THEMIS SIGNAL ANALYSIS STATISTICS RESEARCH PROGRAM

BIVARIATE PROBIT, LOGIT, AND BURRIT ANALYSIS

by

Frederick C. Durling

Technical Report No. 41
Department of Statistics THEMIS Contract

August 8, 1969

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

DEPARTMENT OF STATISTICS
Southern Methodist University

BIVARIATE PROBIT, LOGIT, AND BURRIT ANALYSIS

A Thesis Presented to the Faculty of the Graduate School

of

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

with a

Major in Statistics

by

Frederick C. Durling
(M.S., Southern Methodist University, 1967)

August 6, 1969

Durling, Frederick C.

B.A., Arlington State College, 1965
M.S., Southern Methodist University, 1967

<u>Bivariate Probit, Logit, and Burrit Analysis</u>

Advisor: Professor Donald B. Owen and Associate Professor Wanzer Drane

Doctor of Philosophy degree conferred August 22, 1969

Thesis completed August 6, 1969

The problem of a mixture of two stimulants in a biological quantal assay is investigated from a mathematical standpoint. The basic assumption is made that the response region does not depend on biological considerations - i.e., given a specified mixture of stimulants $\underline{z}$, the response region is defined by the point $\underline{z}'$ in the p-variate space where there are p stimulants under consideration; instead, the probability functions, themselves, may take on different forms. A general form is proposed and investigated. Three analytic models (one utilizing the bivariate normal distribution, one a bivariate logistic distribution developed by Gumbel (1961), and one a bivariate Burr distribution developed by this author) are employed in this investigation. The investigation includes the analysis of data, under the three analytic models, which had been classified by previous investigators as examples of synergistic action, simple similar action, independent action, and additive action. The residual analyses are included as well as the FORTRAN IV subroutines used in evaluating the functions, the partial derivatives and the weights.

The investigation lends some support to the assumption of a constant response region for a diversity of mixtures of stimulants. The analytic

model incorporating the bivariate Burr distribution is recommended for all cases unless the number of parameters to be estimated is a primary concern, in which case the analytic model utilizing the bivariate normal distribution is recommended. The bivariate Burr distribution developed in this paper is found to be more useful in application than that developed by Takahasi (1965).

ACKNOWLEDGMENTS

I wish to express my deep appreciation to Professor Donald B. Owen and Associate Professor J. Wanzer Drane for their kind and able assistance throughout the work on this dissertation. I am also indebted to the faculty of the Department of Statistics for reading and constructively commenting on the final draft. An added word of appreciation must go to Associate Professor C. H. Kapadia for his most generous interest and support during my years as a graduate student. A special word of appreciation is due Professor Paul D. Minton, Chairman of the Department of Statistics at Southern Methodist University, for awarding to me a National Institutes of Health fellowship which made graduate training possible and for his generous giving of himself to my needs as a student and fellow human being. I would also like to thank Mrs. Linda White for the excellent job of typing the final form of this dissertation.

I am indeed grateful for the encouragement and support given to me during my graduate training by my many friends.

TABLE OF CONTENTS

CHAPTER I

1. Introduction

The joint action of mixtures of stimulants in a biological assay has been investigated by Bliss (1939), Finney (1942), Plackett and Hewlett (1967), Ashford and Smith (1966), and others. Plackett and Hewlett have made their investigations largely from the standpoint of biological considerations such as the physiology of the biological organism being used in experimentation. Ashford and Smith, on the other hand, have dealt with the problem somewhat more within a mathematical framework. In this paper, the problem will be approached mathematically.

For the purposes of this paper a biological assay of a mixture of two stimulants will be conducted as follows: A population of N organisms is divided at random into t groups, where the i^{th} group is of size n_i, $n_1 + n_2 + \cdots + n_t = N$. The i^{th} group receives a treatment of a predetermined mixture (z_{1i}, z_{2i}) of two drugs, where z_{ji} is the quantity of stimulant j measured in any convenient units. r_i is the observed number which manifest a prescribed quantal response. The observed relative frequency of response $p_i = r_i/n_i$ is an estimate of the probability of an organism responding if picked at random from the population. The probability that this organism picked at random will respond when treated by the mixture (z_{1i}, z_{2i}) may be assumed to take on a general form, say, $P(z_{1i}, z_{2i}, \underline{\Theta})$.

Now the probability of r_i responses with the i^{th} combination of levels of drugs can be written as

$$P(r_i) = \frac{n_i!}{r_i!(n_i-r_i)!}\left[P(z_{1i}, z_{2i}, \underline{\Theta})\right]^{r_i}\left[1 - P(z_{1i}, z_{2i}, \underline{\Theta})\right]^{n_i-r_i}$$

$$r_i = 0, 1, 2, \cdots, t \qquad (1)$$

$$= 0 \qquad \text{elsewhere.}$$

A series of t combinations of doses is tested in an experiment. The probability of a particular set of r_i's is equal to exp(L) , the likelihood, where

$$L = \sum_{i=1}^{t} r_i \ln (P_i) + \sum_{i=1}^{t} (n_i-r_i)\ln(Q_i) + \sum_{i=1}^{t} \ln[n_i!/r_i!(n_i-r_i)!] \qquad (2)$$

and $P_i = P(z_{1i}, z_{2i}, \underline{\Theta})$, $Q_i = 1 - P_i$. The maximum likelihood estimator $\hat{\Theta}$ of a parameter Θ , Θ an element of $\underline{\Theta}$, must satisfy the relation

$$0 = \frac{\partial L}{\partial \Theta} = \sum_{i=1}^{t} \frac{r_i}{P_i}\frac{\partial P_i}{\partial \Theta} + \sum_{i=1}^{t} \frac{(n_i-r_i)}{Q_i}\frac{\partial Q_i}{\partial \Theta} = \sum_{i=1}^{t} \frac{n_i(p_i-P_i)}{P_iQ_i}\frac{\partial P_i}{\partial \Theta} \; . \qquad (3)$$

Direct solution for Θ is not in general possible, but iterative techniques are available which give a convergent series of approximations to the solution.

The following procedure for two parameters Θ and ϕ , Θ and ϕ elements of $\underline{\Theta}$, is of completely general applicability and may easily be extended for the estimation of a greater number of parameters. By the Taylor-Maclaurin expansion of $\frac{\partial L}{\partial \Theta}$, $\frac{\partial L}{\partial \phi}$ (See relation (3).) ignoring quantities containing terms of higher than the first degree

$$\frac{\partial L}{\partial \Theta_1} + \delta_\Theta \frac{\partial^2 L}{\partial \Theta_1^2} + \delta_\phi \frac{\partial^2 L}{\partial \Theta_1 \partial \phi_1} = 0 \qquad (4)$$

$$\frac{\partial L}{\partial \phi_1} + \delta_\phi \frac{\partial^2 L}{\partial \phi_1^2} + \delta_\Theta \frac{\partial^2 L}{\partial \Theta_1 \partial \phi_1} = 0 \quad ,$$

where the addition of the suffix 1 to Θ , ϕ indicated that the first approximations are to be substituted after differentiation. The solutions δ_Θ , δ_ϕ are adjustments to Θ_1 , ϕ_1 which give the improved approximations $\Theta_2 = \Theta_1 + \delta_\Theta$, $\phi_2 = \phi_1 + \delta_\phi$.

Equations (4) may be simplified through the following procedure which will be illustrated by means of the first of equations (4).

$$\frac{\partial L}{\partial \Theta_1} + \delta_\Theta \frac{\partial^2 L}{\partial \Theta_1^2} + \delta_\phi \frac{\partial^2 L}{\partial \Theta_1 \partial \phi_1} = 0 \quad ,$$

or

$$- \delta_\Theta \frac{\partial^2 L}{\partial \Theta_1^2} - \delta_\phi \frac{\partial^2 L}{\partial \Theta_1 \partial \phi_1} = \frac{\partial L}{\partial \Theta_1} \quad .$$

$$- \delta_\Theta \frac{\partial}{\partial \Theta_1}\left[\frac{\partial L}{\partial \Theta_1}\right] - \delta_\phi \frac{\partial}{\partial \phi_1}\left[\frac{\partial L}{\partial \Theta_1}\right] = \frac{\partial L}{\partial \Theta_1} \quad ;$$

$$- \delta_\Theta \frac{\partial}{\partial \Theta_1}\left\{\sum_{i=1}^{t} \frac{n_i p_i}{P_i Q_i} \frac{\partial P_i}{\partial \Theta_1} - \sum_{i=1}^{t} \frac{n_i}{Q_i} \frac{\partial P_i}{\partial \Theta_1}\right\}$$

$$- \delta_\phi \frac{\partial}{\partial \phi_1}\left\{\sum_{i=1}^{t} \frac{n_i p_i}{P_i Q_i} \frac{\partial P_i}{\partial \Theta_1} - \sum_{i=1}^{t} \frac{n_i}{Q_i} \frac{\partial P_i}{\partial \Theta_1}\right\} = \frac{\partial L}{\partial \Theta_1} \quad ;$$

$$- \delta_\Theta \left\{\sum_{i=1}^{t} \frac{n_i p_i}{P_i Q_i} \frac{\partial^2 P_i}{\partial \Theta_1^2} - \sum_{i=1}^{t} \frac{n_i p_i}{P_i^2 Q_i} \left(\frac{\partial P_i}{\partial \Theta_1}\right)^2 + \sum_{i=1}^{t} \frac{n_i p_i}{P_i Q_i^2} \left(\frac{\partial P_i}{\partial \Theta_1}\right)^2 \right.$$

$$\left. - \sum_{i=1}^{t} \frac{n_i}{Q_i} \frac{\partial^2 P_i}{\partial \Theta_1^2} - \sum_{i=1}^{t} \frac{n_i}{Q_i^2} \left(\frac{\partial P_i}{\partial \Theta_1}\right)^2 \right\}$$

$$- \delta_\phi \left\{ \sum_{i=1}^{t} \frac{n_i p_i}{P_i Q_i} \frac{\partial^2 P_i}{\partial \Theta_1 \partial \phi_1} - \sum_{i=1}^{t} \frac{n_i p_i}{P_i^2 Q_i} \frac{\partial P_i}{\partial \Theta_1} \frac{\partial P_i}{\partial \phi_1} \right.$$

$$+ \sum_{i=1}^{t} \frac{n_i p_i}{P_i Q_i^2} \frac{\partial P_i}{\partial \Theta_1} \frac{\partial P_i}{\partial \phi_1} - \sum_{i=1}^{t} \frac{n_i}{Q_i} \frac{\partial^2 P_i}{\partial \Theta_1 \partial \phi_1}$$

$$\left. - \sum_{i=1}^{t} \frac{n_i}{Q_i^2} \frac{\partial P_i}{\partial \Theta_1} \frac{\partial P_i}{\partial \phi_1} \right\} = \frac{\partial L}{\partial \Theta_1} \quad .$$

At this stage the equation may be simplified by putting $p_i = P_i$ in the coefficients of $\frac{\partial^2 L}{\partial \Theta_1^2}$, $\frac{\partial^2 L}{\partial \Theta_1 \partial \phi_1}$, i.e., on the left hand side of the last equation to give expected instead of empirical values. The last equation then reduces to

$$\delta_\Theta \sum_{i=1}^{t} \frac{n_i}{P_i Q_i} \left(\frac{\partial P_i}{\partial \Theta_1}\right)^2 + \delta_\phi \sum_{i=1}^{t} \frac{n_i}{P_i Q_i} \left(\frac{\partial P_i}{\partial \Theta_1}\right)\left(\frac{\partial P_i}{\partial \phi_1}\right) = \frac{\partial L}{\partial \Theta_1} \quad .$$

The latter equation in (4) can be reduced by means of a similar procedure, i.e., putting $p_i = P_i$ in the coefficients of $\frac{\partial^2 L}{\partial \phi_1^2}$, $\frac{\partial^2 L}{\partial \Theta_1 \partial \phi_1}$. Thus equations (4) are simplified to

$$\delta_\Theta \sum_{i=1}^{t} \frac{n_i}{P_{i1} Q_{i1}} \left(\frac{\partial P_i}{\partial \Theta_1}\right)^2 + \delta_\phi \sum_{i=1}^{t} \frac{n_i}{P_{i1} Q_{i1}} \left(\frac{\partial P_i}{\partial \Theta_1}\right)\left(\frac{\partial P_i}{\partial \phi_1}\right) = \sum_{i=1}^{t} \frac{n_i (p_i - P_{i1})}{P_{i1} Q_{i1}} \left(\frac{\partial P_i}{\partial \Theta_1}\right) ;$$

$$(5)$$

$$\delta_\Theta \sum_{i=1}^{t} \frac{n_i}{P_{i1} Q_{i1}} \left(\frac{\partial P_i}{\partial \Theta_1}\right)\left(\frac{\partial P_i}{\partial \phi_1}\right) + \delta_\phi \sum_{i=1}^{t} \frac{n_i}{P_{i1} Q_{i1}} \left(\frac{\partial P_i}{\partial \phi_1}\right)^2 = \sum_{i=1}^{t} \frac{n_i (p_i - P_{i1})}{P_{i1} Q_{i1}} \left(\frac{\partial P_i}{\partial \phi_1}\right) .$$

Here the addition of the suffix 1 to P_{i1} , Q_{i1} indicates that the first approximations are to be used in the evaluation of $P(z_{1i}, z_{2i}, \underline{\Theta})$. Equations (5) illustrate that only first derivatives are needed in this iterative procedure.

2. Methodology for Obtaining Estimates

Now, it will be seen from what follows that relation (3) can be solved by using a modified non-linear least squares [Moore and Zeigler (1967)]. Assume that the data corresponds to the mathematical model

$$y_i = h(\underline{z}_i \ , \underline{\alpha}) + \varepsilon_i \ , \quad i = 1, 2, \cdots , t \tag{6}$$

where the y_i are observed random variables, $\underline{z}_i$ is a vector of known independent variables, $\underline{\alpha}$ is a vector of unknown parameters, and ε_i is a random variable such that $E(\varepsilon_i) = 0$, $E(\varepsilon_i^2) = \sigma_i^2$, and $E(\varepsilon_i \varepsilon_j) = 0$ for all $i \neq j$. Then the vector of unknown parameters may be estimated by minimizing the weighted sum of squares,

$$S = \sum_{i=1}^{t} \Big(y_i - h(\underline{z}_i \ , \underline{\alpha})\Big)^2 W_i \ , \tag{7}$$

where W_i is an appropriate weight. If the usual procedure is modified so that the partial derivatives are taken ignoring W_i , the normal equations are

$$\frac{\partial S}{\partial \alpha_k} = - 2 \sum_{i=1}^{t} W_i [y_i - h(\underline{z}_i \ , \underline{\alpha})] \frac{\partial h(\underline{z}_i \ , \underline{\alpha})}{\partial \alpha_k} = 0 \ , \tag{8}$$

for $k = 1, 2, \cdots , \ell$, where ℓ is the number of unknown parameters. Now by letting $W_i = n_i / P_i Q_i$ (the reciprocal of the variance of p_i) , $y_i = p_i$, $h(\underline{z}_i \ , \underline{\alpha}) = P_i$, $Q_i = 1 - P_i$, and $\alpha_k = \theta$ it can be seen that relation (3) and equation (8) are equivalent. Thus the maximum likelihood estimate can be obtained by means of a modified weighted non-linear least squares. Relations (5) and their equivalent extensions are used in the modified non-linear least squares fitting of equation (6).

3. Consideration of Necessary Conditions on $P(z_{1i}, z_{2i}, \underline{\theta})$

From this point onward the vector $\underline{z}_i = \begin{pmatrix} z_{1i} \\ z_{2i} \end{pmatrix}$ will be considered from the standpoint of a mixture of stimulants where a transformation has been applied to the original dosage levels so that $-\infty$ is equivalent to zero dosage and $+\infty$ is equivalent to an infinite dosage. For the purposes of this paper $P(z_{1i}, z_{2i}, \underline{\theta})$ must satisfy the following conditions

$$P(z_{1i}, +\infty, \underline{\theta}) = P(+\infty, z_{2i}, \underline{\theta}) = 1 \quad , \tag{9}$$

$$P(z_{1i}, -\infty, \underline{\theta}) = P_1(z_{1i}, \underline{\theta}_1) \quad , \tag{10}$$

and

$$P(-\infty, z_{2i}, \underline{\theta}) = P_2(z_{2i}, \underline{\theta}_2) \quad , \tag{11}$$

where $P_1(z_{1i}, \underline{\theta}_1)$ and $P_2(z_{2i}, \underline{\theta}_2)$ are not in general zero, but rather are marginal probabilities, i.e., the probability of a random individual biological organism responding if it is given a dosage of stimulant j corresponding to z_{ji}. Conditions (9), (10), and (11) imply the conditions

$$P(-\infty, -\infty, \underline{\theta}) = 0 \quad , \tag{12}$$

and

$$P(+\infty, +\infty, \underline{\theta}) = 1 \quad . \tag{13}$$

All five of these conditions are necessary in a bioassay of quantal response data involving a mixture of two stimulants. Natural extensions of these conditions for a mixture of more than two stimulants are now obvious.

Plackett and Hewlett (1967) proposed that

$$P = \int_R f(z_{1i}, z_{2i})dz_{1i}dz_{2i} \tag{14}$$

where $f(z_{1i}, z_{2i})$ is a bivariate density with the usual properties and R is defined on the basis of biological considerations, thus implying that the region of integration may, arbitrarily, be changed due to biological considerations. Their papers do not indicate any homogeneity in the regions. Nowhere is there a general formulation for $P(z_{1i}, z_{2i}, \underline{\theta})$ where the form of the region of integration is homogeneous, much less constant. It would appear that the region of integration should be constant except possibly for simple monotonic transformations of the original dosage levels, such as a logarithmic transformation. The bivariate function itself might be, in specific instances, of different types but still retaining a constant response region.

Let $F_1(z_1, \underline{\theta}_1)$, $F_2(z_2, \underline{\theta}_2)$ be univariate distributions where the parameter vectors $\underline{\theta}_1$, $\underline{\theta}_2$ are not, in general, equal. Note that $F_1(z_1, \underline{\theta}_1)$, $F_2(z_2, \underline{\theta}_2)$ are not necessarily even from the same family of distributions, e.g., the family of normal distributions. Let $F_3(z_1, z_2, \underline{\theta})$ be a bivariate distribution such that $F_3(z_1, +\infty, \underline{\theta}) = F_1(z_1, \underline{\theta}_1)$ and $F_3(+\infty, z_2, \underline{\theta}_2) = F_2(z_2, \underline{\theta}_2)$, where $\underline{\theta}' = (\underline{\theta}_1', \underline{\theta}_2', \underline{\theta}_3')$. Now, what is needed is a function which satisfies conditions (9) through (13).

Let

$$H(z_1, z_2, \underline{\theta}) = F_1(z_1, \underline{\theta}_1) + F_2(z_2, \underline{\theta}_2) - F_3(z_1, z_2, \underline{\theta}) \quad . \tag{15}$$

Then $H(-\infty, z_2, \underline{\theta}) = F_2(z_2, \underline{\theta}_2)$, $H(z_1, -\infty, \underline{\theta}) = F_1(z_1, \underline{\theta}_1)$, $H(+\infty, z_2, \underline{\theta}) = 1 = H(z_1, +\infty, \underline{\theta})$, $H(-\infty, -\infty, \underline{\theta}) = 0$, and $H(+\infty, +\infty, \underline{\theta}) = 1$. Thus $H(z_1, z_2, \underline{\theta})$ does satisfy conditions (9) through (13) which suggests that

$$P(z_{1i}, z_{2i}, \underline{\theta}) = F_1(z_1, \underline{\theta}_1) + F_2(z_2, \underline{\theta}_2) - F_3(z_1, z_2, \underline{\theta}) \tag{16}$$

is a general formulation for $P(z_{1i}, z_{2i}, \underline{\theta})$ where the response region is constant. Note that the forms $F_1(z_{1i}, \underline{\theta}_1)$, $F_2(z_{2i}, \underline{\theta}_2)$, $F_3(z_{1i}, z_{2i}, \underline{\theta})$

are completely general distributions whose forms can depend on biological, chemical, or other considerations. At the same time the region of integration is constant and easily understood from a geometrical standpoint as well as from other standpoints. It is noted here that the general formulation for $P(z_{1i}, z_{2i}, \underline{\theta})$ can easily be extended for a vector of more than two stimulants. The utility of this form is quite general; the only restrictions being conditions (9) through (13) which have been imposed in the development of the general form in equation (16).

CHAPTER II

It is natural in the study of a mixture of two stimulants to consider a bivariate probit or normit. Probit analysis has no advantage over a normit analysis if the analysis is run on a high-speed computer. Also, the analyses are equivalent. Almost all of the work that has been done to date has been along the lines of a bivariate normit.

Bliss (1939) was among the first to study the action of mixtures of two stimulants. He classified the joint action of two stimulants into three biological categories: independent joint action, similar joint action, and synergistic action. Independent joint action occurs whenever two components act on different vital systems in the organism and do not interact with one another. Similar joint action is observed whenever two components act independently of one another but on the same vital system. Synergistic action is characterized by a larger frequency of response than could be predicted from experiments using the individual stimulants. He mentions antagonistic action but did not treat this concept at all. He stated that it is the reverse of synergistic action. Finney (1942) suggested that antagonism is negative synergism and can thus be treated in the same category as synergism.

Bliss (1939), for the category of independent joint action, plotted expected response in probits against dosage of mixture in logarithms. At each point the ratio of the amount of a given stimulant to the amount of the other stimulant was held constant. These curves were not smooth but rather fell into two segments each of which appeared to be a straight line. The transition from one straight line to the other was relatively abrupt. He suggested the equation

$$p_C = p_A + p_B(1 - p_A)(1 - r) \tag{17}$$

where $p_A > p_B$, p_A is the probability of response due to the effect of stimulant A , p_B is the probability of response due to stimulant B , r is a measure of "association of susceptibilities," and p_C is the probability of death due to the combination of stimulants A and B . He did not indicate what, if any, relation he assumed between equation (17) and the plots of data.

For the category of similar joint action Bliss suggested equation

$$Y_C = a' + b \log(D_A + kD_B) \tag{18}$$

for the dosage response curve, where D_A and D_B are the respective doses of stimulants A and B in the mixture and k is the ratio of the frequency of response of the individual stimulants. The plots for this case are thus straight lines.

Bliss suggested two possible equations for synergistic action. The first, which relates the total amount of active material $(D_A + D_B)$ and the amount of the more active stimulant, say A , is

$$(D_A + D_B)D_A^i = k \quad , \tag{19}$$

where D_A and D_B are in original dosage units, which implies the probability of response to the combination of the two stimulants is determined by the sum of the ingredients multiplied by some power of the amount of the more active stimulant. The second equation, again with A being the more active stimulant, is

$$(1 + k_1 D_A) D_B^i = k_2 \tag{20}$$

which was suggested for the cases where the proportion of A approaches zero. It should be noted that these suggested equations do not bear any clear logical relation one to another.

Plackett and Hewlett (1961) utilize the following biological classification of joint drug actions:

	Similar	Dissimilar
Non-interactive	Simple Similar	Independent
Interactive	Complex Similar	Dependent

Here the suggestion is that the actions of the stimulants are similar or dissimilar respectively as the stimulants act on the same biological site or on different ones, and as interactive or non-interactive depending on the presence or absence of synergism (or antagonism). They, then, propose mathematical equations (some in an implicit form) based on the above biological classifications which, again, do not bear any clear logical relation one to another. Finally, they introduce a statistical concept into their presentation by making an assumption as to the bivariate distribution of $\tilde{z}_1$, $\tilde{z}_2$ where $\tilde{z}_1$, $\tilde{z}_2$ are the respective tolerances to stimulants A and B . They suggested that a reasonable assumption would be that the log tolerances $\log \tilde{z}_1$, $\log \tilde{z}_2$ are distributed bivariate normally. They did not give examples of data fitting any of the proposed models.

Ashford and Smith (1964) approached the problem somewhat differently. They classified the mathematical model as interactive or non-interactive rather than attempting to classify on the basis of biological considerations. They define non-interaction as being equivalent to the condition on $P = P(z_1, z_2, \underline{\theta})$, the probability of response, where z_1 and z_2 are the logarithms of dose, such that

$$W_{12}(P) = P_1P_2(P_2P_{12} - P_1P_{122}) + P_{12}(P_1^2P_{22} - P_2^2P_{11}) = 0 \tag{21}$$

where $P_\alpha = \dfrac{\partial P}{\partial z_\alpha}$, $P_{\alpha\beta} = \dfrac{\partial^2 P}{\partial z_\alpha \partial z_\beta}$, and $P_{\alpha\beta\gamma} = \dfrac{\partial^3 P}{\partial z_\alpha \partial z_\beta \partial z_\gamma}$. Their mathematical classification is not equivalent to Plackett and Hewlett's. Ashford and Smith remarked that no valid distinction can be made between similar and dissimilar action purely on the basis of quantal response data.

Ashford and Smith published some trivariate data on exposure to coal dust for which the response was the prevalence of pneumoconiosis for groups of mine workers. The three dosage variables, respectively, were the time spent in years at coalface coal-getting, coalface preparation, and elsewhere underground. They assumed that the tolerances were normally distributed. They then compared two models where the regions of response were not only different but were each complicated functions of the dosage levels. They applied chi-square goodness-of-fit tests (each with fifteen degrees of freedom) to the models obtaining chi-square values of 12.73 and 16.86, respectively, from which they quote the corresponding approximate significance levels. They do not indicate explicitly the form of the probability function used but rather only the functional forms indicating the response regions.

Zeigler and Moore (1966) presented a paper at the 126th Annual Meeting of the American Statistical Association on "Multivariate Quantal Response Analysis Using Regression Methods." In this paper, in addition to showing that weighted least squares can be used to converge on maximum likelihood estimates, they fitted a bivariate normal distribution to toxicity data involving the direct sprays of Pyrethrins and D.D.T. in Shell Oil P31 applied to flour beetles (Tribolium castaneum). Using a chi-square goodness-of-fit test with nineteen degrees of freedom, they obtained a value of 12.17 and reached the conclusion that the fit was satisfactory.

None of the investigations up to this point have utilized the general form suggested in Chapter I, although the specific form utilized by Zeigler and Moore (1966) is equivalent for the special case where the tolerances $\tilde{z}_1$ and $\tilde{z}_2$ to drugs A and B are each distributed normally.

It would seem useful to do some numerical studies utilizing some of the data in the literature with some analytic models which conform to the general form in equation (16). For this purpose, seven sets of data were utilized. Included among these were sets that have been classified in the following categories by previous investigators: synergistic action, simple similar action, independent action, and additive action.

Data set one, classified as synergistic by Bliss (1939), was first published by Kagy and Richardson (1936). This set is from a study of the combined action of 2-4-dinitro-6-cyclohexylphenol and petroleum oil sprayed in emulsions against eggs of a plant bug (Lygueus kalmii Stål.). Data set two, published by Plackett and Hewlett (1952), was classified by them as simple similar action. This data set is from a study of the combined action of D.D.T. and methoxychlor applied in Shell Oil P31 to flour beetles. Data set three, published by Hewlett and Plackett (1950), was classified by them as independent action. This is the data set which Zeigler and Moore (1966) fitted to a bivariate normal by means of weighted least squares. Data sets four, five, and six, published by Martin (1942), were not classified by the investigator into any category. Data set four is from a study of the toxicity of the combined action of rotenone and a dequelin concentrate in a medium of 0.5% saponin containing 5% of alcohol applied to chrysanthemum aphides (Macrosiphoniella sanborni). Data set five is from a study of the toxicity of the combined action of rotenone and ℓ-elliptone under the same laboratory conditions as data set four. Data set six is from a study of the toxicity of the combined action of rotenone and ℓ-α-toxicarol under like laboratory conditions. These three data sets showed some signs of synergism to the investigator, but he did not find it to be significant in any one of the data sets. Data set seven, published by Ashford and Smith (1964), is from a study of the prevalence of pneumoconiosis in groups of mine workers where the years spent on "coal-getting" is one imput and the other imput is years spent in "haulage." This data set was classified as an example of additive action by the investigators.

A bivariate normit analysis was run on the above seven sets of data. The analytic model for the bivariate normit analysis was

$$P(z_1, z_2, \underline{\Theta}) = \int_{-\infty}^{a_1+B_1z_1} (2\pi)^{-\frac{1}{2}} \exp(-\tfrac{1}{2}t^2)dt + \int_{-\infty}^{a_2+B_2z_2} (2\pi)^{-\frac{1}{2}} \exp(-\tfrac{1}{2}s^2)ds - \int_{-\infty}^{a_1+B_1z_1}\int_{-\infty}^{a_2+B_2z_2} (2\pi\sqrt{1-\rho^2})^{-1} \exp[-(t^2 - 2\rho ts + s^2)/2(1-\rho^2)]dtds, \quad -\infty < z_1, z_2 < \infty \quad . \tag{22}$$

A modified least squares (see Chapter I) FORTRAN IV Computer program was utilized on a Model 44 IBM 360 system. A resumé of the results is given in Table 1.

The following is a brief explanation of the items listed in Table 1 as well as the next two tables: N is the number of stimulant combinations. SSE is the weighted sum of squares due to error which is approximately distributed as a chi-square. SSR is the weighted sum of squares due to regression and is computed as SST - SSE where SST is the weighted sum of squares adjusted for the weighted mean. SSR is approximately distributed as a chi-square. R^2 , which is computed as SSR/SST , tells what portion of SST is due to regression. Computing SSR as SST - SSE and R^2 as SSR/SST gives both a conservative estimate of the significance of regression and a conservative coefficient of determination R^2 . The column entitled "No. of significant chi-squares" tells how many of the chi-square statistics computed at each dosage level (stimulant combination) exceeded 3.84 , the .95 value of a chi-square with one degree of freedom.

Data Set No.	N	No. of Significant Chi-squares	SSE	d.f.	SSR	d.f.	R^2
1	18	5	67.029	13	66041	4	.99899
2	10	1	21.775	5	598.86	4	.96491
3	24	0	11.805	19	11176	4	.99894
4	17	2	27.147	12	1656.0	4	.98387
5	12	2	28.947	7	921.36	4	.96954
6	15	0	10.145	10	30672	4	.99967
7	40	2	38.141	35	217.75	4	.85095

TABLE 1

For all of the data sets the regression is found to be significant using SSR as the indicator. However, the chi-square for departure from the model is insignificant in only three of the cases, namely data sets three, six, and seven, which include the cases of independent action and additive action.

