AN ABSTRACT OF THE THESIS OF

EVAN PAUL SMOUSE for the degree of DOCTOR OF PHILOSOPHY

in STATISTICS presented on May 4, 1979

Title: A BAYESIAN APPROACH TO NONRESPONSE IN SAMPLE SURVEYS

Redacted for Privacy

Abstract approved: ______________________

H. Dan Brunk

A common problem in postal surveys of human populations arises when some of the sampled people fail to respond to the questionnaire. The problem is that the answers provided by the respondents may not be typical of the answers the nonrespondents would provide if these had been solicited more successfully. This creates a "nonresponse bias." One solution is to subsample the nonrespondents intensively and obtain their answers by some more costly method (e.g., phone calls or personal interviews).

This thesis presents a Bayesian approach to the prediction of a finite population total when there is nonresponse. It is assumed that the population is grouped on the basis of one or more known auxiliary variables, that the characteristic of interest is linearly related to the auxiliary variable(s), and that the probability of response is constant within a group but possibly varies among groups.

A Bayesian Approach to Nonresponse in Sample Surveys

by

Evan Paul Smouse

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

June 1979

APPROVED:

Redacted for Privacy

Professor of Statistics

in charge of major

Redacted for Privacy

Chairman of Department of Statistics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented May 4, 1979

Typed by Clover Redfern for Evan Paul Smouse

ACKNOWLEDGMENT

The author would like to express his deepest appreciation to Dr. H. Dan Brunk, who served as his major professor, for the ample amount of time, energy and encouragement provided.

The author is especially grateful for generous financial support during his graduate studies. This support was in the form of a research assistantship through the O.S.U. Survey Research Center, an N.I.H. Biometry traineeship administered by the Department of Statistics, and a grant of computing funds from the O.S.U. Computing Center.

Final thanks go to the author's wife, Diana, for her loving patience, understanding and moral support. If not for her inspiration, this thesis would have gone unwritten.

TABLE OF CONTENTS

A BAYESIAN APPROACH TO NONRESPONSE IN SAMPLE SURVEYS

I. INTRODUCTION

In this chapter we first introduce the fixed finite population model of Godambe (1955). We then give the Bayesian approach to finite population inference as proposed by Ericson (1969). Finally, we consider a problem of nonresponse in sample surveys and mention previous solutions.

1.1. Fixed Finite Population Model

According to Cassel, Särndal and Wretman (1977), "The essence of survey sampling consists of the selection of a part of a finite collection of units, followed by the making of statements about the entire collection on the basis of the selected part." Different models, or frameworks of inference, have been proposed to aid in the "making of statements" about a finite population based on observing a portion of that population. One of the most basic, a fixed finite population model, is due essentially to Godambe (1955) with later refinements by Godambe and others. This basic model has become a point of departure for many writers; see, for example, review papers by Basu (1971), J.N.K. Rao (1975) and Solomon and Zacks (1970). We list in this section the basic features of this model.

Consider a finite collection (population), U, of N identifiable units where N is known and finite. An identifiable unit is a physical entity which can be located and uniquely labeled. This identifiability is a feature which separates survey sampling from traditional statistical inference. It allows us to choose our own sampling design (described later). Without loss of generality, assume the units in U are labeled so that $U = \{1, 2, \ldots, N\}$.

Assume there is a characteristic of interest associated with each unit and that the value of this characteristic is denoted by y_j for $j \in U$. Let $y = (y_1, y_2, \ldots, y_N)'$. Then the vector y is called a parameter of the finite population. We denote the parameter space by Ω. Frequently, $\Omega = \mathbb{R}^N$, the N-dimensional Euclidean space.

An important element of the fixed finite population model is the sample design, $(\Delta, \phi(\cdot))$, where Δ is the set of all possible samples from U and where $\phi : \Delta \to [0, 1]$ is such that $\sum_{\delta \in \Delta} \phi(\delta) = 1$. The value of $\phi(\delta)$ for $\delta \in \Delta$ assigns the probability of choosing the sample δ. The sample design plays a major role in making and justifying inferences under this model. A particular design is called noninformative if the function $\phi(\cdot)$ is independent of y.

A data point, d_δ, corresponding to a sample $\delta \in \Delta$ is given by $d_\delta = \{(j, y_j) : j \in \delta\}$. The sample space then is $\{d_\delta : \delta \in \Delta,\ y \in \Omega\}$.

The likelihood function for this fixed finite population model was first considered by Godambe (1966). For a given sample $\delta \in \Delta$ and a data point d_δ, the value of the likelihood function at the parameter value $y \in \Omega$ is

$$L_{d_\delta}(y) = \begin{cases} \phi(\delta), & y \in \Omega_{d_\delta} \\ 0, & \text{otherwise} \end{cases} \tag{1.1.1}$$

where

$$\Omega_{d_\delta} = \{y_0 = (y_{0_1}, \ldots, y_{0_N})' \in \Omega : y_{0_j} \text{ equals the observed value } y_j \text{ for all } j \in \delta\}.$$

Thus, the likelihood function is "flat." It is $\phi(\delta)$ over the region Ω_{d_δ} and is zero elsewhere. Some methods of traditional statistical inference (e.g. maximum likelihood estimation) fail to give useful results.

1.2. Superpopulation Models and the Bayesian Approach

In contrast to the fixed finite population model where the parameter vector y is considered to be a fixed point in $\Omega \subset \mathbb{R}^N$, superpopulation models assume that $y = (y_1, y_2, \ldots, y_N)'$ is a realization of a vector random variable $Y = (Y_1, Y_2, \ldots, Y_N)'$. The joint probability distribution of $Y_1, Y_2, \ldots, Y_N$ in $\mathbb{R}^N$ is denoted by ξ.

The superpopulation concept has been the subject of much recent discussion ranging from distrust of model-based inference (e.g. Neyman (1971)) to the opinion that such inference is almost necessary (e.g. Royall (1971)). While ξ may reflect prior subjective belief about Y as in the Bayesian approach, its interpretation can be completely non-Bayesian (Royall (1971)). For example, the finite population U may be thought of as being randomly drawn from some larger universe (the "superpopulation") or ξ might model some random mechanism in the real world (e.g. econometric modeling).

We will now develop the Bayesian approach, following Ericson (1969). Assuming a noninformative design, the resulting inference will be independent of the design $(\Delta, \phi(\cdot))$ used to select the sample.

Let $f(\underset{\sim}{Y})$ denote the prior density of Y with respect to Lebesgue measure in $\mathbb{R}^N$. The function $f(\cdot)$ is arrived at by standard Bayesian methods of assessing prior subjective belief about Y.

Note: The notation, $f(\underset{\sim}{Y})$, will be used throughout this thesis in place of the more common notation f_Y. The collection of symbols, $f(\underset{\sim}{Y})$, is the name of a function; the probability density of the random vector Y. No symbol for an argument appears, even when $f(\underset{\sim}{Y})$

appears on the left hand side of an equation, and say, $(2\pi)^{-1/2}\exp\{-\frac{1}{2}Y^2\}$ appears on the right. Also, whenever it is clear from the context that a particular density is continuous or discrete, explicit mention will not be made.

For any $\delta \in \Delta$, let $\delta^* := U \setminus \delta$ be the set complement of δ in U. That is, δ^* is the set of units in U not included in the sample, δ.

Assume $(\Delta, \phi(\cdot))$ is noninformative. The likelihood for Y is given by (1.1.1). Thus, the posterior density for Y given d_δ is

$$f(\underset{\sim}{Y} \mid d_\delta) \propto \begin{cases} \phi(\delta) f(\underset{\sim}{Y}), & Y \in \Omega_{d_\delta} \\ 0, & \text{otherwise}. \end{cases} \tag{1.2.1}$$

Since $(\Delta, \phi(\cdot))$ is noninformative, the reciprocal of the omitted proportionality constant is

$$\int \phi(\delta) f(\underset{\sim}{Y}) dY_{\delta^*} = \phi(\delta) \int f(\underset{\sim}{Y}) dY_{\delta^*} \tag{1.2.2}$$

where Y_{δ^*} is the subvector of Y corresponding to units in δ^* and where integration is over $\{Y_{\delta^*} : Y_k \in \mathbb{R} \text{ for } k \in \delta^*\}$. Then $\int f(\underset{\sim}{Y}) dY_{\delta^*}$ is simply the marginal density, $f(\underset{\sim}{Y}_\delta)$, of the subvector, Y_δ, which corresponds to units in δ. Thus, (1.2.2) becomes

$$\int \phi(\delta) f(\underset{\sim}{Y}) dY_{\delta *} = \phi(\delta) f(\underset{\sim}{Y}_{\delta}). \tag{1.2.3}$$

Expressions (1.2.1) and (1.2.3) combine to give

$$f(\underset{\sim}{Y} \mid d_\delta) = \begin{cases} f(\underset{\sim}{Y}_{\delta *} \mid Y_\delta), & Y \in \Omega_{d_\delta} \\ 0, & \text{otherwise} \end{cases} \tag{1.2.4}$$

since $f(\underset{\sim}{Y}_{\delta *} \mid Y_\delta) := \dfrac{f(\underset{\sim}{Y})}{f(\underset{\sim}{Y}_\delta)}$. Notice that $\phi(\cdot)$ does not play a role in Bayesian inference about Y, since that inference is based entirely on the posterior distribution of Y given the data.

This lack of dependence on the particular random sampling plan for selecting a sample δ should not be mistaken to imply that the Bayesian approach has no place for random sampling. Some form of random sampling may indeed be employed to help insure the validity of the superpopulation model being used. However, this randomization does not play the direct role in inference that it plays under the fixed finite population model.

Suppose we are interested in predicting the value of the population total, T, where

$$T := \sum_{j=1}^{N} Y_j.$$

As is well known, the Bayesian inference under squared error loss simply involves finding the posterior mean of T given the data and any known parameters; see, for example, Godambe (1969). The Bayes predictor of T under squared error loss is thus $\hat{T}$ where

$$\hat{T} := \sum_{j \in \delta} Y_j + E\left[\sum_{k \in \delta^*} Y_k \middle| d_\delta\right]. \tag{1.2.5}$$

1.3. Nonresponse Problem and Hansen-Hurwitz Sampling

A common method of administering a sample survey of a human population makes use of a printed questionnaire that is mailed to every unit (person) in the sample. Each sampled person is asked to write his or her value of "Y_j" on the questionnaire and return the completed questionnaire to the survey organization.

A growing problem in such surveys is that not all persons sampled actually respond to the survey. Studies by statisticians and non-statisticians alike have attempted to identify methods of eliciting a high degree of response; see, for example, Kanuk and Berenson (1975). However, in most postal surveys there is some portion of the sample that does not respond. The problem is that an unknown bias, a "nonresponse bias," may be involved if we assume that those responding are representative of the combined total of respondents and nonrespondents.

A commonly used technique for obtaining information about nonrespondents employs a Hansen-Hurwitz sampling plan; see Hansen and Hurwitz (1946). The procedure is to mail questionnaires in excess of the number expected to be returned and to follow up by subsampling the nonrespondents by some more costly method (e.g. telephone calls, personal interviews, use of rewards) which provides complete response from the subsampled nonrespondents. Hansen-Hurwitz sampling thus allows for the economy of a postal survey while providing some information about possible nonresponse bias.

1.4. Previous Approaches

The literature contains many ideas for reducing, estimating and adjusting for nonresponse bias in survey sampling. Two recent review papers are those of Kanuk and Berenson (1975) and Armstrong and Overton (1977). Most of these approaches have assumed a fixed finite population model and have used the concepts of that model to justify inferences. Some approaches have utilized a Hansen-Hurwitz sampling plan; others have not.

One of the most basic approaches was proposed by Hansen and Hurwitz (1946) themselves. They suggest estimation of the population total after subsampling nonrespondents by multiplying the estimated population mean by the population size, N. The population mean is estimated by a weighted sum of the sample means for respondents and nonrespondents, the weights being the proportions of respondents and nonrespondents in the original sample. This estimator is "design unbiased"; that is, its expected value over all $\delta \epsilon \Delta$ is equal to the true population total.

Ericson (1967) has proposed a similar estimator in a Bayesian setting using squared error loss. Instead of using the sample means for respondents and nonrespondents, his method uses posterior estimates of the population means for respondents and nonrespondents. These estimates are often close to the sample means, but they adjust for the effect of prior assumptions.

A more recent Bayesian approach by Singh and Sedransk (1975) is similar to the work presented in Chapters II and III in that auxiliary information is used in a regression setting to improve the Bayes estimates of unknown parameters in the model. However, Sedransk and Singh concentrate on estimating regression coefficients and assume that all members of a population respond with the same probability. The present work has the goal of predicting the population total when probabilities of response possibly vary.

Another Bayesian approach can be found in Rubin (1977). This author also makes use of auxiliary information, but he does not make inferences about the entire population. Instead, he provides subjective probability interval estimates for a statistic which would have been calculated and used if there had been no nonresponse.

