a simplification of $T(z)$ by the "unconscious" mathematical manipulation of averaging denominators to add numerators. Likewise $R(z)$ derives as a simplification of $S(z)$ by letting $n(z) = r-1$. How good an approximation Γ_R will be of Γ_T when the roots of $(I - \frac{1}{r} J)R^{-1}(I - \frac{1}{r} A'A)^{-1}R^{-1}(I - \frac{1}{r} J)$ differ cannot be determined without a great deal of calculation. But limits can be placed on the approximation of Γ_R to Γ_S by solving z' when $S(z') = \sigma^2 K_{n_m}$ where n_m is the value of $n(z)$ which results in the maximum (or possibly minimum) value of K_n over the range of z. [In general, the limit on the approximation is found by letting $S(z') = \sigma^2 K_{n(\infty)}$.]

The sets Γ_T, Γ_S, and Γ_R are obtainable from Ω_T, Ω_S, and Ω_R only when $x_{\cdot j} = x_{\cdot j'}$ for all blocks. Finding Γ_T, Γ_S, and Γ_R in the general case requires some additional calculation.

CHAPTER VII

THE GENERAL CASE

1. Development of a Parametric Statement of the Null Hypothesis

As in the special case, the experimental model can be written as

$$y_{ij} = \mu + \theta_i + \kappa_j + \beta_j r_j \frac{(x_{ij} - x_{\cdot j})}{\sqrt{\sum (x_{ij} - x_{\cdot j})^2}} + \psi r_j \frac{(x_{ij} - x_{\cdot j})}{\sqrt{\sum (x_{ij} - x_{\cdot j})^2}} + \varepsilon_{ij}$$

$$i = 1, 2, \cdots , t; \ j = 1, 2, \cdots , r; \ r_j = \sqrt{\sum_i (x_{ij} - x_{\cdot j})^2} \quad \sum_i \theta_i = 0 \quad \sum_j \kappa_j = 0$$

however, the κ_j used here may appear to be different from the κ_j of Chapter II in terms of the parameters of Chapter I. That is, here $\kappa_j = \nu_j + \phi_j x_{\cdot j} + \psi(x_{\cdot j})$, whereas in the special case $\kappa_j = \nu_j + \psi(x_{\cdot j})$. However, since the special case uses the additional inestimable equation $\sum_j \phi_j = 0$ to solve the normal equations, it should be evident that the previous κ_j is, indeed, a "special case" of the κ_j defined here.

When $X = x_0$ the variation in $E(\underset{\sim}{y})$ from block to block is contained in the expression $\kappa_j + (x_0 - x_{\cdot j})\beta_j + (x_0 - x_{\cdot j})\psi$. From this expression we can eliminate $x_0\psi$, since it is the same in all blocks, and for mathematical convenience we add $x_{\cdot\cdot}\psi$. Then the expression incorporating all block differences becomes

$$\kappa_j + (x_0 - x_{\cdot j})\beta_j - (x_{\cdot j} - x_{\cdot\cdot})\psi \quad .$$

Let

$$P_j(x_0) = (x_0 - x_{\cdot j})\beta_j - (x_{\cdot j} - x_{\cdot\cdot})\psi \quad .$$

The notation $P_j(x_0)$ is intended to imply that P_j is a function of x_0, not a product.

By taking the inestimable condition $\sum_j^r (x_0 - x_{\cdot j})\beta_j = 0$ to distinguish block regression effects from the mean regression effect, we see that $\sum_j^r P_j(x_0) = 0$. It should again be evident that the special case condition $\sum_j \beta_j = 0$ is included in the condition $\sum_j P_j(x_0) = 0$ when $x_{\cdot j} = x_{\cdot j'}$ for all j and j' .

As before $\hat{\kappa}_j = y_{\cdot j} - y_{\cdot\cdot}$ and $\hat{\underline{\kappa}} \sim N[\underline{\kappa}, \frac{\sigma^2}{t}(I - \frac{1}{r} J)]$. Just as in the special case, we find that

$$\hat{\underline{\beta}} + \hat{\psi}\underline{1} = R^{-1}(I - \frac{1}{r} A'A)^{-1}A_d'[I - \frac{1}{r}(I_t \otimes J_r)\underline{Y} \quad .$$

Then

$$\underline{P}(x_0) = (I - \frac{1}{r} J)[x_0 I - \text{diag}(x_{\cdot j})](\underline{\beta} + \underline{1}\psi) = [x_0 I - \text{diag}(x_{\cdot j})]\underline{\beta} - [(\text{diag } x_{\cdot j}) - x_{\cdot\cdot} I]\underline{1}\phi$$

and

$$\hat{\underline{P}}(x_0) = (I - \frac{1}{r} J)[x_0 I - \text{diag}(x_{\cdot j})](\hat{\underline{\beta}} + \underline{1}\hat{\psi}) \quad ,$$

from which we see that

$$\hat{\underline{\kappa}} + \hat{\underline{P}}(x_0) \sim N\left\{\underline{\kappa} + \underline{P}(x_0),\ \sigma^2(I - \frac{1}{r} J)\left(\frac{1}{t} I + [x_0 I\text{-diag}(x_{\cdot j})]M[x_0 I\text{-diag}(x_{\cdot j})]\right)(I - \frac{1}{r} J)\right\}$$

where

$$M = R^{-1}(I - \frac{1}{r} A'A)^{-1}R^{-1} \quad .$$

Let

$$W_{x_0} = (x_0 I - D) M (x_0 I - D)$$

or

$$W_{x_0} = x_0^2 M - x_0 [DM + MD] + DMD \quad ,$$

where $D = \text{diag}(x_{\cdot j})$. (Note that W_{x_0} reduces to $(x_0 - x_{\cdot j})^2 M$ in the special case.) Then

$$\hat{\underline{\kappa}} + \hat{\underline{P}}(x_0) \sim N[\underline{\kappa} + \underline{P}(x_0), \sigma^2 (I - \frac{1}{r} J)(\frac{1}{t} I + W_{x_0})(I - \frac{1}{r} J)] \quad .$$

The hypothesis of no block difference when $X = x_0$ becomes

$$H_o: \ \underline{\kappa} + \underline{P}(x_0) = \underline{0}$$

$$\text{vs.} \quad H_a: \ \underline{\kappa} + \underline{P}(x_0) \neq \underline{0}$$

2. An Exact Test

Again, an exact χ^2 statistic can be found to test this hypothesis.

$$\frac{1}{\sigma^2} T_\gamma(x_0) = \frac{1}{\sigma^2}[\hat{\underline{\kappa}} + \hat{\underline{P}}(x_0)]'[\frac{1}{t} I + (I - \frac{1}{r} J) W_{x_0} (I - \frac{1}{r} J)]^{-1}[\hat{\underline{\kappa}} + \hat{\underline{P}}(x_0)] \sim \chi^2_{r-1,\lambda}$$

where

$$\lambda = \frac{1}{2\sigma^2}[\underline{\kappa} + \underline{P}(x_0)]'[\frac{1}{t} I + (I - \frac{1}{r} J) W_{x_0} (I - \frac{1}{r} J)]^{-1}[\underline{\kappa} + \underline{P}(x_0)] \quad ,$$

and $\lambda = 0$ under the hypothesis H_o .

Unfortunately when $(I - \frac{1}{r} J)(W_{x_0})(I - \frac{1}{r} J)$ is expanded in eigenvectors the vectors are not independent of x_0 as in the special case. To emphasize this point the subscript x will be used for the roots and vectors. The functional notation $\underline{v}_i(\underline{x})$ becomes cumbersome as does the retention of the notation x_0 which has been used to emphasize the fact that the test is made at a specific point and then the inquiries are made as to which X values would cause the test statistic to be higher than some critical function. Hopefully this point has been made well enough to allow the terminology $\underline{v}_{xi}$, and γ_{xi} to adequately communicate the idea that these roots and vectors are dependent on a specific value of X at which the test is to be made. Since

$$(I - \frac{1}{r} J)W_x(I - \frac{1}{r} J) = \sum_{i=1}^{r-1} \gamma_{xi}\underline{v}_{xi}\underline{v}'_{xi}$$

then

$$\frac{1}{\sigma^2} T_\gamma(x) = \frac{1}{\sigma^2} [\hat{\underline{K}} + \hat{\underline{P}}(x)]' \sum_{i=1}^{r-1} \left[\frac{\underline{v}_{xi}\underline{v}'_{xi}}{\gamma_{xi} + \frac{1}{t}}\right] [\hat{\underline{K}} + \hat{\underline{P}}(x)] \quad .$$

It should be apparent that it is possible for $\gamma_{xi} = \gamma_{xi'}$ for all i and i' and for all x values only in the special case previously considered.

Hence, solving for Γ_T, the set of x's for which $T_\gamma(x) > \sigma^2 C_{r-1,\alpha}$ becomes a trial and error task which generally involves the inversion of the matrix

$$\frac{1}{t} I + (I - \frac{1}{r} J)[x^2 M - x(DM + MD) + DMD](I - \frac{1}{r} J)$$

at each trial.

Recall that we cannot assume that Γ_T is of the interval or interval complement form so that finding two end points of intervals in Γ_T does not mean we have found all of Γ_T .

Obviously the T_γ statistic has the theoretical disadvantages (resulting from the uneven weighting of contrasts) that the T statistic was found to have. But in addition, the determination of Γ_T will be tedious and expensive. If there was reason to seek a second test in the special case, there is even more reason in the general case.

3. An Approximate Test

Again an approximate test for the hypothesis

$$H_o:\ \underline{\kappa} + \underline{P}(x_0) = 0$$

$$\text{vs.}\ \ H_a:\ \underline{\kappa} + \underline{P}(x_0) \neq 0$$

can be formulated, which depends on $[\underline{\kappa} + \underline{P}(x_0)]'[\underline{\kappa} + \underline{P}(x_0)]$.

Let

$$S^*_\gamma(x_0) = [\hat{\underline{\kappa}} + \hat{\underline{P}}(x_0)]'[\hat{\underline{\kappa}} + \hat{\underline{P}}(x_0)] = \underline{y}'[H + L(x_0)]'[H + L(x_0)]\underline{y}$$

where

$$H = \frac{1}{t}\,\underline{1} \otimes (I_r - \frac{1}{r}J)$$

and

$$L(x_0) = (I - \frac{1}{r}J)(x_0 I - D)MRA_d'(I_{rt} - [I_t \otimes \frac{1}{r}J_r]) \quad .$$

$T^*_\gamma(x_0)$ is a central χ^2 under the null hypothesis. $S^*_\gamma(x_0)$ is not a χ^2 but it is "central" in the sense that $\underline{\mu}'[H + L(x_0)]'[H + L(x_0)]\underline{\mu} = [\underline{\kappa} + \underline{P}(x_0)]'[\underline{\kappa} + \underline{P}(x_0)] = 0$ only when $\underline{\kappa} + \underline{P}(x_0) = 0$ and again $\underline{\mu} = E(\underline{y})$.

Like $S^*(z)$ the statistic $S^*_\gamma(x)$ can be decomposed as a sum of independent χ^2 statistics. Let

$$[H + L(x_0)]'[H + L(x_0)] = \sum_{i=1}^{r-1} \xi_{xi}\underline{\ell}_{xi}\underline{\ell}_{xi}'$$

where

$$\underline{\ell}_{xi}'\underline{\ell}_{xi'} = 1 \qquad i = i'$$

$$= 0 \qquad i \neq i'$$

As in the special case we note that the r-1 non-zero ξ_{xi}'s are identical to the non-zero $(\frac{1}{t} + \gamma_{xi})$'s, thus

$$[H + L(x_0)][H + L(x_0)]' = \frac{1}{t}(I - \frac{1}{r} J) + (I - \frac{1}{r} J)W_{x_0} (I - \frac{1}{r} J) = \sum_{i=1}^{r-1} (\gamma_{xi} + \frac{1}{t}) \underline{v}_{xi}\underline{v}'_{xi} ,$$

$$W_{x_0} = (x_0 I - D) M(x_0 I - D) .$$

Hence,

$$S_\gamma^*(x) = \underline{Y}' \left[\sum_{i}^{r-1} (\gamma_{xi} + \frac{1}{t}) \underline{\ell}_{xi}\underline{\ell}'_{xi} \right] \underline{Y}$$

and

$$\frac{1}{\sigma^2} \underline{Y}'\underline{\ell}_{xi}\underline{\ell}'_{xi}\underline{Y} \sim \chi^2_{1,\lambda_i}$$

where $\lambda_i = 0$ under the null hypothesis.

