APPLICATION OF A BIVARIATE $t$ DISTRIBUTION
TO HYPOTHESIS TESTING IN CROSSOVER EXPERIMENTS

by

Paul Wilder Stewart

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1401

May 1982
APPLICATION OF A BIVARIATE $t$ DISTRIBUTION
TO HYPOTHESIS TESTING IN CROSSOVER EXPERIMENTS

by

Paul Wilder Stewart

A Thesis submitted to the faculty of The University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biostatistics

Chapel Hill
1981

Approved by:

Ronald W. Helms
Advisor

James L. Johnson
Reader

Dana Quade
Reader
PAUL WILDER STEWART. Application of a Bivariate t Distribution to Hypothesis Testing in Crossover Designs. Under the direction of RONALD W. HELMS, Ph.D.

The joint distribution of a pair of Student-t variates, 
\((t_1, t_2)\), is investigated and methods of computing \(\Pr(t_1 \in \text{set}_1\) and \(t_2 \in \text{set}_2)\) are developed. The resulting algorithms are used to evaluate the power and actual significance level of two-sided tests which are based on the test statistic \(\max |t_i|\) and the Bonferroni inequality. It is assumed that the numerators of \(t_1\) and \(t_2\) are dependent, that the denominators are dependent, and that numerators are independent of denominators.

This research is motivated by an alternative 'by-period' analysis for crossover experiments recently suggested by Helms, Kempthorne and Shen (1980). Conceptual advantages of the by-period analysis over the traditional analysis are discussed and a comparison is made with respect to the power of tests against alternative hypotheses.

An original crossover experiment is report and analyzed according to the by-period method. The power of the treatment effects test is computed. A messy-data subset of the complete data is then re-analyzed and power is computed. The power of the analogous test in the traditional analysis is also computed and comparisons are made between the two analyses. It is found that the by-period testing procedure is more powerful than that in the traditional analysis for this problem.
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>INTRODUCTION AND SUMMARY</td>
<td>v</td>
</tr>
<tr>
<td>1. Review of the Literature</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Summary</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Multivariate t Distributions</td>
<td>1</td>
</tr>
<tr>
<td>1.3 Crossover Designs: Theory and Applications</td>
<td>10</td>
</tr>
<tr>
<td>2. Bivariate-t Distribution Theory and Power Computations: The Case of Identical Design and Matrices in the Linear Models for Periods 1 and 2</td>
<td>16</td>
</tr>
<tr>
<td>2.1 Summary</td>
<td>16</td>
</tr>
<tr>
<td>2.2 Formulation of the Power Function $P(t)$</td>
<td>16</td>
</tr>
<tr>
<td>2.3 A Maclaurin Series Expansion</td>
<td>26</td>
</tr>
<tr>
<td>2.4 Two Algorithms for Computing Power</td>
<td>45</td>
</tr>
<tr>
<td>2.5 Comparison of the Two Algorithms</td>
<td>53</td>
</tr>
<tr>
<td>3. Bivariate-t Distribution Theory for Power Computations: The Case of Different Design Matrices in the Linear Models for Periods 1 and 2</td>
<td>55</td>
</tr>
<tr>
<td>3.1 Summary</td>
<td>55</td>
</tr>
<tr>
<td>3.2 The Structure of the Data</td>
<td>56</td>
</tr>
<tr>
<td>3.3 Transformation of the Data to Canonical Form</td>
<td>56</td>
</tr>
<tr>
<td>3.4 An Algorithm for Computing Power</td>
<td>71</td>
</tr>
<tr>
<td>4. Examples of Data Analyses with Power Calculations</td>
<td>73</td>
</tr>
<tr>
<td>4.1 Summary</td>
<td>73</td>
</tr>
<tr>
<td>4.2 A Bean Sprout Experiment</td>
<td>74</td>
</tr>
<tr>
<td>4.3 By-Period Analysis for the Complete Set of Data</td>
<td>75</td>
</tr>
<tr>
<td>4.4 Messy Data Analyses</td>
<td>84</td>
</tr>
<tr>
<td>4.4.1 By-Period Analysis</td>
<td>85</td>
</tr>
<tr>
<td>4.4.2 Power Calculations for the Traditional Model</td>
<td>91</td>
</tr>
<tr>
<td>4.4.3 Comparisons of Power</td>
<td>96</td>
</tr>
<tr>
<td>5. Recommendations for Further Research</td>
<td>98</td>
</tr>
<tr>
<td>Appendix</td>
<td>101</td>
</tr>
<tr>
<td>A.1 Bean-Sprout Experiment Protocol</td>
<td>101</td>
</tr>
<tr>
<td>A.2 Algorithm: Power via Maclaurin Polynomials</td>
<td>108</td>
</tr>
</tbody>
</table>
A.3 Algorithm: Power via Monte Carlo Integration........ 116
A.4 Algorithm: Power and Significance Level via Monte Carlo Simulation................................. 120

References.................................................................................................................. 124
ACKNOWLEDGEMENTS

I wish to thank Dr. Ronald W. Helms for the guidance and cheerful encouragement that gave this study the potential of its success. I also wish to thank the other members of my Doctoral Committee for their individual contributions: Dr. N.L. Johnson, Dr. Dana Quade, Dr. Dennis B. Gillings and Dr. Seth A. Rudnick. I am grateful to Mrs. Beryl Glover who so carefully typed the final draft. Finally, I wish to express appreciation to special friends, Dr. Patrick W. Crockett and Senorita Susan G. Navey, for their support and encouragement.
INTRODUCTION AND SUMMARY

The joint distribution of a pair of Student-t variates, \((t_1, t_2)\), is investigated and methods of computing \(p r t_1 \in \text{set}_1\) and \(t_2 \in \text{set}_2\) are developed. The results are used to evaluate the power and actual significance level of two-sided tests which are based on the test statistic \(\max|t_i|\) and the Bonferroni inequality. It is shown that some of the results can be used to evaluate power and significance of one-sided tests as well.

In this research it is assumed that \(t_1\) and \(t_2\) are defined as follows:

\[
t_i = \frac{C_i \left[ x_i^T x_i \right]^{-1} x_i^T Y_i}{\left[ Y_i^T P_i Y_i / (n_i - r_i) \right]^{1/2}}
\]

where

\[
P_i(n_i \times n_i) = I - x_i x_i^T x_i^{-1} x_i^T,
\]

\(C_i(1 \times r_i)\) is a matrix of constants,

and \(Y_i(n_i \times 1) \sim N_{n_i}(x_i \beta_i, \sigma_i^2 I)\).

\(X_i\) is the design matrix of the general linear univariate model denoted GLUM \((X_i \beta_i, \sigma_i^2 I)\). It is assumed that the numerators of \(t_1\) and \(t_2\) are dependent, that the denominators of \(t_1\) and \(t_2\) are dependent, and that numerators are independent of denominators.
This assumption is true when the only dependence between \( Y_1 \) and \( Y_2 \) occurs as dependence between some \((Y_{1i}, Y_{2i})\) pairs. It follows that \( t_1 \) and \( t_2 \) are dependent and follow different distributions.

This research is motivated by an alternative 'by-period' strategy for crossover experiments recently suggested by Helms, Kempthorne and Shen (1980). We will focus on two-period, four-sequence, two-treatment crossover designs (denoted \( CO(2,4,2) \)). The by-period analysis avoids some difficulties associated with traditional analyses but requires the straightforward use of the Bonferroni inequality and \( \max |t_i| \) for certain two-sided tests. It is therefore of interest to know how the power of tests by this method of analysis compares with the power of analogous tests by traditional methods.

Both analytic and numerical methods are used to construct algorithms for computing power and actual significance level for the by-period method. The algorithms are then used as tools in making a comparison between traditional analysis and by-period analysis of crossover data.

For the case of \( X_1 = X_2 = \chi(n \times r) \), an algorithm based on a Taylor series is developed in Chapter 2 along with an algorithm using Monte Carlo integration of a fourfold integral. The Taylor series expression is algebraically complicated and its unknown interval of convergence is a function of several parameters. The Taylor algorithm and the Monte Carlo integration algorithm seem to perform equally well whenever the Taylor series converges.
For the more general case of $X_1(n_1 \times r_1) \neq X_2(n_2 \times r_2)$, a Monte Carlo simulation algorithm is developed in Chapter 3. The algorithm is simple in design and efficient in practice. Since only about two significant digits of accuracy are required for purposes of this research, the algorithm is inexpensive. Furthermore it is sufficiently general to handle power and significance level calculations for both two-sided and one-sided tests.

All three algorithms developed are generally useful for certain probability calculations and are not restricted to use with crossover data.

Finally, an original crossover experiment, conducted for the purpose of illustrating the theoretical results of this study, is reported and analyzed in Chapter 4. In addition, a subset of the data is taken as an example of messy data. The by-period and traditional analyses of the messy data are compared. It is shown that the by-period strategy is in some ways more advantageous and that it has superior power for testing treatment effects in the particular messy data analyzed.
CHAPTER 1

REVIEW OF THE LITERATURE

1.1 Summary

The best known of the many possible multivariate $t$ distributions are defined and discussed. It appears that the multivariate theory needed has not yet been reported in the literature. Important contributions to the literature on crossover designs are discussed and outstanding problems concerning analysis are noted.

1.2 Multivariate $t$ Distributions

Generalization of the univariate $t$ distribution can take many forms. Several multivariate $t$ distributions arise naturally in practice. The development of such distributions has a short history and much of the work is incomplete. Although multivariate theory research was gaining momentum at the turn of the century when 'Student' published a description of the $t$ distribution, much of the theory could not be applied until the coming of the electronic computer. The development of multivariate $t$ distributions began circa 1953.

Johnson and Kotz (1972) collate and summarize the literature on a variety of multivariate $t$ ($t_\nu$) distributions. Most $t_\nu$ random vectors (r.v.s.) are of the form $t_\nu = S^{-1} * z$ where the
non-zero elements of $S*S'$ are gamma variates and $Z$ follows an independent multinormal distribution. For this reason the study of a given $t$ distribution often involves concomitant study of such distributions as the Wishart and multivariate chi-square. In this research the $t$-distribution of interest can be characterized by

$$\mathbf{t} = \begin{pmatrix} Q_1 & 0 \\ 0 & Q_2 \end{pmatrix}^{-\frac{1}{2}} \mathbf{Z} = S^{-1/2}Z$$

where $Z(2x1) \sim N_2(\mu, \Sigma)$ independent of $Q_1, Q_2$. The diagonal matrix $S*S'$ is such that $Q_1\nu_1/\sigma_1^2$ is a chi-square r.v. with $\nu_1$ degrees of freedom. Furthermore $|Q_1, Q_2|$ forms the diagonal of a two-dimensional Wishart matrix.

Johnson and Kotz (1972) summarize the still incomplete study of the many multivariate extensions of the gamma distributions. That study did not begin until about the time Wishart derived his well-known distribution in 1928.

1.2.1 Multivariate $t$ with Common Denominator

The most completely developed $t$-distribution has the form

$$\mathbf{t}_{mx1} = S^{-1/2}Z_{mx1} = |Q|^{-\frac{1}{2}}Z$$

where $Z \sim N(\mu, \sigma^2R)$ is independent of $Q \sim \chi^2_{\nu} \sigma^2/\nu$. $R$ is a correlation matrix. This distribution is often referred to as "the multivariate $t$ distribution." It is not particularly useful for appli-
cation to the analysis of crossover (CO) designs since it is not usually appropriate to assume that $\Xi = \sigma^2 \mathcal{R}$.

Most contributions to the literature on t distributions deal specifically with $\frac{1}{S^2}$. Johnson and Kotz (1972) trace the development of this distribution beginning with fundamental results attributed to Cornish (1954) and Dunnett and Sobel (1954). Several applications, algorithms and tabulations have been contributed to the literature on $\frac{1}{S^2}$ since 1971. Freeman and Kuzman (1972) tabulate the distribution of $t = \max_{i} \left| \frac{X_i - S_{i+1}}{\sqrt{2} S_{i+1}} \right|$. Dutt (1972, 1975) discusses rapidly convergent algorithms for evaluating probability integrals of $t$. Wynn (1975) and Bohrer (1973) evaluate probability integrals and provide tables for $\Pr[t \in A_r(c)]$ where $A_r(c)$ define polyhedral cones. Nagarsenkar (1975) obtains the distribution of certain quadratic forms in $t$. Ghosh (1975) obtains the distribution of $(t_1 - t_2)$. Zellner (1976) exemplifies the application of $\frac{1}{S^2} = t_2$ by studying regressions with $t$ error terms. Dutt, et al. (1976) suggest an approximation to the distribution of $\max_{i=1,2,3} |t_i|$. Bishop, et al. (1978) describe and compare four methods for approximating the percentage points of quadratic forms in independent t variates. Amos (1978) demonstrates the evaluation of c.d.f.'s of $t_2$ by numerical quadrature. Chen (1979) tabulates percentage points for $\frac{1}{S^2}$ where $\sigma = 0$. 
1.2.2 Matric-t distributions

Ando and Kaufman (1965) obtained an equivalent representation for \( t = \frac{1}{\sqrt{Z}} \). They showed that

\[
\text{If } t \sim \left[ \frac{1}{\nu} \chi_{\nu}^{1/2} \cdot Y \right]_{m \times 1} \text{ with } Y \sim \mathcal{N}(Q, \Sigma) \\
\text{then } \frac{1}{\sqrt{\nu}} J^{-1} \cdot Z \sim \mathcal{N}(Q, \Sigma^{-1}) \\
\text{and } J \cdot \cdot t \sim \mathcal{W}_m(\nu + m - 1, \Sigma^{-1}).
\]

Dickey (1967) generalized this result by computing the joint p.d.f. of

\[
T = J^{-1} * X \quad \text{where } X_i \sim \mathcal{N}(\overline{Q}, \Sigma) \quad i = 1, 2, \ldots, m \\
\text{and } J \cdot \cdot t \sim \mathcal{W}_m(\nu + m - 1, \Sigma^{-1})
\]

where \( X_i \) and \( X_j \) are independent if \( i \neq j \), and \( \Sigma \) and \( Q \) are positive definite. This distribution is called the matric-t distribution. Tan (1969) obtained the p.d.f. of a special case given by

\[
T = Y \sim \mathcal{W}_m(\nu, \Sigma \cdot \Sigma^{-1}) \quad \text{independent of } Y \sim \mathcal{N}(\overline{Q}, \Sigma \cdot \Sigma^{-1}) \quad \text{singular}
\]

subject to the restrictions \( B \cdot Y = Q \). Dickey (1976) reports a new representation of \( t \) and matric-variate \( t \) as functions of independent \( t \) or \( T \) variates. Juritz and Troskie (1976) discuss non-central \( \mathcal{T} \) distributions that can be applied to 2-stage least squares estimators.
1.2.3 Multivariate Behrens-Fisher Distribution

Dickey (1966) investigated

\[ T = \sum_{j=1}^{k} \frac{B_j}{\sigma_j^2} \cdot t_j, \quad |B_j| \neq 0 \quad \forall \quad j \]

\[ t_j = \frac{1}{s_j} \cdot Z_j \]

\[ Z_j \sim N(\mu, I_m) \quad \text{iid} \]

\[ s_j^2 \sim \frac{1}{\nu_j} \cdot \chi^2_{\nu} \]

Cornish (1966) provides an example of application. Walker and Saw (1978) discuss the distribution for the case of \( m=1 \).

1.2.4 2-step t Distribution

Bulgren, et al., (1974) suggests the use of a 2-step test for data of the form

\[ \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \sim N(\mu, \sigma^2 \cdot I) \]

The t-variates for the 2 steps are

\[ t_1 = \frac{1}{\frac{1}{n} \cdot Y_1 \cdot \left[ Y_1^T \left[ I - \frac{1}{n} \right] Y_1 \right]^{(m)(m-1)}} \]

\[ t_2 = \frac{1}{\frac{1}{m+n} \cdot \left[ Y_1 \right] \cdot \left[ Y_2 \right] \cdot \left[ I - \frac{1}{m} \right] \cdot \left[ Y_2 \right] \cdot \left[ I - \frac{1}{n} \right] \cdot \left[ Y_1 \right] \cdot \left[ Y_2 \right] \cdot \left[ \left( m+n \right) \left( m+n-2 \right) \right]} \]
The p.d.f. of $\bar{t}$ is known and $\bar{t}$ converges in law to $\frac{\sqrt{m}}{\sigma} \left( \frac{1}{\sqrt{m+n}} \bar{Y}^* \right)$ where $\bar{Y}^* = \frac{1}{n+m} \cdot I^T \left[ \begin{array}{c} Y_1 \\ Y_2 \end{array} \right]$. 

Spurrer and Hewett (1975) provide tables for this distribution. Spurrer and Hewett (1976) extend the distribution to the two-population problem.

These distributions are reminiscent of the stepwise $t$ distributions discussed by Steffens (1969), Derflinger and Stappler (1976), and Siotani (1976) for stepwise regression analysis.

1.2.5 Non-student Multivariate $t$ Distribution

Johnson and Kotz (1972) discuss

$$t_i = \frac{\bar{x}_i}{\frac{1}{n} I^T \bar{x}_i \bar{x}_i} \quad i=1,2,\ldots,m$$

where

$$\bar{x}_i = \frac{1}{n} I^T x_i$$

and $x_i \sim N(0, I)$

$$\bar{x}_i \sim N(0, \frac{1}{n} I)$$

$$\bar{S} = \frac{1}{m} \sum_{i=1}^{m} S_i$$

$$S_i = \bar{x}_i^T [I - \frac{1}{n} I] \bar{x}_i$$

The distribution of $\bar{S}$ is a mixture of $\chi^2$ distributions and the p.d.f. of $\bar{t}$ has a relatively simple form. The marginals of this distribution are mixtures of Student-$t$ distributions.
1.2.6 General Multivariate t Distribution

The t distribution of interest in this research is presented by

\[
t = \left[ Q_1 \cdots Q_m \right]^{-1/2} \* Z_{mx1} \\text{with} \ Z_{mx1} \ \overset{d}{\sim} N(\mu, \Sigma) \]

independent of \([Q_1, \ldots, Q_m]\)

\[
Q_i \overset{\nu_i}{\sim} \frac{1}{\alpha_i^2} \overset{\nu_i}{\sim} X^{-2}_i \] is not necessarily mutually independent.

Johnson and Kotz (1972) refer to this form as type (III). They refer to the special case of \(Q_i \overset{\nu_i}{\sim} X^{-2}_i\) as type (II).

The marginal p.d.f.s for type (III) distributions are defined as follows: If \(Y \overset{\nu}{\sim} N(\mu, I)\) independent of \(Q^{-2}_i\), then \(T = Y / \sqrt{Q_i}\) is distributed as a noncentral t with \(\nu\) degrees of freedom and noncentrality parameter \(\mu\). The p.d.f. of this ratio is

\[
f_T(t) = \frac{\begin{bmatrix} \nu+1 \\ \nu \end{bmatrix}^{\nu+1} \ \exp[-\frac{1}{2} \mu^2]}{\sqrt{\nu} \ \Gamma(\frac{\nu}{2}) \ [\nu+t^2]^{-\nu/2}} \ \sum_{k=0}^{\infty} \frac{\Gamma(\nu+k+1)}{(k)! \ (\nu+t^2)^{k/2}}
\]

Halperin (1967) provides some inequalities for type (II) of the form \(t_i \overset{\alpha_i}{\sim} t_i \overset{\alpha_i}{\sim} N(\mu_i, R_i)\) for probabilities involving \(|t_i|\).

Miller (1968) discusses a class of distributions which can be constructed as

\[
t_i = \left[ \frac{\chi_i}{1 + \chi_i} \right]^{\nu_i/2} \ \text{i}=1, \ldots, m
\]

\(Y \overset{d}{\sim} N(\mu, R)\)

\(X_i = \frac{Z_i^T Z_i}{\nu_i} \overset{\nu_i}{\sim} X_i^2\)

\(Z_i \overset{d}{\sim} N_{\nu_i}(0, I)\) not necessarily independent w.r.t. \(i\).
Siddiqui (1967) derives the joint density of $|t_1, t_2|$ of type (III) for the bivariate case when $\Sigma$ is a correlation matrix (i.e., $\sigma_1^2 = \sigma_2^2 = 1$). He computes the joint density of $t_1, t_2$ and $R$ (the sample correlation) and then integrates out $R$ to reach an approximation of the p.d.f. of $t_1$ in terms of the hypergeometric function. 

Sidak (1971) proves the following corollary for $t$ probabilities. If: (1) the rows of $Y$ are i.i.d $N(y, \Sigma)$, (2) the correlation matrix of $\Sigma$ is $R$ and $r_{ij} = J_i \cdot J_j$ (if $i \neq j$) where $|T_i| \leq 1$, (3) $[Y_1, \ldots, Y_p] = \frac{1}{n} 1_n 1_{nxp}$, (4) $[S_1^2 \ldots S_p^2] = \text{diag} \left[ n^{-1} Y_T Y_N^{-1} \right]$, then

$$p_r \left[ \frac{|Y_i - \mu_{i1}|}{S_i / \sqrt{N}} \leq C_i, \forall i = 1, \ldots, p \right] \geq \prod_{i=1}^{p} p_r \left[ \frac{|Y_i - \mu_{i1}|}{S_i / \sqrt{N}} \leq C_i \right] .$$

Krishnan (1972) derives the central and singly-noncentral p.d.f. and c.d.f. of $t$ of type (III) for the case of $\Sigma = \sigma^2 R$, where $R$ is a correlation matrix. He expresses the p.d.f. in terms of the integrated coerror function, $i(b)erfc(a)$, and expresses the c.d.f. in terms of a double sum of incomplete beta or hypergeometric functions.

1.2.7 Related Multivariate Gamma Distributions

Jensen (1970) discusses the joint distribution of
\[ Q_1, Q_2 \] where \( Q_i = Z_i^T \cdot Z_i \) \((i=1, 2)\) and \([Z_1^T, Z_2^T]\) is normally distributed with zero mean and variance-covariance matrix

\[
\begin{bmatrix}
I_{V_1} & \rho [\Delta, 0] \\
\rho [\Delta, 0] & I_{V_2}
\end{bmatrix}
\]

(when \( V_2 \leq V_1 \)). He derives a series expression for the bivariate p.d.f. via the characteristic function of \(|Q_1, Q_2|\). (These results are discussed more fully in Chapter 3.)

Lukacs and Laha (1964) and others have shown that the joint characteristic function of \([S_1^2, \ldots, S_m^2]\) is \( \phi(t) = |I - 2iR \cdot \text{Diag}(t_1, \ldots, t_m)| \) whenever \([S_1, \ldots, S_m] \overset{d}{\sim} N_m(\mathbf{0}, \mathbf{R})\).

In addition to the literature on multivariate gamma and Wishart distributions summarized by Johnson and Kotz (1972), the following more recent contribution is also of interest. Krishnan (1976) derives the moment generating function and the p.d.f. for \([S_1^2, S_2^2]\) when \( S_i^2 = \chi_i^T A \chi_i \) and \( \chi_i \overset{d}{\sim} \text{i.i.d.} N_p(\mathbf{0}, \sigma^2 I_p) \).

\[ 1.2.8 \] Other Publications of Interest

Kerridge and Cook (1976) compare historic algorithms for computing normal probability integrals and suggest an alternative. If the p.d.f., \( f(x) \), is expressed as a Taylor series and then integrated, then the Hermite polynomials involved can be operated on recursively to yield a fast computer algorithm.
Numerical results indicate that this algorithm can be sufficiently accurate and advantageous when speed is important.

Duff (1976) gives an integral representation technique for calculating general multivariate probabilities. He applies it to the multivariate t with common denominator. The algorithm requires knowledge of the characteristics functions of the distribution in question and its marginals.

Tan and Wong (1978) provide a finite series approximation in terms of Laguerre polynomials for central and non-central multivariate gamma p.d.f.s. Use of the approximation requires knowledge of all the mixed moments.

1.3 Crossover Designs: Theory and Applications

All crossover designs are repeated-measure designs: treatments are applied sequentially to each study unit. A repeated-measure design is a CO design if at least one treatment sequence contains two or more distinct treatments. In CO designs each study unit receives one treatment during each time period. The consecutive periods are usually separated by washout time in order to minimize treatment carryover from one period to the next. According to the selected design, each study unit receives one of the possible treatment sequences. Often the selected design includes only a subset of the possible sequences. Conventionally, CO (v,s,p) denotes a crossover design with v treatments, s sequences and p periods. For the CO(2,s,2) design there are four possible
sequences of treatments: A then B, B then A, A then A again, B then B again. The CO(2,2,2) design with sequences (A,B) and (B,A) can be equated to a second-order latin square design. In order to avoid introducing sequence effects into the experiment, the units of study are usually assigned at random to the sequences. Those assigned to the jth sequence form a group of $n_j$ units of study. Balance of the design (i.e., $n_1 = n_2 = \ldots$) will not be required here. It will be assumed that treatment carryover may affect responses in spite of washout time between periods. Such residual effects are characterized by the nature of their duration and decay.

Crossover designs can be applied in a great variety of situations. For example, experiments in business, marketing, engineering, psychology, and nutrition can be appropriate. The earliest applications were made in agriculture; e.g., crop rotation, livestock feeding, etc. These applications continue with success. The most controversial and perhaps most popularly made application has been in clinical trials.

In clinical trials it is often difficult to organize and control subjects and experimentation sufficiently to meet all the assumptions of the usual analysis of CO experiments. Therefore it can be difficult to know in advance whether the most efficient design will be a CO or a completely randomized (CR) design, for example. Frequently the number of study units in the trial is minimal. This can happen because of budget restriction and/or because only a few eligible subjects are available. It may happen
that the investigator cannot randomize the subjects to the treatment sequences. Sequence effects must then be accounted for.

Kershner (1979) has compiled an extensive bibliography on crossover designs. He also surveys in detail the large family of CO designs and their analyses. Currently there is some controversy over the proper application and analysis of CO (2,s,p) designs. Kershner (1979) and Kershner and Federer (1981) examine the difficulties and propose a full-rank general linear univariate model as a solution for a large class of designs. They present comparisons of estimators, some of which are inappropriate because of imprecise interpretations of parameters in similar but distinct models. They also discuss three-period two-treatment designs, CO(2,s,3).

Crossover designs came into use soon after the introduction of latin squares by Fisher in 1925. Brandt (1938) introduced three-period designs. Cochran and Autrey and Cannon (1941) demonstrated the first CO experiment with carryover effects. Williams (1949) extended their work by constructing balanced designs. Patterson (1951) noted that some CO designs can be viewed as factorial experiments, and that repeating the second period altered efficiencies of estimators of effects. Patterson and Lucus (1959, 1962) were first to discuss the serious estimability problems that can arise in CO (2,s,2) designs. Grizzle (1965) discusses the conventional analysis of the CO (2,2,2) design and its associated estimability problems.
Some difficulties with CO(2,s,2) designs remain unresolved. Brown (1980) has urged that CO(2,s,2) designs (as analyzed conventionally) should never be used if carryover effects exist in the experiment. Helms, Kempthorne and Shen (1980) have exposed the nature of the confounding in CO(2,s,2) and CO(2,s,3) designs. They point out that analysis is complicated by several difficulties: the analysis often must account for confounding; customary analyses by less-than-full-rank ANOVA models adds to the complexity; in unbalanced designs it is especially important to give considerable attention to the meaning of model parameters in order to properly interpret results. They also note that the conventional model is a mixed model with uncertainties about proper analysis. They demonstrate a new analysis which is based on the idea that it is advantageous to perform a separate full-rank linear model analysis on each period and then combine results. By this approach confounding is reduced and interpretation of results is simplified. It is, however, necessary to conduct the experiment according to the CO(2,4,2) design. Fortunately balance is not required so that the numbers of study units assigned to sequences (A,A) and (B,B) may be small.

The advantages of the "by-period" strategy proposed by Helms, Kempthorne and Shen (1980) are several: Confounding is reduced, interpretation of results is simplified and improved, messy data can be analyzed without discarding any observations, period*treatment interaction is allowed in the model, and carryover effects do not
effect the validity of the analysis since they too are allowed.