1.5. Preview of Thesis

This thesis presents a Bayesian approach to the nonresponse problem of Section 1.3. We will assume that useful auxiliary information is available about every unit in the population and that response probabilities possibly vary for different units. Our goal is to predict the population total for some characteristic of interest.

We will develop models and derive Bayes predictors under squared error loss in Chapters II and III. In Chapter II we assume that the precisions (or, equivalently, the variances) of the response variables for respondents and nonrespondents are known. In Chapter III we relax this assumption. Chapter IV contains results for the special case of regression through the origin. A simple simulation example there suggests a convenient approximation to the posterior density function of Chapter II.

II. A BAYESIAN APPROACH: PRECISIONS KNOWN

In this chapter we will formally model the situation described in Chapter I and then analyze the model from a Bayesian point of view.

2.1. Specification of Model and Prior Density Function

Consider a finite population, U, of N identifiable units where N is a known positive integer. An identifiable unit is a physical entity which can be uniquely identified and labeled. Some common examples include the population of all 1979 Oregon annual resident angling permit holders and the population consisting of all United States counties having at least one public health clinic on December 31, 1978. In the first example the physical entities are humans while the units are geographic areas in the second example.

The listing of units and/or unit labels is often referred to as the frame or sampling frame.

Suppose the units in U are grouped into g groups, $1 \leq g \leq N$, so that there are n_i units in group i, $i = 1, 2, \ldots, g$. Without loss of generality we can assume the units in U are labeled as

$$U = \{11, 12, \ldots, 1n_1, 21, 22, \ldots, 2n_2, \ldots, g1, g2, \ldots, gn_g\}$$

where $N = \sum_{i=1}^{g} n_i$ and where the unit labeled ij is the j^{th} unit in the i^{th} group; $i = 1, 2, \ldots, g$; $j = 1, 2, \ldots, n_i$.

The basis for grouping may be arbitrary, but the grouping will most often be due to some auxiliary information about the units in U that is known at the time the frame is constructed. This information will often be demographic (e.g. age, sex) when studying a human population.

Suppose there is a numerical characteristic of interest associated with each unit in U. Let the value of this characteristic be denoted by Y_{ij} for $ij \in U$. Our goal is to make inference about the value of the population total, T, where we define

$$T := \sum_{i=1}^{g} \sum_{j=1}^{n_i} Y_{ij} .$$

In order to gain information about T we conduct a survey using a Hansen-Hurwitz, non-informative sampling plan as described in Sections 1.2 and 1.3. That is, we sample n units from U and obtain the response Y_{ij} for r of these n units. Call these r survey respondents "SR's," and call the $(n-r)$ nonrespondents "NR's." We then subsample m, $0 \leq m \leq n-r$, of these NR's and obtain Y_{ij} for each of the m subsampled NR's. The sampling probabilities are assumed not to depend on the values of the Y_{ij}'s.

The population U can now be partitioned as

$$U = A \cup B \cup C \cup D$$

where A is the set of labels for the r SR's, B is the set of labels for the m subsampled NR's, C is the set of labels for the $n-r-m$ NR's that are not subsampled, and where $D = U \setminus (A \cup B \cup C)$ is the set of labels for the $N-n$ units that are never sampled. The particular units comprising A, B, C and D are unknown before the survey. Define $S := A \cup B \cup C$. Then S is the set of labels for the n units originally sampled. Further, let a_i, b_i, c_i, d_i and s_i denote the numbers of units from group i in the sets A, B, C, D and S, respectively, for $i = 1, 2, \ldots, g$.

We assume that every unit in U would be either an SR or an NR (but not both) if sampled. We model this SR/NR status by associating an indicator random variable, Z_{ij}, with each $ij \in U$ where

$$Z_{ij} = \begin{cases} 0 & \text{if unit } ij \text{ would be an NR}, \quad ij \in U. \\ 1 & \text{if unit } ij \text{ would be an SR} \end{cases}$$

This SR/NR status is important, since we believe that SR's and NR's might differ with respect to the characteristic of interest and its relationship with auxiliary information.

While the values of Y_{ij} and Z_{ij}, $ij \in U$, are <u>unknown</u> before the survey, we assume that the characteristic of interest is linearly related to one or more auxiliary variables that is/are known for every unit in U. We further assume that the value of any one of these auxiliary variables is the same for all units in the i^{th} group, $i = 1, 2, \ldots, g$. Indeed, it will often be the case that this feature determines the grouping of units in U. Let X_i, $i = 1, 2, \ldots, g$, denote the $x \times 1$ column vector whose entries are the values for the i^{th} group of the x auxiliary variables. Note that some entries in X_i may be transformations of other entries as in the case of polynomial regression.

Let h_1 and h_2 be specified constants. (They will be seen to represent known precisions for SR's and NR's respectively.) Define $h := (h_1, h_2)$. Let β_1 and β_2 be $x \times 1$ vectors of parameters, and denote $\beta := (\beta_1, \beta_2)$. Define

$$X := \begin{bmatrix} X_1' \\ \vdots \\ X_1' \\ \vdots \\ X_g' \\ \vdots \\ X_g' \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} X \\ \vdots \\ X \end{matrix}}\right\} n_1 \\ \vphantom{\vdots} \\ \left.\vphantom{\begin{matrix} X \\ \vdots \\ X \end{matrix}}\right\} n_g \end{matrix}$$

where X_i' is the transpose of X_i. Let

$$Z := (Z_{11}, Z_{12}, \ldots, Z_{1n_1}, \ldots, Z_{g1}, \ldots, Z_{gn_g})'.$$

Given X, h, β and Z, assume that

$Y_{11}, Y_{12}, \ldots, Y_{1n_1}, \ldots, Y_{g1}, \ldots, Y_{gn_g}$ are independent random variables such that

$$Y_{ij} \sim N\left(Z_{ij}X_i'\beta_1 + (1-Z_{ij})X_i'\beta_2, \left(\frac{1}{h_1}\right)^{Z_{ij}}\left(\frac{1}{h_2}\right)^{1-Z_{ij}}\right)$$

for $i = 1, 2, \ldots, g$ and $j = 1, 2, \ldots, n_i$. Thus, if ij is an SR, Y_{ij} has mean $X_i'\beta_1$ and precision h_1; if ij is an NR, Y_{ij} has mean $X_i'\beta_2$ and precision h_2. Note that $Y \sim N(\mu, V)$ means that the random variable Y has a normal probability distribution in $\mathbb{R}$ with mean μ and precision V^{-1}.

Assume that Z, β_1 and β_2 are independent random vectors. Given the vector $\pi := (\pi_1, \pi_2, \ldots, \pi_g)'$, assume that $Z_{11}, Z_{12}, \ldots, Z_{1n_1}, \ldots, Z_{g1}, Z_{g2}, \ldots, Z_{gn_g}$ are independent random variables such that

$$Z_{ij} \sim \text{Bernoulli } (\pi_i), \quad ij \in U.$$

That is,

$$E(Z_{ij}) = \Pr\{Z_{ij} = 1\} = \Pr\{\text{unit } ij \text{ is an SR}\} = \pi_i$$

is the same for all units in group i; $i = 1, 2, \ldots, g$. Note that $0 \leq \pi_i \leq 1$ for $i = 1, 2, \ldots, g$.

Let λ_1 and λ_2 be specified $x \times 1$ vectors of hyperparameters. Let τ_1 and τ_2 be $x \times x$ positive definite matrices. Given λ_1, λ_2, h_1, h_2, τ_1 and τ_2, assume that β_1 and β_2 are independent random vectors such that

$$\beta_1 \sim N(\lambda_1, h_1^{-1}\tau_1^{-1})$$

and

$$\beta_2 \sim N(\lambda_2, h_2^{-1}\tau_2^{-1}).$$

Note that we are assuming β_1 and β_2 are independent of π.

Finally, we complete the specification of the model by assuming that, given $p = (p_1, p_2, \ldots, p_g)'$ and $q = (q_1, q_2, \ldots, q_g)'$, $\pi_1, \pi_2, \ldots, \pi_g$ are independent random variables such that

$$\pi_i \sim \text{Beta}\,(p_i, q_i), \quad i = 1, 2, \ldots, g.$$

That is,

$$f(\pi_i \mid p_i, q_i) \propto \pi_i^{p_i - 1} (1-\pi_i)^{q_i - 1}, \quad 0 \leq \pi_i \leq 1,$$

for $i = 1, 2, \ldots, g$ with $\pi_1, \ldots, \pi_g$ independent. Note that

$$E(\pi_i \mid p_i, q_i) = \frac{p_i}{p_i + q_i}, \quad i = 1, 2, \ldots, g.$$

Let $I := \{X, h_1, h_2, \lambda_1, \lambda_2, \tau_1, \tau_2, p, q\}$. Let $Z_N := \text{diag}(Z_{ij})$

be an $N \times N$ diagonal matrix having ij^{th} diagonal element Z_{ij}. Let I_N be an $N \times N$ identity matrix.

We are assuming that, conditional on Z, β, π, and I,

$$Y \sim N(\mu_N, V_N)$$

where

$$\mu_N := (Z_N \mid (I_N - Z_N)) \left[\begin{array}{c|c} X & 0 \\ \hline 0 & X \end{array}\right] \beta$$

$$= (Z_N X \mid (I_N - Z_N)X) \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$$

$$= Z_N X \beta_1 + (I_N - Z_N) X \beta_2$$

and

$$V_N := \mathrm{diag}(h_1^{-Z_{ij}} h_2^{-(1-Z_{ij})}) .$$

We have the following joint prior density function for Y, Z, β and π:

$$f(\underset{\sim}{Y}, \underset{\sim}{Z}, \underset{\sim}{\beta}, \underset{\sim}{\pi} \mid I) = f(\underset{\sim}{Y} \mid Z, \beta, \pi, I) f(\underset{\sim}{Z} \mid \beta, \pi, I) f(\underset{\sim}{\beta} \mid \pi, I) f(\underset{\sim}{\pi} \mid I)$$

$$\propto (|V_N|)^{-1/2} \exp\{-\tfrac{1}{2} (Y - \mu_N)' V_N^{-1} (Y - \mu_N)\}$$

$$\times \prod_{i=1}^{g} \prod_{j=1}^{n_i} [\pi_i^{Z_{ij}} (1 - \pi_i)^{1 - Z_{ij}}]$$

$$\times (|h_1 \tau_1|)^{1/2} \exp\{-\tfrac{1}{2} (\beta_1 - \lambda_1)'(h_1 \tau_1)(\beta_1 - \lambda_1)\} \times$$

$$\times (|h_2\tau_2|)^{1/2}\exp\{-\tfrac{1}{2}(\beta_2-\lambda_2)'(h_2\tau_2)(\beta_2-\lambda_2)\}$$

$$\times \prod_{i=1}^{g} [\pi_i^{p_i-1}(1-\pi_i)^{q_i-1}] \qquad (2.1.1)$$

for $Y \in \mathbb{R}^N$, $\beta_1 \in \mathbb{R}^x$, $\beta_2 \in \mathbb{R}^x$, $Z \in [0,1] \times \ldots \times [0,1]$ and for $0 \leq \pi_1, \pi_2, \ldots, \pi_g \leq 1$.

2.2. Posterior Density Function

Let Y_A, Y_B, Y_C and Y_D be those subvectors of Y corresponding to units in A, B, C and D respectively. Similarly, Z_A, Z_B, Z_C and Z_D are subvectors of Z while X_A, X_B, X_C and X_D are submatrices of X. Further, let

$$Y_S := \begin{bmatrix} Y_A \\ Y_B \\ Y_C \end{bmatrix} \quad \text{and let} \quad Z_S := \begin{bmatrix} Z_A \\ Z_B \\ Z_C \end{bmatrix}.$$

Define $R := \sum_{ij \in C \cup D} Y_{ij}$ where, for any subset U^* of U, $\sum_{ij \in U^*}$ indicates the sum over all $i \in \{1, 2, \ldots, g\}$ and all j such that $ij \in U^*$. Thus R is the sum of all the values of Y_{ij} which are not observed in the survey.

We observe Y_A, Y_B and Z_S in the survey. Z_S specifies A, and B may depend on A. However, for any noninformative design, B is ancillary, given Z_S, since the probability of observing a set B, given Z_S, is independent of Y, Z_C, Z_D, β and π. In the remainder of this thesis, we assume that A and B are known after the survey, but we do not mention this explicitly in the notation. Define $II := I \cup \{Y_A, Y_B, Z_S\}$. Then II is the set of variable and parameter values known after the survey.

Our goal is to predict T, the population total for the characteristic of interest. Section 1.2 shows that the Bayes predictor of T under squared error loss is $\hat{T}$ where

$$\hat{T} := \sum_{ij \in A \cup B} Y_{ij} + \hat{R}$$

for $\hat{R} := E[R \mid II]$.

In this section we will find the posterior density $f(\underset{\sim}{R} \mid II)$. We will first find $f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} \mid II)$. This density will be transformed to $f(\underset{\sim}{R}, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} \mid II)$. Then Z_D, β and π will be treated as nuisance parameters, and the marginal density $f(\underset{\sim}{R} \mid II)$ will be found.

