A RAND NOTE

USING THE PISE CRITERION TO MEASURE THE EFFECTS OF IMBALANCE IN THE ANALYSIS OF COVARIANCE

S. James Press

February 1983

N-1890-MRAL

Prepared for

The Office of the Assistant Secretary of Defense/Manpower, Reserve Affairs and Logistics

35th Year

Rand
SANTA MONICA, CA. 90406

The research described in this report was sponsored by the Office of the Assistant Secretary of Defense/Manpower, Reserve Affairs and Logistics under Contract No. MDA903-80-C-0652.

Published by The Rand Corporation

A RAND NOTE

USING THE PISE CRITERION TO MEASURE THE EFFECTS OF IMBALANCE IN THE ANALYSIS OF COVARIANCE

S. James Press

February 1983

N-1890-MRAL

Prepared for

The Office of the Assistant Secretary of Defense/Manpower, Reserve Affairs and Logistics

35th Year

Rand
SANTA MONICA, CA. 90406

PREFACE

This research describes a new statistical technique developed at Rand to assist in designing an efficient experiment associated with recruiting for the All Volunteer Force. The experiment is intended to measure the effects of various cash bonus option incentives as inducements to "high-quality" young men to join the Army in hard-to-fill occupational specialties. This Note fully describes the research related to this new experimental design technique, and can therefore stand alone. Later it may be merged with one or more other Rand Notes into an integrated report on methodology associated with the Rand project on the Enlistment Bonus Test.

SUMMARY

This Note describes a new statistical technique for comparing unbalanced experimental designs which will be modeled by the univariate analysis of covariance. We propose minimizing a design criterion variable called PISE (percent inflation of the standard error of a contrast).

The research was motivated by the need to design an experiment to measure the effectiveness of a potential new Army recruitment policy. The policy would provide greater management flexibility in paying cash bonuses to eligible "high-quality" young men who agree to enlist in the U.S. Army.

We provide results for both the standard Gauss-Markov model (constant error variance) and the model with heteroscedasticity. We also discuss the problem of attributing the increased variance caused by imbalance in a design to particular covariates.

When implemented, the proposed PISE criterion will generate a design which has greater sensitivity to treatment effect differences.

ACKNOWLEDGMENTS

The author benefited from conversations about the experiment and the research with Rand colleagues Robert M. Bell, Richard L. Fernandez, Jeffrey Garfinkle, Gus W. Haggstrom, and J. Michael Polich, and with University of California, Riverside, colleague and Professor Subir Ghosh. He is also indebted to Robert M. Bell for his excellent suggestions for improving the presentation.

CONTENTS

I. INTRODUCTION

This research was motivated by the need to minimize the cost and to maximize the sensitivity of an experiment Rand is designing to test a new U.S. Army military recruitment policy. Since 1973 the Army has been an all volunteer force, and the Army has therefore been faced with the problem of how to induce eligible people to volunteer for service. At the same time, it has been trying to ensure that an appropriate fraction of those who volunteer will be of "high quality," will be agreeable to being assigned to the specialty areas of interest to the Army at the current time, and will be agreeable to serving for a reasonable length of time. To achieve these goals, the Army has been studying various incentive systems for recruiting, and has been testing these systems on a small scale before implementing them all over the country.

The incentive plan which stimulated the research reported on here is called the bonus test. An Army recruit is paid a given dollar bonus if he agrees to enlist for a designated number of years (we use the term "he" since women have been excluded from the test because of their small numbers). In the experiment the continental United States is subdivided into three (unequal) parts. One part, the "control cell," will continue to employ the current recruiting incentive system. The other two parts of the country, called "test cells," will each employ a different type of bonus system. After letting the experiment run for a year or more, we will compare the numbers of recruits in the test cells with the number in the control cell to try to determine the differential effects of the different bonus plans.

We will of course have to account, and adjust, for different economic and social conditions in the different cells. Differences across the cells are compared statistically with "the analysis of variance" model. To "correct" the results of such a model for differential effects across the cells that are not bonus-induced (treatment) differences involves a slight modification of the model. The modified model is called "analysis of covariance" (ANOCOVA). The correcting variables are called covariates. Ideally, we would understand enough about the underlying (recruitment) process to select covariates that will assign most of the variability in the dependent variable to specific causes, so that the residuals in the model are only randomly associated with the treatments (see Rubin, 1974). Usually, we do not have complete understanding, and we try to compensate for the resulting model misspecification error through a pre-experiment scheme called "balancing," discussed below.

In designing an experiment that we plan to analyze using ANOCOVA we usually try to assign subjects to treatments so that there is balance across cells (in our problem, there is a sample mean vector of covariates for each of the three cells and we try to effect an assignment which will equalize the three mean vectors). The reason for this design objective is that while estimators of the unknown parameters (and contrasts) remain unbiased regardless of imbalance, so long as the assumptions of the classical Gauss-Markov model are satisfied (see Haggstrom, 1975, P-5449, p. 37), the variances of all contrasts are inflated when there is imbalance. As a result, a greatly unbalanced design can make it very difficult to detect meaningful differences across test cells (since the difference in effects observed in the experiment might be attributed to sampling variation instead of to differences caused by treatment effects). In experiments in which the covariates may be readily controlled by the experimenter, such as those that typically take place in a physical science context, it is often not too difficult to achieve the objective of a balanced design. In social experimentation, however, some of the covariates are usually not very easily controlled, so that it is rarely possible to achieve perfect balance.