The synergistic data (data set 1) had sample sizes ranging from 240 to 479 (see Appendix I) at its eighteen data points. The bivariate normit analysis indicated that five of these points differed significantly from the bivariate normal model. Some of these points were marginal data points and some were not. One of the data points contributed 34.266 to the cumulative chi-square, slightly more than half of the total, but the chi-square would still be significant even without this particular data point. Upon examination of the residuals, the fit does look good with the exception of the one data point, but with the large sample size at each point, the fit would have to be extremely close in order for the cumulative chi-square to be insignificant. On the whole, it is felt that the bivariate normit analysis did quite well with the data and that the model does describe the phenomenon reasonably well, considering the significance of regression (SSR), the weighted sum of squares due to error (SSE), along with the sample sizes, and the coefficient of determination R^2 .

The simple similar action data (data set 2) had sample sizes ranging from 148 to 200 (see Appendix II) at its ten data points. The analysis indicated that one of these points differed significantly from the bivariate normal model. Again upon examination of the residuals, the fit does look good although not quite as good as the previous data set. The conclusion based on the analysis of the data is that the model does describe the phenomenon fairly well, with the exception of the one data point.

The independent action data (data set 3) had sample sizes ranging from 48 to 50 (see Appendix III) at its twenty four data points. The model does fit the data well and none of the data points differed significantly from the model. The weighted sum of squares due to error is 11.805 . Zeigler and Moore (1966) fitted this same data set and the weighted sum of squares due to error for their model is 12.17 , thus indicating the similarity of the fit.

Data sets four and five are quite similar. They had sample sizes ranging from 28 to 51 (see Appendices IV and V) at their data points. Each had two data points that differed significantly from the bivariate normal model and examination of the residuals does not indicate as good a fit as for any of the previous data sets. The model still does describe most of the data points well, but it does not seem to do as well as for the earlier cases.

Data set six had sample sizes ranging from 48 to 51 (see Appendix VI) at its fifteen data points. The model does fit the data well and none of the data points exhibit a significant deviation from the model. Two bivariate normit analyses were run on this data set using slightly different convergence criteria. The first run utilized the relative change in the unweighted sum of squares due to error and the second the relative change in the weighted sum of squares due to error. The first run after convergence had the sum of squares due to error as 0.031265, while the weighted sum of squares due to error was 0.78159×10^{15} . The second run after convergence had the sum of squares due to error as 0.032040 while the weighted sum of squares due to error was 10.145 . Which criteria produces the best fit becomes questionable at this point. It would seem that either set of parameter estimates would have to be considered acceptable despite the large chi-square value attributed to the first fit.

Data set seven had sample sizes ranging from 2 to 135 (see Appendix VII) at its forty data points. The model does fit the data well although there are two data points which deviate significantly from the model. Ashford and Smith (1964), who classified this data as an example of additive action, fitted the data to a model, assuming the marginals to be logistic, using a rather complicated response region which does not seem to have been necessary.

In general the bivariate normit analysis seems to do quite well with a diversity of mixtures of stimulants, as is evidenced by the seven sets of data analyses here. These analyses appear to lend support to the assumption that the form of the response region should remain constant irrespective of the biological considerations, at least in relation to a bivariate normit.

CHAPTER III

A bivariate logit is, perhaps, as natural to consider in the study of a mixture of two stimulants as a bivariate normit, even though very little work has been done along these lines.

Ashford and Smith (1964) ran an analysis on data set seven assuming the marginals to be logistic. They fitted the data to a model using a complicated response region without explicitly defining the mathematical model. There does not appear to have been any other examinations of data by means of a bivariate logit in the literature

In the case of a bivariate logit, the first consideration is the form of the bivariate distribution to be used. The bivariate logistic distribution utilized in this study was

$$F_3(x,y) = \begin{cases} [1 + \exp(-x)]^{-1}[1 + \exp(-y)]^{-1} \\ \cdot \{1 + a_0[1 + \exp(-x)]^{-1}[1 + \exp(-y)]^{-1} \\ \cdot \quad \exp(-x-y)\} \quad , \quad -\infty < x \, , \, y < \infty \end{cases} \tag{23}$$

which was developed by Gumbel (1961). The density function is

$$f_3(x,y) = \begin{cases} \{\exp(-x-y) \cdot [1 + \exp(-x)]^{-2} \\ \cdot \; [1 + \exp(-y)]^{-2}\} \cdot \{1 + a_0[1 - \exp(-x) \\ - \exp(-y) + \exp(-x-y)]/[1 + \exp(-x) \\ + \exp(-y) + \exp(-x-y)]\} \quad , -\infty < x \, , \, y < \infty \end{cases} \tag{24}$$

The correlation coefficient is

$$\rho = 3a_0/\pi^2 \quad , \tag{25}$$

where $-1 \leq a_0 \leq 1$; thus $|\rho| \leq 3/\pi^2$.

A bivariate logit analysis was run on the seven data sets utilizing the Gumbel bivariate logistic distribution. The analytic model for the bivariate logit analysis was

$$\begin{aligned} P(z_1 , z_2 , \underline{\Theta}) = & \{1 + \exp[B_1(z_1 + a_1)]\}^{-1} + \{1 + \exp[B_2(z_2 + a_2)]\}^{-1} \\ & - \{1 + \exp[B_1(z_1 + a_1)]\}^{-1}\{1 + \exp[B_2(z_2 + a_2)]\}^{-1} \\ & \cdot \Big\{1 + a_0\{1 + \exp[B_1(z_1 + a_1)]\}^{-1}\{1 + \exp[B_2(z_2 + a_2)]\}^{-1} \\ & \cdot \exp[B_1(z_1 + a_1) + B_2(z_2 + a_2)]\Big\} \ . \\ & -\infty < z_1 , z_2 < \infty \quad . \end{aligned} \tag{26}$$

A resumé of the results is given in Table 2. The entries of Table 2 are the same as those of Table 1.

Data Set No.	N	No. of Significant Chi-squares	SSE	d.f.	SSR	d.f.	R^2
1	18	5	76.403	13	48773	4	.99844
2	10	4	29.262	5	574.90	4	.95157
3	24	0	15.603	19	3165.4	4	.99509
4	17	2	29.616	12	1365.2	4	.97877
5	12	3	30.169	7	675.69	4	.95724
6	15	0	11.976	10	20170	4	.99941
7	40	2	39.020	35	221.36	4	.85014

TABLE 2

As was the case for the bivariate normit, the regression was found to be significant for each of the seven data sets, and data sets three, six, and seven have nonsignificant chi-squares indicating no significant departure from regression.

For each data set, the SSE from the bivariate logit analysis was larger than the corresponding SSE from the bivariate normit analysis. Similarly R^2 from the bivariate logit analysis for each data set was smaller than the corresponding R^2 from the bivariate normit analysis. The bivariate logit analysis indicated that the same number of data points differed significantly from the bivariate logit model as was the case with bivariate normit model for each data set with the exception of data set two (simple similar action), and data set five. With data set two, the bivariate logit analysis indicated that four out of the ten data points differed significantly from the bivariate logit model as compared to one out of ten in the bivariate normit analysis. With data set five, the bivariate logit analysis indicated that three out of the twelve data points differed significantly from the bivariate logit model as compared to two out of twelve in the bivariate normit analysis.

On the whole, the bivariate logit analysis did not do as well as the bivariate normit analysis, although it did nearly as well with six out of seven of the data sets. It would seem likely that the main reason that the bivariate logit model did not do as well was due to the fact that the correlation coefficient of the model employed was restricted so that $|\rho| \leq 0.30396$, approximately, and $|\hat{\rho}| > 0.30396$ for all seven data sets in the bivariate normit analysis. It would be useful to extend this investigation to include a bivariate logit model where the correlation coefficient is not so restricted, i.e., where $-1 \leq \rho \leq 1$ inclusive.

CHAPTER IV

In this chapter the assumption that the marginals follow the Burr distribution will be made. This is a somewhat more general assumption than the assumption that the marginals are normal (or logistic).

The general system of distributions referred to here was first given by Burr (1942). Using as an expression for the distribution function

$$F(x) = \begin{cases} 1 - (1 + x^b)^{-p} & x \geq 0 \,;\, b \,,\, p > 0 \\ 0 & x < 0 \end{cases} \tag{27}$$

F(x) covers an important region of the standardized third and fourth central moments in the following sense. Figure 1 shows that the system covers a large portion of the curve-shape characteristics for Types I, III, IV, and VI of the Pearson system. Figure (1) is drawn with coordinates $\alpha_3^2 = \beta_1$ and $\delta = (2\alpha_4 - 3\alpha_3^2 - 6)/(\alpha_4 + 3)$, where α_i is the i^{th} standardized central moment. The regions covered by the Pearson Types I (or beta), IV, and VI are indicated, as well as Type III (or gamma) which lies on a curve, and the normal, logistic,rectangular, and exponential distributions which are represented by points. The subscript B refers to bell shaped functions and J to J shaped functions. It can be seen from Figure (1) that this system of distributions is quite general.

Takahasi (1965) developed a multivariate Burr distribution by using the fact that a Burr distribution is a compound Weibull distribution with a gamma-distribution as a compounder. That is, if

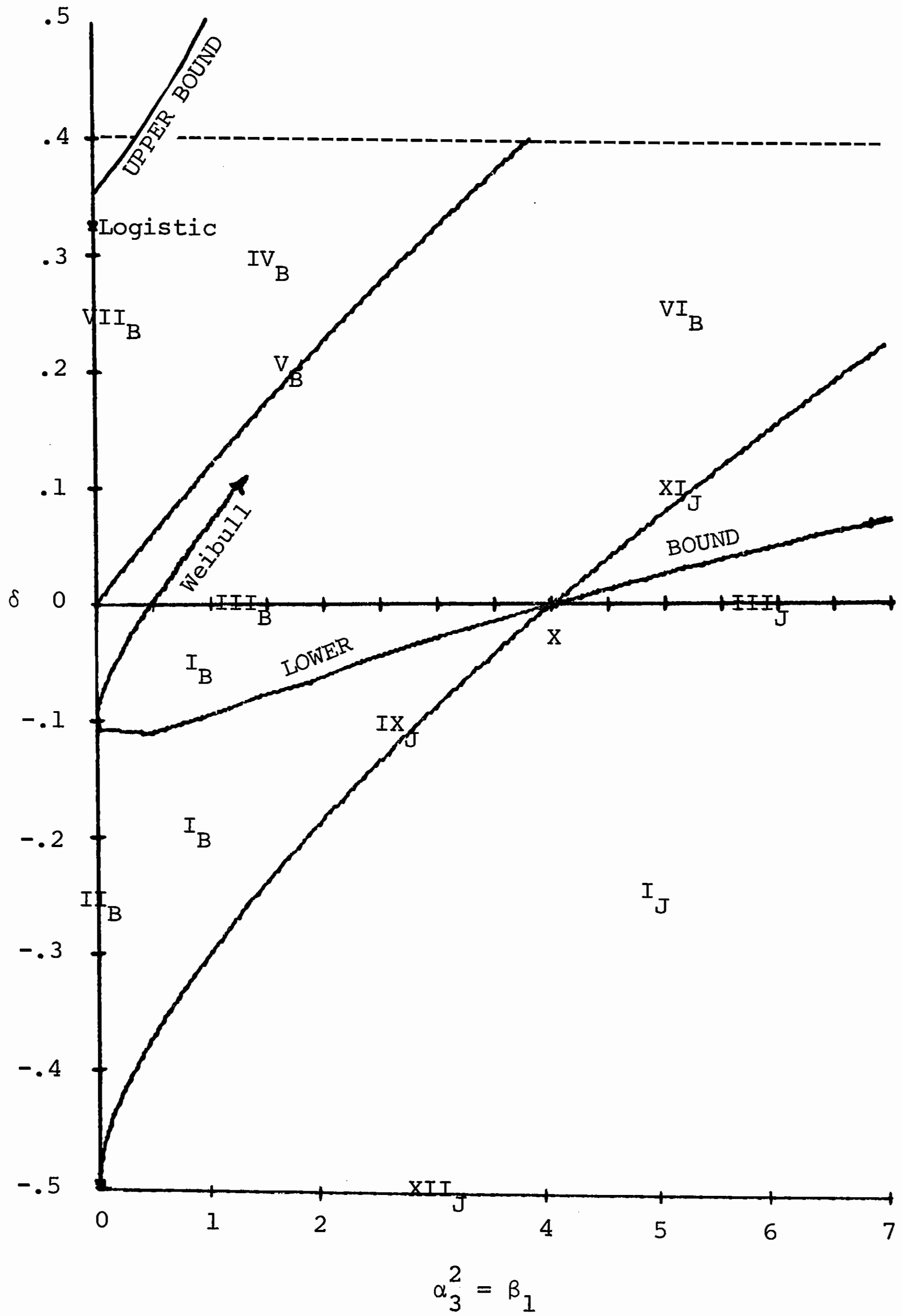


Figure 1. Upper and lower bounds of coverage in β_1 , δ space for the general system of distributions as given by Burr (1968).

$$w(x;b,\theta) = \begin{cases} \theta b x^{b-1} e^{-\theta x^b} & x > 0 \\ 0 & x \leq 0 \end{cases} \tag{28}$$

and θ is a random variable such that

$$g(\theta;p,1) = \begin{cases} \theta^{p-1} e^{-\theta}/\Gamma(p) & \theta > 0 \\ 0 & \theta \leq 0 \end{cases} \tag{29}$$

then the resultant probability density function is Burr. The special case of the bivariate density is

$$f(x_1, x_2) = \frac{\Gamma(p+2)}{\Gamma(p)} b_1 b_2 r_1 r_2 x_1^{b_1-1} x_2^{b_2-1} (1 + r_1 x_1^{b_1} + r_2 x_2^{b_2})^{-(p+2)} \quad x_i > 0 \ (i=1,2)$$

$$= 0 \qquad \text{elsewhere.} \tag{30}$$

The bivariate distribution is

$$F(x_1, x_2) = 1 - (1 + r_1 x_1^{b_1})^{-p} - (1 + r_2 x_2^{b_2})^{-p} + (1 + r_1 x^{b_1} + r_2 y^{b_2})^{-p} \quad x_i \leq 0$$

$$= 0 \qquad \text{elsewhere.} \tag{31}$$

It should be noted at this point that the r_i are equal to one in the Burr distribution as given by Burr (1942). If x_i is set equal to $B_i(z_i + a_i)$, it is easily seen that the r_i's are redundant. In addition, if the b_i's and p are held constant, e.g., the third and fourth standardized central moments can be set equal to those of the normal distribution by proper choice of the b_i's and p , then the correlation coefficient is a constant.

It was attempted to find a form of a bivariate Burr distribution such that the correlation coefficient would not be a fixed constant. The form developed by the author is

$$F(x_1, x_2) = 1 - (1 + x_1^{b_1})^{-p} - (1 + x_2^{b_2})^{-p} + (1 + x_1^{b_1} + x_2^{b_2} + rx_1^{b_1}x_2^{b_2})^{-p} .$$

$$x_i \geq 0$$

$$0 \leq r \leq p + 1$$

$$= 0 \qquad \text{elsewhere .} \tag{32}$$

The bivariate density is

$$f(x_1, x_2) = p(p + 1)(1 + rx_2^{b_2})(1 + rx_1^{b_1})b_1 b_2 \; x_1^{b_1-1} x_2^{b_2-1} (1 + x_1^{b_1} + x_2^{b_2} + rx_1^{b_1}x_2^{b_2})^{-(p+2)} - prb_1 b_2 x_1^{b_1-1} x_2^{b_2-1} (1 + x_1^{b_1} + x_2^{b_2} + rx_1^{b_1}x_2^{b_2})^{-(p+1)} .$$

$$x_i \geq 0$$

$$0 \leq r \leq p + 1$$

$$= 0 \qquad \text{elsewhere .} \tag{33}$$

The marginals are of the form given by Burr (1942). The conditional distribution of x_i given x_j $i \neq j$ is

$$F(x_i | x_j) = 1 - r\left(\frac{1 + x_j^{b_j}}{1 + rx_j^{b_j}}\right)\left[1 + \left(\frac{1 + rx_j^{b_j}}{1 + x_j^{b_j}}\right) x_i^{b_i}\right]^{-p} + \left(\frac{r - 1}{1 + rx_j^{b_j}}\right)\left[1 + \left(\frac{1 + rx_j^{b_j}}{1 + x_j^{b_j}}\right) x_i^{b_i}\right]^{-(p+1)} . \quad x_i \geq 0$$

$$= 0 \qquad \text{elsewhere .} \tag{34}$$

The conditional density of x_i given x_j is

$$f(x_i | x_j) = (p+1)(1+rx_i^{b_i})\left(\frac{1 + rx_j^{b_j}}{1 + x_j^{b_j}}\right) b_i x_i^{b_i-1}\left[1 + \left(\frac{1 + rx_j^{b_j}}{1 + x_j^{b_j}}\right)x_i^{b_i}\right]^{-(p+2)} - rb_i x_i^{b_i-1}\left[1 + \left(\frac{1 + rx_j^{b_j}}{1 + x_j^{b_j}}\right)x_i^{b_i}\right]^{-(p+1)} . \qquad x_i \geq 0$$

$$= 0 \qquad \text{elsewhere.} \tag{35}$$

The correlation coefficient is

$$\rho_{x_1x_2} = \Gamma\left(p - \frac{1}{b_1}\right)\Gamma\left(p - \frac{1}{b_2}\right)\Gamma\left(1 + \frac{1}{b_1}\right)\Gamma\left(1 + \frac{1}{b_2}\right) \cdot$$

$$\left\{\left(1 - \frac{1}{pb_1}\right) {}_2F_1\left(p + 1 - \frac{1}{b_1} , 1 + \frac{1}{b_2} ; p + 1 ; 1 - r\right)\right.$$

$$\left. + \frac{r}{pb_1} {}_2F_1\left(p - \frac{1}{b_1} , 1 + \frac{1}{b_2} ; p + 1 ; 1 - r\right)\right\} \Bigg/$$

$$\left\{\left[\Gamma\left(1 + \frac{2}{b_1}\right)\Gamma\left(p - \frac{2}{b_1}\right)\Gamma(p) - \Gamma^2\left(1 + \frac{1}{b_1}\right)\Gamma^2\left(p - \frac{1}{b_1}\right)\right] \cdot\right.$$

$$\left.\left[\Gamma\left(1 + \frac{2}{b_2}\right)\Gamma\left(p - \frac{2}{b_2}\right)\Gamma(p) - \Gamma^2\left(1 + \frac{1}{b_2}\right)\Gamma^2\left(p - \frac{1}{b_2}\right)\right]\right\}^{\frac{1}{2}} , \qquad (36)$$

where ${}_2F_1(\alpha,\beta;\gamma,z)$ is Gauss' hypergeometric function. If $r = 1$, then $\rho_{x_1x_2} = 0$.

A bivariate Burrit analysis was run on the seven sets of data using the bivariate Burr distribution described above. The analytic model for the bivariate Burrit analysis was

$$P(z_1 , z_2 , \underline{\Theta}) = 1 - \left\{ 1 + [B_1(z_1 + a_1)]^{b_1} + [B_2(z_2 + a_2)]^{b_2}\right.$$

$$\left. + r[B_1(z_1 + a_1)]^{b_1}[B_2(z_2 + a_2)]^{b_2}\right\}^{-p} . \quad \begin{array}{l} -a_1 \leq z_1 < \infty \\ -a_2 \leq z_2 < \infty \end{array}$$

$$= 0 \qquad \text{elsewhere} . \qquad (37)$$

A resume of the results is given in Table 3. The rows corresponding to eight parameters to be estimated constitute the general case of the Burrit

Data Set No.	No. of Parameters To be Estimated	N	No. of Significant Chi-squares	SSE	d.f.	SSR	d.f.	R^2
1	8	18	2	41.646	10	67964	7	.99939
	7		5	64.694	11	45745	6	.99859
	5		5	70.017	13	70566	4	.99901
	4		5	92.619	14	34582	3	.99733
2	8	10	3	23.027	2	573.70	7	.96146
	7		4	38.279	3	544.20	6	.93380
	5		3	26.296	5	577.20	4	.95643
	4		5	46.538	6	559.09	3	.92316
3	8	24	0	13.227	16	4262.0	7	.99690
	7		2	29.737	17	1178.8	6	.97539
	5		2	32.579	19	8668.4	4	.99696
	4		10	114.17	20	1309.0	3	.91978
4	8	17	2	27.125	9	1655.7	7	.98388
	7		3	32.782	10	1683.0	6	.98089
	5		2	27.130	12	1673.8	4	.98450
	4		3	34.593	13	1483.8	3	.97722
5	8	12	2	29.221	4	1081.9	7	.97370
	7		2	31.901	5	1508.3	6	.97929
	5		2	29.379	7	862.11	4	.96705
	4		3	36.603	8	905.50	3	.96115
6	8	15	1	9.660	7	67305	7	.99986
	7		2	13.648	8	13398	6	.99898
	5		0	12.047	10	23681	4	.99949
	4		1	23.480	11	6862.7	3	.99659
7	8	40	2	38.293	32	217.30	7	.85018
	7		3	38.765	33	219.84	6	.85010
	5		2	38.338	35	218.53	4	.85075
	4		2	38.766	36	219.06	3	.84964

TABLE 3

analysis; the rows corresponding to seven parameters to be estimated correspond to the special case with r = 0 which reduces to a Burrit analysis using the bivariate Burr distribution developed by Takahasi (1965). The rows corresponding to five and four parameters to be estimated have $\alpha_3 = 0$, $\alpha_4 = 3$ (the third and fourth standardized central moments), which are the same as the normal distributions' α_3 and α_4 . The first of these is a special case of the general Burrit analysis and the second, a special case of the Burrit analysis using the Takahasi bivariate Burr distribution.

As was the case for both the bivariate normit and the bivariate logit, the regression was found to be significant for each of the seven data sets for all of the bivariate Burrit analyses (four on each data set). The chi-square test was insignificant, indicating no significant departure from regression for data set three with the general Burrit analysis, for data set six for all but the analysis with four parameters to be estimated, and for data set seven for all four of the analyses.

The SSE from the general case of the bivariate Burrit analysis was significantly smaller than that from the bivariate normit analysis only with the synergistic data (data set one). In no other case is there any indication that the bivariate Burrit model is better than the bivariate normit model in the actual fitting of these data to a model.

Each SSE from the bivariate Burrit analyses utilizing the bivariate Burr developed in this paper is significantly smaller than the corresponding SSE from the analyses utilizing the Takahasi bivariate Burr distribution in all but three cases: both cases with data set seven, and the first case with data set five (the case corresponding to the two analyses with eight and seven parameters to be estimated). On the basis of these analyses it would seem that the bivariate Burr developed in this paper would be, in general, more useful in application than the form developed by Takahasi.

In addition it may be noted that the marginal distributions for the synergistic data, as characterized by α_3 and α_4 , do not lie in the same Pearson curve area (see figure 2). The marginals for data set three also display this characteristic but not to as high a degree. The marginals for data sets four through seven are all clustered around the normal distribution. The fact that the assumption that the marginals are Burr distribution does allow given marginal to have curve shape characteristics different from that of the other marginal suggests that the bivariate Burrit analysis may be well adapted for the analysis of data where the marginal distributions do not belong to the same family, e.g. the family of normal distributions.

In summary, the bivariate analyses utilizing the general form indicated by equation (16) seem to do quite well with a diversity of mixtures of stimulants as is evidenced by the seven sets of data which have been analyzed in this paper. The bivariate normit model and the bivariate Burrit model (general case, i.e., the case with eight parameters to be estimated) seem to be best suited for these types of analyses. The bivariate normit model would have to be recommended if the number of parameters to be estimated is of concern, but otherwise the bivariate Burrit model could well be the best model for these types of analyses.

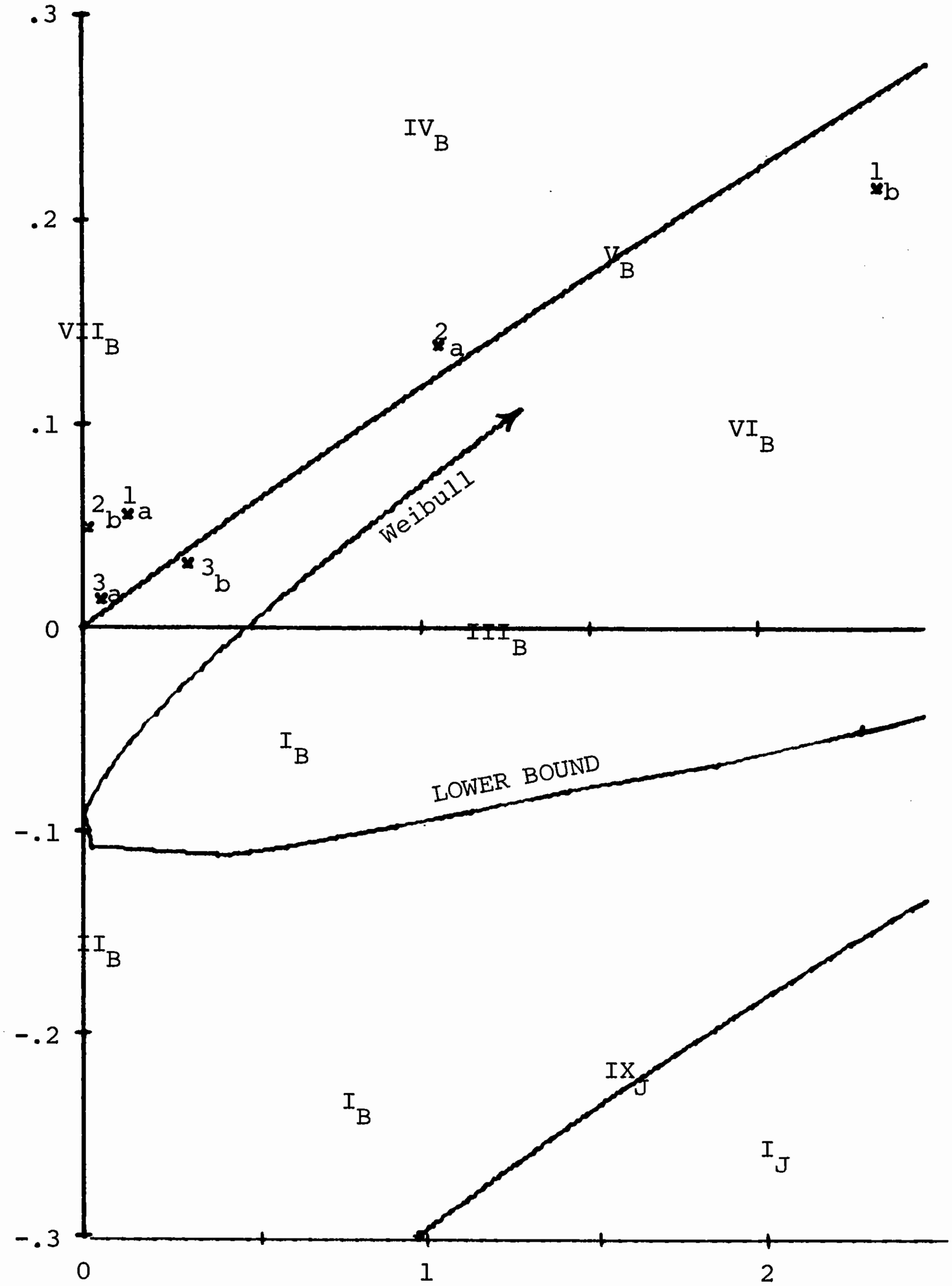


Figure 2. Expanded portion of the coverage β_1 , δ space. The x's mark six of the sample population points (β_1 , δ), from the data sets analyzed in this paper. N_i (N=1, 2, 3; i=a, b) refers to the i^{th} marginal of the N^{th} data set.

APPENDIX I

Data of Kagy and Richardson (1936): The combined action of 2-4 dinitro-6-cyclohexylphenol and petroleum oil sprayed in emulsions against eggs of a plant bug (Lygaeus Kalmii Stål). The data as described by Kagy and Richardson, the translated data, and the analyses on this set of data (data set one) are in this appendix.

DATA AS DESCRIBED IN TEXT

CONCENTRATION OF Phenol in Oil Mixture %	CONCENTRATION OF Mixture in Spray %	N_i Number of Eggs	Net Kill %
0	1	240	6.5
0	2	479	40.1
0	3	479	58.7
0.1	1	240	9.9
0.1	2	479	59.7
0.1	3	479	72.3
0.5	1	288	30.1
0.5	2	479	73.7
0.5	3	479	90.4
1.0	1	288	58.6
1.0	2	384	94.0
1.0	3	288	97.22
2.0	1	288	81.2
2.0	2	384	97.13
2.0	3	288	99.65
3.0	1	288	86.8
3.0	2	384	99.48
5.0	1	240	96.66

TRANSLATED DATA

X(1)	X(2)	p_i
0	.01	.0667
0	.02	.4008
0	.03	.5866
.00001	.00999	.1000
.00002	.01998	.5971
.00003	.02997	.7223
.00005	.00995	.3021
.0001	.0199	.7370
.00015	.02985	.9040
.0001	.0099	.5868
.0002	.0198	.9401
.0003	.0297	.9722
.0002	.0098	.8125
.0004	.0196	.9714
.0006	.0294	.9965
.0003	.0097	.8681
.0006	.0194	.9948
.0005	.0095	.9667

Here (N_i(i^{th} Net kill %)/100) was rounded off to the nearest integer -- which should be r_i, the number that responded to the i^{th} mixture of stimulants, and then p_i was computed as r_i/N_i.