As before the distribution of $S_\gamma^*(x)$ under the null hypothesis can be approximated as a factor times a χ^2 with m degrees of freedom. Now

$$E[S_\gamma^*(x)] = \sigma^2 \sum_{i=1}^{r-1} (\gamma_{xi} + \frac{1}{t})$$

and

$$\mathrm{Var}[S_\gamma^*(x)] = 2\sigma^2 \sum (\gamma_{xi} + \frac{1}{t})^2 .$$

So

$$E\left[\frac{S_\gamma^*(x) m(x)}{\sigma^2 \sum (\gamma_{xi} + \frac{1}{t})} \right] = m(x)$$

and we wish to choose m(x) such that

$$\mathrm{Var}\left[\frac{S_\gamma^*(x) m(x)}{\sigma^2 \sum (\gamma_{xi} + \frac{1}{t})} \right] = 2m(x)$$

or

$$\frac{m^2(x)\,2\sigma^4 \sum(\gamma_{xi} + \frac{1}{t})^2}{\sigma^4\left[\sum \gamma_{xi} + \frac{1}{t}\right]^2} = 2m(x) \quad .$$

Thus

$$m(x) = \frac{\left[\sum_i^{r-1} (\gamma_{xi} + \frac{1}{t})\right]^2}{\sum_i^{r-1} (\gamma_{xi} + \frac{1}{t})^2} = r-1\left[\frac{(\bar{\gamma}_x + \frac{1}{t})^2}{(\bar{\gamma}_x + \frac{1}{t})^2 + (\bar{\bar{\gamma}}_x - \bar{\gamma}_x^2)}\right] = r-1\left[\frac{(\bar{\gamma}_x + \frac{1}{t})^2}{(\bar{\gamma}_x + \frac{1}{t})^2 + \delta_x}\right]$$

where

$$\bar{\gamma}_x = \frac{1}{r-1} \sum_i^{r-1} \gamma_{xi}$$

$$\bar{\bar{\gamma}}_x = \frac{1}{r-1} \sum_i^{r-1} \gamma_{xi}^2$$

$$\delta_x = \bar{\bar{\gamma}}_x - \bar{\gamma}_x^2$$

$$0 \leq \delta_x < (r-2)\bar{\gamma}_x^2 \quad ,$$

hence

$$1 < m(x) \leq r-1 \quad .$$

Let

$$S_\gamma(x) = \frac{S_\gamma^*(x)}{\bar{\gamma}_x + \frac{1}{t}} \quad ,$$

then

$$\frac{1}{\sigma^2} S_\gamma(x) \;\dot{\sim}\; \frac{r-1}{m(x)} \chi^2_{m(x)}$$

recalling that

(1) $$S_\gamma(x) = \frac{[\hat{\underline{\kappa}} + \hat{\underline{P}}(x)]'[\hat{\underline{\kappa}} + \hat{\underline{P}}(x)]}{\bar{\gamma}_x + \frac{1}{t}}$$

$S_\gamma(x) = T_\gamma(x)$ only for those x's, if any, such that $\gamma_{xi} = \gamma_{xi'}$ for all i and i'. Now

$$(r-1)\bar{\gamma}_x = \text{trace of}(I - \frac{1}{r}J)W_x(I - \frac{1}{r}J) > 0$$

$$= tr(I - \frac{1}{r}J)[x^2M - x(MD + DM) + DMD](I - \frac{1}{r}J)$$

$$= x^2\ tr(I - \frac{1}{r}J)M - 2x\ tr(I - \frac{1}{r}J)MD + tr(I - \frac{1}{r}J)DMD \quad .$$

Let

$$\bar{\gamma}_x = \lambda_2 x^2 + \lambda_1 x + \lambda_0 \quad ,$$

where

$$\lambda_2 = \frac{1}{r-1}\, tr(I - \frac{1}{r}J)M$$

$$\lambda_1 = -\frac{2}{r-1}\, tr(I - \frac{1}{r}J)MD$$

$$\lambda_0 = \frac{1}{r-1}\, tr(I - \frac{1}{r}J)DMD \quad ;$$

and

$$\hat{\underline{P}}(x) = (I - \frac{1}{r}J)(xI - D)Gy$$

where

$$G = R^{-1}(I - \frac{1}{r}A'A)A_d'[I - \frac{1}{r}(I_t \otimes J_r)]$$

$$\hat{\underline{P}}(x) = x(I - \frac{1}{r}J)G\underline{y} - (I - \frac{1}{r}J)DG\underline{y} \quad .$$

Let

$$\hat{\underline{\Pi}}_1 = (I - \frac{1}{r}J)G\underline{y}$$

and

$$\underline{\hat{\Pi}}_0 = -(I - \frac{1}{r} J) DG\underline{y}$$

thus

$$\underline{\hat{P}}(x) = \underline{\hat{\Pi}}_1 x + \underline{\hat{\Pi}}_0 \quad .$$

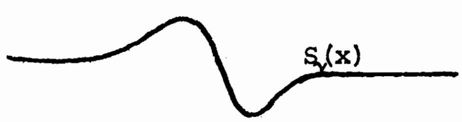
Then

$$(2) \qquad S_\gamma(x) = \frac{x^2 \underline{\hat{\Pi}}_1' \underline{\hat{\Pi}}_1 + 2x\underline{\hat{\Pi}}_1'(\underline{\hat{\Pi}}_0 + \underline{\hat{\kappa}}) + (\underline{\hat{\Pi}}_0 + \underline{\hat{\kappa}})'(\underline{\hat{\Pi}}_0 + \underline{\hat{\kappa}})}{x^2\lambda_2 + x\lambda_1 + (\lambda_0 + \frac{1}{t})} \quad .$$

When $S_\gamma(x)$ is expressed in the first form (1), it is difficult to see how values of x affect the value of $S_\gamma(x)$. However, in the second form (2), it is apparent that $S_\gamma(x)$ is again a ratio of two quadratic functions in x.

Again $S_\gamma(x)$ is continuous and has at most two extreme points, although those two points can no longer be expected to lie on either side of 0 as with S(z).

Looking at $\frac{\partial S_\gamma(x)}{\partial x}$ it can be seen that $S_\gamma(x)$ takes its "shape" from the sign of $Q = \lambda_1 \underline{\hat{\Pi}}_1' \underline{\hat{\Pi}}_1 - 2\lambda_2 \underline{\hat{\Pi}}_1'(\underline{\hat{\Pi}}_0 + \underline{\hat{\kappa}})$ rather than simply from $-\underline{\hat{\kappa}}'\underline{\hat{\beta}}$ as in the special case. If $Q > 0$, then $\frac{\partial S_\gamma(x)}{\partial x} > 0$ for very large values of $|x|$ which implies that S(x) has a maximum to the left of its minimum.

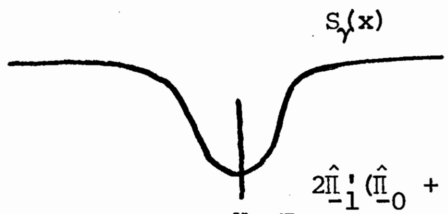


$$--- \frac{1}{\lambda_2} \underline{\hat{\Pi}}_1' \underline{\hat{\Pi}}_1$$

Likewise, if $Q < 0$,

$$--- \frac{1}{\lambda_2} \underline{\hat{\Pi}}_1' \underline{\hat{\Pi}}_1$$

And if $Q = 0$, and $\frac{1}{\lambda_2}\hat{\underline{\Pi}}_1'\hat{\underline{\Pi}}_1 > \frac{1}{\lambda_0 + \frac{1}{t}}(\hat{\underline{\Pi}}_0 + \hat{\underline{\kappa}})'(\hat{\underline{\Pi}}_0 + \hat{\underline{\kappa}}) = S_\gamma(0)$



$$x_c = \frac{2\hat{\underline{\Pi}}_1'(\hat{\underline{\Pi}}_0 + \hat{\underline{\kappa}})}{\hat{\underline{\Pi}}_1'\hat{\underline{\Pi}}_1}$$

Similarly, if $Q = 0$ and $S_\gamma(\infty) < S_\gamma(0)$, $S_\gamma(x)$ is again symmetric about x_c, but concave downward. It will be noted that

$$S_\gamma(\infty) = \lim_{x \to \infty} S_\gamma(x) = \lim_{z \to \infty} S_\gamma(z) = S_\gamma(\infty) = \frac{y'G'(I - \frac{1}{r}J)Gy}{\lambda_2} ,$$

again the test for regression effects to which Cox [1] refers.

However, $S_\gamma(0)$ is not the test statistic for fixed effects. If $x_{\cdot j} = x_{\cdot j'}$, for all j and j', then $S_\gamma(x_{\cdot j})$ would be equal to S(0). Thus $S_\gamma(x)$ is shaped much as S(x), and can be easily solved for its two intersections (if any) with any horizontal line. The general critical function $\sigma^2 \frac{r-1}{m(x)} \chi^2_{m(x)}$ is not, however, shaped as the critical function for the special case. In the special case, n(z) decreased monotonically with $|z|$ so that when $\frac{1}{n} C_n$ was monotonic with n, the critical function was monotonic with $|z|$. Here the functions $\frac{1}{n} C_n$ are the same, but m(x) is not unimodal.

Recall that

$$m(x) = (r-1) \frac{\left(\frac{1}{r-1}\sum \gamma_{xi} + \frac{1}{t}\right)^2}{\left(\frac{1}{r-1}\sum \gamma_{xi} + \frac{1}{t}\right)^2 + \frac{1}{r-1}\sum \gamma_{xi}^2 - \left(\frac{1}{r-1}\sum \gamma_{xi}\right)^2}$$

where $\sum \gamma_{xi}$ and $\sum \gamma_{xi}^2$ are the sum of roots and sum of squares of roots of the matrix $(I - \frac{1}{r} J)[x^2M - x(MD + DM) + DMD)(I - \frac{1}{r} J)$. These roots will vary with X and certain X values will produce more unbalanced roots than others. In order to evaluate the variation of m(x) with X it will be necessary to rewrite m(x) in a form which incorporates the functional dependence of the roots on X. This can be done by noting that the sum of the roots is the trace of the matrix and the sum of the squares of the roots is the trace of the square of the matrix.

$$\frac{1}{r-1} \sum \gamma_{xi} = \bar{\gamma}_x = \lambda_2 x^2 + \lambda_1 x + \lambda_0$$

where the λ_0, λ_1, and λ_2 are defined for establishing the second form (2) of $S_\gamma(x)$. Now

$$\frac{1}{r-1} \sum \gamma_{xi}^2 = \bar{\bar{\gamma}}_x = \frac{1}{r-1} \operatorname{tr}(I - \frac{1}{r} J)W_x(I - \frac{1}{r} J)W_x$$

$$W_x = x^2M - x(MD + DM) + DMD$$

$$\bar{\bar{\gamma}}_x = \frac{1}{r-1} \operatorname{tr}(I - \frac{1}{r} J)[x^2M - x(MD + DM) + DMD](I - \frac{1}{r} J)[x^2M - x(MD + DM) + DMD]$$

Let

$$(I - \frac{1}{r} J) = T$$

$$\bar{\bar{\gamma}}_x = \frac{1}{r-1}\{x^4 \operatorname{tr} TMTM - 4x^3 \operatorname{tr} TMTMD + 2x^2[\operatorname{tr} TMTDMD + \operatorname{tr} TMDT(MD + DM)]$$

$$- 4x \operatorname{tr} TMDTDMD + \operatorname{tr} TDMDTDMD\} \quad .$$

Thus

$$\bar{\bar{\gamma}} = x^4\psi_4 + x^3\psi_3 + x^2\psi_2 + x\psi_1 + \psi_0 \quad ,$$

where

$$\psi_4 = \frac{1}{r-1}\,\mathrm{tr}[(I-\tfrac{1}{r}J)M(I-\tfrac{1}{r}J)M] \qquad \psi_3 = -\frac{4}{r-1}\,\mathrm{tr}[(I-\tfrac{1}{r}J)M(I-\tfrac{1}{r}J)MD]$$

$$\psi_2 = \frac{2}{r-1}\{\mathrm{tr}[(I-\tfrac{1}{r}J)M(I-\tfrac{1}{r}J)DMD] + \mathrm{tr}[(I-\tfrac{1}{r}J)MD(I-\tfrac{1}{r}J)(MD+DM)]\}$$

$$\psi_1 = -\frac{4}{r-1}\,\mathrm{tr}[(I-\tfrac{1}{r}J)MD(I-\tfrac{1}{r}J)DMD] \qquad \psi_0 = \frac{1}{r-1}\,\mathrm{tr}[(I-\tfrac{1}{r}J)DMD(I-\tfrac{1}{r}J)DMD]$$

Then

$$m(x) = r-1\,\frac{[\lambda_2x^2+\lambda_1x+\lambda_0+\frac{1}{t}]^2}{[\lambda_2x^2+\lambda_1x+\lambda_0+\frac{1}{t}]^2+g(x)} \quad \text{where } g(x) = (\psi_4-\lambda_2^2)x^4+(\psi_3-2\lambda_1\lambda_2)x^3 + (\psi_2-\lambda_2\lambda_0-\lambda_1^2)x^2+(\psi_1+\lambda_1\lambda_0)x+(\psi_0+\lambda_0^2)$$

We note that $m(x)$ is the ratio of two expressions which are 4^{th} degree in X. As such it can have four distinct critical, or extreme, points. It is bounded above by $r-1$ but it is difficult to establish a useful lower bound. Since $\sum_i^{r-1} \gamma_{xi}^2 < \left(\sum_i^{r-1} \gamma_{xi}\right)^2$ for any x, we know that $m(x) > 1$. It was shown in Chapter VI that $\alpha = .25$, the critical function is insensitive to changes in $m(x)$ and the lower bound $m(x) = 1$ may be adequate to bound $m(x)$ and hence bound $\frac{r-1}{m(x)} C_{m(x)}$ and hence bound Γ_S, the set of x's for which blocks are significantly different. But when α is small and there are a large number of blocks, the $m(x) = 1$ bound may well be too low to be of value. The $m(x) = 1$ result is equivalent to the extreme case of all roots equal to 0 except one. That any value of X could cause this sort of distortion of the matrix W_x becomes more and more unlikely as r, the size of the matrix, becomes larger.