The following lemma will be used repeatedly. It is widely known. A concise proof is given by Lindley and Smith (1972).

<u>Lemma 2.2.1.</u> Let A_1 and A_2 be known matrices of dimensions $n^* \times p_1^*$ and $p_1^* \times p_2^*$ respectively. Let C_1 and C_2 be known positive definite matrices of dimensions $n^* \times n^*$ and $p_1^* \times p_1^*$ respectively. Suppose, given a $p_1^* \times 1$ vector of parameters, θ_1, that

$$Y \sim N(A_1 \theta_1, C_1)$$

and that, given θ_2, a $p_2^* \times 1$ vector of hyperparameters,

$$\theta_1 \sim N(A_2 \theta_2, C_2) .$$

Then,

(1) the marginal distribution of Y given A_1, A_2, C_1, C_2, θ_2 is $N(A_1 A_2 \theta_2, C_1 + A_1 C_2 A_1')$, and

(2) the distribution of θ_1 given $Y = y, A_1, A_2, C_1, C_2, \theta_2$ is $N(Bb, B)$ where $B^{-1} = A_1' C_1^{-1} A_1 + C_2^{-1}$ and $b = A_1' C_1^{-1} y + C_2^{-1} A_2 \theta_2$.

Note that A_1 need not have full column rank for B to exist, since C_2^{-1} is positive definite while $A_1' C_1^{-1} A_1$ is nonnegative definite. $\square$

Clearly,

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} | II) = f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D | Z_D, \beta, \pi, II) f(\underset{\sim}{Z}_D | \beta, \pi, II) \times f(\underset{\sim}{\beta} | \pi, II) f(\underset{\sim}{\pi} | II) \qquad (2.2.2)$$

Lemma 2.2.3.

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, | Z_D, \beta, \pi, II) = f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D | Z_C, Z_D, \beta, \pi, I).$$

Proof.

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D | Z_D, \beta, \pi, II)$$

$$= f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D | Z_C, Z_D, \beta, \pi, I, Y_A, Y_B, Z_A, Z_B)$$

$$= \frac{f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{Y}_A, \underset{\sim}{Y}_B | Z_C, Z_D, \beta, \pi, I, Z_A, Z_B)}{f(\underset{\sim}{Y}_A, \underset{\sim}{Y}_B | Z_C, Z_D, \beta, \pi, I, Z_A, Z_B)}$$

$$= \frac{f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D | Z_C, Z_D, \beta, \pi, I, Z_A, Z_B) f(\underset{\sim}{Y}_A, \underset{\sim}{Y}_B | Z_C, Z_D, \beta, \pi, I, Z_A, Z_B)}{f(\underset{\sim}{Y}_A, \underset{\sim}{Y}_B | Z_C, Z_D, \beta, \pi, I, Z_A, Z_B)}$$

by the conditional independence of $\binom{Y_C}{Y_D}$ and $\binom{Y_A}{Y_B}$. That is,

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D | Z_D, \beta, \pi, II)$$

$$= f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D | Z_C, Z_D, \beta, \pi, I, Z_A, Z_B)$$

$$= \frac{f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{Z}_A, \underset{\sim}{Z}_B | Z_C, Z_D, \beta, \pi, I)}{f(\underset{\sim}{Z}_A, \underset{\sim}{Z}_B | Z_C, Z_D, \beta, \pi, I)}$$

$$= \frac{f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D | Z_C, Z_D, \beta, \pi, I) f(\underset{\sim}{Z}_A, \underset{\sim}{Z}_B | Z_C, Z_D, \beta, \pi, I)}{f(\underset{\sim}{Z}_A, \underset{\sim}{Z}_B | Z_C, Z_D, \beta, \pi, I)}$$

by the conditional independence of $\binom{Y_C}{Y_D}$ and $\binom{Z_A}{Z_B}$. $\square$

Lemma 2.2.4. $f(\underset{\sim}{Z}_D \mid \beta, \pi, II) = f(\underset{\sim}{Z}_D \mid \pi, I)$.

Proof.

$$f(\underset{\sim}{Z}_D \mid \beta, \pi, II) = f(\underset{\sim}{Z}_D \mid \beta, \pi, I, Y_A, Y_B, Z_S)$$

$$= f(\underset{\sim}{Z}_D \mid \pi, I)$$

since, conditional on π and I, Z_D is independent of Z_S, Y_A, Y_B and β. $\square$

Lemma 2.2.5. $f(\underset{\sim}{\beta} \mid \pi, II) = f(\underset{\sim}{\beta} \mid II)$.

Proof.

$$f(\underset{\sim}{\beta} \mid \pi, II) = f(\underset{\sim}{\beta} \mid \pi, Y_A, Y_B, Z_S, I)$$

$$= \frac{f(\underset{\sim}{\beta}, \underset{\sim}{Y}_A, \underset{\sim}{Y}_B \mid Z_S, I, \pi)}{f(\underset{\sim}{Y}_A, \underset{\sim}{Y}_B \mid \pi, Z_S, I)}$$

$$= \frac{f(\underset{\sim}{\beta}, \underset{\sim}{Y}_A, \underset{\sim}{Y}_B \mid Z_S, I)}{f(\underset{\sim}{Y}_A, \underset{\sim}{Y}_B \mid Z_S, I)}$$

since, conditional on Z_S and I, the vectors β, Y_A and Y_B are independent of π. $\square$

Applying the results of Lemmas 2.2.3, 2.2.4 and 2.2.5 to the expression (2.2.2) shows that

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} \mid II) = f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D \mid Z_C, Z_D, \beta, \pi, I) f(\underset{\sim}{Z}_D \mid \pi, I)$$

$$\times\ f(\underset{\sim}{\beta} \mid II) f(\underset{\sim}{\pi} \mid II). \qquad (2.2.6)$$

Let $Z_D := \mathrm{diag}_D(Z_{ij})$ be an $(N-n) \times (N-n)$ diagonal matrix having diagonal elements Z_{ij} for $ij \in D$. Then, conditional on Z_C, Z_D, β, π and I,

$$\begin{bmatrix} Y_C \\ Y_D \end{bmatrix} \sim N(\mu_{CD}, V_{CD})$$

where

$$\mu_{CD} := \left[\begin{array}{c|c} 0 & X_C \\ \hline Z_D X_D & X_D - Z_D X_D \end{array}\right] \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$$

and where

$$V_{CD} := \left[\begin{array}{c|c} h_2^{-1} I_C & 0 \\ \hline 0 & V_D \end{array}\right]$$

for the $(n-m-r) \times (n-m-r)$ identity matrix I_C and for

$$V_D := \mathrm{diag}_D(h_1^{-Z_{ij}} h_2^{-(1-Z_{ij})})$$

an $(N-n) \times (N-n)$ diagonal matrix having diagonal elements

$h_1^{-Z_{ij}} h_2^{-(1-Z_{ij})}$ for $ij \in D$. Thus,

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D \mid Z_C, Z_D, \beta, \pi, I) \tag{2.2.7}$$

$$\propto (|V_{CD}|)^{-1/2} \exp\left\{ -\frac{1}{2} [\binom{Y_C}{Y_D} - \mu_{CD}]' V_{CD}^{-1} [\binom{Y_C}{Y_D} - \mu_{CD}] \right\} .$$

Clearly,

$$f(\underset{\sim}{Z}_D \mid \pi, I) \propto \prod_{ij \in D} [\pi_i^{Z_{ij}} (1-\pi_i)^{1-Z_{ij}}] . \tag{2.2.8}$$

We have yet to find $f(\underset{\sim}{\beta} \mid II)$ and $f(\underset{\sim}{\pi} \mid II)$. Now, $f(\underset{\sim}{\beta} \mid II) = f(\underset{\sim}{\beta} \mid Y_A, Y_B, Z_S, I)$. Assume Z_S and I are known. (After the survey, we do know that Z_A is a vector of 1's while Z_B and Z_C are vectors of 0's.) Given β,

$$\begin{bmatrix} Y_A \\ Y_B \end{bmatrix} \sim N(A_1\beta, C_1)$$

where

$$A_1 = \left[\begin{array}{c|c} X_A & 0 \\ \hline 0 & X_B \end{array}\right] \quad \text{and where} \quad C_1 = \left[\begin{array}{c|c} h_1^{-1} I_A & 0 \\ \hline 0 & h_2^{-1} I_B \end{array}\right]$$

for the $r \times r$ and $m \times m$ identity matrices I_A and I_B.

Further,

$$\beta \sim N(A_2 \binom{\lambda_1}{\lambda_2}, C_2)$$

where A_2 is a $(2x) \times (2x)$ identity matrix and where

$$C_2 = \left[\begin{array}{c|c} h_1^{-1}\tau_1^{-1} & 0 \\ \hline 0 & h_2^{-1}\tau_2^{-1} \end{array}\right] .$$

By Lemma 2.2.1, $f(\underset{\sim}{\beta} | \mathrm{II})$ is a normal density with mean $B_\beta b_\beta$ and variance B_β where $B_\beta^{-1} := A_1' C_1^{-1} A_1 + C_2^{-1}$ and

$$b_\beta := A_1' C_1^{-1} \binom{Y_A}{Y_B} + C_2^{-1} \binom{\lambda_1}{\lambda_2} .$$

Now,

$$C_2^{-1} = \left[\begin{array}{c|c} h_1\tau_1 & 0 \\ \hline 0 & h_2\tau_2 \end{array}\right] ,$$

$$A_1' C_1^{-1} A_1 = \left[\begin{array}{c|c} X_A' & 0 \\ \hline 0 & X_B' \end{array}\right] \left[\begin{array}{c|c} h_1 I_A & 0 \\ \hline 0 & h_2 I_B \end{array}\right] \left[\begin{array}{c|c} X_A & 0 \\ \hline 0 & X_B \end{array}\right]$$

$$= \left[\begin{array}{c|c} h_1 X_A' X_A & 0 \\ \hline 0 & h_2 X_B' X_B \end{array}\right] ,$$

$$A_1' C_1^{-1} \binom{Y_A}{Y_B} = \begin{bmatrix} h_1 X_A' Y_A \\ h_2 X_B' Y_B \end{bmatrix} ,$$

and

$$C_2^{-1}\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix} = \begin{bmatrix}h_1\tau_1\lambda_1\\ h_2\tau_2\lambda_2\end{bmatrix} .$$

Thus,

$$b_\beta = \begin{bmatrix}h_1(X_A'Y_A+\tau_1\lambda_1)\\ h_2(X_B'Y_B+\tau_2\lambda_2)\end{bmatrix}$$

and

$$B_\beta = \left[\begin{array}{c|c} h_1^{-1}(X_A'X_A+\tau_1)^{-1} & 0 \\ \hline 0 & h_2^{-1}(X_B'X_B+\tau_2)^{-1} \end{array}\right].$$

The posterior density of β given II is now seen to be

$$f(\underset{\sim}{\beta}|\text{II}) \propto (|B_\beta|)^{-1/2}\exp\{-\tfrac{1}{2}(\beta-B_\beta b_\beta)'B_\beta^{-1}(\beta-B_\beta b_\beta)\} .$$

The Bayes estimate of β under squared error loss is $\hat{\beta}$ where

$$\hat{\beta} := \begin{pmatrix}\hat{\beta}_1\\ \hat{\beta}_2\end{pmatrix} := B_\beta b_\beta = \begin{bmatrix}(X_A'X_A+\tau_1)^{-1}(X_A'Y_A+\tau_1\lambda_1)\\ (X_B'X_B+\tau_2)^{-1}(X_B'Y_B+\tau_2\lambda_2)\end{bmatrix}.$$

We thus will write

$$f(\underset{\sim}{\beta}|\text{II}) \propto (|B_\beta|)^{-1/2}\exp\{-\tfrac{1}{2}(\beta-\hat{\beta})'B_\beta^{-1}(\beta-\hat{\beta})\}. \tag{2.2.9}$$

We now turn to the task of finding $f(\underset{\sim}{\pi}|\text{II})$. Clearly,

$$f(\underset{\sim}{\pi}|II) = \frac{f(\underset{\sim}{\pi}, \underset{\sim}{Y}_A, \underset{\sim}{Y}_B, \underset{\sim}{Z}_S|I)}{\int_0^1 \cdots \int_0^1 f(\underset{\sim}{\pi}, \underset{\sim}{Y}_A, \underset{\sim}{Y}_B, \underset{\sim}{Z}_S|I)d\pi_1 d\pi_2 \cdots d\pi_g} .$$

But $f(\underset{\sim}{\pi}, \underset{\sim}{Y}_A, \underset{\sim}{Y}_B, \underset{\sim}{Z}_S|I) = f(\underset{\sim}{Y}_A, \underset{\sim}{Y}_B|\pi, Z_S, I)f(\underset{\sim}{Z}_S|\pi, I)f(\underset{\sim}{\pi}|I)$ where $f(\underset{\sim}{Y}_A, \underset{\sim}{Y}_B|\pi, Z_S, I)$ is independent of π (conditional on Z_S). That is, $f(\underset{\sim}{Y}_A, \underset{\sim}{Y}_B|\pi, Z_S, I)$ is the same function for every value of the vector π. This means that

$$f(\underset{\sim}{\pi}|II) = \frac{f(\underset{\sim}{\pi}, \underset{\sim}{Z}_S|I)}{\int_0^1 \cdots \int_0^1 f(\underset{\sim}{\pi}, \underset{\sim}{Z}_S|I)d\pi_1 d\pi_2 \cdots d\pi_g} . \tag{2.2.10}$$