BIVARIATE NORMIT ANALYSIS

Parameter Estimates

$\hat{a}_1 = 8.640$

$\hat{B}_1 = 0.935$

$\hat{a}_2 = 5.583$

$\hat{B}_2 = 1.489$

$\hat{\rho} = -0.379$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	66041
Departure from Model	SSE	13	67.029
TOTAL	SST	17	66107

Coefficient of Determination $R^2 = .99899$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0667	.1016	-.0349	3.201
.4008	.4049	-.0041	0.033
.5866	.6417	-.0551	6.317
.1000	.1180	-.0180	0.750
.5971	.4637	.1334	34.266
.7223	.7231	-.0008	0.002
.3021	.3596	-.0575	4.139
.7370	.7675	-.0305	2.490
.9040	.9277	-.0237	4.024
.5868	.5859	-.0009	0.001
.9401	.8986	.0415	7.266
.9722	.9768	-.0046	0.274
.8125	.7986	-.0139	0.346
.9714	.9691	.0023	0.069
.9965	.9950	.0015	0.137
.8681	.8865	-.0184	0.973
.9948	.9870	.0078	1.809
.9667	.9536	.0131	0.931

BIVARIATE LOGIT ANALYSIS

Parameter Estimates

$\hat{a}_0 = -0.938$

$\hat{a}_1 = 9.220$

$\hat{B}_1 = -1.690$

$\hat{a}_2 = 3.762$

$\hat{B}_2 = -2.413$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	48773
Departure from Model	SSE	13	76.402
TOTAL	SST	17	48849

Coefficient of Determination $R^2 = .99844$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0667	.1155	-.0488	5.597
.4008	.4103	-.0095	0.177
.5866	.6492	-.0626	8.246
.1000	.1352	-.0352	2.540
.5971	.4601	.1370	36.200
.7223	.7120	.0103	0.247
.3021	.3438	-.0417	2.220
.7370	.7627	-.0257	1.749
.9040	.9303	-.0263	5.117
.5868	.5837	.0031	0.011
.9401	.9013	.0388	6.503
.9722	.9774	-.0052	0.348
.8125	.8087	.0038	0.027
.9714	.9658	.0056	0.365
.9965	.9929	.0036	0.520
.8681	.8917	-.0236	1.654
.9948	.9821	.0127	3.527
.9667	.9504	.0163	1.354

BIVARIATE BURRIT ANALYSES

1: General Case - Eight Parameters to be Estimated

Parameter Estimates

$\hat{r} = 3.556$

$\hat{b}_1 = 9.643$

$\hat{b}_2 = 1.773$

$\hat{p} = 4.813$

$\hat{a}_1 = 18.073$

$\hat{B}_1 = 0.094$

$\hat{a}_2 = 4.877$

$\hat{B}_2 = 0.318$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	7	67964
Departure from Model	SSE	10	41.646
TOTAL	SST	17	68006

Coefficient of Determination $R^2 = .99939$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0667	.0602	.0065	0.181
.4008	.4273	-.0265	1.377
.5866	.6289	-.0423	3.676
.1000	.1018	-.0018	0.008
.5971	.5054	.0917	16.107
.7223	.7202	.0021	0.010
.3021	.3471	-.0450	2.575
.7370	.7743	-.0373	3.821
.9040	.9211	-.0171	1.935
.5868	.5627	.0241	0.682
.9401	.9010	.0390	6.580
.9722	.9755	-.0033	0.130
.8125	.7823	.0302	1.540
.9714	.9711	.0003	0.014
.9965	.9950	.0015	0.137
.8681	.8781	-.0100	0.271
.9948	.9883	.0066	1.420
.9667	.9516	.0151	1.194

2: Takahasi Burr - r = 0 ; Seven Parameters to be Estimated

Parameter Estimates

$\hat{b}_1 = 8.008$

$\hat{b}_2 = 1.799$

$\hat{p} = 6.239$

$\hat{a}_1 = 16.156$

$\hat{B}_1 = 0.113$

$\hat{a}_2 = 4.869$

$\hat{B}_2 = 0.291$

Chi-square Analysis Table

source		d.f.	
Due to model	SSR	6	45745
Departure from Model	SSE	11	64.694
TOTAL	SST	17	45810

Coefficient of Determination $R^2 = .99859$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0667	.0592	.0075	0.245
.4008	.4480	-.0472	4.318
.5866	.6599	-.0733	11.463
.1000	.0921	.0079	0.181
.5971	.5001	.0970	18.015
.7223	.7115	.0108	0.272
.3021	.3552	-.0531	3.544
.7370	.7466	-.0096	0.231
.9040	.8874	.0166	1.319
.5868	.5958	-.0090	0.097
.9401	.8844	.0557	0.001
.9722	.9581	.0141	1.420
.8125	.8261	-.0136	0.370
.9714	.9670	.0044	0.230
.9965	.9904	.0061	1.111
.8681	.9151	-.0469	8.167
.9948	.9875	.0073	1.656
.9667	.9733	-.0066	0.403

3: $\alpha_3 = 0$; $\alpha_4 = 3$ (Third and Fourth Standardized Central Moments) ;

$b_1 = b_2 = 4.874$; $p = 6.158$; Five Parameters to be Estimated

Parameter Estimates

$\hat{r} = 4.383$

$\hat{a}_1 = 13.485$

$\hat{B}_1 = 0.153$

$\hat{a}_2 = 6.426$

$\hat{B}_2 = 0.242$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	70566
Departure from Model	SSE	13	70.017
TOTAL	SST	17	70636

Coefficient of Determination $R^2 = .99901$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0667	.1057	-.0390	3.861
.4008	.4058	-.0050	0.049
.5866	.6443	-.0577	6.967
.1000	.1221	-.0221	1.090
.5971	.4603	.1368	36.090
.7223	.7168	.0056	0.073
.3021	.3594	-.0573	4.107
.7370	.7592	-.0222	1.296
.9040	.9271	-.0231	3.794
.5868	.5829	.0039	1.873
.9401	.8964	.0437	7.906
.9722	.9782	-.0060	0.483
.8125	.7985	.0140	0.353
.9714	.9691	.0023	0.068
.9965	.9954	.0011	0.073
.8681	.8876	-.0195	1.100
.9948	.9869	.0079	1.847
.9667	.9543	.0124	0.843

4: Takahasi Burr - $r = 0$; $\alpha_3 = 0$; $\alpha_4 = 3$ (Third and Fourth Standardized Central Moments); $b_1 = b_2 = 4.874$; $p = 6.158$; Four Parameters to be Estimated

Parameter Estimates

$\hat{a}_1 = 13.460$
$\hat{B}_1 = 0.160$
$\hat{a}_2 = 6.265$
$\hat{B}_2 = 0.259$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	3	34582
Departure from Model	SSE	14	92.619
TOTAL	SST	17	34674

Coefficient of Determination $R^2 = .99733$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0667	.0957	-.0290	2.333
.4008	.4125	-.0117	0.270
.5866	.6674	-.0808	14.102
.1000	.1140	-.0140	0.463
.5971	.4596	.1375	36.464
.7223	.7144	.0079	0.147
.3021	.3799	-.0778	7.395
.7370	.7378	-.0008	14.954
.9040	.8911	.0129	0.822
.5868	.6188	-.0319	1.246
.9401	.8789	.0612	13.499
.9722	.9564	.0158	1.723
.8125	.8321	-.0196	0.796
.9714	.9611	.0103	1.086
.9965	.9877	.0088	1.844
.8681	.9128	-.0446	7.208
.9948	.9830	.0118	3.208
.9667	.9680	.0013	0.013

APPENDIX II

Data of Plackett and Hewlett (1952): The toxity to <u>Tribolium castaneum</u> of D.D.T., methoxychlor (MOC), and combinations of the two applied in Shell Oil P31 as films on filter paper, six-day exposures. The data as described by Plackett and Hewlett, the translated data, and the analyses on this set of data (data set two) are in this appendix.

DATA AS DESCRIBED BY PLACKETT AND HEWLETT

D.D.T. Percent w/v	MOC Percent w/v	N_i Number of Beetles	Observed Mortality Percent
0.0	0.4	199	7.5
0.0	0.8	148	29.7
0.0	1.6	199	77.9
0.2	0.0	200	14.5
0.2	0.4	150	26.0
0.2	0.8	151	63.6
0.4	0	149	43.6
0.4	0.4	148	66.2
0.4	0.8	150	78.7
0.8	0.0	199	70.9

TRANSLATED DATA

X(1)	X(2)	p_i
0.0	0.004	.0754
0.0	0.008	.2973
0.0	0.016	.7789
0.002	0.0	.1450
0.002	0.004	.2600
0.002	0.008	.6358
0.004	0.0	.4362
0.004	0.004	.6622
0.004	0.008	.7867
0.008	0.0	.7085

BIVARIATE NORMIT ANALYSIS

Parameter Estimates

$\hat{a}_1 = 5.787$

$\hat{B}_1 = 1.071$

$\hat{a}_2 = 6.925$

$\hat{B}_2 = 1.503$

$\hat{\rho} = -0.9999$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	598.86
Departure from Model	SSE	5	21.775
TOTAL	SST	9	620.63

Coefficient of Determination $R^2 = .96491$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0754	.0844	-.0090	0.210
.2973	.3693	-.0720	3.292
.7789	.7606	.0183	0.365
.1450	.1920	.0470	2.848
.2600	.2764	-.0164	0.203
.6358	.5613	.0745	3.405
.4362	.4491	-.0129	0.100
.6622	.5335	.1287	9.845
.7867	.8184	-.0317	1.013
.7085	.7306	-.0221	0.494

BIVARIATE LOGIT ANALYSIS

Parameter Estimates

$\hat{a}_0 = -1.000$

$\hat{a}_1 = 5.448$

$\hat{B}_1 = -1.776$

$\hat{a}_2 = 4.645$

$\hat{B}_2 = -2.559$

Chi-Square Analysis Table

source		d.f.	
Due to Model	SSR	4	574.90
Departure from Model	SSE	5	29.262
TOTAL	SST	9	604.16

Coefficient of Determination $R^2 = .95157$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0754	.0958	-.0204	0.960
.2973	.3846	-.0873	4.760
.7789	.7865	-.0076	0.068
.1450	.2041	-.0591	4.297
.2600	.2944	-.0344	0.856
.6358	.5486	.0872	4.638
.4362	.4675	-.0313	0.586
.6622	.5401	.1221	8.811
.7867	.7312	.0555	2.351
.7085	.7504	-.0419	1.864

BIVARIATE BURRIT ANALYSES

1: General Case - Eight Parameters to be Estimated

Parameter Estimates

$\hat{r} = 5.271$

$\hat{b}_1 = 2.351$

$\hat{b}_2 = 6.755$

$\hat{p} = 4.271$

$\hat{a}_1 = 7.252$

$\hat{B}_1 = 0.270$

$\hat{a}_2 = 8.206$

$\hat{B}_2 = 0.216$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	7	573.70
Departure from Model	SSE	2	23.027
TOTAL	SST	9	596.73

Coefficient of Determination $R^2 = .96141$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0754	.1009	-.0255	1.430
.2973	.3814	-.0841	4.440
.7789	.7764	.0025	0.007
.1450	.1889	-.0439	2.513
.2600	.2862	-.0262	0.504
.6358	.5423	.0935	5.321
.4362	.4835	-.0473	1.332
.6622	.5643	.0979	5.768
.7867	.7559	.0308	0.770
.7085	.7387	-.0302	0.943

2: Takahasi Burr - r = 0 ; Seven Parameters to be Estimated

Parameter Estimates

$\hat{b}_1 = 0.961$
$\hat{b}_2 = 5.064$
$\hat{p} = 3.265$
$\hat{a}_1 = 6.399$
$\hat{B}_1 = 0.298$
$\hat{a}_2 = 7.263$
$\hat{B}_2 = 0.290$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	6	544.20
Departure from Model	SSE	3	38.249
TOTAL	SST	9	582.45

Coefficient of Determination $R^2 = .93433$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0754	.0987	-.0233	1.217
.2973	.4096	-.1123	7.723
.7789	.7928	-.0139	0.234
.1450	.1774	-.0324	1.440
.2600	.2542	.0058	0.026
.6358	.5004	.1354	11.071
.4362	.5490	-.1128	7.662
.6622	.5844	.0778	3.685
.7867	.7037	.0830	4.953
.7085	.7240	-.0155	0.239

3: $\alpha_3 = 0$; $\alpha_4 = 3$ (Third and Fourth Standardized Central Moments) ;

$b_1 = b_2 = 4.874$; $p = 6.158$; Five Parameters to be Estimated

Parameter Estimates

$\hat{r} = 7.158$

$\hat{a}_1 = 9.108$

$\hat{B}_1 = 0.176$

$\hat{a}_2 = 7.272$

$\hat{B}_2 = 0.245$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	577.20
Departure from Model	SSE	5	26.296
TOTAL	SST	9	603.50

Coefficient of Determination $R^2 = .95643$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0754	.0946	-.0192	0.854
.2973	.3869	-.0896	5.005
.7789	.7805	-.0016	0.003
.1450	.2018	-.0568	4.002
.2600	.2928	-.0328	0.780
.6358	.5586	.0772	3.653
.4362	.4628	-.0266	0.424
.6622	.5410	.1212	8.757
.7867	.7490	.0377	1.133
.7085	.7484	-.0399	1.685

4: Takahasi Burr - $r = 0$; $\alpha_3 = 0$; $\alpha_4 = 3$ (Third and Fourth Standardized Central Moments); $b_1 = b_2 = 4.874$; $p = 6.158$; Four Parameters to be Estimated

Parameter Estimates

$\hat{a}_1 = 9.196$

$\hat{B}_1 = 0.174$

$\hat{a}_2 = 7.270$

$\hat{B}_2 = 0.248$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	3	559.09
Departure from Model	SSE	6	46.537
TOTAL	SST	9	605.63

Coefficient of Determination $R^2 = .92316$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0754	.0990	-.0236	1.239
.2973	.4013	-.1040	6.666
.7789	.7952	-.0163	0.324
.1450	.2165	-.0715	6.027
.2600	.2912	-.0312	0.707
.6358	.5219	.1139	7.855
.4362	.4798	-.0436	1.135
.6622	.5264	.1358	10.953
.7867	.6727	.1140	8.857
.7085	.7590	-.0505	2.775

APPENDIX III

Data of Hewlett and Plackett (1950): A study of six day toxicity to beetles (*Tribolium castaneum*) of direct sprays of Pyrethins, D.D.T., and the two together in Shell Oil P31. The data as reproduced by Zeigler and Moore (1966), and the analyses on this set of data (data set three) are in this appendix.

DATA AS REPRODUCED BY ZEIGLER AND MOORE

Insecticide	DEPOSIT (mg./10 sq. cm.)	X(1)	X(2)	N_i	p_i
1.2% w/v Pyrethins	2.52	.03024	0	48	.0625
	3.30	.03960	0	48	.0625
	4.25	.05100	0	50	.1800
	5.33	.06396	0	50	.3200
	7.15	.08580	0	50	.4000
	9.53	.11436	0	50	.6000
	12.28	.14739	0	49	.7551
	15.58	.18696	0	50	.7000
2.0% w/v D.D.T.	2.45	0	.0490	49	.1633
	3.18	0	.0636	50	.1600
	4.25	0	.0850	50	.3200
	5.48	0	.1096	50	.4200
	7.24	0	.1448	50	.5000
	9.54	0	.1908	50	.5600
	12.36	0	.2472	50	.7000
	15.54	0	.3108	50	.7400
1.2% w/v Pyrethins plus 2.0% w/v D.D.T.	2.74	.02964	.0494	50	.2800
	3.20	.03840	.0640	49	.3673
	4.10	.04920	.0820	50	.4400
	5.34	.06408	.1068	50	.7200
	7.11	.08532	.1422	50	.8400
	9.60	.11520	.1920	50	.9000
	12.45	.14940	.2490	50	1.0000
	15.65	.18780	.3130	50	1.0000

BIVARIATE NORMIT ANALYSIS

Parameter Estimates

$\hat{a}_1 = 2.827$

$\hat{B}_1 = 1.231$

$\hat{a}_2 = 1.698$

$\hat{B}_2 = 0.882$

$\hat{\rho} = -0.686$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	11176
Departure from Model	SSE	19	11.805
TOTAL	SST	23	11188

Coefficient of Determination $R^2 = .99894$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0625	.0696	-.0071	0.037
.0625	.1257	-.0632	1.744
.1800	.2017	-.0217	0.146
.3200	.2888	.0312	0.237
.4000	.4225	-.0225	0.104
.6000	.5629	.0371	0.280
.7551	.6809	.0742	1.241
.7000	.7773	-.0773	1.727
.1633	.1677	-.0044	0.007
.1600	.2318	-.0718	1.447
.3200	.3166	.0034	0.003
.4200	.4002	.0198	0.082
.5000	.4972	.0028	0.002
.5600	.5934	-.0334	0.231
.7000	.6790	.0210	0.101
.7400	.7476	-.0076	0.015
.2800	.2358	.0442	0.542
.3673	.3506	.0167	0.060
.4400	.4896	-.0496	0.492
.7200	.6559	.0642	0.912
.8400	.8215	.0185	0.117
.9000	.9364	-.0364	1.109
1.0000	.9816	.0184	0.939
1.0000	.9954	.0046	0.231

BIVARIATE LOGIT ANALYSIS

Parameter Estimates

$\hat{a}_0 = -1.000$

$\hat{a}_1 = 2.338$

$\hat{B}_1 = -2.090$

$\hat{a}_2 = 1.960$

$\hat{B}_2 = -1.520$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	3165.4
Departure from Model	SSE	19	15.603
TOTAL	SST	23	3181.0

Coefficient of Determination $R^2 = .99509$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0625	.0813	-.0188	0.227
.0625	.1346	-.0720	2.140
.1800	.2087	-.0287	0.250
.3200	.2975	.0225	0.121
.4000	.4390	-.0390	0.309
.6000	.5879	.0121	0.030
.7551	.7079	.0472	0.528
.7000	.7994	-.0994	3.082
.1633	.1674	-.0041	0.006
.1600	.2301	-.0701	1.385
.3200	.3171	.0029	0.002
.4200	.4059	.0141	0.041
.5000	.5105	-.0105	0.022
.5600	.6133	-.0533	0.600
.7000	.7016	-.0016	0.001
.7400	.7690	-.0290	0.237
.2800	.2442	.0358	0.347
.3673	.3493	.0180	0.070
.4400	.4754	-.0354	0.252
.7200	.6266	.0934	1.865
.8400	.7816	.0584	0.999
.9000	.9002	-.0002	0.000
1.0000	.9578	.0422	2.201
1.0000	.9826	.0174	0.888

BIVARIATE BURRIT ANALYSES

1: General Case - Eight Parameters to be Estimated

Parameter Estimates

$\hat{r} = 6.323$

$\hat{b}_1 = 3.933$

$\hat{b}_2 = 2.941$

$\hat{p} = 5.323$

$\hat{a}_1 = 4.863$

$\hat{B}_1 = 0.239$

$\hat{a}_2 = 4.760$

$\hat{B}_2 = 0.183$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	7	4258.4
Departure from Model	SSE	16	13.227
TOTAL	SST	23	4271.6

Coefficient of Determination $R^2 = .99690$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0625	.0622	.0003	0.000
.0625	.1216	-.0591	1.569
.1800	.2025	-.0225	0.156
.3200	.2947	.0253	0.154
.4000	.4346	-.0346	0.244
.6000	.5784	.0216	0.095
.7551	.6963	.0588	0.802
.7000	.7898	-.0898	2.429
.1633	.1666	-.0033	0.004
.1600	.2384	-.0784	1.691
.3200	.3298	-.0098	0.022
.4200	.4161	.0039	0.003
.5000	.5120	-.0120	0.029
.5600	.6034	-.0434	0.393
.7000	.6820	.0180	0.075
.7400	.7438	-.0038	0.037
.2800	.2261	.0540	0.832
.3673	.3475	.0198	0.085
.4400	.4875	-.0475	0.451
.7200	.6463	.0737	1.187
.8400	.8007	.0393	0.484
.9000	.9135	-.0135	0.115
1.0000	.9660	.0340	1.759
1.0000	.9873	.0127	0.644

2: Takahasi Burr - $r = 0$; Seven Parameters to be Estimated

Parameter Estimates

$\hat{b}_1 = 3.831$

$\hat{b}_2 = 2.939$

$\hat{p} = 6.740$

$\hat{a}_1 = 4.675$

$\hat{B}_1 = 0.243$

$\hat{a}_2 = 4.701$

$\hat{B}_2 = 0.176$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	6	1178.8
Departure from Model	SSE	17	29.737
TOTAL	SST	23	1208.5

Coefficient of Determination $R^2 = .97539$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0625	.0540	.0085	0.067
.0625	.1148	-.0523	1.292
.1800	.2011	-.0211	0.138
.3200	.3017	.0184	0.080
.4000	.4552	-.0552	0.614
.6000	.6110	-.0109	0.025
.7551	.7346	.0205	0.106
.7000	.8282	-.1282	5.772
.1633	.1707	-.0074	0.019
.1600	.2469	-.0869	2.032
.3200	.3444	-.0244	0.132
.4200	.4361	-.0161	0.053
.5000	.5373	-.0373	0.280
.5600	.6326	-.0725	1.132
.7000	.7132	-.0132	0.042
.7400	.7752	-.0352	0.355
.2800	.2137	.0663	1.310
.3673	.3256	.0417	0.387
.4400	.4503	-.0103	0.022
.7200	.5888	.1312	3.554
.8400	.7267	.1134	3.234
.9000	.8406	.0594	1.318
1.0000	.9088	.0912	5.019
1.0000	.9478	.0522	2.754

3: $\alpha_3 = 0$; $\alpha_4 = 3$ (Third and Fourth Standardized Central Moments) ;

$b_1 = b_2 = 4.874$; $p = 6.158$; Five Parameters to be Estimated

Parameter Estimates

$\hat{r} = 7.129$

$\hat{a}_1 = 5.499$

$\hat{B}_1 = 0.200$

$\hat{a}_2 = 5.056$

$\hat{B}_2 = 0.207$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	8668.4
Departure from Model	SSE	19	32.579
TOTAL	SST	23	8701.0

Coefficient of Determination $R^2 = .99626$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0625	.0679	-.0054	0.022
.0625	.1215	-.0590	1.566
.1800	.1937	-.0137	0.060
.3200	.2771	.0429	0.460
.4000	.4076	-.0076	0.012
.6000	.5483	.0517	0.540
.7551	.6695	.0856	1.624
.7000	.7699	-.0699	1.378
.1633	.0884	.0749	3.408
.1600	.1525	.0075	0.022
.3200	.2532	.0668	1.178
.4200	.3654	.0546	0.644
.5000	.5046	-.0046	0.004
.5600	.6438	-.0838	1.530
.7000	.7602	-.0602	0.995
.7400	.8431	-.1031	4.014
.2800	.1542	.1258	6.062
.3673	.2656	.1018	2.601
.4400	.4094	.0306	0.194
.7200	.5903	.1297	3.477
.8400	.7803	.0597	1.041
.9000	.9197	-.0197	0.263
1.0000	.9767	.0233	1.193
1.0000	.9942	.0058	0.291

4: Takahasi Burr - $r = 0$; $\alpha_3 = 0$; $\alpha_4 = 3$ (Third and Fourth Standardized Central Moments); $b_1 = b_2 = 4.874$; $p = 6.158$; Four Parameters to be Estimated

Parameter Estimates

$\hat{a}_1 = 4.932$
$\hat{B}_1 = 0.246$
$\hat{a}_2 = 4.585$
$\hat{B}_2 = 0.237$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	3	1309.0
Departure from Model	SSE	20	114.16
TOTAL	SST	23	1423.2

Coefficient of Determination $R^2 = .91978$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0625	.0377	.0248	0.813
.0625	.0849	-.0224	0.309
.1800	.1588	.0212	0.168
.3200	.2533	.0667	1.176
.4000	.4114	-.0114	0.027
.6000	.5849	.0151	0.047
.7551	.7280	.0271	0.182
.7000	.8357	-.1357	6.706
.1633	.0482	.1151	14.133
.1600	.0989	.0611	2.094
.3200	.1906	.1294	5.423
.4200	.3043	.1157	3.162
.5000	.4569	.0431	0.374
.5600	.6174	-.0573	0.696
.7000	.7533	-.0533	0.764
.7400	.8480	-.1080	4.521
.2800	.0827	.1973	25.672
.3673	.1693	.1978	13.626
.4400	.2940	.1460	5.132
.7200	.4632	.2568	13.259
.8400	.6569	.1831	7.440
.9000	.8241	.0759	1.987
1.0000	.9172	.0828	4.513
1.0000	.9625	.0375	1.947

APPENDIX IV

Data of J. T. Martin (1942): The toxicities to _Macrosiphoniella sanborni_ of rotenone, a deguelin concentrate, and of a mixture. Tests of 17 November 1938. Fivefold replication. Results one day after spraying. Medium 0.5% saponin, containing 5% alcohol. Tattersfield apparatus. The data as described by Martin, the translated data, and the analyses of this set of data (data set four) are in this appendix.

DATA AS DESCRIBED BY MARTIN

TRANSLATED DATA

CONCENTRATIONS (mg./ℓ.) X(1) Rotenone	X(2) Deguelin Concentrate	N_i Number of Insects Used	Percent Mortality	p_i
10.2	0.0	50	88.0	.8800
7.7	0.0	49	85.7	.8571
5.1	0.0	46	52.2	.5217
3.8	0.0	48	33.3	.3333
2.6	0.0	50	12.0	.1200
0.0	50.5	48	100.0	1.0000
0.0	40.4	50	94.0	.9400
0.0	30.3	49	95.9	.9592
0.0	20.2	48	70.8	.7083
0.0	10.1	48	37.5	.3750
0.0	5.1	49	32.6	.3265
5.1	20.3	50	96.0	.9600
4.0	16.3	46	93.5	.9348
3.0	12.2	48	79.2	.7917
2.0	8.1	46	58.7	.5870
1.0	4.1	46	47.8	.4783
0.5	2.0	47	14.9	.1489

BIVARIATE NORMIT ANALYSIS

Parameter Estimates

$\hat{a}_1 = -2.775$

$\hat{B}_1 = 1.762$

$\hat{a}_2 = -1.645$

$\hat{B}_2 = 0.823$

$\rho = -0.530$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	1656.0
Departure from Model	SSE	12	27.146
TOTAL	SST	16	1683.1

Coefficient of Determination $R^2 = .98387$

Residual Analysis

p_i	P_i	Residual	Chi-square
.8800	.9059	-.0259	0.392
.8571	.7940	.0631	1.193
.5217	.5377	-.0160	0.047
.3333	.3359	-.0026	0.001
.1200	.1374	-.0174	0.128
1.0000	.9432	.0568	2.893
.9400	.9190	.0210	0.297
.9592	.8773	.0819	3.053
.7083	.7962	-.0879	2.284
.3750	.6017	-.2267	10.297
.3265	.3805	-.0540	0.605
.9600	.9649	-.0049	0.036
.9348	.9084	.0264	0.386
.7917	.7893	.0024	0.002
.5870	.5826	.0044	0.004
.4783	.3170	.1613	5.528
.1489	.1506	-.0017	0.001

BIVARIATE NORMIT ANALYSIS

Parameter Estimates

$\hat{a}_1 = -4.184$

$\hat{B}_1 = 2.545$

$\hat{a}_2 = -4.978$

$\hat{B}_2 = 1.618$

$\hat{\rho} = -0.743$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	921.36
Departure from Model	SSE	7	28.947
TOTAL	SST	11	950.30

Coefficient of Determination $R^2 = .96954$

Residual Analysis

p_i	P_i	Residual	Chi-square
1.0000	.9577	.0423	1.236
.8500	.8396	.0104	0.032
.3750	.4846	-.1096	2.306
.0816	.0904	-.0088	0.046
.9216	.9145	.0071	0.032
.9184	.8172	.1012	3.359
.3878	.5467	-.1589	4.993
.0652	.2830	-.2178	10.753
.9200	.9217	-.0017	0.002
.6667	.6173	.0492	0.491
.3182	.2258	.0924	2.149
.1020	.0458	.0562	3.548

BIVARIATE LOGIT ANALYSIS

Parameter Estimates

$\hat{a}_0 = -1.000$

$\hat{a}_1 = -1.611$

$\hat{B}_1 = -3.509$

$\hat{a}_2 = -3.136$

$\hat{B}_2 = -3.145$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	675.39
Departure from Model	SSE	7	30.169
TOTAL	SST	11	705.56

Coefficient of Determination $R^2 = .95724$

Residual Analysis

p_i	P_i	Residual	Chi-square
1.0000	.9240	.0760	2.303
.8500	.8159	.0341	0.310
.3750	.5164	-.1414	3.844
.0816	.1510	-.0694	1.841
.9216	.9221	-.0005	0.000
.9184	.8275	.0909	2.836
.3878	.5095	-.1217	2.904
.0652	.2133	-.1481	6.010
.9200	.8549	.0651	1.707
.6667	.5743	.0924	1.678
.3182	.2141	.1041	2.835
.1020	.0441	.0579	3.902

BIVARIATE BURRIT ANALYSES

1: General Case - Eight Parameters to be Estimated

Parameter Estimates

$\hat{r} = 7.345$

$\hat{b}_1 = 6.078$

$\hat{b}_2 = 5.388$

$\hat{p} = 6.345$

$\hat{a}_1 = 0.224$

$\hat{B}_1 = 0.373$

$\hat{a}_2 = -0.483$

$\hat{B}_2 = 0.260$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	7	1081.9
Departure from Model	SSE	4	29.221
TOTAL	SST	11	1111.1

Coefficient of Determination $R^2 = .97370$

Residual Analysis

p_i	P_i	Residual	Chi-square
1.0000	.9691	.0309	0.893
.8500	.8526	-.0025	0.002
.3750	.4718	-.0967	1.803
.0816	.0896	-.0080	0.039
.9216	.9361	-.0145	0.180
.9184	.8434	.0750	2.087
.3878	.5585	-.1707	5.794
.0652	.2823	-.2171	10.704
.9200	.8784	.0416	0.811
.6667	.5871	.0796	1.254
.3182	.2258	.0924	2.151
.1020	.0460	.0560	3.503

2: Takahasi Burr - $r = 0$; Seven Parameters to be Estimated

Parameter Estimates

$\hat{b}_1 = 9.296$
$\hat{b}_2 = 7.696$
$\hat{p} = 77.198$
$\hat{a}_1 = 1.655$
$\hat{B}_1 = 0.183$
$\hat{a}_2 = 0.700$
$\hat{B}_2 = 0.142$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	6	1508.3
Departure from Model	SSE	5	31.901
TOTAL	SST	11	1540.2

Coefficient of Determination $R^2 = .97929$

Residual Analysis

p_i	P_i	Residual	Chi-square
1.0000	.9804	.0196	0.559
.8500	.8623	-.0123	0.051
.3750	.4921	-.1171	2.631
.0816	.1317	-.0501	1.074
.9216	.9509	-.0293	0.935
.9184	.8434	.0751	2.089
.3878	.5262	-.1384	3.767
.0652	.2614	-.1962	9.171
.9200	.7874	.1326	5.253
.6667	.5364	.1304	3.280
.3182	.2376	.0806	1.577
.1020	.0602	.0418	1.515

3: α_3 = ; α_4 = 3 (Third and Fourth Standardized Central Moments) ;

$b_1 = b_2 = 4.874$; $p = 6.158$; Five Parameters to be Estimated

Parameter Estimates

$\hat{r} = 7.158$

$\hat{a}_1 = -0.085$

$\hat{B}_1 = 0.414$

$\hat{a}_2 = -0.588$

$\hat{B}_2 = 0.260$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	862.11
Departure from Model	SSE	7	29.379
TOTAL	SST	11	891.49

Coefficient of Determination $R^2 = .96705$

Residual Analysis

p_i	P_i	Residual	Chi-square
1.0000	.9598	.0402	1.174
.8500	.8426	.0074	0.016
.3750	.4811	-.1061	2.166
.0816	.0925	-.0109	0.069
.9216	.9181	.0035	0.009
.9184	.8217	.0697	3.126
.3878	.5500	-.1622	5.205
.0652	.2887	-.2235	11.188
.9200	.8762	.0438	0.883
.6667	.5942	.0725	1.046
.3182	.2333	.0849	1.773
.1020	.0504	.0516	2.723

4: Takahasi Burr - $r = 0$; $\alpha_3 = 0$; $\alpha_4 = 3$ (Third and Fourth Standardized Central Moments); $b_1 = b_2 = 4.874$; $p = 6.158$; Four Parameters to be Estimated

Parameter Estimates

$\hat{a}_1 = -0.080$

$\hat{B}_1 = 0.418$

$\hat{a}_2 = -0.488$

$\hat{B}_2 = 0.254$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	3	905.59
Departure from Model	SSE	8	36.603
TOTAL	SST	11	942.19

Coefficient of Determination $R^2 = .96115$

Residual Analysis

p_i	P_i	Residual	Chi-square
1.0000	.9659	.0341	0.989
.8500	.8588	-.0088	0.025
.3750	.5037	-.1287	3.179
.0816	.0993	-.0177	0.172
.9216	.9207	.0009	0.001
.9184	.8294	.0890	2.741
.3878	.5697	-.1819	6.615
.0652	.3118	-.2466	13.032
.9200	.7923	.1277	4.954
.6667	.5576	.1091	2.317
.3182	.2484	.0698	1.147
.1020	.0611	.0409	1.430

APPENDIX VI

Data of J. T. Martin (1942): The toxicities to *Macrosiphoniella sanborni* of rotenone, ℓ-α-toxicarol, and of a mixture. Tests of 24 September 1941. Fivefold replication. Results one day after spraying. Medium 0.5% saponin, containing 5% of alcohol. Tattersfield apparatus. The data as described by Martin, the translated data, and the analyses of this set of data (data set six) are in this appendix.