Although it is not in general a lower bound, it can be pointed out that $m(\infty) = n(\infty)$. That is

$$\lim_{x \to \infty} m(\infty) = r-1 \frac{\lambda_2^2}{\psi_4}$$

where

$$\lambda_2 = \frac{1}{r-1} \operatorname{tr}(I - \frac{1}{r} J)M$$

and

$$\psi_4 = \frac{1}{r-1} \operatorname{tr}[(I - \frac{1}{r} J)M]^2 \quad .$$

This should not be surprising since as x becomes increasingly large any $x_{\cdot j}$ subtracted from it would become immaterial.

Hence, $m(x)$ is asymptotic to $m(\infty)$ and the corresponding critical function is bounded by $\sigma^2 C_{r-1}$ and asymptotic to $\sigma^2 \frac{r-1}{m(\infty)} C_{m(\infty)}$. The function $m(x)$ can have a minimum lower than $m(\infty)$, though it will be greater than 1. A great deal of effort can be expended finding the intersection of the $S_\gamma(x)$ function and the critical function. Consider an example.

Example 4.

$$r = 3,\ t = 6,\ \sigma^2 = 1,\ \alpha = .05,\ m = \frac{1}{24}\begin{pmatrix} 501 \\ 060 \\ 105 \end{pmatrix}$$

$$x_{\cdot 1} = 0,\ x_{\cdot 2} = \frac{1}{2},\ x_{\cdot 3} = -\frac{1}{2},\ D = \operatorname{diag}(x_{\cdot j}) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{2} \end{pmatrix}$$

Then

$$\lambda_2 = \frac{1}{2} \operatorname{tr}(I - \frac{1}{3} J)M = \frac{1}{2} \operatorname{tr} \frac{1}{24}\left[2\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}(1\ 0\ -1) + \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}(1\ -2\ 1)\right] = \frac{5}{24} = .208$$

$$\lambda_1 = -\frac{2}{2}\,\mathrm{tr}(I - \frac{1}{3}J)MD = -\mathrm{tr}\,\frac{1}{48}\left[\begin{pmatrix}2\\0\\-2\end{pmatrix}_{(0\ \ 0\ \ +1)} + \begin{pmatrix}1\\-2\\1\end{pmatrix}_{(0\ \ -2\ \ -1)}\right] = -\frac{1}{48} = -.0208$$

$$\lambda_0 = \frac{1}{2}\,\mathrm{tr}(I - \frac{1}{3}J)DMD = \frac{1}{2}\,\mathrm{tr}\,\frac{1}{96}\,(I - \frac{1}{3}J)\begin{pmatrix}0&0&0\\0&6&0\\0&0&5\end{pmatrix} = \frac{11}{288} = .0382$$

$\psi_4 = .0452$	$\psi_4 - \lambda_2^2 = .00174$
$\psi_3 = -.0174$	$\psi_3 - 2\lambda_1\lambda_2 = -.0087$
$\psi_2 = .0318$	$\psi_2 - 2\lambda_2\lambda_0 + \lambda_1^2 = .0163$
$\psi_1 = -.00232$	$\psi_1 - 2\lambda_1\lambda_0 = -.00073$
$\psi_0 = .00163$	$\psi_0 - \lambda_0^2 = .00017$

$$\delta_x = \overline{\overline{\gamma}}_x - \overline{\gamma}_x^2 = .00174x^4 - .00864x^3 + .0163x^2 - .00073x + .00017$$

$$m(x) = \frac{2(.208x^2 - .0208x + .2049)^2}{(.208x^2 - .0208x + .2049)^2 + .00174x^4 - .00864x^3 + .0163x^2 - .00073x + .00017}$$

The function m(x) is plotted in Figure VIII (Appendix). We see that m(0) is very close to r-1, that is, 2 . This is because the $x_{.j}$'s of the example sum to 0 . The curve is asymptotic to 1.92 . However, the minimum value of the curve occurs between -1 and -2 . This means that X values between -1 and -2 cause the roots of the W_x matrix to be relatively further apart than any other X values. Such a plot may be disconcerting but its effect on the problem at hand, i.e., the finding of Γ_{S_γ}, can be very small.

In Figure VIII, above the m(x) curve, are the $\frac{r-1}{m(x)}\,C_{m(x)}$ curves for $\alpha = .05$ and $\alpha = .10$. Since σ^2 is taken equal to 1 in this example, these are the critical curves for evaluating $S_\gamma(x)$ at the .05 and .10 levels, respectively. When $\alpha = .10$ the total range of the critical function is .012 units, and clearly it would be much less for $\alpha = .25$. When $\alpha = .05$, the range is .24 units. Even if $\alpha = .05$, this slightly undulating critical function may not cause appreciable trouble.

Consider the following values for the estimates of the parameters which comprise $S_\gamma(x)$. Let

$$G\underline{y} = MRA_d'(I_{rt} - [I_t \otimes \frac{1}{r} J_r])\underline{y} = \begin{pmatrix} -1 \\ 0 \\ +1 \end{pmatrix}$$

$$\hat{\underline{\Pi}}_1 = (I - \frac{1}{3} J)G\underline{y} = \begin{pmatrix} -1 \\ 0 \\ +1 \end{pmatrix} \qquad \hat{\underline{\Pi}}_0 = -(I - \frac{1}{3} J)DG\underline{y} = \frac{1}{6}\begin{pmatrix} -1 \\ -1 \\ +2 \end{pmatrix}$$

$$\hat{\underline{\kappa}} = \begin{pmatrix} 1/2 \\ 1/2 \\ -1 \end{pmatrix} \quad \hat{\underline{\Pi}}_1'\hat{\underline{\Pi}}_0 = \frac{1}{2} \quad \hat{\underline{\Pi}}_1'\hat{\underline{\kappa}} = -\frac{3}{2} \quad \hat{\underline{\Pi}}_0'\hat{\underline{\kappa}} = -\frac{1}{2} \quad ,$$

then

$$S_\gamma(x) = \frac{x^2\hat{\underline{\Pi}}_1'\hat{\underline{\Pi}}_1 + x[2\hat{\underline{\Pi}}_1'(\hat{\underline{\Pi}}_0 + \hat{\underline{\kappa}})] + (\hat{\underline{\Pi}}_0 + \hat{\underline{\kappa}})'(\hat{\underline{\Pi}}_0 + \hat{\underline{\kappa}})}{x^2\lambda_2 + x\lambda_1 + \lambda_0 + \frac{1}{t}}$$

and

$$S_\gamma(x) = \frac{2x^2 - 2x + \frac{2}{3}}{.208x^2 - .0208x + .2049}$$

in Example 4 (see Figure IX).

As a first step in finding Γ_S we find Γ_R by equating $S_\gamma(x)$ to $\sigma^2 C_{r-1}$ for $\alpha = .05$.

$$S_\gamma(x') = 5.992 = \frac{2x'^2 - 2x' + .667}{.208x'^2 - .0280x' + .2049}$$

(Recall that r-1 = 2 .)

$$x' = 2.76, -.27$$

Noting that $S_\gamma(0) < 5.99$ and $.27 < 0 < 2.76$ we state that

$$\Gamma_R = \overline{U} - (-.27,\ 2.76) \quad .$$

Γ_R is an upper bound on Γ_S in the sense that $\Gamma_S \subset \Gamma_R$. Note that use of Γ_R means that the true α-level for the test is more than the stated .05 level.

A second step toward finding Γ_S would be to solve $S_\gamma(x") = (r-1)C_1$. Again, $\alpha = .05$.

$$S_\gamma(x") = \frac{2x"^2 - 2x" + .667}{.208x"^2 - .0208x" + .2049} = 7.68$$

$$x" = 5.05,\ -.45$$

$$\Gamma \text{ Lower Bound}_{m(x)=1} = \overline{U} - (-.45,\ 5.05) \quad .$$

The fact that this lower bound is not close to Γ_R means that it will be necessary to find m(-.27) and m(2.76). Knowing that the interval found by using m(x) = 1 is a very poor limiting value, we can expect the true set, Γ_S, to be close to Γ_R .

Finding the values of m(-.27) and m(2.76) requires the determination of the ψ_0, ψ_1, ψ_2, ψ_3, and ψ_4 values. Had they not previously been determined in order to plot Figure VIII, it would be necessary to find them at this time.

$$m(-.27) = 1.94$$

$$m(2.76) = 1.97$$

$$S_\gamma(x) = \frac{2}{1.94} C_{1.94} = 6.00 = \frac{2x^2 - 2x + \frac{2}{3}}{.208x^2 - .0208x + .2049}$$

$$x = (-.27,\ 2.77)$$

This result would be sufficient, in most cases, for us to state that $\Gamma_S \simeq \overline{U}$ -(-.27, 2.76) .

Figure IX shows the critical curve for Example 4 for α values of .25, .10, .05, .025, and .01 . These are presented in order to show the relatively small amount of difference in the α-level which would result from using Γ_R as an approximation to Γ_S . However, this is admittedly only one example and a great deal of work can be done on the matter of finding a realistic lower limit for the critical function. It must be remembered that if the eigenroots of the M matrix are all near zero except for one, the critical function could rise very nearly to $\sigma^2(r-1)C_1$. Hence, it would be very desirable to find a quick and easy lower limit for the critical function in terms of the matrix M and the diagonal matrix D.

At this point it can be said only that the use of the critical function $\sigma^2 C_{r-1}$ seems to give very good approximations to $\sigma^2 \frac{r-1}{m(x)} C_{m(x)}$ unless the M matrix has only one root appreciably different from zero.

CHAPTER VIII

UNKNOWN VARIANCE

When the error variance σ^2 must be estimated from the experimental data, an F-statistic can be formed with $T_\gamma(x)$ and MSE, or an approximate F-statistic can be formed with $S_\gamma(x)$ and MSE.

Both $T_\gamma(x)$ and $S_\gamma(x)$ are independent of MSE, since the matrix of the quadratic form of SSE is orthogonal to the $\hat{\underline{\kappa}}$, $\hat{\underline{\theta}}$, and $\hat{\underline{\beta}}$ functions of the y_{ij}'s from which $T_\gamma(x)$ and $S_\gamma(x)$ are formed.

$T_\gamma(x)$, being a true χ^2_{r-1}, can be used to find a true F-statistic, namely

$$F_T(x) = \frac{T_\gamma(x)/r-1}{MSE} \sim F_{r-1,rt-2r-t+1,\lambda}$$

where

$$\lambda = \frac{1}{2}[\underline{\kappa} + \underline{P}(x)]'[\frac{1}{t} I + (I - \frac{1}{r} J)W_x(I - \frac{1}{r} J)]^{-1}[\underline{\kappa} + \underline{P}(x)] \quad .$$

Everything that was said about $T_\gamma(x)$ can be reiterated for $F_T(x)$. Unless conditions for the special case hold, the calculation of each point of $F_T(x)$ will require inversion of an $r \times r$ matrix. The power of $F_T(x)$ will be greater against some specific alternate hypotheses than others, as determined by the eigenvectors and eigenroots of the matrix of the quadratic form.

At the cost of knowing the exact distribution of the statistic, these faults can be corrected by formulating, with $S_\gamma(x)$, an approximate central

F-statistic under the null hypothesis, namely

$$F_S(x) = \frac{S_\gamma(x)\ \frac{m(x)}{r-1}\ /\ m(x)}{MSE} = \frac{S_\gamma(x)/r-1}{MSE} \mathrel{\dot\sim} F_{m(x),q}$$

where

$$q = rt - 2r - t + 1 \quad .$$

As in the consideration of $S_\gamma(x)$, the two values of x (if any) at which $F_S(x)$ equals any critical value of an F-statistic can readily be found. However, the critical function for evaluating $F_S(x)$ is not a horizontal line although it may be nearly so in most cases. To consider the variation of the critical function with x, it is necessary to consider the variation of the critical points of a F-statistic with ν_1, the numerator degrees of freedom; and the variation of the numerator degrees of freedom $\nu_1 = m(x)$ with x .

Define $f_{\nu_1,q}$ by $\int_{f_{\nu_1,q}}^{\infty} F_{\nu_1,q} = \alpha$. Note that $C_n = f_{n,\infty}$. We have considered the variation of $f_{\nu_1,\infty}$ with ν_1 in Chapter VI. It was found that $f_{\nu_1,\infty}$ was a monotonic decreasing function function of ν_1 for small α, and a monotonic increasing function of ν_1 for α larger than .25 . When $\alpha = .25$ the largest value of $f_{\nu_1,\infty}$ is found at $\nu_1 = 2$.

In Figures X, XI and XII, values of $f_{\nu_1,q}$ are plotted against $\frac{1}{\nu_1}$. It will be noted that the slopes of these curves are "quite constant," the variation being in the same direction as before when $\nu_2 = q$ is not small. When ν_2 is 1 or 2, the critical function actually increases with ν_1 . The constant slopes of these curves indicated that linear interpolation should be done with $\frac{1}{\nu_1}$ rather than with ν_1 .

Having established that the critical function for $F_S(x)$ will be generally monotonic with m(x), the only exception being $\alpha = .25$ with a large number of degrees of freedom for error, attention is again focused on the variation of $m(x)$ with x. In the special case of equal covariate block means, $m(x)$ is monotonic with $|x - x_{\cdot j}|$, so that limits can be placed on the critical function by $m(\infty)$ and $r-1$. However, the same problem of finding a lower limit for $m(x)$ in the general case that was discussed in Chapter VII occurs here. When α is small, the critical function $f_{m(x),q}$ will reflect the shape of the $m(x)$ curve. The closeness of Γ_S to Γ_R can be established only by solving, iteratively, for the intersections of $F_S(x)$ and $f_{m(x),q}$. Consider an example.