Now

$$f(\underset{\sim}{\pi}, \underset{\sim}{Z}_S|I) = f(\underset{\sim}{Z}_S|\pi, I)f(\underset{\sim}{\pi}|I)$$

$$\propto \prod_{ij \in S} [\pi_i^{Z_{ij}}(1-\pi_i)^{1-Z_{ij}}] \prod_{i=1}^{g} [\pi_i^{p_i - 1} (1-\pi_i)^{q_i - 1}] .$$

Since Z_A is a vector of 1's while $\binom{Z_B}{Z_C}$ is a vector of 0's,

$$\prod_{ij \in S} [\pi_i^{Z_{ij}}(1-\pi_i)^{1-Z_{ij}}] = \prod_{i=1}^{g} [\pi_i^{a_i}(1-\pi_i)^{s_i - a_i}] .$$

Thus,

$$f(\underset{\sim}{\pi}, \underset{\sim}{Z}_S|I) \propto \prod_{i=1}^{g} [\pi_i^{p_i' - 1} (1-\pi_i)^{q_i' - 1}]$$

where $p_i' := p_i + a_i$ and $q_i' := q_i + s_i - a_i$. For Z_S known, the integral in the denominator of (2.2.10) is a known constant. We can now write

$$f(\underset{\sim}{\pi}\,|\,\mathrm{II}) \propto \prod_{i=1}^{g} [\pi_i^{p_i'-1}(1-\pi_i)^{q_i'-1}] .$$

The constant of proportionality is easily seen to give

$$f(\underset{\sim}{\pi}\,|\,\mathrm{II}) = \prod_{i=1}^{g} \left\{ \frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \pi_i^{p_i'-1}(1-\pi_i)^{q_i'-1} \right\} \tag{2.2.11}$$

where

$$\Gamma(u) := \int_0^\infty t^{u-1}e^{-t}dt \quad \text{for} \quad u > 0.$$

We now apply the results in (2.2.7), (2.2.8), (2.2.9) and (2.2.11) to (2.2.6) to obtain

$$\begin{aligned} &f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi}\,|\,\mathrm{II}) \\ &\propto (|V_{CD}|)^{-1/2} \exp\left\{ -\frac{1}{2} [\binom{Y_C}{Y_D} - \mu_{CD}]' V_{CD}^{-1} [\binom{Y_C}{Y_D} - \mu_{CD}] \right\} \\ &\quad \times \prod_{ij \in D} [\pi_i^{Z_{ij}} (1-\pi_i)^{1-Z_{ij}}] (|B_\beta|)^{-1/2} \exp\{-\frac{1}{2}(\beta-\hat{\beta})' B_\beta^{-1} (\beta-\hat{\beta})\} \\ &\quad \times \prod_{i=1}^{g} \left\{ \frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \pi_i^{p_i'-1}(1-\pi_i)^{q_i'-1} \right\} . \end{aligned} \tag{2.2.12}$$

The next step is to transform $f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} | II)$ to $f(\underset{\sim}{R}, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} | II)$. Clearly,

$$f(\underset{\sim}{R}, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} | II) = f(\underset{\sim}{R} | Z_D, \beta, \pi, II) f(\underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} | II)$$

where

$$f(\underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} | II) = f(\underset{\sim}{Z}_D | \beta, \pi, II) f(\underset{\sim}{\beta} | \pi, II) f(\underset{\sim}{\pi} | II).$$

(Compare (2.2.2).)

Now $f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D | Z_D, \beta, \pi, II)$ is a multivariate normal density. (See (2.2.7), and note that, conditional on I, Z_D, β and π, $\binom{Y_C}{Y_D}$ is independent of Y_A, Y_B and Z_S.) It follows that, conditional on Z_D, β, π and II,

$$\binom{Y_C}{Y_D} \sim N(\mu_{CD}, V_{CD}).$$

But

$$R := 1'\binom{Y_C}{Y_D}$$

where 1 is an $(N-r-m) \times (N-r-m)$ column vector of 1's. It follows from well known properties of sums of normally distributed random variables that, conditional on Z_D, β, π and II, $R \sim N(\mu_R, V_R)$ where

$$\mu_R := 1'\mu_{CD} = 1'\left[\begin{array}{c|c} 0 & X_C \\ \hline Z_D X_D & X_D - Z_D X_D \end{array}\right]\binom{\beta_1}{\beta_2}$$

and

$$V_R := 1'V_{CD}1 = 1'\left[\begin{array}{c|c} h_2^{-1} I_C & 0 \\ \hline 0 & V_D \end{array}\right]1.$$

Thus,

$$f(\underset{\sim}{R} \,|\, Z_D, \beta, \pi, II) \propto V_R^{-1/2} \exp\{-\frac{1}{2V_R}(R-\mu_R)^2\}. \tag{2.2.13}$$

This means that

$$f(\underset{\sim}{R}, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} \,|\, II) \propto V_R^{-1/2} \exp\{-\frac{1}{2V_R}(R-\mu_R)^2\}$$

$$\times \; f(\underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} \,|\, II). \tag{2.2.14}$$

The results of Lemmas 2.2.3, 2.2.4 and 2.2.5 show that

$$f(\underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} \,|\, II) = f(\underset{\sim}{Z}_D \,|\, \pi, I) f(\underset{\sim}{\beta} \,|\, II) f(\underset{\sim}{\pi} \,|\, II) . \tag{2.2.15}$$

Thus, by (2.2.8), (2.2.9), (2.2.11) and (2.2.14),

$$f(\underset{\sim}{R}, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} \,|\, II) \propto V_R^{-1/2} \exp\{-\frac{1}{2V_R}(R-\mu_R)^2\}$$

$$\times \prod_{ij \in D} [\pi_i^{Z_{ij}} (1-\pi_i)^{1-Z_{ij}}]$$

$$\times \; (|B_\beta|)^{-1/2} \exp\{-\frac{1}{2}(\beta-\hat{\beta})' B_\beta^{-1} (\beta-\hat{\beta})\} \; \times$$

$$\times \prod_{i=1}^{g} \left\{ \frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \pi_i^{p_i'-1} (1-\pi_i)^{q_i'-1} \right\}. \quad (2.2.16)$$

Before "integrating out" the nuisance parameters Z_D, β and π, we will make another transformation that will simplify the presentation. For $i = 1, 2, \ldots, g$ define $S_i := \sum_{j: ij \in D} Z_{ij}$. That is, S_i is the number of SR's from group i among that portion (D) of the population that is never sampled.

We will now transform $f(\underset{\sim}{R}, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} | II)$ to $f(\underset{\sim}{R}, \underset{\sim}{S}_1, \ldots, \underset{\sim}{S}_g, \underset{\sim}{\beta}, \underset{\sim}{\pi} | II)$. Note that

$$f(\underset{\sim}{R}, \underset{\sim}{S}_1, \ldots, \underset{\sim}{S}_g, \underset{\sim}{\beta}, \underset{\sim}{\pi} | II) = f(\underset{\sim}{R} | S_1, \ldots, S_g, \beta, \pi, II) \quad (2.2.17)$$

$$\times f(\underset{\sim}{S}_1, \ldots, \underset{\sim}{S}_g | \beta, \pi, II) f(\underset{\sim}{\beta} | II) f(\underset{\sim}{\pi} | II)$$

using Lemma 2.2.5.

By Lemma 2.2.4, conditionally on π and I, Z_D is independent of β, Y_A, Y_B and Z_S. But $S_1, \ldots, S_g$ are functions of Z_D. Thus,

$$f(\underset{\sim}{S}_1, \ldots, \underset{\sim}{S}_g | \beta, \pi, II) = f(\underset{\sim}{S}_1, \ldots, \underset{\sim}{S}_g | \pi, I). \quad (2.2.18)$$

Suppose π and I are known. Then the elements of Z_D are independent random variables. This means that $S_1, \ldots, S_g$ are independent. So, for $i = 1, 2, \ldots, g$, S_i is the sum of d_i

independent random variables each having a Bernoulli (π_i) distribution. Thus S_i has a Binomial (d_i, π_i) distribution. The joint density of $S_1, \ldots, S_g$, given π and I, is the product

$$f(\underset{\sim}{S}_1, \ldots, \underset{\sim}{S}_g \mid \pi, I) = \prod_{i=1}^{g} [\binom{d_i}{S_i} \pi_i^{S_i} (1-\pi_i)^{d_i - S_i}] . \qquad (2.2.19)$$

Now

$$f(\underset{\sim}{R} \mid Z_D, \beta, \pi, II) \propto V_R^{-1/2} \exp\{- \frac{1}{2V_R} (R - \mu_R)^2\}$$

where

$$\mu_R = 1' \left[\begin{array}{c|c} 0 & X_C \\ \hline Z_D X_D & X_D - Z_D X_D \end{array}\right] \binom{\beta_1}{\beta_2}$$

$$= 1' \left[\begin{array}{c} X_C \beta_2 \\ \hline Z_D X_D \beta_1 + (X_D - Z_D X_D) \beta_2 \end{array}\right]$$

$$= \sum_{i=1}^{g} c_i X_i' \beta_2 + \sum_{ij \in D} [Z_{ij} X_i' \beta_1 + (1 - Z_{ij}) X_i' \beta_2]$$

$$= \sum_{i=1}^{g} c_i X_i' \beta_2 + \sum_{i=1}^{g} [S_i X_i' \beta_1 + (d_i - S_i) X_i' \beta_2]$$

and where

$$V_R = 1' \left[\begin{array}{c|c} h_2^{-1} I_C & 0 \\ \hline 0 & V_D \end{array}\right] 1$$

$$= (n-m-r)h_2^{-1} + \sum_{ij \in D} [(\frac{1}{h_1})^{Z_{ij}}(\frac{1}{h_2})^{1-Z_{ij}}]$$

$$= (n-m-r)h_2^{-1} + \sum_{ij \in D} [Z_{ij}(\frac{1}{h_1})+(1-Z_{ij})(\frac{1}{h_2})]$$

$$= (n-m-r)h_2^{-1} + \sum_{i=1}^{g} [S_i(\frac{1}{h_1})+(d_i-S_i)(\frac{1}{h_2})] .$$

Thus, $f(\underset{\sim}{R}|Z_D, \beta, \pi, II)$ depends on Z_D only through $S_1, \ldots, S_g$ and

$$f(\underset{\sim}{R}|Z_D, \beta, \pi, II) = f(\underset{\sim}{R}|S_1, \ldots, S_g, \beta, \pi, II). \qquad (2.2.20)$$

If we now combine (2.2.17) with (2.2.20), (2.2.13), (2.2.18), (2.2.19), (2.2.9) and (2.2.11) we see that

$$f(\underset{\sim}{R}, \underset{\sim}{S}_1, \ldots \underset{\sim}{S}_g, \underset{\sim}{\beta}, \underset{\sim}{\pi}|II)$$

$$\propto V_R^{-1/2} \exp\{-\frac{1}{2V_R}(R-\mu_R)^2\} \prod_{i=1}^{g} [\binom{d_i}{S_i}\pi_i^{S_i}(1-\pi_i)^{d_i-S_i}]$$

$$\times (|B_\beta|)^{-1/2} \exp\{-\frac{1}{2}(\beta-\hat{\beta})'B_\beta^{-1}(\beta-\hat{\beta})\}$$

$$\times \prod_{i=1}^{g} \left\{\frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \pi_i^{p_i'-1}(1-\pi_i)^{q_i'-1}\right\} . \qquad (2.2.21)$$