DATA AS DESCRIBED BY MARTIN | TRANSLATED DATA

CONCENTRATIONS (mg./ℓ.) X(1) Rotenone	X(2) ℓ-α-Toxicarol	N_i Number of Insects Used	Percent Mortality	p_i
1.06	0.0	51	100.0	1.0000
0.85	0.0	48	97.9	.9792
0.64	0.0	48	93.8	.9375
0.42	0.0	48	62.5	.6250
0.21	0.0	48	12.5	.1250
0.0	9.75	49	100.0	1.0000
0.0	7.80	48	97.9	.9792
0.0	5.85	52	98.1	.9808
0.0	3.90	49	87.7	.8776
0.0	1.95	48	50.0	.5000
0.53	4.88	48	100.0	1.0000
0.42	3.90	48	100.0	1.0000
0.32	2.93	49	89.8	.8980
0.21	1.95	50	82.0	.8200
0.11	0.98	50	30.0	.3000

BIVARIATE NORMIT ANALYSIS

1: Using Relative Change in Weighted Sum of Squares Due To Error as Part of the Convergence Criteria

Parameter Estimates

$\hat{a}_1 = 2.372$

$\hat{B}_1 = 2.212$

$\hat{a}_2 = -0.580$

$\hat{B}_2 = 1.328$

$\hat{\rho} = -0.432$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	30672
Departure from Model	SSE	10	10.144
TOTAL	SST	14	30682

Coefficient of Determination $R^2 = .99967$

Unweighted Sum of Squares Due to Error = .03204

Residual Analysis

p_i	P_i	Residual	Chi-square
1.0000	.9938	.0062	0.318
.9792	.9779	.0013	0.004
.9375	.9169	.0206	0.267
.6250	.6747	-.0497	0.541
.1250	.1401	-.0151	0.090
1.0000	.9928	.0072	0.357
.9792	.9845	-.0053	0.089
.9808	.9614	.0194	0.528
.8776	.8903	-.0127	0.081
.5000	.6207	-.1207	2.971
1.0000	.9985	.0015	0.072
1.0000	.9893	.0107	0.518
.8980	.9385	-.0405	1.391
.8200	.7130	.1070	2.796
.3000	.2780	.0220	1.210

BIVARIATE NORMIT ANALYSIS

2: Using Relative Change in Unweighted Sum of Squares

Due to Error as Part of the Convergence Criteria

Parameter Estimates

$\hat{a}_1 = 2.279$

$\hat{B}_1 = 2.063$

$\hat{a}_2 = -0.589$

$\hat{B}_2 = 1.290$

$\hat{\rho} = -0.9999$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	$.19 \cdot 10^{18}$
Departure from Model	SSE	10	$.78 \cdot 10^{15}$
TOTAL	SST	14	$.19 \cdot 10^{18}$

Coefficient of Determination $R^2 = .99592$

Unweighted Sum of Squares Due to Error = .03127

Residual Analysis

p_i	P_i	Residual	Chi-square
1.0000	.9918	.0082	0.422
.9792	.9740	.0052	0.050
.9375	.9129	.0246	0.367
.6250	.6878	-.0628	0.882
.1250	.1736	-.0496	0.790
1.0000	.9906	.0094	0.465
.9792	.9808	-.0016	0.006
.9808	.9545	.0263	0.828
.8776	.8784	-.0008	0.000
.5000	.6074	-.1074	2.321
1.0000	1.0000	.0000	0.000
1.0000	1.0000	.0000	0.000
.8980	1.0000	-.1020	$0.78 \cdot 10^{15}$
.8200	.7810	.0390	0.445
.3000	.2806	.0194	0.928

BIVARIATE LOGIT ANALYSIS

Parameter Estimates

$\hat{a}_0 = -1.000$

$\hat{a}_1 = 1.053$

$\hat{B}_1 = -4.027$

$\hat{a}_2 = -0.436$

$\hat{B}_2 = -2.332$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	20170
Departure from Model	SSE	10	11.975
TOTAL	SST	14	20182

Coefficient of Determination $R^2 = .99941$

Residual Analysis

p_i	P_i	Residual	Chi-square
1.0000	.9888	.0113	0.580
.9792	.9731	.0061	0.069
.9375	.9201	.0174	0.198
.6250	.6786	-.0536	0.632
.1250	.1146	.0104	0.051
1.0000	.9865	.0135	0.670
.9792	.9779	.0014	0.004
.9808	.9570	.0238	0.717
.8776	.8962	-.0186	0.183
.5000	.6317	-.1317	3.579
1.0000	.9979	.0021	0.102
1.0000	.9869	.0131	0.635
.8980	.9286	-.0306	0.690
.8200	.6976	.1225	3.554
.3000	.2652	.0348	0.311

BIVARIATE BURRIT ANALYSES

1: General Case - Eight Parameters to be Estimated

Parameter Estimates

$\hat{r} = 8.817$

$\hat{b}_1 = 5.045$

$\hat{b}_2 = 5.469$

$\hat{p} = 7.826$

$\hat{a}_1 = 2.869$

$\hat{B}_1 = 0.339$

$\hat{a}_2 = 3.055$

$\hat{B}_2 = 0.184$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	7	67305
Departure from Model	SSE	7	9.6604
TOTAL	SST	14	67314

Coefficient of Determination $R^2 = .99986$

Residual Analysis

p_i	P_i	Residual	Chi-square
1.0000	.9948	.0052	0.266
.9792	.9797	-.0005	0.001
.9375	.9144	.0231	0.327
.6250	.6434	-.0184	0.070
.1250	.1202	.0049	0.011
1.0000	.9932	.0068	0.335
.9792	.9851	.0059	0.113
.9808	.9612	.0196	0.536
.8776	.8850	-.0074	0.026
.5000	.6025	-.1025	2.107
1.0000	.9994	.0006	0.030
1.0000	.9920	.0080	0.385
.8980	.9355	-.0375	1.140
.8200	.6866	.1334	4.135
.3000	.2734	.0266	0.179

2: Takahasi Burr - r = 0 ; Seven Parameters to be Estimated

Parameter Estimates

$\hat{b}_1 = 4.451$

$\hat{b}_2 = 5.250$

$\hat{p} = 7.940$

$\hat{a}_1 = 2.746$

$\hat{B}_1 = 0.346$

$\hat{a}_2 = 2.909$

$\hat{B}_2 = 0.192$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	6	13398
Departure from Model	SSE	8	13.648
TOTAL	SST	14	13412

Coefficient of Determination $R^2 = .99898$

Residual Analysis

p_i	P_i	Residual	Chi-square
1.0000	.9932	.0068	0.349
.9792	.9769	.0023	0.011
.9375	.9139	.0236	0.339
.6250	.6637	-.0387	0.321
.1250	.1385	-.0135	0.073
1.0000	.9953	.0047	0.229
.9792	.9894	-.0102	0.472
.9808	.9708	.0100	0.183
.8776	.9067	-.0291	0.492
.5000	.6402	-.1402	4.095
1.0000	.9854	.0146	0.711
1.0000	.9590	.0410	2.051
.8980	.8865	.0115	0.064
.8200	.6844	.1356	4.255
.3000	.2967	.0033	0.003

3: $\alpha_3 = 0$; $\alpha_4 = 3$ (Third and Fourth Standardized Central Moments) ;

$b_1 = b_2 = 4.874$; $p = 6.158$; Five Parameters to be Estimated

Parameter Estimates

$\hat{r} = 5.265$

$\hat{a}_1 = 2.907$

$\hat{B}_1 = 0.357$

$\hat{a}_2 = 3.215$

$\hat{B}_2 = 0.179$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	23681
Departure from Model	SSE	10	12.047
TOTAL	SST	14	23693

Coefficient of Determination $R^2 = .99949$

Residual Analysis

p_i	P_i	Residual	Chi-square
1.0000	.9943	.0057	0.294
.9792	.9808	-.0016	0.006
.9375	.9267	.0108	0.083
.6250	.6934	-.0684	1.056
.1250	.1561	-.0311	0.353
1.0000	.9826	.0174	0.869
.9792	.9697	.0095	0.148
.9808	.9397	.0411	1.549
.8776	.8637	.0139	0.081
.5000	.6229	-.1229	3.086
1.0000	.9980	.0020	0.095
1.0000	.9868	.0132	0.642
.8980	.9297	-.0317	0.755
.8200	.7131	.1069	2.792
.3000	.3326	.0326	0.240

4: Takahasi Burr - $r = 0$; $\alpha_3 = 0$; $\alpha_4 = 3$ (Third and Fourth Standardized Central Moments); $b_1 = b_2 = 4.874$; $p = 6.158$; Four Parameters to be Estimated

Parameter Estimates

$\hat{a}_1 = 2.741$

$\hat{B}_1 = 0.357$

$\hat{a}_2 = 3.171$

$\hat{B}_2 = 0.189$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	3	6862.7
Departure from Model	SSE	11	23.480
TOTAL	SST	14	6886.2

Coefficient of Determination $R^2 = .99659$

Residual Analysis

p_i	P_i	Residual	Chi-square
1.0000	.9857	.0143	0.741
.9792	.9567	.0225	0.586
.9375	.8604	.0771	2.374
.6250	.5544	.0706	0.969
.1250	.0863	.0387	0.911
1.0000	.9914	.0086	0.423
.9792	.9840	-.0048	0.070
.9808	.9648	.0160	0.393
.8776	.9088	-.0312	0.575
.5000	.6934	-.1934	8.444
1.0000	.9773	.0227	1.114
1.0000	.9479	.0521	2.640
.8980	.8801	.0179	0.148
.8200	.7154	.1046	2.685
.3000	.3814	-.0814	1.405

APPENDIX VII

Data of Ashford and Smith (1964): Exposure to dust and prevalence of pneumoconiosis for groups of mine workers. The data as described by Ashford and Smith, the computed p_i , and the analyses on this set of data (data set seven) are in this appendix.

DATA AS PRESENTED BY ASHFORD AND SMITH | COMPUTED DATA

Period Spent (Years) X(1) Coal-Getting	X(2) Haulage	N_i Number of Men	r_i Number Observed With Pneumoconiosis	p_i
2.1	0.5	135	3	.0222
1.9	6.6	18	2	.1111
1.6	12.0	16	1	.0625
1.4	16.9	17	3	.1765
0.7	21.6	14	2	.1429
1.1	27.6	12	3	.2500
1.2	32.4	22	5	.2273
1.5	37.2	31	7	.2258
2.4	41.6	25	5	.2000
1.4	47.1	17	5	.2941
6.6	0.4	80	7	.0875
6.3	6.7	10	1	.1000
7.1	12.0	14	5	.3571
6.4	17.5	8	2	.2500
6.3	21.9	21	11	.5238
6.9	27.2	14	5	.3571
6.2	32.3	13	7	.5385
7.2	37.3	10	7	.7000
12.2	0.2	71	19	.2676
12.0	6.9	8	1	.1250
11.8	11.8	4	2	.5000
11.0	16.7	7	2	.2857
11.5	22.5	6	3	.5000
12.8	29.5	10	6	.6000
12.5	37.8	4	2	.5000
17.0	0.3	106	53	.5000
16.2	6.6	5	2	.4000
16.8	13.2	5	2	.4000
19.5	17.0	6	4	.6667
17.2	21.5	4	1	.2500
21.8	0.2	58	34	.5862
24.7	7.7	3	0	0.0000
26.0	10.8	4	1	.2500
22.0	23.7	3	1	.3333
26.8	0.2	66	43	.6515
27.5	18.2	4	3	.7500
32.5	13.0	2	2	1.0000
31.7	0.2	33	22	.6667
36.8	0.2	20	11	.5500
42.2	1.0	10	8	.8000

BIVARIATE NORMIT ANALYSIS

Parameter Estimates

$\hat{a}_1 = -2.818$

$\hat{B}_1 = 0.937$

$\hat{a}_2 = -2.446$

$\hat{B}_2 = 0.501$

$\hat{\rho} = -0.320$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	217.75
Departure from Model	SSE	35	38.141
TOTAL	SST	39	255.89

Coefficient of Determination $R^2 = .85095$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0222	.0195	.0027	0.053
.1111	.0800	.0311	0.236
.0625	.1236	-.0611	0.551
.1765	.1579	.0186	0.044
.1429	.1834	-.0405	0.153
.2500	.2200	.0300	0.063
.2273	.2451	-.0179	0.038
.2258	.2703	-.0445	0.311
.2000	.3030	-.1030	1.256
.2941	.3091	-.0150	0.018
.0875	.1488	-.0613	2.375
.1000	.2025	-.1025	0.651
.3571	.2719	.0852	0.514
.2500	.2886	-.0386	0.058
.5238	.3120	.2118	4.390

Residual Analysis--*Continued*

p_i	P_i	Residual	Chi-square
.3571	.3573	-.0002	0.000
.5385	.3611	.1773	1.772
.7000	.4099	.2901	3.480
.2676	.3185	-.0509	0.847
.1250	.3735	-.2485	2.112
.5000	.4047	.0953	0.151
.2857	.4146	-.1289	0.479
.5000	.4580	.0420	0.043
.6000	.5186	.0815	0.266
.5000	.5417	-.0417	0.028
.5000	.4365	.0635	1.737
.4000	.4717	-.0717	0.103
.4000	.5267	-.1267	0.322
.6667	.5947	.0720	0.129
.2500	.5744	-.3243	1.721
.5862	.5287	.0575	0.771
0.0000	.6260	-.6260	5.022
.2500	.6594	-.4094	2.985
.3333	.6586	-.3253	1.412
.6515	.6047	.0468	0.606
.7500	.7069	.0431	0.036
1.0000	.7376	.2624	0.712
.6667	.6638	.0029	0.001
.5500	.7132	-.1632	2.603
.8000	.7591	.0410	0.092

BIVARIATE LOGIT ANALYSIS

Parameter Estimates

$\hat{a}_0 = -1.000$

$\hat{a}_1 = -2.996$

$\hat{B}_1 = -1.627$

$\hat{a}_2 = -4.527$

$\hat{B}_2 = -1.049$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	221.36
Departure from Model	SSE	35	39.020
TOTAL	SST	39	260.38

Coefficient of Determination $R^2 = .85014$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0222	.0291	-.0069	0.226
.1111	.0802	.0309	0.233
.0625	.1211	-.0586	0.516
.1765	.1567	.0197	0.050
.1429	.1828	-.0399	0.150
.2500	.2280	.0220	0.033
.2273	.2592	-.0320	0.117
.2258	.2912	-.0654	0.643
.2000	.3297	-.1297	1.903
.2941	.3416	-.0475	0.170
.0875	.1446	-.0571	2.111
.1000	.1909	-.0909	0.535
.3571	.2575	.0996	0.727
.2500	.2787	-.0287	0.033
.5238	.3063	.2175	4.674

Residual Analysis--*Continued*

p_i	P_i	Residual	Chi-square
.3571	.3564	.0007	0.000
.5385	.3674	.1710	1.636
.7000	.4202	.2796	3.209
.2676	.3106	-.0429	0.612
.1250	.3586	-.2336	1.898
.5000	.3897	.1103	0.205
.2857	.4020	-.1163	0.394
.5000	.4515	.0485	0.057
.6000	.5213	.0787	0.248
.5000	.5533	-.0533	0.046
.5000	.4362	.0638	1.753
.4000	.4631	-.0631	0.080
.4000	.5200	-.1200	0.288
.6667	.5945	.0722	1.298
.2500	.5748	-.3248	1.727
.5862	.5361	.0501	0.586
0.0000	.6290	-.6290	5.086
.2500	.6632	-.4132	3.058
.3333	.6666	-.3333	1.500
.6515	.6178	.0338	0.318
.7500	.7145	.0355	0.025
1.0000	.7447	.2553	0.686
.6667	.6798	-.0132	0.026
.5500	.7302	-.1802	3.295
.8000	.7746	.0254	0.037

BIVARIATE BURRIT ANALYSES

1: General Case - Eight Parameters to be Estimated

Parameter Estimates

$\hat{r} = 8.582$

$\hat{b}_1 = 4.714$

$\hat{b}_2 = 5.933$

$\hat{p} = 7.582$

$\hat{a}_1 = 1.127$

$\hat{B}_1 = 0.147$

$\hat{a}_2 = 3.533$

$\hat{B}_2 = 0.082$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	7	217.30
Departure from Model	SSE	32	38.293
TOTAL	SST	39	255.59

Coefficient of Determination $R^2 = .85018$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0222	.0183	.0039	0.113
.1111	.0734	.0377	0.377
.0625	.1170	-.0545	0.459
.1765	.1528	.0236	0.073
.1429	.1809	-.0381	0.137
.2500	.2208	.0292	0.060
.2273	.2489	-.0217	0.055
.2258	.2774	-.0516	0.412
.2000	.3145	-.1145	1.520
.2941	.3223	-.0282	0.062
.0875	.1500	-.0625	2.448
.1000	.1997	-.0997	0.622
.3571	.2711	.0860	0.524
.2500	.2908	-.0408	0.064
.5238	.3171	.2067	4.142

Residual Analysis--*Continued*

p_i	P_i	Residual	Chi-square
.3571	.3670	-.0099	0.006
.5385	.3737	.1648	1.509
.7000	.4274	.2726	3.037
.2676	.3174	-.0498	0.811
.1250	.3704	-.2454	2.065
.5000	.4050	.0950	0.150
.2857	.4188	-.1331	0.510
.5000	.4677	.0323	0.025
.6000	.5355	.0645	0.167
.5000	.5653	-.0653	0.069
.5000	.4344	.0656	1.856
.4000	.4689	-.0689	0.095
.4000	.5303	-.1303	0.341
.6667	.6034	.0632	0.100
.2500	.5866	-.3366	1.869
.5862	.5275	.0587	0.802
0.0000	.6271	-.6271	5.045
.2500	.6642	-.4142	3.076
.3333	.6754	-.3420	1.601
.6515	.6048	.0467	0.603
.7500	.7200	.0300	0.018
1.0000	.7466	.2534	0.679
.6667	.6653	.0013	0.000
.5500	.7160	-.1660	2.712
.8000	.7619	.0381	0.080

2: Takahasi Burr - r = 0 ; Seven Parameters to be Estimated

Parameter Estimates

$\hat{b}_1 = 5.074$

$\hat{b}_2 = 5.026$

$\hat{p} = 6.351$

$\hat{a}_1 = 1.369$

$\hat{B}_1 = 0.149$

$\hat{a}_2 = 3.850$

$\hat{B}_2 = 0.075$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	6	219.84
Departure from Model	SSE	33	38.765
TOTAL	SST	39	258.60

Coefficient of Determination $R^2 = .85010$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0222	.0222	.0000	0.000
.1111	.0980	.0131	0.035
.0625	.1436	-.0811	0.856
.1765	.1779	-.0014	0.000
.1429	.2034	-.0605	0.317
.2500	.2383	.0117	0.009
.2273	.2622	-.0349	0.139
.2258	.2857	-.0598	0.544
.2000	.3144	-.1144	1.517
.2941	.3224	-.0282	0.062
.0875	.1510	-.0635	2.519
.1000	.2117	-.1117	0.748
.3571	.2752	.0819	0.471
.2500	.2908	-.0408	0.064
.5238	.3117	.2121	4.402

Residual Analysis--*Continued*

p_i	P_i	Residual	Chi-square
.3571	.3507	.0064	0.003
.5385	.3554	.1831	1.903
.7000	.3960	.3040	3.865
.2676	.3165	-.0489	0.784
.1250	.3683	-.2433	2.036
.5000	.3937	.1064	0.190
.2857	.4001	-.1144	0.382
.5000	.4364	.0636	0.099
.6000	.4876	.1125	0.506
.5000	.5059	-.0059	0.001
.5000	.4342	.0658	1.867
.4000	.4615	-.0615	0.076
.4000	.5057	-.1057	0.223
.6667	.5663	.1004	0.246
.2500	.5428	-.2928	1.382
.5862	.5276	.0586	0.798
0.0000	.6110	-.6110	4.712
.2500	.6388	-.3888	2.620
.3333	.6218	-.2885	1.062
.6515	.6056	.0460	0.584
.7500	.6760	.0740	0.100
1.0000	.7149	.2851	0.798
.6667	.6666	.0001	0.000
.5500	.7176	-.1676	2.773
.8000	.7632	.0368	0.075

3: $\alpha_3 = 0$; $\alpha_4 = 3$ (Third and Fourth Standardized Central Moments) ;

$b_1 = b_2 = 4.874$; $p = 6.158$; Five Parameters to be Estimated

Parameter Estimates

$\hat{r} = 4.742$

$\hat{a}_1 = 1.259$

$\hat{B}_1 = 0.151$

$\hat{a}_2 = 3.405$

$\hat{B}_2 = 0.078$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	4	218.53
Departure from Model	SSE	35	38.338
TOTAL	SST	39	256.87

Coefficient of Determination $R^2 = .85075$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0222	.0211	.0011	0.008
.1111	.0905	.0206	0.093
.0625	.1340	-.0715	0.704
.1765	.1670	.0094	0.011
.1429	.1913	-.0485	0.212
.2500	.2261	.0239	0.039
.2273	.2499	-.0226	0.060
.2258	.2740	-.0481	0.361
.2000	.3061	-.1061	1.326
.2941	.3103	-.0162	0.021
.0875	.1528	-.0653	2.634
.1000	.2139	-.1139	0.772
.3571	.2820	.0752	0.391
.2500	.2978	-.0478	0.087
.5238	.3199	.2039	4.013

Residual Analysis--*Continued*

p_i	P_i	Residual	Chi-square
.3571	.3631	-.0060	0.002
.5385	.3660	.1725	1.666
.7000	.4127	.2873	3.405
.2676	.3187	-.0511	0.854
.1250	.3793	-.2543	2.198
.5000	.4103	.0897	0.133
.2857	.4198	-.1341	0.517
.5000	.4615	.0385	0.036
.6000	.5201	.0800	0.256
.5000	.5420	-.0419	0.028
.5000	.4353	.0647	1.806
.4000	.4749	-.0749	0.112
.4000	.5294	-.1294	0.336
.6667	.5964	.0703	0.123
.2500	.5760	-.3260	1.740
.5862	.5273	.0589	0.807
0.0000	.6278	-.6278	5.059
.2500	.6612	-.4112	3.018
.3333	.6598	-.3265	1.425
.6515	.6039	.0477	0.626
.7500	.7086	.0414	0.033
1.0000	.7396	.2604	0.704
.6667	.6638	.0029	0.001
.5500	.7139	-.1639	2.632
.8000	.7605	.0395	0.086

4: Takahasi Burr - $r = 0$; $\alpha_3 = 0$; $\alpha_4 = 3$ (Third and Fourth Standardized Central Moments); $b_1 = b_2 = 4.874$; $p = 6.158$; Four Parameters to be Estimated

Parameter Estimates

$\hat{a}_1 = 1.313$
$\hat{B}_1 = 0.150$
$\hat{a}_2 = 3.783$
$\hat{B}_2 = 0.074$

Chi-square Analysis Table

source		d.f.	
Due to Model	SSR	3	219.06
Departure from Model	SSE	36	38.766
TOTAL	SST	39	257.83

Coefficient of Determination $R^2 = .84964$

Residual Analysis

p_i	P_i	Residual	Chi-square
.0222	.0240	-.0018	0.018
.1111	.1007	.0104	0.021
.0625	.1456	-.0831	0.888
.1765	.1791	-.0026	0.001
.1429	.2037	-.0608	0.319
.2500	.2380	.0120	0.010
.2273	.2613	-.0341	0.132
.2258	.2845	-.0587	0.524
.2000	.3135	-.1135	1.496
.2941	.3200	-.0259	0.052
.0875	.1566	-.0691	2.894
.1000	.2179	-.1179	0.815
.3571	.2809	.0763	0.403
.2500	.2956	-.0455	0.080
.5238	.3159	.2080	4.202

Residual Analysis--*Continued*

p_i	P_i	Residual	Chi-square
.3571	.3542	.0029	0.001
.5385	.3581	.1804	1.841
.7000	.3983	.3017	3.799
.2676	.3210	-.0534	0.930
.1250	.3735	-.2485	2.110
.5000	.3984	.1016	0.172
.2857	.4045	-.1188	0.410
.5000	.4399	.0601	0.088
.6000	.4899	.1101	0.485
.5000	.5075	-.0075	0.001
.5000	.4363	.0637	1.750
.4000	.4645	-.0645	0.084
.4000	.5079	-.1079	0.233
.6667	.5669	.0997	0.243
.2500	.5441	-.2941	1.395
.5862	.5271	.0591	0.814
0.0000	.6096	-.6096	4.685
.2500	.6369	-.3869	2.588
.3333	.6209	-.2876	1.054
.6515	.6027	.0489	0.658
.7500	.6734	.0766	0.107
1.0000	.7108	.2892	0.814
.6667	.6619	.0048	0.003
.5500	.7115	-.1615	2.542
.8000	.7563	.0437	0.104

APPENDIX VIII

Listings of FORTRAN subroutines used in evaluating the functions, partial derivatives, and the weights are in this appendix.