<u>Example 4.</u> (Continued from Chapter VII).

$$r = 3,\ t = 6,\ \sigma^2 \text{ unknown, MSE} = .60,\ \alpha = .05$$

$$S_\Upsilon(x) = \frac{2x^2 - 2x + .667}{.208x^2 - .0208x + .2049}$$

$$F_S(x) = \frac{1}{2(.6)} S_\Upsilon(x) \overset{\cdot}{\sim} F_{m(x),rt-2r-t+1} \quad \text{under } H_0$$

$$f_{r-1,rt-2r-t+1} = f_{2,7} = 4.74$$

To find Γ_R solve $S_\Upsilon(x) = 1.2(4.74) = 5.69$ for x; this yields $x = -.24,\ 2.57$. Hence, $\Gamma_R = \overline{U} - (-.24,\ 2.57)$.

Next find $m(-.24)$ and $m(2.57)$.

$$m(-.24) = \frac{2[.208(-.24)^2 - .0208(-.24) + .2049]^2}{[.208(-.24)^2 - .0208(-.24) + .2049]^2 + \delta_{-.24}}$$

$$\delta_{-.24} = .00174(-.24)^4 - .00864(-.24)^3 + .0163(-.24)^2 - .00073x + .00017$$

$$\left[\begin{array}{l} m(-.24) = \dfrac{2(.2219)^2}{(.2219)^2 + .0014} = 1.94 \\[2ex] f_{1.94,7} = 4.76 \\[2ex] \quad S_\gamma(x) = 1.2(4.76) \qquad \text{solving, } x = -.24 \end{array}\right.$$

$$\left[\begin{array}{l} m(2.57) = \dfrac{2(1.524)^2}{(1.524)^2 + .0353} = \dfrac{2(2.32)}{2.36} = 1.97 \\[2ex] f_{1.97,7} = 4.75 \\[2ex] \quad S_\gamma(x) = 1.2(4.75) = 5.70 \qquad \text{solving, } x = 2.55 \end{array}\right.$$

Then $\qquad \Gamma_S \simeq \overline{\cup} - (-.24,\ 2,56)$.

CHAPTER IX

EVALUATION OF THE TESTS

It is desired to find a test for block differences for an experimental model which has block effects with a fixed portion added to a regression portion.

In looking for a test one turns naturally to a χ^2 test when σ^2 is known or an F-test when σ^2 is estimated. To study the variation of the test statistic with the covariate, the matrix of the quadratic form is expanded in eigenvectors.

$$\frac{1}{\sigma^2} T_\gamma(x) = \frac{1}{\sigma^2} \sum_{i=1}^{r-1} \frac{\{[\hat{\underline{\kappa}} + \hat{\underline{P}}(x)]'\underline{v}_{xi}\}^2}{\gamma_{xi} + \frac{1}{t}} \sim \chi^2_{r-1}$$

It is noted that the numerators of the fractions would add nicely to $[\hat{\underline{\kappa}} + \hat{\underline{P}}(x)]'[\hat{\underline{\kappa}} + \hat{\underline{P}}(x)]$, but with unequal γ_{xi}'s in the denominators of the fractions such an operation is not permissible. Clearly an upper limit on $T_\gamma(x)$ could be found using the minimum γ_{xi} and a lower limit would result using the maximum of the γ_{xi}'s . This assumes, of course, that the γ_{xi}'s are known. The temptation to average the γ_{xi}'s would occur to an applied person who had not been too thoroughly influenced by the mathematical quest for exactness.

The applied mathematician might very well average the γ_{xi}'s to get an approximation of $T_\gamma(x)$. Then if he were asked "Does your approximation of $T_\gamma(x)$ have the same critical point (or function) $\sigma^2 C_{r-1}$?" he might say

"Only if the γ_{xi}'s are exactly alike, otherwise I hedge a little. That is, where the γ_{xi}'s are very different it takes more to surprise me at the same α-level. However, in all honesty, if I have no reason to believe the roots are very different, I just go ahead and use $\sigma^2 C_{r-1}$."

The statistic

$$S_\gamma(x) = [\hat{\underline{\kappa}} + \hat{\underline{P}}(x)]'[\hat{\underline{\kappa}} + \hat{\underline{P}}(x)] \mathrel{\dot{\sim}} \frac{r-1}{m(x)} \chi^2_{m(x)}$$

where

$$m(x) = \frac{\left(\sum \gamma_{xi} + \frac{r-1}{t}\right)^2}{\left(\sum \gamma_{xi} + \frac{r-1}{t}\right)^2 + (r-1) \sum \gamma_{xi}^2 - \left(\sum \gamma_{xi}\right)^2}$$

does the averaging and hedging with the mathematical justification of fitting a multiple of a central χ^2 with a certain number of degrees of freedom. With this justification for what the applied mathematician might do instinctively, the test is found to be insensitive to differences in roots when α is large. When α is small, checks (finding the ψ_0, ψ_1, ψ_2, ψ_3, ψ_4, and hence m(x) values defined in Chapter VI) are tedious, but not nearly so tedious as solving $T_\gamma(x) = \sigma^2 C_{r-1}$ for x .

The expansion of $T_\gamma(x)$ has another result. It shows that the $T_\gamma(x)$ statistic is weighting the various contrasts of the block unevenly as we have more or less information about them. Granting the general wisdom of such a move, if we are equally interested in all of the blocks, $T_\gamma(x)$ is not really telling us what we want to know. That is, $T_\gamma(x)$ gives an exact answer to the approximate problem at hand. $S_\gamma(x)$, on the other hand, gives an approximate answer to the exact problem at hand, which is more in line with the desires of the statistician.

The $T_\gamma(x)$ test is uniformly most powerful against a set of alternatives which may not be of particular interest. Of interest in regard to this point is work by Reisch and Webster [6], Scheffe [8] and Hsu [4]. The redirection of power offered by $S_\gamma(x)$ should more than compensate for the errors incurred by the approximate nature of the test.

The $R_\gamma(x)$ test which is the $S_\gamma(x)$ statistic evaluated against the critical function of $T_\gamma(x)$ is a further approximation. $R_\gamma(x)$ is a very good approximation of $S_\gamma(x)$ for high α-levels and seems to also give good results when α is low. $R_\gamma(x)$ has the further intuitively appealing characteristic of always yielding a set of X values for which the test is significant which is of the interval or interval-complement form.

The test $R_\gamma(x)$ and the resulting set of X values, Γ_R, are simple and practical tools which should find use in industry and in the social sciences. They are well suited to the large α situations in which one is particularly concerned (worried) about not recognizing block differences when they occur. The usefulness of these tools in scientific research would be increased immeasurably with a lower bound for m(x) when the covariate block means are not equal and with some study into the closeness of the approximating distribution.

Although this work has been developed for a two-way classification model with a covariate, it is equally valid when the treatment effects are 0. Then we are talking about comparing simple regressions based on t replications from r different sources and the M matrix is diagonal in that case.

Work is currently being done by the author to extend these results to compare multiple regressions from different sources. In these cases, the S-type statistic addresses itself more directly to the problem at hand than does a T-type statistic, and it will greatly simplify the problem of finding sets of points for which the test is significant. However, there may be a number of other cases in which one would rather not accept the unequally weighted contrasts of the conventional sum of squares for regression test. Armed with some facts about the closeness of the approximations it is possible than one might wage war on many fronts against the wide-spread use of $\hat{\underline{\beta}}'(X'X)\hat{\underline{\beta}}$ to test the effects of $\underline{\beta}$ when $\underline{y} = x\underline{\beta} + \underline{\varepsilon}$.

APPENDIX I

THE MATRIX $(I - \frac{1}{r} J)(I - \frac{1}{r} A'A)^{-1}(I - \frac{1}{r} J)$

The matrix $(I - \frac{1}{r} J)(I - \frac{1}{r} A'A)^{-1}(I - \frac{1}{r} J)$ where $A = [a_{ij}]_{t\times r}$ and $a_{ij} = \frac{x_{ij} - x_{\cdot j}}{(qx)_j}$ where $(qx)_j = \sqrt{\sum_i (x_{ij} - x_{\cdot j})^2}$ occurs in this work as a special case of the matrix $(I - \frac{1}{r} J)M(I - \frac{1}{r} J)$ where $M = R^{-1}(I - \frac{1}{r} A'A)^{-1}R^{-1}$ and R is a diagonal matrix with j^{th} element $(qx)_j$. If the covariate values are controllable, the elements of R can be expected to be made very close to the same values. In Cox's example [1] the covariate values in each block are the digits 1, 2, 3, $\cdots$, t so that R would be the constant $\sqrt{\frac{t(t+1)(t-1)}{12}}$ times an identity matrix.

The R matrix is clearly nonsingular so the nonsingularity of M^{-1} depends on the matrix $(I - \frac{1}{r} A'A)$. Singularity of this matrix could result only from A'A being of rank one. This follows from the positive semi-definite nature of A'A and that its trace equals r .

Let $\frac{1}{r} A'A = \sum_i^r \lambda_i \underline{v}_i \underline{v}_i'$, then $(I - \frac{1}{r} A'A) = \sum_{i=1}^r (1 - \lambda_i)\underline{v}_i\underline{v}_i'$, where $\sum_i \lambda_i = 1$ and $0 \leq \lambda_i \leq 1$ for all i. Hence the only way to produce a 0-root of $(I - \frac{1}{r} A'A)$ is for one $_i$ to be equal to 1 and the others to be 0. In that case $\frac{1}{r} A'A = 1\underline{v}_i\underline{v}_i'$ where $v_{ik}^2 = 1/r$ for all k. That is, $\left|\sum_i^t a_{ij}a_{i'j}\right| = 1$ for all i and i'. In other words, each $\underline{a}_j$ vector, with elements a_{ij}, is identical to all other $\underline{a}_{j'}$ vectors, or to their negative value. Symbolically, $A = (\underline{a}, -\underline{a}, -\underline{a}, \underline{a}, \cdots, \underline{a})$. If this were the case, one linear function of β_j's and ψ would be hopelessly confounded with one linear function of treatments, thus accounting for the singularity.

If the covariate is uncontrollable, such an occurrence is most improbable, and if the covariate is controllable the situation will be avoided.

One desirable case of singularity should be noted here. If $A'A = J$ then ψ will be confounded with treatments leaving the β_j's estimable. Such mathematical neatness is only possible if the covariate is controllable, and only desirable if there is no interest in examining treatments or mean regression.

Under the above very rigid restrictions the A matrix could be written as $[\mathrm{diag}(a_i)]J_{t\times r}$ or $A = (\underline{a}, \underline{a}, \underline{a}, \underline{a}, \cdots, \underline{a})$ where the elements of the vector $\underline{a}$ are a_i. Then equation (3) of Chapter II becomes

$$\omega \begin{bmatrix} \sum_i a_i y_{i1} & & \sum_i a_i \sum_j y_{ij} \\ \sum_i a_i y_{i2} & & \sum_i a_i \sum_j y_{ij} \\ \vdots & -\frac{1}{\gamma} & \vdots \\ \sum_i a_i y_{ir} & & \sum_i a_i \sum_j y_{\cdot j} \end{bmatrix} = \omega^2 (I - \frac{1}{r} J)(\hat{\beta} + \hat{\psi}\underline{1})$$

where

$$R = \omega I$$

and

$$\frac{1}{\omega}[\underline{a}' \otimes (I - \frac{1}{r}J)]\underline{y} = \hat{\underline{\beta}}$$

variance

$$\hat{\underline{\kappa}} + z\hat{\underline{\beta}} = \frac{\sigma^2}{t}(I - \frac{1}{r}J) + \frac{\sigma^2}{\omega^2}(I - \frac{1}{r}J)$$

Hence,

$$\hat{\underline{\kappa}} + z\hat{\underline{\beta}} \sim N[\underline{\kappa} + z\underline{\beta}, \sigma^2(\frac{1}{t} + \frac{1}{\omega^2})(I - \frac{1}{r}J)]$$

when $A'A = J$ and $R = \omega I$.

Assuming $(I - \frac{1}{r}A'A)$ to be nonsingular by design or happenstance, one need be wary only of near singularity when the primary eigenvector of the $A'A$ matrix, that is, the vector corresponding to the largest eigenroot is not proportional to $\underline{1}$ or nearly so. To clarify this point consider first a desirable $A'A$ matrix of the following form.

$$\frac{1}{r}\begin{bmatrix} 1 & \rho & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho & \rho \\ \rho & \rho & 1 & \rho & \rho \\ \rho & \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & \rho & 1 \end{bmatrix}$$

Then

$$\frac{1}{r}A'A = \frac{1}{r}(1-\rho)I + \frac{1}{r}\rho J = \frac{1+(r-1)\rho}{r}[\frac{1}{r}\underline{1}\,\underline{1}'] + \sum_{i=1}^{r-1}\frac{r-(1-\rho)}{r}\underline{v}_i\underline{v}_i'$$

where

$$\underline{v}_i'\underline{1} = 0 \quad \text{for all } i$$

and

$$\underline{v}_i'\underline{v}_{i'} = 1 \quad \text{if } i = i'$$

$$\underline{v}_i'\underline{v}_{i'} = 0 \quad \text{if } i \neq i'$$

Then

$$I - \frac{1}{r} A'A = \frac{(r-1)(1-\rho)}{r} [\frac{1}{r} \underline{1}\,\underline{1}'] + \sum_{i=1}^{r-1} \frac{r-(1-\rho)}{r} \underline{v}_i \underline{v}_i'$$

and

$$(I - \frac{1}{r} A'A)^{-1} = \frac{r}{(r-1)(1-\rho)} [\frac{1}{r} \underline{1}\,\underline{1}'] + \sum_{i=1}^{r-1} \frac{r}{r-(1-\rho)} \underline{v}_i \underline{v}_i'$$

The full matrix $(I - \frac{1}{r} J)(I - \frac{1}{r} A'A)^{-1}(I - \frac{1}{r} J)$ then becomes

$$\sum_{i=1}^{r-1} \frac{r}{r-(1-\rho)} \underline{v}_i \underline{v}_i' \quad .$$

The matrix has r-1 identical eigenroots of magnitude $\frac{r}{r-(1-\rho)}$. These roots are between 1 and $\frac{r-1}{r-2}$ resulting in a trace between r-1 and $\frac{(r-1)^2}{r-2}$.