The problem of balance in social experiments was studied by Morris, 1979, in the context of a social experiment on health insurance designed at The Rand Corporation. His procedure, called the Finite Selection Model, is designed to select an optimal subset of subjects from a finite population available for potential experimentation, on the basis of the values of covariates each subject is likely to have during the experiment. These subjects are selected one at a time by his model using a nonlinear integer programming procedure with a steepest descents algorithm. The criterion for selection of subjects involves minimizing a weighted sum of variances of linear functions of the estimated parameters of interest, subject to constraints. This minimization is carried out using pre-experimental data comparable to the anticipated experimental data. The Morris approach follows the approach used earlier by Conlisk and Watts, 1969, although the latter procedure did not employ pre-experimental data.

The solution proposed here is appropriate for quite a different context. We describe in this Note a procedure for minimizing imbalance in experimental design in a context in which all subjects are potentially available for experimentation, regardless of their anticipated covariate values. Our problem is how to allocate all available subjects to treatments, and how to set tolerances on the differences of cell means of covariates, so that the design is almost balanced. The criterion function is the minimum average PISE, or percent inflation of the standard error of a linear function of the parameters of interest (averaged over all of the contrasts of interest). Whereas both the Morris and Press procedures for balancing include both first and second moments, the Press procedure compares the PISEs for various possible configurations of subject-to-treatment allocations of the entire population; the Morris procedure adds subjects to the design one by one. So the implementing algorithms are quite different (aside from the fact that one criterion uses inflated standard deviation, while the other uses total variance). The Press procedure has the advantage over other procedures of providing a "natural scale" for the effects of imbalance. The imbalance scale we use measures the percent increase of the standard errors of the quantities of interest over what the standard errors would be in a perfectly balanced situation.

In a social experiment in which we are attempting to detect an effect attributable to some treatment (as in the bonus test), we often want to minimize the standard errors of the contrasts of interest (in order to maximize the power of tests of significance for the contrasts). Such a problem involves both selection of appropriate covariates and balancing across cells. Given a set of covariates, the balancing problem reduces to minimizing the PISE.

Using the PISE criterion, percent inflation of standard errors for the various designs considered for the bonus test ranged from about 3 percent, for the best design, to as much as 25 percent, depending upon which design, which covariate cell mean inequality tolerances, and which contrasts of interest were selected. Average PISE over all three contrasts of potential interest in the bonus test ran as high as 18 percent for some designs considered (before we selected the best one).

In the remainder of this Note we develop statistical models for evaluating the effects of imbalance on the standard errors of the contrasts across the cells. The PISE criterion can be used for comparing potential unbalanced designs on the basis of how much a given design inflates the standard errors of the contrasts (and we propose selecting that design which effects the best compromise between close balance on some covariates and not such close balance on others, so that, overall, inflation of the standard error is minimized for the average of all contrasts of interest). We adopt some of the notation used in Haggstrom, 1975. Our main result is given in Theorem (1).

Section II examines the problem for the classical Gauss-Markov model. Section III provides a result for the more general heteroscedastic model (which is the model we plan to use for the bonus test design). Section IV discusses the problems associated with attempting to attribute the difficulties that arise in connection with imbalance to particular covariates.

II. INFLATION OF STANDARD ERRORS IN CLASSICAL ANOCOVA

MODEL

We adopt the classical ANOCOVA, one way layout, fixed effects model (for given x_{ij} and z_i),

$$y_i = \sum_{j=1}^{p} \beta_j x_{ij} + \gamma' z_i + e_i , \tag{2.1}$$

$i = 1,\ldots,n$, where y_i denotes the response of subject i, there are p groups or cells with $n_1, n_2, \ldots, n_p$ numbers of subjects assigned to each of the p cells, $n = \sum_j n_j$ denotes the total number of subjects, x_{ij} is one, or zero, depending upon whether or not the i^{th} subject is assigned to the j^{th} group, z_i denotes an (h x 1) vector of covariates for subject i, (β_j, γ) are unknown coefficients that must be estimated from the data, and e_i denotes mutually uncorrelated disturbance terms with means zero and variances not depending upon i. Prime denotes the transposed vector.

An alternative formulation involves writing

$$\beta_j = \tau_j + \alpha, \tag{2.2}$$

where α denotes the ambient effect, or the effect in the absence of any treatments, and τ_j denotes the effect of treatment j on a subject in cell j. In this formulation,

$$y_i = \sum_{j=1}^{p} (\tau_j + \alpha) x_{ij} + \gamma' z_i + e_i$$

$$= \alpha + \sum_{j=1}^{p} \tau_j x_{ij} + \gamma' z_i + e_i . \tag{2.3}$$

In our problem of bonus test recruiting, the subjects in the experiment are recruiting centers called Armed Forces Entrance and Examining Stations (AFEES). There are 64 AFEES, so that n = 64. There are three groups (cells), so that p = 3. The continental United States is subdivided into the 64 AFEES named and delineated in Fig. 1.