1. Bivariate Normal Subroutines

```
      DOUBLE PRECISION FUNCTION F(X,A)
C     BIVARIATE NORMAL FUNCTION
      DIMENSION X(1),A(1)
      DOUBLE PRECISION X,A,AA,B,R,S,GOFU,THA ,C,D
      AA=A(1)+A(2)*X(1)
      B=A(3)+A(4)*X(2)
      IF (DABS(A(5)).LE.0.9999D+00) GO TO 16
      A(5)=DSIGN(0.9999D+00,A(5))
   16 R=A(5)
      S=DSQRT(1.-R*R)
      C=GOFU(AA)
      D=GOFU(B)
      F=C+D-THA(AA,B/AA)-THA(B,AA/B)
     1+THA(AA,(B-R*AA)/(AA*S))+THA(B,(AA-R*B)/(B*S))-C*D
      RETURN
      END
```

```
      SUBROUTINE PD(X,A,FXA,P)
      DIMENSION X(1),A(1),P(1)
      DOUBLE PRECISION X,A,AA,R,S,B,FXA,P,WATE,GOFU,GPRIME,C,D
C     A1=ALPHA1,A2=BETA1,A3=ALPHA2,A4=BETA2,A5=RHO
C     BIVARIATE NORMAL PARTIALS
      AA=A(1)+A(2)*X(1)
      B=A(3)+A(4)*X(2)
      R=A(5)
      S=DSQRT(1.-R*R)
      C=GPRIME(AA)
      D=(B-R*AA)/S
      P(1)=C*(1.-GOFU(D))
      P(2)=X(1)*P(1)
      P(3)=GPRIME(B)*(1.-GOFU((AA-R*B)/S))
      P(4)=X(2)*P(3)
      P(5)=-C*GPRIME(D)/S
      RETURN
      END
```

```
      DOUBLE PRECISION FUNCTION GPRIME(X)
      DOUBLE PRECISION X,A1,A2,A3
      A1 =X*X*(-0.5D 0)
      IF (A1.LE.-150) GO TO 15
      A2=DEXP(A1)
      A3 = .398942280401433D+00
      GPRIME = A2*A3
      GO TO 16
   15 GPRIME=0
   16 RETURN
      END
```

```
      DOUBLE PRECISION FUNCTION GOFU(U)
      DIMENSION Y(160)
      DOUBLE PRECISION U,GU,X,XSET,DELTA,GPR,DELK,DELTAX,Y,SUM,TOP
      IF(Y(17)-.69146246127D+00)2,5,2
C
 2    Y(  1) =      .50000000000D+00
      Y(  2) =      .39894228040D+00
      Y(  3) =      .00000000000D+00
      Y(  4) =     -.66490380066D-01
      Y(  5) =      .00000000000D+00
      Y(  6) =      .99735570100D-02
      Y(  7) =      .00000000000D+00
      Y(  8) =     -.11873282155D-02
      Y(  9) =      .59870632568D+00
      Y( 10) =      .38666811680D+00
      Y( 11) =     -.48333514600D-01
      Y( 12) =     -.60416893250D-01
      Y( 13) =      .11831641595D-01
      Y( 14) =      .84709519078D-02
      Y( 15) =     -.19305085421D-02
      Y( 16) =     -.93949992204D-03
      Y( 17) =      .69146246127D+00
      Y( 18) =      .35206532676D+00
      Y( 19) =     -.88016331691D-01
      Y( 20) =     -.44008165845D-01
      Y( 21) =      .20170409346D-01
      Y( 22) =      .45841839423D-02
      Y( 23) =     -.30714032413D-02
      Y( 24) =     -.32635023779D-03
      Y( 25) =      .77337264763D+00
      Y( 26) =      .30113743216D+00
      Y( 27) =     -.11292653706D+00
      Y( 28) =     -.21957937761D-01
      Y( 29) =      .22938202840D-01
      Y( 30) =     -.14703976180D-03
```

```
Y( 31) =     -.30400470751D-02
Y( 32) =      .34322406302D-03
Y( 33) =      .84134474606D+00
Y( 34) =      .24197072452D+00
Y( 35) =     -.12098536226D+00
Y( 36) =      .00000000000D+00
Y( 37) =      .20164227043D-01
Y( 38) =     -.40328454087D-02
Y( 39) =     -.20164227043D-02
Y( 40) =      .76816103022D-03
Y( 41) =      .89435022633D+00
Y( 42) =      .18264908539D+00
Y( 43) =     -.11415567837D+00
Y( 44) =      .17123351755D-01
Y( 45) =      .13674898971D-01
Y( 46) =     -.59872275061D-02
Y( 47) =     -.57598079906D-03
Y( 48) =      .81561889341D-03
Y( 49) =      .93319279874D+00
Y( 50) =      .12951759567D+00
Y( 51) =     -.97138196750D-01
Y( 52) =      .26982832430D-01
Y( 53) =      .60711372969D-02
Y( 54) =     -.58687660536D-02
Y( 55) =      .65770654049D-03
Y( 56) =      .55772550961D-03
Y( 57) =      .95994084314D+00
Y( 58) =      .86277318827D-01
Y( 59) =     -.75492653974D-01
Y( 60) =      .29657828346D-01
Y( 61) =     -.39319090611D-03
Y( 62) =     -.43110574348D-02
Y( 63) =      .13098172060D-02
Y( 64) =      .18576682170D-03
Y( 65) =      .97724986805D+00
Y( 66) =      .53990966513D-01
Y( 67) =     -.53990966513D-01
Y( 68) =      .26995483257D-01
Y( 69) =     -.44992472094D-02
Y( 70) =     -.22496236047D-02
Y( 71) =      .13497741628D-02
Y( 72) =     -.11783742691D-03
Y( 73) =      .98777552735D+00
Y( 74) =      .31739651836D-01
Y( 75) =     -.35707108315D-01
Y( 76) =      .21490389264D-01
Y( 77) =     -.61371592417D-02
Y( 78) =     -.46183673081D-03
Y( 79) =      .99147667294D-03
Y( 80) =     -.26370836740D-03
```

```
Y( 81) =      .99379033467D+00
Y( 82) =      .17528300494D-01
Y( 83) =     -.21910375617D-01
Y( 84) =      .15337262932D-01
Y( 85) =     -.59340600629D-02
Y( 86) =      .66644059169D-03
Y( 87) =      .51352442853D-03
Y( 88) =     -.26273974729D-03
Y( 89) =      .99702023677D+00
Y( 90) =      .90935625017D-02
Y( 91) =     -.12503648440D-01
Y( 92) =      .99460839861D-02
Y( 93) =     -.47539913338D-02
Y( 94) =      .11227826357D-02
Y( 95) =      .11925680315D-03
Y( 96) =     -.18051548644D-03
Y( 97) =      .99865010197D+00
Y( 98) =      .44318484120D-02
Y( 99) =     -.66477726180D-02
Y(100) =      .59091312160D-02
Y(101) =     -.33238863090D-02
Y(102) =      .11079621030D-02
Y(103) =     -.11079621030D-03
Y(104) =     -.84416160226D-04
Y(105) =      .99942297496D+00
Y(106) =      .20290480573D-02
Y(107) =     -.32972030931D-02
Y(108) =      .32337953413D-02
Y(109) =     -.20779248659D-02
Y(110) =      .86558186169D-03
Y(111) =     -.19180019295D-03
Y(112) =     -.13995370141D-04
Y(113) =      .99976737091D+00
Y(114) =      .87268269505D-03
Y(115) =     -.15271947163D-02
Y(116) =      .16362800532D-02
Y(117) =     -.11772125938D-02
Y(118) =      .57860680771D-03
Y(119) =     -.18055895865D-03
Y(120) =      .21397716502D-04
Y(121) =      .99991158271D+00
Y(122) =      .35259568237D-03
Y(123) =     -.66111690444D-03
Y(124) =      .76763018349D-03
Y(125) =     -.60946714628D-03
Y(126) =      .34195583218D-03
Y(127) =     -.13246010895D-03
Y(128) =      .30251745009D-04
Y(129) =      .99996832876D+00
Y(130) =      .13383022576D-03
```

```
      Y(131) =     -.26766045153D-03
      Y(132) =      .33457556441D-03
      Y(133) =     -.28996548916D-03
      Y(134) =      .18178605667D-03
      Y(135) =     -.82528639221D-04
      Y(136) =      .25518025191D-04
      Y(137) =      .99998931147D+00
      Y(138) =      .47718636540D-04
      Y(139) =     -.10140210265D-03
      Y(140) =      .13569987266D-03
      Y(141) =     -.12728076426D-03
      Y(142) =      .87833668723D-04
      Y(143) =     -.45244746779D-04
      Y(144) =      .17013635696D-04
      Y(145) =      .99999660233D+00
      Y(146) =      .15983741107D-04
      Y(147) =     -.35963417490D-04
      Y(148) =      .51281169385D-04
      Y(149) =     -.51697412642D-04
      Y(150) =      .38835495972D-04
      Y(151) =     -.22233633626D-04
      Y(152) =      .96697768581D-05
      Y(153) =      .99999898292D+00
      Y(154) =      .50295072886D-05
      Y(155) =     -.11945079810D-04
      Y(156) =      .18074791818D-04
      Y(157) =     -.19472968649D-04
      Y(158) =      .15788101444D-04
      Y(159) =     -.99025178234D-05
      Y(160) =      .48400297796D-05
C
    5 IF(U) 11,10,12
   10 GU=.5D+00
      GO TO 100
   11 X=DABS(U)
      GO TO 13
   12 X=U
   13 IF(X-7.0D+00)15,14,14
   14 IF(U)141,10,142
  141 GU=0.0D+00
      GO TO 100
  142 GU=1.0D+00
  100 GOFU=GU
      RETURN
   15 IF(X-4.87499D+00)16,16,40
   16 XSET=X*4.0D+00
      XSST=XSET+.5D+00
      I=IFIX(XSST)
      XSET=DFLOAT(I)
      DELTA=X-(XSET*.25D+00)
  201 K=I*8+1
      I=K+7
      SUM=0.0D+00
```

```
      DO 20 J=K,I,1
      L=K+I-J
      SUM=SUM*DELTA
      SUM=Y(L)+SUM
   20 CONTINUE
      IF(U) 21,10,22
   21 GU=1.0D+00 -SUM
      GO TO 100
   22 GU=SUM
      GO TO 100
   40 XSET=-X*X
      IF (XSET.LE.-300) GO TO 101
      GPR=(DEXP(XSET*.5D+00))*.398942280401433D+00
      GO TO 102
  101 GPR=0
  102 DELTA=1.0D+00/X
      SUM=DELTA
      TOP=1.0D+00
   41 DELK=TOP/XSET
      DELTAX=DELTA*DELK
      IF (DABS(DELTA)-DABS(DELTAX))45,45,43
   43 DELTA=DELTAX
      SUM=SUM+DELTA
      IF(DABS(GPR*DELTA)-.5D-9)45,45,42
   42 TOP=TOP+2.0D+00
      GO TO 41
   45 SUM=GPR*SUM
      IF(U) 22,10,21
      END
```

```
      DOUBLE PRECISION FUNCTION THA(HX,AX)
      DOUBLE PRECISION HX,AX,AA,U,H,A,SUM,C,DA,TA,TX,X,Y,Z ,GOFU
      DIMENSION AA(9),U(9)
      DATA AA(1),AA(9),AA(2),AA(8),AA(3),AA(7),AA(4),AA(6),AA(5)/2*.4063
     17194181E-1,2*.90324080347E-1,2*.1303053482,2*.15617353852,.1651196
     2775/,U(1),U(2),U(3),U(4),U(5),U(6),U(7),U(8),U(9)/.15919880246E-1,
     3.81984446337E-1,.19331428365,.3378732883,.5,.6621267117,.806685716
     435,.91801555366,.98408011975/
      H=HX
      A=AX
      IF(DABS(H).LE.5.77)GO TO 10
   11 THA=0.
      RETURN
   10 IF (DABS(A).LE.1.)GO TO 13
   12 H=A*H
      IF (DABS(H).LE.5.77)GO TO 15
      GO TO 16
   15 A=1./A
```

```
13 SUM=0.
   DO 61 M=1,9,1
   DA=1.+A**2*U(M)**2
   C=-.5*H**2*DA
   TA=DEXP (C)/DA * AA(M)
   SUM=SUM+TA
61 CONTINUE
   TX=A/6.2831853072*SUM
   GO TO 17
16 TX=0.
17 IF (DABS (AX).LE.1.)GO TO 20
14 X=GOFU(HX)
   Y=GOFU(H)
   Z=X*Y
   TX=.5*X+.5*Y-Z-TX
18 IF(AX)21,20,20
21 TX=TX-.5D+00
20 THA=TX
   RETURN
   END
```

2. Bivariate Logistic Subroutines

```
      DOUBLE PRECISION FUNCTION F(X,A)
C     BIVARIATE LOGISTIC (GUMBEL) FUNCTION
      DIMENSION X(1),A(1)
      DOUBLE PRECISION X,A,C,CC,B,BB,D,E,BC,DE
      IF (DABS(A(1)).GE.1.0D+00) A(1)=DSIGN(1.0D+00,A(1))
      C=A(3)*(X(1)+A(2))
      IF (DABS(C).GT.150.0D+00) C=DSIGN(150.0D+00,C)
      CC=DEXP(C)
      B=A(5)*(X(2)+A(4))
      IF (DABS(B).GT.150.0D+00) B=DSIGN(150.0D+00,B)
      BB=DEXP(B)
      D=1./(1.+CC)
      E=1./(1.+BB)
      BC=BB*CC
      DE=D*E
      F=D+E-DE*(1.+A(1)*BC*DE)
      RETURN
      END
```

```
      SUBROUTINE PD(X,A,FXA,P)
C     BIVARIATE LOGISTIC (GUMBEL) PARTIALS
C     A(1)=ALPHA 0,A(2)=ALPHA1,A(3)=BETA1,A(4)=ALPHA2,A(5)=BETA2
      DIMENSION X(1),A(1),P(1)
      DOUBLE PRECISION X,A,P,FXA,WATE,C,CC,B,BB,D,DD,E,EE,BC,DE,R,S,T,Z
      C=A(3)*(X(1)+A(2))
      IF (DABS(C).GT.150.0D+00) C=DSIGN(150.0D+00,C)
      CC=DEXP(C)
      B=A(5)*(X(2)+A(4))
      IF (DABS(B).GT.150.0D+00) B=DSIGN(150.0D+00,B)
      BB=DEXP(B)
      D=1./(1.+CC)
      DD=D*D
      E=1./(1.+BB)
      EE=E*E
      BC=BB*CC
      DE=D*E
      Z=1.+A(1)*BC*DE
      P(1)=-DD*EE*BC
      T=A(1)*P(1)
      R=DD*CC*(E*Z-1.)+T*(1.-D*CC)
      P(2)=A(3)*R
      P(3)=(X(1)+A(2))*R
      S=EE*BB*(D*Z-1.)+T*(1.-E*BB)
      P(4)=A(5)*S
      P(5)=(X(2)+A(4))*S
      RETURN
      END
```

3. Bivariate Burr Subroutines

```
      DOUBLE PRECISION FUNCTION F(X,A)
C     BIVARIATE BURR R (FD) FUNCTION
C     A1=R,A2=B1,A3=B2,A4=P,A5=ALPHA1,A6=BETA1,A7=ALPHA2,A8=BETA2
      DOUBLE PRECISION X,A,B,BB,C,CC,D,DD
      DIMENSION X(1),A(1)
      IF(A(1).GT.A(4)+1.0D+00) A(1)=A(4)+1.0D+00
      IF(A(1).LE.0.0D+00) A(1)=0.0
      B=A(6)*(X(1)+A(5))
      IF (B.GT.0.0D+00) GO TO 17
      BB=0
      GO TO 18
   17 IF (A(2)*DLOG(B).LT.150.0) GO TO 19
      BB=DEXP(150.0D+00)
      GO TO 18
   19 IF (A(2))13,14,15
   13 BB=1./(B**DABS(A(2)))
      GO TO 18
   14 BB=1.0
      GO TO 18
   15 BB=B**A(2)
   18 C=A(8)*(X(2)+A(7))
      IF (C.GT.0.0D+00) GO TO 16
      CC=0
      GO TO 10
   16  IF(A(3)*DLOG(C).LT.150.0) GO TO 20
      CC=DEXP(150.0D+00)
      GO TO 10
   20 IF (A(3))9,11,12
    9 CC=1./(C**DABS(A(3)))
      GO TO 10
   11 CC=1.0
      GO TO 10
   12 CC=C**A(3)
   10 D=(1.0+BB+CC+A(1)*BB*CC)
      IF (A(4)*DLOG(D).LT.150.0) GO TO 21
      DD=DEXP(-150.0D+00)
      GO TO 22
   21 DD=1.0/(D**A(4))
   22 F=1.0D+00-DD
      RETURN
      END
```

```
      SUBROUTINE PD(X,A,FXA,P)
C     BIVARIATE BURR R PARTIALS
C     A1=R,A2=B1,A3=B2,A4=P,A5=ALPHA1,A6=BETA1,A7=ALPHA2,A8=BETA2
      DIMENSION X(1),A(1),P(1)
      DOUBLE PRECISION X,A,FXA,P,WATE,B,BB,C,CC,D,DD,E,EE,G,GG,H,HH,I
      IF(A(1).GT.A(4)+1.0D+00) A(1)=A(4)+1.0D+00
      IF(A(1).LE.0.0D+00) A(1)=0.0
      B=A(6)*(X(1)+A(5))
      IF (B.GT.0.0D+00) GO TO 17
      BB=0
      GO TO 18
   17 IF (A(2)*DLOG(B).LT.150.0) GO TO 30
      BB=DEXP(150.0D+00)
      GO TO 18
   30 IF (A(2))13,14,15
   13 BB=1./(B**DABS(A(2)))
      GO TO 18
   14 BB=1.0
      GO TO 18
   15 BB=B**A(2)
   18 C=A(8)*(X(2)+A(7))
      IF (C.GT.0.0D+00) GO TO 16
      CC=0
      GO TO 10
   16  IF(A(3)*DLOG(C).LT.150.0) GO TO 31
      CC=DEXP(150.0D+00)
      GO TO 10
   31 IF (A(3))9,11,12
    9 CC=1./(C**DABS(A(3)))
      GO TO 10
   11 CC=1.0
      GO TO 10
   12 CC=C**A(3)
   10 D=(1.0+BB+CC+A(1)*BB*CC)
      IF((A(4)+1.0)*DLOG(D).LT.150.0) GO TO 32
      DD=A(4)*DEXP(-150.0D+00)
      GO TO 33
   32 DD=A(4)*((1.0/D)**(A(4)+1.0))
   33 E=DD*BB
      G=1.0+A(1)*BB
      GG=1.0+A(1)*CC
      H=DD*A(2)*GG
      HH=DD*A(3)*G
      P(1)=E*CC
      IF (B.GT.0.0D+00) GO TO 1
      P(2)=0
      GO TO 2
    1 P(2)=E*DLOG(B)*GG
    2 IF (C.GT.0.0D+00) GO TO 3
      P(3)=0
      GO TO 4
    3 P(3)=DD*CC*DLOG(C)*G
    4 IF (A(4)*DLOG(D).LT.150.0) GO TO 34
      P(4)=(DLOG(D))*(DEXP(-150.0D+00))
```

```
      GO TO 35
   34 P(4)=DLOG(D)/(D**A(4))
   35 IF (B.GT.0.0D+00) GO TO 23
      P(5)=0.0
      P(6)=0.0
      GO TO 24
   23 IF ((A(2)-1.0)*DLOG(B).LT.150.0) GO TO 36
      EE=H*DEXP(150.0D+00)
      GO TO 8
   36 IF (A(2)-1.0)5,6,7
    5 EE=H/(B**DABS(A(2)-1.0))
      GO TO 8
    6 EE=H
    7 EE=H*(B**(A(2)-1.0))
    8 P(5)=EE*A(6)
      P(6)=EE*X(1)
   24 IF (C.GT.0.0D+00) GO TO 25
      P(7)=0.0
      P(8)=0.0
      GO TO 26
   25 IF ((A(3)-1.0)*DLOG(C).LT.150.0) GO TO 37
      I=HH*DEXP(150.0D+00)
      GO TO 22
   37 IF (A(3)-1.0)19,20,21
   19 I=HH/(C**DABS(A(3)-1.0))
      GO TO 22
   20 I=HH
      GO TO 22
   21 I=HH*(C**(A(3)-1.0))
   22 P(7)=I*A(8)
      P(8)=I*X(2)
   26 CONTINUE
      RETURN
      END
```

4. Weight Subroutine

```
DOUBLEPRECISIONFUNCTIONWATE(X,FXA)
DOUBLEPRECISIONX,FXA
DIMENSIONX(1)
WATE=DABS(X(3)/((1.0D+00-FXA)*FXA))
RETURN
END
```

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

Bivariate Probit, Logit, and Burrit Analysis

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

Technical Report

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

Frederick C. Durling

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
August 8, 1969	114	32

8a. CONTRACT OR GRANT NO. N00014-68-A-0515 b. PROJECT NO. NR 042-260 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 41 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.

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

13. ABSTRACT

The problem of a mixture of two stimulants in a biological quantal assay is investigated from a mathematical standpoint. The basic assumption is made that the response region does not depend on biological considerations - i.e., given a specified mixture of stimulants $\underline{z}$, the response region is defined by the point $\underline{z}'$ in the p-variate space where there are p stimulants under consideration; instead, the probability functions, themselves, may take on different forms. A general form is proposed and investigated. Three analytic models (one utilizing the bivariate normal distribution, one a bivariate logistic distribution developed by Gumbel (1961), and one a bivariate Burr distribution developed by this author) are employed in this investigation. The investigation includes the analysis of data, under the three analytic models, which had been classified by previous investigators as examples of synergistic action, simple similar action, independent action, and additive action. The residual analyses are included as well as the FORTRAN IV subroutines used in evaluating the functions, the partial derivatives and the weights.

The investigation lends some support to the assumption of a constant response region for a diversity of mixtures of stimulants. The analytic model incorporating the bivariate Burr distribution is recommended for all cases unless the number of parameters to be estimated is a primary concern, in which case the analytic model utilizing the bivariate normal distribution is recommended. The bivariate Burr distribution developed in this paper is found to be more useful in application than that developed by Takahasi (1965).

DD FORM 1 NOV 65 1473

UNCLASSIFIED
Security Classification

THEMIS SIGNAL ANALYSIS STATISTICS RESEARCH PROGRAM

ON THE STEPWISE CONSTRUCTION OF A PREDICTION EQUATION

by

Paul Terrell Pope

Technical Report No. 37
Department of Statistics THEMIS Contract

June 25, 1969

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

DEPARTMENT OF STATISTICS
Southern Methodist University

ON THE STEPWISE CONSTRUCTION

OF A PREDICTION EQUATION

A Dissertation Presented to the Faculty of the Graduate School

of

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

with a

Major in Statistics

by

Paul Terrell Pope
(B.A., Abilene Christian College, 1964;
M.A., The University of Texas, 1966)

June 17, 1969

Pope, Paul Terrell

B.A., Abilene Christian College, 1964
M.A., The University of Texas, 1966

On the Stepwise Construction of a Prediction Equation

Adviser: Associate Professor J. T. Webster

Doctor of Philosophy degree conferred August 22, 1969

Dissertation completed June 17, 1969

In the analysis of data, it is often desired to construct an equation of the form $y = b_0 + b_1x_1 + b_2x_2 + \cdots + b_kx_k$, where y is the response variable and x_1 , x_2 , $\cdots$, x_k are input variables. A familiar method of constructing such an equation is known as the forward selection procedure.

At some stage of the forward selection procedure when p variables are being considered for entry into the equation, the ratio $F_i = SSR_i/MSE$ is computed for each variable being considered. If the ratio $F = \max(F_1 , F_2 , \cdots , F_p)$ is greater than some constant F_0 , the variable x_i associated with this maximum ratio is included in the equation.

The joint distribution of F_1 , F_2 , $\cdots$, F_p is determined under the usual assumptions for the regression model; including normality of the error. Through this the distribution of F is given. Different philosophies exist concerning the choice of the value for F_0 . In this paper the probability integral of F is computed for $p = 3$ and comparisons are made among the critical points F_0 for these different philosophies.

ACKNOWLEDGMENTS

I would like to thank my adviser Dr. J. T. Webster for his helpful advice during the preparation of this paper. I am indebted to all my professors for their many contributions to my graduate studies. I am especially grateful to Dr. Paul Minton for his helping me obtain financial aid through a Research Assistantship and a National Defense Education Act Title IV Fellowship.

I should also thank Mrs. Linda White for her excellent job in typing this manuscript and my wife Gayla for her encouragement during all my studies.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

In the analysis of data, it is often desired to construct a linear prediction equation of the form

$$y = b_0 + b_1x_1 + b_2x_2 + \cdots + b_kx_k$$

where y is the response variable; $x_1 , x_2 , \cdots , x_k$ are the input variables; and $b_0 , b_1 , \cdots , b_k$ are the estimated coefficients. The true model with which we will be working is assumed to be of the form

$$y_i = \sum_{j=0}^{k} x_{ij}\beta_j + \varepsilon_i \quad ; \quad i = 1, 2, \cdots , m$$

where $x_{i0} = 1$ and $\sum_{i=1}^{m} x_{ij} = 0$ for $j = 1, 2, \cdots , k$. That is, $x_{ij} = X_{ij} - \bar{X}_j$ is the sample value adjusted by $\bar{X}_j = \sum_{i=1}^{m} X_{ij}/m$. The error ε_i is assumed to be distributed as $N(0,\sigma^2)$ with $E[\varepsilon_i\varepsilon_j] = 0$ for $i \neq j$. In matrix notation the model is represented as

$$\underline{Y} = \underline{1}\ \beta_0 + X\underline{\beta} + \underline{\varepsilon}$$

where $\underline{Y}$ is a $m \times 1$ vector of observations on the response variable, $\underline{1}$ is a $m \times 1$ vector with each element being one, β_0 is an unknown constant, X is a $m \times k$ matrix of data values for the input variables, $\underline{\beta}$ is a $k \times 1$ vector of unknown constant coefficients, and $\underline{\varepsilon}$ is a $m \times 1$ vector of randon error terms with $\underline{\varepsilon}$ being distributed as $N(\phi,\sigma^2 I)$. The estimates of coefficients are obtained by minimizing the error sum of squares (see Graybill [6]).

The question then arises as to whether an adequate equation could have been constructed using only a subset of the input variables x_1, x_2, $\cdots$, x_k. There is no unique statistical procedure for determining which input variables to include in the equation and all procedures rely somewhat on personal judgment. The purpose of this paper is to consider the choice of the "critical point" in what is known as the forward selection procedure, but first a review is made of other selection procedures which are in use (e.g., Draper and Smith [2]).

One could consider all possible regressions involving any of the input variables. Since each x_i $(i = 1, 2, \cdots, k)$ can either be, or not be, in the equation, there are 2^k possible regressions. The 2^k regressions are then divided into sets of runs which involve p variables, $p = 1, 2, \cdots, k$. Each set of equations involving p variables is ordered according to the value of R^2, the ratio of the sum of squares due to regression to the total sum of squares. After viewing this ordering, a few equations will appear superior to the others and one of these is chosen as the best equation. For a large k, the effort required by this procedure is quite large.

Hocking and Leslie [10] have made an effort to reduce the computation necessary in examining all possible regressions. They have developed a statistic for comparing the $\binom{k}{p}$ possible regressions which involve p variables. This statistic is

$$C_p = \frac{RSS_p}{\hat{\sigma}^2} - (k - 2p) \quad ,$$

where p is the number of variables in the prediction equation, RSS_p is the error sum of squares for that particular p-variate equation, and $\hat{\sigma}^2$ is an estimate of σ^2 . The subset of the $\binom{k}{p}$ equations whose C_p value is small is then considered of interest. They have developed a procedure for identifying this smaller subset after having computed the error sum of squares for only a small fraction of the possible $\binom{k}{p}$ subsets. Also Beale, Kendall and Mann [1] have considered this problem from a similar standpoint.

A procedure known as backward elimination is also an improvement on the "all possible regressions" method from the standpoint of computations involved. First, a regression equation containing all k variables is computed. Then the partial F-test value is calculated for every variable as though it were the last variable to enter the regression equation (i.e., the ratio of the mean square for x_i adjusted for all other variables to the error mean square). The lowest partial F-test value, F_L , is compared with a preselected significance level point F_0 . If $F_L > F_0$, adopt the regression as calculated. If $F_L < F_0$, remove the variable x_L , which gave rise to F_L , from the equation and recompute the equation in the remaining variables. Hamaker [9] has illustrated this procedure in some useful examples.

Another consideration for constructing an equation involving only a subset of the input variables is given by Webster [20]. It is often desired to estimate $E[Y|\underline{X}_0]$, the expected value of the response variable, Y , for a given point in the input variable space, $\underline{X}_0' = [1 , x_1 , \cdots , x_k]$. The usual estimator is the minimum variance unbiased estimator. The question arises as to whether $E[Y|\underline{X}_0]$ could be estimated as well using an estimator involving only a subset of the input variables, a reduced estimator. It is known that in general this reduced estimator will be biased (see Larson and Bancroft [12]). Webster considers two properties which seem of bias terms. Thus if the difference between the MSE matrix for $\underline{b}$, and the MSE matrix for $\hat{\underline{\beta}}$ is positive definite, then $\hat{\underline{\beta}}$ is the better estimator according to the mean square error criterion. Wallace and Toro-Vizcarrondo present a uniformily most powerful testing procedure for this criterion.

As mentioned earlier, we will be primarily concerned with the forward selection procedure. This procedure inserts variables one at a time until the regression equation is satisfactory. Each input variable x_i is tried in the first position and the variable with the largest sum of squares is retained as the first variable in the equation. With the residuals, we continue and enter the variable which, of those remaining, has the largest sum of squares. At each stage a partial F-test is made to see if the variable entering the equation at that stage has taken up a significant amount of the variation. This partial F-test value is the ratio of the mean square for that variable adjusted for all variables in the equation to the error mean square when all variables are put in the equation. When the partial F-test value related to the most recently entered variable becomes nonsignificant the process is terminated. Hamaker [9] has illustrated this procedure through several examples. A slight modification of this procedure is known as the stepwise procedure (see Draper and Smith [2]). With this procedure, at each stage when a new variable enters the equation, a partial F-test is made on each variable already in the equation as though it were the last to enter. If the partial F-test value for any of the variables is nonsignificant, that variable is removed from the equation. A good discussion of this procedure is given by Efroymson [17].

At some stage in most of these selection procedures, a statistic $F_i = SSR_i/MSE$ is compared to a constant F_0 in order to determine whether or not to include the variable x_i in the prediction equation. The philosophy is that if $F_i > F_0$, then the sum of squares SSR_i, corresponding to the variable x_i, represents a large amount of the total variation and x_i should be included in the equation. The constant F_0 is usually chosen as the critical point of the F distribution for some preselected significance level. However, as will be seen later, the statistic F_i does not follow an F distribution, hence the significance level corresponding to the critical point F_0 will not be what it is thought to be. The significance level (i.e., the probability of including a variable x_i in the equation when in reality it should be excluded) chosen depends somewhat on personal judgment. The error of excluding a variable that should be included is generally more serious than the error of including a variable that should be excluded, therefore, the significance level should be chosen so as not to make the power (i.e., the probability of including a necessary variable) too low. Usually the significance level, α, is chosen to be .05, .10 or .25 . Later in this paper the philosophy of choosing the constant F_0 from the F table will be discussed, and a discussion will be given concerning the proper choice of significance levels.

CHAPTER II

FORWARD SELECTION PROCEDURE

In the preceding chapter, the forward selection procedure was described in some detail. At each stage when new variables are being considered for entry into the equation, the sums of squares for each of the x_i's being considered are computed. If any variable x_i enters the equation, it is the variable having the largest sum of squares at that stage. Thus if considering p variables, p sums of squares, SSR_1 , SSR_2 , $\cdots$, SSR_p , would be computed. The statistic F_i used to determine if the new variable x_i should be in the equation is of the form $F_i = SSR_i/MSE$. Let us consider an expression of the model at a certain stage. It is given by

$$\underline{Y} = \underline{1}\beta_0 + X_1\underline{\beta}_1 + \underline{X}_i\beta_i + X_2\underline{\beta}_2 + \underline{\varepsilon}$$

where β_0 is an unknown constant, X_1 is a $m \times (k-p)$ matrix of the data values for those input variables which are already in the equation, $\underline{\beta}_1$ is a $(k-p) \times 1$ vector of unknown constant coefficients, $\underline{X}_i$ is a $m \times 1$ vector of the data values for the variable x_i under immediate consideration, β_i is an unknown constant, X_2 is a $m \times (p-1)$ matrix of the data values for those variables, except x_i , which are not in the equation, and $\underline{\beta}_2$ is a $(p-1) \times 1$ vector of unknown constant coefficients.

At the stage when the variable x_i is considered for entry, the sum of squares due to regression can be expressed as

$$SSR = \underline{Y}'X_I(X_I'X_I)^{-1}X_I'\underline{Y}$$

where $X_I = [X_1\ ,\ \underline{X}_i]$ is a $m \times (k-p+1)$ matrix. Let K be a $(k-p+1) \times (k-p+1)$ non-singular matrix and let $Z = X_I K$. Then

$$SSR = \underline{Y}'Z(Z'Z)^{-1}Z'\underline{Y} \quad .$$

Represent Z by $[Z_1\ ,\ \underline{Z}_i]$, where $\underline{Z}_i$ is a $m \times 1$ vector and Z_1 is a $m \times (k-p)$ matrix.