The by-period analysis involves separate tests of the form

$$H_{0i}: \tau_i = 0 \text{ vs } H_{ai}: \tau_i \neq 0 \quad (i=1,2)$$

In order to test

$$H_0: (\tau_1 = 0) \cap (\tau_2 = 0) \text{ vs } H_a: (\tau_1 \neq 0) \cup (\tau_2 \neq 0)$$

the Bonferroni inequality can be used as follows. If at least one of the separate two-sided t-tests is significant at the $\alpha/2$ level then reject $H_0$. The test statistic is then $\max |t_i|$ where $t_1$ and $t_2$ are the individual t-statistics. For this test, the true probability of type I error is $\alpha^*$ which is less than $\alpha$. The difference is

$$(\alpha - \alpha^*) = \Pr[|t_1| > t(\nu_1, \frac{\alpha}{4}) \cap |t_2| > t(\nu_2, \frac{\alpha}{4})]$$

where $t(\nu, \frac{\alpha}{4})$ is defined by

$$1 - \frac{\alpha}{4} = \Pr[t_\nu < t(\nu, \frac{\alpha}{4})]$$

with $t_\nu$ a student-t variate with $\nu$ degrees of freedom. The power of this test procedure depends on the noncentral distribution of the two $t$-statistics, $(t_1, t_2)$.

$$1 - \text{Power} = \int_{-b}^{a} \int_{-a}^{b} f(t_1, t_2) \, dt_1 \, dt_2$$

where $a = t(\nu_1, \frac{\alpha}{4})$, $b = t(\nu_2, \frac{\alpha}{4})$ and $f(\cdot, \cdot)$ is the bivariate p.d.f. for the case of $(\tau_1 \neq 0) \cup (\tau_2 \neq 0)$.

The distribution of $(t_1, t_2)$ depends on the parameters of the GLUMS for the two periods: It is assumed that

$$E[Y_i] = X_i \ast \beta_i \quad (i=1,2)$$

$$n_i x_{pi} n_i x_{p_i} \quad p_i x_i$$
and

\[ V \left( \begin{array}{c} Y_1 \\ N_1 x_1 \end{array} \right) = \sigma_i^2 I_{N_1}. \]

Some \((Y_1, Y_2)\) pairs are assumed and have covariance \(\sigma_1 \sigma_2 \rho\). The exact definition of \(X_i\) can take many forms. Here it is sufficient to know that \(X_i\) is a full-rank matrix. In the case of \(N_1 = N_2 = N\) with all \(N\) study units observed in both periods, it happens that \(X_1\) and \(X_2\) may be identical. In the analysis of complete data it will often happen that this is true and is no restriction. It is never required that \(\hat{\theta}_1 = \hat{\theta}_2\). More generally, it may happen that there is missing data or that covariable adjustment is desired in one or both periods. Then \(X_1\) and \(X_2\) may have different dimensions and entries.

In each period, secondary parameters \(\hat{\theta}_i = \zeta_i \hat{\theta}_1\) will be defined and the hypothesis \(H_0: \theta_1 = 0\) will be tested against a one- or two-sided alternative. Estimate \(\hat{\theta}_i = M_i Y_i\) is defined by \(M_i = \zeta_i [X_i^T X_i]^{-1} X_i\). Without loss of generality we may assume that \(\zeta_i\) is scaled so that \(M_i M_i^T = 1\). The numerator of the \(t\)-statistic is \(\hat{\theta}_i\). The denominator is the square-root of a sum of squares for error. The distribution of \((\hat{\theta}_1, \hat{\theta}_2)\) is bivariate normal. The distribution of the pair of sums of squares is bivariate gamma. Relaxation of the restriction that \(X_1 = X_2\) complicates the deviation of the bivariate gamma distribution.
CHAPTER 2

BIVARIATE-\( t \) DISTRIBUTION THEORY FOR POWER COMPUTATIONS:

THE CASE OF IDENTICAL DESIGN MATRICES IN THE
LINEAR MODELS FOR PERIODS 1 AND 2

2.1 Summary

In this chapter distribution theory results are obtained which are necessary for power calculation for the hypothesis tests made in the by-period analysis of \( CO(2,4,2) \) designs where design matrix \( X_1 \) equals \( X_2 \). The theoretical results culminate in two algorithms for computing \( \Pr[|t_1| \leq t \text{ and } |t_2| \leq t] \) where \( t_i (i=1,2) \) are \( t \)-statistics which have dependent numerators and dependent denominators. These algorithms are used in Chapter 4 as tools for comparing analyses on the basis of power against alternative hypotheses.

2.2 Formulation of the Power Function \( P(t) \)

In order to evaluate the power of general linear hypothesis tests based on two \( t \)-statistics and the Bonferroni inequality it is necessary to express the power function in a form suitable for numerical computation. For two-sided tests this means computing \( P(t) = \Pr[|t_1| \leq t \text{ and } |t_2| \leq t] \) where \( t_1 \) and \( t_2 \) are the
t-statistics. The t-statistics are assumed to be constructed from data which follow a bivariate linear model with mean \( \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \). The resulting joint distribution of \([t_1, t_2]\) in its general form is quite complicated. It is, however, possible to derive the joint distribution of \([t_1, t_2, x_1, x_2]\) where \(x_i\) is proportional to the denominator of \(t_i\) \((i=1,2)\). Integration with respect to \(\chi(2x1)\) over \(R^2\) and integration with respect to \(t(2by1)\) over the square region \(t_i \in [-r, r]\) \(i=1,2\) yields \(P(t)\). An expression for \(P(t)\) is given in Theorem 2.1. The proof of Theorem 2.1 is constructed from a series of results given by Lemmas 2.1-2.3.

Lemma 2.1 sets forth the assumed distribution of the data and the resulting distributions of the r.v.s. which form the numerators and denominators of the t-statistics in the most general case.

In Lemma 2.2 it is additionally assumed that the two design matrices \(X_1\) and \(X_2\) are identical and concomitantly that there is no missing data, i.e., all the observations occur in pairs. The joint distribution of the error-sums-of-squares \([S_1^2, S_2^2]\) is shown to be a bivariate gamma with known p.d.f.

Lemma 2.3 gives the resulting joint p.d.f. of \(|t_1, t_2, x_1, x_2|\) where \(x_i = [S_i^2/\sigma_i^2(1-\rho^2)]^{1/2}\).

The combined results from the proof of Theorem 2.1. The resulting expression for \(P(t)\) is useful and can be evaluated numerically by several methods which are discussed later.
Lemma 2.1:

Let

\[
\begin{bmatrix}
  Y_1 \\
  Y_2
\end{bmatrix}
\begin{bmatrix}
  X_1 \\
  0
\end{bmatrix}
= \begin{bmatrix}
  \beta_1 \\
  \beta_2
\end{bmatrix},
\]

with \( X_i \) (1 by \( n_i \)) by \( r_i \)

\[
\Sigma = \begin{bmatrix}
  \sigma_1^2 \I_{n_1} & \sigma_1 \sigma_2 \rho \I_{n_1} \\
  \sigma_1 \sigma_2 \rho \I_{n_1} & \sigma_2^2 \I_{n_2}
\end{bmatrix}
\]

where

\[
B(n_1 \text{ by } n_2) = [I_{q_1}, 0] \begin{bmatrix}
  \Sigma \\
  \Sigma
\end{bmatrix}
\begin{bmatrix}
  I_{q_1} \\
  0
\end{bmatrix}
\]

\( P_i = I_{n_1} - X_i [X_i^T X_i]^{-1} X_i^T \)

which has rank \( v_i \) \((v_2 \leq 1)\);

\( M_i \) (1 by \( n_i \)) = \( C_i [X_i^T X_i]^{-1} X_i^T \), such that \( M_i M_i^T = I \)

\( M_i B M_i^T = M \);

and \( \hat{\beta} = \begin{bmatrix}
  C_1 \\
  C_2
\end{bmatrix} \begin{bmatrix}
  \beta_1 \\
  \beta_2
\end{bmatrix} \).

Then \( \hat{\gamma} = \begin{bmatrix}
  M_i \\
  C_2
\end{bmatrix} \begin{bmatrix}
  Y_1 \\
  Y_2
\end{bmatrix}\)

\( S_i^2 = \frac{Y_i^T P_i Y_i}{\nu_i} \begin{bmatrix}
  \sigma_1^2 \\
  \sigma_2^2
\end{bmatrix} \begin{bmatrix}
  \sigma_2 \rho M \\
  \sigma_2
\end{bmatrix} \)

and the particular bivariate chi-square distribution of

\( [S_1^2, S_2^2] \) is independent of \( \hat{\beta} \).

Proof of Lemma 2.1:

The application of well-known multinormal distribution theory yields the distributions of \( \hat{\beta} \) and of \( S_1 \) and \( S_2 \). Since \( S_i^2 \) is a function of the data which is orthogonal to both \( \hat{\beta}_i \) and \( \hat{\beta}_2 \), the two bivariate distributions must be independent.
Lemma 2.2

With the conditions of Lemma 2.1, if $X_1 = X_2 = X(n$ by $\rho)$, $B = I_n$,

$v = n - \rho = \text{trace}(I - X[X^T X]^{-1}X^T)$ so that $M = C_i [X^T X]^{-1} C_2$

then
1.) $S_1^2 \overset{d}{=} \sigma_1^2 \chi_v^2$ i=1,2 $E[S_1^2] = \sigma_1^2 \nu$, $V[S_1^2] = \sigma_1^4 \nu$

2.) $\text{COV}[S_1^2, S_2^2] = 2\sigma_1^2 \sigma_2^2 \rho \nu$

3.) the joint p.d.f. of $S_1^2$ and $S_2^2$ is

$$f(S_1^2, S_2^2) = \sum_{k=0}^{\infty} c_k \cdot \chi_2^v \left( \frac{S_1^2}{\nu^2} \right) \cdot \chi_2^{v+2k} \left( \frac{S_2^2}{\sigma_1^2 (1-\rho^2)} \right)$$

where $C \equiv \left[ \frac{\Gamma \left( \frac{v+2k}{2} \right) \rho^{2k} (1-\rho^2)^{v/2}}{\Gamma \left( \frac{v}{2} \right) (k)!} \right]$, $k=0,1,...$ are terms from

the negative binomial expansion of $\left( \frac{1}{1-\rho^2} - \frac{\rho^2}{1-\rho^2} \right)^{n-1}$.

Proof of Lemma 2.2

The outline for this proof is found in Johnson and Kotz (1972), p. 221. The joint distribution is a mixture of joint distributions since the sum of the weights, $[C_k]$, is unity. The covariance of the two quadratic forms $S_1^2$ and $S_2^2$ can be derived using the procedures developed by Evans (1969), (See Searle (1971), p. 65).

$$\text{COV}(X_1^T A X_1, X_2^T A X_2) = 2 \cdot \text{trace}(A C A C) + 4 \nu_1^T A \zeta A \nu_2$$

where $\zeta = E[(X_1 - \nu_1)(X_2 - \nu_2)^T]$. For the case of $S_1^2$ and $S_2^2$, ...
\( C = \sigma_1 \sigma_2 \rho I_n \) and \( u_1^T \omega = 0 = A u_2 \). Thus

\[
\text{COV}(S_1^2, S_2^2) = 2 \cdot \sigma_1^2 \sigma_2^2 \rho^2 \cdot \text{Trace}(I - X \text{[}X^T X\text{]}^{-1} X) = 2 \sigma_1 \sigma_2 \rho^2 \nu. 
\]

The coefficient of correlation follows since the moments of the marginal \((X_0^2)^2\) distributions are known; in particular,

\[
E[S_1^2] = \sigma_1^2 \nu \text{ and } V[S_1^2] = \sigma_1^4 \nu.
\]

The joint p.d.f. of \([S_1^2, S_2^2]\) is of the form

\[
f(S_1^2, S_2^2) = \sum_{k=0}^{\infty} a_k f_{1,k}(S_1^2) f_{2,k}(S_2^2)
\]

where \( \sum_{k=0}^{\infty} a_k = 1 \), \( a_k > 0 \) and \( f_{i,k}(\cdot) \) is a gamma density function.

Hence the joint distribution of \([S_1^2, S_2^2]\) is a mixture of bivariate distributions within each of which the marginal distributions are independent. A sample from the joint distribution of \([S_1^2, S_2^2]\) can be obtained by sampling from the distribution of \( f_{1,k}(S_1^2) \cdot f_{2,k}(S_2^2) \) with probability \( a_k \). It is also interesting to note that the mixed moments, \( E[S_1^{2a} S_2^{2b}] \), are easily computed as a weighted sum of moment products.
Lemma 2.3

With the conditions of Lemma 2.2, if \( X_1 = \left[ \frac{S_1^2}{\sigma_1^2} \right]^{\frac{1}{2}} \left( 1 - \rho_1^2 \right)^{\frac{1}{2}} \)

and

\[
\begin{bmatrix}
T_1 \\
T_2
\end{bmatrix} = \begin{bmatrix}
S_1^2 / \nu & 0 \\
0 & S_2^2 / \nu
\end{bmatrix}^{-\frac{1}{2}} \begin{bmatrix}
\theta_1 \\
\theta_2
\end{bmatrix}
\]

with \( A = \begin{bmatrix} 1 - \rho M & (1 - \rho_2^2) / \nu \left( 1 - \rho_2^2 \right) M \end{bmatrix} \), \( |A| = \left( \frac{(1 - \rho_2^2)}{\nu (1 - \rho_2^2) M} \right)^{\frac{1}{2}} \), and

\[
a = \begin{bmatrix} \theta_1 / \sigma_1 \\
\theta_2 / \sigma_2
\end{bmatrix} \begin{bmatrix} \nu \\
((1 - \rho_2^2)^2)
\end{bmatrix}^{\frac{1}{2}}
\]

then the joint p.d.f. of \([T_1, T_2, X_1, X_2]\) is

\[
f(t', x') = \frac{\left( 1 - \rho_2^2 \right)^{\nu/2} |A|^{1/2}}{\pi^{\nu/2} \Gamma(\nu/2)} \cdot \sum_{k=0}^{\infty} \frac{\left[ \begin{bmatrix} t_1 & 0 \\
0 & t_2
\end{bmatrix} \right]^{2k} \left[ X_1 X_2 \right]^{\nu+2k}}{\Gamma((\nu+2k)/2) \Gamma((2+2k)/2)}
\]

\[
\cdot f(t, x) \exp^{-\frac{1}{2}X'X}
\]

where \( f(t; x) = \exp \left[ -\frac{1}{2} \begin{bmatrix} t_1 & 0 \\
0 & t_2
\end{bmatrix} \begin{bmatrix} X - a \\
X - a
\end{bmatrix} \right] \begin{bmatrix} t_1 & 0 \\
0 & t_2
\end{bmatrix} \begin{bmatrix} X - a \\
X - a
\end{bmatrix} \]

The parameters of the joint distribution are

\[
\frac{\theta_1}{\sigma_1}, \frac{\theta_2}{\sigma_2}, \nu, \rho, \text{ and } M.
\]

Proof of Lemma 2.3:

By independence the joint p.d.f. can be written as

\[
f(\hat{\theta}_1, \hat{\theta}_2, S_1^2, S_2^2) = f(\hat{\theta}_1, \hat{\theta}_2) \cdot f(S_1^2, S_2^2)
\]

The joint p.d.f. of the t-distributed r.v.'s and their denominators can be obtained and simplified using the following transformation:
\[
T_1 = \frac{\hat{b}}{(s_1^2/\nu)^{\frac{1}{2}}} \text{ so } \hat{a}_1 = x_1 T_1 \hat{a}_1 (\nu/(1-\rho^2)^{\frac{1}{2}})
\]
\[
T_2 = \frac{\hat{b}}{(s_2^2/\nu)^{\frac{1}{2}}} \text{ so } \hat{a}_2 = x_2 T_2 \hat{a}_2 (\nu/(1-\rho^2)^{\frac{1}{2}})
\]
\[
x_1 = (s_1^2/\sigma_1^2 (1-\rho^2))^{\frac{1}{2}} \text{ so } s_1^2 = x_1^2 \sigma_1^2 (1-\rho^2)
\]
\[
x_2 = (s_2^2/\sigma_2^2 (1-\rho^2))^{\frac{1}{2}} \text{ so } s_2^2 = x_2^2 \sigma_2^2 (1-\rho^2)
\]

The Jacobian of this transformation is
\[
J = \frac{\sigma_1 \sigma_2 (1-\rho^2)}{\nu} \left[ x_1 x_2 \right]^2 4\sigma_1 \sigma_2^2 (1-\rho^2)^2.
\]

The desired p.d.f. is then
\[
f_{t_1,x_1, t_2,x_2} = f_0(x_1 t_1 (\sigma_1^2 (1-\rho^2)/\nu)^{\frac{1}{2}}, x_2 t_2 (\sigma_2^2 (1-\rho^2)/\nu)^{\frac{1}{2}})
\]
\[
\cdot \frac{\sigma_1 \sigma_2 (1-\rho^2)}{\nu} \cdot f_S(x_1^2 \sigma_1^2 (1-\rho^2), x_2^2 \sigma_2^2 (1-\rho^2))
\]
\[
\cdot [x_1 x_2]^2 4\sigma_1 \sigma_2^2 (1-\rho^2)^2
\]

Define \( \hat{a}^T = [\hat{\theta}_1/\sigma_1, \hat{\theta}_2/\sigma_2] \cdot [\nu/(1-\rho^2)]^{\frac{1}{2}} \) and

\[
A = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \cdot \hat{a}^{-1} \cdot \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \cdot \left(\frac{1-\rho^2}{\nu}\right) \text{ with}
\]

\[
|A|^{\frac{1}{2}} = |\hat{a}|^{-\frac{1}{2}} \cdot \frac{\sigma_1 \sigma_2 (1-\rho^2)}{\nu}
\]

Then
\[
f_{t_1,x_1, t_2,x_2} = f_{N_2(a', A^{-1})}(x_1 y, x_2 t_2) \cdot
\]
\[
\cdot f_s(x_1^2 \sigma_1^2 (1-\rho^2), x_2^2 \sigma_2^2 (1-\rho^2)) \cdot [x_1 x_2]^2 4\sigma_1 \sigma_2^2 (1-\rho^2)^2
\]
\[ f(t_1, t_2, x_1, x_2) = f_{\Lambda_2(a, \Lambda^{-1})}(x_1 t_1, x_2 t_2) \cdot 4[x_1 x_2]^2 \]

\[ = \sum_{k=0}^{\infty} C_k f_2(x_1^2) \frac{\sigma_1^2 \sigma_2^2 (1-\rho^2)}{\sigma_1^2 \sigma_2^2 (1-\rho^2)} \left( f_2(\frac{(x_2^2)^2}{x_2^2 + 2k}) \right) \left( f_2(\frac{(x_1^2)^2}{x_1^2 + 2k}) \right)^{\frac{(x_1^2 + x_2^2)^2}{x_1^2 + x_2^2}} \]

\[ = \sum_{k=0}^{\infty} 2^2 C_k f_2(x_1^2) \frac{\sigma_1^2 \sigma_2^2 (1-\rho^2)}{\sigma_1^2 \sigma_2^2 (1-\rho^2)} \left( \frac{\Gamma^2(\frac{v+2k}{2})}{\Gamma^2(\frac{v+2k}{2}) \Gamma(k+1)} \right)^v \frac{\Gamma^2(\frac{v}{2})}{\Gamma^2(\frac{v}{2}) \Gamma(k+1)} \left( \frac{1-\rho^2}{2} \right)^{\frac{v}{2}} \frac{e^{-\frac{v}{2}(x_1^2 + x_2^2)}}{\Gamma(\frac{v}{2})} \]

\[ = \sum_{k=0}^{\infty} \frac{\Gamma^2(\frac{v+2k}{2})}{\Gamma^2(\frac{v+2k}{2}) \Gamma(k+1)} \left( \frac{1-\rho^2}{2} \right)^{\frac{v}{2}} \frac{e^{-\frac{v}{2}(x_1^2 + x_2^2)}}{\Gamma(\frac{v}{2})} \]

\[ \cdot f(t; \xi) \quad e^{-\frac{v}{2}[\xi^T \xi]} \]

where \( f(t; \xi) = \exp \left[ -\frac{1}{2} \left( \begin{array}{cc} t_1 & 0 \\ 0 & t_2 \end{array} \right) \xi - a \right] \Lambda \left( \begin{array}{cc} t_1 & 0 \\ 0 & t_2 \end{array} \right) \xi - a \right] \]
It is easy to show that the series

\[ S = \sum_{k=0}^{\infty} \frac{[\rho/2]^{2k} [X_1 X_2]^{\nu+2k}}{\Gamma(\frac{\nu+2k}{2}) \Gamma(\frac{2+2k}{2})}, \quad X_i \in [0, \infty) \]

converges. The partial sums are non-decreasing and bounded above by \( S_n = \sum_{k=0}^{n} \frac{[X_1 X_2]^{\nu+2k}}{\Gamma(\frac{2+2k}{2})} \). The limit of the ratio \( U_{k+1}/U_k \) is zero. Therefore by the ratio test \(|S_n|\) is a convergent sequence. It follows that series \( S \) converges.

For the special case of \( \beta = 0 \) and \( \rho = 0 \), we have

\[ f(t_1, t_2, x_1, x_2) = f(t_1, x_1) f(t_2, x_2) \]

and integration over \( x_1 \) verifies that \( f(t_i) = (1 + t_i^2 / \nu)^{-\frac{1}{2}(\nu+1)} / \sqrt{\beta(\nu, \nu/2)}. \)

---

**Theorem 2.1**

With the conditions of Lemma 2.2 the probability,

\[ P(t) = \int_{-t}^{t} \int_{-t}^{t} f(t_1, t_2) \, dt_1 \, dt_2 = \Pr[|t_1^*| \leq t \text{ and } |t_2^*| \leq t] \]

can be computed as

\[ P(t) = \left[ \frac{(1-t^2)^{\frac{\nu}{2}}}{\pi^{\nu-1} \Gamma(\frac{\nu}{2})} \right] |A|^{\frac{1}{2}} \sum_{k=0}^{\infty} \frac{(\frac{1}{2})^{2k} t^k \Gamma(\frac{\nu+2k}{2})}{\Gamma(\frac{2+2k}{2})} \]

\[ = \int_{0}^{\infty} \int_{0}^{\infty} [X_1 X_2]^{\nu+2k} \gamma_X(t) e^{-\frac{1}{2} [X^T X]} \, dx_1 \, dx_2 \]

where \( \gamma_X(t) = \int_{-t}^{t} \int_{-t}^{t} F(t; x) \, dt_1 \, dt_2 \)

and \( F(t; x) = \exp \left[ -\frac{1}{2} \begin{bmatrix} t_1 & 0 \\ t_2 & 0 \end{bmatrix} X - a \right] \)
Proof of Theorem 2.1:

Lemma 2.3 gives $f(t_1, t_2, x_1, x_2)$ as the joint p.d.f. for $t_i \in (-\infty, \infty)$ and $x_j \in [0, \infty)$. The bivariate density $f(t_1, t_2)$ is obtained by integrating $f(t, x)$ with respect to $x$ over the region defined by $x_1 > 0, x_2 > 0$. $P(t)$ is then obtained by integrating $f(t)$ over the square region $|t_i| \leq t, i = 1, 2$. Since $f(t, x)$ is a p.d.f. the order of integration may be changed. $P(t)$ is a function of $t$ having parameters $\frac{\theta_1}{\theta_1}, \frac{\theta_2}{\theta_2}, \nu, \rho, \text{ and } M$. Alternative expressions for the joint p.d.f. $f(t, x)$ do not yield a more useful result. For example, $f(t, x)$ can be expressed in terms of bivariate normal p.d.f.s with respect to $t$ or with respect to $x$. For example if $f^*(x)$ is the p.d.f. of the $N_2(\mathbf{D} \ast \mathbf{d}, \mathbf{A})$ distribution with $\mathbf{D}, \mathbf{d}, \text{ and } \mathbf{A}$ certain matrices involving $t$, then the p.d.f. of $t$ is

$$f(t) = C \cdot \sum_{k=0}^{\infty} W_k \int_{0}^{\infty} \int_{0}^{\infty} [X_1 X_2]^\nu [X_1 X_2]^2 \cdot f^*(x) \, dx_1 \, dx_2$$

where

$$C = \frac{\sigma_1^2 \sigma_2^2 (1-\rho^2)^{\frac{\nu+2}{2}}}{\nu \cdot 2^{\nu-2} \Gamma(\frac{\nu}{2})} \left| \Sigma_0 \right|^{\frac{1}{2}} \cdot \left| \mathbf{A} \right|^{\frac{1}{2}} \left| \mathbf{I} \right|^{-\frac{1}{2}} \mathbf{d}^T \mathbf{D} \mathbf{d}$$

and

$$W_k = [\nu/2]^{2k} / \Gamma\left(\frac{\nu+2k}{2}\right) \cdot k!$$

Thus the joint p.d.f. of $[t_1, t_2]$ can be expressed in terms of incomplete moments of the $N_2(\mathbf{D} \ast \mathbf{d}, \mathbf{A})$ distribution (or absolute moments if $\nu$ is even). A useful recurrence relationship for
such moments is not available. It is possible to express the incomplete moments as an iterated integral:

\[
\int_0^\infty \int_0^\infty [x_1x_2]^{\nu+2k} f^*(x_1, x_2) \, dx_1 \, dx_2
\]

\[
= \int_0^\infty x_2^{\nu+2k} f^*(x_2) \left[ \int_0^\infty x_1^{\nu+2k} f^*(x_1 | x_2) \, dx_1 \right] \, dx_2
\]

The equation \([2\pi\sigma^2]^{-1/2} \int_0^\infty e^{-\frac{1}{2}(x-\mu)^2/\sigma^2} \, dx = M(k)\]
can be solved recursively to obtain \(M(k) = a_k + b_k M(0)\).

Unfortunately \(a_k\) and \(b_k\) are known but algebraically involved functions of \(\sigma\) and \(\mu\). This renders the iterated integral approach cumbersome. Use of the \(H_{2k}(\cdot)\) function is equally cumbersome.

### 2.3 A MacLaurin Series Expansion

Theorem 2.1 provides an expression for \(P(t)\) in terms of

\[
g_X(t) = \int_{-t}^t \int_{-t}^t \exp\left[-\frac{1}{2} \begin{bmatrix} t_1 & 0 \\ 0 & t_2 \end{bmatrix} \begin{bmatrix} x_1-a_1 \\ x_2-a_2 \end{bmatrix} \right] dt_1 dt_2
\]

By expressing \(g_X(t)\) as a MacLaurin polynomial in \(t\) it is possible to reduce \(P(t)\) to terms of the form

\[
\int_0^\infty \int_0^\infty x_1^r x_2^s e^{-\frac{1}{2}(x_1^2 + x_2^2)} \, dx_1 \, dx_2
\]

which can be easily integrated. \(P(t)\) is then reduced to an algebraically involved but discrete summation of terms. Theorem 2.2 gives an expression for the coefficients of the MacLaurin polynomial in \(t\). The proof of Theorem 2.2 is constructed from
Lemmas 2.4 - 2.10.

Lemma 2.4 gives the kth derivative of \( \int_0^t F(t,u)du \) with respect to \( t \) when \( F \) is integrable. Lemma 2.5 gives the kth derivative in the case of \( F(t,u) = \int_0^t f(v,u)dv \). Lemma 2.6 gives the kth derivative of \( g(t) = \int_{-t}^t \int_{-t}^t f(v,u)dv \) du. Lemma 2.7 provides a recurrence equation needed for Lemma 2.8. Lemma 2.8 gives the kth derivative with respect to \( t \) of \( f(t) = \exp(g(t)) \). Lemma 2.9 adds the assumption that \( f(v,u) \) is of the form \( \exp[q] \) with \( q \) a quadratic form in \( u \) and \( v \). The resulting derivatives are used in Lemma 2.10 to obtain the derivatives needed to compute the coefficients of the Maclaurin polynomial in \( t \). Theorem 2.2 uses the results of Lemmas 2.4-2.10 in order to compute the coefficients.