That is, even a ρ value very near 1 does not destroy the balance of the covariance matrix although it makes the $I - \frac{1}{r} A'A$ matrix very nearly singular.

However, this property is lost if the A'A matrix is of a form such as

$$\frac{1}{r} \begin{bmatrix} 1 & -\rho & -\rho & -\rho \\ -\rho & 1 & \rho & \rho \\ -\rho & \rho & 1 & \rho \\ -\rho & \rho & \rho & 1 \end{bmatrix}$$

Then

$$\frac{1}{r} A'A = \frac{1}{r} (1-\rho) I + \frac{1}{r} \rho \left[\begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \end{pmatrix} (-1 \; 1 \; 1 \; 1) \right]$$

And

$$(I - \frac{1}{r} A'A)^{-1} = \frac{r}{(r-1)(1-\rho)} \left[\frac{1}{r} \begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \end{pmatrix} (-1\ 1\ 1\ 1) \right] + \sum_{i=1}^{r-1} \frac{r}{r-(1-\rho)} \underset{-i}{v} \underset{-i}{v}'$$

$$= \frac{r}{r-(1-\rho)} I + \frac{r^2 \rho}{(r-1)(1-\rho)[r-(1-\rho)]} \left[\frac{1}{r} \begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \end{pmatrix} (-1\ 1\ 1\ 1) \right]$$

The $(I - \frac{1}{r} J)(I - \frac{1}{r} A'A)^{-1}(I - \frac{1}{r} J)$ matrix becomes

$$\frac{r}{r-(1-\rho)} (I - \frac{1}{r} J) - \frac{r^2}{\left(\frac{r}{\rho} - \frac{1}{\rho}\right)\left(\frac{1}{\rho} - 1\right)\left[\frac{r}{\rho} - \left(\frac{1}{\rho} - 1\right)\right]} \left[\frac{1}{r}(I - \frac{1}{r} J) \begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \end{pmatrix} (-1\ 1\ 1\ 1) (I - \frac{1}{r} J) \right]$$

The roots of this matrix are not balanced and the trace of this matrix is

$$\frac{(r-1)r}{r-(1-\rho)} + \frac{r^2 - (r-2)^2}{\left(\frac{r}{\rho} - \frac{1}{\rho}\right)\left(\frac{1}{\rho} - 1\right)\left[\frac{r}{\rho} - \left(\frac{1}{\rho} - 1\right)\right]} .$$

In this form, it is apparent that the trace of the covariance matrix can be made as large as anyone chooses by making ρ very close to 1 . If ρ is very close to 1, the primary eigenvector of the A'A matrix would be $\frac{1}{\sqrt{r}}$ (-1 1 1 1) .

These two illustrations of near singularity are presented to aid the reader in understanding the types of A'A matrices which will yield unbalanced eigenroots and large trace values for the covariance matrices of block regression effects. It is these conditions which will cause the tests for block effects to give poor results. By way of summary, these conditions are near singularity, caused by a nearly degenerate A'A matrix, unless the primary vector of the A'A matrix is proportional to a vector of +1's .

Furthermore, the r-1 eigenroots of the matrix $(I - \frac{1}{r} J)(I - \frac{1}{r} A'A)^{-1} \times (I - \frac{1}{r} J)$ are all greater than 1 . To prove this point let

$$Q = (I - \frac{1}{r} A'A)^{-1} = \sum_{i=1}^{r} \frac{1}{1-\lambda_1} \underline{v}_i \underline{v}_i' \quad ,$$

where $0 < \lambda_i < 1$ and hence $0 < 1-\lambda_i < 1$. That is, Q has r roots greater than 1 . Then

$$I = \sum_i \underline{v}_i \underline{v}_i' \quad ,$$

and

$$Q-I = \sum_i \left[\frac{1}{1-\lambda_i} - 1\right] \underline{v}_i \underline{v}_i' \quad .$$

Note that

$$\left[\frac{1}{1-\lambda_i} - 1\right] > 0 \quad \text{for all } i \ .$$

Hence, Q-I is a positive definite matrix. It then follows that $(I - \frac{1}{r} J)(Q-I)(I - \frac{1}{r} J)$ is a positive semi-definite matrix with r-1 positive roots and one zero root corresponding to the eigenvector $\frac{1}{\sqrt{r}} \underline{1}$. Therefore,

$$(I - \frac{1}{r} J)(Q-I)(I - \frac{1}{r} J) = 0 \frac{1}{r} \underline{1}\, \underline{1}' + \sum_{i=1}^{r-1} \eta_i \underline{\mu}_i \underline{\mu}_i'$$

$$\eta_i > 0$$

$$(I - \frac{1}{r} J)(Q)(I - \frac{1}{r} J) - (I - \frac{1}{r} J) = 0 \frac{1}{r} \underline{1}\, \underline{1}' + \sum_{i=1}^{r-1} \eta_i \underline{\mu}_i \underline{\mu}_i'$$

$$(I - \frac{1}{r} J)(Q)(I - \frac{1}{r} J) + \frac{1}{r} J = \frac{1}{r} \underline{1}\ \underline{1}' + \sum_{i=1}^{r-1} (1 + \eta_i)\underline{\mu}_i\underline{\mu}_i'$$

and

$$(I - \frac{1}{r} J)Q(I - \frac{1}{r} J) = \sum_{i=1}^{r-1} (1 + \eta_i)\underline{\mu}_i\underline{\mu}_i' \quad \text{and} \quad 1 + \eta_i > 1 \quad .$$

Hence, $(I - \frac{1}{r} J)(I - \frac{1}{r} A'A)^{-1}(I - \frac{1}{r} J)$ has r-1 non-zero roots all greater than 1 . And if $R = \omega I$, then $(I - \frac{1}{r} J)M(I - \frac{1}{r} J)$ has r-1 roots all greater than $\frac{1}{\omega^2}$.

With this lower bound on the roots and the upper bound furnished by the sum of the roots being equal to the trace of the matrix, we can often be assured of a high degree of balance among the roots without actually finding them.

APPENDIX II

CHARTS AND GRAPHS

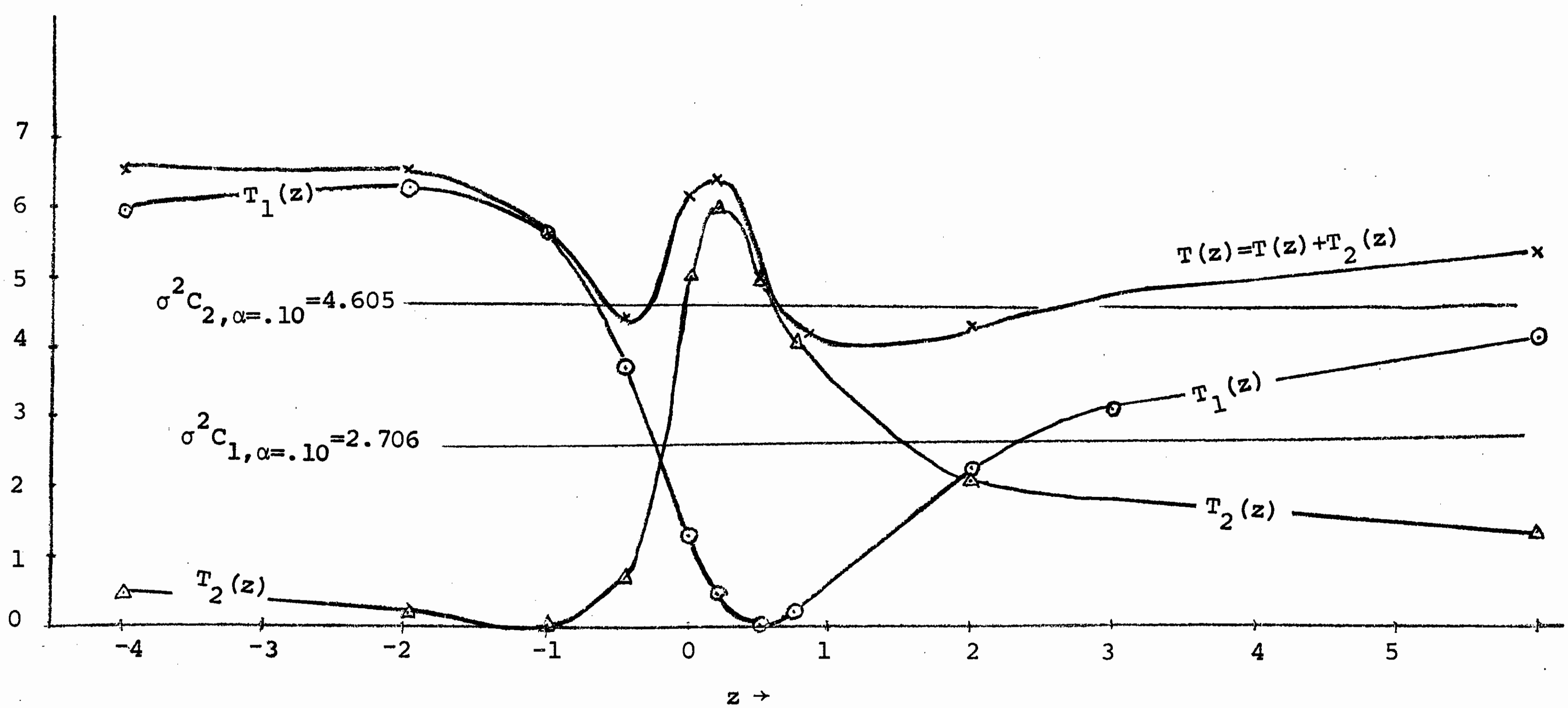

7
6
5
4
3
2
1
0
-4
-3
-2
-1
0
1
2
3
4
5
z →
T(z)=T(z)+T2(z)
T1(z)
T2(z)
σ2C2,α=.10=4.605
σ2C1,α=.10=2.706

FIGURE I.

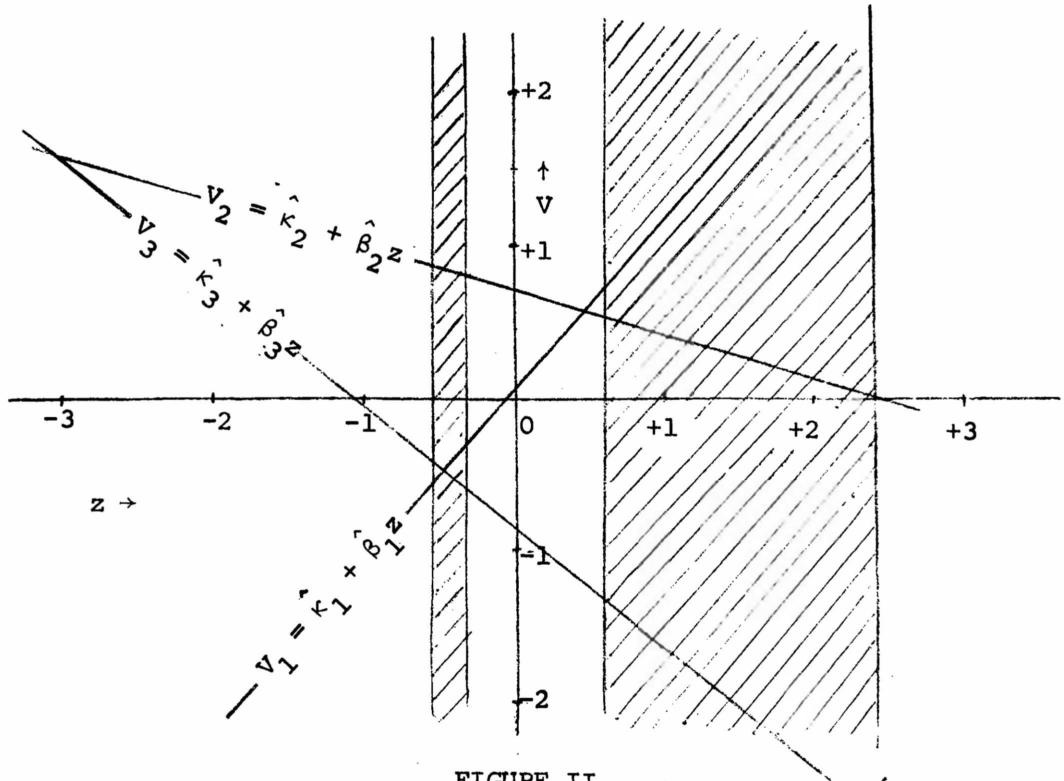


FIGURE II.