Let $\underline{Z}_i = \underline{X}_i - X_1\underline{k}_i$, where $\underline{k}_i$ is such that $X_1'\underline{Z}_i = 0$. Then

$$X_1'\underline{X}_i - X_1'X_1\underline{k}_i = 0 \quad \text{and} \quad \underline{k}_i = (X_1'X_1)^{-1}X_1'\underline{X}_i$$

Note that $Z_1 = X_1$.

The sum of squares due to regression can now be written as

$$SSR = \underline{Y}'Z_1(Z_1'Z_1)^{-1}Z_1'\underline{Y} + \underline{Y}'\underline{Z}_i(\underline{Z}_i'\underline{Z}_i)^{-1}\underline{Z}_i'\underline{Y}$$

where the second term on the right represents the sum of squares due to the variable x_i adjusted for all other variables already in the equation.

Write

$$SSR_i = \underline{Y}'\underline{Z}_i(\underline{Z}_i'\underline{Z}_i)^{-1}\underline{Z}_i'\underline{Y} = \underline{Y}'\underline{a}_i\underline{a}_i'\underline{Y}$$

where

$$\underline{a}_i = \frac{\underline{Z}_i}{\sqrt{\underline{Z}_i'\underline{Z}_i}} \quad .$$

Let $W_i = \underline{Y}'\underline{a}_i$. W_i will follow a normal distribution since it is a linear combination of the elements of $\underline{Y}$, which are normally distributed.

$$E[W_i] = E[\underline{Y}'\underline{a}_i] = E[(\beta_0\underline{1}' + \underline{\beta}_1'X_1' + \beta_i\underline{X}_i' + \underline{\beta}_2'X_2' + \underline{\varepsilon}')\underline{a}_i]$$

$$= \beta_i\underline{X}_i'\underline{a}_i + \underline{\beta}_2'X_2'\underline{a}_i$$

$$\mathrm{Var}[W_i] = E[\underline{Y}'\underline{a}_i\underline{a}_i'\underline{Y}] - \left\{E[\underline{Y}'\underline{a}_i]\right\}^2$$

which is the matrix associated with the covariance matrix of the W_i's that was given by expression (1).

The matrix $(X'X)^{-1}$ will be computed in finding SSE , so it will be known from the start. At each stage, the elements of $(X'X)^{-1}$ can be permuted to correspond to the situation at that particular stage. Then the matrix B_{22} can be inverted to give the covariance matrix of the W_i's .

Either of the above methods for computing the covariance matrix requires a matrix inversion. If $p > k/2$, invert $(X_1'X_1)$ and use expression (1). If $p < k/2$, invert B_{22} .

Now let us consider the statistic $F_i = SSR_i/MSE$ used in determining whether or not to include the variable x_i in the equation. Note that $SSR_i = W_i^2$. Recall that the expected value of W_i is a linear combination of β_i and the elements of $\underline{\beta}_2$. When considering whether or not x_i should enter the equation, we are actually considering the hypothesis that this linear combination is equal to zero. If F_i is large and x_i is included in the equation, it is concluded that $\beta_i \neq 0$. Intuitively, this seems to be a logical conclusion and in looking at the properties of $SSR_i = W_i^2$ it is seen that the coefficient of β_i in the linear combination of $E[W_i]$ is larger than the coefficient of β_i in any other linear combination, $E[W_j]$, associated with the other variables being considered for entry.

When considering the variable x_i , we have $E[W_i] = \beta_i \underline{X}_i' \underline{a}_i + \underline{\beta}_2' X_2' \underline{a}_i$, and when considering the variable x_j , we have $E[W_j] = \beta_j \underline{X}_j' \underline{a}_j + \tilde{\underline{\beta}}_2' \tilde{X}_2' \underline{a}_j$. It will be shown that

$$|\underline{X}_i' \underline{a}_i| > |\underline{X}_i' \underline{a}_j| \quad , \text{ for } i \neq j \quad .$$

$$\underline{X}_i' \underline{a}_i = \frac{\underline{X}_i' (\underline{X}_i - X_1 (X_1'X_1)^{-1} X_1' \underline{X}_i)}{\sqrt{\underline{X}_i' \underline{X}_i - \underline{X}_i' X_1 (X_1'X_1)^{-1} X_1' \underline{X}_i}} = \sqrt{\underline{X}_i' \underline{X}_i - \underline{X}_i' X_1 (X_1'X_1)^{-1} X_1' \underline{X}_i}$$

$$\underline{X}_i'\underline{a}_j = \frac{\underline{X}_i'(\underline{X}_j - X_1(X_1'X_1)^{-1}X_1'\underline{X}_j)}{\sqrt{\underline{X}_j'\underline{X}_j - \underline{X}_j'X_1(X_1'X_1)^{-1}X_1'\underline{X}_j}}$$

Let $X_1(X_1'X_1)^{-1}X_1' = R$. This is a symmetric idempotent matrix. Thus if $|\underline{X}_i'\underline{a}_i| > |\underline{X}_i'\underline{a}_j|$,

$$+\sqrt{\underline{X}_i'\underline{X}_i - \underline{X}_i'R\underline{X}_i} > \left| \frac{\underline{X}_i'\underline{X}_j - \underline{X}_i'R\underline{X}_j}{\sqrt{\underline{X}_j'\underline{X}_j - \underline{X}_j'R\underline{X}_j}} \right|$$

Squaring both sides and multiplying by the denominator on the right, we obtain

$$\underline{X}_i'\underline{X}_i\underline{X}_j'\underline{X}_j - \underline{X}_i'\underline{X}_i\underline{X}_j'R\underline{X}_j - \underline{X}_i'R\underline{X}_i\underline{X}_j'\underline{X}_j + \underline{X}_i'R\underline{X}_i\underline{X}_j'R\underline{X}_j > [\underline{X}_i'\underline{X}_j - \underline{X}_i'R\underline{X}_j]^2$$

$$\underline{X}_i'(I - R)\underline{X}_i\underline{X}_j'(I - R)\underline{X}_j > [\underline{X}_i'(I - R)\underline{X}_j]^2$$

Since $(I - R)$ is symmetric idempotent, this can be written as

$$\underline{X}_i'(I - R)(I - R)\underline{X}_i\underline{X}_j'(I - R)(I - R)\underline{X}_j > [\underline{X}_i'(I - R)(I - R)\underline{X}_j]^2$$

Let $\underline{X}_i'(I - R) = \underline{b}'$ and $\underline{X}_j'(I - R) = \underline{c}'$. Then we have

$$\underline{b}'\underline{b}\ \underline{c}'\underline{c} > (b'c)^2 \quad .$$

This is Schwarz's inequality, therefore $|\underline{X}_i'\underline{a}_i|$ is greater than $|\underline{X}_i'\underline{a}_j|$.

Under the hypothesis that $E[W_i] = 0$, W_i/σ is distributed as $N(0,1)$ and $W_i^2/\sigma^2 = SSR_i/\sigma^2$ is distributed as $\chi^2(1)$.

Now consider the MSE , which is the denominator of the statistic F_i . It is a practice at times to use the MSE* computed only from those variables in the equation. This has one real advantage in that the matrices to be inverted in computing MSE will be of smaller order than if all variables were used. This becomes increasingly important for a large number of degrees of freedom and noncentrality parameters $\lambda_1 = \frac{1}{2\sigma^2}(E[W_i])^2$ and λ_2 .

On the other hand, if the SSE is computed by placing all variables in the equation, we have

$$SSE = \underline{Y}'\left[I - \frac{1}{m}[1] - X(X'X)^{-1}X'\right]\underline{Y}$$

$\left[I\sigma^2\right]\left[I - \frac{1}{m}[1] - X(X'X)^{-1}X'\right]/\sigma^2$ is idempodent of rank $m - k - 1$. Therefore SSE/σ^2 is distributed as $\chi'^2(m - k - 1, \lambda)$, where

$$\begin{aligned}\lambda &= \frac{1}{2\sigma^2}\left(\beta_0\underline{1}' + \underline{\beta}'X'\right)\left(I - \frac{1}{m}[1] - X(X'X)^{-1}X'\right)\left(\beta_0\underline{1} + X\underline{\beta}\right) \\ &= 0\end{aligned}$$

Therefore if the SSE is computed by placing all variables in the equation, the marginal distribution of any one of the statistics F_i is the noncentral F distribution with $(1, m - k - 1)$ degrees of freedom and noncentrality parameter λ_1 .

Now suppose the statistic F_i corresponding to the variable x_i is found to be the largest such ratio among the variables being considered at this particular time. It is in general then compared to a value F_0 , a critical point from the central F table. If $F_i > F_0$, the variable x_i is included in the equation. But note that if the SSE* is computed using only those variables already in the equation, the denominator of F_i will tend to be larger than if SSE is used and new variables which should be included in the equation may be excluded. An illustration of what might happen is given by the following example.

Suppose we have three input variables x_1 , x_2 , x_3 and, for the sake of simplicity, let $X'X$ be a diagonal matrix. The situation can then be illustrated by the following numerical example:

AOV

Source	d.f.	S.S.	M.S.
x_1	1	1525	1525
x_2	1	1550	1550
x_3	1	1575	1575
Error	8	3400	425
Total	11	8050	

The sum of squares for error computed by putting all variables in the equation is 3400 with 8 degrees of freedom. Now in constructing the prediction equation the variable x_3 is entered first since it has the largest sum of squares. The statistic $F_3 = SSR_3/MSE = 3.71$. This is greater than $F_0 = F(1, 8; .10) = 3.46$, so the variable x_3 would be included in the equation. However, if SSE* had been computed for only x_3 in the equation, we would get, taking advantage of X'X being diagonal,

$$SSE^* = 1525 + 1550 + 3400 = 6475$$

and

$$F_3 = \frac{SSR_3}{MSE^*} = \frac{1575}{\frac{6475}{10}} = 2.43 \quad .$$

This is less than $F(1, 10; .10)$, which says that the variable x_3 should be excluded. Thus it is important to obtain a good estimate of error before considering whether to include or exclude variables.

The statistic $F = \max_i\{F_i\}$ is treated as though it follows a univariate F distribution with $(1, m - k - 1)$ degrees of freedom. However this is not true. In order to find the constant F_0 which would correspond to a given significance level, it is necessary to determine the distribution of $F = \max_i\{F_i\}$, where the distribution of the F_i's is what will be referred to as a p-variate F distribution with $(1, m - k - 1)$ degrees of freedom. This will be discussed in the next chapter.

CHAPTER III

MULTIVARIATE t^2 DISTRIBUTION

As mentioned in the preceding chapter, it is desired to compute the probability integral of $F = \max(F_1, F_2, \cdots, F_p)$, where each F_i is of the form

$$F_i = \frac{W_i^2}{\sigma^2} \cdot \frac{n\sigma^2}{SSE} .$$

Recall that W_i is distributed $N(0,\sigma^2)$ and SSE/σ^2 is distributed as $\chi^2(n)$, and W_i and SSE are independent. The joint density of $F_1, F_2, \cdots, F_p$ will be derived as a step toward the computation of the desired probability integral.

The $W_i (i = 1, 2, \cdots, p)$ follow a p-variate normal distribution with mean zero and covariance $\Sigma = \{\sigma_{ij}\}$, where $\sigma_{ii} = \sigma^2$ and $\sigma_{ij} = \underline{a}_i'\underline{a}_j\sigma^2$. Let the matrix A be the inverse of the correlation matrix of the W_i's, that is, $A = \{a_{ij}\} = \{\underline{a}_i'\underline{a}_j\}^{-1}$. Then the joint density function of $W_1, W_2, \cdots, W_p$ is given by

$$\frac{|A|^{1/2}}{(2\pi)^{p/2}\sigma^p} e^{-\frac{1}{2\sigma^2}(\underline{W}'A\underline{W})}$$

where $\underline{W}' = [W_1, W_2, \cdots, W_p]$.

Let $s^2 = MSE = SSE/n$. Then ns^2/σ^2 is distributed as chi-square with n degrees of freedom and independent of W_i for $i = 1, 2, \cdots, p$. The density function of s is given by

$$f(s) = \frac{n^{n/2} s^{n-1}}{2^{n/2-1}\Gamma\left(\frac{n}{2}\right)\sigma^n} e^{-ns^2/\sigma^2} \quad , \quad 0 < s < \infty$$

The joint density function of W_1 , W_2 , $\cdots$, W_p , s is given by

$$\frac{2|A|^{1/2}}{(n\pi)^{p/2}\Gamma\left(\frac{n}{2}\right)} \left(\frac{n}{2\sigma^2}\right)^{\frac{1}{2}(n+p)} s^{n-1}\left\{\exp - \frac{1}{2\sigma^2}\left(\sum_{i,j}^{p} a_{ij}W_iW_j + ns^2\right)\right\} \tag{1}$$

Let $t_i = W_i/s$, for i = 1, 2, $\cdots$, p . Making this transformation in (1), we obtain

$$\frac{2|A|^{1/2}}{(2\pi)^{p/2}\Gamma\left(\frac{n}{2}\right)} \left(\frac{n}{2\sigma^2}\right)^{\frac{1}{2}(n+p)} s^{n+p-1}\left\{\exp - \frac{ns^2}{2\sigma^2}\left(1 + \frac{1}{n}\sum_{i,j}^{p} a_{ij}t_it_j\right)\right\}$$

Integrating out s , we obtain the joint density function of t_1 , t_2 , $\cdots$, t_p

$$\frac{|A|^{1/2}\Gamma\left[\frac{1}{2}(n+p)\right]}{(n\pi)^{p/2}\Gamma\left(\frac{n}{2}\right)} \left\{1 + \frac{1}{n}\sum_{i,j}^{p} a_{ij}t_it_j\right\}^{-\frac{1}{2}(n+p)} \tag{2}$$

This is the multivariate analogue of Student's t-distribution with n degrees of freedom. The expression (2) was derived by Dunnett and Sobel [4].

Note that $F_i = W_i^2/s^2 = t_i^2$. So make the transformation $F_i = t_i^2$ (i = 1, 2, $\cdots$, p) in expression (2). By this transformation, each of the 2^p disjoint regions in the p-dimensional space of the t_i's are mapped one-to-one onto the region $0 < F_i < \infty$ (i = 1, 2, $\cdots$, p) . The Jacobian of this transformation is $\left(\frac{1}{2}\right)^p \prod_{i=1}^{p} (F_i)^{1/2}$ for all 2^p regions. Thus the joint density function of F_1 , F_2 , $\cdots$, F_p is given by

$$\frac{|A|^{1/2}\Gamma\left[\frac{1}{2}(n+p)\right]}{2^p (n\pi)^{p/2}\Gamma\left(\frac{n}{2}\right)} \left(\prod_{i=1}^{p} F_i\right)^{-\frac{1}{2}} \left\{ \sum_{k=1}^{2^p} \left[1 + \frac{1}{n}\left(\sum_{i}^{p} a_{ii}F_i \right.\right.\right.$$

$$\left.\left.\left. + \sum_{\substack{i \\ i \neq j}}^{p} \sum_{i}^{p} \pm a_{ij} \sqrt{F_i} \sqrt{F_j} \right)\right]_k^{-\frac{1}{2}(n+p)} \right\} \qquad (3)$$

This expression, for $p = 3$, is given in detail as expression (1) in Chapter IV. The notation $\sum_{k=1}^{2^p}$ indicates the sum of 2^p terms of this form in which the signs of the a_{ij} will differ depending on the region from which the t_i, t_j are mapped.

The expression (3) is the density function of the p-variate t^2 distribution with n degrees of freedom. This is also the density function of the p-variate F distribution with (1, n) degrees of freedom.

The cumulative distribution of $F = \max(F_1, F_2, \cdots, F_p)$ is given by

$$\int_0^c f(F)\,dF = \int_0^c \cdots \int_0^c f(F_1, F_2, \cdots, F_p) \prod_{i=1}^{p} dF_i$$

This integral has been computed for some special cases in the literature. For $p = 1$, the expression (3) is just the univariate F density with (1, n) degrees of freedom and its probability integral is well tabulated (e.g., Owen [16]).

One special case of this problem is the case when the numerators of all the F_i's are independent. For this case, Nair [15] has tabulated the upper 5 percent and 1 percent points of the distribution of $F = \max(F_1, F_2, \cdots, F_p)$ for $p = 1(1)10$, and $n = 10, 12, 15, 20, 30, 60,$ and ∞. Also for this special case of independent numerators, Gupta and Sobel [8] have derived exact expressions, as well as approximations and bounds, for the probability integral, percentage points, and moments of the statistic $F_L = \min(F_1, F_2, \cdots, F_p)$.

The case in which the numerators of the F_i's are equally correlated has been considered by Krishnaiah [11]. He has tabulated the upper 10 percent, 5 percent, 2.5 percent, and 1 percent points of the distribution of $F = \max(F_1, F_2, \cdots, F_p)$ for $p = 1(1)10$, $n = 5(1)35$, and correlations $\rho = .1(.1).9$.

This paper will tabulate the probability integral of $F = \max(F_1, F_2, F_3)$ with no restrictions made on the correlations. The results of these computations will be presented in the next chapter.

CHAPTER IV

COMPUTATION OF THE PROBABILITY INTEGRAL OF THE TRIVARIATE t^2 DISTRIBUTION

As discussed in the preceding chapter, it is desired to compute the probability integral of $F = \max_i \{F_i\}$, where the F_i follow a p-variate F distribution with (1, n) degrees of freedom. The cumulative distribution of F is given by

$$\int_0^c f(F)dF = \int_0^c \cdots \int_0^c f(F_1 , F_2 , \cdots , F_p) \prod_{i=1}^{p} dF_i$$

where $f(F_1 , F_2 , \cdots , F_p)$ is the density function of the p-variate F distribution with (1, n) degrees of freedom. An expression for this density function was given in the preceding chapter.

The computation of the integral $\int_0^c f(F)dF$ will be presented for the case of $p = 3$. The joint density of F_1, F_2, F_3 is given by

$$\frac{|A|^{1/2}\Gamma\left[\frac{1}{2}(n + 3)\right]}{4(n\pi)^{1/2}\Gamma\left(\frac{n}{2}\right)} (F_1F_2F_3)^{-\frac{1}{2}}\left\{\left[1 + \frac{1}{n}(a_{11}F_1 + a_{22}F_2 + a_{33}F_3 + 2a_{12}\sqrt{F_1}\sqrt{F_2} + 2a_{13}\sqrt{F_1}\sqrt{F_3} + 2a_{23}\sqrt{F_2}\sqrt{F_3})\right]^{-\frac{1}{2}(n+3)}\right.$$

$$+ \left[1 + \frac{1}{n}(a_{11}F_1 + a_{22}F_2 + a_{33}F_3 - 2a_{12}\sqrt{F_1}\sqrt{F_2} - 2a_{13}\sqrt{F_1}\sqrt{F_3} + 2a_{23}\sqrt{F_2}\sqrt{F_3})\right]^{-\frac{1}{2}(n+3)}$$

$$+\left[1 + \frac{1}{n}(a_{11}F_1 + a_{22}F_2 + a_{33}F_3 - 2a_{12}\sqrt{F_1}\sqrt{F_2} + 2a_{13}\sqrt{F_1}\sqrt{F_3} - 2a_{23}\sqrt{F_2}\sqrt{F_3})\right]^{-\frac{1}{2}(n+3)}$$

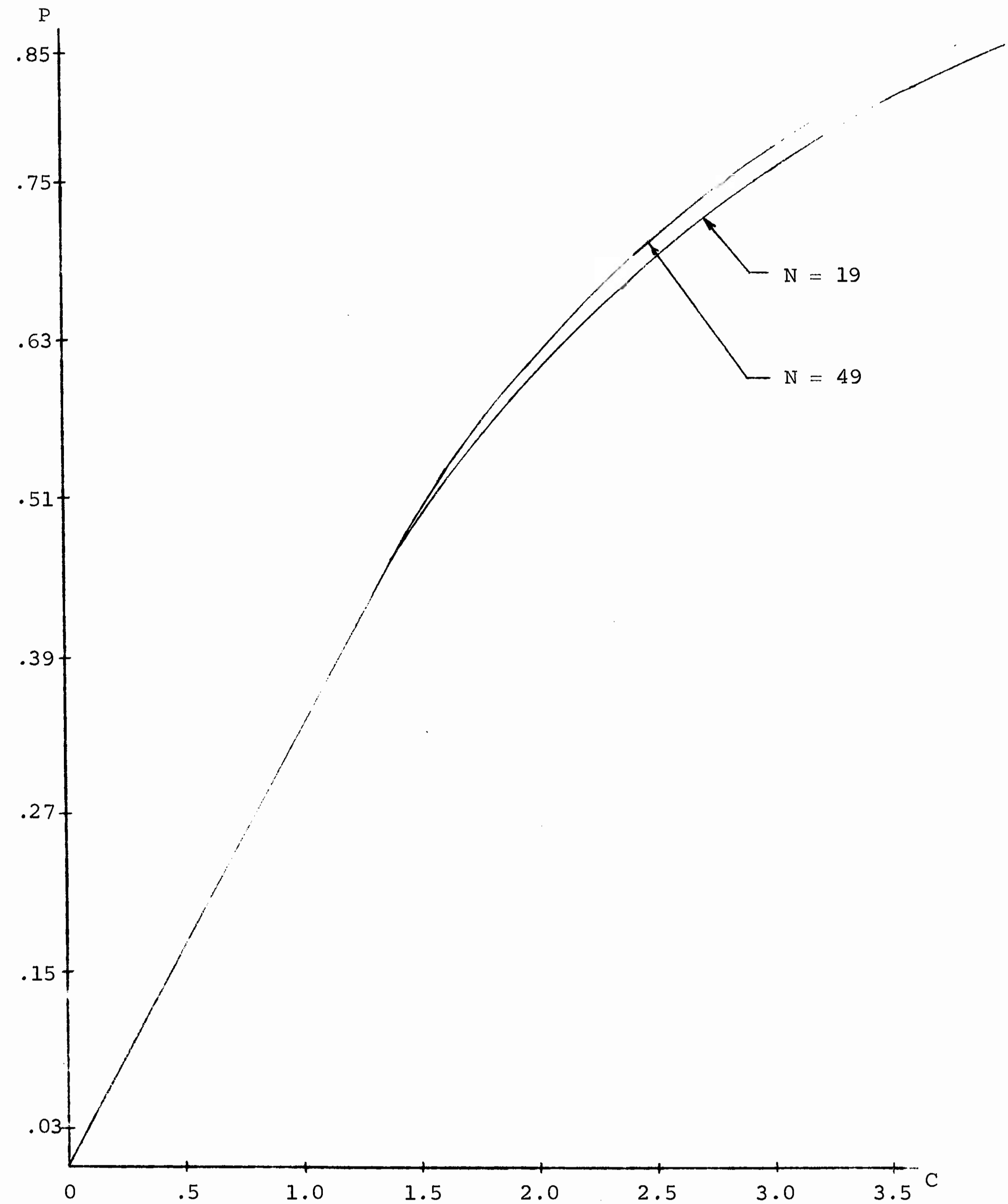
$$+\left[1 + \frac{1}{n}(a_{11}F_1 + a_{22}F_2 + a_{33}F_3 + 2a_{12}\sqrt{F_1}\sqrt{F_2} - 2a_{13}\sqrt{F_1}\sqrt{F_3} - 2a_{23}\sqrt{F_2}\sqrt{F_3})\right]^{-\frac{1}{2}(n+3)}\Bigg\}$$

This computation was performed by numerically integrating the expression for $f(F_1, F_2, F_3)$. The numerical integration technique used was a generalization of the trapezoidal rule. McCormick and Salvadori [14] discuss using the trapezoidal rule to perform multiple numerical integration. The computer program was written in Fortran IV and the computations performed on an IBM System 360 Model 44 computer. More details concerning the computer program will be given in the Appendix. The computations were checked for accuracy by comparing the results to those of Krishnaiah [11] for the case of equicorrelated variables.

The value of $\int_0^c f(F)dF$ will be graphed for the region $0 \leq c \leq 11.5$, for n = 19, 25, 31, and 49 and various values of $\rho = (\rho_{12}, \rho_{13}, \rho_{23})$. In order to interpolate for critical points associated with values of n between those given here, the interpolation scheme given by Laubscher [13] may be used. It appears that the value of the probability integral for $\rho = (\rho_{12}, \rho_{13}, \rho_{23})$ is well approximated by the equally correlated case with correlation $\rho = \Sigma\rho_{ij}/3$, hence permitting a wider use of Krishnaiah's tables. The value of $\int_0^c f(F)dF$ was not investigated for values of p greater than 3 because of the large amount of computation time required to do so. The following graphs present the results of the computations with the value $P = \int_0^c f(F)dF$ plotted versus values of c. The value of $P = \int_0^c f(F)dF$ is graphed in two parts; the first being for values of c such that $0 < P \leq .85$, the second being for values of c such that $.85 \leq P < 1$.

FIGURE 1a: DISTRIBUTION FUNCTION OF F

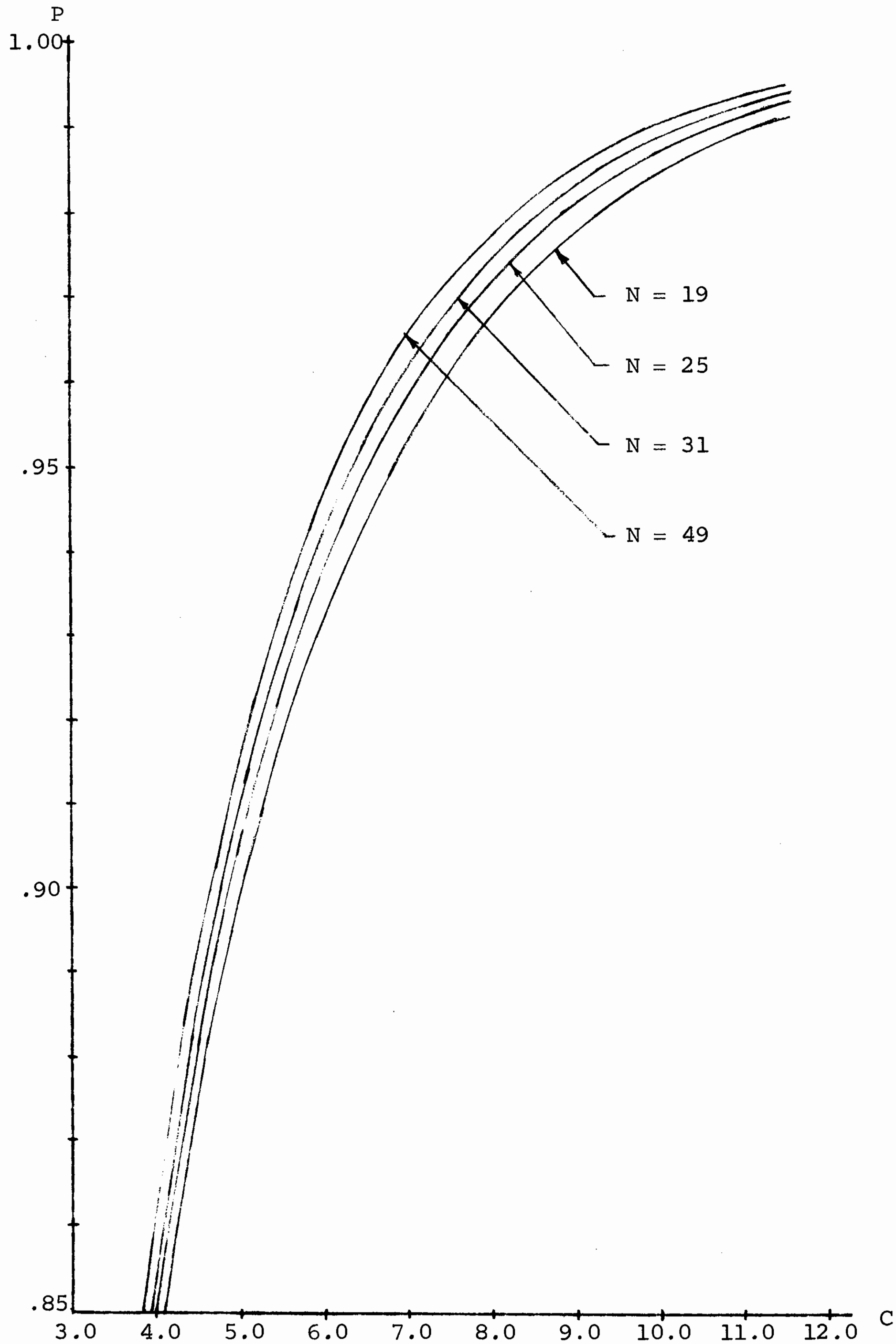
FOR $\rho = (.2, .2, .5)$



P = Prob. [F ≤ C]