Finally a corollary to Theorem 2.2 shows that the end result can be generalized from integration over a square region to integration over a rectangular region.
Lemma 2.4

\[
\frac{\partial^{k-1}}{\partial t^{k-1}} \int_0^t F(t,u) \, du = \sum_{i=0}^{k-2} \frac{\partial^{i}}{\partial t^{i}} \left[ \frac{\partial}{\partial t} F(t,u) \right]_{u=t} + \int_0^t \frac{\partial^{k-1}}{\partial t^{k-1}} F(t,u) \, du
\]

where \( \frac{\partial^0}{\partial t^0} F(t,u) = F(t,u) \) by definition.

Proof of Lemma 2.4:

The equation is true for \( k=2 \); i.e.,

(1.1.1) \( \frac{\partial}{\partial t} \int_0^t F(t,u) \, du = \left[ F(t,u) \right]_{u=t} + \int_0^t \frac{\partial}{\partial t} F(t,u) \, du \)

Equation (1.1.1) can be used to compute \( \frac{\partial^2}{\partial t^2} \int_0^t F(t,u) \, du \) for \( k=3 \).

Suppose the lemma is true for \( k=n \). Then

(1.1.2) \( \frac{\partial^n}{\partial t^n} \int_0^t F(t,u) \, du = \sum_{i=0}^{n-2} \frac{\partial^{n-1-i}}{\partial t^{n-1-i}} \left[ \frac{\partial}{\partial t} F(t,u) \right]_{u=t} + \int_0^t \frac{\partial^{n-1}}{\partial t^{n-1}} F(t,u) \, du \)

and application of equation (1.1.1) gives

(1.1.3) \( \frac{\partial}{\partial t} \int_0^t \frac{\partial^{n-1}}{\partial t^{n-1}} F(t,u) \, du = \frac{\partial^0}{\partial t^0} \left[ \frac{\partial^{n-1}}{\partial t^{n-1}} F(t,u) \right]_{u=t} + \int_0^t \frac{\partial^n}{\partial t^n} F(t,u) \, du \).
Combining (1.1.2) and (1.1.3) shows that the lemma is true for \( k=n+1 \). Thus, by mathematical induction, we conclude that Lemma 1.1 is true for \( k=2,3,\ldots,n,n+1,\ldots \).

**Lemma 2.5**

If \( F(t,u) = \int_0^t f(v,u)\,dv \)

then \( \frac{\partial}{\partial t} \int_0^t F(t,u)\,du = \int_0^t f(t,u)\,du + \int_0^t f(v,t)\,dv \)

\( \frac{\partial^k}{\partial t^k} \int_0^t F(t,u)\,du = \int_0^t \frac{\partial^{k-1}}{\partial t^{k-1}} f(t,u)\,du + \int_0^t \frac{\partial^{k-1}}{\partial t^{k-1}} f(v,t)\,dv \)

\( + \sum_{i=0}^{k-2} \frac{\partial^{k-2-i}}{\partial t^{k-2-i}} \left[ \int_0^t f(t,u) + f(v,t) \right] \bigg|_{u=v=t} \)

Of course, \( \int_0^t F(t,u)\,du = \int_0^t \int_0^t f(v,u)\,dv\,du \).

**Proof of Lemma 2.5**

Part (a.) follows as a direct corollary of equation (1.1.1) of Lemma 2.4. Part (b.) gives the \((k-1)^{th}\) derivative of

\( \int_0^t f(t,u)\,du + \int_0^t f(v,t)\,dv \) and is therefore a direct application of Lemma 1.1.
Lemma 2.6

If \( g(t) = \int_{-t}^{t} \int_{-t}^{t} f(v,u) \, dv \, du \) and \( k = 2, 3, 4, \ldots \)

Then (a.) \( g^{(1)}(t) = \int_{0}^{t} f(t,u) + f(-t,u) + f(t,-u) + f(-t,-u) \, du \)

\( + \int_{0}^{t} f(v,t) + f(-v,t) + f(v,-t) + f(-v,-t) \, dv \)

(b.) \( g^{(k)}(t) = \int_{0}^{t} \frac{\partial^{k-1}}{\partial t^{k-1}} \left[ f(t,u) + \ldots + f(-t,-u) \right] \, du \)

\( + \int_{0}^{t} \frac{\partial^{k-1}}{\partial t^{k-1}} \left[ f(v,t) + \ldots + f(-v,-t) \right] \, dv \)

\( + \sum_{i=0}^{k-2} \frac{\partial^{k-2-i}}{\partial t^{k-2-i}} \int_{0}^{t} f(t,u) + \ldots + f(-t,-u) + f(v,t) + \ldots + f(-v,-t) \left. \, du \right|_{u=v=t} \)

Proof of Lemma 2.6:

The double integral can be evaluated as an iterated integral (see Purcell (1976), p. 777). By changing variables,

\( g(t) = \int_{0}^{t} \int_{0}^{t} f(v,u) + f(-v,u) + f(v,-u) + f(-v,-u) \, dv \, du \) (1.3.1).

Direct application of Lemma 2.5 to equation (1.3.1) gives the above results.
Lemma 2.7

If \( C_{k,i} = \begin{cases} \frac{(k)! \cdot 2^{-i}}{(k-2i)!i!} & \text{if } i \leq k/2 \text{ and } i \in \{1,2,3,...\} \\ 0 & \text{if otherwise} \end{cases} \)

Then \( C_{k+1,i} = C_{k,i} + k \cdot C_{k-1,i-1} \)

and \( C_{k+1,i+1} = C_{k,i} \) if \( i = \frac{k-1}{2} \) and \( k \) is odd

Proof of Lemma 2.7:

This result is easily verified in several ways.

Lemma 2.8

If \( f(t) = \exp[g(t)], t \in (0, \infty) \)

and \( g^{(k)}(t) = \begin{cases} \text{finite if } k=1,2 \\ 0 & \text{if } k \geq 3 \end{cases} \)

Then (1.) \( f^{(k)}(t) = f^{(k-1)}(t) \cdot g^{(1)}(t) + (k-1) \cdot f^{(k-2)}(t) \cdot g^{(2)}(t) \)

(2.) \( f^{(k)}(t) = f(t) \cdot \sum_{i=0}^{k} C_{k,i} [g^{(1)}(t)]^{k-2i} [g^{(2)}(t)]^{i} \)

Note: If \( f(t) = \exp[-\frac{1}{2}t^{2}] \) then \( \frac{f^{(n)}(t)}{f(t)} = (-1)^{n} \cdot H_{n}(t) \), the Hermite Polynomial.

Proof of Lemma 2.8:

Begin with \( f^{(1)}(t) = f(t) \cdot g^{(1)}(t) \). Leibniz' formula is

\[
\frac{d^{k}}{dt^{k}} h_{1}(t) \cdot h_{2}(t) = \sum_{i=0}^{k} \binom{k}{i} h_{1}^{(i)}(t) \cdot h_{2}^{(k-i)}(t)
\]

and can be applied to
\[ f(k+1)(t) = \sum_{i=0}^{k} (k-i)f(k-i)(t) g(i+1)(t) \]
\[ = \frac{1}{2} \sum_{i=0}^{k} (k-i)f(k-i)(t) g(i+1)(t) \quad \text{since} \]
\[ g(i+1)(t) = 0 \quad \text{if} \quad i \geq 2. \quad \text{Thus part (1.) is proven.} \]

Part (2.) is proved inductively by using part (1.). Let \( S = \{ \text{integers } k : \text{part (2) is true} \} \). Then \( 1 \in S \) and \( 2 \in S \) because

\[ f(1)(t) = f(t) \cdot g(1)(t) \]
\[ f(2)(t) = f(t) \cdot [g(1)(t)]^2 + [g(2)(t)] \]

are of the form of part (2.). Assume that \( n \in S \) and \( (n-1) \in S \). Then

\[ f(n+1)(t) = f(t) \left[ \sum_{i=0}^{\frac{n-w}{2}} C_n,i [g(1)(t)]^{n-2i+1} [g(2)(t)]^i \right. \]
\[ \left. + \sum_{j=0}^{\frac{n-2+w}{2}} n \cdot C_{n-1,j} [g(1)(t)]^{n-2j+1} [g(2)(t)]^j \right] \]

by the recurrence relation (part (1.)) already proved. Here, \( w \) is the indicator variable for \( n \) being odd. By considering the two cases "\( n \) is odd" and "\( n \) is even", \( f(n+1)(t) \) can be expressed as

\[ f(n+1)(t) = f(t) \cdot \sum_{i=0}^{\left\lfloor \frac{n+1}{2} \right\rfloor} D_{n+1,i} [g(1)(t)]^{n+1-2i} [g(2)(t)]^i \]

where

\[ D_{n+1} = \begin{cases} 1 = C_{n,0} & \text{if } i = 0 \\ C_{n,i} n \cdot C_{n-1,i-1} & \text{if } i = 1, 2, \ldots, \left\lfloor \frac{n+1}{2} \right\rfloor \\ n \cdot C_{n-1,i-1} & \text{if } i = \frac{n+1}{2} \end{cases} \]
By Lemma 2.7 \( D_{n+1} = C_{n+1} \). Thus \((n+1) \in S\) whenever \(n \in S\) and \((n-1) \in S\). By the axiom of mathematical induction, \( S = \{1,2,3,\ldots\} \).

**Lemma 2.9**

Assume \( f(cv, bu) = \exp[-\frac{1}{2}(cvx_1, bx_2 - a')] A (\frac{cvx_1, bx_2}{2} - a) ]^T a \)

where \( b = \pm 1 \) and \( c = \pm 1 \)

with \( Q_1 = [cx_1, 0] A [cx_1, 0]' \), \( Q_2 = [cx_1, 0] A a \),

\[ Q_3 = [cx_1, 0] A [cx_1, bx_2]', Q_4 = [cx_1, bx_2] A a, \]

\[ Q_5 = [cx_1, bx_2] A [cx_1, bx_2]', m=k-2-i, \text{ and } i=0,1,\ldots,k-2. \]

(1.) Then \( \frac{\partial^m}{\partial t^m} \left[ \frac{\partial^j}{\partial t^j} f(ct, bx) \right] \bigg|_{u=t} \)

\[ \min \{ m, i \} = \sum_{j=0}^{\min \{ m, i \}} \left( \begin{array}{c} m-j \cr j \end{array} \right) \frac{(-1)^{i-j}}{(i-2)!(i-2-j)!} \sum_{l=0}^{i-l} \frac{(-1)^{l-1}}{(i-l)! (i-2-l-j)!} \]

\[ [-Q_2]^{i-l-1}[-Q_3]^j \]

\[ [Q_2-tQ_3]^{i-2l-j}[-Q_3]^j \] \quad \quad (1.6.1)

and \( \frac{\partial^m}{\partial t^m} \left[ \frac{\partial^j}{\partial t^j} f(cv, bt) \right] \bigg|_{v=t} \) is obtained from (1.6.1) above by replacing \([cx_1, 0]\) with \([0, bx_2]\) in the definitions of \( Q_1, Q_2 \) and \( Q_3 \).

(2.) If \( i=0 \) then \( \frac{\partial^m}{\partial t^m} \left[ \frac{\partial^j}{\partial t^j} f(ct, bu) \right] \bigg|_{u=t} = \frac{\partial^m}{\partial t^m} \left[ \frac{\partial^j}{\partial t^j} f(cv, bt) \right] \bigg|_{v=t} \)

\[ \sum_{n=0}^{\lfloor m/2 \rfloor} C_{m,n} [Q_4-tQ_5] \cdot [-Q_5] = \frac{\partial^m}{\partial t^m} f(ct, bt) \] \quad \quad (1.6.2)

(3.) \( \frac{\partial^m}{\partial t^m} \left[ \frac{\partial^j}{\partial t^j} f(ct, bu) \right] \bigg|_{u=t} \bigg|_{t=0} \)

\[ = (-1)^k \cdot \frac{\partial^m}{\partial t^m} \left[ \frac{\partial^j}{\partial t^j} f(-ct, -bu) \right] \bigg|_{u=t} \bigg|_{t=0} \] \quad \quad (1.6.3)

and similarly for \( f(cv, bt) \).
essentially repeated; 
\[ \frac{\partial}{\partial t} \log f(cv, bt) = [0, bx_2] \sum_{j=1}^{m-j} \left[ \frac{\partial^m}{\partial t^m} f(cv, bt) \right] \left[ 0, bx_2 \right] \] 
and the above computation is 

For \( m=1 \), the above computation is 

\[ \frac{\partial}{\partial t} \log f(cv, bt) = [0, bx_2] \sum_{j=1}^{m-j} \left[ \frac{\partial^m}{\partial t^m} f(cv, bt) \right] \left[ 0, bx_2 \right] \] 

Taking the differential operation inside the summation indexed by \( j \) 

The formula (see the proof of Lemma 2.8) as 

It follows that 

\[ \frac{\partial}{\partial t} \log f(cv, bt) = \left[ \frac{\partial}{\partial t} \frac{1}{u} \right] + \frac{1}{u} \sum_{j=1}^{m-j} \left[ \frac{\partial^m}{\partial t^m} f(cv, bt) \right] \left[ 0, bx_2 \right] \] 

and \( \frac{\partial}{\partial t} \log f(cv, bt) = [0, bx_2] \sum_{j=1}^{m-j} \left[ \frac{\partial^m}{\partial t^m} f(cv, bt) \right] \left[ 0, bx_2 \right] \) 

Evaluation at \( u=0 \) gives 

\[ \frac{\partial}{\partial t} \log f(cv, bt) = \left[ \frac{\partial}{\partial t} \frac{1}{u} \right] + \frac{1}{u} \sum_{j=1}^{m-j} \left[ \frac{\partial^m}{\partial t^m} f(cv, bt) \right] \left[ 0, bx_2 \right] \] 

where \( \frac{\partial}{\partial t} \log f(cv, bt) = [0, bx_2] \sum_{j=1}^{m-j} \left[ \frac{\partial^m}{\partial t^m} f(cv, bt) \right] \left[ 0, bx_2 \right] \) 

By Lemma 2.8 part (2), we have 

\[ \frac{\partial}{\partial t} \log f(cv, bt) = \left[ \frac{\partial}{\partial t} \frac{1}{u} \right] + \frac{1}{u} \sum_{j=1}^{m-j} \left[ \frac{\partial^m}{\partial t^m} f(cv, bt) \right] \left[ 0, bx_2 \right] \]
from equation (1.6.1) by replacing \([cx_1, 0]\) with \([0, bx_2]\) in the definitions of \(Q_1, Q_2,\) and \(Q_3\). Thus, part (1.) of Lemma 1.6 is proved.

The proof of part (2.) is trivial: it (equation (1.6.2)) is a result of the direct application of Lemma 2.8--part (2.). The proof of part (3.) is as follows.

By part (2.), 
\[
\left[ \frac{a^{m-j}}{a^j t^{m-j}} f(ct, bt) \right]_{t=0} = f(0,0) \sum_{n=0}^{\left\lceil \frac{m-j}{2} \right\rceil} c_{m-j,n} [Q_4] [-Q_5]^n [Q_4] [-Q_5]^n
\]

and 
\[
\left[ \frac{a^{m-j}}{a^j t^{m-j}} f(-ct, -bt) \right]_{t=0} = f(0,0) \sum_{n=0}^{\left\lceil \frac{m-j}{2} \right\rceil} c_{m-j,n} (-Q_4) [Q_2] \]

\[
= (-1)^{m-j} \left[ \frac{a^{m-j}}{a^j t^{m-j}} f(ct, bt) \right]_{t=0}
\]

By part (1.), 
\[
\left[ \frac{a^m}{a^j t^j} \left[ \frac{a^i}{a^j t^j} f(ct, bu) \right]_{u=t} \right]_{t=0}
\]

\[
= \min_{j=0}^{[m,i]} \left[ \frac{a^{m-j}}{a^j t^{m-j}} f(ct, bt) \right]_{t=0} \cdot \sum_{l=0}^{\left\lceil \frac{i-j}{2} \right\rceil} c_{i,1} (i-2l)! [-Q_1] [Q_2] \]

and 
\[
\left[ \frac{a^m}{a^j t^j} \left[ \frac{a^i}{a^j t^j} f(-ct, -bu) \right]_{u=t} \right]_{t=0} = \min_{j=0}^{[m,i]} \left[ \frac{a^{m-j}}{a^j t^{m-j}} f(-ct, -bt) \right]_{t=0} \cdot \sum_{l=0}^{\left\lceil \frac{i-j}{2} \right\rceil} c_{i,1} (i-2l)! [-Q_1] [Q_2] [Q_3]^{i-j-2l} \]
\[ \begin{aligned}
= \min_{j=0}^{m} (m_j) \cdot (-1)^{m-j} \left[ \frac{a_{m-j}}{a_t} f(ct, bt) \right] \bigg|_{t=0} \cdot (-1)^{i-j}.
\end{aligned} \]

\[ \begin{aligned}
\left[ t^j \right] \\
\cdot \sum_{l=0}^{(1-21)!} C_i, l \left[ \begin{array}{c}
(1-21)! \ \left[ -Q_1 \right] \ [Q_2] \ [Q_3] \end{array} \right]^{i-21-j} \]
\[\begin{aligned}
= (-1)^{m+i} \left[ \frac{a_m}{a_t} \left[ \frac{a_i}{a_t} f(ct, bu) \right] \right] \bigg|_{u=t} \bigg|_{t=0}.
\end{aligned}\]

Since \( m = k - 2 - i \), \( (-1)^{m+i} = (-1)^k \) and equation (1.63) is proved.

Note: If \( Q_4 = 0 \) then \( \left[ \frac{a_{m-j}}{a_t} f(ct, bt) \right] \bigg|_{t=0} = \begin{cases} 
\text{zero if } m-j \text{ is odd} & \\
\text{not necessarily zero} & \\
\text{if } m-j \text{ is even} & 
\end{cases} \)

If \( Q_2 = 0 \) then \( \sum_{l=0}^{(1-21)!} C_i, l \left[ -Q_1 \right] [Q_2] [Q_3]^{i-21-j} \)

Here \( A_{ij} = C_i, l \left[ \begin{array}{c}
(1-21)! \ [Q_1] \ [Q_2] \ [Q_3] \end{array} \right]^{i-21-j} \).
Lemma 2.10

If \( m = k - 2 - i \) and \( i = 0, 1, \ldots, k - 2 \) and \( f(v, u) \) is as given in Lemma 2.9

\[
\left. \left[ \frac{\partial^m}{\partial t^m} \left[ \frac{\partial^i}{\partial t^i} f(t, u) + f(-t, u) + f(v, t) + f(v, t) \right] \right|_{u=v=t} \right|_{t=0}
\]

is

\[
f(0, 0) \cdot \sum_{j=0}^{m} \sum_{i=0}^{i} \sum_{n=0}^{n} \left( \begin{array}{c} m \\ j \end{array} \right) c_{i, 1} \left( \begin{array}{c} i-21 \\ j \end{array} \right) ! \cdot c_{m-i, n} (-1)^{j+n}
\]

\[
\sum_{j=0}^{m} \sum_{i=0}^{i} \sum_{n=0}^{n} \left( \begin{array}{c} m \\ j \end{array} \right) c_{i, 1} \left( \begin{array}{c} i-21 \\ j \end{array} \right) ! \cdot c_{m-i, n} (-1)^{j+n}
\]

Proof of Lemma 2.10:

Define

\[
Q_1(cX_j) = X_j^2 A_{jJ}
\]

\[
Q_2(cX_j) = c \cdot X_j \left( A_{j1} x_1 + A_{j2} x_2 \right)
\]

\[
Q_3(cX_j) = X_j^2 A_{JJ} + cX_j X_1 A_{12}
\]

\[
Q_4(cX_j) = c \cdot X_j \left( A_{j1} x_1 + A_{j2} x_2 \right) + X_1 \left( A_{11} x_1 + A_{12} x_2 \right)
\]

\[
Q_5(cX_j) = X_j^2 A_{JJ} + X_1^2 A_{11} + c \cdot X_j X_1 \cdot 2A_{12}
\]

where \( c = +1 \) and \( (J, I) \in \{(1, 2), (2, 1)\} \)
By Lemma 2.9
\[
\left[ \left[ f(w_1, w_2) \right]_{w_j = ct} \right]_{w_i = t=0} t=0
\]
is
\[
\min_{j=0}^{m} \left\{ \left( \frac{m-j}{2} \right)^{m-j} \right\} \sum_{n=0}^{\frac{m-j}{2}} c_{m-j,n} \left[ \sum_{j=0}^{m} \left( \frac{m-j}{2} \right)^{m-j} \right] \left( \frac{m-j}{2} \right) c_{i,1,1} \left( \frac{m-j}{2} \right) ! \left[ \begin{array}{c}
-Q_1(c x_j) \\
Q_2(c x_j) \\
-Q_3(c x_j)
\end{array} \right]^{j} \left[ \begin{array}{c}
Q_4(c x_j) \\
-Q_5(c x_j)
\end{array} \right]^{n}
\]

\[
= f(0,0) \cdot \sum_{j=0}^{m} \left( \frac{m-j}{2} \right)^{m-j} \sum_{n=0}^{\frac{m-j}{2}} c_{m-j,n} \left( \frac{m-j}{2} \right) c_{i,1,1} \left( \frac{m-j}{2} \right) ! \left[ \begin{array}{c}
-Q_1(c x_j) \\
Q_2(c x_j) \\
-Q_3(c x_j)
\end{array} \right]^{j} \left[ \begin{array}{c}
Q_4(c x_j) \\
-Q_5(c x_j)
\end{array} \right]^{n}
\]

The product of the "Q"-terms is
\[
\left[ (-1)^{j+1+n} \right] \cdot \left[ \sum_{j=0}^{m} \left( \frac{m-j}{2} \right)^{m-j} \right] \cdot \sum_{n=0}^{\frac{m-j}{2}} c_{m-j,n} \left( \frac{m-j}{2} \right) c_{i,1,1} \left( \frac{m-j}{2} \right) ! \left[ \begin{array}{c}
-Q_1(c x_j) \\
Q_2(c x_j) \\
-Q_3(c x_j)
\end{array} \right]^{j} \left[ \begin{array}{c}
Q_4(c x_j) \\
-Q_5(c x_j)
\end{array} \right]^{n}
\]

\[
\cdot \left[ \sum_{p=0}^{j} \left( \frac{j}{p} \right) c_{i,1,1}^{p} x_{1}^{j-p} \right] \cdot \left[ \sum_{q=0}^{m-j-2n} \left( \frac{m-j-2n}{q} \right) c_{i,1,1}^{q} x_{1}^{m-j-2n-q} \right]
\]
\[
\left[ \sum_{r=0}^{n} \binom{n}{r} \sum_{s=0}^{r} \binom{r}{s} \left[ x_j^2 A_{jj} \right]^2 \left[ x_i^2 A_{II} \right]^{-s} \left[ 2^r c x_j x_i A_{12} \right]^{n-r} \right]
\]

which can be written as

\[
(-1)^{j+1+n+i+n-p+q-r-j} \sum_{p=0}^{j} \sum_{q=0}^{m-j-2n} \sum_{r=0}^{n} \binom{j}{p} \binom{m-j-2n}{q} \binom{n}{r} \binom{p}{q} \binom{r}{s}.
\]

\[
\left[ A_{jj} \right] \left[ A_{II} \right] \left[ A_{12} \right] \left[ 2 \right].
\]

\[
\left[ A_{j1} a_1 + A_{j2} a_2 \right] \left[ A_{II} a_1 + A_{I2} a_2 \right] \left[ x_j \right] \left[ x_i \right].
\]

The above result for (J,I) with c = ±1 can be used by noting that

\[
\left[ \frac{\partial^m}{\partial t^m} \frac{\partial^j}{\partial t^j} \right] f(t,u) + f(-t,u) + f(v,t) + f(v,-t) \bigg|_{u=v=t} \bigg|_{t=0} =
\]

\[
= \left[ \frac{\partial^m}{\partial t^m} \frac{\partial^j}{\partial t^j} \right] f(w_1,w_2) \bigg|_{w_1=t} + f(w_1,w_2) \bigg|_{w_1=-t} + f(w_1,w_2) \bigg|_{w_2=t} + f(w_1,w_2) \bigg|_{w_2=-t} \bigg|_{w_1=w_2=t} \bigg|_{t=0}.
\]

The result of Lemma 1.7 is immediate.

Note: Suppose \( Q_2 = 0 = Q_4 \), then terms in the sum indexed by \( j \) ( \( \sum_{j=0}^{\infty} \) ) say, term(j), depends on \( m=2(h-1)-i \) as well as \( j \). If \( Q_2 \) and \( Q_4 \) are both zero then

\[
\text{term(j)} = \begin{cases} 
0 & \text{if } i-j \text{ is odd} \\
\text{not necessary zero if } i-j \text{ is even} &
\end{cases}
\]
Thus for fixed $i$, the sum indexed by $j$ may be computed as

\[
\sum_{j=1}^{J} \text{term}(2j-2) \text{ if } i \text{ is even with } J = \frac{\min[m, i]}{2} + 2
\]

\[
\sum_{j=1}^{J} \text{term}(2j-1) \text{ if } i \text{ is odd with } J = \frac{\min[m, i]}{2} + 1
\]
Theorem 2.2

If \( g(t) = \int_{-t}^{t} f(v,u) dv \ du \)

where \( f(v,u) = \exp[-\frac{1}{2}(v^{T}0 + 0^{T}u)X^{-1}a] \)

\( A \in \mathbb{R}^{2 \times 2} \) is symmetric, and \( X = [X_1, X_2] \) has non-negative entries,
then \( g(t) \) can be expressed as a MacLaurin series:

\[
g(t) = \sum_{k=0}^{N} \frac{g^{(k)}(0) t^k}{k!} + R_N(t).
\]

The coefficients of this polynomial in \( t \) are given by

\[
g^{(k)}(0) = \left. \frac{\partial^k g(t)}{\partial t^k} \right|_{t=0}
\]

\[
= \begin{cases} 
2 \cdot \frac{k-2}{k} \cdot D_i & \text{if } k=2,4,6,\ldots \\
0 & \text{if } k=0,1,3,5,\ldots 
\end{cases}
\]

where

\[
D_i = \left. \frac{\partial^k f(t,u) + f(-t,u) + f(v,t) + f(v,-t)}{\partial t^k} \right|_{u=v=t=0}
\]

\[
= f(0,0) \sum_{j=0}^{i-2} \sum_{m=0}^{j} \sum_{r=0}^{m-j} \sum_{p=0}^{j+1+n} \sum_{q=0}^{m-j-2n} \sum_{s=0}^{n-r} \left[ \begin{array}{c} i+n-p+q-r \\ j+n-p-r \\ 1 \\ 2 \\ [A_{12}] \\ [X_2] \\ [X_1] \end{array} \right] + \left[ \begin{array}{c} i-j-21+q \\ n-j-2n-q \\ i+n+p+q+r+2s \\ m-(n+p+q+r+2s) \\
[A_{11}] \end{array} \right]
\]

\[
+ \left[ \begin{array}{c} i-j-21+q \\ n-j-2n-q \\ i+n+p+q+r+2s \\ m-(n+p+q+r+2s) \\
[A_{22}] \end{array} \right].
\]

and \( m=k-2-i \)
The proof of Theorem 2.2 is easily shown by using the results of Lemmas 2.4-2.10. The integral \( g(t) \) can be differentiated any number of times at all points of the range of \( t, [0, \infty) \). Lemma 2.6 which is derived from Lemmas 2.4 and 2.5 gives the form of the kth derivative of \( g(t) \) as

\[
g^{(1)}(t) = \int_0^t f(t,u) + f(-t,u) + f(t,-u) + f(-t,-u) \, du
\]
\[
+ \int_0^t f(v,t) + f(-v,t) + f(v,-t) + f(-v,-t) \, dv
\]
\[
g^{(k)}(t) = \int_0^t \left[ \frac{\partial^k}{\partial t^k} f(t,u) + \ldots + f(-t,-u) \right] \, du +
\]
\[
+ \int_0^t \left[ \frac{\partial^k}{\partial t^k} f(v,t) + \ldots + f(-v,-t) \right] \, dv +
\]
\[
+ \sum_{i=0}^{k-2} \left[ \frac{\partial^{k-2-i}}{\partial t^{k-2-i}} \left[ \frac{\partial^i}{\partial t^i} f(t,u) + \ldots + f(-t,-u) + f(v,t) + \ldots + f(-v,-t) \right] \right]_{u=v=t}
\]

where \( k = 2, 3, 4, \ldots \) and \( \frac{\partial^0}{\partial t^0} f = f \).