The balancing problem in the bonus test is to allocate each of the 64 AFEES into one of three cells (which will receive three distinct treatments, that is, three different bonus plans will be used) so that the mean vectors of the covariates of the AFEES in the three cells are equal. We will see below that the effect of such balancing is to minimize the standard deviations of all estimated contrasts (comparisons of treatment effects).

We now reformulate (2.1) in vector terms, to simplify the algebra. Let $\underset{(p \times 1)}{\beta} = (\beta_j)$, $\underset{(p \times 1)}{x_i} = (x_{ij})$, $\underset{(p + h) \times 1}{\delta} = (\beta', \gamma')'$, $\underset{(p + h) \times 1}{w_i} = (x_i', z_i')'$. Now, (2.1) becomes

$$y_i = \delta' w_i + e_i \ , \tag{2.4}$$

$i = 1,\ldots,n$. If we let $\underset{(n \times 1)}{e} = (e_i)$, and write the model as

$$\mathcal{E} = e'e = \sum_{i=1}^{n} (y_i - \delta' w_i)^2 \ , \ E(e) = 0, \ \text{var}\ (e) = \sigma^2 I \ ,$$

it is readily seen that $\mathcal{E}$ is minimized by taking

$$\hat{\delta} = A^{-1} g \ , \tag{2.5}$$

where, for A_{11}: (p x p) , A_{22}: (h x h) ,

$$A \equiv \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} \Sigma x_i x_i' & \Sigma x_i z_i' \\ \Sigma z_i x_i' & \Sigma z_i z_i' \end{pmatrix} , \tag{2.6}$$

Fig. 1 – Armed Forces Entrance and Examining Stations

$$g \equiv \sum_{i=1}^{n} y_i w_i \equiv \begin{pmatrix} \sum_i x_i y_i \\ \\ \sum_i z_i y_i \end{pmatrix} \equiv \begin{pmatrix} g_1 \\ \\ g_2 \end{pmatrix} , \tag{2.7}$$

and $\hat{\delta}$ is the least squares estimator of δ. We must, of course, have $A > 0$; i.e., A is positive definite (and symmetric). It is also easy to find (from (2.5)) that since

$$\text{var}\ (g) = \sigma^2 A ,$$

$$\text{var}\ (\hat{\delta}) = \sigma^2 A^{-1} . \tag{2.8}$$

Define

$$A^{-1} = B \equiv \begin{pmatrix} B_{11} & B_{12} \\ \\ B_{21} & B_{22} \end{pmatrix} .$$

Then,

$$\text{var}\ (\hat{\delta}) = \text{var} \begin{pmatrix} \hat{\beta} \\ \\ \hat{\gamma} \end{pmatrix} = \sigma^2 B .$$

Thus, using, for example, Press, 1982, equations (2.6.3) and (2.6.4),

$$\text{var}\ (\hat{\beta}) = \sigma^2 B_{11} = \sigma^2 (A_{11} - A_{12} A_{22}^{-1} A_{21})^{-1} \equiv \Sigma_{\hat{\beta}} , \tag{2.9}$$

and

$$\text{var}\ (\hat{\gamma}) = \sigma^2 B_{22} = \sigma^2 (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} \equiv \Sigma_{\hat{\gamma}} . \tag{2.10}$$

We also note in passing that cov $(\hat{\beta},\hat{\gamma}) = \sigma^2 B_{12}$.

Let ψ denote a contrast in the effects of the experiment; i.e., $\psi = c'\beta$, where $c \equiv (c_j)$, $\sum_{j=1}^{p} c_j = 0$. For example, for $p = 3$, if we take $c' = (-1,1,0)$, $\psi = \beta_2 - \beta_1$. If $\hat{\beta}$ denotes the least squares estimator of β, $\hat{\psi} = c'\hat{\beta}$ denotes the least squares estimator of ψ. Thus,

$$\Delta \equiv \text{var}\ (\hat{\psi}) = c'[\text{var}\ (\hat{\beta})]c = c'\Sigma_{\hat{\beta}} c . \tag{2.11}$$

We next evaluate $\Sigma_{\hat{\beta}}$.

From (2.5) and (2.7), $A\hat{\delta} = g$ implies $A_{11}\hat{\beta} + A_{12}\hat{\gamma} = g_1$,

or

$$\hat{\beta} = A_{11}^{-1}g_1 - A_{11}^{-1}A_{12}\,\hat{\gamma}. \tag{2.12}$$

It is well known that the error and estimation spaces in the general linear model (under the Gauss-Markov assumptions that we have made) are orthogonal (see, e.g. Scheffé, 1959, p. 23). So $\hat{\gamma}$ and g_1 (depending only on the y_i's, or the errors, for given w_i's) are uncorrelated. Therefore, from (2.12),

$$\Sigma_{\hat{\beta}} = \text{var}\,(A_{11}^{-1}g_1) + \text{var}\,(A_{11}^{-1}A_{12}\hat{\gamma}),$$

or, from (2.11),

$$\Delta = c'A_{11}^{-1}(\text{var } g_1)A_{11}^{-1}c + c'A_{11}^{-1}A_{12}(\text{var } \hat{\gamma})A_{21}A_{11}^{-1}c\ . \tag{2.13}$$