FIGURE 1b: DISTRIBUTION FUNCTION OF F

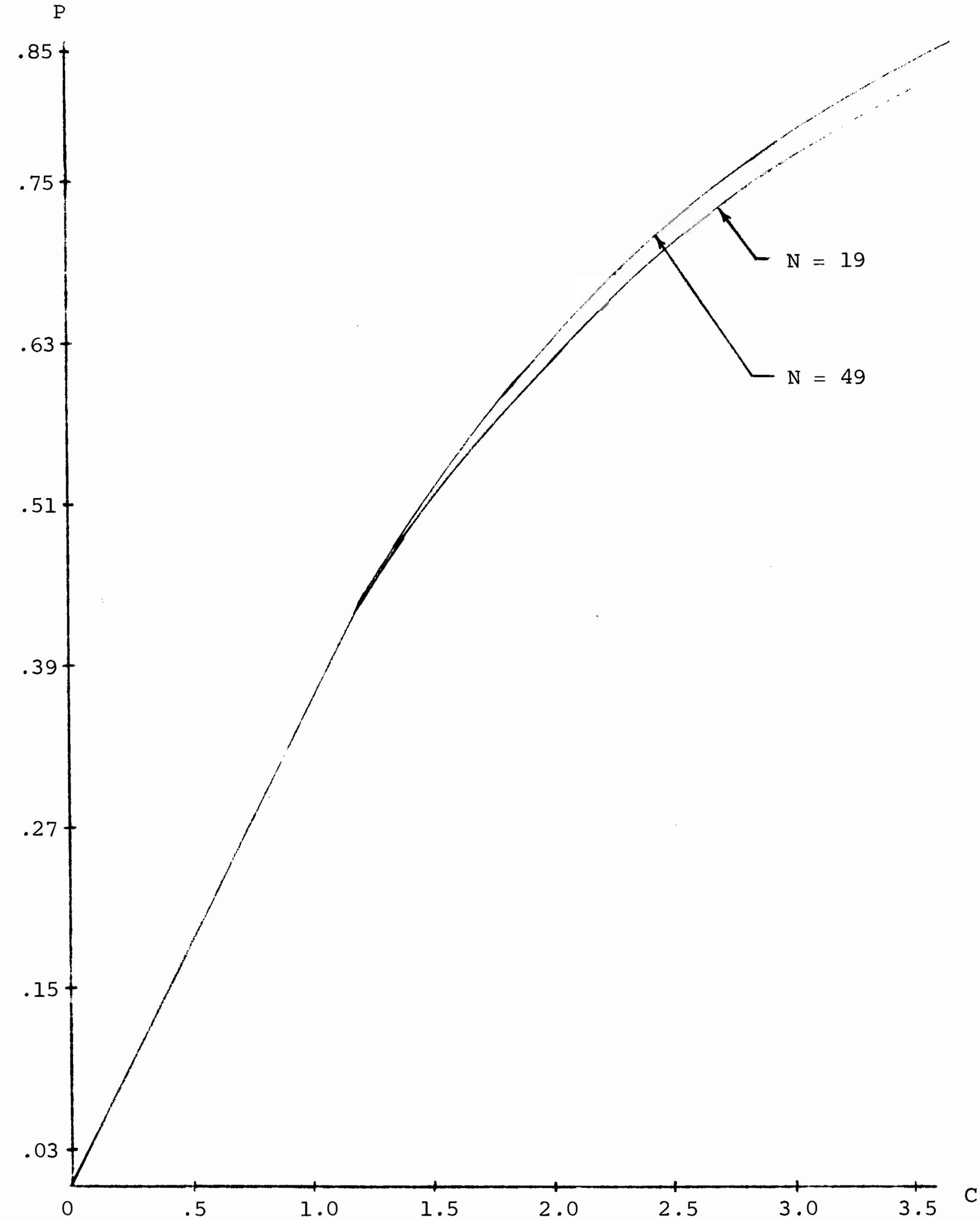
FOR ρ = (.2, .2, .5)



P = Prob.[F ≤ C]

FIGURE 2a: DISTRIBUTION FUNCTION OF F

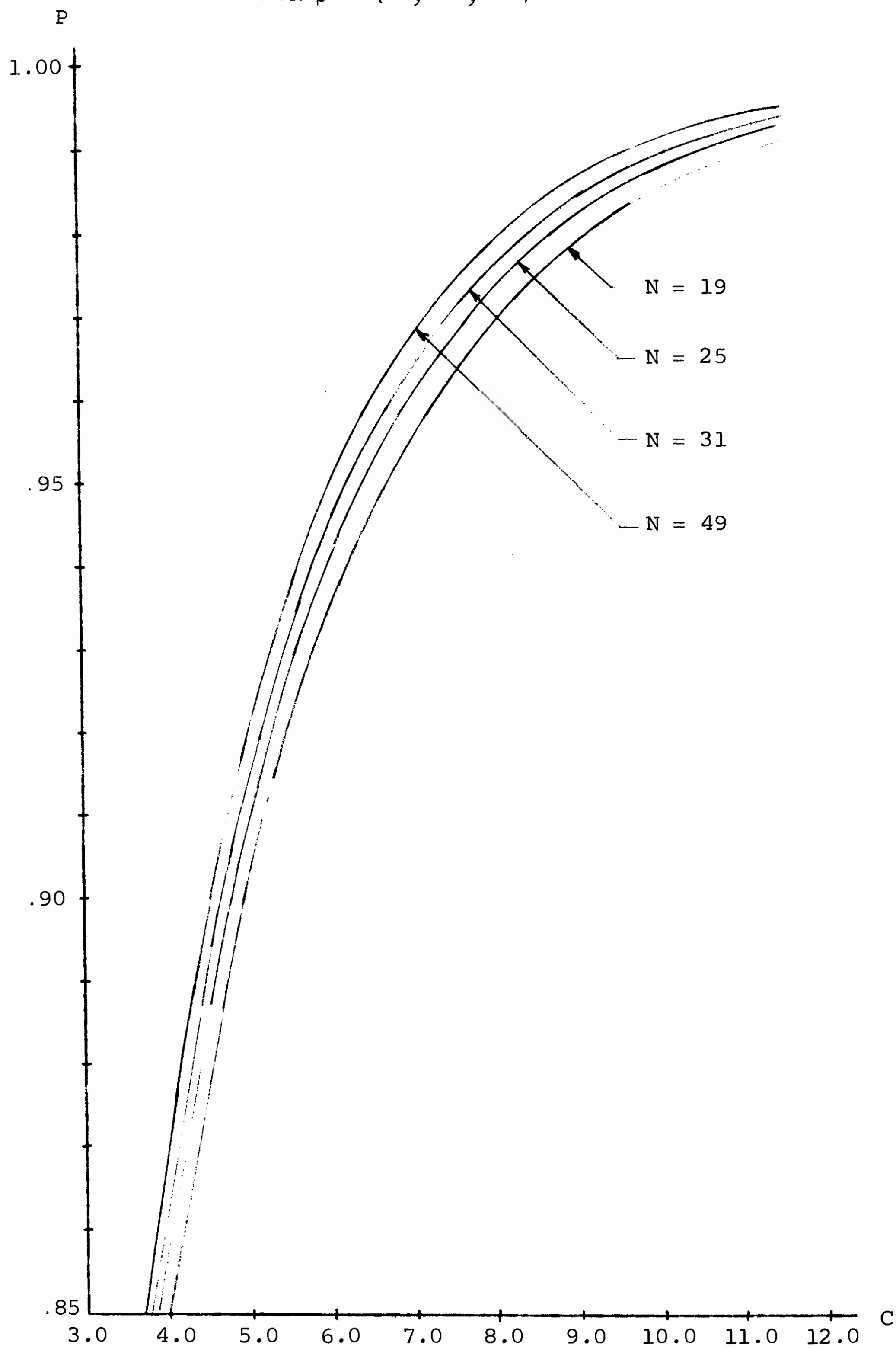
FOR $\rho = (.2, .5, .5)$



P = Prob. [F ≤ C]

FIGURE 2b: DISTRIBUTION FUNCTION OF F

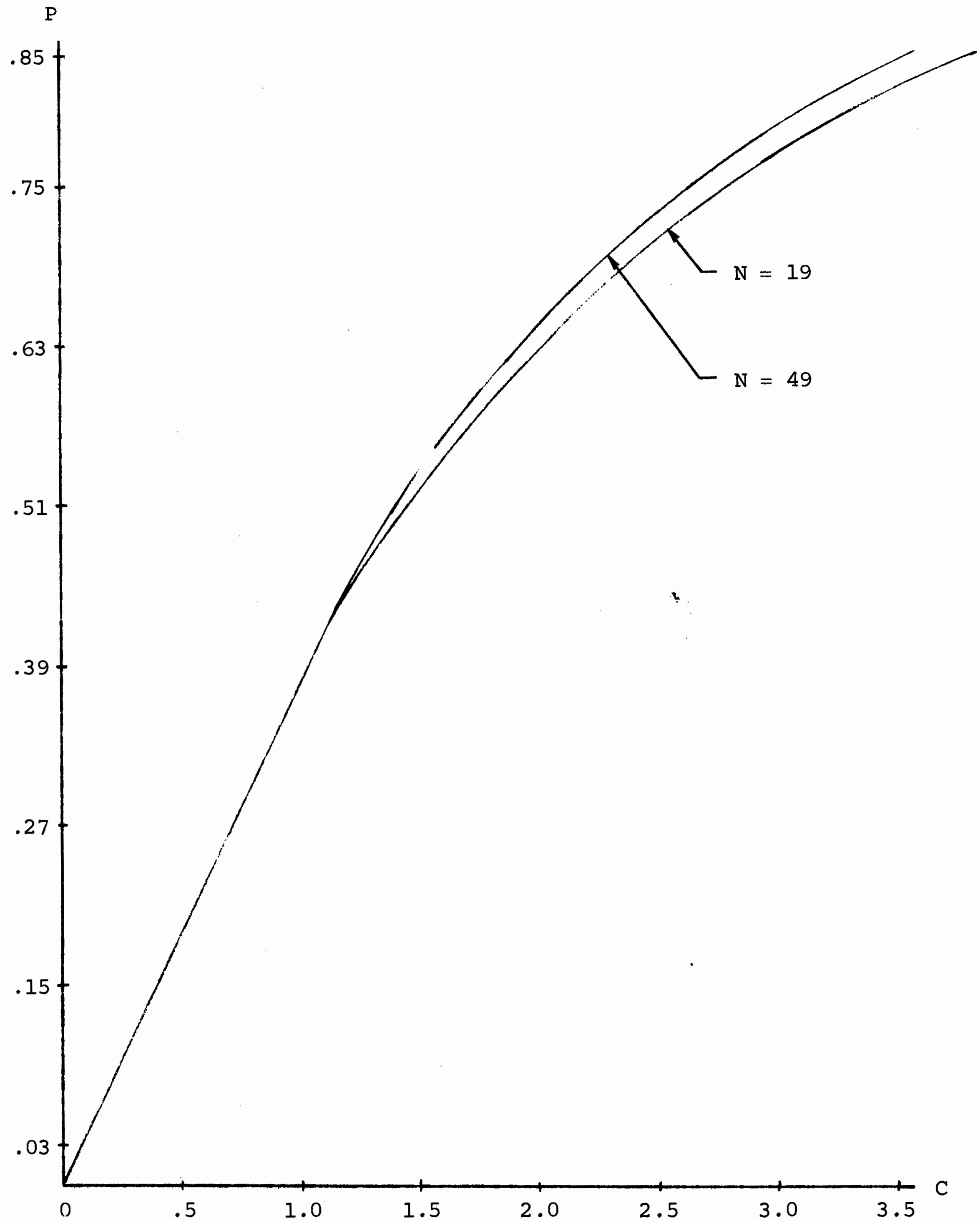
FOR ρ = (.2, .5, .5)



P = Prob.[F ≤ C]

FIGURE 3a: DISTRIBUTION FUNCTION OF F

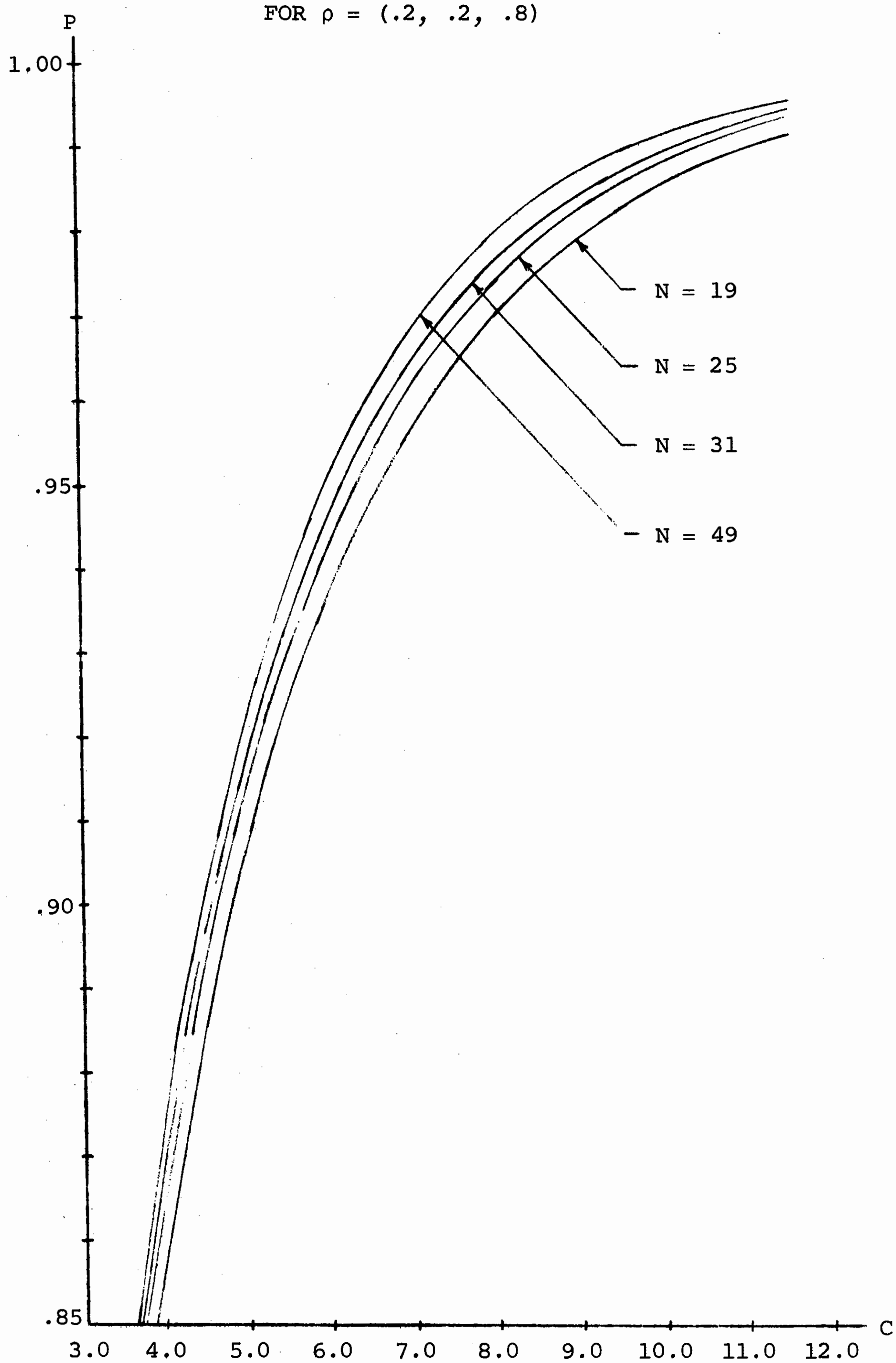
FOR ρ = (.2, .2, .8)



P = Prob.[F ≤ C]

FIGURE 3b: DISTRIBUTION FUNCTION OF F

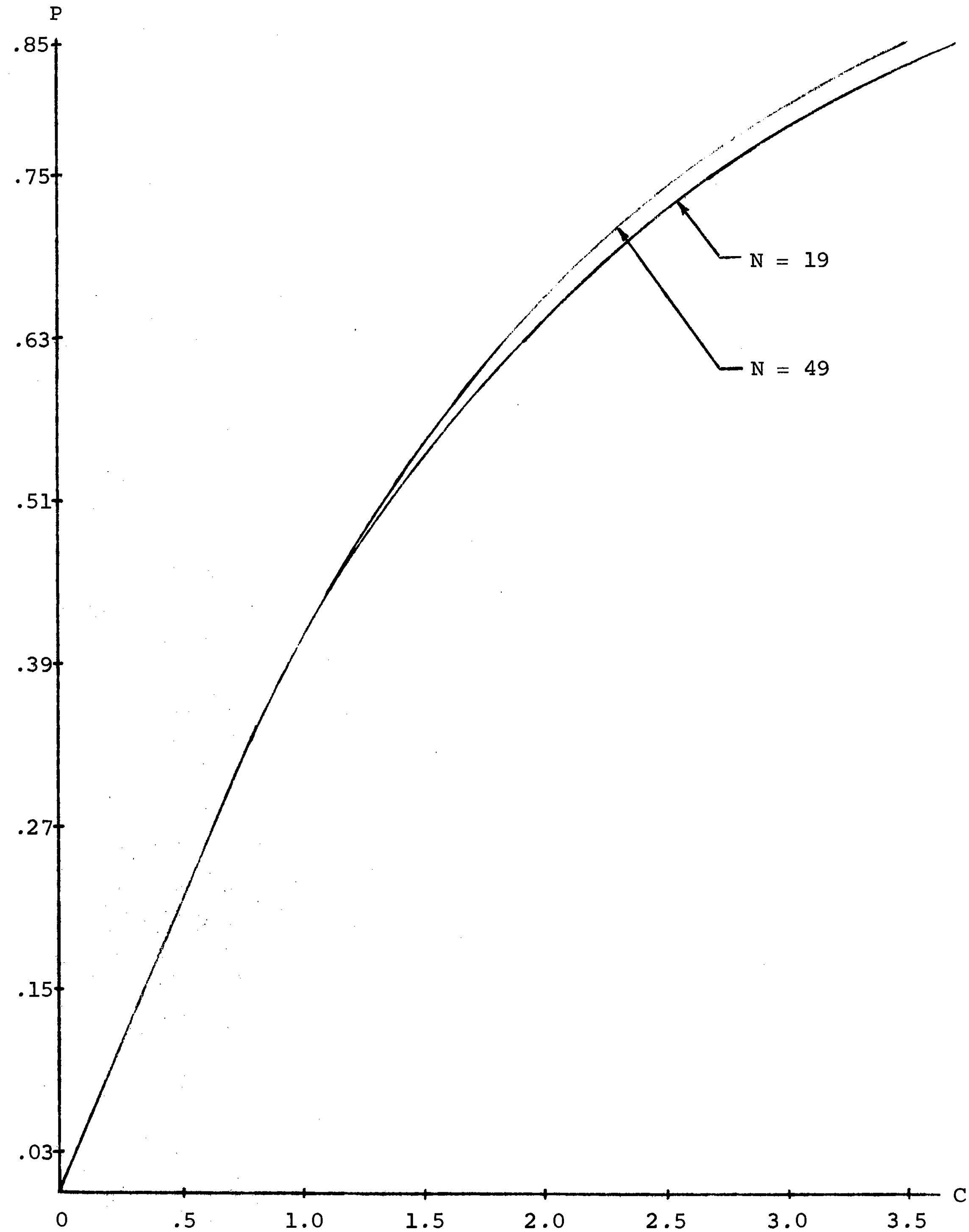
FOR ρ = (.2, .2, .8)



P = Prob.[F ≤ C]

FIGURE 4a: DISTRIBUTION FUNCTION OF F

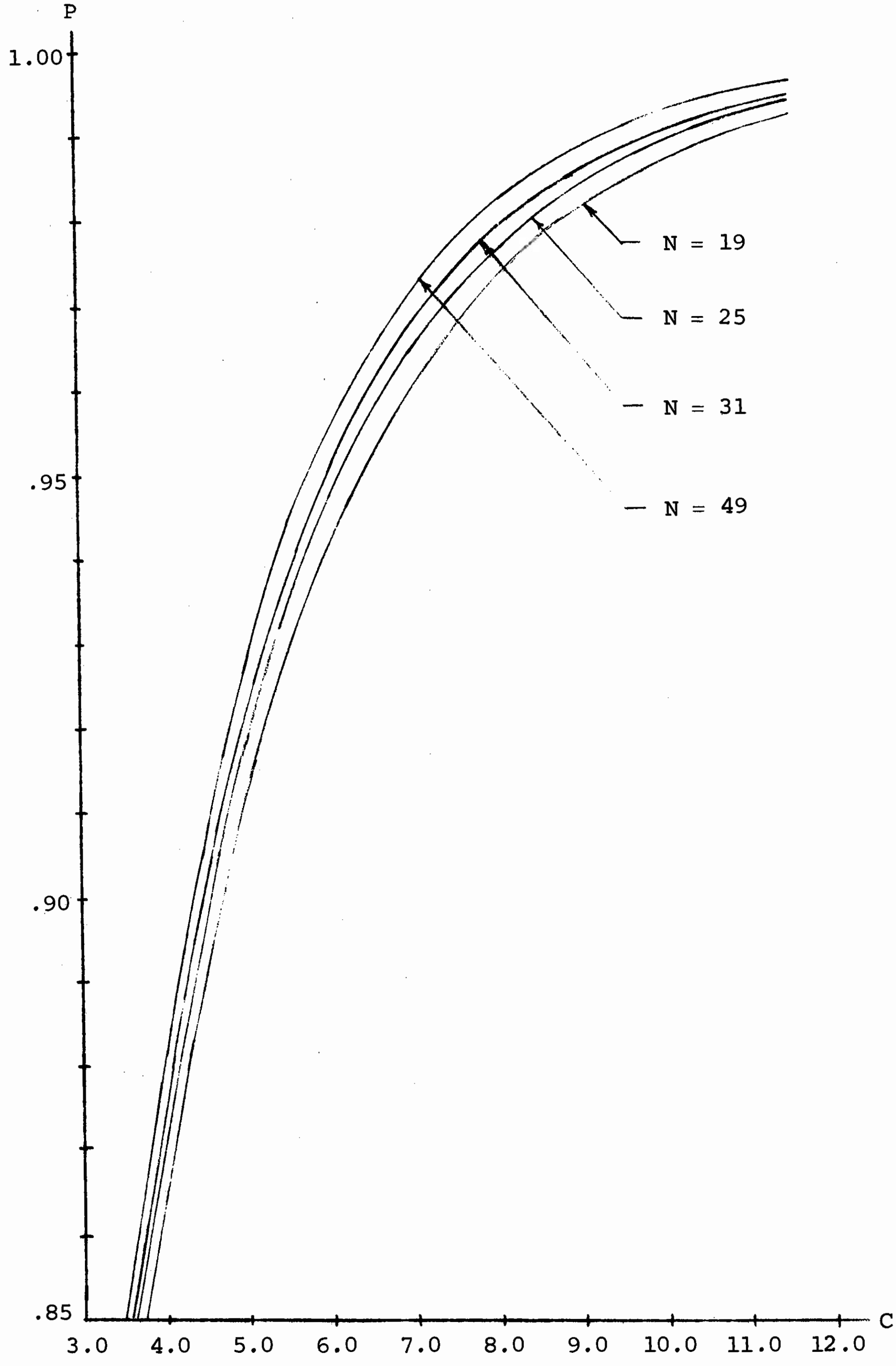
FOR $\rho = (.2, .5, .8)$



$P = \text{Prob.}[F \leq C]$

FIGURE 4b: DISTRIBUTION FUNCTION OF F

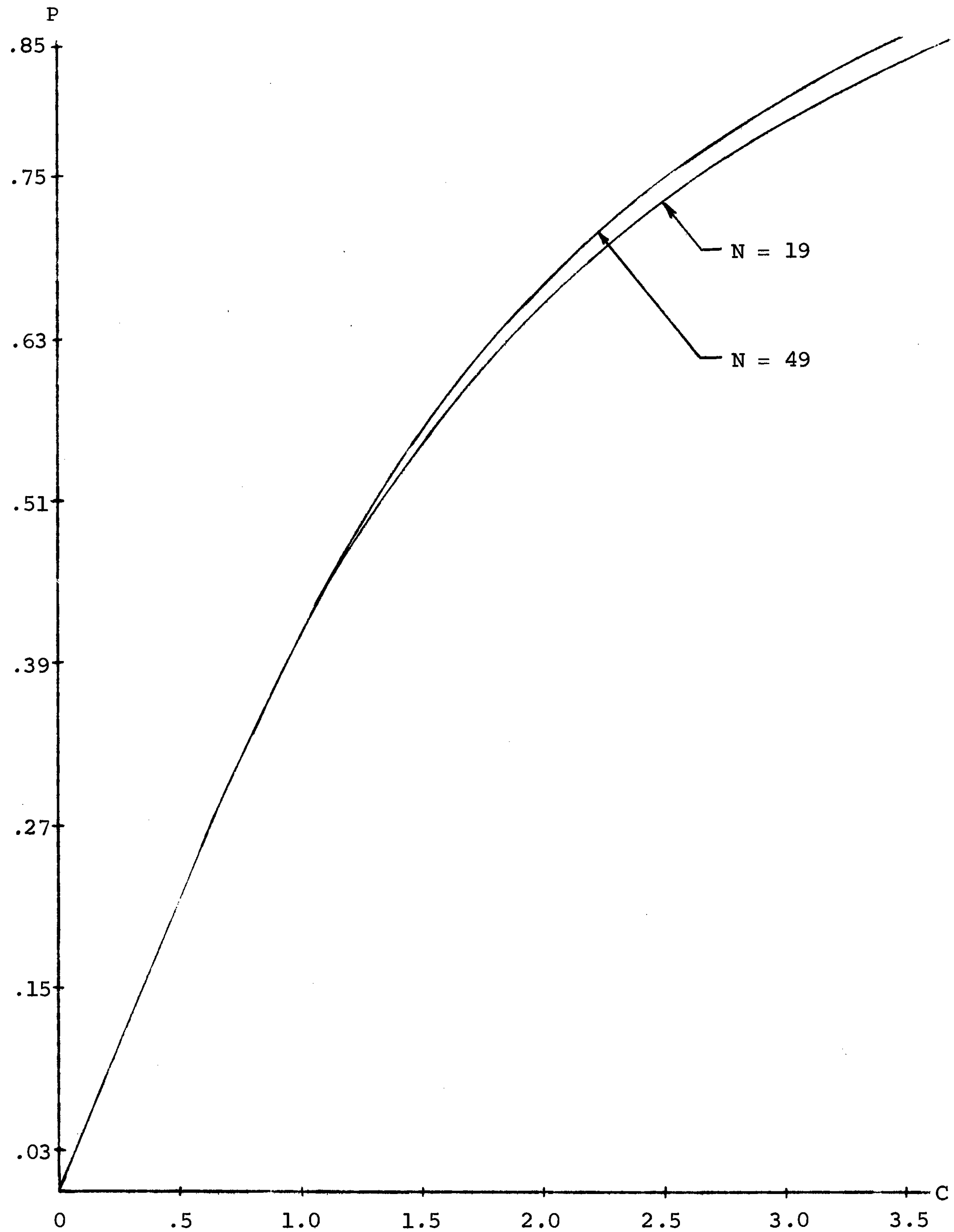
FOR ρ = (.2, .5, .8)



P = Prob.[F ≤ C]

FIGURE 5a: DISTRIBUTION FUNCTION OF F

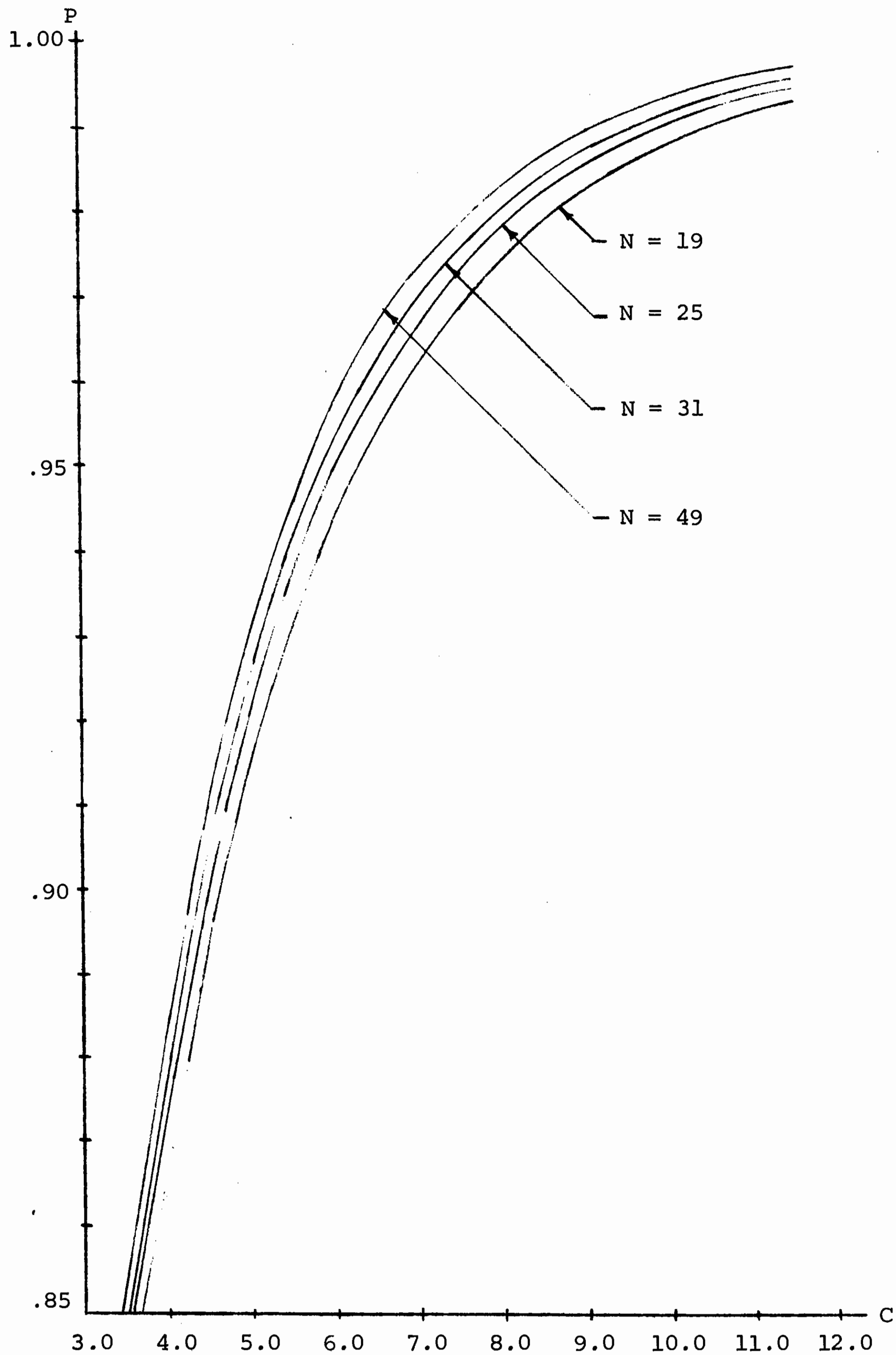
FOR $\rho = (.5, .5, .8)$



P = Prob.[F ≤ C]