Lemma 2.9 which is based on Lemmas 2.7 and 2.8 derives the form of

\[
\left. \frac{\partial^m}{\partial t^m} \left[ \frac{\partial^i}{\partial t^i} f(ct, bu) \right] \right|_{u,t} \quad \text{and} \quad \left. \frac{\partial^m}{\partial t^m} \left[ \frac{\partial^i}{\partial t^i} f(cv, bt) \right] \right|_{v,t}
\]

where \( c = \pm 1, b = \pm 1, \) and \( m = k-2-i \). Lemma 2.9 also shows that

\[
\left. \frac{\partial^m}{\partial t^m} \left[ \frac{\partial^i}{\partial t^i} f(ct, bu) \right] \right|_{u=t} = (-1)^k \left. \frac{\partial^m}{\partial t^m} \left[ \frac{\partial^i}{\partial t^i} f(-ct, -bu) \right] \right|_{u=t}
\]

(and similarly for \( f(cv, bt) \)). It follows that if \( k \) is odd then
\[ D_i = \left[ \frac{\partial^{k-2-i}}{\partial t^{k-2-i}} \frac{\partial^i}{\partial t^i} f(t, u) + f(-t, -u) + f(t, -u) + f(-t, -u) \right. \]
\[ + f(v, t) + f(-v, -t), + f(-v, t) + f(v, -t) \bigg|_{u=v=t} \bigg| \bigg|_{t=0} \]
is zero for all valid values of \( i \). Thus \( g(k)(0) = 0 \) if \( k \) is odd. When \( k \) is even then
\[ \eta D_i = \left[ \frac{\partial^{k-2-i}}{\partial t^{k-2-i}} \frac{\partial^i}{\partial t^i} f(t, u) + f(-t, -u) + f(v, t) + f(v, -t) \bigg|_{u=v=t} \bigg| \bigg|_{t=0}. \]
Finally, Lemma 2.10 gives an expression for \( D_i \) in terms of \( X_1 \) and \( X_2 \).

Since \( g(t) \) is continuous together with all its derivatives on the interval \([0, \infty)\), \( g(t) \) can be represented as a Taylor polynomial in \( t \). For practical applications \( t \) will be small (e.g., \( t \in (0, 10) \)) and therefore a Taylor series expansion of \( g(t) \) about \( t=0 \) is most appropriate. Thus Theorem 2.2 gives the explicit form of \( g(t) \) as a MacLaurin series.

The remainder term for the MacLaurin series is
\[ R_N(t) = \int_0^t (t-w)^N \frac{g^{(n+1)}(w)}{N!} \, dw, \]
Lagrange's form of the remainder is
\[ R_N(t) = \frac{g^{(N+1)}(c)}{(N+1)!} \, t^{N+1} \]
where \( c \) is an unknown value only restricted by \( c \in [0, t] \). An explicit expression for \( R_N(t) \) in terms of the underlying parameters \( \theta_1, \theta_2, \rho, \nu, M \) and \( t \) is so algebraically involved that no usefully sharp bounds on \( R_N(t) \) could be derived. Preliminary
computer trials were run in order to determine empirically whether or not the series converges numerically. It was discovered that the series is an alternating series and that whenever it does converge it does so quickly and accurately. Unfortunately the interval of convergence is an unknown function of the above parameters.

**Corollary to Theorem 2.2**

\[
\begin{align*}
V_t & \quad V_t \\
\text{If } g(t) &= \int_{-V_t}^{V_t} \int_{-V_t}^{V_t} f(v,u) \, dv \, du \text{ where } U > 0 \text{ and } V > 0 \\
& \quad -U_t \quad -V_t
\end{align*}
\]

with \( z_1 = Vx_1 \) and \( z_2 = Ux_2 \)

then \( g(t) \) can be expressed as a Maclaurin Series:

\[
g(t) = UV \sum_{k=0}^{N} \frac{g^{(k)}(0)}{(k)!} t^k + R_N'(t)
\]

where the definition of \( g^{(k)}(0) \) is as in Theorem 1 but with \([x_1, x_2]\) replaced by \([z_1, z_2]\).

**Proof of the corollary:**

By a simple change of variables \( g(t) = UV \int_{-t}^{t} \int_{-t}^{t} f(v,u) \, dv \, du. \)

Since \( f(v,u) = \exp\left(-\frac{1}{2}\begin{bmatrix} v & 0 \\ 0 & u \end{bmatrix} (X-a)^T A \begin{bmatrix} v \\ 0 \end{bmatrix} (X-a) \right) \), \( f(Vv,Uu) \) may be written as \( f(Vv,Uu) = \exp\left(-\frac{1}{2}\begin{bmatrix} v & 0 \\ 0 & u \end{bmatrix} (Z-a)^T A \begin{bmatrix} v \\ 0 \end{bmatrix} (Z-a) \right) . \) As in Theorem 2.2, the remainder term, \( R_N'(t) \), cannot be adequately expressed in terms of underlying parameters to yield useful information about the interval of convergence.
2.4  Two Algorithms for Computing Power

Theorem 2.2 can be used to provide a computational algorithm for \( P(t) \) as represented in Theorem 2.1. Theorem 2.2 gives a polynomial approximation for \( g(t) \):

\[
g(t) = \sum_{h=1}^{N} \frac{g^{(2h)}(0)}{(2h)!} t^{2h} + R_{2N}(t)
\]

\[
= \sum_{h=1}^{N} \frac{g^{(2h)}(0)}{(2h)!} t^{2h} = g_N(t)
\]

where \( N \) may be taken to be as large as desired.

Thus \( P(t) \) can be approximated by

\[
P(t;N_1,N_2) = \left[ \frac{(1-\beta)^{v/2}}{\pi^{v-1} \Gamma(v/2)} \right] A_1^{1/2} \sum_{k=0}^{N_1} \left( \frac{(v)^{2k}}{\Gamma(v/2)^2 \Gamma(2+2k)} \right)
\]

\[
\int_0^\infty \int_0^\infty [x_1 x_2] g_{N_2}(t) \exp\left(-\frac{t}{\chi_2}\right) dx
\]

The integral in \( P(t;N_1,N_2) \) can be written as

\[
\sum_{h=1}^{N_2} \frac{t^{2h}}{(2h)!} \int_0^\infty \int_0^\infty [x_1 x_2] g^{(2h)}(0) \exp\left(-\frac{t}{\chi_2}\right) dx
\]

where \( g^{(2h)}(0) \) is a polynomial in \( \chi_1 \) and \( \chi_2 \). Therefore, the integrals in the resulting sum are of the form

\[
c \cdot \int_0^\infty \int_0^\infty [x_1]^r [x_2]^s \exp\left(-\frac{t}{\chi_2}\right) dx_1 dx_2
\]

which can be written as the product of two integrals equal to

\[
c \cdot \frac{r+s-2}{2} \Gamma\left(\frac{r+1}{2}\right) \Gamma\left(\frac{s+1}{2}\right).
\]
It has been observed numerically that \( g(2h)(0) \) alternates in sign:

\( \{g^{(2)}(0), g^{(6)}(0), g^{(10)}(0)\ldots\} \) are positive while

\( \{g^{(4)}(0), g^{(8)}(0),\ldots\} \) are negative.

The integral in \( P(t;N_1,N_2) \) can be written as

\[
\sum_{h=1}^{N_2} \left( \frac{t}{2h} \right)^{2h} \cdot 2 \cdot \sum_{i=0}^{2h-2} \sum_{j=0}^{\infty} \left[ x_1 x_2 \right] D_i e^{-\frac{i}{2} \left[ \frac{x}{x} \right]} dx_1 \ dx_2
\]

where the integral can be expressed as

\[
f(0,0) \sum_{j=0}^{m} \sum_{1=0}^{n} \sum_{h=0}^{i+n-p+q-r} \left( \frac{m-j}{2} \right) C_{m-j,n} \left[ -1 \right]^{j+1+n-j+n-p+q-r} \times
\]

\[
\sum_{p=0}^{j} \sum_{q=0}^{m-j-2n} \sum_{r=0}^{n} \sum_{s=0}^{r} \left( \frac{m-j-2n}{2} \right) \left[ -1 \right]^{p+q+r} \times
\]

\[
\left[ A_{12} \right]^{j+n-p-r} + p+r
\]

\[
\left[ A_{11} \right]^{i-j-21+q} \left[ A_{12} a_1 + A_{11} a_2 \right]^{m-j-2n-q}\]

\[
\left[ A_{12} a_1 + A_{11} a_2 \right]^{i-j-21+q} \left[ A_{11} a_1 + A_{12} a_2 \right]^{m-j-2n-q}\}
\]

\[
\times \Gamma \left( \frac{(i+n+p+q-r+2s) + v+2k+1}{2} \right) \times \frac{1}{2} \nn \times \Gamma \left( \frac{m-j-2n-q + v+2k+1}{2} \right)
\]

where \( m = (2h-2-i) \) and \( A_{11} = A_{22} \) for the case in Theorem 2.1.

Thus \( P(t;N_1,N_2) \) is

\[
\left[ f(0,0)(1-\rho) \right]^{v/2} \left[ A_{11} \right]^{\left[ \frac{N_2}{2} \right]} \sum_{k=0}^{\left[ \frac{N_2}{2} \right]} \left( \frac{t}{2h} \right)^{2k} \Gamma \left( \frac{v+2k}{2} \right) \Gamma \left( \frac{2+2k}{2} \right) \sum_{h=1}^{N_1} \frac{t^{2h}}{(2h)!} \times
\]
\[
2(h-1) \min \left[ m,n \right] \left[ \begin{array}{cc}
\frac{i-j}{2} & \frac{m-j}{2} \\
\frac{i-j}{2} & \frac{m-j}{2}
\end{array} \right] \\
\sum_{i=0}^{j+1+n} \sum_{j=0}^{m-j-2n} \sum_{n=0}^{r} \sum_{r=0}^{\left(\frac{j}{p}\right)} (-1)^{i+n-p+q-r} n-r^{1+p+r} j+n-p-r \\
\left( \begin{array}{c}
m-j-2n \\
q
\end{array} \right) \left( \begin{array}{c}
r \\frac{r}{s}
\end{array} \right) \left[ \begin{array}{c}
i+j-21+q \\
m-j-2n-q
\end{array} \right] \\
\left[ A_{11} + A_{12} a_1 + A_{11} a_2 \right] \\
\left[ A_{12} a_1 + A_{11} a_2 \right] \\
+ \left[ A_{12} a_1 + A_{11} a_2 \right] \\
\left[ A_{11} a_1 + A_{12} a_2 \right]
\right\}
\]
\[
\frac{2(h-2)+2(v+2k)}{2} \\
\max_{\frac{v}{2}} \left( \frac{i+n+p+q-r+2s}{2} \right) \\
\left( \frac{i+n+p+q-r+2s}{2} \right) \\
\right]\]

where \( m = (2h-2-i) \).

Notice that \( \frac{i+n+p+q-r+2s}{2} \) is an integer since \( \left( \begin{array}{c}
i+n-p+q-r \\
2
\end{array} \right) +1 \) is zero otherwise. Also, both \( i \) and \( i+n+p+q-r+2s \) are \( \leq 2(h-1) \).

Furthermore,
\[
f(0,0) = \exp \left\{ -\frac{1}{2} \left[ 1 - \frac{M^2}{m^2} \right] - \frac{1}{2} \left[ \frac{1}{\sigma_1} - \frac{1}{\sigma_2} \right] \right\} \\
\left[ \begin{array}{cc}
1 & -\rho M \\
-\rho M & 1
\end{array} \right] \left( \frac{1-\theta^2}{\nu(1-\rho^2 M^2)} \right), \\
A = \left[ \begin{array}{cc}
1 & -\rho M \\
-\rho M & 1
\end{array} \right] \left( \frac{1-\theta^2}{\nu(1-\rho^2 M^2)} \right), \\
A = \left[ \begin{array}{cc}
\frac{1}{\nu} & \frac{-\theta}{\nu} \\
\frac{\theta}{1+\theta^2} & 1
\end{array} \right], \\
A = \left[ \begin{array}{cc}
\frac{1}{\nu} & \frac{-\theta}{\nu} \\
\frac{\theta}{1+\theta^2} & 1
\end{array} \right], \\
A = \left[ \begin{array}{cc}
\frac{1}{\nu} & \frac{-\theta}{\nu} \\
\frac{\theta}{1+\theta^2} & 1
\end{array} \right]
\]

and
\[
\left| A \right|^{\frac{1}{2}} = \frac{(1-\theta^2)}{\nu(1-\rho^2 M^2)^{\frac{1}{2}}}
\]

Then \( P(t; N_1, N_2) = f(0,0) \left[ \frac{(1-\rho^2)^{\nu/2}}{\pi^{\nu}(1-\rho^2 M^2)^{3\nu}} \right] * \)
\[ \sum_{h=1}^{N_2} \left[ \frac{2t^2(1-p^2)}{v(1-p^2)^2} \right]^h \frac{1}{(2h)!} \]

\[ \sum_{i=0}^{2(h-1)} 1 \]

\[ \sum_{j=0}^{\min[M,i]} \binom{M}{j} [\rho M(1-p^2m^2)]^j \]

\[ \sum_{j=1}^{\binom{M-j}{2}} \frac{(i)!}{(1-j-2i)!} \frac{1}{[(1-p^2m^2)^{1.5}]^{1+i-j-2i}} \]

\[ \sum_{n=0}^{\binom{M-j}{2}} \frac{(M-j)!}{(M-j-2n)!}\frac{1}{(n)!} [\rho M(1-p^2m^2)]^n \]

\[ \sum_{p=0}^{j} \binom{j}{p} [-1/\rho M]^p \]

\[ \sum_{q=0}^{M-j-2n} \binom{M-j-2n}{q} \left\{ \frac{\theta_1}{\sigma_1} - \rho M \frac{\theta_2}{\sigma_2} \right\} \left\{ \frac{\theta_2}{\sigma_2} - \rho M \frac{\theta_1}{\sigma_1} \right\} + \]

\[ + \left\{ \frac{\theta_2}{\sigma_2} - \rho M \frac{\theta_1}{\sigma_1} \right\} \left\{ \frac{\theta_1}{\sigma_1} - \rho M \frac{\theta_2}{\sigma_2} \right\} \]

\[ \sum_{r=0}^{n} \binom{n}{r} \left[ (-1)^{i+n-p+q-r+1} \right] \left[ -1/\rho M \right]^r \]

\[ \sum_{s=0}^{r} \binom{r}{s} \]

\[ \sum_{k=0}^{N_1} \left[ \rho \right]^k \frac{\Gamma(k+\frac{\nu+1}{2}+\theta)\Gamma(k+\frac{\nu+1}{2}+(h-1)-\theta)}{\Gamma(k+\frac{\nu}{2})\Gamma(k+h)} \]

where \( \theta = (i+n+p+q-r+2s)/2 \)
Aside: Notice that the sum indexed by $k$ can be written as

$$S_{N_1} = \sum_{k=0}^{N_1} [\rho^2]^k \frac{\Gamma(k+a)\Gamma(k+b)}{\Gamma(k+c)(k)!}$$

where

$$a = \frac{\nu}{2} + \frac{1}{2} + \theta$$

$$b = \frac{\nu}{2} + \frac{1}{2} + h - 1 - \theta$$

$$c = \frac{\nu}{2}$$

$$\theta = (1+n+p+q-r+2s) \leq h - 1$$

and

$$\lim_{N_1 \to \infty} S_{N_1} = \frac{\Gamma(a)\Gamma(b)}{\Gamma(c)} \cdot F(a,b;c,\rho^2) < \infty$$

$$= \int_0^1 t^{b-1} (1-t)^{-b-1+c} (1-t\rho^2)^{-a} \, dt$$

where $F(a,b;c,\rho^2)$ is the hypergeometric function.

By using

$$\frac{\Gamma(a+b)}{\Gamma(b)} = \frac{\Gamma(a)}{\Gamma(b)} \left(\frac{b-1+u}{u}\right)_{u=1}^{a}$$

$$\frac{\Gamma(a+b+1)}{\Gamma(b+1)} = \frac{\Gamma(a)}{\Gamma(b)} \left(\frac{b-1+u}{u}\right)_{u=1}^{a}$$

$$\binom{b}{a} = \frac{\Gamma(b+1)}{\Gamma(b)} \left(\frac{b-1+u}{u}\right)_{u=1}^{a}$$

$$\frac{\Gamma(a+b)}{\Gamma(b)} = \frac{\Gamma(a)}{\Gamma(b)} \left(\frac{b-1+u}{u}\right)_{u=1}^{a}$$

then $P(t; N_1, N_2)$ can be written as
\[
P(t; N_1, N_2) = f(0,0) \left[ \frac{(1-\rho^2)^{3/2}}{\pi \Gamma(\frac{3}{2}) \sqrt{1-\rho^2}} \right] \star
\]
\[
\times \sum_{h=1}^{N_2} \left[ \frac{2t^2(1-\rho^2)}{\sqrt{1-\rho^2}} \right]^h \left[ \frac{h}{\sum_{u=1}^{h} \frac{1}{(2u)(2u-1)}} \right] \star
\]
\[
2(h-1) \sum_{i=0}^{\min(M,i)} \left[ (1-\rho^2 M^2) \right]^j \left[ \frac{i}{\pi} \frac{(2h-2-i+1-u)(i+1-u)}{(u)} \right] \star
\]
\[
\left[ \frac{i-j}{2} \right] \sum_{1=0}^{\frac{i-j}{2}} \left[ (1-\rho^2 M^2) \right]^1 \left[ \frac{1}{\pi} \frac{(i-j+1-2u)(i-j+2-2u)}{(u)} \right] \star
\]
\[
\left[ \frac{M-j}{2} \right] \sum_{n=0}^{\frac{M-j}{2}} \left[ (1-\rho^2 n^2) \right]^n \left[ \frac{n}{\pi} \frac{(M-j+1-2u)(M-j+2-2u)}{(u)} \right] \star
\]
\[
\frac{j}{\pi} \left[ \frac{1}{\rho M} \right] \left[ \frac{1}{\pi} \frac{(i+1-u)}{(u)} \right] \star
\]
\[
\sum_{p=0}^{M-j-2h} \left[ \frac{\rho^2}{\sigma_2} \right] \left[ \frac{\rho^2}{\sigma_1} \right] \left[ \frac{\rho^2}{\sigma_1} \right] \left[ \frac{\rho^2}{\sigma_2} \right] \left[ \frac{\rho^2}{\sigma_1} \right] \star
\]
\[
\sum_{q=0}^{M-j-2h} \left[ \frac{\rho^2}{\sigma_2} \right] \left[ \frac{\rho^2}{\sigma_1} \right] \left[ \frac{\rho^2}{\sigma_2} \right] \left[ \frac{\rho^2}{\sigma_1} \right] \left[ \frac{\rho^2}{\sigma_2} \right] \left[ \frac{\rho^2}{\sigma_1} \right] \star
\]
\[
\sum_{r=0}^{n} \left[ \frac{1}{\rho M} \right] \left[ \frac{1}{\pi} \frac{(n+1-u)}{(u)} \right] \star
\]
\[
\frac{r}{\pi} \left[ \frac{1}{\rho M} \right] \left[ \frac{1}{\pi} \frac{(n+1-u)}{(u)} \right] \star
\]
\[
\frac{r}{\pi} \left[ \frac{1}{\rho M} \right] \left[ \frac{1}{\pi} \frac{(n+1-u)}{(u)} \right] \star
\]
\[
\frac{r}{\pi} \left[ \frac{1}{\rho M} \right] \left[ \frac{1}{\pi} \frac{(n+1-u)}{(u)} \right] \star
\]
This expression is suitable for translation to FORTRAN.

The expression for \( P(t) \) in Theorem 2.1 can be evaluated by methods other than that of the MacLaurin polynomial. Since

\[
P(t) = \int_{-t}^{t} \int_{-t}^{t} f_{1,2}(t_1, t_2, x_1, x_2) \, dx_1 \, dx_2 \, dt_1 \, dt_2
\]

is a 4-fold integral with known integrand, 4-fold numerical integration is possible by Monte Carlo integration. This method is straightforward and inexpensive when only a few (e.g., 2) significant digits of accuracy are necessary. As a function of the number of significant digits desired, cost grows rapidly. Fortunately only one or two digits are required for the purposes of this research.

The probability integral

\[
P(t) = \int_{-t}^{t} \int_{-t}^{t} f_{1,2}(t_1, t_2, x_1, x_2) \, dx_1 \, dx_2 \, dt_1 \, dt_2
\]

is prepared for monte carlo computation by the following change of variables:

\[Y_1 = \frac{x_1}{1+x_1}, \quad Y_2 = \frac{x_2}{1+x_2}, \quad Y_3 = \frac{1}{2}(1+t_1/t), \quad Y_4 = \frac{1}{2}(1+t_2/t)\].

The Jacobian of this transformation is \( J = \left(2t/(1-Y_1)(1-Y_2)\right)^2 \) and so

\[
P(t) = \int_{0}^{1} \int_{0}^{1} \int_{0}^{1} \int_{0}^{1} f(t; Y_1, Y_2, Y_3, Y_4) \, dy_1 \, dy_2 \, dy_3 \, dy_4
\]

where

\[f(t; y) = f_{1,2}(t(2Y_3-1), t(2Y_4-1), Y_1/(1-Y_1), Y_2/(1-Y_2))* J\].

The p.d.f. \( f_{1,2}(t_1, t_2, x_1, x_2) \) can be written as
\[ \left( \frac{1 - \rho^2}{\pi^2 \nu - 1} \right)^\frac{\nu}{2} \exp \left\{ -\frac{\nu}{2} \left( \begin{array}{cc} 0 & \frac{\rho}{\sigma_1} \\ \frac{\rho}{\sigma_2} & 0 \end{array} \right) \right\} \cdot \left[ \prod_{i=1}^N f(t_i; Y) \, dy_1 \, dy_2 \, dy_3 \, dy_4 \right] \]

The integral \( P(t) = \int \int \int \int f(t; Y) \, dy_1 \, dy_2 \, dy_3 \, dy_4 \)

can be approximated by
\[ P_n(t) = \frac{N}{V} \sum_{i=1}^N f(t; Y_i) \]

where \( V = \int \int \int \int 1 \, dy_1 = 1 \) is the volume of the region of integration and
\[ Y_i = [Y_{1i}, Y_{2i}, Y_{3i}, Y_{4i}] \quad i=1,2,3,\ldots,N \]

are \( N \) randomly selected points uniformly distributed in the region of integration. The random variables \( Y_{ij} \) and \( Y_{kJ} \) are independent for all \((i,j) \neq (k,l)\) with \( N(0,1) \) marginal.
distributions. It follows that \( P_N(t) \) is a random variable with mean \( P(t) \) and variance \( \left( E[f(t;Y)] - [P(t)]^2/N \right) \) where

\[
E_Y[f(t;Y)] = \int_0^1 \int_0^1 \int_0^1 \int_0^1 f(t;Y) \, dy.
\]

For numerical evaluation, \( N \) 4-tuples of pseudo-random numbers, \( \{Y_i\}_{i=1}^N \), are generated and \( f(t;Y_i) \) must be evaluated for each of the \( N \) 4-tuples.

2.5 **Comparison of the two Algorithms**

In order to debug the algorithms and determine how well they each perform preliminary trial computations were attempted.

In the first trial, the following arbitrary parameter values were used: \( \frac{\theta_1}{\sigma_1} = \frac{\theta_2}{\sigma_2} = 1.0, \rho = 0.5, \nu = 10, m = 0.5 \) and \( t = 1 \). Using 24 seconds of c.p.u. time, the Monte Carlo algorithm quickly converged to \( \hat{P}(t) = .27 \). The MacLaurin algorithm converged slowly as an alternating series: after 12 seconds \( \hat{P}(t) = .24 \), after another 3 minutes and 15 seconds \( \hat{P}(t) = .25 \). It appears that the alternating series would eventually converge to a value near .27; however, this cannot be proven since no usefully sharp bounds could be derived for the residual term of the MacLaurin expansion used.

For the special case of \( \rho = 0 \) the correct value of \( P(t) \) can be computed from tables of the student-t distribution. Several examples were tried. The results for the MacLaurin algorithm are as follows:
<table>
<thead>
<tr>
<th>RUN</th>
<th>$\theta_1$</th>
<th>$\theta_2$</th>
<th>$\rho$</th>
<th>$M$</th>
<th>$v$</th>
<th>$t$</th>
<th>$P(t)$</th>
<th>$\hat{P}(t)$</th>
<th>Converged?</th>
<th>Seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0</td>
<td>10</td>
<td>1.093</td>
<td>.4900</td>
<td>.4920</td>
<td>apparently</td>
<td>12</td>
</tr>
<tr>
<td>2.</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>10</td>
<td>0.879</td>
<td>.3600</td>
<td>.3601</td>
<td>apparently</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>10</td>
<td>1.900</td>
<td>.8100</td>
<td>&gt;1</td>
<td>rapidly diverging</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>1</td>
<td>6.314</td>
<td>.8100</td>
<td>&gt;1</td>
<td>rapidly diverging</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>6</td>
<td>2.000</td>
<td>.9025</td>
<td>&gt;1</td>
<td>rapidly diverging</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>10</td>
<td>2.228</td>
<td>.9025</td>
<td>&gt;1</td>
<td>rapidly diverging</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

The Monte Carlo algorithm was also run for these examples and performed as well as or better than the MacLaurin algorithm in all cases. While the MacLaurin algorithm does compute an analytic formulation of $P(t)$, it must perform many arithmetic operations to do so. The Monte Carlo Integration method only estimates $P(t)$ as the mean of a distribution. However to do so it needs relatively few operations.
CHAPTER 3

BIVARIATE-t DISTRIBUTION THEORY FOR POWER COMPUTATIONS:
THE CASE OF DIFFERENT DESIGN MATRICES IN THE
LINEAR MODELS FOR PERIODS 1 AND 2

3.1 **Summary**

As in Chapter 2, theoretical results are developed in order
to construct an algorithm for computing $\Pr[|t_1| \leq t \text{ and } |t_2| \leq t]$. In this chapter the theory is additionally complicated by the
removal of the assumption that period 1 and period 2 linear
models in the by-period analysis have identical design matrices.
Hence algorithms for direct integration or numerical integration
or Monte Carlo integration are not derived. Instead a method
of Monte Carlo simulation is presented which handles the most
general case of messy data with different design matrices. Fur-
thermore this algorithm can be used to compute $\Pr[t_1 \in R_1 \text{ and } t_2 \in R_2]$ where $R_1$ and $R_2$ are intervals on the real line or sets of
real numbers. This algorithm is therefore useful for the power
calculations and for calculation of the difference, $(\alpha - \alpha^*)$, be-
tween the selected level of significance and the actual level
when using the Bonferroni inequality. The algorithm is useful
for both two-sided and one-sided tests. Of course this
algorithm can also be used to perform the computations of Chapter 2 algorithms as special cases. The theoretical results on this chapter hinge on the use of canonical variates without loss of generality.

3.2 The Structure of the Data

In Chapter 2 it is assumed that the two t-statistics are constructed from data which follow a bivariate general linear model (GLM) having mean

$$E[Y_i] = X_i \beta_i, \ i=1,2$$

In this chapter the requirement of identical design ($X$) matrices is relaxed to handle the most general case;

$$E[Y_i] = X_i \beta_i, \ i=1,2$$

$$X_i \ (n_i \ by \ r_i)$$

$$\beta_i \ (r_i \ by \ 1)$$

Here $X_1$ and $X_2$ may have different dimensions as well as different elements. It is assumed that some (or all) of the $Y_1$ and $Y_2$ observations are pairwise correlated; i.e., $Y_{1i}$ and $Y_{2i}$ are dependent for some $i$. This model is applicable to experiments which suffer dropouts and/or admit new study subjects between periods 1 and 2.