From (2.6) and (2.7), note that

$$\text{var}\,(g_1) = \sum_i x_i[\text{var}\,(y_i)]\,x_i' = \sigma^2\sum_i x_ix_i' = \sigma^2A_{11}\ . \tag{2.14}$$

Substituting (2.14) into (2.13) gives

$$\Delta = \sigma^2c'A_{11}^{-1}c + c'A_{11}^{-1}A_{12}(\text{var } \hat{\gamma})A_{21}A_{11}^{-1}c\ . \tag{2.15}$$

Now note that since x_{ij} denotes an indicator variable, $x_{ij}^2 = x_{ij}$, so that (using (2.6))

$$A_{11} = \begin{pmatrix} n_1 & & 0 \\ & \ddots & \\ 0 & & n_p \end{pmatrix}. \tag{2.16}$$

Thus,

$$c'A_{11}^{-1}c = \sum_{j=1}^{p}\left(\frac{c_j^2}{n_j}\right)\ . \tag{2.17}$$

Substituting (2.17) into (2.15) gives

$$\Delta = \sigma^2 \sum_{j=1}^{p} \left(\frac{c_j^2}{n_j} \right) + \varepsilon, \qquad (2.18)$$

where

$$\varepsilon \equiv c' A_{11}^{-1} A_{12} \Sigma_{\hat{\gamma}} A_{21} A_{11}^{-1} c \ . \qquad (2.19)$$

A similar formulation was given by Harville (1975), p. 220. The value of ε is given in the theorem below. First define

$$\underset{(p \times h)}{Z} \equiv (\bar{z}_1, \ldots, \bar{z}_p)' \ , \qquad (2.20)$$

$$\underset{(h \times h)}{\Phi} \equiv \sum_{i=1}^{n} z_i z_i' - \sum_{j=1}^{p} n_j \bar{z}_j \bar{z}_j' \qquad (2.21)$$

$$= \sum_{j=1}^{p} \sum_{i=1}^{n_j} (z_{ij} - \bar{z}_j)(z_{ij} - \bar{z}_j)' \ ,$$

and

$$\underset{(h \times 1)}{\bar{z}_j} \equiv (\bar{z}_{j1}, \ldots, \bar{z}_{jh})' \ , \ \bar{z}_{jk} \equiv \frac{1}{n_j} \sum_{i=1}^{n} z_{ik} x_{ij} \ . \qquad (2.22)$$

Theorem (1): If ε is the portion of the variance of $\hat{\psi}$ defined in (2.19), ε is given by the quadratic form

$$\varepsilon = \sigma^2 c' W c \equiv \sigma^2 c' (Z \Phi^{-1} Z') c \ . \qquad (2.23)$$

Proof: See the Appendix.

Remark: Note that Φ is the matrix of sums of squares and crossproducts within cells. Therefore, ε is a measure of imbalance of the design Z as measured in its natural metric. ε is also the inflated portion of var $(\hat{\psi})$.

Referring back to the variance of any contrast, given in (2.18), note that the first term of that expression is strictly positive. Therefore, Δ is minimized when $\varepsilon = 0$ (since, by definition in (2.11), Δ must be positive). This case is made specific in the corollary below.

Corollary (1): If the sample means of the covariate vectors are equal, i.e.,

$$\bar{z}_1 = \bar{z}_2 = \ldots = \bar{z}_p = \bar{z} ,$$

ε, defined in (2.19), vanishes.

Remark (1): This well known result (cf. Haggstrom, 1975, P-5449, p. 37) occurs when the design is balanced. Thus, when the design is balanced, W, the matrix of the positive semidefinite quadratic form in (2.23), is not of full rank.

Remark (2): Note that the "balanced design" condition of Corollary (1) is only a sufficient condition for $\varepsilon = 0$. There are other conditions that would also make $\varepsilon = 0$ (we merely need $Z'c = \sum_1^p c_j\bar{z}_j = 0$; so, for example, the combination of $p = 3$, $c = (-1,1,0)'$, and $\bar{z}_1 = \bar{z}_2$, and $\bar{z}_3$ = anything would also yield $\varepsilon = 0$). If, however, we wanted a balanced design regardless of which contrast was of interest, i.e., we wanted $\varepsilon = 0$ for every possible contrast, we would need $Z = (\bar{z},\ldots,\bar{z})'$.

Proof of Corollary (1): Under the hypothesis of the corollary,

$$Z = (\bar{z},\ldots,\bar{z})' .$$

Substituting into (2.23) gives (since ψ is a contrast)

$$Z'c = \bar{z} \sum_{i=1}^{p} c_i = 0 . \qquad (2.24)$$

So $\varepsilon = 0$, and Δ is minimized. This result becomes transparent by noting that $c'(Z\Phi^{-1}Z')c$ in (2.23) does not change if we subtract a constant vector $\bar{z}$ from each column of Z', and replace it by $Z_0' \equiv (\bar{z}_1 - \bar{z},\ldots, \bar{z}_p - \bar{z})$.