We are now ready to treat $S_1, \ldots, S_g$, β and π as nuisance parameters in (2.2.21). Suppose $S_1, \ldots, S_g$, π and Π are known. Given β, $R \sim N(\mu_R, V_R)$ while, given b_β, $\beta \sim N(B_\beta b_\beta, B_\beta)$. Note that

$$\mu_R = \sum_{i=1}^{g} c_i X_i' \beta_2 + \sum_{i=1}^{g} [S_i X_i' \beta_1 + (d_i - S_i) X_i' \beta_2]$$

$$= \sum_{i=1}^{g} [S_i X_i' \beta_1 + (c_i + d_i - S_i) X_i' \beta_2]$$

$$= 1' \left[\begin{array}{c|c} S_* & 0 \\ \hline 0 & S^* \end{array}\right] \left[\begin{array}{c|c} X_* & 0 \\ \hline 0 & X_* \end{array}\right] \beta \qquad (2.2.22)$$

where

$$S_* = \begin{bmatrix} S_1 & & & 0 \\ & S_2 & & \\ & & \ddots & \\ 0 & & & S_g \end{bmatrix},$$

$$S^* = \begin{bmatrix} (c_1 + d_1 - S_1) & & 0 \\ & \ddots & \\ 0 & & (c_g + d_g - S_g) \end{bmatrix},$$

and

$$X_* = \begin{bmatrix} X_1' \\ X_2' \\ \vdots \\ X_g' \end{bmatrix} .$$

Lemma 2.2.1 shows that, conditional on $S_1, \ldots, S_g$, π and II, the marginal density of R is normal with mean

$$\mu_R^* = 1' \left[\begin{array}{c|c} S_* & 0 \\ \hline 0 & S^* \end{array}\right] \left[\begin{array}{c|c} X_* & 0 \\ \hline 0 & X_* \end{array}\right] \hat{\beta}$$

$$= \sum_{i=1}^{g} [S_i X_i' \hat{\beta}_1 + (c_i + d_i - S_i) X_i' \hat{\beta}_2] \qquad (2.2.23)$$

and with variance

$$V_R^* = V_R + 1' \left[\begin{array}{c|c} S_* & 0 \\ \hline 0 & S^* \end{array}\right] \left[\begin{array}{c|c} X_* & 0 \\ \hline 0 & X_* \end{array}\right] B_\beta \left[\begin{array}{c|c} X_*' & 0 \\ \hline 0 & X_*' \end{array}\right] \left[\begin{array}{c|c} S_*' & 0 \\ \hline 0 & (S^*)' \end{array}\right] 1$$

$$= V_R + 1' \left[\begin{array}{c|c} S_* X_* & 0 \\ \hline 0 & S^* X_* \end{array}\right] \left[\begin{array}{c|c} h_1^{-1}(X_A'X_A+\tau_1)^{-1} & 0 \\ \hline 0 & h_2^{-1}(X_B'X_B+\tau_2)^{-1} \end{array}\right]$$

$$\times \left[\begin{array}{c|c} X_*'S_*' & 0 \\ \hline 0 & X_*'(S^*)' \end{array}\right] 1$$

$$= V_R + 1' \left[\begin{array}{c|c} h_1^{-1} S_* X_* (X_A'X_A+\tau_1)^{-1} X_*'S_*' & 0 \\ \hline 0 & h_2^{-1} S^* X_* (X_B'X_B+\tau_2)^{-1} X_*'(S^*)' \end{array}\right] 1 =$$

$$= (n-m-r)h_2^{-1} + \sum_{i=1}^{g} [h_1^{-1}S_i + h_2^{-1}(d_i - S_i)]$$

$$+ h_1^{-1} \sum_{i=1}^{g} S_i^2 X_i'(X_A'X_A+\tau_1)^{-1}X_i + h_1^{-1} \sum_{i=1}^{g} \sum_{\substack{k=1 \\ k\neq i}}^{g} S_i S_k X_i'(X_A'X_A+\tau_1)^{-1}X_k$$

$$+ h_2^{-1} \sum_{i=1}^{g} (c_i+d_i-S_i)^2 X_i'(X_B'X_B+\tau_2)^{-1}X_i$$

$$+ h_2^{-1} \sum_{i=1}^{g} \sum_{\substack{k=1 \\ i\neq i}}^{g} (c_i+d_i-S_i)(c_k+d_k-S_k)X_i'(X_B'X_B+\tau_2)^{-1}X_k \,. \qquad (2.2.24)$$

We now have

$$f(\underset{\sim}{R}, \underset{\sim}{S}_1, \ldots, \underset{\sim}{S}_g, \underset{\sim}{\pi} \mid II) = f(\underset{\sim}{R} \mid S_1, \ldots, S_g, \pi, II)$$

$$\times f(\underset{\sim}{S}_1, \ldots, \underset{\sim}{S}_g \mid \pi, II)\, f(\underset{\sim}{\pi} \mid II). \qquad (2.2.25)$$

We have just shown that $f(\underset{\sim}{R} \mid S_1, \ldots S_g, \pi, II)$ is a $N(\mu_R^*, V_R^*)$ density. We showed before that, conditional on π and I, $S_1, \ldots, S_g$ are independent of Y_A, Y_B and Z_S. Then, using (2.2.19), we have that $f(\underset{\sim}{S}_1, \ldots, \underset{\sim}{S}_g \mid \pi, II)$ is a product of Binomial densities. Finally, by (2.2.11) we see that

$$f(\underset{\sim}{R},\underset{\sim}{S}_1,\ldots,\underset{\sim}{S}_g,\underset{\sim}{\pi}\,|\,\mathrm{II}) = (2\pi V_R^*)^{-1/2}\exp\{-\frac{1}{2V_R^*}(R-\mu_R^*)^2\}$$
$$\times \prod_{i=1}^{g}\left[\binom{d_i}{s_i}\pi_i^{S_i}(1-\pi_i)^{d_i-S_i}\right] \qquad (2.2.26)$$
$$\times \prod_{i=1}^{g}\left\{\frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')}\pi_i^{p_i'-1}(1-\pi_i)^{q_i'-1}\right\}.$$

Next we "integrate out" $S_1,\ldots,S_g$ to obtain

$$f(\underset{\sim}{R},\underset{\sim}{\pi}\,|\,\mathrm{II}) = \sum_{S_g=0}^{d_g}\cdots\sum_{S_1=0}^{d_1}\left\{\left[\prod_{i=1}^{g}\frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')}\binom{d_i}{s_i}\pi_i^{(S_i+p_i')-1}(1-\pi_i)^{(d_i-S_i+q_i')-1}\right]\right.$$
$$\left.\times\ (2\pi V_R^*)^{-1/2}\exp\{-\frac{1}{2V_R^*}(R-\mu_R^*)^2\}\right\}. \qquad (2.2.27)$$

Finally, we observe that, for $\alpha_1,\alpha_2 > 0$,

$$\int_0^1 t^{\alpha_1-1}(1-t)^{\alpha_2-1}\,dt = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)}\ .$$

This fact shows that

$$f(\underset{\sim}{R}\,|\,\mathrm{II}) = \sum_{S_g=0}^{d_g}\cdots\sum_{S_1=0}^{d_1}\left\{\left[\prod_{i=1}^{g}\frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')}\binom{d_i}{S_i}\frac{\Gamma(S_i+p_i')\Gamma(d_i-S_i+q_i')}{\Gamma(p_i'+q_i'+d_i)}\right]\right.$$
$$\left.\times\ (2\pi V_R^*)^{-1/2}\exp\{-\frac{1}{2V_R^*}(R-\mu_R^*)^2\}\right\}. \qquad (2.2.28)$$

This is the posterior density function used for making inferences about R and hence about T.

2.3. Posterior Prediction

Section 1.2 showed that the Bayes predictor of T under squared error loss is

$$\hat{T} := \sum_{ij \in A \cup B} Y_{ij} + \hat{R} \tag{2.3.1}$$

where $\hat{R} := E(R|\Pi)$. In this section we will find $E(R|\Pi)$ as well as $Var(R|\Pi)$, the posterior variance of R.

Recall from (2.2.11) that, given Π, $\pi_1, \pi_2, \ldots, \pi_g$ are independent random variables such that

$$\pi_i \sim \text{Beta}(p_i', q_i'), \quad i = 1, 2, \ldots, g.$$

That is,

$$f(\underset{\sim}{\pi}_i | \Pi) = \frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \pi_i^{p_i'-1} (1-\pi_i)^{q_i'-1}, \quad 0 \le \pi_i \le 1; \quad i = 1, 2, \ldots, g,$$

with $\pi_1, \ldots \pi_g$ independent. Under squared error loss, the Bayes estimate of π_i is $\hat{\pi}_i$ where

$$\hat{\pi}_i := E(\pi_i \mid II) = \frac{p_i'}{p_i'+q_i'} . \tag{2.3.2}$$

Note also that

$$\operatorname{Var}(\pi_i \mid II) = \hat{\pi}_i(1-\hat{\pi}_i)\left(\frac{1}{p_i'+q_i'+1}\right) \tag{2.3.3}$$

and

$$E(\pi_i^2 \mid II) = \hat{\pi}_i\left[\frac{p_i'+1}{p_i'+q_i'+1}\right] . \tag{2.3.4}$$

Recall from (2.2.18) and (2.2.19) that, given $\pi = (\pi_1, \ldots, \pi_g)'$ and II, $S_1, S_2, \ldots, S_g$ are independent random variables such that $S_i \sim \text{Binomial}(d_i, \pi_i)$, $i = 1, 2, \ldots, g$. That is,

$$f(\underset{\sim}{S}_i \mid \pi, II) = \binom{d_i}{S_i}\pi_i^{S_i}(1-\pi_i)^{d_i-S_i} ; \quad S_i = 0, 1, 2, \ldots, d_i , \quad i = 1, 2, \ldots, g,$$

with $S_1, \ldots, S_g$ independent. Thus,

$$E(S_i \mid \pi, II) = d_i\pi_i , \tag{2.3.5}$$

$$\operatorname{Var}(S_i \mid \pi, II) = d_i\pi_i(1-\pi_i) = d_i\pi_i - d_i\pi_i^2 , \tag{2.3.6}$$

and

$$E(S_i^2 \mid \pi, II) = d_i\pi_i + d_i(d_i-1)\pi_i^2 . \tag{2.3.7}$$

Recall that, for any random variable F and any random vector G,

$$E(F) = E_G[E(F|G)] \tag{2.3.8}$$

and

$$\mathrm{Var}(F) = E_G[\mathrm{Var}(F|G)] + \mathrm{Var}_G[E(F|G)] \tag{2.3.9}$$

where $E_G[\cdot]$ and $\mathrm{Var}_G[\cdot]$ are defined to be expectation and variance operators with respect to the joint distribution of the random variables that are the components of G. Note that $E(F|G)$ and $\mathrm{Var}(F|G)$ are both real-valued functions of the components of G.

We will use (2.3.2) through (2.3.9) to find $E(R|\mathrm{II})$ and $\mathrm{Var}(R|\mathrm{II})$. We have

$$E(R|\mathrm{II}) = E_\pi\{E(R|\pi, \mathrm{II})\}$$
$$= E_\pi\{E_{S_1,\ldots,S_g}[E(R|S_1,\ldots,S_g,\pi,\mathrm{II})]\}.$$

In Section 2.2 we showed that, given $S_1,\ldots,S_g$, π and II, $R \sim N(\mu_R^*, V_R^*)$. (See (2.2.23) and (2.2.24).) Thus,

$$E(R|\mathrm{II}) = E_\pi\{E_{S_1,\ldots,S_g}[\mu_R^*|\pi,\mathrm{II}]\}$$

$$= E_\pi\left\{E_{S_1,\ldots,S_g}\left[\sum_{i=1}^{g}(S_iX_i'\hat{\beta}_1+(c_i+d_i-S_i)X_i'\hat{\beta}_2)\Big|\pi,\mathrm{II}\right]\right\}$$

$$= E_\pi\left\{\sum_{i=1}^{g}E_{S_1,\ldots,S_g}[S_iX_i'\hat{\beta}_1+(c_i+d_i-S_i)X_i'\hat{\beta}_2|\pi,\mathrm{II}]\right\} =$$

$$= E_{\pi}\left\{ \sum_{i=1}^{g} [d_i \pi_i X_i' \hat{\beta}_1 + (c_i + d_i - d_i \pi_i) X_i' \hat{\beta}_2] \middle| \text{II} \right\}$$

$$= \sum_{i=1}^{g} [d_i \hat{\pi}_i X_i' \hat{\beta}_1 + (c_i + d_i - d_i \hat{\pi}_i) X_i' \hat{\beta}_2] \ , \qquad (2.3.10)$$

using (2.3.2) and (2.3.5).

By re-writing (2.3.10) we see that

$$\hat{R} := E(R|\text{II}) = \sum_{i=1}^{g} c_i X_i' \hat{\beta}_2 + \sum_{i=1}^{g} d_i [\hat{\pi}_i X_i' \hat{\beta}_1 + (1-\hat{\pi}_i) X_i' \hat{\beta}_2]. \qquad (2.3.11)$$

Similar arguments show that the Bayes predictor of Y_{ij} is $X_i' \hat{\beta}_2$ if $ij \in C$ or $\hat{\pi}_i X_i' \hat{\beta}_1 + (1-\hat{\pi}_i) X_i' \hat{\beta}_2$ if $ij \in D$. If $ij \in C$ then we know the unit is an NR, and the parameter vector β_2 applies. Not knowing β_2, we use $\hat{\beta}_2$, the posterior mean of β_2. If $ij \in D$, we do not know whether β_1 or β_2 applies, since we do not know whether the unit is an SR or an NR. We thus use a weighted average of $X_i' \hat{\beta}_1$ and $X_i' \hat{\beta}_2$. The weights are $\hat{\pi}_i$ and $(1-\hat{\pi}_i)$, the posterior estimates of the probabilities that unit ij is an SR or an NR respectively.