3.3 Transformation of the Data to Canonical Form

As in Chapter 2, the computation of the probability

$$P(t) = Pr[|T_1| \leq t \ and \ |T_2| \leq t]$$

is desired. Here $T_i \ i=1,2$ are
t-statistics based on period 1 and period 2 data respectively. The joint distribution of \((T_1, T_2)\) in its general form is quite complicated. It is however possible to derive a useful expression for the joint distribution of \([T_1, T_2, Q_1, Q_2]\) where \(Q_i\) is proportional to the square of the denominator of \(T_i\) (\(i=1,2\)). (Integration with respect to \(Q(2\) by \(1\)) over \(R^2\) and integration with respect to \(T(2\) by \(1\)) over the square region \(T_i \in [-t, t]\) \((i=1,2)\) yields \(P(t)\).

As a first step in deriving the distribution of \([T, Y]\) it is quite beneficial to transform the data to canonical form. The result is simplified distribution theory without loss of generality. The t-statistic is the ratio of a normal r.v. to the square-root of an error-sum-of-squares, and may be written

\[
T_i = (M_i * Y_i) / [Y_i^T P_i Y_i / \nu_i]^{1/2}
\]

\[
= 1 / \alpha_i (\hat{\theta}_i) / [Y_i^T P_i Y_i / \sigma_i^2 \nu_i]^{1/2}
\]

\[
= W_i / [Z_i^T Z_i / \sigma_i^2 \nu_i]^{1/2}
\]

where \(M_i\) is 1 by \(n_i\), \(P_i\) is \(n_i\) by \(n_i\), \(W_i \overset{d}{\sim} N(u_i, 1)\) and

\[
1 / \sigma_i^2 (\nu_i \text{ by } 1) \overset{d}{\sim} N(Q_i, I) \text{ independent of } W_i (i=1,2). \text{ The } Z_i \text{ are canonical variates computed from the values of } Y_i \text{ and } X_i. \text{ The simplifying transformations from } Y_i \text{ to } W_i \text{ and from } Y_i \text{ to } Z_i \text{ can be performed simultaneously: the joint distribution of } [W_1, W_2] \text{ is normal and independent of the distribution of } [Q_1, Q_2] \text{ where } Q_i = Z_i^T Z_i = Y_i^T P_i Y_i. \text{ The bivariate chi-square distribution of}
\]
\([Q_1, Q_2]\) is complicated, but useful expressions for the characteristic function and for the p.d.f. can be derived from the joint distribution of \(Z_1(v_1 \text{ by 1})\) and \(Z_2(v_2 \text{ by 1})\). The canonical variates \(Z_i = [Z_{i1}, Z_{i2}, \ldots, Z_{i\nu_i}]^T\) are such that the only dependencies are pairwise:

\[
[Z_{j1}, Z_{j2}] \sim N_2(Q, \begin{bmatrix} 1 & \rho_{d_j} \\ \rho_{d_j} & 1 \end{bmatrix})
\]

for all \(j \in \{1, 2, \ldots, \nu_2\}\). Canonical correlations, \(d_j\), are defined in Lemma 3.3. For \(j \in \{\nu_2 + 1, \ldots, \nu_1\}\), \(Z_{j1}\) is independent of all other \(Z_{ji}\). It is unfortunate but in general true that the \([Z_{j1}, Z_{j2}]\) pairs do not share a common covariance. Some of the \(d_j\) may be zero, and \(d_j = 0\) implies \(Z_{j1}\) and \(Z_{j2}\) are independent.

Lemma 2.1 describes the structure and assumed distribution of the data \((Y_1, Y_2)\) in general. Lemma 3.3 uses the results of Lemmas 3.1 and 3.2 to construct the linear transformation from response variables \([Y_1^T, Y_2^T]\) to canonical variables \([Z_1^T, Z_2^T]\) and to give the distribution of \([Z_1^T, Z_2^T]\). Lemma 3.4 shows that dividing the numerator and denominator of \(T_i\) by \(\sigma_i\) poses no distributional problems when performed simultaneously for \(i = 1, 2\). Lemma 3.5 proves that \([T_1^\ast, T_2^\ast]\) constructed from \([W_1, W_2, Z_1^T, Z_2^T]\) has the same distribution as \([T_1, T_2]\) constructed directly from \([Y_1^T, Y_2^T]\). It also shows that the distribution of \([T_1, T_2]\) depends on \(\theta_1, \theta_2, \sigma_1, \) and \(\sigma_2\) only through \(\theta_1/\sigma_1\) and \(\theta_2/\sigma_2\).

Lemma 3.6 provides a result useful for computing \(P(t)\) by Monte Carlo simulation. Pseudo-random variables having the
joint distribution of \([W_1, W_2, Q_1, Q_2]\) can be easily generated from standard normal variable values. The method is most efficient when \(\nu_1, \nu_2\) is small.

Lemma 3.7 derives the bivariate chi-square distribution of \([Q_1, Q_2]\) following the work of Jensen (1970). Lemma 3.8 gives an expression for the p.d.f. of the four-variate distribution of \([T_1, T_2, Q_1, Q_2]\). Theorem 3.1 presents a formula for \(P(t)\) which can easily be numerically evaluated using Monte Carlo integration or by using the Maclaurin series method described in Chapter 2.

**Lemma 3.1**

If \(P = [I_n - XX^T X^{-1}]\) where \(X(n \times 4)\) has rank \(r\) so that \(P\) has rank \(= n-r\) then \(\exists M\) s.t. \(MM^T = P, MM^T = I_\nu,\) and \(MX = 0.\)

**Proof of Lemma 3.1:**

The spectral decomposition of \(P\) is \(P = V\Lambda V^T\) where \(\Lambda\) is a diagonal matrix of eigenvalues and \(V\) is an orthonormal matrix of eigenvectors. Since \(P\) is idempotent the eigenvalues are zeros and ones. Since \(\text{rank}(P) = \nu\), only \(\nu\) of the eigenvalues are not zero. Define \(\Delta = (\Lambda\) with null rows deleted). Then \(\Delta\nu = I_\nu,\)

\(\Delta\nu = A,\) and \(A^T A = \Delta.\) So \(P = \Delta \nu \Delta \nu^T = M^T M\) where \(M = A \Delta \nu^T\) is a subset of the rows of \(\Delta \nu^T\). Furthermore \(M^T = \Delta \nu^T = I_\nu.\) Finally, \(P \Delta \nu = 0\) implies \(M^T M \Delta \nu = 0\) implies \(X^T M \Delta \nu = 0\) implies \(X^T M = 0\) implies \(MX = 0.\)
Lemma 3.2

If \( P_i = I_{n_i} - X_i X_i^T X_i^{-1} X_i^T \) where \( X_i(n_i \times r_i) \) has rank \( r_i \) (i=1,2) so that \( P_i \) has rank \( v_i = n_i - r_i \) and \( v_1 \geq v_2 \) and

\[
B(n_1 \times n_2) = \begin{bmatrix} I_{q} & 0 \\ 0 & 0 \end{bmatrix} \text{ where } q \in \{1,2,\ldots,v_2\}
\]

Then there exists \( A_1(v_1 \times n_1) \) and \( A_2(v_2 \times 2) \) such that \( A_i^T A_i = P_i \), \( A_i A_i^T = I_{v_i} \), \( A_i X_i = 0 \), and

\[
A_1 B A_2^T = \begin{bmatrix} \Delta \\ 0 \end{bmatrix} \text{ where } \Delta(v_2 \times v_2) \text{ is a diagonal matrix having rank } \leq \min\{q,v_2\}.
\]

Proof of Lemma 3.2:

By Lemma 4.1 there exist \( M_i(v_1 \times n_1) \) s.t. \( M_i^T M_i = P_i \) and \( M_i M_i^T = I_{v_i} \). The singular value decomposition of \( M_0 = M_1 B M_2^T \) is

\[
M_0 = Q_i^T \begin{bmatrix} \Delta \\ 0 \end{bmatrix} Q_2 \text{ where } Q_i(v_i \times v_i) \text{ is orthonormal and } \Delta(v_2 \times v_2) \text{ is diagonal. Define } A_i = Q_i M_i. Then } A_i A_i^T = Q_i M_i M_i^T Q_i^T = Q_i Q_i^T = I,
\]

\[
A_i^T A_i = M_i^T Q_i^T Q_i M_i = M_i^T M_i = P_i, \ A_i X_i = Q_i M_i X_i = Q_i 0 = 0,
\]

and \( A_1 B A_2^T = Q_1 M_1 B M_2^T Q_2^T = Q_1 Q_1^T \begin{bmatrix} \Delta \\ 0 \end{bmatrix} Q_2 Q_2^T = \begin{bmatrix} \Delta \\ 0 \end{bmatrix} \). Since \( B \) has rank \( q \), the rank of \( \Delta \) must be \( \leq \min \{q,v_2\} \).
Lemma 3.3

If \[
\begin{bmatrix}
    y_1 \\
    y_2
\end{bmatrix}
\] d N \( n_1 + n_2 \) \begin{bmatrix}
    \begin{bmatrix}
        X_1 & 0
    \end{bmatrix} & \beta_1 \\
    0 & X_2
\end{bmatrix} \begin{bmatrix}
    \beta_1 \\
    \beta_2
\end{bmatrix}, \quad \Sigma_y \), with rank \( X_i (n_i \text{ by } r_i) = r_i \)

\[
\Sigma_y = \begin{bmatrix}
\sigma_1^2 I_{n_1} & \sigma_1 \sigma_2 \rho B \\
\sigma_1 \sigma_2 \rho B & \sigma_2^2 I_{n_2}
\end{bmatrix}
\]

where \( B (n_1 \text{ by } n_2) = \begin{bmatrix}
I_q & 0 \\
0 & 0
\end{bmatrix} = \begin{bmatrix}
I_q \\
0
\end{bmatrix} \quad \begin{bmatrix}
I_q, 0
\end{bmatrix} \]

and \( p_i = [I_{n_1} - X_1 [X_1 X_1^T]^{-1} X_1] \) has rank \( v_i \) (i = 1, 2)

and \( v_2 \leq v_1 \)

Then there exists a transformation to canonical variates

\[
\begin{bmatrix}
    z_1 \\
    z_2
\end{bmatrix}
\] d N \( v_1 + v_2 \) \begin{bmatrix}
    \begin{bmatrix}
        A_1 & 0
    \end{bmatrix} & \begin{bmatrix}
        y_1 \\
        y_2
    \end{bmatrix} \\
    0 & A_2
\end{bmatrix} \begin{bmatrix}
    \begin{bmatrix}
        A_1 & 0
    \end{bmatrix} & \begin{bmatrix}
        y_1 \\
        y_2
    \end{bmatrix} \\
    0 & A_2
\end{bmatrix}\]

where \( A_1 (v_1 \text{ by } n_1) \) and \( A_2 (v_2 \text{ by } n_2) \) are matrix functions of \( X_1, X_2 \) and \( B \) such that

\[
\begin{bmatrix}
    z_1 \\
    z_2
\end{bmatrix}
\] d N \( v_1 + v_2 \) \begin{bmatrix}
    \begin{bmatrix}
        0 \\
        0
    \end{bmatrix} \\
    \begin{bmatrix}
        \sigma_1^2 I_{v_1} & \sigma_1 \sigma_2 \rho [\Delta \Omega]^T \\
        \sigma_1 \sigma_2 [\Delta \Omega] & \sigma_2^2 I_{v_2}
    \end{bmatrix}
\end{bmatrix}
\]

where \( \Delta \) is diagonal with elements \([d_1, \ldots, d_{v_2}]\) on the diagonal, and

\[
Z_1^T Z_1 = Y_1^T A_1^T A_1 Y_1 = Y_1^T p_i Y_1 \quad (i=1, 2).
\]
Proof of Lemma 3.3:

By lemma 3.2 there exist $A_1$ and $A_2$ such that $A_1 A_1^T = I_{\nu_1}$, $A_1^T A_2 = P_1$, $A_1 X_1 = Q$, and $A_1 BA_2^T = [Q]_v^T$. The joint distribution of $Z_1$ and $Z_2$ is obtained immediately by applying well-known multivariate normal theory.

In the special case of $X_1 = X_2$ it happens that $\Delta = I_{\nu_1}$
where $\nu = \nu_1 = \nu_2$.

---

**Lemma 3.4**

If $\hat{\theta} \sim N_2(0, \begin{bmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \\ \sigma_1 \sigma_2 & 2 \sigma_2 \end{bmatrix})$ and $[Q_1, Q_2]$ are non-negative variates which are independent of $\hat{\theta}$ and for $i=1,2$

$$T_i = (\hat{\theta}_i / \sigma_i) / (Q_i / \sigma_i^2 \nu_i)^{1/2} \quad \text{and} \quad T_i^* = (\hat{\theta}_i / \sigma_i) / (Q_i / \sigma_i^2 \nu_i)^{1/2}$$

Then the joint distributions of $[T_1, T_2, Q_1, Q_2]$ and of $[T_1^*, T_2^*, Q_1, Q_2]$ are identical, and those of $[T_1, T_2]$ and $[T_1^*, T_2^*]$ are identical.

---

**Proof of Lemma 3.4**

Assume that the p.d.f. of $[T_1, T_2, Q_1, Q_2]$ exists. Then it is $f(T, Q) = f(T | Q) f(Q)$

where the conditional distribution of $[T$ given $Q]$ is bivariate normal. It is only necessary to prove that $f(T | Q) = f(T^* | Q)$.

Given $Q$, the transformation from $\hat{\theta}$ to $T$ is of the form

$$T = \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix} \begin{bmatrix} \hat{\theta} \end{bmatrix}$$
and the transformation from $\hat{\theta}$ to $\mathcal{I}^*$ is of the form

$$
\mathcal{I}^* = \begin{bmatrix}
 b_1 & 0 \\
 0 & b_2 
\end{bmatrix}
$$

Since $a_i = b_i$, the transformations are identical. Integration with respect to $Q_1$ and $Q_2$ yields the distributions of $\mathcal{I}$ and $\mathcal{I}^*$ which it follows must be identical. The usefulness of this lemma is that it shows that $T$ can be constructed from underlying variates which may have a simpler variance-covariance structure.

**Lemma 3.5**

If the conditions of Lemma 2.1 are assumed and

$$
T_i = \frac{\hat{\theta}_i}{[Q_i/\nu_i]^{1/2}} \quad i=1,2
$$

and

$$
\begin{bmatrix}
 W_1 \\
 W_2 
\end{bmatrix}
 \overset{d}{\sim}
 N_2 \left( \begin{bmatrix} 0/\sigma_1 \\ 0/\sigma_2 \end{bmatrix}, \begin{bmatrix} I & \rho M \\ \rho M & 1 \end{bmatrix} \right)
$$

is independent of

$$
\begin{bmatrix}
 Z_1 \\
 Z_2 
\end{bmatrix}
 \overset{d}{\sim}
 N_{\nu_1+\nu_2} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} I_{\nu_1} & \rho [\Delta,0]^T \\ \rho [\Delta,0] & I_{\nu_2} \end{bmatrix} \right)
$$

and

$$
T_i^* = \frac{W_i}{[Z_i^T Z_i/\nu_i]^{1/2}} \quad i=1,2
$$

Then $\mathcal{I}$ and $\mathcal{I}^*$ have the same bivariate distribution, and $[\mathcal{I},Q]$ and $[\mathcal{I}^*,Q]$ have the same 4-variate distribution.

**Proof of Lemma 3.5:**

By Lemma 3.3, $Z_i^T Z_i \overset{d}{\sim} Q_i/\sigma_i$ and $W_i \overset{d}{\sim} \hat{\theta}_i/\sigma_i$. Then by Lemma 3.4, $[\mathcal{I},Q] \overset{d}{\sim} [\mathcal{I}^*,Q]$. 
This Lemma shows that a sample from the distribution of $T$ can be obtained from the distribution of the $\nu_1 + \nu_2 + 2$ variates $Z_1, Z_2, W_1,$ and $W_2$. The parameters of the $T$ distribution must depend only on $\rho m$, $\rho A$, $\theta / \sigma$, and $\nu_i$ ($i=1,2$).

**Lemma 3.6**

If $W_1, W_2, Z_1,$ and $Z_2$ have the normal distributions given in Lemma 3.5 and $W_1^*, W_2^*$ and the elements of $Z_1^*$ and $Z_2^*$ are mutually independent standard-normal variates

$$\begin{bmatrix} W_1 \\ W_2 \end{bmatrix} \overset{d}{\sim} \begin{bmatrix} 1 & 0 \\ \rho m & [1-\rho^2 m^2]^{1/2} \end{bmatrix} \begin{bmatrix} W_1^* \\ W_2^* \end{bmatrix} + \begin{bmatrix} 0_{1/\sigma_1} \\ 0_{2/\sigma_2} \end{bmatrix}$$

and

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \overset{d}{\sim} \begin{bmatrix} I_{\nu_1} & 0 \\ \rho [A, 0] & [I-\rho^2 A^2]^{1/2} \end{bmatrix} \begin{bmatrix} Z_1^* \\ Z_2^* \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and thus $Q_1 = Z_1^T Z_1 \overset{d}{\sim} Z_1^* Z_1^*$

while $Q_2 = Z_2^T Z_2 \overset{d}{\sim} \begin{bmatrix} \rho^2 Z_1^T Z_1^* & 0 \\ 0 & 0 \end{bmatrix} + Z_2^T [I-\rho^2 \Delta^2] Z_2^* + 2\rho Z_2^T [I-\rho^2 \Delta^2]^{1/2} [A, 0] Z_2^*$

If $\nu_1 = \nu_2 = \nu$ and $\Delta = d I$ then $Q_2 = \rho^2 d^2 Z_1^T Z_1^* + (1-\rho^2 d^2) Z_2^T Z_2^* + 2\rho d (1-\rho^2 d^2) Z_2^T Z_1^*$
Proof of Lemma 3.6:

Matrix multiplication yields

\[
\begin{bmatrix}
1 & \rho M \\
\rho M & 1
\end{bmatrix}
= \begin{bmatrix}
1 & 0 \\
\rho M & [1 - \rho^2 M^2]^{\frac{1}{2}}
\end{bmatrix}
\cdot
\begin{bmatrix}
1 & \rho M \\
0 & [1 - \rho^2 M^2]^{\frac{1}{2}}
\end{bmatrix}
\]

and

\[
\begin{bmatrix}
I_{v_1} & \rho [\Delta, 0]^T \\
\rho [\Delta, 0] & I_{v_2}
\end{bmatrix}
= \begin{bmatrix}
I_{v_1} & 0 \\
0 & [1 - \rho^2 \Delta^2]^{\frac{1}{2}}
\end{bmatrix}
\cdot
\begin{bmatrix}
I_{v_1} & \rho [\Delta, 0]^T \\
\rho [\Delta, 0] & [1 - \rho^2 \Delta^2]^{\frac{1}{2}}
\end{bmatrix}
\]

By applying well-known multivariate normal theory it is easily shown that the linear transformations of \([W_1^*, W_2^*]\) and of \([Z_1^*, Z_2^*]^T\) implied in the lemma yield normally distributed variates which are distributed as \([W_1, W_2]\) and as \([Z_1, Z_2]^T\).

Since identical functions of identically distributed variates are distributed identically, the remaining results of the lemma follow.
Lemma 3.7

If $Z_1$ and $Z_2$ have the normal distribution given in Lemma 3.5 and

$$ Q_i = Z_i^T Z_i \ (i=1,2) $$

Then the joint distribution of $Q_1$ and $Q_2$ has characteristic function

$$ \phi(t_1,t_2) = \left[1-2it_1 \right]^{-\nu_1/2} \left[1-2it_2 \right]^{-\nu_2/2} \pi_{j=1}^2 \left[1-c_2 \rho^2 d_j 2^{-1} \right] $$

where $c = (2it_1)(2it_2)/(1-2it_1)(1-2it_2) \ 
and \ A = \text{diag}(d_1,...,d_2)$ as in Lemma 3.5.

The bivariate p.d.f. of $Q_1$ and $Q_2$ can be written (a.e.) as

$$ f_{Q1,Q2}(Q_1,Q_2) = \frac{Q_1^{\nu_1/2} \psi(\frac{\nu_1}{2}; \frac{Q_1}{2}) \psi(\frac{\nu_2}{2}; \frac{Q_2}{2})}{\psi(\frac{\nu_1+\nu_2}{2}; \frac{Q_1+Q_2}{2})} \times $$

$$ \times \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\nu_1+2m} \sum_{j=0}^{\nu_2+2n} g_k g_{k'} L_k^L \left( \frac{\nu_1}{2} \right) L_{k'}^L \left( \frac{\nu_2}{2} \right) Q_1 \psi(\frac{\nu_1+2m}{2}; \frac{Q_1}{2}) \psi(\frac{\nu_2+2n}{2}; \frac{Q_2}{2}) $$

where $\psi(\frac{\nu_i}{2}; \frac{Q_i}{2}) = \left[ Q_i \right]^{\nu_i/2} - \frac{Q_i}{2} \ e^{-Q_i/2}$ is the marginal $\chi^2$ p.d.f.

$$ L_k^L(\alpha)(u) = \sum_{i=0}^{K} \binom{K}{K+\alpha} \left[-u\right]^i/(K)! \text{ is a Laguerre polynomial,}$$

$$ g_k = (K)!(K)! \Gamma(\frac{\nu_1}{2}) \Gamma(\frac{\nu_2}{2}) / \Gamma(\frac{\nu_1+2K}{2}) \Gamma(\frac{\nu_2+2K}{2}) \ ,$$

and

$$ G_k = \sum_{j_1+j_2+\ldots+j_{2N}} a_{j_1} \rho_{d_1} \ldots a_{j_{2N}} \rho_{d_{2N}} $$

where

$$ a_{j_1} = [\rho_{d_1}]^{2j_1} \Gamma(j_1+\frac{1}{2}) / \sqrt{\pi} \Gamma(j_1+1) $$
Proof of Lemma 3.7:

The results of this lemma are given by Jensen (1970). An outline-proof of these results is as follows. For each $j \in \{1, 2, \ldots, \nu_2\}$ the pair $[Z_{1j}, Z_{2j}]$ has a standard-bivariate-normal distribution with correlation coefficient p.d.j. For each $j \in \{\nu_2+1, \ldots, \nu_1\}$ $Z_{1j}$ is a standard normal variate. Except for pairwise $\rho \cdot d_j$ correlations all the $Z_{1j}$'s are independent. It has been shown that the joint characteristic function of $[Z_{1j}^2/2, Z_{2j}^2/2]$ is

$$
\phi(t_{1j}, t_{2j}) = [(1-it_{1j})(1-it_{2j}) - \rho^2 d_j^2(1-it_{1j})(1-it_{2j})]^\frac{-\nu_2}{2}
$$

for $j \leq \nu_2$. For $j > \nu_2+1$ the characteristic function of $|Z_{1j}^1/2|$ is

$$
\phi(t_{1j}) = [1-it_{1j}]^{-\frac{\nu_1}{2}}.
$$

By independence the joint c.f. of $[\{Z_{1j}^2/2\}, \{Z_{2j}^2/2\}]$ is

$$
\phi(t_{11}, \ldots, t_{1\nu_1}, t_{21}, \ldots, t_{2\nu_2}) = \prod_{j=1}^{\nu_2} \phi(t_{1j}, t_{2j}) \prod_{k=1}^{\nu_1} \phi(t_{1k})
$$

which can be written as

$$
\prod_{k=1}^{\nu_1} (1-it_{1k})^{-\frac{\nu_1}{2}} \cdot \prod_{j=1}^{\nu_2} (1-it_{2j})^{-\frac{\nu_2}{2}} \cdot \prod_{L=1}^{\nu_2} [1-C_j \rho^2 d_j^2]^{-\frac{\nu_2}{2}}
$$

where

$$
C_j = (it_{1j})(it_{2j})/(1-it_{1j})(1-it_{2j})
$$

Since $Q_1/2$ is a convolution of $[Z_{1j}^2]$ (and similarly $Q_2/2$) the joint c.f. of $[Q_1, Q_2]$ is as given in the lemma. In order
to obtain an expression for the p.d.f. based on the c.f. Jensen (1970) uses the fact that

\[(1-cp_j^2d_j^2)^{-\frac{1}{2}} = \sum_{k=0}^{\infty} [cpc_j^2d_j^2]^k \Gamma(k+\frac{1}{2})/\Gamma(k+1)\Gamma(\frac{1}{2}).\]

Collecting terms of the product in \(\phi(t_1,t_2)\) gives

\[\phi(t_1,t_2) = \sum_{k=0}^{\infty} G_k \left(1-it_1\right)^{-\frac{v_1}{2}} \left[\frac{it_1}{1-it_1}\right]^k \left(1-it_2\right)^{-\frac{v_2}{2}} \left[\frac{it_2}{1-it_2}\right]^k\]

where

\[G_k = \sum_{j_1}^{\ldots} \sum_{j_2}^{\ldots} \sum_{j_3}^{\ldots} a_{j_1} \ldots a_{j_2} \ldots (\rho d^2) \frac{2^k}{j_1!j_2!j_3!} \]

is a sum taken over all integer partitions of \(k\) in terms of the \(j_i\)'s.

It has been shown that \((1-it)^{-g}(it/(1-it))^h\) has as its inverse Fourier transform

\[L_h^{(g-1)}(\chi) \psi(\chi;g) \Gamma(g)\Gamma(h+1)/\Gamma(g+h).\]

Hence the p.d.f. of \((Q_1, Q_2)\) is as given in the Lemma. This series expression for the bivariate p.d.f. uses an orthogonal system of functions (the Laguerre polynomials) which have the marginal p.d.f.s as their weight functions. Jensen (1970) shows that the series is absolutely convergent almost everywhere (a.e.).

For \(K=0\), \(G_0=1\). For \(K=1\), \(G_1 = \frac{1}{2}\rho^2 \sum_{i=1}^{\nu_2} d_i^2\)

is of the form \(\sum_i b_i^2\) and for \(K=2\) \(G_2 = (\rho b_i b_j + \sum b_{i1} b_{j1})^2 - 2b_{i1} b_{j1}\).
Lemma 3.8

If \( W \overset{d}{=} N_2(\mu, R_W) \) is independent of

\[
\begin{bmatrix}
Z_1 \\
Z_2
\end{bmatrix} \sim N_{\nu_1 + \nu_2}(0, R_Z)
\]

where

\[
R_W = \begin{bmatrix}
1 & \rho M \\
\rho M & 1
\end{bmatrix} \quad \text{and} \quad R_Z = \begin{bmatrix}
\nu_1 & \rho [\lambda, 0] \\
\rho [\lambda, 0] & \nu_2
\end{bmatrix}
\]

and \( T_i = W_i/(Q_i/\nu_i)^{1/2} \) \((i = 1, 2)\)

where \( Q_i = Z_i^T Z_i \), \( f_{X^2} (\cdot) \) is the chi-square p.d.f. with \( p \) degrees of freedom, and

\[
q = \begin{bmatrix}
\nu_1/\nu_1 - \mu_1 \\
\nu_2/\nu_2 - \mu_2
\end{bmatrix}^T R_W^{-1} \begin{bmatrix}
\nu_1/\nu_1 - \mu_1 \\
\nu_2/\nu_2 - \mu_2
\end{bmatrix}
\]

\[
= (1 - \rho^2 M^2)^{-1} \left\{ Q_1 T_1^2/\nu_1 + 2 T_1 \sqrt{Q_1/\nu_1} \left[ \rho M \nu_2 - \mu_2 \right] \right\} + \\
\mu_1^2 + \mu_2^2 - 2 \mu_1 \mu_2 - 2 \rho M Q_1 Q_2 T_1 T_2 + \\
Q_2 T_2^2/\nu_2 + 2 T_2 \sqrt{Q_2/\nu_2} \left[ \rho M \nu_1 - \mu_1 \right] \right\}
\]

Then the joint p.d.f. of \([T_1, T_2, Q_1, Q_2]\) can be written as

\[
f_{T, Q}(T, Q) = f_{T/Q}(T/Q) \cdot f_Q(Q)
\]

where

\[
f_{T/Q}(T/Q) = \frac{(1 - \rho^2 M^2)^{-1/2}}{2\pi} \exp \left[ -\frac{1}{2} q \right]
\]

and

\[
f_Q(Q) = \sum_{K=0}^{\infty} \sum_{m=0}^{K} \sum_{n=0}^{K} \left( \begin{array}{c}
K \\
m
\end{array} \right) \left( \begin{array}{c}
K \\
n
\end{array} \right) \frac{m+n}{\sqrt{\nu_1 \nu_2}} \chi_{\nu_1 + 2m}^2 (Q_1) \chi_{\nu_2 + 2n}^2 (Q_2)
\]

almost everywhere.
Proof of Lemma 3.8:

The conditional p.d.f. $f_{T|Q}$ is a bivariate normal p.d.f. easily computed using well-known normal distribution theory. The bivariate p.d.f. $f_Q$ is derived in Lemma 3.7.

\[ P(t) = \int_{-t}^{t} \int_{-t}^{t} f_{T|Q}(T_1, T_2) \,dT = \]

\[ \left( \frac{\nu_1 \nu_2 (1 - \rho^2 \nu^2)}{2\pi} \right)^{-\frac{1}{4}} \sum_{k=0}^{\infty} \frac{K^k}{k!} \sum_{m=0}^{\infty} \frac{(K)^k}{m!} \int_{m+n}^{\infty} \frac{Q_1(t) \,dQ}{\sqrt{\nu_1 + 2m}} \frac{Q_2(t) \,dQ}{\sqrt{\nu_2 + 2n}} \]

where

\[ g_Q(t) = \int_{-t}^{t} \exp\left\{ -\frac{1}{2} \begin{bmatrix} \sqrt{\nu_1/\nu_1} & 0 \\ 0 & \sqrt{\nu_2/\nu_2} \end{bmatrix} T_{-y} \right\} \begin{bmatrix} T_{-y}^{-1} \left( \begin{array}{c} 0 \\ \sqrt{\nu_2/\nu_2} \end{array} \right) \right\} d\Gamma. \]

and $G_k$ is defined in Lemma 3.7.