INFLATION OF STANDARD ERRORS

Suppose we wish to design an experiment in which we know we will have to tolerate some degree of imbalance. How do we decide among various unbalanced designs? We propose below a natural criterion for making such a decision, in light of Theorem (1).

Define Δ_B as the value of Δ (defined in (2.11)) produced in the case of a balanced design. That is,

$$\Delta_B = \sigma^2 \sum_{j=1}^{p} \left(\frac{c_j^2}{n_j} \right) . \qquad (2.25)$$

Define the percent inflation of the standard error (PISE) of any contrast $\psi = c'\beta$, for any covariate design $Z = (\bar{z}_1, \ldots, \bar{z}_p)'$, as

$$\text{PISE} = \left[\frac{\sqrt{\Delta} - \sqrt{\Delta_B}}{\sqrt{\Delta_B}} \right] (100) . \qquad (2.26)$$

It should be noted that PISE in Eq. (2.26) does not depend upon σ^2.

Suppose we are trying to compare various designs which have the same contrast of interest, but we are free to vary Z so that subjects can be allocated to different treatments in various ways. If we can predetermine Z (at least approximately), based upon external data sources, for each contending design, we can also evaluate PISE for each such design. We might then select that design possessing the minimum value of PISE. Alternatively, by careful study of PISE for various competing designs, we might decide that although one design had a higher PISE than another, the difference was small enough to be tolerable, and the one with the slightly higher PISE should be selected in light of other economic, political, or social factors; by examining the PISE we can evaluate the "cost" of such a tradeoff, in terms of effectiveness of the experiment (loss of probability of finding the effect for which we are designing).

Suppose, alternatively, that there are r contrasts of interest, and we are free to vary Z so that subjects can be allocated to different treatments in various ways. We assume that some contrasts may be more important than others. Let q_i denote the weight to be placed on the i^{th} contrast, $0 < q_i < 1$. We might now select a design that possesses the minimum (weighted) average value of PISE:

$$\overline{PISE} = \sum_{i=1}^{r} \left(\frac{\sqrt{\Delta_i} - \sqrt{\Delta_{B_i}}}{\sqrt{\Delta_{B_i}}} \right) (100q_i) \ , \qquad (2.27)$$

where $\Delta_i \equiv \text{var}(\hat{\psi}_i) \equiv \text{var}(C_i'\hat{\beta})$, and C_i: (p x 1) denotes the weight vector for the i^{th} contrast. This type of averaging is in the spirit of Cox, 1982, p. 197.

For example, suppose there are three cells (two test cells and a control cell) and two simple contrasts of interest, namely

$$\psi_1 = \beta_2 - \beta_1, \quad \psi_2 = \beta_3 - \beta_1 \ .$$

Then, $C_1 = (-1, 1, 0)'$ and $C_2 = (-1, 0, 1)'$. Suppose further that ψ_1 and ψ_2 are of equal interest and importance, so that $q_1 = q_2 = \frac{1}{2}$. Then, from (2.27), for r = 2, if $\Delta_{B_1} = \Delta_{B_2} = \Delta_B$,

$$\overline{PISE} = \frac{(\sqrt{\Delta_1} + \sqrt{\Delta_2} - 2\sqrt{\Delta_B})(100)}{2\sqrt{\Delta_B}} \ .$$

III. INFLATION OF STANDARD ERRORS IN WEIGHTED ANOCOVA

The model treated in Sec. II adopted the Gauss-Markov assumptions $E(e) = 0$, $E(e_i e_j) = 0$, $i \neq j$, and $\text{var}\ (e_i) = \sigma^2$: $i = 1,\ldots,n$. But suppose, alternatively, that

$$\text{var}\ (e_i) = a_i \sigma^2 , \tag{3.1}$$

while the other assumptions remain the same. That is, we have heteroscedasticity. Suppose the a_i's are known constants, however. We see below that the results of Sec. II are readily extended to cover this case as well. (This is the case of interest in the bonus experiment).

Adopt the model of (2.1), but with the heteroscedasticity assumption of (3.1). Define the transformed variables:

$$y_i^* = y_i (a_i)^{-1/2},\ e_i^* = e_i (a_i)^{-1/2},\ z_i^* = z_i (a_i)^{-1/2},$$

$$x_{ij}^* = x_{ij} (a_i)^{-1/2} . \tag{3.2}$$

The transformed model becomes (for given x_{ij}^* and z_i^*),

$$y_i^* = \sum_{j=1}^{p} \beta_j x_{ij}^* + \gamma' z_i^* + e_i^* , \tag{3.3}$$

for $i = 1,\ldots,n$, with

$$E(e_i^*) = 0,\ E(e_i^* e_j^*) = 0,\ i \neq j,\ \text{var}\ (e_i^*) = \sigma^2 . \tag{3.4}$$