The posterior variance, $\text{Var}(R|\text{II})$, will be seen in Chapter IV to be helpful in making interval estimates of R (hence of T). By (2.3.9), $\text{Var}(R|\text{II}) = E_{\pi}\{\text{Var}(R|\pi, \text{II})\} + \text{Var}_{\pi}\{E(R|\pi, \text{II})\}$. Again by (2.3.9),

$$\mathrm{Var}(R|\pi, II) = E_{S_1,\ldots,S_g}[\mathrm{Var}(R|S_1,\ldots,S_g,\pi,II)]$$
$$+ \mathrm{Var}_{S_1,\ldots,S_g}[E(R|S_1,\ldots,S_g,\pi,II)].$$

By (2.3.8),

$$E(R|\pi, II) = E_{S_1,\ldots,S_g}[E(R|S_1,\ldots,S_g,\pi,II)].$$

Thus,

$$\mathrm{Var}(R|II) = E_\pi\{E_{S_1,\ldots,S_g}[\mathrm{Var}(R|S_1,\ldots,S_g,\pi,II)]\}$$
$$+ E_\pi\{\mathrm{Var}_{S_1,\ldots,S_g}[E(R|S_1,\ldots,S_g,\pi,II)]\}$$
$$+ \mathrm{Var}_\pi\{E_{S_1,\ldots,S_g}[E(R|S_1,\ldots,S_g,\pi,II)]\}. \tag{2.3.12}$$

Using (2.2.24), (2.3.5), (2.3.7) and the conditional independence of $S_1,\ldots,S_g$, we have

$$E_\pi\{E_{S_1,\ldots,S_g}[\mathrm{Var}(R|S_1,\ldots,S_g,\pi,II)]\}$$
$$= E_\pi\{E_{S_1,\ldots,S_g}[V_R^*|\pi, II]\}$$
$$= E_\pi\Bigg\{(n-m-r)h_2^{-1}+\sum_{i=1}^{g}[h_1^{-1}d_i\pi_i+h_2^{-1}d_i(1-\pi_i)]$$
$$+ h_1^{-1}\sum_{i=1}^{g}(d_i\pi_i+d_i(d_i-1)\pi_i^2)\delta_{ii} + h_1^{-1}\sum_{i=1}^{g}\sum_{\substack{k=1\\k\neq i}}^{g} d_i\pi_i d_k\pi_k\delta_{ik} +$$

$$+ h_2^{-1} \sum_{i=1}^{g} [(c_i+d_i)^2 - 2(c_i+d_i)d_i\pi_i + d_i\pi_i + d_i(d_i-1)\pi_i^2]\epsilon_{ii}$$

$$+ h_2^{-1} \sum_{i=1}^{g} \sum_{\substack{k=1 \\ k\neq i}}^{g} (c_i+d_i-d_i\pi_i)(c_k+d_k-d_k\pi_k)\epsilon_{ij} \Big| \mathrm{II} \Big\} \qquad (2.3.13)$$

where, for $i, k = 1, 2, \ldots, g$,

$$\delta_{ik} := X_i'(X_A'X_A + \tau_1)^{-1}X_k \qquad (2.3.14)$$

and

$$\epsilon_{ik} := X_i'(X_B'X_B + \tau_2)^{-1}X_k \,. \qquad (2.3.15)$$

Using (2.3.2), (2.3.4) and the conditional independence of $\pi_1, \ldots, \pi_g$, expression (2.3.13) simplifies to become

$$E_\pi\{E_{S_1,\ldots,S_g}[\mathrm{Var}(R|S_1,\ldots,S_g,\pi,\mathrm{II})]\}$$

$$= (n-m-r)h_2^{-1} + \sum_{i=1}^{g} [h_1^{-1}d_i\hat{\pi}_i + h_2^{-1}d_i(1-\hat{\pi}_i)]$$

$$+ h_1^{-1} \sum_{i=1}^{g} [d_i\hat{\pi}_i + d_i(d_i-1)\hat{\pi}_i(\frac{p_i'+1}{p_i'+q_i'+1})]\delta_{ii} + h_1^{-1} \sum_{i=1}^{g} \sum_{\substack{k=1 \\ k\neq i}}^{g} d_i d_k \hat{\pi}_i \hat{\pi}_k \delta_{ik}$$

$$+ h_2^{-1} \sum_{i=1}^{g} [(c_i+d_i)^2 - 2(c_i+d_i)d_i\hat{\pi}_i + d_i\hat{\pi}_i + d_i(d_i-1)\hat{\pi}_i(\frac{p_i'+1}{p_i'+q_i'+1})]\epsilon_{ii} \quad +$$

$$+ h_2^{-1} \sum_{i=1}^{g} \sum_{\substack{k=1 \\ k \neq i}}^{g} [c_i + d_i(1-\hat{\pi}_i)][c_k + d_k(1-\hat{\pi}_k)]\epsilon_{ik} \, . \tag{2.3.16}$$

Using (2.2.23), (2.3.6) and the conditional independence of $S_1, \ldots, S_g$, we have

$$\mathrm{Var}_{S_1,\ldots,S_g}[E(R|S_1,\ldots,S_g,\pi,\mathrm{II})]$$

$$= \mathrm{Var}_{S_1,\ldots,S_g}[\mu_R^*|\pi,\mathrm{II}]$$

$$= \mathrm{Var}_{S_1,\ldots,S_g}\left\{\sum_{i=1}^{g}(c_i+d_i)X_i'\hat{\beta}_2 + \sum_{i=1}^{g} S_i[X_i'(\hat{\beta}_1-\hat{\beta}_2)]\middle|\pi,\mathrm{II}\right\}$$

$$= \mathrm{Var}_{S_1,\ldots,S_g}\left\{\sum_{i=1}^{g} S_i[X_i'(\hat{\beta}_1-\hat{\beta}_2)]\middle|\pi,\mathrm{II}\right\}$$

$$= \sum_{i=1}^{g}[d_i\pi_i - d_i\pi_i^2][X_i'(\hat{\beta}_1-\hat{\beta}_2)]^2 \, .$$

Thus, using (2.3.5),

$$E_\pi\{\mathrm{Var}_{S_1,\ldots,S_g}[E(R|S_1,\ldots,S_g,\pi,\mathrm{II})]\}$$

$$= E_\pi\left\{\sum_{i=1}^{g}[d_i\pi_i - d_i\pi_i^2][X_i'(\hat{\beta}_1-\hat{\beta}_2)]^2\middle|\mathrm{II}\right\} \quad =$$

$$= \sum_{i=1}^{g} [d_i\hat{\pi}_i - d_i\hat{\pi}_i(\frac{p_i'+1}{p_i'+q_i'+1})][X_i'(\hat{\beta}_1 - \hat{\beta}_2)]^2. \qquad (2.3.17)$$

Finally, using (2.2.23), (2.3.5), (2.3.3) and the conditional independence of $\pi_1, \ldots, \pi_g$, we have

$$\mathrm{Var}_\pi\{E_{S_1,\ldots,S_g}[E(R|S_1,\ldots,S_g,\pi,\mathrm{II})]\}$$

$$= \mathrm{Var}_\pi\{E_{S_1,\ldots,S_g}[\mu_R^*|\pi,\mathrm{II}]\}$$

$$= \mathrm{Var}_\pi\left\{E_{S_1,\ldots,S_g}\left[\sum_{i=1}^{g}(c_i+d_i)X_i'\hat{\beta}_2 + \sum_{i=1}^{g} S_i X_i'(\hat{\beta}_1-\hat{\beta}_2)\middle|\pi,\mathrm{II}\right]\right\}$$

$$= \mathrm{Var}_\pi\left\{\sum_{i=1}^{g}(c_i+d_i)X_i'\beta_2 + \sum_{i=1}^{g} d_i\pi_i X_i'(\hat{\beta}_1-\hat{\beta}_2)\middle|\mathrm{II}\right\}$$

$$= \mathrm{Var}_\pi\left\{\sum_{i=1}^{g} d_i\pi_i X_i'(\hat{\beta}_1-\hat{\beta}_2)\middle|\mathrm{II}\right\}$$

$$= \sum_{i=1}^{g} \mathrm{Var}(d_i\pi_i X_i'(\hat{\beta}_1-\hat{\beta}_2)|\mathrm{II})$$

$$= \sum_{i=1}^{g} d_i^2\hat{\pi}_i(1-\hat{\pi}_i)(\frac{1}{p_i'+q_i'+1})[X_i'(\hat{\beta}_1-\hat{\beta}_2)]^2 . \qquad (2.3.18)$$

If we now combine (2.3.16), (2.3.17) and (2.3.18) with (2.3.12) and simplify the resulting expression, we obtain

$$\begin{aligned}
\operatorname{Var}(R|II) = {} & (n-m-r)h_2^{-1} + \sum_{i=1}^{g} [h_1^{-1} d_i \hat{\pi}_i + h_2^{-1} d_i (1-\hat{\pi}_i)] \\
& + \sum_{i=1}^{g} d_i \hat{\pi}_i \left[\frac{q_i' + d_i (p_i'+1)}{p_i' + q_i' + 1} \right] [h_1^{-1} \delta_{ii} + h_2^{-1} \epsilon_{ii}] \\
& + h_1^{-1} \sum_{i=1}^{g} \sum_{\substack{k=1 \\ k \neq i}}^{g} d_i d_k \hat{\pi}_i \hat{\pi}_k \delta_{ik} + h_2^{-1} \sum_{i=1}^{g} (c_i + d_i)(c_i + d_i - 2 d_i \hat{\pi}_i) \epsilon_{ii} \\
& + h_2^{-1} \sum_{i=1}^{g} \sum_{\substack{k=1 \\ k \neq i}}^{g} [c_i + d_i (1-\hat{\pi}_i)][c_k + d_k (1-\hat{\pi}_k)] \epsilon_{ik} \\
& + \sum_{i=1}^{g} d_i \hat{\pi}_i (1-\hat{\pi}_i) \left[\frac{p_i' + q_i' + d_i}{p_i' + q_i' + 1} \right] [X_i'(\hat{\beta}_1 - \hat{\beta}_2)]^2 \qquad (2.3.19)
\end{aligned}$$

where, for $i, k = 1, 2, \ldots, g$, δ_{ik} and ϵ_{ik} are given by (2.3.14) and (2.3.15) respectively.

2.4. Special Case: π Known

In this chapter we have assumed that the vector π is random. Given the $g \times 1$ parameter vectors $p = (p_1, \ldots, p_g)'$ and $q = (q_1, \ldots, q_g)'$, we assume that $\pi_1, \ldots, \pi_g$ are independent random variables such that

$$\pi_i \sim \operatorname{Beta}(p_i, q_i), \quad i = 1, 2, \ldots, g.$$

Then

$$E(\pi_i \mid p_i, q_i) = \frac{p_i}{p_i+q_i}$$

and

$$\mathrm{Var}(\pi_i \mid p_i, q_i) = \frac{p_i q_i}{(p_i+q_i+1)(p_i+q_i)^2}.$$

Suppose we prefer to treat π as a known vector of constants. That is, suppose we assume that $\pi_i = \phi_i$ for known $\phi_i \in [0, 1]$, $i = 1, 2, \ldots, g$. We can incorporate this change in the model of this chapter by letting (p_i+q_i) approach $+\infty$ while keeping $\frac{p_i}{p_i+q_i} = \phi_i$ fixed. Then $\mathrm{Var}(\pi_i \mid p_i, q_i)$ becomes zero while $E(\pi_i \mid p_i, q_i) = \phi_i$. The prior density $f(\underset{\sim}{\pi}_i \mid p_i, q_i)$ becomes degenerate at ϕ_i for $i = 1, 2, \ldots, g$ so that (2.1.1) becomes

$$\begin{aligned} f(\underset{\sim}{Y}, \underset{\sim}{Z}, \underset{\sim}{\beta}, \underset{\sim}{\pi} \mid I) \propto\ & (|V_N|)^{-1/2} \exp\{-\tfrac{1}{2}(Y-\mu_N)' V_N^{-1} (Y-\mu_N)\} \\ & \times \prod_{i=1}^{g} \prod_{j=1}^{n_i} [\pi_i^{Z_{ij}} (1-\pi_i)^{1-Z_{ij}}] \\ & \times (|h_1 \tau_1|)^{1/2} \exp\{-\tfrac{1}{2}(\beta_1-\lambda_1)'(h_1\tau_1)(\beta_1-\lambda_1)\} \\ & \times (|h_2 \tau_2|)^{1/2} \exp\{-\tfrac{1}{2}(\beta_2-\lambda_2)'(h_2\tau_2)(\beta_2-\lambda_2)\} \end{aligned} \tag{2.4.1}$$

for $\pi_i = \phi_i$, $i = 1, 2, \ldots, g$.