FIGURE 5b: DISTRIBUTION FUNCTION OF F

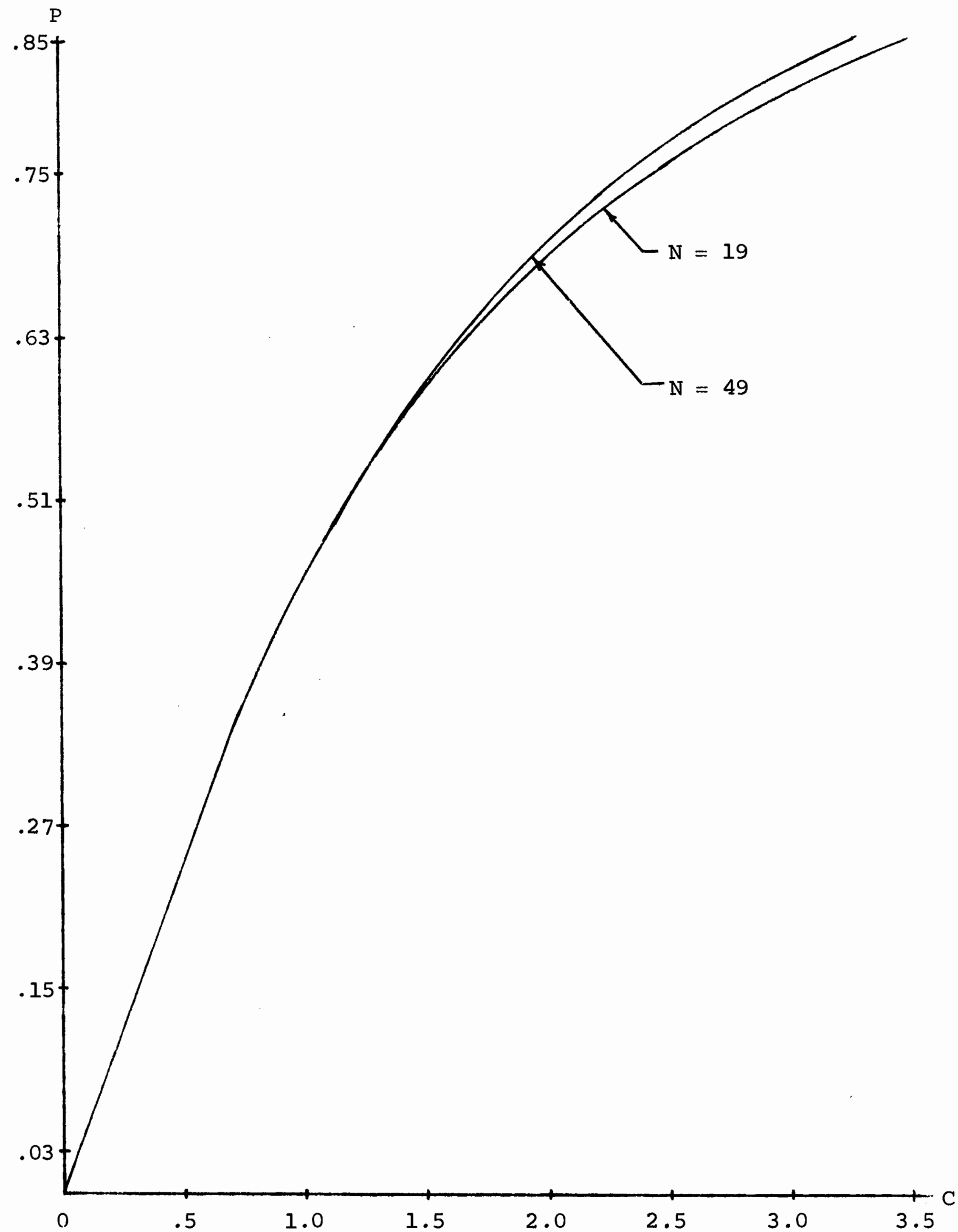
FOR ρ = (.5, .5, .8)



P = Prob.[F ≤ C]

FIGURE 6a: DISTRIBUTION FUNCTION OF F

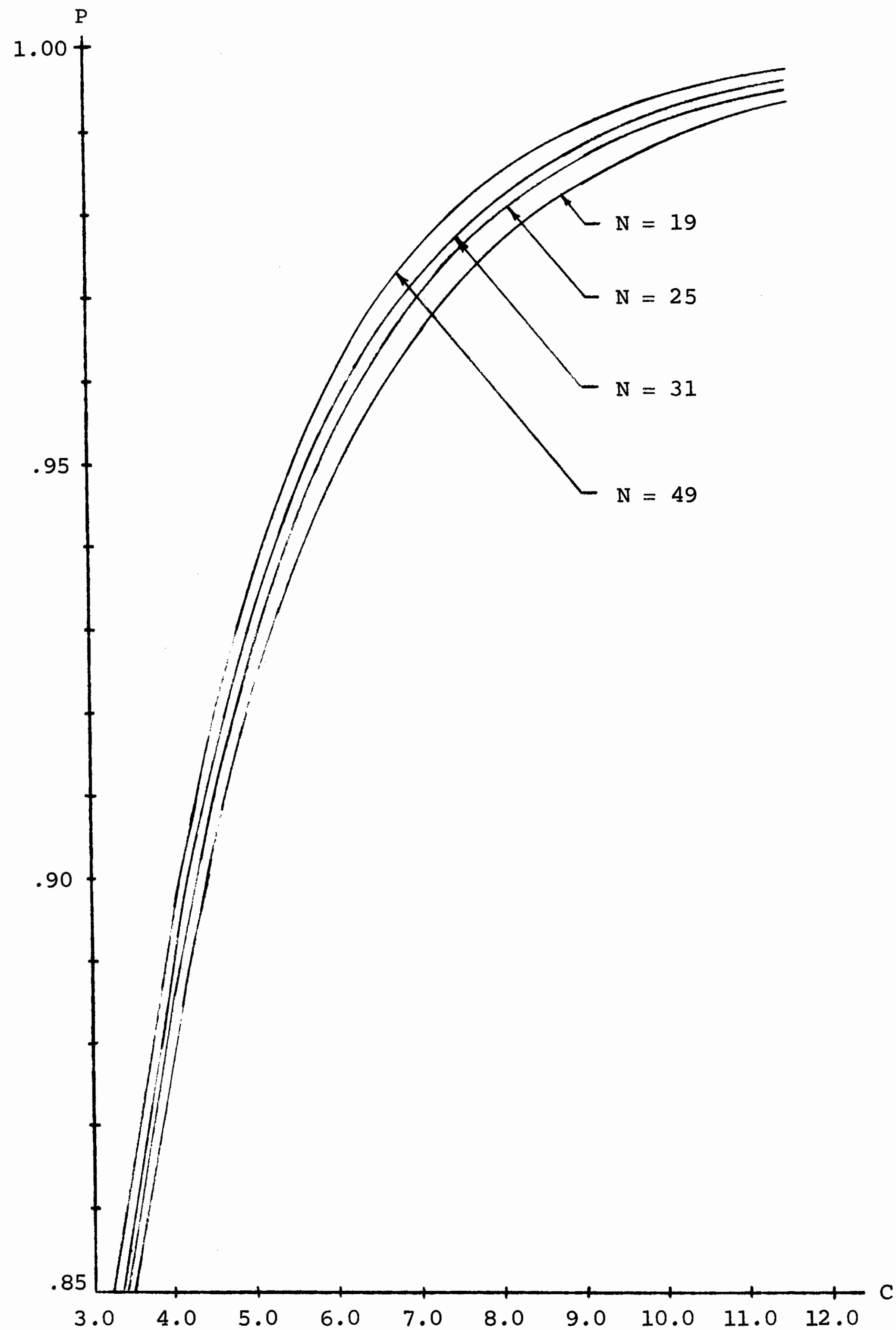
FOR ρ = (.5, .8, .8)



P = Prob.[F ≤ C]

FIGURE 6b: DISTRIBUTION FUNCTION OF F

FOR ρ = (.5, .8, .8)



P = Prob.[F ≤ C]

CHAPTER V

SELECTION OF THE CRITICAL POINT

An important phase of the forward selection procedure, or for that matter any of the procedures, is the selection of the constant F_0 to which the statistic F_i is compared when the variable x_i is being considered for entry into the prediction equation. This constant F_0 should be chosen so that the probabilities of excluding necessary variables and including unnecessary variables are at desirable levels. The desired level for these error probabilities depends largely on personal judgment. Unfortunately, because of the stepwise feature of making tests at each step when considering new variables for entry, these error probabilities are not known. However, by necessity, we shall act as though the F_i are jointly distributed as correlated F's at each step. Several philosophies regarding the problem of choosing F_0 will be considered and the consequences of each compared.

It is possible to consider our problem in the framework of a simultaneous test of hypothesis (see Ghosh [5]). At a certain stage when considering p variables for entry into the equation, there are p hypotheses of the form

$$H_i: \quad \beta_i \underline{X}_i' \underline{a}_i + \underline{\beta}_2' X_2' \underline{a}_i = 0$$

which are tested simultaneously. The joint hypothesis H that all these p linear combinations are equal to zero is rejected if any one of the H_i is rejected. The type I error associated with H is the rejection of at least one of the p hypotheses when in fact they are all true. This is equivalent to all the β_i, associated with those variables being considered, being equal to zero. Recall that the test statistic for the hypothesis H_i is $F_i = SSR_i/MSE$. If $F_i > F_0$, where F_0 is some constant, then the hypothesis H_i is rejected. Intuition then leads us to look at the statistic $F = \max(F_1 , F_2 , \cdots , F_p)$. If $F > F_0$, the joint hypothesis H is rejected.

The question arises as to what level of protection is desired against making type I and type II errors. The probability of making a type I error is decreased at the expense of increasing the probability of making a type II error, therefore, some thought should be given to the importance of the alternative action taken if the hypothesis H is rejected. The rejection of the joint hypothesis results in the entry of the variable x_i which corresponds to the largest ratio F .

In general, the error of excluding variables that should be included in the equation is more serious than the error of including variables that should be excluded. The equation is used for prediction, so more variables than are necessary will be of little harm. However, if certain necessary variables are excluded, the equation's ability to predict will be lessened. Also, since the estimator of E[Y] is biased when variables are omitted, the omission of necessary variables might cause the bias to be rather large. Thus the probability, which we shall denote as α , of making a type I error should be chosen with this in mind.

The classical approach to the problem of finding the constant F_0 has been to select a desirable value for α , usually in the range from $\alpha = .25$ to $\alpha = .01$, and then use the critical value from the F table, which corresponds to this α , for the constant F_0 . This procedure has some practical appeal since the critical points from an F distribution are well tabulated. However, as has been observed earlier, the statistic $F = \max(F_1, F_2, \cdots, F_p)$ does not follow an F distribution. Given H, the probability that F will exceed the constant F_0 is greater than α. The true probability that $F > F_0$, when F_0 comes from the F table, will be discussed later in this chapter.

Another philosophy concerning the choice of the critical point F_0 (or the type I error probability, α) is discussed by Duncan [3]. He draws a parallel between testing the joint hypothesis H and making independent tests of the p hypotheses. If circumstances validate considering the p tests as independent, then p α-level tests could be performed. If the p null hypotheses in the individual tests were true, the probability of not rejecting this joint hypothesis H, when it is true, would be $(1 - \alpha)^p$. Thus if p were equal to 3 and the individual test made at the 5% level, the probability of not rejecting the joint hypothesis H, when it is true, would be $(.95)^3 = .857$. Thus we would need to find the constant F_0 such that $P[F < F_0|H] = (1 - \alpha)^p$.

A final philosophy which will be discussed is that given by Scheffé [18]. He suggests that when testing p hypotheses simultaneously, α should be chosen to correspond to the probability of rejecting the least likely hypothesis when in fact it is true. In our case all the H_i could be considered equally likely. Then according to this, the constant F_0 could be chosen such that $P[F > F_0|H) = \alpha$, where $F = \max(F_1, F_2, \cdots, F_p)$.

The computations presented in the preceding chapter enable us to make comparisons among these philosophies concerning the choice of the critical point F_0 for the case when $p = 3$. Table 1 and Table 2 at the end of this chapter will give the value of the critical point F_0, for values of α = .25, .10, .05, and .01, for the different approaches discussed above. The value in parentheses in the "univariate F" column is the true value of $P[F < F_0|H]$, when it is thought to be $(1 - \alpha)$. Note the value of $P[F < F_0|H]$ equals $(1 - \alpha)^3$ for Duncan's approach.

From the results presented in the tables, some interesting observations can be made. Note that the critical points suggested by Duncan and by the univariate F are close together when the correlations among the three variables are small. These points become closer together as the value of α decreases. The critical points suggested by Scheffé which correspond to the α level stated are always greater than those of Duncan or the univariate F. This difference decreases as α decreases. Another interesting note is, that for small correlations among the variables, the critical points of Scheffé for $\alpha = .25$ are very close to the critical points of the univariate F for $\alpha = .10$. As mentioned earlier, when using the univariate F, the true value of α is much larger than it is thought to be. For example, when α is thought to be .10 it may actually be .25.

The column headed "Krishnaiah" gives the critical points from Krishnaiah's tables for the correlation equal to $\bar{\rho} = \Sigma\rho_{ij}/3$. For example, for $\rho = (.2, .2, .5)$, the critical points from Krishnaiah's table would be those for $\rho = .3$. Note that these critical points are, for practical use, good approximations to those points given by the graphs (Scheffé).

Table 3 gives a comparison of the critical points suggested by Nair (all numerators of the F_i's considered independent) and by Krishnaiah (all numerators considered equally correlated) for $\alpha = .05$ and $.01$. Note that Nair's values are always greater than the corresponding values of Krishnaiah and these differences increase as the correlations among the variables increase. Hence if F_0 were chosen from Nair's tables, the $P[F > F_0|H] < \alpha$. Thus as α decreases, the probability of a type II error might become quite large.

For values of p greater than 3 it would seem that the same relations among the critical points would appear. Further investigation needs to be done in this regard.

TABLE 1: COMPARISON OF CRITICAL POINTS FOR α = .25 AND .10

$\rho_{12}, \rho_{13}, \rho_{23}$	d.f.	SCHEFFÉ		DUNCAN		UNIVARIATE F		KRISHNAIAH($\bar{\rho}$)
		75%	90%	42%	73%	75%	90%	90%
.2,.2,.5	19	2.92	5.02	1.27	2.78	1.40	2.99	5.02
	25	2.88	4.86	1.27	2.71	1.39 (.46)	2.92 (.75)	4.86
	31	2.86	4.76	1.27	2.70	1.38	2.87	4.76
	49	2.81	4.63	1.27	2.63	1.36	2.81	----
.2,.5,.5	19	2.85	4.92	1.17	2.71	1.40	2.99	4.94
	25	2.80	4.78	1.17	2.64	1.39 (.49)	2.92 (.77)	4.78
	31	2.78	4.65	1.17	2.63	1.38	2.87	4.69
	49	2.71	4.51	1.17	2.57	1.36	2.81	----
.2,.2,.8	19	2.77	4.81	1.13	2.61	1.40	2.99	4.94
	25	2.73	4.61	1.13	2.56	1.39 (.50)	2.92 (.78)	4.78
	31	2.71	4.51	1.13	2.55	1.38	2.87	4.69
	49	2.63	4.44	1.13	2.49	1.36	2.81	----
.2,.5,.8	19	2.67	4.67	1.05	2.57	1.40	2.99	4.83
	25	2.62	4.51	1.05	2.46	1.39 (.52)	2.92 (.79)	4.69
	31	2.60	4.43	1.05	2.45	1.38	2.87	4.60
	49	2.54	4.31	1.05	2.41	1.36	2.81	----
.5,.5,.8	19	2.63	4.60	1.02	2.48	1.40	2.99	4.70
	25	2.59	4.48	1.02	2.41	1.39 (.53)	2.92 (.79)	4.56
	31	2.57	4.37	1.02	2.40	1.38	2.87	4.47
	49	2.50	4.25	1.02	2.36	1.36	2.81	----
.5,.8,.8	19	2.42	4.39	.90	2.27	1.40	2.99	4.52
	25	2.39	4.27	.90	2.22	1.39 (.56)	2.92 (.81)	4.39
	31	2.38	4.17	.90	2.21	1.38	2.87	4.31
	49	2.33	4.09	.90	2.18	1.36	2.81	----

TABLE 2: COMPARISON OF CRITICAL POINTS FOR α = .05 AND .01

$\rho_{12},\rho_{13},\rho_{23}$	d.f.	SCHEFFÉ		DUNCAN		UNIVARIATE F		KRISHNAIAH($\bar{\rho}$)	
		95%	99%	86%	97%	95%	99%	95%	99%
.2,.2,.5	19	6.75	11.02	4.20	8.13	4.38	8.18	6.69	11.11
	25	6.44	10.48	4.08	7.78	4.24 (.87)	7.77 (.97)	6.42	10.42
	31	6.27	10.04	4.00	7.55	4.16	7.53	6.26	10.03
	49	6.08	9.69	3.90	7.28	4.04	7.18	----	----
.2,.5,.5	19	6.64	11.00	4.13	8.02	4.38	8.18	6.61	11.04
	25	6.33	10.42	3.94	7.67	4.24 (.87)	7.77 (.97)	6.35	10.36
	31	6.16	10.03	3.87	7.44	4.16	7.53	6.19	9.97
	49	5.94	9.62	3.78	7.13	4.04	7.18	----	----
.2,.2,.8	19	6.57	10.97	4.00	7.98	4.38	8.18	6.61	11.04
	25	6.29	10.37	3.80	7.57	4.24 (.88)	7.77 (.97)	6.35	10.36
	31	6.14	10.00	3.78	7.30	4.16	7.53	6.19	9.97
	49	5.88	9.55	3.75	7.05	4.04	7.18	----	----
.2,.5,.8	19	6.34	10.70	3.85	7.70	4.38	8.18	6.51	10.94
	25	6.10	10.13	3.75	7.40	4.24 (.89)	7.77 (.97)	6.25	10.27
	31	5.94	9.81	3.70	7.10	4.16	7.53	6.10	9.89
	49	5.70	9.23	3.60	6.82	4.04	7.18	----	----
.5,.5,.8	19	6.28	10.53	3.79	7.63	4.38	8.18	6.37	10.79
	25	6.01	10.00	3.69	7.29	4.24 (.89)	7.77 (.98)	6.12	10.15
	31	5.87	9.57	3.60	7.04	4.16	7.53	5.97	9.78
	49	5.67	9.12	3.53	6.75	4.04	7.18	----	----
.5,.8,.8	19	6.02	10.27	3.63	7.37	4.38	8.18	6.18	10.58
	25	5.79	9.67	3.53	7.00	4.24 (.90)	7.77 (.98)	5.94	9.96
	31	5.64	9.33	3.49	6.87	4.16	7.53	5.81	9.61
	49	5.50	8.89	3.37	6.58	4.04	7.18	----	----

TABLE 3: CRITICAL POINTS GIVEN BY NAIR AND KRISHNAIAH FOR $\alpha = .05$ AND .01

$\rho_{12},\rho_{13},\rho_{23}$	d.f.	KRISHNAIAH		NAIR	
		95%	99%	95%	99%
.2,.2,.2	19	6.75	11.16	6.77	11.24
	25	6.47	10.46	6.49	10.51
	31	6.30	10.06	6.33	10.08
.5,.5,.5	19	6.51	10.94	6.77	11.24
	25	6.25	10.27	6.49	10.51
	31	6.10	9.89	6.33	10.08
.8,.8,.8	19	5.92	10.27	6.77	11.24
	25	5.70	9.69	6.49	10.51
	31	5.57	9.35	6.33	10.08

CHAPTER VI

SUMMARY

It has been our aim to take a close look at the forward selection procedure as a method of constructing a prediction equation. The basic approach to the problem and the different philosophies regarding the choice of critical values have been discussed in order to shed more light on what is actually being done and, hopefully, to enable one to construct the equation in an optimum manner.

The distribution of F , the statistic used to determine whether or not a certain variable should be included in the equation, was studied. The statistic F is often treated as though it follows the univariate F distribution, however, this is not true. The joint distribution of p correlated F-distributed variables was given by (3) in Chapter III. Then the distribution function of $F = \max(F_1 , F_2 , \cdots , F_p)$ was obtained numerically, with the results presented in Chapter IV.

From the results presented in the preceding chapter, some conclusions may be drawn concerning the choice of significance levels and corresponding critical points. As discussed earlier, it is important to protect against excluding necessary variables from the equation, therefore the value for α should not be chosen too small. Generally, $\alpha = .10$ or $\alpha = .25$ would be considered acceptable values. If one wishes to use the table of critical points of the univariate F distribution in choosing F_0 , he must remember that the true α value corresponding to such an F_0 is larger than expected.

Thus if using the F table, an F_0 should be chosen which corresponds to an α value smaller than the α value desired. For example, if an α of .25 were desired, the F_0 corresponding to $\alpha = .10$ might be used.

In order to use Krishnaiah's tables in choosing F_0 , it appears that the values of the correlations ρ_{ij} should be averaged and the critical point from Krishnaiah's table for $\rho = \bar{\rho}$ used for F_0 . The α value chosen in this manner will be approximately equal to the desired α value.

There is more work that needs to be done in the future. The probability integral of F needs to be calculated for values of p greater than 3 and comparisons made similar to those that have been made here. Also other data reduction methods, such as the backward solution, could be analyzed in much the same way that the forward selection procedure was done here.

APPENDIX

NUMERICAL INTEGRATION

The numerical integration was performed with the aid of a computer program written in Fortran IV. The integration technique used was a generalization of the trapezoidal rule. When performing the integration

$$\int_0^c \int_0^c \int_0^c f(F_1\ , F_2\ , F_3)dF_1dF_2dF_3 \quad ,$$

one is integrating over a cube with sides of length c . This large cube was divided into segments with smaller volumes to increase the accuracy of the approximation.

Consider a three dimensional figure within this cube of side c . Let ΔF_1 , ΔF_2 , and ΔF_3 be the lengths of the sides of this figure and represent the corners of the figure by the points (i, i, i), (i, i, i+1), (i, i+1, i+1), (i, i+1, i), (i+1, i, i), (i+1, i, i+1), (i+1, i+1, i), and (i+1, i+1, i+1). Then according to the trapezoidal rule,

$$\int_i^{i+1} \int_i^{i+1} \int_i^{i+1} f(F_1\ , F_2\ , F_3)dF_1dF_2dF_3 \sim$$

$$\frac{\Delta F_1 \Delta F_2 \Delta F_3}{8}\Big\{f(i,\ i,\ i) + f(i,\ i,\ i+1) + f(i,\ i+1,\ i+1)$$
$$+ f(i,\ i+1,\ i) + f(i+1,\ i,\ i) + f(i+1,\ i,\ i+1)$$
$$+ f(i+1,\ i+1,\ i) + f(i+1,\ i+1,\ i+1)\Big\}$$

The function $f(F_1, F_2, F_3)$ increases without bound as the arguments approach zero. This property requires us to choose regions of small volume for integration near any of the three axes. For regions near the axes, volumes as small as .000625 were used. For less crucial regions, volumes as large as .125 were used. The value of the integral over the entire cube of side c is then computed by summing the values of the integrals over these smaller disjoint regions.

As an illustration of the technique, consider evaluating

$$\int_2^3 \int_2^3 \int_2^3 f(F_1, F_2, F_3)\, dF_1 dF_2 dF_3 \quad .$$

We shall divide this unit cube over which we are integrating into disjoint cubes with sides of length .1 . Then there are 1000 such small cubes which compose the unit cube. As mentioned above, the values of the integral over these 1000 small cubes would be added to give the value over the unit cube.

First the function $f(F_1, F_2, F_3)$ must be evaluated at the corners of each of these small cubes. These values are stored in a three-dimensional array denoted by F(I, J, K) where F(1, 1, 1) = f(2, 2, 2), F(2, 1, 1) = f(2.1, 2, 2), etc. The order of this array is denoted by IO × JO × KO . For our particular example this would be 11 × 11 × 11 . Once the function values are stored, the following Fortran statements complete the integration.

```
   DO  1  I=2,IOO                    (where  IOO=IO-1
   DO  1  J=2,JOO                            JOO=JO-1
   DO  1  K=1,KO,KOO                         KOO=KO-1)
1  F(I,J,K)=F(I,J,K)*4.0
   DO  2  I=1,IO,IOO
   DO  2  J=2,JOO
   DO  2  K=2,KOO
2  F(I,J,K)=F(I,J,K)*4.0
   DO  3  I=2,IOO
   DO  3  J=1,JO,JOO
   DO  3  K=2,KOO
```

```
3  F(I,J,K)=F(I,J,K)*4.0
   DO  4  I=2,IOO
   DO  4  J=2,JOO
   DO  4  K=2,KOO
4  F(I,J,K)=F(I,J,K)*8.0
   DO  5  I=2,IOO
   DO  5  J=1,JO,JOO
   DO  5  K=1,KO,KOO
5  F(I,J,K)=F(I,J,K)*2.0
   DO  6  I=1,IO,IOO
   DO  6  J=2,JOO
   DO  6  K=1,KO,KOO
6  F(I,J,K)=F(I,J,K)*2.0
   DO  7  I=1,IO,IOO
   DO  7  J=1,JO,JOO
   DO  7  K=2,KOO
7  F(I,J,K)=F(I,J,K)*2.0
   ANS(N)=0.0
   DO  8  I=1,IO
   DO  8  J=1,JO
   DO  8  K=1,KO
8  ANS(N)=ANS(N)+F(I,J,K)
   ANS(N)=ANS(N)*CO(N)/8.0
```

CO(N) = $\Delta F_1 \cdot \Delta F_2 \cdot \Delta F_3$. IO, JO, and KO are constants which denote the order of the array of stored functional values. These values must be defined before the above operations are carried out. ANS(N) is the value of the integral over this region (N being a subscript to denote a particular region).

As mentioned earlier, this function is unbounded as the arguments approach zero. In the integration, the function is evaluated at points very close to zero to take the place of the function evaluated at zero. The function increases faster as the values of the correlations become larger. Therefore the increments for integration were chosen so that the approximation was accurate for ρ = (.8, .8, .8). For cases which involve smaller correlations, a correction term was added to the value of the integral near the three axes, since the use of the increments for ρ = (.8, .8, .8) gives a value smaller than the true value. For

ρ = (.5, .5, .5), a correction term of .91 was added, and for ρ = (.2, .2, .2), a correction term of .01 was added. These results were checked against those results of Krishnaiah and were found to be very accurate. For other values of ρ , proportionate adjustment was made.

LIST OF REFERENCES

1. Beale, E. M. L., Kendall, M. G., and Mann, D. W. "The discarding of variables in multivariate analysis." Biometrika, Vol. 54 (1967), pp. 357-366.

2. Draper, N. R. and Smith, H. Applied Regression Analysis. New York: John Wiley and Sons, Inc., 1966.

3. Duncan, David B. "Multiple range and multiple F test." Biometrics, Vol. 11 (1955), pp. 1-42.

4. Dunnett, C. W. and Sobel, M. "A bivariate generalization of Student's t distribution, with tables for certain special cases." Biometrika, Vol. 41 (1954), pp. 153-169.

5. Ghosh, M. N. "Simultaneous test of linear hypotheses." Biometrika, Vol. 42 (1955), pp. 441-449.

6. Graybill, F. A. An Introduction to Linear Statistical Models, Volume I. New York: McGraw-Hill Book Company, 1961.

7. Gröbner, Wolfgang and Hofreiter, Nikolaus. Integraltafel. Part 2, third edition, Vienna: Springer-Verlag, 1961, p. 64.

8. Gupta, S. S. and Sobel, M. "On the smallest of correlated F-statistics." Biometrika, Vol. 49 (1962), pp. 509-523.

9. Hamaker, H. C. "On multiple regression analysis." Statistica Neerlandica, Vol. 16 (1962), pp. 31-56.

10. Hocking, R. R. and Leslie, R. N. "Selection of the best subset in regression analysis." Technometrics, Vol. 9, November 1967, pp. 531-540.

11. Krishnaiah, P. R. and Armitage, J. V. Probability Integrals of the Multivariate F Distribution, with Tables and Applications. ARL 65-236, November 1965, Aerospace Research Laboratories, Wright-Patterson Air Force Base, Ohio.

12. Larson, H. J. and Bancroft, T. A. "Biases in prediction by regression for certain incompletely specified models." Biometrika, Vol. 50 (1963), pp. 391-402.

13. Laubscher, Nico F. "Interpolation in an F table." The American Statistician, Vol. 19, February 1965, pp. 28 and 40.

14. McCormick, John M. and Salvadori, Mario G. Numerical Methods in Fortran. Englewood Cliffs, New Jersey: Prentice-Hall, 1964.

15. Nair, K. R. "The studentized form of the extreme mean square test in the analysis of variance." Biometrika, Vol. 35 (1948), pp. 16-31.

16. Owen, D. B. Handbook of Statistical Tables. Reading, Mass.: Addison-Wesley Publishing Company, Inc., 1962.

17. Ralston, A. and Wilf, H. S. Mathematical Methods for Digital Computers. New York: John Wiley and Sons, Inc., 1962. Article: "Multiple regression analysis" by M. A. Efroymson.

18. Scheffé, H. "A method for judging all contrast in the analysis of variance." Biometrika, Vol. 40 (1953), pp. 87-104.

19. Wallace, T. D. and Toro-Vizcarrondo, Carlos. "A test of the mean square error criterion for restrictions in linear regression." Journal of the American Statistical Association, Vol. 63, June 1968, pp. 558-572.

20. Webster, J. T. "On the use of a biased estimator in linear regression." Journal of the Indian Statistical Association, Vol. 3, April-July 1965, pp. 82-90.

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

On the Stepwise Construction of a Prediction Equation

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

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

Paul Terrell Pope

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
June 25, 1969	53	20

8a. CONTRACT OR GRANT NO. N00014-68-A-0515	9a. ORIGINATOR'S REPORT NUMBER(S) 37
b. PROJECT NO. NR 042-260	
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	

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

In the analysis of data, it is often desired to construct an equation of the form $y = b_0 + b_1x_1 + b_2x_2 + \cdots + b_kx_k$, where y is the response variable and x_1 , x_2 , $\cdots$, x_k are input variables. A familiar method of constructing such an equation is known as the forward selection procedure.

At some stage of the forward selection procedure when p variables are being considered for entry into the equation, the ratio $F_i = SSR_i/MSE$ is computed for each variable being considered. If the ratio $F = \max(F_1 , F_2 , \cdots , F_p)$ is greater than some constant F_0 , the variable x_i associated with this maximum ratio is included in the equation.

The joint distribution of F_1 , F_2 , $\cdots$, F_p is determined under the usual assumptions for the regression model; including normality of the error. Through this the distribution of F is given. Different philosophies exist concerning the choice of the value for F_0 . In this paper the probability integral of F is computed for p = 3 and comparisons are made among the critical points F_0 for these different philosophies.

DD FORM 1 NOV 65 **1473**

UNCLASSIFIED

Security Classification