The model now has the general form assumed in Sec. II, in terms of transformed variables. So the corresponding results are immediately obtainable. The basic result in Theorem (1) becomes (instead of (2.23))

$$\varepsilon^* = \sigma^2 c' W^* c \equiv \sigma^2 c' (Z^* \Phi^{*-1} Z^{*\prime}) c, \tag{3.5}$$

where

$$Z^* \equiv (\bar{z}_1^*, \ldots, \bar{z}_p^*)' , \qquad (3.6)$$

$$\underset{(h \times 1)}{\bar{z}_j^*} = (\bar{z}_{j1}^*, \ldots, \bar{z}_{jh}^*)' , \quad \bar{z}_{jk}^* \equiv \frac{1}{n_j^*} \sum_{i=1}^{n} z_{ik}^* x_{ij}^* , \qquad (3.7)$$

$$n_j^* = \sum_{i=1}^{n} (x_{ij}^{*2}) = \sum_{i=1}^{n} \left(\frac{x_{ij}^2}{a_i}\right) = \sum_{i=1}^{n} \left(\frac{x_{ij}}{a_i}\right) , \qquad (3.8)$$

and

$$\Phi^* = \sum_{i=1}^{n} z_i^* z_i^{*\prime} - \sum_{j=1}^{p} n_j^* \bar{z}_j^* \bar{z}_j^{*\prime} . \qquad (3.9)$$

Note that Eqs. (3.7) through (3.9) are weighted averages, weighted sums, or weighted sums of squares, as contrasted with their unweighted analogs in Sec. II.

The analog of (2.18) becomes

$$\Delta = \sigma^2 \sum_{j=1}^{p} \left(\frac{c_j^2}{n_j^*}\right) + \varepsilon^* , \qquad (3.10)$$

with ε^* defined in (3.5), and $\Delta = \text{var}\,(\hat{\psi}) = \text{var}\,(c'\hat{\beta})$. So if

$$\Delta_B^* = \sigma^2 \sum_{j=1}^{p} \left(\frac{c_j^2}{n_j^*}\right) \qquad (3.11)$$

denotes the variance of a contrast for a balanced design (one with $\varepsilon^* = 0$), the selection criterion becomes, for the case of heteroscedasticity,

$$\text{PISE} = \left[\frac{\sqrt{\Delta} - \sqrt{\Delta_B^*}}{\sqrt{\Delta_B^*}}\right] (100) , \qquad (3.12)$$

where Δ and Δ_B^* are defined in (3.10) and (3.11). Note again, as in (2.26), PISE does not depend upon σ^2.

To apply the PISE criterion in the case of heteroscedasticity we first transform the covariate vectors to form Z^* and Φ^*. Then, for a given contrast of interest (choice of c), we can evaluate PISE for various competing designs.

σ^2 is usually not known, but in some situations we may wish to estimate it. For example, suppose our original dependent variables are Y_i, and the Y_i's are independently distributed Poisson variates with means λ_i. If we transform the model so that $y_i \equiv \log(Y_i + 1/2)$, where y_i is given by (2.1), it may be shown (see, e.g., Cox, 1955) that approximately,

$$\text{var}\ (y_i) = \text{var}\ [\log\ (Y_i + 1/2)] \cong \frac{1}{\lambda_i}\ . \tag{3.13}$$

But in the Poisson distribution, we can use the approximation

$$\lambda_i = E(Y_i) \cong Y_i\ .$$

Substituting in (3.13) gives:

$$\text{var}\ (y_i) \cong \frac{1}{Y_i}\ .$$

But from (3.1) we can take

$$\text{var}\ (y_i) = \text{var}\ (e_i) \equiv a_i\sigma^2 \cong \frac{1}{Y_i}\ .$$

So we can take $\sigma^2 \equiv 1$ in this instance, but take

$$a_i = \frac{1}{Y_i}\ . \tag{3.14}$$

Thus, we adopt the model in (3.3), with a_i as in (3.14) and $\sigma^2 = 1$. We approximate Y_i, prior to an experiment, by using previously obtained data comparable to Y_i.

IV. THE PROBLEM OF ATTRIBUTING INFLATED STANDARD ERRORS TO A GIVEN SOURCE

Once the degree of inflation of the standard errors of contrasts caused by an unbalanced design has been determined, it is tempting to try to separate the sources of the inflation by attributing these sources to the distinct covariates. This is very difficult to do, except in unusual circumstances. For simplicity, we examine the nature of this problem in terms of variances instead of standard deviations, and exhibit the behavior of the inflation explicitly for several sharply defined situations relating to a single contrast. Other cases are more complicated.

The proportion of a contrast variance attributable to imbalance of the design is given in Theorem (1), (2.23), for the classical Gauss-Markov model as

$$\varepsilon = \sigma^2 c'(Z\Phi^{-1}Z')c \; . \qquad (4.1)$$

That is, ε denotes the inflation of the variance.