Similarly,

$$f(\underset{\sim}{\pi} | II) = \prod_{i=1}^{g} 1_{\{\phi_i\}}(\pi_i)$$

is now a discrete density where

$$1_{\{\phi_i\}}(\pi_i) = \begin{cases} 0, & \pi_i \neq \phi_i \\ 1, & \pi_i = \phi_i \end{cases} .$$

Use of this fact in (2. 2. 25) and (2. 2. 26) leads to

$$f(\underset{\sim}{R}, \underset{\sim}{S}_1, \ldots, \underset{\sim}{S}_g, \underset{\sim}{\pi} | II) = (2\pi V_R^*)^{-1/2} \exp\{-\frac{1}{2V_R^*} (R-\mu_R^*)^2\} \tag{2.4.2}$$

$$\times \prod_{i=1}^{g} [\binom{d_i}{S_i} \pi_i^{S_i} (1-\pi_i)^{d_i - S_i}] \prod_{i=1}^{g} 1_{\{\phi_i\}}(\pi_i) .$$

To "integrate out" π, we simply replace π_i by ϕ_i, $i = 1, 2, \ldots, g$. Then

$$f(\underset{\sim}{R}, \underset{\sim}{S}_1, \ldots, \underset{\sim}{S}_g | II) = (2\pi V_R^*)^{-1/2} \exp\{-\frac{1}{2V_R^*} (R-\mu_R^*)^2\}$$

$$\times \prod_{i=1}^{g} [\binom{d_i}{S_i} \phi_i^{S_i} (1-\phi_i)^{d_i - S_i}] . \tag{2.4.3}$$

Finally,

$$f(\underset{\sim}{R} \mid \text{II}) = \sum_{S_g=0}^{d_g} \cdots \sum_{S_1=0}^{d_1} \left\{ \left[\prod_{i=1}^{g} \binom{d_i}{S_i} \phi_i^{S_i} (1-\phi_i)^{d_i - S_i} \right] \times (2\pi V_R^*)^{-1/2} \exp\{- \frac{1}{2V_R^*} (R - \mu_R^*)^2\} \right\} . \quad (2.4.4)$$

A slightly different approach is to examine $f(\underset{\sim}{R} \mid \text{II})$, given by (2.2.28), in the case where $p_i \to +\infty$ and $q_i \to +\infty$ while $\frac{p_i}{p_i + q_i} = \phi_i$ remains fixed. Note that $\frac{p_i}{p_i+q_i} = \phi_i$ remains fixed while $p_i \to +\infty$ and $q_i \to +\infty$ if and only if $p_i \to +\infty$ and $q_i \to +\infty$ while $\frac{q_i}{p_i} = \frac{1-\phi_i}{\phi_i}$ remains fixed.

Now, for any $i = 1, 2, \ldots, g$,

$$\frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \cdot \frac{\Gamma(S_i+p_i')\Gamma(d_i-S_i+q_i')}{\Gamma(p_i'+q_i'+d_i)}$$

$$= \int_0^1 t^{S_i}(1-t)^{d_i-S_i} \left[\frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} t^{p_i'-1} (1-t)^{q_i'-1} \right] dt$$

$$= E[\pi_i^{S_i}(1-\pi_i)^{d_i-S_i}]$$

where $\pi_i \sim \text{Beta}(p_i', q_i')$. Suppose $p_i \to +\infty$ and $q_i \to +\infty$ while $\frac{q_i}{p_i} = \frac{1-\phi_i}{\phi_i}$ remains fixed. Consider the $\text{Beta}(p_i', q_i')$ density. Its mean is

$$\frac{p_i'}{p_i'+q_i'} = \frac{p_i+a_i}{p_i+q_i+s_i} = \frac{1+\frac{a_i}{p_i}}{1+\frac{q_i}{p_i}+\frac{s_i}{p_i}}$$

which becomes

$$\frac{1}{1+\frac{q_i}{p_i}} = \frac{p_i}{p_i+q_i} = \phi_i \ .$$

Its variance is

$$\frac{p_i'q_i'}{(p_i'+q_i'+1)(p_i'+q_i')^2} = \frac{p_i'}{p_i'+q_i'} \cdot \frac{q_i'}{p_i'+q_i'} \cdot \frac{1}{p_i'+q_i'+1}$$

which becomes zero. Thus the $\mathrm{Beta}(p_i', q_i')$ density becomes degenerate at ϕ_i so that

$$E[\pi_i^{S_i}(1-\pi_i)^{d_i-S_i}] = \phi_i^{S_i}(1-\phi_i)^{d_i-S_i} \ .$$

We have just shown that the case of π known leads to replacing

$$\frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \cdot \frac{\Gamma(S_i+p_i')\Gamma(d_i-S_i+q_i')}{\Gamma(p_i'+q_i'+d_i)}$$

by $\phi^{S_i}(1-\phi)^{d_i-S_i}$ for $i = 1, 2, \ldots, g$. Use of this fact in (2.2.28) again gives (2.4.4).

We have also shown that $\hat{\pi}_i := \dfrac{p_i'}{p_i'+q_i'}$ becomes ϕ_i when we know that $\pi_i = \phi_i$, $i = 1, 2, \ldots, g$. If we now replace $\hat{\pi}_i$ by ϕ_i in (2.3.11), we obtain

$$\hat{R} = \sum_{i=1}^{g} c_i X_i'\hat{\beta}_2 + \sum_{i=1}^{g} d_i[\phi_i X_i'\hat{\beta}_1 + (1-\phi_i)X_i'\hat{\beta}_2] . \tag{2.4.5}$$

Similar changes occur in the expression for $\mathrm{Var}(R|\mathrm{II})$ given by (2.3.19). Note that

$$\frac{q_i'+d_i(p_i'+1)}{p_i'+q_i'+1} = \frac{q_i+s_i-a_i+d_i(p_i+a_i+1)}{p_i+q_i+s_i}$$

$$= \frac{\dfrac{q_i}{p_i} + \dfrac{(s_i-a_i)}{p_i} + d_i\left(1+\dfrac{a_i}{p_i}+\dfrac{1}{p_i}\right)}{1+\dfrac{q_i}{p_i}+\dfrac{s_i}{p_i}}$$

which becomes

$$\frac{\dfrac{q_i}{p_i} + d_i}{1+\dfrac{q_i}{p_i}} = \frac{q_i+d_ip_i}{p_i+q_i} = (1-\phi_i) + d_i\phi_i .$$

Also note that

$$\frac{p_i'+q_i'+d_i}{p_i'+q_i'+1} = \frac{p_i+a_i+q_i+s_i-a_i+d_i}{p_i+a_i+q_i+s_i-a_i+1}$$

$$= \frac{1+\frac{a_i}{p_i}+\frac{q_i}{p_i}+\frac{(s_i-a_i+d_i)}{p_i}}{1+\frac{a_i}{p_i}+\frac{q_i}{p_i}+\frac{(s_i-a_i+1)}{p_i}}$$

becomes

$$\frac{1+\frac{q_i}{p_i}}{1+\frac{q_i}{p_i}} = 1.$$

We thus have

$$\begin{aligned} \mathrm{Var}(R\,|\,II) = {} & (n-m-r)h_2^{-1} + \sum_{i=1}^{g} h_2^{-1} d_i + \sum_{i=1}^{g} d_i\phi_i(h_1^{-1}-h_2^{-1}) \\ & + \sum_{i=1}^{g} d_i\phi_i[1-\phi_i+d_i\phi_i][h_1^{-1}\delta_{ii}+h_2^{-1}\epsilon_{ii}] \\ & + \sum_{i=1}^{g} (c_i+d_i)(c_i+d_i-2d_i\phi_i)(h_2^{-1}\epsilon_{ii}) \\ & + \sum_{i=1}^{g} d_i\phi_i(1-\phi_i)[X_i'(\hat{\beta}_1-\hat{\beta}_2)]^2 \quad + \end{aligned}$$

$$+ 2\sum_{i=1}^{g}\sum_{k>i} d_i d_k \phi_i \phi_k h_1^{-1}\delta_{ik}$$

$$+ 2\sum_{i=1}^{g}\sum_{k>i}[c_i+d_i(1-\phi_i)][c_k+d_k(1-\phi_k)](h_2^{-1}\epsilon_{ik}). \tag{2.4.6}$$

2.5. Special Case: Population not Grouped

Suppose the population cannot be grouped on the basis of auxiliary information; that is, suppose $g = N$. Then U can be relabeled as $U = \{1, 2, \ldots, N\}$. Each a_i, b_i, c_i, d_i and s_i, $i = 1, 2, \ldots, g$, is either one or zero depending on whether or not the i^{th} unit in U is in A, B, C, D and S respectively. We relabel Y, Z, X, π, p and q to correspond to the relabeling of U, and we write $T = \sum_{i=1}^{N} Y_i$.

The sums $S_1, \ldots, S_g$ are now such that

$$S_i = \begin{cases} 0, & \text{unit } i \text{ not in } D \\ Z_i, & \text{unit } i \text{ in } D\,. \end{cases}$$

Similarly,

$$R = \sum_{i\in C\cup D} Y_i\,.$$

The posterior density $f(\underset{\sim}{\beta}|II)$ remains as given by (2.2.9), and $\hat{\beta} = \begin{pmatrix}\hat{\beta}_1\\ \hat{\beta}_2\end{pmatrix}$ is still as in (2.2.9). Note, however, that X_A and X_B are relabeled as appropriate.

The posterior density, $f(\underset{\sim}{\pi}|II)$, given by (2.2.11) now becomes

$$f(\underset{\sim}{\pi}|II) = \prod_{i=1}^{N} \left\{ \frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \pi_i^{p_i'-1} (1-\pi_i)^{q_i'-1} \right\}$$

where

$$p_i' = p_i + a_i = \begin{cases} p_i+1, & \text{unit } i \text{ in } A \\ p_i, & \text{unit } i \text{ not in } A \end{cases}$$

and where

$$q_i' = q_i + s_i - a_i = \begin{cases} q_i+1, & \text{unit } i \text{ in } B \cup C \\ q_i, & \text{unit } i \text{ not in } B \cup C. \end{cases}$$

Suppose $i \in A \cup B \cup C$. Then $d_i = S_i = 0$, $\binom{d_i}{S_i} = 1$, and $h_i(S_i) = 1$ where, for $i = 1, 2, \ldots, g$,

$$h_i(S_i) := \frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \binom{d_i}{S_i} \frac{\Gamma(p_i'+S_i)\Gamma(q_i'+d_i-S_i)}{\Gamma(p_i'+q_i'+d_i)}.$$

If $i \in D$ then $d_i = 1$, $p_i' = p_i$, $q_i' = q_i$ and

$$h_i(S_i) = \frac{\Gamma(p_i+q_i)}{\Gamma(p_i)\Gamma(q_i)} \binom{1}{S_i} \frac{\Gamma(S_i+p_i)\Gamma(1-S_i+q_i)}{\Gamma(p_i+q_i+1)}$$

$$= \frac{\Gamma(p_i+q_i)}{\Gamma(p_i)\Gamma(q_i)} \frac{\Gamma(S_i+p_i)\Gamma(1-S_i+q_i)}{\Gamma(p_i+q_i+1)}$$

If $S_i = 0$,

$$h_i(S_i) = \frac{\Gamma(p_i+q_i)}{\Gamma(p_i)\Gamma(q_i)} \frac{\Gamma(p_i)q_i\Gamma(q_i)}{(p_i+q_i)\Gamma(p_i+q_i)}$$

$$= \frac{q_i}{p_i+q_i}$$

$$= \frac{q_i'}{p_i'+q_i'}$$

$$= 1 - \hat{\pi}_i$$

$$= (\hat{\pi}_i)^{S_i}(1-\hat{\pi}_i)^{1-S_i} .$$

If $S_i = 1$,

$$h_i(S_i) = \frac{p_i}{p_i+q_i} = \hat{\pi}_i = (\hat{\pi}_i)^{S_i}(1-\hat{\pi}_i)^{1-S_i} .$$

The posterior density, $f(\underset{\sim}{R}|\Pi)$, given by (2.2.28) becomes

$$f(\underset{\sim}{R}|\Pi) = \sum_{\substack{Z_i \in \{0,1\} \\ i \in D}} \left\{ \left[\prod_{i \in D} (\hat{\pi}_i)^{Z_i}(1-\hat{\pi}_i)^{1-Z_i} \right] \times \left[(2\pi V_R^*)^{-1/2} \exp\{-\frac{1}{2V_R^*}(R-\mu_R^*)^2\} \right] \right\}, \tag{2.5.1}$$

since $S_i = Z_i$ for $i \in D$.

For $i \in C$, $c_i = 1$ while, for $i \in D$, $d_i = 1$ and $\hat{\pi}_i = \frac{p_i}{p_i+q_i}$, the prior mean of π_i. Thus, (2.3.11) becomes

$$\hat{R} = \sum_{i \in C} X_i'\hat{\beta}_2 + \sum_{i \in D} [\hat{\pi}_i X_i'\hat{\beta}_1 + (1-\hat{\pi}_i)X_i'\hat{\beta}_2] \,, \tag{2.5.2}$$

and we have

$$\hat{T} = \sum_{i \in A \cup B} Y_i + \hat{R} \,. \tag{2.5.3}$$

Similar changes can be made in $Var(R|II)$ given by (2.3.19).