The function $g_Q(t)$ can be written as a Maclaurin series such that

\[ g_Q(t) = \sum_{K=0}^{N} \frac{g(2K)(0)}{(2K)!} t^{2K} + R_N(t) \]

where $g(2K)(0)$ is a polynomial in terms of the form $Q_1^r Q_2^s$. 

\[ \]
Proof of Theorem 3.1:

The series expression for $f_{J,Q}$ is given in Lemma 3.8 and (as stated in Lemma 3.7) is convergent a.e. By Theorem 2.2 $g_Q(t)$ can be expressed as a MacLaurin polynomial in $t$ with known coefficients.

If expressed as a MacLaurin series, $P(t)$ involves no integrals but is algebraically involved with the summation of many terms. As a 4-fold integral, $P(t)$ can be evaluated by crude Monte Carlo integration or other numerical methods.

3.4 An Algorithm for Computing Power

Lemma 3.6 shows that a sample from the joint distribution of $[W_1,W_2,Q_1,Q_2]$ can be obtained by transforming samples from the $N_{\nu_1+\nu_2+2}(0,1)$ distribution of $[W_1*,W_2*,Z_1^*,Z_2^*]$. Thus, one pseudo-random pair $[T_1,T_2]$ from the distribution of $[T_1,T_2]$ can be obtained from $\nu_1+\nu_2+2$ independent $N(0,1)$ values.

The probability $p = \Pr[T_1 \in R_1$ and $T_2 \in R_2]$, where $R_i$ is a real interval or a set of real numbers, can be estimated from a sample of $[T_1,T_2]$ by counting the number of pseudo-random values which fall into the region designated by $\{(t_1,t_2): t_1 \in R_1$ and $t_2 \in R_2\}$. Let $X_n$ be the number of pairs which fall into the region out of a sample of $n$ pairs. Then $X_n$ has a binomial distribution with mean $np$ and variance $np(1-p)$. $\bar{X}_n$ has mean $p$ and variance $p(1-p)/n$. Since
p(1-p) ≤ 0.25, the standard error of $\hat{p} = \frac{\chi_n}{n}$ is less than or equal to $0.5/\sqrt{n}$. For $n=500$ the standard error is $\leq 0.0224$. If $(\nu_1 + \nu_2 + 2)$ is not too large, the computation of $\hat{p}$ is efficient as well as straightforward.

This algorithm has the advantage of being able to compute an estimate of not only power = $1 - \Pr(\left|t_1^*\right| \leq t_1 \text{ and } |t_2^*| \leq t_2)$ but also $(\alpha - \alpha^*) = \Pr(\left|t_1^*\right| > t_1 \text{ and } |t_2^*| > t_2)$. Furthermore, the algorithm can handle similar probability calculations for one-sided tests for which power = $1 - \Pr(t_1^* \leq t_1 \text{ and } t_2^* \leq t_2)$ and $(\alpha - \alpha^*) = \Pr(t_1^* > t_1 \text{ and } t_2^* > t_2)$. 
4.1 Summary

Data from a CO(2,4,2) experiment are analyzed according to the by-period method. In this case the linear model design matrices, $X_1$ and $X_2$, are identical. The hypothesis that the treatment effects are zero in both periods is tested using the Bonferroni inequality. The power of this test is computed.

Having analyzed the complete set of data, a randomly selected subset of the data is presented as a messy data problem. The subset is analyzed according to the by-period method. In this case the design matrices, $X_1$ and $X_2$, have different dimensions and entries. The power of the two-sided test for treatment effects is again computed. Analysis of the messy-data subset according to the model explicated by Grizzle (1965) is discussed and the power of the recommended test for treatment effects is computed. Since many of the messy-data observations are unusable in the usual model, power is lost. It is shown that in this example the test in the by-period analysis is more powerful than the analogous test in the traditional analysis.
The particular set of data analyzed is illustrative of situations in which the traditional model is not appropriate; carryover is prominent, there is a period*treatment interaction, treatment effects in the two periods are not the same, and the two periods do not share a common variance. Other analyses of the same experiment are also possible and some were performed. For example, analysis of rate of growth (rather than cumulative growth) is very different from the analyses presented. The data analyzed in this chapter provide the best example of the need for an alternative analysis and strategy such as that of by-period approach. Furthermore, application of the theoretical methods developed in Chapters 2 and 3 is best demonstrated for the analyses presented below.

4.2 A Bean Sprout Experiment

Details of the bean-sprout experiment involving the herbicide 2,4D are given in the appendix. The protocol is elaborated, the data are listed and simple descriptive statistics are presented. The data are arrayed as 44 observations on 3 variables: Group, Y1 and Y2. The unit of study is a triplet of sprouts and the analysis variables are within-triplet averages of the measurements made on sprout lengths. Y1 is a measure of length in mm at the end of period 1. Y2 is a measure of length at the end of period 2.

Figure 4.2.1 illustrates the group means plotted as
growth curves. The effects of the different treatment sequences are dramatic with respect to absolute sprout length. There appears to be carryover from period 1, and treatment given in period 2 appears to have minimal effect.

The correlations between Y1 and Y2 for the four treatment groups are given below. The treatment sequences associated with the groups refer to the active reagent, "A", and the placebo, "P".

<table>
<thead>
<tr>
<th>Group</th>
<th>Seq.</th>
<th>$r_{Y_1,Y_2}$</th>
<th>(p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>A,A</td>
<td>0.86</td>
<td>(.0006)</td>
</tr>
<tr>
<td>2.</td>
<td>A,P</td>
<td>0.81</td>
<td>(.0028)</td>
</tr>
<tr>
<td>3.</td>
<td>P,A</td>
<td>0.96</td>
<td>(.0001)</td>
</tr>
<tr>
<td>4.</td>
<td>P,P</td>
<td>0.85</td>
<td>(.0009)</td>
</tr>
</tbody>
</table>

These coefficients indicate that sprout lengths as measured at the ends of the periods are highly correlated.

4.3 By-Period Analysis for the Complete Set of Data

The two length variables, Y1 and Y2, are measurements made at the ends of the two periods. We define $u_{p,T1,T2} = E[Y_{p,T1,T2,K}]$ as the mean of the observation on the Kth study unit in period p; this unit received treatment sequence (T1,T2). We assume

$\text{COV}[Y_{p,T1,T2,K}, Y_{p',T1',T2',K'}]$ is zero if $K \neq K'$, and is

$\sigma_{pp'}$ if $K = k'$. 


Figure 4.2.1
MEAN SPROUT LENGTHS

Means are plotted for each of the treatment groups. "A" denotes the active reagent treatment. "P" denotes the placebo treatment.
4.3.1 Period 1 Analysis

The primary parameters are defined as follows:
\[ \mu_1 = \bar{\mu}_1 \quad \text{--- the overall period 1 average,} \]
\[ T_1 = \bar{\mu}_{1p} - \bar{\mu}_{1A} \quad \text{--- the period 1 treatment effect,} \]
\[ S_2^{(1)} = \bar{\mu}_{1p} - \bar{\mu}_{1A} \quad \text{--- a sequence effect.} \]

Because of the randomized treatment design, we assume all sequence effects are zero. The sequence effect \( S_2^{(1)} \) is included in the Period 1 model so that the design matrices \( X_1 \) and \( X_2 \) will be identical and thus the case of \( X_1 = X_2 \) will be illustrated.

The period 1 response vector \( Y_1(44 \times 1) \) has expected value \( X_1 \beta_1 \) where
\[ X_1 = \begin{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} & (A,A) \\ (A,P) \\ (P,A) \\ (P,P) \end{bmatrix} \]

and
\[ \beta_1 = [\mu_1 \quad T_1 \quad S_1^{(1)}]^t. \]

For this general linear univariate model (GLUM), the estimates of the model parameters based on \( N = 44 \) sprout triplets are:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>t-test</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_{11} )</td>
<td>95.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sqrt{\sigma_{11}} )</td>
<td>9.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \mu_1 )</td>
<td>59.77</td>
<td>1.47</td>
<td>4.17</td>
<td>0.000</td>
</tr>
<tr>
<td>( T_1 )</td>
<td>9.71</td>
<td>1.47</td>
<td>6.68</td>
<td>0.000</td>
</tr>
<tr>
<td>( s_2^{(1)} )</td>
<td>-0.94</td>
<td>1.47</td>
<td>-0.65</td>
<td>0.526</td>
</tr>
</tbody>
</table>
The expected values of length in mm at the end of period 1 are estimated as follows:

<table>
<thead>
<tr>
<th>Group</th>
<th>Sequence</th>
<th>Mean Length</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>A,A</td>
<td>( \mu_{1-T_1-S_2}^{(1)} )</td>
<td>49.00mm</td>
<td>2.55</td>
</tr>
<tr>
<td>2.</td>
<td>A,P</td>
<td>( \mu_{1-T_1+S_2}^{(1)} )</td>
<td>47.12mm</td>
<td>2.55</td>
</tr>
<tr>
<td>3.</td>
<td>P,A</td>
<td>( \mu_{1+T_1-S_2}^{(1)} )</td>
<td>68.42mm</td>
<td>2.55</td>
</tr>
<tr>
<td>4.</td>
<td>P,P</td>
<td>( \mu_{1+T_1+S_2}^{(1)} )</td>
<td>66.54mm</td>
<td>2.55</td>
</tr>
</tbody>
</table>

A separate analysis of the residuals indicates that the assumptions of the model are well founded. The normality assumption is supported by the Kolmogorov-Smirnov goodness-of-fit test (p-value > .20) as well as by a probability plot of the residuals.

As may be seen from Figure 4.2.1 and from the p-value for the test of \( H_0: T_1=0 \), the treatment applied clearly has a substantial effect in period 1. As expected, the sequence effect (included in the model solely to illustrate the case \( S_1=S_2 \)) is not significant.

4.3.2 Period 2 Analysis

The primary parameters of interest are defined as follows:
\[ \mu_2 = \bar{\mu}_2 \ldots \]  -- overall period 2 mean

\[ \rho_{12} = \bar{\mu}_2^P - \bar{\mu}_2^A. \]  -- different in treatment P carryover and treatment A carryover

\[ T_2 = \bar{\mu}_2^P - \bar{\mu}_2^A. \]  -- period 2 treatment effect

Because of the experimental design, we assume all sequence effects are zero. The response vector \( Y_2 \) (44 by 1) has mean \( X_2 \beta_2 \) where

\[
X_2 = \begin{bmatrix}
1 & -1 & -1 \\
1 & -1 & 1 \\
1 & 1 & -1 \\
1 & 1 & 1
\end{bmatrix} \quad (A,A) \\
(1 & -1 & 1) \quad (A,P) \\
(1 & 1 & -1) \quad (P,A) \\
(1 & 1 & 1) \quad (P,P)
\]

\[ \beta_2 = [\mu_2 \quad \rho_{12} \quad T_2]' . \]

In this model \( X_1 = X_2 \). Also notice that \( T_2 \) is proportional to the average of the group1--group2 difference and the group3--group4 difference. Since both differences are valid comparisons, their average is also meaningful.

Model parameters (based on N=44) triplets are estimated as follows:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>t-test</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_{22} )</td>
<td>237.61</td>
<td>15.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sqrt{\sigma_{22}} )</td>
<td>( \sqrt{237.61} )</td>
<td>2.32</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>( \mu_2 )</td>
<td>94.31</td>
<td>2.32</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>( \rho_{12} )</td>
<td>23.01</td>
<td>2.32</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>( T_2 )</td>
<td>0.11</td>
<td>2.32</td>
<td>0.961</td>
<td></td>
</tr>
</tbody>
</table>
The expected values of length in mm at the end of period 2 are estimated as follows:

<table>
<thead>
<tr>
<th>Group</th>
<th>Sequence</th>
<th>Mean Length</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>A,A</td>
<td>$\mu_2 - \rho_1 12 - \tau_2$</td>
<td>71.19</td>
<td>4.02</td>
</tr>
<tr>
<td>2.</td>
<td>A,P</td>
<td>$\mu_2 - \rho_1 12 + \tau_2$</td>
<td>71.41</td>
<td>4.02</td>
</tr>
<tr>
<td>3.</td>
<td>P,A</td>
<td>$\mu_2 + \rho_1 12 - \tau_2$</td>
<td>117.21</td>
<td>4.02</td>
</tr>
<tr>
<td>4.</td>
<td>P,P</td>
<td>$\mu_2 + \rho_1 12 + \tau_2$</td>
<td>117.43</td>
<td>4.02</td>
</tr>
</tbody>
</table>

A separate analysis of the residuals indicates that the model assumptions are well founded. The Kolmogorov-Smirnov normality goodness-of-fit test has a p-value greater than .20. A probability plot of the residuals also supports the normality assumption.

The treatment effect in period 2 is negligible whereas the treatment effect in period 1 was found to be quite large. The carryover effect from period 1 into period 2 is clearly substantial. Also note the increase in the estimated variance from 95 mm$^2$ in period 1 to 238 mm$^2$ in period 2.

4.3.3 Combined Over-Periods Analysis

The period effect defined as $\pi = \mu_2 - \mu_1$ can be estimated by $\hat{\pi} = \hat{\mu}_2 - \hat{\mu}_1$. The standard error of $\hat{\pi}$ and a test of Ho: $\pi = 0$ vs. Ha: $\pi \neq 0$ are both obtained by fitting the GLUM
$$E[Y_{2}Y_{1}] = X_{1}^{*}\hat{\beta}_{21}$$ where \( \pi \) is the row 1, column 1 element of \( \hat{\beta}_{21} \) (3 by 1). The computed value of \( \hat{\pi} \) is 36.54 with a standard error of 1.21. The 2-sided t-test has p-value = .000 and therefore the period effect is substantial.

The computational results for periods 1 and 2 and for the combined analysis can all be obtained by fitting the bivariate model

$$E[Y_{1}, Y_{2}] = X_{1}^{*}[\hat{\beta}_{1}, \hat{\beta}_{2}] \quad (X_{1} = X_{2})$$

$$V[Y_{1}, Y_{2}] = I_{44} @ \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

where \( \sigma_{11} = \sigma_{22} \), \( \sigma_{12} = \sigma_{21} \), and \( \sigma_{12} = \sigma_{21} \). An estimate of covariance \( \sigma_{12} \) is also obtained as 134.14. Hence \( \hat{\rho} = .892 \). The computed r-square for the bivariate model is 0.98.

The hypothesis Ho: \( T_{1} = 0 \) and \( T_{2} = 0 \) vs. Ha: \( T_{1} \neq 0 \) and/or \( T_{2} \neq 0 \) can be tested using the Bonferroni inequality and the two individual t-statistics: if at least one of the individual null hypotheses is rejected at the \( \alpha/2 \) level then the joint null hypothesis is rejected. Since Ho: \( T_{1} = 0 \) is significant at any level, the joint test is also significant (i.e., Ho is rejected).

The hypothesis Ho: \( T_{1} = T_{2} \) vs Ha: \( T_{1} \neq T_{2} \) is not readily testable. However, the estimates \( \hat{T}_{1} = 9.71 \pm 2(1.47) \) and \( \hat{T}_{2} = 0.11 \pm 2(2.32) \) indicate that \( T_{1} \) and \( T_{2} \) are probably unequal. This means that the difference between treatments A and P in period 1 is not equal to the difference between A and P in period 2; i.e., there is a period\* treatment interaction. An estimate of this interaction is \( \hat{T}_{2} - \hat{T}_{1} = -9.60 \) but
the standard error is not available.

In summary, it appears that under the conditions of the experiment 2,4D has a dramatic detrimental effect on the sprouts in period 1 (relative to the effect of placebo) which carries over into period 2 as damage sustained. Then in period 2, 2,4D does little damage if any, regardless of the period 1 treatment. This interpretation implies that 2,4D and time period affect the sprouts interactively; i.e., the effect of 2,4D depends on when it is administered.

The significant carryover effect and the difference in period 1 and period 2 treatment effects (i.e., treatment*period interaction) raise serious questions about the validity of the traditional analysis for this experiment. The increase in the estimated variance from period 1 to period 2 is another violation of the assumptions of the usual analysis. By comparison, none of the assumptions of the by-period analysis are violated.

The power of the Bonferroni test of $H_0: T_1=0$ and $T_2=0$ is equal to the probability of rejecting the null hypothesis when $T_1\neq 0$ and/or $T_2\neq 0$; i.e.,

$$\text{Power} = \Pr(|t_1^*| > t(\nu; \alpha/2) \text{ and/or } |t_2^*| > t(\nu; \alpha/4))$$

$$= 1 - \Pr(\max\{t_1^*, t_2^*\} \leq t(\nu; \alpha/4))$$

$$= 1 - \text{P}(t)$$

where $t_1^*$ and $t_2^*$ are the individual t-statistics, $t(\nu; \alpha/4)$ is the $100(1-\alpha/4)$ percentage point of the student-t distribution with $\nu$ degrees of freedom, and $\nu=41$. For $\alpha=.05$
t(\nu; \alpha/4) = 2.3267. Other parameters of P(t) are (in the notation of Lemma 2.1)

\[ \theta_1 = C_1 \cdot \bar{e}_1 = \sqrt{N} \cdot t_1 \quad \text{where} \quad C_1 = \sqrt{N} \cdot [0, 1, 0] \quad \text{and} \quad N = 44 \]

\[ \theta_2 = C_2 \cdot \bar{e}_2 = \sqrt{N} \cdot t_2 \quad \text{where} \quad C_2 = \sqrt{N} \cdot [0, 0, 1] \]

\[ \sigma_1^2 = \sigma_{11} \]

\[ \sigma_2^2 = \sigma_{22} \]

\[ \rho = \sigma_{12} / (\sigma_{11} \sigma_{22})^{1/2} \]

\[ \nu = N \cdot \text{rank}(X_1) = 41 \]

\[ m = C_1 [X_1^T X_1]^{-1} X_1^T X_2 [X_2^T X_2]^{-1} C_2 = 0 \quad \text{since} \quad X_1^T X_1 = N \cdot I \]

The function P(t) depends on \( \theta_1 \) and \( \sigma_1^2 \) only through \( \theta_1 / \sigma_1 \).

If we assume some fixed values for \( T_1 \) and \( T_2 \) and assume \( \sigma_1^2 = 95.13, \sigma_2^2 = 237.61, \) and \( \rho = 0.892 \) (as estimated) then P(t) can be computed as a function of \( \hat{T}_1 \) and \( \hat{T}_2 \). The figure below presents power = 1 - P(t) as a function of \( \hat{T}_1 \) and \( \hat{T}_2 \). In the neighborhood of \( \hat{T}_1 \) and \( \hat{T}_2 \):

Power as a function of \( T_1 \) and \( T_2 \). Computation by Monte Carlo simulation is based on 500 pseudo-random bivariate-t pairs.

| \( \hat{T}_2 \) | 1.00 | .96 | .95 | .90 | .80 | .70 | .60 | .50 | .40 | .30 | .20 | .10 |
| 1.00 | .07 | .19 | .11 | .06 | .18 | .09 | .06 | .06 | .06 | .06 | .06 | .06 |

| \( T_2 \) | .07 | .19 | .11 | .06 | .18 | .09 | .06 | .06 | .06 | .06 | .06 | .06 |
| 1.00 | .07 | .19 | .11 | .06 | .18 | .09 | .06 | .06 | .06 | .06 | .06 | .06 |

| \( \hat{T}_1 \) | 1.00 | .96 | .95 | .90 | .80 | .70 | .60 | .50 | .40 | .30 | .20 | .10 |
| 1.00 | .07 | .19 | .11 | .06 | .18 | .09 | .06 | .06 | .06 | .06 | .06 | .06 |

| \( T_1 \) | .07 | .19 | .11 | .06 | .18 | .09 | .06 | .06 | .06 | .06 | .06 | .06 |
| 1.00 | .07 | .19 | .11 | .06 | .18 | .09 | .06 | .06 | .06 | .06 | .06 | .06 |
If the experiment was analyzed using the model explicated by Grizzle (1965) a test for the effect of treatment would be performed and the power of that test could be computed. The power of that test is discussed below in Section 4.4.2.

Using the Bonferroni inequality with the individual t-statistics assures that the true probability of type I error is \( \alpha^* < \alpha \) where \( \alpha \) has a value chosen by the analyst. The difference \( (\alpha - \alpha^*) \) is equal to \( \Pr(|t_1^*| > t(\nu; \alpha/4) \text{ and } |t_2^*| > t(\nu; \alpha/4)) \) evaluated at \( T_1 = 0 \) and \( T_2 = 0 \). Computation of \( (\alpha - \alpha^*) \) therefore involves a probability integral of a central bivariate-t distribution. For \( \alpha = .05 \) and parameter values as in Figure 4.3.3.1 (with \( T_1 = 0 = T_2 \)), Monte Carlo simulation based on 500 pseudo-random bivariate-t pairs indicates that \( (\alpha - \alpha^*) \) is approximately zero. This is reasonable when we consider that \( m = 0 \) (page 4.12) implies that the numerators of the two t-statistics are independent. Because the denominators are dependent we still need the Bonferroni inequality however.

4.4 Messy-Data Analyses

A subset of the bean-sprout data was selected at random to provide an example of analyses involving missing data and different design matrices. Figure 4.4.1 lists the data available. This set of data is analyzed by period and the test of treatment effect is compared with the analogous test in the traditional analysis with respect to power.
4.4.1 By-Period Analysis

The same linear model used in Section 4.3 applies to the analysis of the messy data except that \( S_2^{(1)} \) is omitted from the model. All of the data (listed in Figure 4.4.1) are usable.

4.4.1.1 Period 1 Analysis

The estimates of the model parameters based on the \( N_1=40 \) available sprout triplets are as follows:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>t-test p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_{11} )</td>
<td>90.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \nu_1 )</td>
<td>57.42</td>
<td>1.50</td>
<td>0.000</td>
</tr>
<tr>
<td>( T_1 )</td>
<td>9.47</td>
<td>1.50</td>
<td>0.000</td>
</tr>
</tbody>
</table>

For this univariate model R-square is 0.98. A separate analysis of the residuals also supports the assumptions of the model.

The expected values of length in mm at the end of period 1 for each group are estimated as follows:

<table>
<thead>
<tr>
<th>Group</th>
<th>Sequence</th>
<th>( N_{11} )</th>
<th>Mean Length</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>( i=1 )</td>
<td>A,A</td>
<td>10</td>
<td>( \nu_1-T_1 )</td>
<td>47.95</td>
<td>2.12</td>
</tr>
<tr>
<td>( i=2 )</td>
<td>A,P</td>
<td>10</td>
<td>( \nu_1-T_1 )</td>
<td>47.95</td>
<td>2.12</td>
</tr>
<tr>
<td>( i=3 )</td>
<td>P,A</td>
<td>11</td>
<td>( \nu_1+T_1 )</td>
<td>66.89</td>
<td>2.12</td>
</tr>
<tr>
<td>( i=4 )</td>
<td>P,P</td>
<td>9</td>
<td>( \nu_1+T_1 )</td>
<td>66.89</td>
<td>2.12</td>
</tr>
</tbody>
</table>
Figure 4.4.1

Observations on Sprout Triples

<table>
<thead>
<tr>
<th>OBS</th>
<th>Period 1</th>
<th>Period 2</th>
<th>Group</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>53.66</td>
<td>-</td>
<td>1</td>
<td>A,A</td>
</tr>
<tr>
<td>2</td>
<td>45.66</td>
<td>-</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>54.33</td>
<td>-</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>44.00</td>
<td>-</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>46.00</td>
<td>-</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>49.33</td>
<td>72.66</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>38.00</td>
<td>58.00</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>50.33</td>
<td>70.00</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>48.33</td>
<td>69.00</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>42.66</td>
<td>58.00</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>-</td>
<td>68.00</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>51.00</td>
<td>-</td>
<td>2</td>
<td>A,P</td>
</tr>
<tr>
<td>13</td>
<td>55.66</td>
<td>-</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>44.00</td>
<td>-</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>41.66</td>
<td>-</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>39.33</td>
<td>55.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>54.33</td>
<td>84.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>52.33</td>
<td>67.33</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>48.66</td>
<td>77.33</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>53.00</td>
<td>87.33</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>46.66</td>
<td>62.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>-</td>
<td>78.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>66.00</td>
<td>-</td>
<td>3</td>
<td>P,A</td>
</tr>
<tr>
<td>24</td>
<td>81.00</td>
<td>-</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>62.66</td>
<td>-</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>89.33</td>
<td>-</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>58.00</td>
<td>-</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>70.33</td>
<td>-</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>82.00</td>
<td>135.33</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>52.33</td>
<td>84.00</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>80.33</td>
<td>133.33</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>74.66</td>
<td>120.33</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>54.33</td>
<td>94.00</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.4.1
(continued)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th>4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>34</td>
<td>69.33</td>
<td>-</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>74.33</td>
<td>-</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>66.66</td>
<td>-</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>77.00</td>
<td>-</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>58.00</td>
<td>-</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>53.00</td>
<td>85.00</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>68.33</td>
<td>120.00</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>61.33</td>
<td>103.66</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>38.66</td>
<td>82.33</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>-</td>
<td>143.00</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>-</td>
<td>129.66</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>
4.4.1.2 Period 2 Analysis

Only $N_2=24$ observations are available in period 2. The estimates of the model parameters are as follows.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>t-test p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_{22}$</td>
<td>303.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>90.70</td>
<td>3.58</td>
<td>0.000</td>
</tr>
<tr>
<td>$\rho_{12}$</td>
<td>21.06</td>
<td>3.57</td>
<td>0.000</td>
</tr>
<tr>
<td>$T_2$</td>
<td>1.27</td>
<td>3.57</td>
<td>0.724</td>
</tr>
</tbody>
</table>

The expected values of length in mm at the end of period 2 are estimated as follows:

<table>
<thead>
<tr>
<th>Group</th>
<th>Sequence</th>
<th>$N_{21}$</th>
<th>Mean Length</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i=1$</td>
<td>A,A</td>
<td>6</td>
<td>$\mu_2 - \rho_{12} - T_2$</td>
<td>68.37</td>
<td>6.20</td>
</tr>
<tr>
<td>$i=2$</td>
<td>A,P</td>
<td>7</td>
<td>$\mu_2 - \rho_{2} + T_2$</td>
<td>70.91</td>
<td>6.20</td>
</tr>
<tr>
<td>$i=3$</td>
<td>P,A</td>
<td>5</td>
<td>$\mu_2 + \rho_{12} - T_2$</td>
<td>110.49</td>
<td>6.20</td>
</tr>
<tr>
<td>$i=4$</td>
<td>P,P</td>
<td>6</td>
<td>$\mu_2 + \rho_{12} + T_2$</td>
<td>113.03</td>
<td>6.20</td>
</tr>
</tbody>
</table>

The R-square value for this model is 0.97. Residual analysis supports the distributional assumptions of the model.