Now suppose that all of the covariates are mutually uncorrelated, so that if σ_α^2 denotes the (sample) variance of the α^{th} covariate,

$$\Phi = n \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_h^2 \end{pmatrix} .$$

Then

$$\varepsilon = \sigma^2 \sum_{\alpha=1}^{h} \sum_{i=1}^{p} \sum_{j=1}^{p} c_i c_j \left(\frac{\bar{z}_{i\alpha}\, \bar{z}_{j\alpha}}{n\sigma_\alpha^2} \right) = \sum_{\alpha=1}^{h} f_\alpha \; ,$$

where

$$f_\alpha = \left(\frac{\sigma^2}{\sigma_\alpha^2} \right) \sum_{i=1}^{p} \sum_{j=1}^{p} c_i c_j \left(\frac{\bar{z}_{i\alpha}\, \bar{z}_{j\alpha}}{n} \right) . \qquad (4.2)$$

Define

$$P_\alpha \equiv \frac{f_\alpha}{\varepsilon} \quad . \tag{4.3}$$

Note that P_α is the proportion of the inflated variance attributable to the $\alpha^{\underline{th}}$ covariate.

Consider a particular (simple) contrast,

$$\underset{(p \times 1)}{c} = (-1,1,0,\ldots,0)' \ .$$

In this case, from (4.2),

$$f_\alpha = \left(\frac{\bar{z}_{1\alpha} - \bar{z}_{2\alpha}}{\sigma_\alpha} \right)^2 \left(\frac{\sigma^2}{n} \right) \ .$$

Let $\theta_\alpha \equiv (\bar{z}_{1\alpha} - \bar{z}_{2\alpha})$ denote the balancing "tolerance" for covariate α. That is, when we attempt to balance the design, we attempt to make the sample means for all covariates the same in all of the cells; θ_α is the amount by which the cell means differ for cells 1 and 2, for covariate α. Thus,

$$f_\alpha = \frac{\theta_\alpha^2}{n\sigma_\alpha^2} \sigma^2 \ ,$$

and, from (4.3),

$$P_\alpha = \frac{(\theta_\alpha^2/\sigma_\alpha^2)}{\sum_{\alpha=1}^{h} (\theta_\alpha^2/\sigma_\alpha^2)} \ . \tag{4.4}$$

Note that for a particular α, say $\alpha = 1$, we may express (4.4) in the form

$$P_1 = \frac{(\theta_1^2/\sigma_1^2)}{(\theta_1^2/\sigma_1^2) + K} = \frac{1}{1 + K(\sigma_1^2/\theta_1^2)} \ , \tag{4.5}$$

where $K = \sum_{\alpha=2}^{h} (\theta_\alpha^2/\sigma_\alpha^2)$. Figure 2 shows how the proportion of inflated variance attributable to the first covariate varies with balancing tolerance and covariate standard deviation. We see that P_1 increases as the square of balancing tolerance, so that the further apart are the covariate cell means,