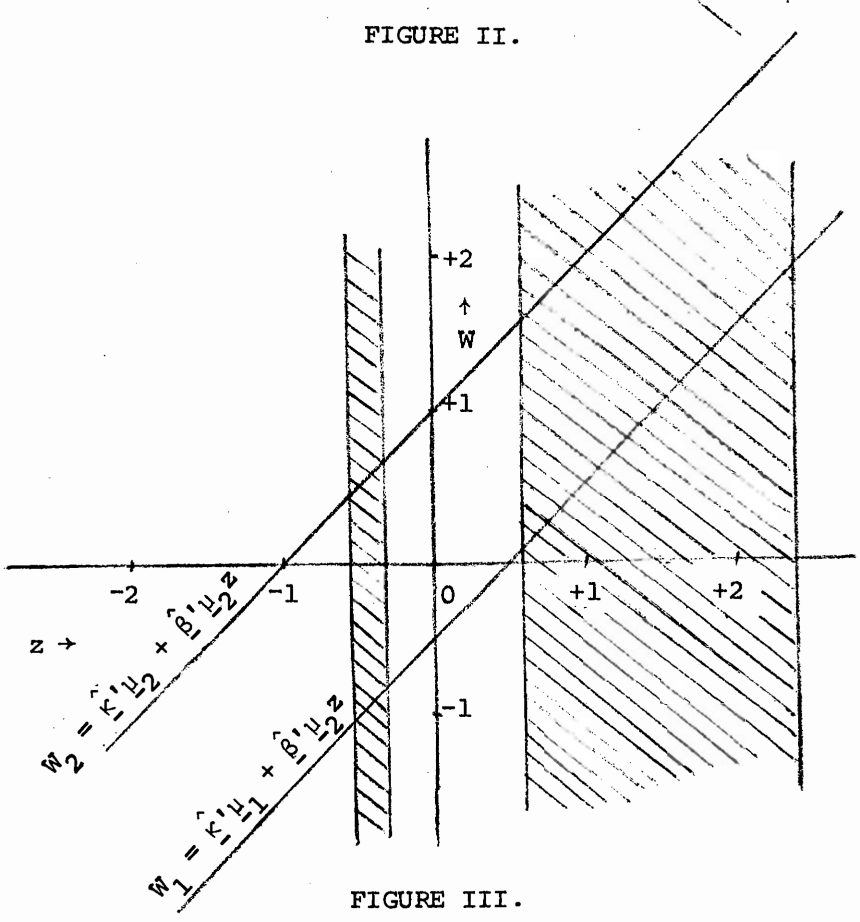
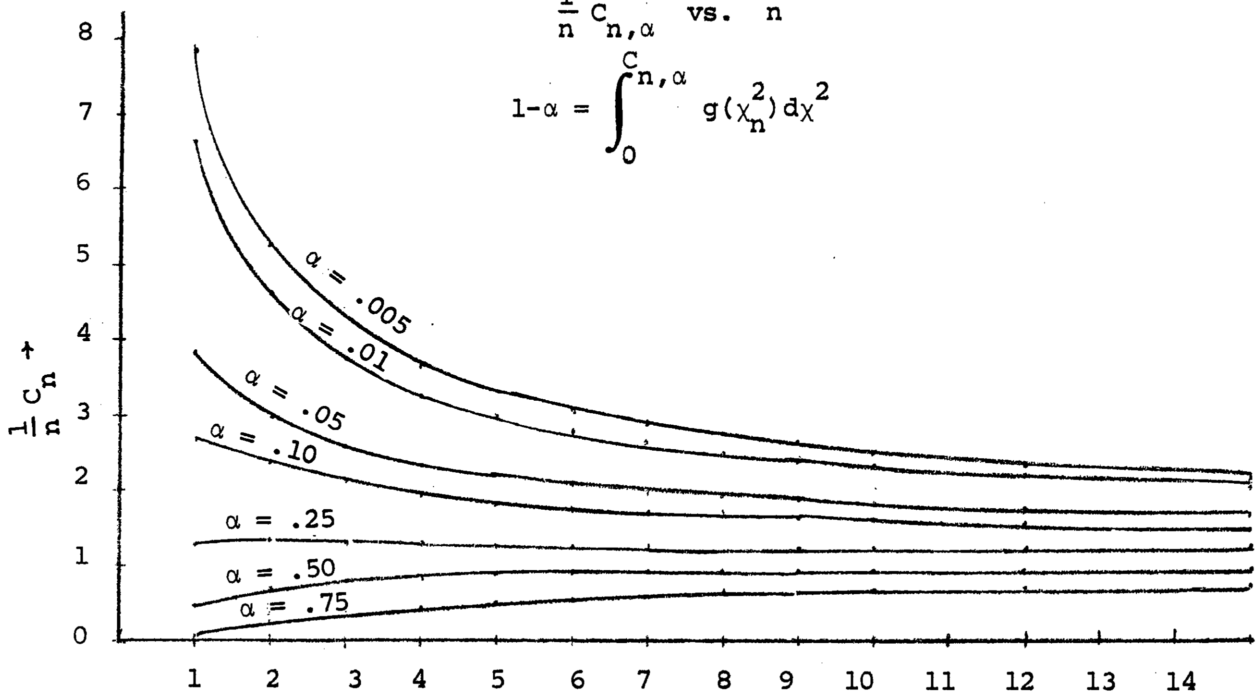


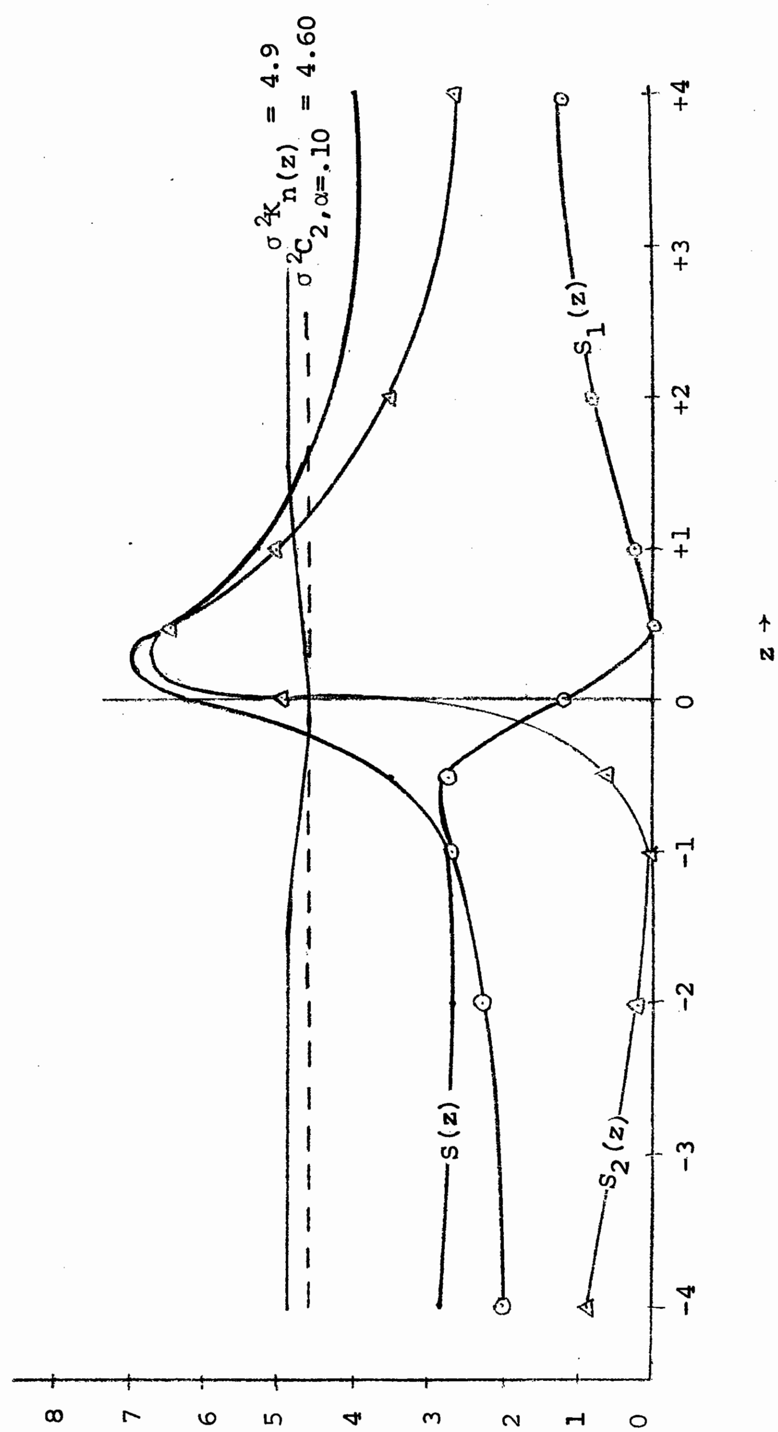
FIGURE III.

FIGURE IV

$$\frac{1}{n} C_{n,\alpha} \quad \text{vs.} \quad n$$

$$1-\alpha = \int_0^{C_{n,\alpha}} g(\chi_n^2)\, d\chi^2$$





$\sigma^2 K_n(z) = 4.9$
$\sigma^2 C_{2,\alpha=.10} = 4.60$
$S_1(z)$
$S(z)$
$S_2(z)$
z →
-4
-3
-2
-1
0
+1
+2
+3
+4
0
1
2
3
4
5
6
7
8

FIGURE V

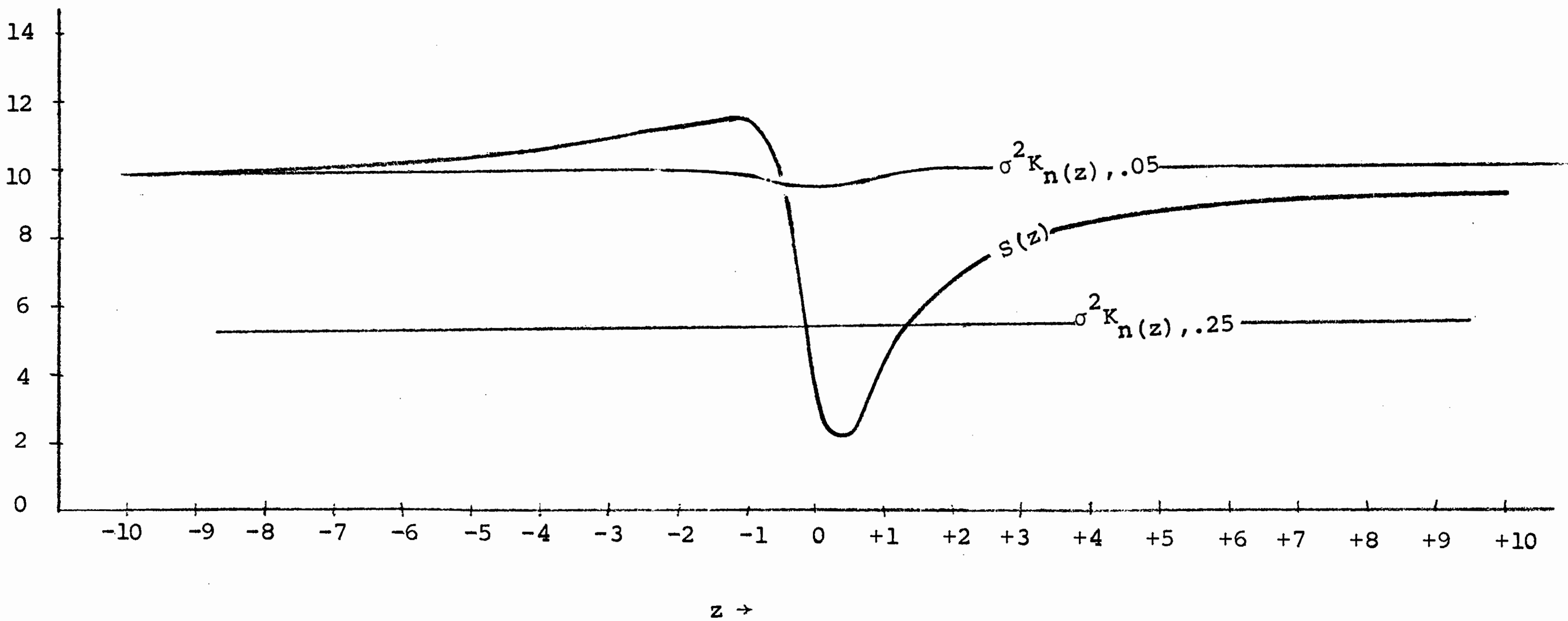


FIGURE VI.

Example 2.