4.4.1.3 Combined-Over-Periods Analysis

The period effect $\eta=\mu_2 - \mu_1$ can be estimated by fitting the linear model for $E[Y_{21} - Y_{11}]$. However only $N_{12}=20$
observations on Y2 -Y1 are available. Based on these data
\( \hat{\mu} = 32.6 \) with a standard error of 1.7. The two-sided test has
p-value \( \leq .000 \) and represents a substantial period effect. Complete
residual analysis supports the assumptions of this linear model.

The Bonferroni method of testing Ho: \( \tau_1=0 \) and \( \tau_2=0 \) vs.
Ha: \( \tau_1 \neq 0 \) and/or \( \tau_2 \neq 0 \) leads to the rejection of Ho since \( \hat{\tau}_1 \) is
significantly different from zero at any level. With estimates
\( \hat{\tau}_1 = 9.47 \pm 2.150 \) and \( \hat{\tau}_2 = 1.27 \pm 2.357 \), \( \tau_1 \) and \( \tau_2 \) are probably
unequal and hence there is a period*treatment interaction. An
estimate of this interaction is \( (\hat{\tau}_2 - \hat{\tau}_1) = -8.20 \) mm.

The power of the Bonferroni test of Ho: \( \hat{\tau}_1=0 \) and \( \hat{\tau}_2=0 \) was
computed using the following parameter values:

\[ N_1 = 40 = \text{the number of observations in period } 1 \]
\[ \nu_1 = 38 = \text{the degrees of freedom in period } 1 \]
\[ N_2 = 24 = \text{the number of observations in period } 2 \]
\[ \nu_2 = 21 = \text{the degrees of freedom in period } 2 \]
\[ \alpha = .05 = \text{the selected level of significance} \]
\[ t(\nu_1;\alpha/4) = 2.33372 = \text{the } 100(1-\alpha/4) \text{ percentage point of} \]
\[ \text{the } t_{\nu_1} \text{-distribution} \]
\[ t(\nu_2;\alpha/4) = 2.41385 = (\text{similarly}) \]
\[ C_1 = [0, 1] \ast ([0, 1][X_1^T X_1]^{-1}[0, 1]^T)^{-\frac{1}{2}} = [0, 1] \ast 6.32 \]
\[ \theta_1 = C_1 \beta_1 = \tau_1 \ast 6.32 \]
\[ C_2 = [0, 0, 1] \ast ([0, 0, 1][X_2^T X_2]^{-1}[0, 0, 1]^T)^{-\frac{1}{2}} = [0, 0, 1] \ast 4.88 \]
\[ \theta_2 = C_2 \beta_2 = \tau_2 \ast 4.88 \]
\[ m = \mathbf{C}_1 \left( \mathbf{x}_1^T \mathbf{x}_1 \right)^{-1} \mathbf{x}_1^T \left( \mathbf{I}_{2,20} + \mathbf{0} \right) \left( \mathbf{I}_{2,20} + \mathbf{0} \right) \mathbf{x}_2 \left( \mathbf{x}_2^T \mathbf{x}_2 \right)^{-1} \mathbf{C}_2 \]
\[ = 0.0684 \]

\[ \Delta = \text{diag}[0,1,1] \]

The figure below presents power as a function of \( T_1 \) and \( T_2 \) and the above parameter values:

Power as a function of \( T_1 \) and \( T_2 \):

Computation by Monte Carlo simulation is based on 500 pseudo-random bivariate-t pairs.

\[
\begin{array}{cccccc}
\hat{T}_2 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 \\
1 & .66 & .85 & 1.00 & .99 & 1.00 & .64 & .99 & 1.00 & .63 & .96 & .96 & .96 & 1.00 & .61 & .96 & 1.00 & .61 & .96 & .96 & 1.00 & .61 & .96 & .96 & 1.00 \\
2 & .09 & .62 & .96 & 1.00 & .96 & 1.00 & .08 & .61 & .96 & 1.00 & .96 & .96 & .96 & 1.00 & .96 & .96 & .96 & 1.00 & .96 & .96 & .96 & 1.00 & .96 & .96 & .96 & 1.00 \\
3 & .11 & .63 & .96 & 1.00 & .96 & 1.00 & .09 & .96 & .96 & 1.00 & .96 & .96 & .96 & 1.00 & .96 & .96 & .96 & 1.00 & .96 & .96 & .96 & 1.00 & .96 & .96 & .96 & 1.00 \\
4 & .36 & .73 & .99 & 1.00 & .99 & 1.00 & .36 & .73 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 \\
5 & .16 & .64 & .99 & 1.00 & .99 & 1.00 & .16 & .64 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 \\
6 & .76 & .85 & 1.00 & .99 & 1.00 & .76 & .85 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 \\
7 & .86 & .95 & 1.00 & .99 & 1.00 & .86 & .95 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 \\
8 & .96 & .95 & 1.00 & .99 & 1.00 & .96 & .95 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 \\
9 & .96 & .95 & 1.00 & .99 & 1.00 & .96 & .95 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 \\
10 & .96 & .95 & 1.00 & .99 & 1.00 & .96 & .95 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 & .99 & 1.00 \\
\end{array}
\]

The power of the test against the alternative hypothesis for \( \alpha = 0.05 \) is approximately 1.00 when \( T_1 = \hat{T}_1 \) and \( T_2 = \hat{T}_2 \), due primarily
to the large first period treatment effect.

4.4.2 Power Calculations for the Traditional Model

The messy data could be analyzed using the model explicated by Grizzle (1965). In this section, we compute the power of the treatment effects test for that model. To do so it is only necessary to estimate $\sigma^2$ and $\alpha$ which are defined below. The following model is under consideration (see Grizzle (1965)):

$$E[Y_{pijk}] = \mu \quad \text{-- a general mean}$$

$$+ \xi_{ij} \quad \text{-- effect of subject } j \text{ within sequence } i$$

which is distributed $N(0, \sigma^2_s)$. 

$$i = 2, 3 \text{ and } j \in \{1, \ldots, n_i\}. \text{ Also assume}$$

$$\xi_i = 0.$$ 

$$+ \pi_p \quad \text{-- period effect } p, p = 1, 2 \text{ and assume}$$

$$\pi_1 + \pi_2 = 0.$$ 

$$+ \phi_k \quad \text{-- treatment effect } K, K = 1, 2 \text{ and assume}$$

$$\phi_1 + \phi_2 = 0.$$ 

$$+ \lambda_{kp} \quad \text{-- carryover effect } k \text{ is zero unless } p = 2$$

$$+ e_{ijp} \quad \text{-- random variation distributed as}$$

$$N(0, \sigma^2_e) \text{ independent of } \xi_{ij}.$$ 

$$V[Y_{pijk}] = \sigma^2_3 + \sigma^2_s = \sigma^2$$

$$\text{COV}[Y_{pijk}, Y_{p'ij'k}] = \sigma^2_s$$

Note that this model would utilize only the data in the (A,P) and (P,A) sequence groups and within those groups only those experimental units with data from both periods.
When carryover effects ($\lambda_k$) are included in this model, the period contrast ($\pi_2 - \pi_1$) is not estimable. Grizzle (1965) recommends that the contrasts ($\phi_2 - \phi_1$) and ($\lambda_2 - \lambda_1$) be estimated and then if $H_0: \lambda_2 = \lambda_1$ is not rejected at a moderate level (e.g., $\alpha = .10$), the carryover effects ($\lambda_k$) are deleted from the model.

Unlike the by-period analysis, this model assumes that the Kth treatment has the same effect in period 2 that it had in period 1; i.e., period*treatment interaction is assumed nil. For the messy data at hand it is obvious (from preceding analysis and data plots) that this assumption is not true.

The non-centrality parameter of the t-distribution of the test statistic for $H_0: \phi_1 = \phi_2$ vs. $H_a: \phi_1 \neq \phi_2$ depends on whether $\lambda_1 = \lambda_2$; i.e., whether there is equal carryover from the two treatments, A and P. It is obvious from previous analyses that $\lambda_1 \neq \lambda_2$ and so the non-centrality parameter is

$$\Delta = (\phi_2 - \phi_1)/\sigma \cdot [1/n_2 + 1/n_3]^{1/2}$$

where $n_i$ is the number of sequence i observations (See Grizzle (1965)). The degrees of freedom for the t-statistic are $(n_2 + n_3 - 2) = 9$ since the number of observation pairs available on study units which received sequence A, P is $n_2 = 6$ while for P, A $n_3 = 5$. Grizzle (1965) provides a method of computing the amount of experimentation necessary to obtain 0.95 power against the alternative hypothesis, $H_a: \phi_2 - \phi_1 \neq 0$, for a 0.05 significance level. The method defines $\Delta^* = (\phi_2 - \phi_1)/\sigma \sqrt{2}$ and uses the graph in section 2.3 (p. 39) of Owen (1962). Estimates of $\sigma^2$ and
\((\phi_2 - \phi_1)\) for the messy data are given by:

\[
2 \cdot (\phi_2 - \phi_1) = (Y_{12} - Y_{22})^T [I_1 | T_1]^{-1} - (Y_{13} - Y_{23})^T [T_1 | I_1]^{-1}
\]

\[
(n_1 + n_2 - 2) \sigma^2 = Y_{12}^T [I - I_1 | I_1]^{-1} T_1 Y_{12} + \]

\[
+ Y_{13}^T [I - I_1 | I_1]^{-1} T_1 Y_{13}
\]

where \(Y_{ji}(n_i \text{ by } 1)\) are the observations in period \(j\) on group \(i\). Computation yields \(\sigma^2 = 96.56\) and \((\phi_2 - \phi_1) = 4.28\). The noncentrality is \(\Lambda = 0.72\). Notice that the estimate of \(\phi_2 - \phi_1\) uses period 1 and period 2 data. The estimate of \(\sigma^2\) is based only on period 1 data, as recommended by Grizzle. The following table of approximate sample sizes needed is calculated from Owen's graph. (See Grizzle (1965).)

<table>
<thead>
<tr>
<th>(\sigma^2 = 96.56) assumed</th>
<th>((\phi_2 - \phi_1))</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n_2 + n_3) for power 0.95</td>
<td>100</td>
<td>76</td>
<td>54</td>
<td>40</td>
<td>33</td>
<td>28</td>
<td>23</td>
<td>20</td>
<td>17</td>
<td>15</td>
<td>13</td>
<td>12</td>
<td>11</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

Since \((n_2 + n_3) = 11\), 0.95 power is anticipated for any \(\phi_2 - \phi_1 > 16\). However, the actual estimate of \(\phi_2 - \phi_1\) is 4.28.

Power calculation based on \(\Lambda\) is actually conservative. The true non-centrality is \(\Lambda\) only when the assumption of constant variance across periods is valid. The traditional model assumes that \(\sigma^2\) is common to both periods. But from previous analyses we know that variance is much larger in period 2 than in period 1.
Thus the period 1 estimate of $\sigma^2$ leads to an overestimate of non-centrality which in turn gives a "conservative" calculation of power: power is even smaller than the above table implies. For the $(n_2+n_3)=11$ observation pairs actually available the power of the test can be computed as

$$\Pr [\text{Ho rejected given } \Delta \neq 0] = 1 - \Pr \left[ |t^*(\Delta)| \leq t(v; \alpha/2) \right]$$

where $t^*$ is the test statistic with $v=9$ degrees of freedom and non-centrality parameter $\Delta$. Computation of the probability yields the following table. (Remember, since $\Delta$ was overestimated so is power.)

<table>
<thead>
<tr>
<th>$\sigma^2=96.56$ and $n_2=6$, $n_3=5$ assumed</th>
<th>$\phi_2 - \phi_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 4 7 10 13 16 19 22</td>
</tr>
</tbody>
</table>

| non-centrality parameter | 0.17 0.67 1.18 1.68 2.18 2.59 3.19 3.70 |
| Power                    | 0.053 0.093 0.184 0.324 0.496 0.669 0.811 0.907 |

For $(\phi_2 - \phi_1) = \overline{(\phi_2 - \phi_1)} = 4.28$ power is small.

In the by-period analysis the power of test was boosted by the usefulness of Group 1 and Group 4 data. If we had given Group 1 the treatment sequence $(A, P)$ and had given Group 4 the treatment sequence $(P, A)$ then the same missing-value pattern would have yielded $n_2=11$ and $n_3=9$. Then $v=18$ and the following
table would be obtained.

\[
\begin{array}{|c|ccccccccc|}
\hline
\sigma^2=96.56 & & (\phi_2-\phi_1) \\
n_2=11, n_3=9 & 1 & 4 & 7 & 10 & 13 & 16 & 19 & 22 \\
\text{assumed} & & & & & & & & \\
\hline
\text{non-centrality} & 0.23 & 0.91 & 1.58 & 2.26 & 2.94 & 3.62 & 4.30 & 4.98 & \\
\text{parameter} & & & & & & & & \\
\text{power} & 0.054 & 0.138 & 0.324 & 0.573 & 0.795 & 0.928 & 0.982 & 0.997 & \\
\hline
\end{array}
\]

For \((\phi_2-\phi_1) = 4.28\) power is again small: Power improved with the hypothetical alteration in the experimental design but not enough to equal the power of the by-period analysis test procedure. Furthermore power was overestimated because of use of \(\Delta\).

Generally, the CO(2,2,2) design is expected to decrease the amount of necessary experimentation whenever carryover effects can be omitted and \(\sigma_s^2/\sigma^2>0\). When \(\sigma_s^2/\sigma^2=0\) a completely randomized design (1 period) is equally efficient. Here, even though \(\sigma_s^2/\sigma^2>0\), so many observations are unusable in the traditional model that substantial power is lost.

If there were no missing observations and only the (A,P) and (P,A) sequences had been applied then we would have \(n_2=22, n_3=22, \nu=42\) and \(\Delta\) still approximately equal to 0.72 under the model's assumptions. It is only necessary to check the results from Owen's graph to see that 0.95 power will be obtained only if \((\phi_2-\phi_1)>8\). The estimate of \((\phi_2-\phi_1)\) is 4.28.
Yet another analysis alternative is the use of only period 1 data (see Brown (1980).) This is the same as using only the period 1 model in the by-period analysis. This method does not allow an investigation of period*treatment interaction or of the period 2 treatment effect. We cannot say that this method has any advantage over the by-period analysis with respect to power since (1) it is a part of the by-period analysis, and (2) a comparison of power would concern $H_0: (T_1=0)$ vs $H_a: (T_1 \neq 0)$ and $H_0: (T_1=0)_{\Delta} (T_2=0)$ vs $H_a: (T_1 \neq 0)_{\Delta} (T_2 \neq 0)$.

4.4.3 Comparisons of Power

It is possible to examine power as a function of $(T_1, T_2)$ in the by-period analysis and power as a function of $\Delta$ in the traditional analysis. The two analyses can be compared on the basis of power of the "treatment effect(s)" test by assuming that $\Delta$ is in the neighborhood of 0.72 and $(T_1, T_2)$ is in the neighborhood of $(\hat{T}_1, \hat{T}_2) = (9.47, 1.27)$. For these values of $T_1, T_2$, and $\Delta$ power $(T_1, T_2) \approx 1.00$ and power $(\Delta) \approx .10$ based on the available messy data. If we hypothetically consider the case of Group 1 and Group 2 receiving sequence $(A, P)$ and Group 3 and Group 4 receiving sequence $(P, A)$ while the pattern of missing values remains unchanged, then more data would be usable in the traditional analysis. Then we would have power $(\Delta) \approx .15$ while power $(T_1, T_2) \approx 1.00$ as before.

Thus for these particular data which involve missing
values and an unbalanced design, the power of the "treatment
effects" test in the by-period analysis is greater than the
power of the analogous test in the traditional analysis.

For the complete set of data with no values missing it is
also easy to see from the results of Owen's graph that there are
not nearly enough degrees of freedom to give 0.95 power against
the alternative hypothesis in the traditional model. By comparison,
power in the by-period analysis is well above 0.95 for this
particular experiment.
CHAPTER 5
RECOMMENDATIONS FOR FURTHER RESEARCH

Several outstanding problems are readily apparent. These will require both analytic and numerical methods for further study.

In Chapter 2 the MacLaurin expansion was seen to converge for some parameter values but not others. Convergence is an unknown function of the parameters. For completeness, some usefully sharp bounds on the residual term would be desirable. This however is not an important problem since the Monte Carlo integration algorithm performs as well as the MacLaurin algorithm when only a few significant digits of accuracy are required. A larger contribution would be the development of an analytic expression for the power function, \( P(t) \), which could be translated into an efficient numerical algorithm. Another contribution to be made is the extension of the various results to the trivariate case (i.e., three-period CO designs.) The trivariate distribution of the denominators of \( t_1, t_2, \) and \( t_3 \) is not readily available although Johnson and Kotz (1972) have suggested an approximation of the p.d.f. Extensions to the trivariate case could also be made in Chapter 3 and probably
with more success than in Chapter 2.

The distribution of the test statistic \( \max |t_i| \) is not known explicitly. Further study could yield results on that distribution or could produce a new test statistic having superior qualities and characterization.

In Chapter 3 further study using the Monte Carlo simulation algorithm as a tool could take many forms. The algorithm is sufficiently general to make a wide variety of probability calculations involving the distribution of \((t_1, t_2)\). Further work regarding the distribution theory in Chapter 3 should also provide rewards.

In Chapter 4 we have not addressed the question of how to perform and analyze four-sequence experiments with respect to the number of study units per group. How do balanced designs compare with unbalanced designs? Some of the theoretical tools developed in Chapters 2 and 3 could be used to study the relative merits of different \(CO(2, 4, 2)\) designs.

It should be noted that other alternative by-period models could be fitted to (possibly messy) data. For example, since the design matrices may be different, the following model offers no new problems with respect to estimation, testing, power calculation and significance level calculation:
\[ E[Y_1] = [X_1, Z_1] \cdot [\beta_1] \]
\[ E[Y_2] = [X_2, Z_2] \cdot [\beta_2] \]

\[ \text{COV} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 I_{n_1} & \sigma_1 \sigma_2 \rho B \\ \sigma_1 \sigma_2 \rho B & \sigma_2^2 I_{n_2} \end{bmatrix} \]

where \( B \) is of the form \([I_q, 0]^T \cdot [I_q, 0] \).

This model gives an example of covariate adjustment: the end-of-period measurement is conditional on experimental design and covariates given by \( Z_1 \) and \( Z_2 \). Further investigation into the usefulness of this idea is recommended.
APPENDIX

A.1 Bean-Sprout Experiment Protocol

The objective of this crossover study was to determine the effect of 2,4 di-chloro-phenoxy acetic acid (2,4D) on the growth of 132 mung bean sprouts. Because of their young and fast-growing tissues, sprouts have been given an important role in biological and biomedical research. 2,4D is a hormone-like reagent which can dramatically influence growth in young plant tissues. Like many other growth-control agents, low doses of 2,4D have positive effects on growth while high doses are detrimental. 2,4D is popularly used as a herbicide. In this study sprouts were exposed to a moderately detrimental dose. The response observed was the length of the root sprouted by each bean.

The study was conducted as a two-period, two-treatment crossover with four sequences. Each period was 28 hours long. The two treatments were tap-water (treatment B) as a placebo, and a $10^{-4}$ molar solution of 2,4D (treatment A). In order to keep the 2,4D in solution the solution was .06% ethanol by volume. The solution had the same appearance as tap-water. The four treatment sequences (A,A), (A,P), (P,A), and (P,P) were applied to groups 1, 2, 3 and 4 respectively. Each group consisted of 34 sprouts which were randomly assigned.

An arbitrary collection of more than 200 dry mung beans was placed in a bowl of tap-water located within a small insulated
Figure A.1.1

Tracing of array of holes on stainless-steel screen used for growing sprouts (actual size).

85 mm
chest. The beans were left to soak at room temperature and in darkness for 24 hours. At the end of the 24th hour (T=24) the germinated seeds had produced tiny root sprouts approximately four mm long. At that point any bean which showed no evidence of germination was rejected as a possible study subject. The beans were then assigned at random to the four groups. The first bean picked up was assigned to group 1, the second picked up was assigned to group 2, etc..., until each group had 33 germinated beans. Each group was placed on one of four identical growing-jars. Each jar was constructed by placing a stainless-steel screen disk over the mouth of a 1-pint, wide-mouth, clear glass, Mason jar. Each screen was perforated with an array of 37 holes which were 5 mm in diameter. The beans were placed sprout-downward over the holes and the jars were filled to the brim with tap water. Thus within each group every bean could be identified by its array location. The four growing-jars with sprouted-beans in place were placed together in the darkness of the room temperature chest mentioned above. The sprouts were left to lengthen and were disturbed only to ascertain that the sprouted roots were all growing downward into the water through the holes in the screens. This assured that the roots would be as straight as possible for easy handling and measurement. At T=45 the shedding hushs were removed. At T=49 a random sample of 24 sprouts (6 per group) was selected for measurement of length.
At T=72 (3:00 pm 15 July 1981) the water in the four jars was replaced by treatments A or P according to design. It was possible to conduct the experiment as a "double-blind" study by enlisting the help of an assistant. Period 1 lasted 28 hours (T=72 until T=100). Sprout length was measured at T=100. Since the sprouts were growing nearly one millimeter per hour it was necessary to randomize the sequence of measurements of the sprouts. Measurement of length was made by removing the sprout from its array location and laying it on a flat surface alongside a rule. Length was accurately measured to the nearest millimeter. If the sprout was bent or curly then the sprout was laid on the flat surface in such a way as to give the maximum measurement with the restriction that it would be straightened only by its own weight. No attempt was made to otherwise straighten crooked sprouts. When not being measured or inspected the four groups were stored in the darkness of the chest.

At T=100 the four groups were simultaneously rinsed with tap-water to thoroughly remove period 1 treatment solutions. The four jars were then refilled with tap-water and replaced in the chest for the remainder of the wash-out time period. At T=121 period 2 was begun. Treatments were applied and methods mentioned above were used again. At T=140 period 2 ended and measurements were made.

Throughout the study the following methods were practiced: (1) all tapwater used was room-temperature; (2) the four groups
were stored together in darkness when not being observed, (3) the four groups were kept in very close proximity to one another at all times, (4) root length was defined as the length of an imaginary line segment beginning where the root emerges from the bean and ending at the root tip.

At the end of period 1 it was patently obvious which treatment each group had received. Treatment A caused the sprouts to have thick, short roots and slightly discolored and wilted leaves. (During the wash-out time period every bean produced the beginnings of a stem and two leaves.)

For analysis purposes the unit of study is taken to be a triplet of sprouts rather than the individual sprout. In each of the four treatment groups there are eleven triplets; and thus, N=44 units of study in all. The variables to be analyzed are triplet-averages of sprout length. The data can therefore be arrayed as 44 observations on the variables Y1, Y2 and GROUP. Figure A.1.3 lists the values of this array.

Figure A.1.4 presents simple descriptive statistics for the 11 triplets in each group. Within each group the Kolmogorov-Smirnov goodness-of-fit test for normality was performed on Y1 and Y2. The p-values for these two tests were both greater than .20. It will henceforth be assumed that these variables are approximately normally distributed within each group.
<table>
<thead>
<tr>
<th>OBS</th>
<th>Y1</th>
<th>Y2</th>
<th>Group</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>53.66</td>
<td>74.33</td>
<td>1</td>
<td>(A,A)</td>
</tr>
<tr>
<td>2</td>
<td>45.66</td>
<td>74.33</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>54.33</td>
<td>84.33</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>44.00</td>
<td>62.00</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>46.00</td>
<td>68.66</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>49.33</td>
<td>72.66</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>38.00</td>
<td>58.00</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>50.33</td>
<td>70.00</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>48.33</td>
<td>69.00</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>42.66</td>
<td>58.00</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>48.33</td>
<td>68.00</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>51.00</td>
<td>75.33</td>
<td>2</td>
<td>(A,P)</td>
</tr>
<tr>
<td>13</td>
<td>55.66</td>
<td>84.33</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>44.00</td>
<td>74.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>41.66</td>
<td>64.66</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>39.33</td>
<td>55.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>54.33</td>
<td>84.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>52.33</td>
<td>67.33</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>48.66</td>
<td>77.33</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>53.00</td>
<td>87.33</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>46.66</td>
<td>62.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>50.00</td>
<td>78.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>66.00</td>
<td>109.66</td>
<td>3</td>
<td>(P,A)</td>
</tr>
<tr>
<td>24</td>
<td>81.00</td>
<td>132.33</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>62.66</td>
<td>115.66</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>89.33</td>
<td>156.00</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>58.00</td>
<td>108.66</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>70.33</td>
<td>123.66</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>82.00</td>
<td>135.33</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>52.33</td>
<td>84.00</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>80.33</td>
<td>133.33</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>74.66</td>
<td>120.33</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>54.33</td>
<td>94.00</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>69.33</td>
<td>124.33</td>
<td>4</td>
<td>(P,P)</td>
</tr>
<tr>
<td>35</td>
<td>74.33</td>
<td>135.33</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>66.66</td>
<td>123.00</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>77.00</td>
<td>117.00</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>58.00</td>
<td>104.66</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>53.00</td>
<td>85.00</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>68.33</td>
<td>120.00</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>61.33</td>
<td>103.66</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>38.66</td>
<td>82.33</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>88.00</td>
<td>143.00</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>59.00</td>
<td>129.66</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>
## Figure A.1.4