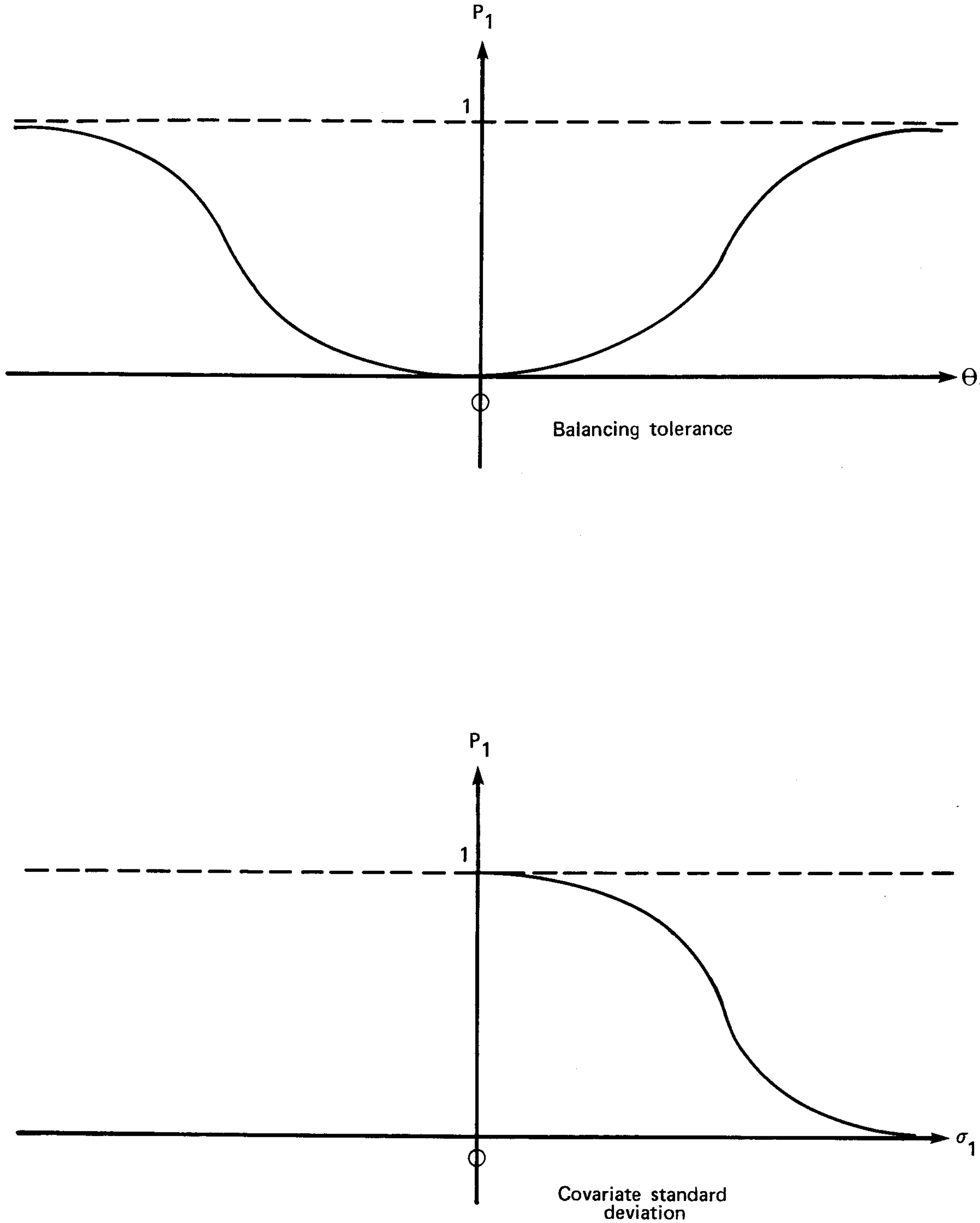

Fig. 2 — Proportion of inflated variance attributable to balancing tolerance and covariate standard deviation

the greater is the variance inflation (increasing by the square of the tolerance). Moreover, P_1 decreases with increasing covariate standard deviation, so P_1 is greatest when the first covariate is not free to vary much. Thus, referring to Eq. (4.2), when the covariates are uncorrelated we can minimize the inflated portion of the total variance by choosing covariates which achieve acceptably low values of $(\sigma^2/\sigma_\alpha^2)$. Accordingly, if there were a choice between two covariates, each of which reduced σ^2 by the same amount, but one had a larger σ_α^2, we should select the one with the larger σ_α^2.

We see that because in most ANOCOVA designs the covariates are correlated, we cannot break out the effects due to each covariate separately. The greater the correlations among covariates the more difficult it will be to analyze the individual covariate effects separately, as in Fig. 2. In the general case, therefore, we must be content to evaluate the PISE criterion for the design as a whole, and then to make PISE as small as possible.

Appendix

Proof of Theorem (1): From the definition of ε, in (2.19), and that of $\Sigma_{\hat{\gamma}}$, in (2.10), we have

$$\varepsilon = \sigma^2 c' A_{11}^{-1} A_{12} (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} A_{21} A_{11}^{-1} c \ . \qquad \text{(A.1)}$$

From (2.6),

$$\underset{(p \times h)}{A_{12}} = \sum_{i=1}^{n} x_i z_i' = \begin{pmatrix} n_1 \bar{z}_1' \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ n_p \bar{z}_p' \end{pmatrix} .$$

Using (2.16) and (2.20), we find the representation

$$A_{12} = A_{11} Z \ ; \ A_{21} = Z' A_{11} \ . \qquad \text{(A.2)}$$

Substituting (A.2) into (A.1) gives

$$\varepsilon = \sigma^2 c' Z (A_{22} - Z' A_{11} Z)^{-1} Z' c \ . \qquad \text{(A.3)}$$

From (2.6) and (2.21), it is readily seen that

$$\Phi = A_{22} - Z' A_{11} Z \ . \qquad \text{(A.4)}$$

Substituting (A.4) into (A.3) gives the result in (2.23).

REFERENCES

Conlisk, John, and Harold Watts, "A Model for Optimizing Experimental Designs for Estimating Response Surfaces," Proceedings of the Social Statistics Section, American Statistical Association, 1969, pp. 150-156.

Cox, D. R., "Randomization and Concomitant Variables in the Design of Experiments," in G. Kallianpur, P. R. Krishnaiah, and J. K. Ghosh (eds.), Statistics and Probability: Essays in Honor of C. R. Rao, North Holland Publishing Co. Inc., Amsterdam, 1982.

Cox, D. R., "Some Statistical Methods Connected with Series of Events," Journal of the Royal Statistical Society (B), Vol. 17, 1955, pp. 129-157.

Haggstrom, G. W., The Pitfalls of Manpower Experimentation, The Rand Corporation, P-5449, April 1975; also in Sinaiko and Laurie A. Broedling (eds.), Perspectives on Attitude Assessment: Surveys and Their Alternatives, Smithsonian Institution, Washington, D.C., 1975.

Harville, D. A., "Computing Optimum Designs for Covariance Models," in J. N. Svivastavia (ed.), A Survey of Statistical Design and Linear Models, North Holland Publishing Co., Inc., Amsterdam, 1975.

Morris, C., "A Finite Selection Model for Experimental Design of the Health Insurance Study," Journal of Econometrics, Vol. 11, 1979, pp. 43-61.

Press, S. J., Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference, 2nd Edition Revised, Krieger Publishing Co., Melbourne, Florida, 1982.

Rubin, Donald B., "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies," Journal of Educational Psychology, Vol. 66, No. 5, 1974, pp. 688-701.

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

RAND/N-1890-MRAL