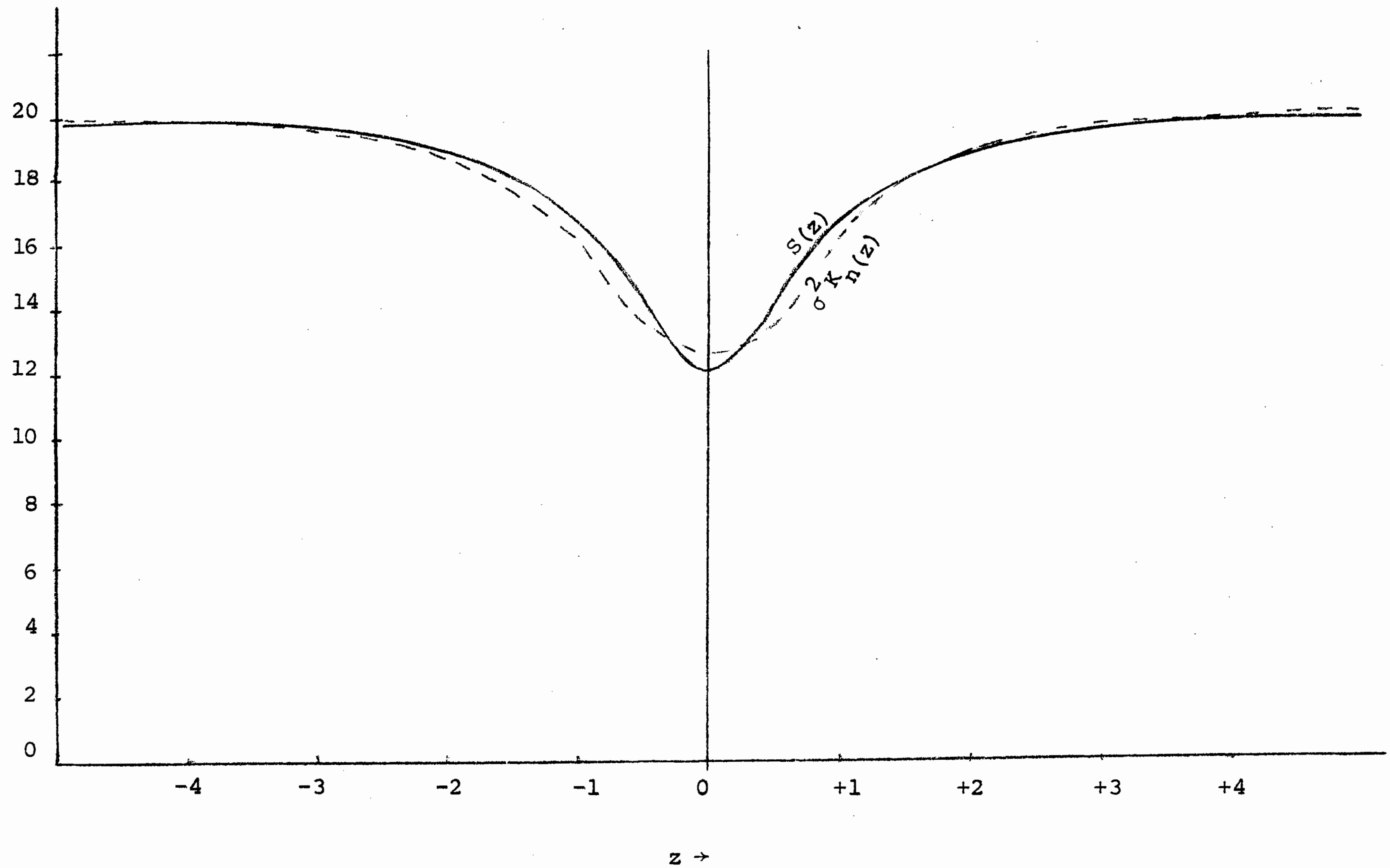


FIGURE VII.

Example 3.

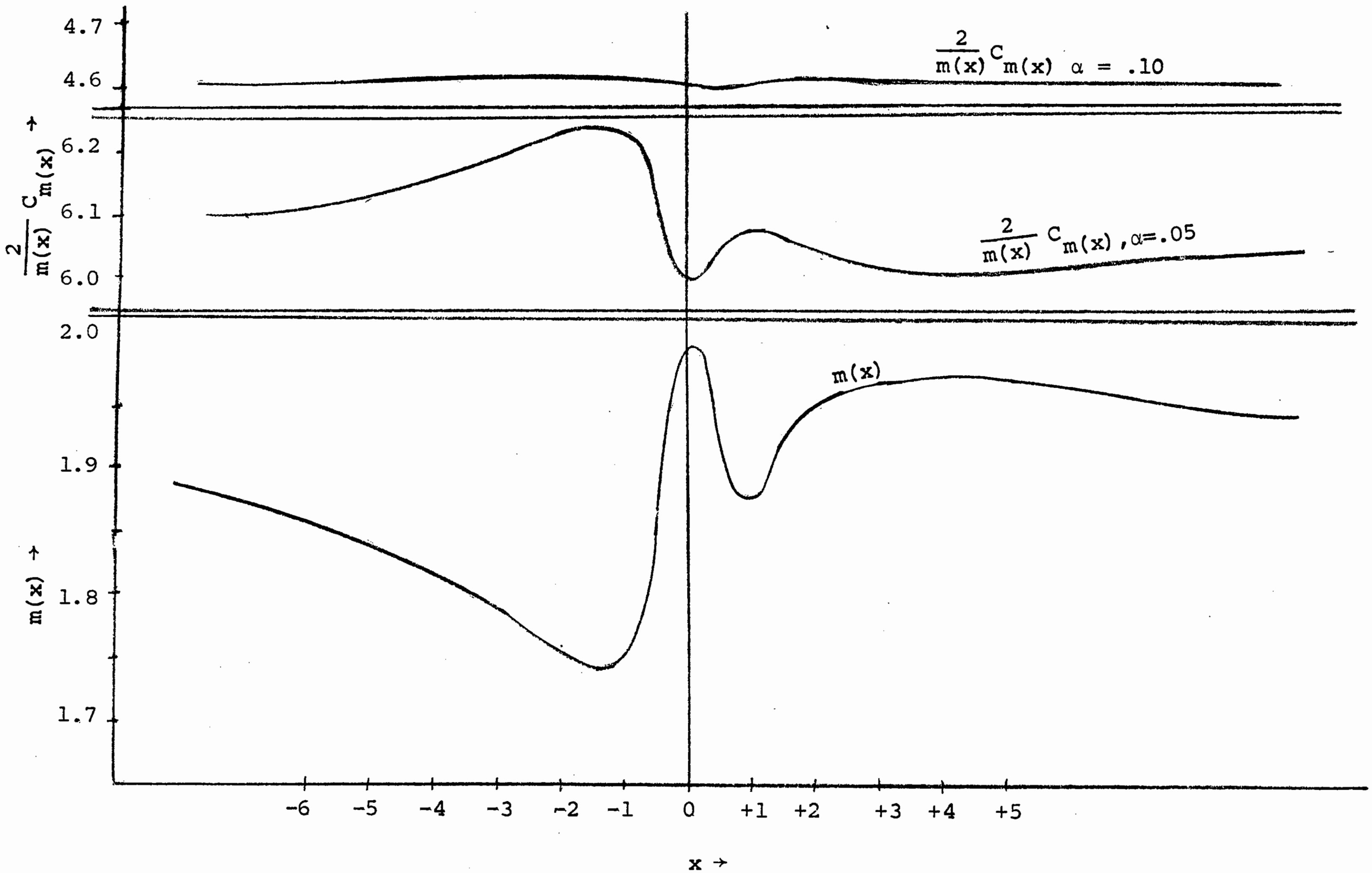

4.7
4.6
6.2
6.1
6.0
2.0
1.9
1.8
1.7
$\frac{2}{m(x)} C_{m(x)}$ →
m(x) →
$\frac{2}{m(x)} C_{m(x)}$ α = .10
$\frac{2}{m(x)} C_{m(x)}$, α=.05
m(x)
-6 -5 -4 -3 -2 -1 0 +1 +2 +3 +4 +5
x →

FIGURE VIII.

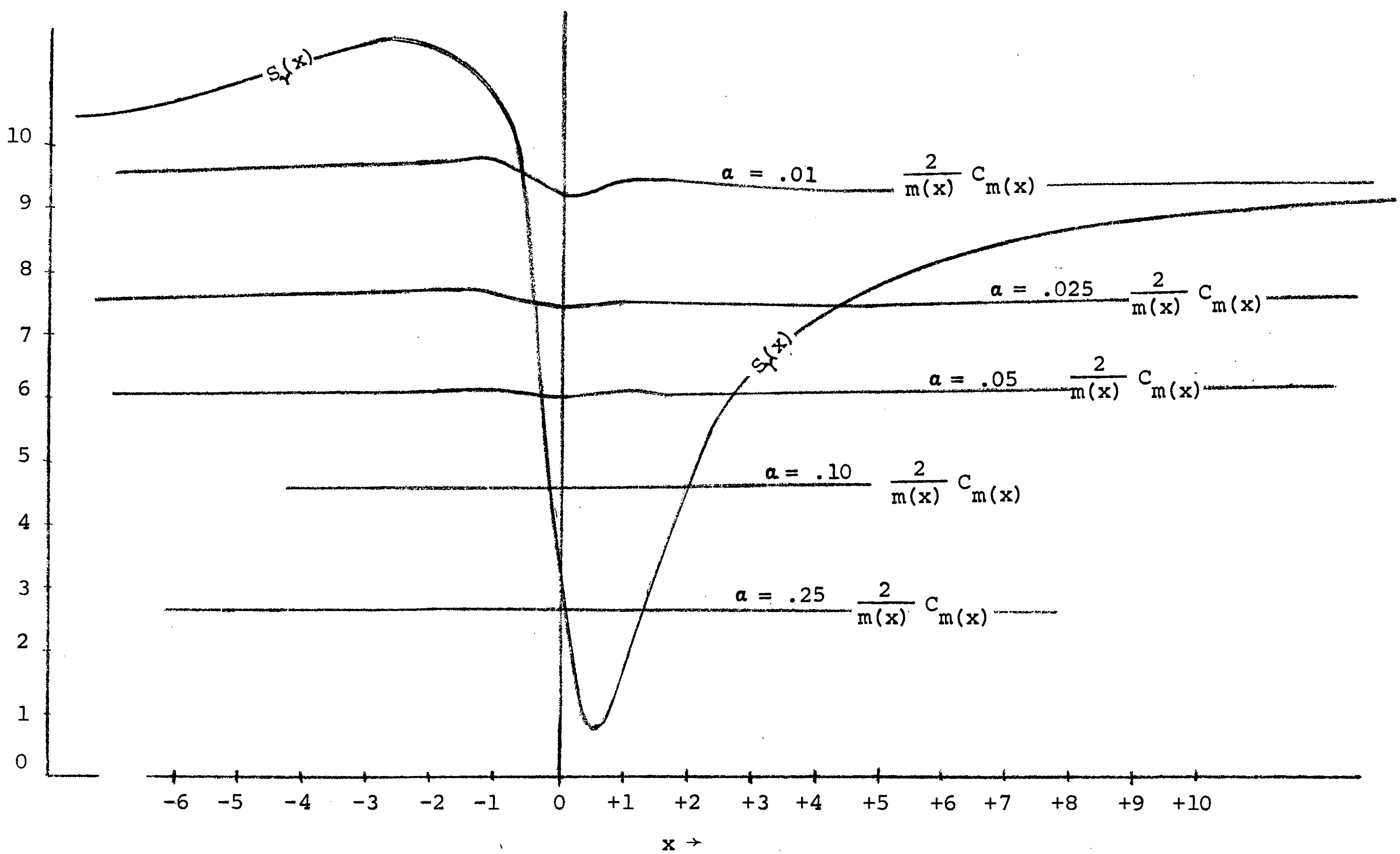


FIGURE IX.

Example 4.