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Standard Error</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y1</td>
<td>11</td>
<td>47.33</td>
<td>4.78</td>
<td>1.44</td>
<td>38.00</td>
<td>54.33</td>
</tr>
<tr>
<td>Y2</td>
<td>11</td>
<td>69.03</td>
<td>7.74</td>
<td>2.33</td>
<td>58.00</td>
<td>84.33</td>
</tr>
<tr>
<td>Group 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y1</td>
<td>11</td>
<td>48.39</td>
<td>5.31</td>
<td>1.60</td>
<td>39.33</td>
<td>55.67</td>
</tr>
<tr>
<td>Y2</td>
<td>11</td>
<td>73.58</td>
<td>10.24</td>
<td>3.09</td>
<td>55.00</td>
<td>87.33</td>
</tr>
<tr>
<td>Group 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y1</td>
<td>11</td>
<td>70.09</td>
<td>12.40</td>
<td>3.74</td>
<td>52.33</td>
<td>89.33</td>
</tr>
<tr>
<td>Y2</td>
<td>11</td>
<td>119.36</td>
<td>20.24</td>
<td>6.10</td>
<td>84.00</td>
<td>156.00</td>
</tr>
<tr>
<td>Group 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y1</td>
<td>11</td>
<td>64.88</td>
<td>13.15</td>
<td>3.96</td>
<td>38.67</td>
<td>88.00</td>
</tr>
<tr>
<td>Y2</td>
<td>11</td>
<td>115.27</td>
<td>19.48</td>
<td>5.87</td>
<td>82.33</td>
<td>143.00</td>
</tr>
</tbody>
</table>
A.2 Algorithm: Power via MacLaurin Polynomials

```plaintext
//POWER1 JOB UNC.B.S665V,STEWART,WASTE=YES,M=1,
// DEST=UNCPREVU,
// PRTY=0,
// PAGES=25,
// T=(2.00)
//PW=PS
//STEP1 EXEC FPGCG,REGION=300K
//SYSIN DD *
C
C...#...1...#...2...#...3...#...4...#...5...#...6...#...7
C
C**** DECLARATION STATEMENTS ***
DOUBLE PRECISION RESULT, THETA1, THETA2, SD1, SD2,
  1 RH0, T, TT2, XM, XME2, RHXME2,
  2 RHOXM, RHOE2, HALFNU, PI, CONST, TS1,
  3 TS2, TERM1, TERM2, TERM3, TERM4, TS1E2,
  4 TERM, TS2E2, EPSLNK, EPSLNH, SUMH, SUMI, SUMJ,
  5 SUMK, SUML, SUMN, SUMP, SUMQ, SUMR, SUMS,
  6 RXXXH, TERMH, TERMIN, TERMJ, TERMS, TERMN, TERMP,
  7 RXXXR, TERMQ, TERMQ, TERMS, TERMK, LASTH, LASTJ,
  8 LASTK, LASTL, LASTN, LASTP, LASTQ, LSTR, LASTS,
  9 RXXXQ1, RXXXQ2, RXXXQ3, RXXXQ4, RXXXS, RXXXK
C**** FUNCTION STATEMENTS ***
C**** DATA STATEMENTS ***
C**** INPUT VALUES ***
THETA1 = 23.065125
THETA2 = THETA1
SD1 = 10.6216750
SD2 = 8.8300622
RH0 = -.117
NU = 133
T = 2.3
XM = 0.0
N1 = 100
N2 = 5
EPSLNK = (10.0)**(-6.0)
EPSLNH = (10.0)**(-6.0)
ISUPPH = 2
C
```
C**** INITIALIZE VALUES ...
PI = 3.1415926535898
TS1 = THETA1/SD1
TS2 = THETA2/SD2
XM2 = XX*XM
RHOE2 = RHO*RHO
RHOXM = RHO*XM
RHXM2 = RHOXM*RHOXM
TS1E2 = TS1*TS1
TS2E2 = TS2*TS2
T2 = T*T*2.
TERM1 = TS1 - RHOXM*TS2
TERM2 = TS2 - RHOXM*TS1
TERM3 = 1.0 - RHOE2
TERM4 = 1.0 - RHXM2
HALFNU = 0.5*FLOAT(NU)

C**** PRINT THE INPUT PARAMETERS ...
PRINT 1
1 FORMAT (10X,'---------INPUT PARAMETERS FOR P(T;PARMS)---------')
PRINT 2, SD1, SD2
2 FORMAT (10X,SD1 = 'D18.10,5X,SD2 = 'D18.10)
PRINT 3, RHO, NU
3 FORMAT (10X, RHO = 'D18.10,5X,NU = 'I17 )
PRINT 4, T, XM
4 FORMAT (10X,T = 'D18.10,5X,XM = 'D18.10)
PRINT 5, THETA1, THETA2
5 FORMAT (10X,THETA1 = 'D18.10,5X,THETA2 = 'D18.10)
PRINT 6, N1, N2
6 FORMAT (10X,N1 = 'I17 ,5X,N2 = 'I17 )
PRINT 7, TS1, TS2
7 FORMAT (10X,TS1 = 'D18.10,5X,TS2 = 'D18.10)
PRINT 8, TERM1, TERM2
8 FORMAT (10X,TERM1 = 'D18.10,5X,TERM2 = 'D18.10)
PRINT 9, TERM3, TERM4
9 FORMAT (10X,TERM3 = 'D18.10,5X,TERM4 = 'D18.10)
PRINT 10, EPSLNK, ISUPFK
PRINT 11, EPSLNH, ISUPPH
10 FORMAT (10X,EPSLNK = 'D18.10,5X,ISUPFK = 'I17 )
11 FORMAT (10X,EPSLNH = 'D18.10,5X,ISUPPH = 'I17 )

C
C**** COMPUTE THE LOG OF THE OVERALL CONSTANT ----
CONST = DEXP(( TERM1*TS1 + TERM2*TS2 ) / ((-2.0)*TERM4))
CONST = CONST*(TERM3**HALFNU)*(TERM4**1.5)/PI
CONST = DLOG (CONST) - 2.0*DLGAMA (HALFNU)
PRINT 12, CONST
12 FORMAT (10X,'LOG (OVERALL CONSTANT) = 'D18.10)
C
SMALL  = DLOG(EPSLNH)
PRINT 13, SMALL
13 FORMAT (10X,'CONVERGE H: LOG( TERM**2 ) < SMALL=*,D10.10)

C SMALL  = 2.0*SML

C
C**** COMPUTE THE NESTED SUMS . . .

C
C**** I.  BEGIN LOOP SUMH . . .

C
SUMH  = 0
RXXXH = TT2*TERM3 / (FLOAT(NU)*(TERM4**2))
ICUMH = 0
LASTH = 1
DO 100 IH=1,N2
     TERMH = LASTH*RXXXH / FLOAT(2*IH*(2*IH - 1))
100

C
C**** II.  BEGIN LOOP SUMI . . .

C
SUMI  = 0
MAXI  = 2*(IH - 1) + 1
DO 9900 II=1,MAXI
   II  = II - 1
9900

C
C**** III.  BEGIN LOOP SUMJ . . .

C
SUMJ  = 0
IXXXJ = 2*(IH - 1) -II
MAXJ  = III
IF (II .GE. IH) MAXJ = IXXXJ + 1
TERMJ = 1
DO 8900 IIJ=1,MAXJ
   IJ  = IIJ - 1
   IF (IJ .EQ. 0) GO TO 8100
   TERMJ = FLOAT( (IXXXJ + 1 - IJ)*(II + 1 - IJ) )
   TERMJ = TERMJ * LASTJ * TERM4 / FLOAT(IJ)
8100 CONTINUE

C
C**** IV. BEGIN LOOP SUML ...
C
SUML = 0.
IXXXL = II - IJ
MAXL = ( IXXXL / 2 ) + 1
TERML = 1.
DO 7900 IIL=1,MAXL
    IL = IIL - 1
    ITMPQL = IXXXL - 2*IL
    IF (IL.EQ.0) GO TO 7100
    TERML = FLOAT( (ITMPQL + 1) * (ITMPQL + 2) )
    TERML = TERML / FLOAT( IL )
    TERML = TERML*(-.5)*TERM4
    TERML = TERML*LASTL
7100 CONTINUE
C
C**** V. BEGIN LOOP SUMN ...
C
SUMN = 0.
IXXXN = IXXXJ - IJ
MAXN = ( IXXXN / 2 ) + 1
TERMN = 1.
DO 6900 IIN=1,MAXN
    IN = IIN - 1
    IF (IN.EQ.0) GO TO 6100
    TERMN = FLOAT(( IXXXN + 1 - 2*IN) * ( IXXXN + 2* (1 - IN) ) )
    TERMN = TERMN / FLOAT(IN)
    TERMN = TERMN * TERM4 * LASTN
6100 CONTINUE
C
C**** VI. BEGIN LOOP SUMP ...
C
SUMP = 0.
MINP = 1
MAXP = IIJ
IF (RH0XM.EQ.0.) MINP = MAXP
TERMP = 1.
DO 5900 IIP=MINP,MAXP
    IP = IIP - 1
    IF ((IP.NE. IJ ).AND. (IP.NE. 0 ))
        1    TERMP = LASTP * FLOAT( IP - IJ - 1 ) / FLOAT(IP)
    LASTP = TERMP
    IF (IP.EQ. IJ) TERMP = 1.0
    IF ((IP.EQ. IJ ).AND. (IP.NE. 0 ))
        1    TERMP = (-1.0)**( IJ )
    IF ((IP.NE. IJ ).AND. (RH0XM.NE. 0.))
        1    TERMP = TERMP * (RH0XM**( IJ - IP ))
C
C**** VII. BEGIN LOOP SUMQ ...

C
SUMQ = 0.
C**** ITMPQL = II - IJ - 2*IL
ITMPQN = IXXXN - 2*IN
INQ = 1
IF (((TERM1.EQ.0.) .OR. (TERM2.EQ.0.).) .AND. (ITMPQN .GE. 1))
1
INQ = ITMPQN
IMAXQ = ITMPQN + 1
DO 4900 IQ = 1, IMAXQ, INCQ
IQ = IQ - 1
IQL = ITMPQL + IQ
IQN = ITMPQN - IQ
TERMQ = 1.
IF (((IQ.EQ.0.).) .OR. (IQ.EQ. ITMPQN)) GO TO 4100
IF (((TERM1.EQ.0.).) .OR. (TERM2.EQ.0.).) GO TO 4100
TERMQ = LASTQ * FLOAT(IQN + 1) / FLOAT(IQ)
4100 CONTINUE
LASTQ = TERMQ
C
IF (IQL.EQ.0.)
IF ((IQL.LE.0.).AND.(TERM1.LE.0.).) RXXQ1 = TERM1**IQL
RXXQ2 = 0.
RXXQ2 = 1.
IF ((IQN.EQ.0.).) RXXQ3 = TERM2**IQN
RXXQ3 = 0.
RXXQ3 = 1.
IF ((IQL.EQ.0.).) RXXQ4 = TERM1**IQL
RXXQ4 = 0.
RXXQ4 = 1.
C
TERMQ = TERMQ * ( RXXQ1*RXXQ2 + RXXQ3*RXXQ4 )
IF (TERM.EQ.0.) GO TO 4200
C
C**** VIII. BEGIN LOOP SUMR ...

C
SUMR = 0.
MINR = 1
MAXR = IIN
IF (RHOMX.EQ.0.) MINR = MAXR
IXXXR = II + IN + IP + IQ
DO 3900 IIR = MINR, MAXR
IR = IIR - 1
TERMR = 1.
IF ((IR.LE.0.).) (IR.LE.IN))
1
TERMR = LASTR * (-.5) * FLOAT(MAXR - IR) / FLOAT(IR)
LASTR = TERMR
IXXX = IXXXR - IR
IF (IXXX.LE. (2*(IXXX/2))) GO TO 3200
RXXXR = DGLAMA( FLOAT(IH) + HALPN - .5*FLOAT(IXXX + 1) )
RXXXR = RXXXR + DGLAMA(HALPN + .5*FLOAT(IXXX + 1) )
RXXXR = 2.*DEXP( RXXXR + CONST )
IF ((IR.LE.IN) .AND. (RHOMX.LE.0.).)
1
RXXXR = RXXXR * (RHOMX**(IN - IR))
IF (IR.EQ.IN) RXXXR = RXXXR *((-5)**IN)
TERMR = (TERMR) * RXXXR
C
C**** IX. BEGIN LOOP SUMS ...
C
SUMS = 0.
MAXS = IIR
RXXXS = .5 * FLOAT( IXXXR - IR )
TERMS = 1.
DO 2900 IIS=1,MAXS
IS = IIS - 1
IF (IS .EQ. 0) GO TO 2100
TERMS = ( HALFNU -.5 + RXXXS + FLOAT(IS) ) / FLOAT( IS )
TERMS = TERMS / ( HALFNU -.5 - RXXXS + FLOAT(IH - IS) )
TERMS = TERMS * LASTS * FLOAT( MAXS - IS )
2100 CONTINUE
C
C**** X. BEGIN LOOP SUMK ...
C
SUMK = 1.
MAXK = 0
IF (RHO .EQ. 0.) GO TO 1400
LASTK = 1.
RXXXK = RXXXS + FLOAT(IS)
DO 1200 IK=1,N1
TERMK = HALFNU -.5 + RXXXK + FLOAT(IK)
TERMK = TERMK * (HALFNU - 1.5 - RXXXK + FLOAT(IH + IK))
TERMK = TERMK * RHOE2 / FLOAT(IK)
TERMK = TERMK / ( HALFNU + FLOAT(IK - 1) )
TERMK = TERMK * LASTK
SUMK = SUMK + TERMK
IF (TERMK .LT. EPSLNK) GO TO 1300
LASTK = TERMK
1200 CONTINUE
PRINT 1201, LASTK, TERMK, SUMK
1201 FORMAT (5X,'LOOP K NOT CONVERGED:LASTK=',D18.10,
     1 ' TERMK=',D18.10,' SUMK=',D18.10)
1300 CONTINUE
IF (MAXIK .LT. IK) MAXIK = IK
1400 CONTINUE
C
C**** (X. END LOOP K ) ...
C
SUMS = SUMS + TERMS*SUMK
LASTS = TERMS
2900 CONTINUE
C
C**** (IX. END LOOP S ) ...
C
SUMR = SUMR + ( TERMR * (SUMS) )
3200 CONTINUE
3900 CONTINUE
C
C**** (VIII. END LOOP R ) ...
SUMQ = SUMQ + TERMQ*SUMR
4200 CONTINUE
4900 CONTINUE

C**** (VII. END LOOP Q) ... 
C
SUMP = SUMP + TERM*SUMQ
5900 CONTINUE

C**** (VI. END LOOP P) ...
C
SUMN = SUMN + TERMN*SUMP
LASTN = TERMN
6900 CONTINUE

C**** (V. END LOOP N) ...
C
SUML = SUML + TERML*SUMN
LASTL = TERML
7900 CONTINUE

C**** (IV. END LOOP L) ...
C
SUMJ = SUMJ + TERMJ*SUML
LASTJ = TERMJ
8900 CONTINUE

C**** (III. END LOOP J) ...
C
SUMI = SUMI + SUMJ
C
9900 CONTINUE
C**** (II. END LOOP I) ...
C
SUMH = SUMH + TERMH*SUMI
TERM = TERMH*SUMI
IF ((DLOG(TERM**2)) .LT. SMALL) ICUMH = ICUMH + 1
C***
IF ((IH .GT. 3) .AND. (ICUMH .EQ. 0)) GO TO 194
PRINT 193, IH, MAXIK, TERM
193 FORMAT (9X,'IH=',I9,' MAXIK=',I9,' TERM=',D18.10)
C*194 CONTINUE
IF ((DLOG(TERM**2)) .GE. SMALL) ICUMH = 0
IF (ICUMH .GE. ISUFFH) GO TO 190
LASTH = TERMH
C
100 CONTINUE
190 CONTINUE
PRINT 191, ICUMH, SUMH
191 FORMAT (10X,'LOOP H ENDED: ICUMH=' ,I17, ', SUMH=' ,D18.10)
PRINT 192, IH, LASTH, TERMH
192 FORMAT (10X,'IH=' ,I9, ', LASTH=' ,D18.10, ', TERMH=' ,D18.10)
PRINT 195, MAXIK
195 FORMAT (10X,'MAXIK=' ,I9)

C
C**** (I. END LOOP H ) ...
C
C
C**** COMPUTE THE END RESULT P(T) ...
RESULT = SUMH
PRINT 9999, RESULT
9999 FORMAT (10X,'RESULT: P(T;PARMS)=' ,D18.10)
C**** END THE PROGRAM ...
END
/*
A.3 Algorithm: Power via Monte Carlo Integration

```plaintext
//POWER3 JOB UNC.B.E518U, STEWART, WASTE=NO, M=1,
// PRTY=0,
// PAGES=100,
// T=(0,30)
//STEP1 EXEC FTGCG, REGION=600K
//SYSIN DD *
C
C 111111 111111 111111 111111 111111 111111 111111
C
C **** DECLARATION STATEMENTS ...
REAL R (40000)
DOUBLE PRECISION RESULT, THETA1, THETA2, SD1, SD2,
1 RHO, T, TT4, XM, XME2, RXME2,
2 RHOXM, RHOE2, HALFNU, PI, CONST, TS1,
3 TS2, TERM1, TERM2, TERM3, TERM4, TS1E2,
4 TS2E2, EPSLN, DSEED, SUMJ, LASTJ, DIFF,
5 RTERM1, RTERM2, SMALL, TEMPO, PCN,
6 X1, X2, Y1, Y2, T1, T2,
7 RSUMK, TERMK, LASTK, SERIES, EPSLNJ
C **** FUNCTION STATEMENTS ...
C **** DATA STATEMENTS ...
C **** INPUT VALUES FOR MONTE CARLO METHOD ...
DSEED = 123457.D0
N = 40000
ICHECK = 2500
EPSLNJ = (100.0)**(-2)
C **** INPUT PARAMETER VALUES ...
THETA1 = 39.799497
THETA2 = 6.63324
SD1 = 9.7534609
SD2 = 15.414603
RHO = 0.892
NU = 41
T = 2.32672
XM = 0.0
N1 = 7
EPSLN = (10.0)**(-2)
C
```
** Initialize Values ... **

```c
PI = 3.1415926535898
TS1 = THETA1/SD1
TS2 = THETA2/SD2
XME2 = XM**2
RHOE2 = RHO**2
RHOXM = RHO*XMIN
RHXME2 = RHOXM**2
TS1E2 = TS1**2
TS2E2 = TS2**2
T4 = T*T4.
TERM1 = TS1 - RHOXM*TS2
TERM2 = TS2 - RHOXM*TS1
TERM3 = 1.0 - RHOE2
TERM4 = 1.0 - RHXME2
HALFNU = 0.5*FLOAT(NU)
RTERM1 = (-.5) * TERM3 / (FLOAT(NU) * TERM4)
RTERM2 = ( (TERM3 / FLOAT(NU))**(.5) ) / TERM4
```

** Print the Input Parameters ... **

```c
C
1 FORMAT (10X,'______INPUT PARAMETERS FOR P(T;PARMS)______')
PRINT 2, SD1, SD2
2 FORMAT (10X,'SD1 = ',D18.10,'5X','SD2 = ',D18.10)
PRINT 3, RHO, NU
3 FORMAT (10X,'RHO = ',D18.10,'5X','NU = ',I17)
PRINT 4, T, XM
4 FORMAT (10X,'T = ',D18.10,'5X','XM = ',D18.10)
PRINT 5, THETA1, THETA2
5 FORMAT (10X,'THETA1 = ',D18.10,'5X','THETA2 = ',D18.10)
PRINT 6, N1, N2
6 FORMAT (10X,'N1 = ',I17,'6X','N2 = ',I17)
PRINT 7, TS1, TS2
7 FORMAT (10X,'TS1 = ',D18.10,'5X','TS2 = ',D18.10)
PRINT 8, TERM1, TERM2
8 FORMAT (10X,'TERM1 = ',D18.10,'5X','TERM2 = ',D18.10)
PRINT 9, TERM3, TERM4
9 FORMAT (10X,'TERM3 = ',D18.10,'5X','TERM4 = ',D18.10)
PRINT 10, RTERM1, RTERM2
10 FORMAT (10X,'RTERM1 = ',D18.10,'5X','RTERM2 = ',D18.10)
PRINT 11, EPSLN
11 FORMAT (10X,'EPSLN = ',D18.10)
```

** Compute the Overall Constant ... **

```c
CONST = DEXP((TERM1*TS1 + TERM2*TS2) / ((-2.0)*TERM4))
CONST = CONST * (TERM3**(HALFNU + 1)) * (TERM4**(.5))
CONST = CONST / FLOAT(NU) * PI * (2.0**(NU-1))
CONST = CONST / ( (DGAMMA(HALFNU))**(2) )
CONST = CONST * T4
PRINT 12, CONST
```

```c
12 FORMAT (10X,'THE OVERALL CONSTANT IS ',D18.10)
```
SMALL = ( EPSLN / CONST )**2
PRINT 13, SMALL

13 FORMAT (10X,'CONVERGENCE: TERMK .LT. SMALL = ',D18.10)

CALL GGBUS (DSEED,N,R)

IMAX = N
IF ( RHO .EQ. 0. ) IMAX = 1
SERIES = 0.0
JJMAX = N/4
LASTK = 1.0
DO 7000 IK=1,IMAX
   IK = IK - 1
   TERMK = 1.0
   IF (IK .EQ. 0) GO TO 50
   TERMK = LASTK * ( (.5)*RHO)**2
   TERMK = TERMK / ( FLOAT(IK) * (HALFNU + FLOAT(IK-1)) )
7000 CONTINUE
LASTK = TERMK

C***** COMPUTE THE INTEGRAL FOR X**IK ...
NU2IK = NU + 2*IK
SUMJ = 0.
KOUNT = 0.
LASTJ = 0.
DO 1000 JJ=1,JJMAX
   J = 4*JJ
   Y1 = R(J-3) X1 = Y1/(1-Y1)
   Y2 = R(J-2) X2 = Y2/(1-Y2)
   Y3 = R(J-1) T1 = T*(2*Y3 - 1)
   Y4 = R(J ) T2 = T*(2*Y4 - 1)
C
   Y1 = R(J-3)
   Y2 = R(J-2)
   X1 = Y1/(1.0 - Y1)
   X2 = Y2/(1.0 - Y2)
   T1 = T *(2.0*R(J-1) - 1.0)
   T2 = T *(2.0*R(J ) - 1.0)
C**** COMPUTE THE INTEGRAND FCN ...
FCN = 0.
TEMPQ = (X1*T1)**2 + (X2*T2)**2 - 2.*X1*X2*T1*T2*RHOKM
TEMPQ = TERM1*TEMPQ
TEMPQ = TEMPQ + RTERM2*( X1*T1*TERM1 + X2*T2*TERM2 )
TEMPQ = TEMPQ + (-5.)*( X1**2 + X2**2 )
TEMPQ = TEMPQ + FLOAT(NU2I?) DLOG((X1*X2))
TEMPQ = TEMPQ + (2.0)*DLOG(((1.+X1)*(1.+X2)))
IF ( TEMPQ - LT. (-180.218) ) GO TO 75
FCN = DEXP( TEMPQ )
KOUNT = KOUNT + 1
75 CONTINUE
SUMJ = SUMJ + FCN
IF ( JJ .LT. 5000 ) GO TO 100
IF ( (ICHECK*(JJ/ICHECK)) .NE. JJ ) GO TO 100
DIFF = (SUMJ/FLOAT(JJ)) - (LASTJ/FLOAT(JJ-1))
IF ( (DIFF**2) .LE. EPSLNJ ) GO TO 2000
LASTJ = SUMJ
100 CONTINUE
1000 CONTINUE
2000 CONTINUE
C
C**** COMPUTE FINAL RIEMANN SUM AS RSUMK ...
RSUMK = SUMJ / FLOAT(JJ)
C**** USE THE INTEGRAL VALUE RSUMK ...
TERMK = TERMK * RSUMK
SERIES = SERIES + TERMK
PRINT 3000, IK, TERMK, JJ, KOUNT
3000 FORMAT (1X,'K = ',I3,' TERMK= ',D18.10,' JJ= ',I10,' #OK= ',I10)
IF ( (TERMK**2) .LE. SMALL ) GO TO 8000
7000 CONTINUE
8000 CONTINUE
C
C**** CUMPUTE THE END RESULT P(T) ...
RESULT = CONST* SERIES
PRINT 9001, RESULT
9001 FORMAT (10X,'THE RESULT IS P(T;PARMS) = ',D18.10)
C**** END THE PROGRAM ...
END
/
A.4 Algorithm: Power and Significance Level via Monte Carlo Simulation

```plaintext
//POWER4 JOB UNC_B=E518U, STEWART, WASTE=YES, M=1,
// PRTY=0,
// PAGES=100,
// T=(0,30)
/*PW=PAPER
//STEP1 EXEC FTGCG, REGION=300K
//SYSIN DD *
C
C
C...
C**** DECLARATION STATEMENTS...
REAL R(4200)
INTEGER D(41)
DOUBLE PRECISION THETA1, THETA2, SD1, SD2,
1 RHO, T1MIN, T1MAX, T2MIN, T2MAX, XM,
2 RHOHM, PERCNT,
3 TS1, TS2, TERM4, TERM5, Q1, Q2,
4 DSEED, W1, W2,
5 Z1K, Z2K, DL, DLRHO, T1, T2

C**** FUNCTION STATEMENTS...
C**** DATA STATEMENTS...
DATA D/24*1/
C**** INPUT VALUES FOR MONTE CARLO METHOD...
J1MAX = 50
J2MAX = 10
DSEED = 123457.D0
ICHECK = 5
MINJ1 = 10
C**** INPUT PARAMETER VALUES...
TAU1 = 7.000
TAU2 = 1.27000
THETA1 = TAU1 * 6.3245554
THETA2 = TAU2 * 4.8818199
SD1 = 9.5231664
SD2 = 17.407383
RHO = 0.892
XM = 0.0684005
NU1 = 38
NU2 = 21
DO 111 ID=1,4
D(ID) = 0
111 CONTINUE
T1MIN = -2.33372
T1MAX = 2.33372
T2MIN = -2.41385
T2MAX = 2.41385
```
C
C**** INITIALIZE VALUES ...
C

PT  = 3.1415926535898
NU  = NU1 + NU2 + 2
TS1 = THETA1/SD1
TS2 = THETA2/SD2
RHOXM = RHO*XM
TERM4 = 1.0 - RHOXM**2
N   = J1MAX * NU
CALL GGNML (DSEED, N, R)

C**** PRINT THE INPUT PARAMETERS ...
PRINT 1
1 FORMAT (10X,'_________INPUT PARAMETERS FOR P(T;PARMS)_______')
PRINT 2, SD1, SD2
2 FORMAT (10X,'SD1  = ',D18.10,' , SD2  = ',D18.10)
PRINT 3, NU1, NU2
3 FORMAT (10X,'NU1  = ',I5,' , NU2  = ',I5)
PRINT 4, RHO, XM
4 FORMAT (10X,'RHO  = ',D18.10,' , XM  = ',D18.10)
PRINT 5, TAU1, TAU2
5 FORMAT (10X,'TAU1  = ',D18.10,' , TAU2  = ',D18.10)
PRINT 7, TS1, TS2
7 FORMAT (10X,'TS1  = ',D18.10,' , TS2  = ',D18.10)
PRINT 8, T1MIN, T1MAX
8 FORMAT (10X,'T1MIN = ',D18.10,' , T1MAX = ',D18.10)
PRINT 9, T2MIN, T2MAX
9 FORMAT (10X,'T2MIN = ',D18.10,' , T2MAX = ',D18.10)
KOUNT1 = 0
KOB5 = 0
DO 2000 J2=1,J2MAX
   IF (J2.EQ.1) GO TO 101
   DSEED = DSEED - 1.D0
   CALL GGNML (DSEED,N,R)
101 CONTINUE
   DO 1000 J1=1,J1MAX
      K1MIN = NU*(J1-1) + 1
      W1 = R(K1MIN)
      W2 = R(K1MIN + 1)
      W2 = RHOMX*W1 + DSQRT(TERM4)*W2
      W1 = W1 + TS1
      W2 = W2 + TS2
      K1MIN = K1MIN + 2
      K1MAX = K1MIN + NU1 - 1
      K2MIN = K1MAX + 1
      K2MAX = K2MIN + NU2 - 1
      Q1 = 0.0
      DO 100 K1=K1MIN,K1MAX
         Q1 = Q1 + ( R(K1) )**2
      100 CONTINUE
      Q2 = 0.0
      L1 = 1
      DO 200 K2=K2MIN,K2MAX
         Z1K = R( K2-NU1 )
         Z2K = R( K2 )
         DL = D( L1 )
         DLRHO = DL * RHO
         TERM5 = 1.0 - DLRHO**2
         L1 = L1 + 1
         Q2 = Q2 + (Z1K*DLRHO)**2
         Q2 = Q2 + 2*K*Z2K*TERM5
         Q2 = Q2 + Z1K*Z2K*2.0*DLRHO*DSQRT(TERM5)
      200 CONTINUE
      T1 = W1 / DSQRT( Q1/FLOAT(NU1) )
      T2 = W2 / DSQRT( Q2/FLOAT(NU2) )
C***
   IF ( T1 .LT. T1MIN ) GO TO 300
   IF ( T1 .GT. T1MAX ) GO TO 300
   IF ( T2 .LT. T2MIN ) GO TO 300
   IF ( T2 .GT. T2MAX ) GO TO 300
   KOUNT1 = KOUNT1 + 1
300 CONTINUE
C IF ( (T1 .GE. T1MIN) .AND. (T1 .LE. T1MAX) ) GO TO 350
C IF ( (T2 .GE. T2MIN) .AND. (T2 .LE. T2MAX) ) GO TO 350
C KOUNT1 = KOUNT1 + 1
C 350 CONTINUE
KOB5 = KOB5 + 1
1000 CONTINUE
2000 CONTINUE
** END THE PROGRAM **
END
REFERENCES