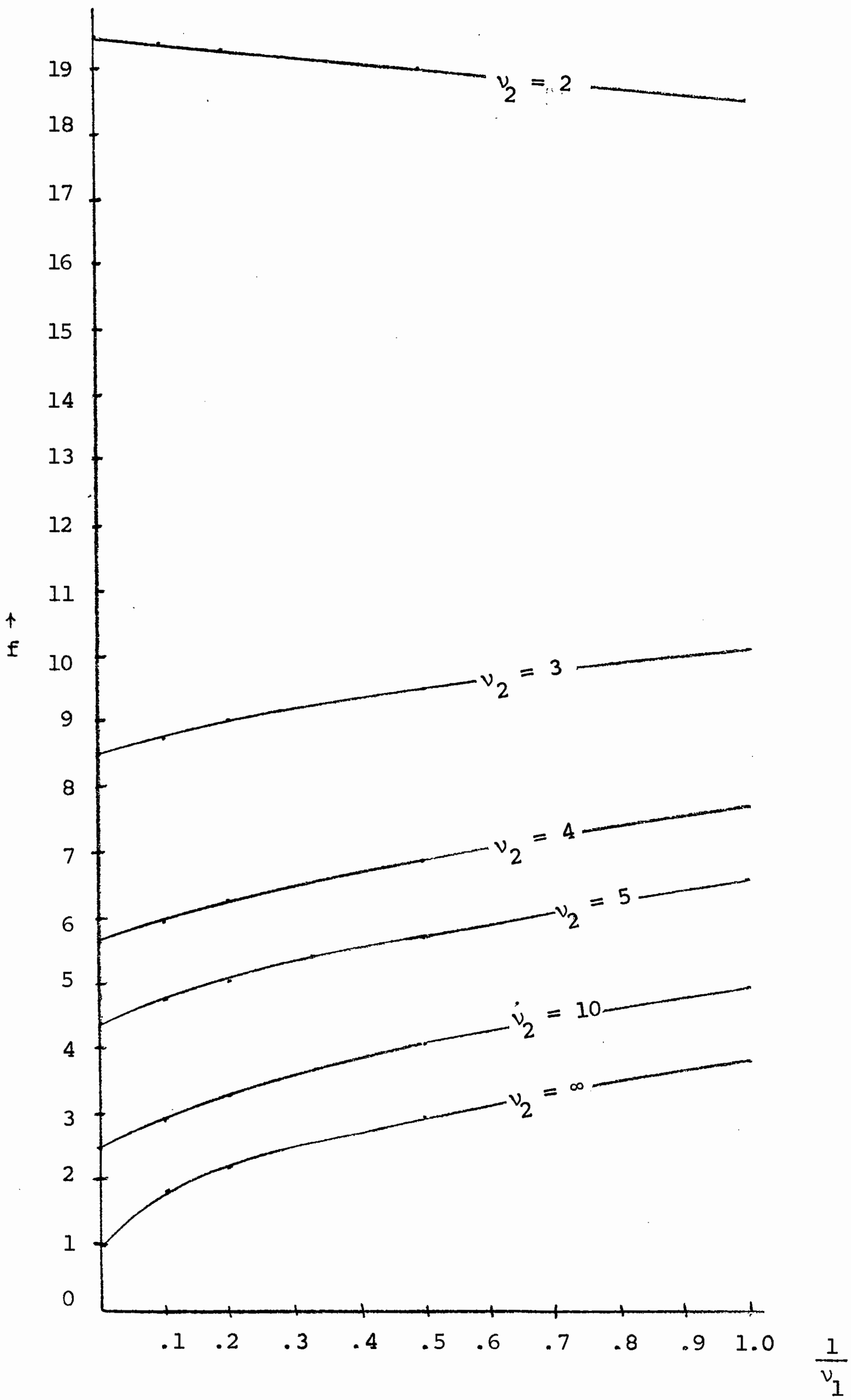


FIGURE X

$\alpha = .05$

f vs. $\frac{1}{\nu_1}$

$$\int_f^\infty F_{\nu_1, \nu_2} = .05$$

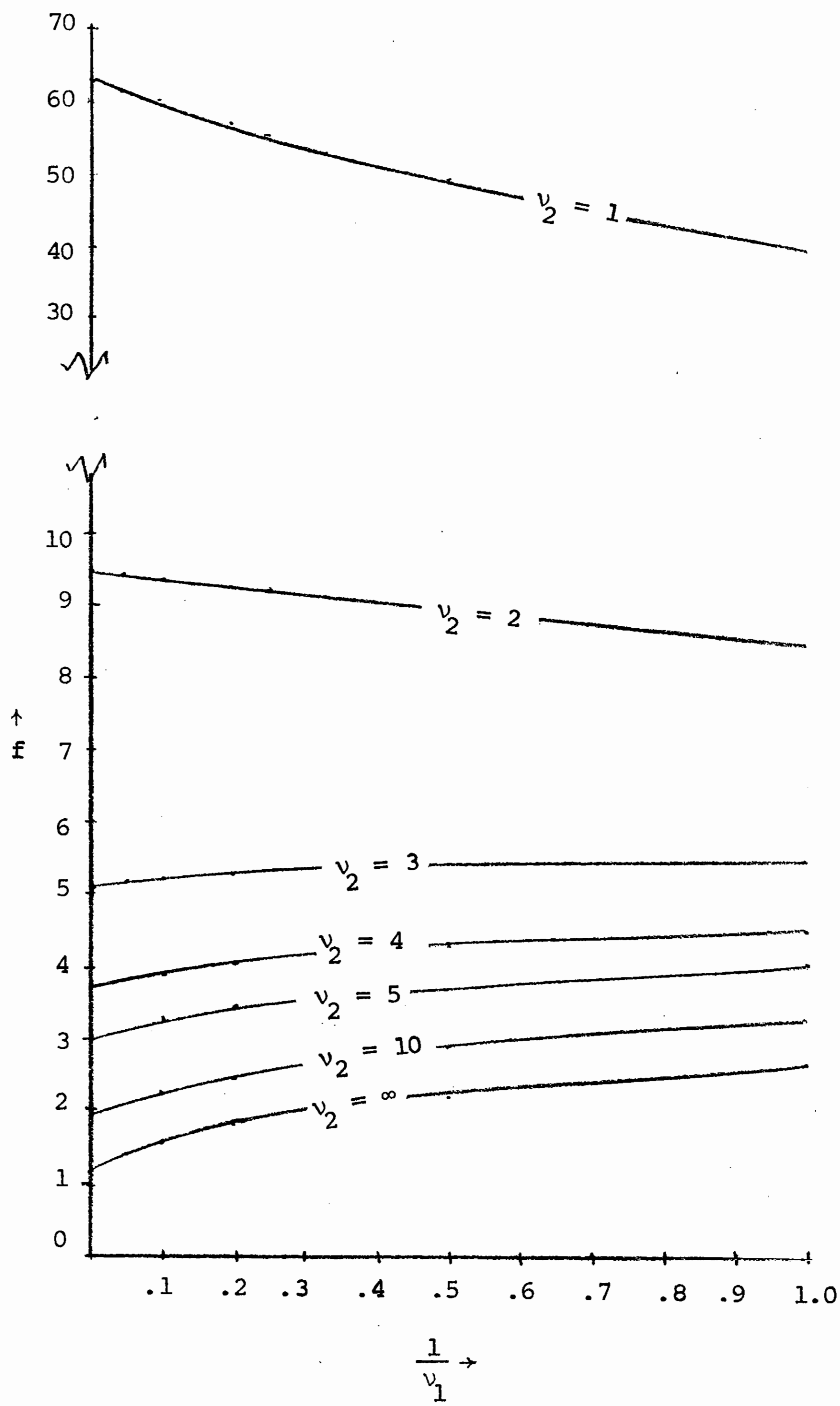


FIGURE XI

$\alpha = .10$

f vs. $\frac{1}{\nu_1}$

$$\int_f^\infty F_{\nu_1,\nu_2} = .10$$

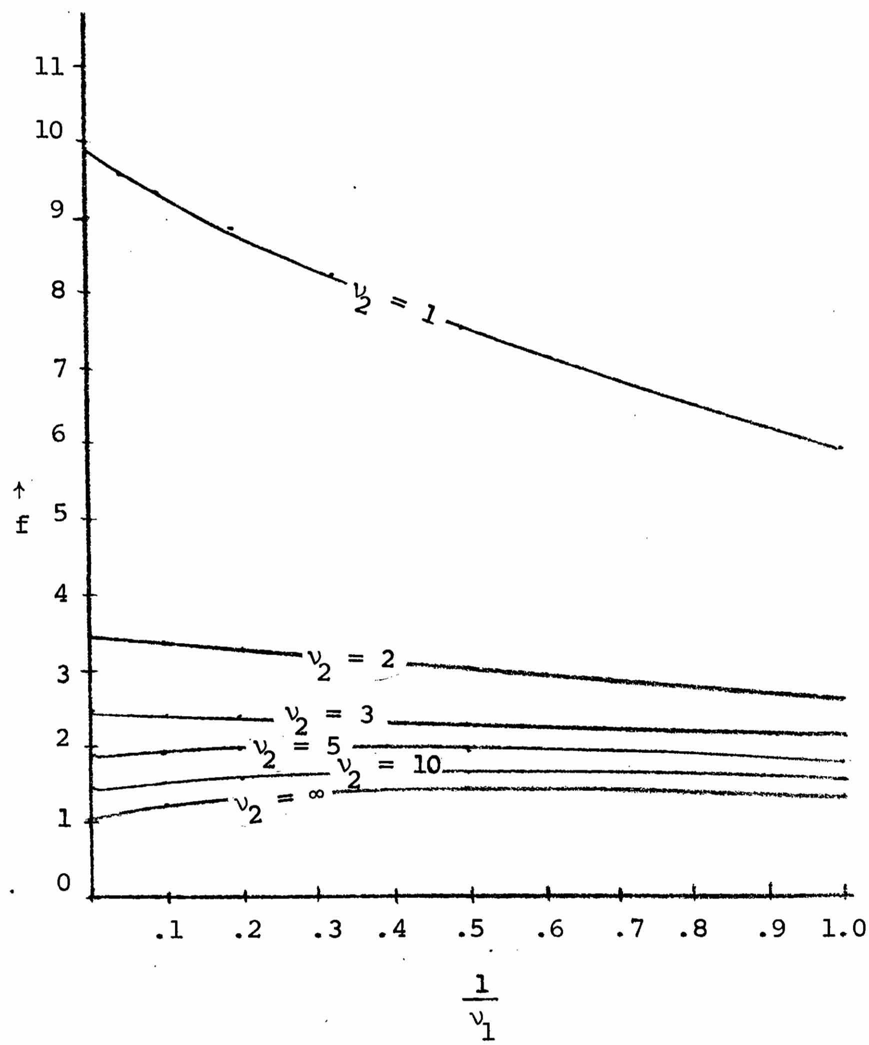


FIGURE XII

$\alpha = .25$

f vs. $\frac{1}{\nu_1}$

$$\int_f^\infty F_{\nu_1, \nu_2} = .25$$

LIST OF REFERENCES

[1] Cox, C. P. (1958). "The analysis of latin square designs with individual curvatures in one direction," J. R. Stat. Soc. B., 20, 193.

[2] ________. (1958). "Experiments with two treatments per experimental unit in the presence of an individual covariate," Biometrics, 14,

[3] Fleiss, J. L. (1971). "Testing for equal slopes in a randomized block design with covariances," Biometrics, 27, 225-228.

[4] Hsu, P. L. (1941). "Analysis of variance from the power function standpoint," Biometrika, 32, 62-69.

[5] Johnson, P. O. and Neyman, J. "Test of certain linear hypotheses and their application to some educational problems," Statistical Research Memoirs, (University of London) 1, 57-93.

[6] Reisch, J. S. and Webster, J. T. (1969). "The power of a test in covariance analysis," Biometrics, 25, 701-714.

[7] Robson, D. and Atkinson, G. (1960). "Individual degrees of freedom for testing homogeneity of regression coefficients in a one-way analysis of covariance," Biometrics, 16, 593.

[8] Scheffe, H. (1959). Analysis of Variance. New York: John Wiley and Sons, 46-48, 130-134.

[9] Tukey, J. W. (1949). "One degree of freedom for non-additivity," Biometrics, 5, 232-242.

[10] Zelen, Marvin (1957). "Incomplete block designs," Biometrics, 13, 320.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) SOUTHERN METHODIST UNIVERSITY	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP UNCLASSIFIED

3. REPORT TITLE

When the coefficient of a covariate changes from block to block.

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report

5. AUTHOR(S) (First name, middle initial, last name)

Elinor S. Pape

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
November 18, 1971	85	10

8a. CONTRACT OR GRANT NO. N00014-68-A-0515 b. PROJECT NO. 042-260 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 108 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office of Naval Research

13. ABSTRACT

When in a two-way classification model the block effects take the form $\nu_j + \phi_j x$ where ν_j and ϕ_j are unknown constants, $j = 1, 2, \ldots, r$ where r is the number of different blocks, and X is a known covariate (measured without error), a test for block differences will be a function of X. Clearly, if X = 0 a test for block differences would test differences in the ν_j. And if $|X|$ is infinitely large, block differences would be caused only by differences in the ϕ_j.

The model in question might be written as $y_{ij} = \mu + \theta_i + \nu_j + \psi x_{ij} + \phi_j x_{ij} + \varepsilon_{ij}$. The μ, θ_i's, ν_j's, ψ, and ϕ_j's ($j = 1, 2, \ldots, r$; $i = 1, 2, \ldots, t$) are unknown constants and constraints are placed on the θ_i's, ν_j's, and ϕ_j's. x_{ij} is the value of the covariate X corresponding to the ij^{th} outcome, y_{ij}. x_{ij} is assumed to be measured without error. The ε_{ij}'s are realizations of random variables assumed to be distributed normally and independently with mean 0 and variance σ^2.

Utilizing this normality assumption and standard least squares techniques, a statistic can be constructed to test block effects which will have an exact χ^2 (or F when σ^2 is unknown) distribution. This statistic is a function of X and is termed $T_\gamma(x)$. $\frac{1}{\sigma^2} T_\gamma(x) \sim \chi^2_{r-1}$.

Continued on additional sheet.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) *Continuation Sheet SOUTHERN METHODIST UNIVERSITY	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP UNCLASSIFIED

3. REPORT TITLE

When the coefficient of a covariate changes from block to block.

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report

5. AUTHOR(S) (First name, middle initial, last name)

Elinor S. Pape

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
November 18, 1971	85	10

8a. CONTRACT OR GRANT NO. N00014-68-A-0515	9a. ORIGINATOR'S REPORT NUMBER(S) 108
b. PROJECT NO. 042-260	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

* see previous sheet

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	OFFICE OF NAVAL RESEARCH

13. ABSTRACT

(Continuation from previ us sheet.)

Efforts to solve $T_\gamma(x) = C$ for X, i.e., efforts to find the X values for which the hypothesis would not be rejected at any particular level of test significance, reveal that T (x) places different weights on different contrasts of the blocks.

In an effort to weight all contrasts of the blocks equally, a second statistic, S (x), is developed which has the same mean and variance as a multiple of a χ^2 statistic with a certain number, m(x), degrees of freedom. That is

$$\frac{1}{\sigma^2} S_\gamma(x) \sim \frac{r-1}{m(x)} \chi^2_{m(x)} .$$

The statistic $S_\gamma(x)$ appears to have higher power against general alternatives of interest than does $T_\gamma(x)$. The equation $S_\gamma(x) = C$ is more easily solved for the values of X such that $S_\gamma(x) = c$, always yielding a set of x's which is a single interval or the complement of a single interval on the real line. The fact that $\frac{r-1}{m(x)} \chi^2_{m(x)}$ is a function of X and not a constant causes little practical difficulty since $\frac{r-1}{m(x)} \chi^2_{m(x)}$ is a close approximation of χ^2_{r-1} in many cases.

Hence, S (x) is recommended as a test statistic for determining whether or not blocks are significantly different at any particular value of the covariate X.

DD FORM 1473 1 NOV 65 (PAGE 1)

UNCLASSIFIED