III. A BAYESIAN APPROACH: PRECISIONS UNKNOWN

In Chapter II we explored a Bayesian approach to the non-response problem of Section 1.3. We assumed that the precisions h_1 and h_2 were known (see Section 2.1). In this chapter we will relax this assumption and present a Bayesian approach when h_1 and h_2 are random variables.

3.1. Revision of Model and Prior Density Function

Recall the model of Section 2.1 where the precisions h_1 and h_2 are known. We now consider a revised model that is the same as the previous model except that, instead of assuming h_1 and h_2 are known, we assume that h_1 and h_2 are random variables. Specifically, we now assume that h_1 and h_2 are independent random variables having gamma prior distributions such that, for known constants η_1, η_2, e_1 and e_2,

$$f(\underset{\sim}{h}_1, \underset{\sim}{h}_2 \mid \eta_1, e_1, \eta_2, e_2) = f(\underset{\sim}{h}_1 \mid \eta_1, e_1) f(\underset{\sim}{h}_2 \mid \eta_2, e_2)$$

$$\propto \{h_1^{(e_1/2)-1} e^{-(e_1 h_1)/2\eta_1}\}$$

$$\times \{h_2^{(e_2/2)-1} e^{-(e_2 h_2)/2\eta_2}\}$$

for $h_1, h_2 > 0$. We again write $h = \binom{h_1}{h_2}$.

Recall that $I := \{X, h_1, h_2, \lambda_1, \lambda_2, \tau_1, \tau_2, p, q\}$. Let $I^* := \{X, \eta_1, e_1, \eta_2, e_2, \lambda_1, \lambda_2, \tau_1, \tau_2, p, q\}$. Then $I^* \cup \{h_1, h_2\} = I \cup \{\eta_1, e_1, \eta_2, e_2\}$.

The joint prior density of β and h is

$$\begin{aligned} f(\underset{\sim}{\beta}, \underset{\sim}{h} \mid I^*) &= f(\underset{\sim}{\beta} \mid h, I^*) f(\underset{\sim}{h} \mid I^*) \\ &= f(\underset{\sim}{\beta} \mid I, \eta_1, e_1, \eta_2, e_2) f(\underset{\sim}{h} \mid I^*) \\ &= f(\underset{\sim}{\beta} \mid I) f(\underset{\sim}{h} \mid \eta_1, e_1, \eta_2, e_2) , \end{aligned}$$

since the prior density of β, given I, does not involve η_1, e_1, η_2 or e_2, and since the prior density of h depends on I^* only through η_1, e_1, η_2, e_2. Thus,

$$\begin{aligned} f(\underset{\sim}{\beta}, \underset{\sim}{h} \mid I^*) &\propto (|h_1 \tau_1|)^{1/2} \exp\{-\tfrac{1}{2}(\beta_1 - \lambda_1)'(h_1 \tau_1)(\beta_1 - \lambda_1)\} \\ &\quad \times (|h_2 \tau_2|)^{1/2} \exp\{-\tfrac{1}{2}(\beta_2 - \lambda_2)'(h_2 \tau_2)(\beta_2 - \lambda_2)\} \\ &\quad \times [h_1^{(e_1/2)-1} e^{-(e_1 h_1)/2\eta_1}][h_2^{(e_2/2)-1} e^{-(e_2 h_2)/2\eta_2}] \\ &= h_1^{x/2} |\tau_1|^{1/2} \exp\{-\tfrac{h_1}{2}(\beta_1 - \lambda_1)'\tau_1(\beta_1 - \lambda_1)\} \\ &\quad \times h_2^{x/2} |\tau_2|^{1/2} \exp\{-\tfrac{h_2}{2}(\beta_2 - \lambda_2)'\tau_2(\beta_2 - \lambda_2)\} \\ &\quad \times h_1^{(e_1/2)-1} \exp\{-(\tfrac{e_1}{2\eta_1})h_1\} h_2^{(e_2/2)-1} \exp\{-(\tfrac{e_2}{2\eta_2})h_2\} \\ &\propto h_1^{(x+e_1)/2-1} \exp\{-\tfrac{1}{2}[\tfrac{e_1}{\eta_1} + (\beta_1 - \lambda_1)'\tau_1(\beta_1 - \lambda_1)]h_1\} \times \end{aligned}$$

$$\times\ h_2^{(x+e_2)/2-1} \exp\{-\frac{1}{2}[\frac{e_2}{\eta_2}+(\beta_2-\lambda_2)'\tau_2(\beta_2-\lambda_2)]h_2\}\ . \tag{3.1.1}$$

The marginal prior density of β in this revised model is

$$f(\underset{\sim}{\beta}|I^*) = \int_0^\infty \int_0^\infty f(\underset{\sim}{\beta}, \underset{\sim}{h}|I^*)dh_1 dh_2\ .$$

By noting that

$$\int_0^\infty t^{\alpha_1-1} e^{-\alpha_2 t} dt = (\alpha_2)^{-\alpha_1}\Gamma(\alpha_1)$$

for any $\alpha_2 > 0$, we have

$$f(\underset{\sim}{\beta}\ |I^*) \propto \{\frac{1}{2}[\frac{e_1}{\eta_1} + (\beta_1-\lambda_1)'\tau_1(\beta_1-\lambda_1)]\}^{-(x+e_1)/2}\Gamma(\frac{x+e_1}{2})$$

$$\times\ \{\frac{1}{2}[\frac{e_2}{\eta_2} + (\beta_2-\lambda_2)'\tau_2(\beta_2-\lambda_2)]\}^{-(x+e_2)/2}\Gamma(\frac{x+e_2}{2})$$

$$\propto [1+(\beta_1-\lambda_1)'\frac{\eta_1}{e_1}\tau_1(\beta_1-\lambda_1)]^{-(x+e_1)/2}$$

$$\times\ [1+(\beta_2-\lambda_2)'\frac{\eta_2}{e_2}\tau_2(\beta_2-\lambda_2)]^{-(x+e_2)/2}\ .$$

Clearly, $f(\underset{\sim}{\beta}|I^*) = f(\underset{\sim}{\beta}_1|I^*)f(\underset{\sim}{\beta}_2|I^*)$ where $f(\underset{\sim}{\beta}_1|I^*)$ is a $t(x, e_1, \lambda_1, \eta_1\tau_1)$ density and where $f(\underset{\sim}{\beta}_2|I^*)$ is a $t(x, e_2, \lambda_2, \eta_2\tau_2)$ density (see Definition A.3).

The revised model thus implies that, conditional on the known prior information contained in I^*, β_1 and β_2 have independent multivariate-t distributions. The choice of prior parameters might be made simpler by noting that $E(\beta_1|I^*) = \lambda_1$, $E(\beta_2|I^*) = \lambda_2$,

$$\mathrm{Cov}(\beta_1|I^*) = \frac{e_1}{e_1-2}(\eta_1^{-1}\tau_1^{-1}) \quad \text{and} \quad \mathrm{Cov}(\beta_2|I^*) = \frac{e_2}{e_2-2}(\eta_2^{-1}\tau_2^{-1})$$

(see A.3). Further, Fact A.7 shows that $f(\underset{\sim}{\beta}_1|I^*)$ tends to a multivariate normal density with mean λ_1 and covariance matrix $\eta_1^{-1}\tau_1^{-1}$ as $e_1 \to +\infty$ while $f(\underset{\sim}{\beta}_2|I^*)$ tends to a multivariate normal density with mean λ_2 and covariance matrix $\eta_2^{-1}\tau_2^{-1}$ as $e_2 \to +\infty$. Thus, "asymptotically," $f(\underset{\sim}{\beta}_1|I^*)$ is the same density as $f(\underset{\sim}{\beta}_1|I)$ except that h_1 is replaced by its prior mean, η_1. A similar statement holds for $f(\underset{\sim}{\beta}_2|I^*)$ and $f(\underset{\sim}{\beta}_2|I)$. Also note that the prior variances, $2\eta_1^2/e_1$ and $2\eta_2^2/e_2$, of h_1 and h_2 approach zero as $e_1 \to +\infty$ and $e_2 \to +\infty$.

We end this section by finding the joint prior density function, $f(\underset{\sim}{Y}, \underset{\sim}{Z}, \underset{\sim}{\beta}, \underset{\sim}{\pi}, \underset{\sim}{h}|I^*)$. Clearly, since η_1, e_1, η_2 and e_2 are constants,

$$\begin{aligned} f(\underset{\sim}{Y}, \underset{\sim}{Z}, \underset{\sim}{\beta}, \underset{\sim}{\pi}|h, I^*) &= f(\underset{\sim}{Y}, \underset{\sim}{Z}, \underset{\sim}{\beta}, \underset{\sim}{\pi}|I, \eta_1, e_1, \eta_2, e_2) \\ &= f(\underset{\sim}{Y}, \underset{\sim}{Z}, \underset{\sim}{\beta}, \underset{\sim}{\pi}|I). \end{aligned}$$

This density is given by (2. 1. 1). Thus,

$$
\begin{aligned}
f(\underset{\sim}{Y}, \underset{\sim}{Z}, \underset{\sim}{\beta}, \underset{\sim}{\pi}, \underset{\sim}{h} \mid I^{*}) &= f(\underset{\sim}{Y}, \underset{\sim}{Z}, \underset{\sim}{\beta}, \underset{\sim}{\pi} \mid h, I^{*}) f(\underset{\sim}{h} \mid I^{*}) \\
&\propto (|V_N|)^{-1/2} \exp\{-\tfrac{1}{2}(Y-\mu_N)' V_N^{-1} (Y-\mu_N)\} \\
&\quad \times \prod_{i=1}^{g} \prod_{j=1}^{n_i} [\pi_i^{Z_{ij}} (1-\pi_i)^{1-Z_{ij}}] \\
&\quad \times (|h_1 \tau_1|)^{1/2} \exp\{-\tfrac{1}{2}(\beta_1-\lambda_1)'(h_1\tau_1)(\beta_1-\lambda_1)\} \\
&\quad \times (|h_2 \tau_2|)^{1/2} \exp\{-\tfrac{1}{2}(\beta_2-\lambda_2)'(h_2\tau_2)(\beta_2-\lambda_2)\} \\
&\quad \times \prod_{i=1}^{g} [\pi_i^{p_i-1} (1-\pi_i)^{q_i-1}] \\
&\quad \times [h_1^{(e_1/2)-1} \exp(-\frac{e_1 h_1}{2\eta_1})][h_2^{(e_2/2)-1} \exp(-\frac{e_2 h_2}{2\eta_2})] \ .
\end{aligned}
\tag{3.1.2}
$$

3.2. Posterior Density Function

Define $II^{*} := I^{*} \cup \{Y_A, Y_B, Z_S\}$. Still treating h_1 and h_2 as random variables, we will find the posterior density function, $f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D \mid II^{*})$, in this section. In Section 2.2 we were able to find $f(\underset{\sim}{R} \mid II)$. While we will not transform $f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D \mid II^{*})$ to $f(\underset{\sim}{R} \mid II^{*})$, we will still be able to make posterior inference about R under squared error loss in Section 3.3.

Clearly,

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi}, \underset{\sim}{h} | II^*) = f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} | h, II^*) f(\underset{\sim}{h} | II^*) . \qquad (3.2.1)$$

Now,

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} | h, II^*) = f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} | II, \eta_1, e_1, \eta_2, e_2)$$

$$= f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi} | II) , \qquad (3.2.2)$$

since, conditional on II, the joint density of Y_C, Y_D, Z_D, β and π does not involve the constants η_1, e_1, η_2 and e_2. This density is given by (2.2.12).

It remains to find $f(\underset{\sim}{h} | II^*)$. Note that

$$f(\underset{\sim}{h} | II^*) = f(\underset{\sim}{h} | I^*, Y_A, Y_B, Z_S)$$

$$= \frac{f(\underset{\sim}{h}, \underset{\sim}{Y}_A, \underset{\sim}{Y}_B | I^*, Z_S)}{f(\underset{\sim}{Y}_A, \underset{\sim}{Y}_B | I^*, Z_S)}$$

where $f(\underset{\sim}{h}, \underset{\sim}{Y}_A, \underset{\sim}{Y}_B | I^*, Z_S) = f(\underset{\sim}{Y}_A, \underset{\sim}{Y}_B | I^*, Z_S, h) f(\underset{\sim}{h} | Z_S, I^*)$. But, given II^*, $f(\underset{\sim}{Y}_A, \underset{\sim}{Y}_B | I^*, Z_S)$ is a function of known value, since Y_A and Y_B are known. Thus,

$$f(\underset{\sim}{h} | II^*) \propto f(\underset{\sim}{Y}_A, \underset{\sim}{Y}_B | I^*, Z_S, h) f(\underset{\sim}{h} | Z_S, I^*) . \qquad (3.2.3)$$

To find $f(\underset{\sim}{Y}_A, \underset{\sim}{Y}_B | I^*, Z_S, h)$, we argue conditionally on I^*, Z_S and h as follows: Given β (and I^*, Z_S and h),

$$\binom{Y_A}{Y_B} \sim N\left(\left[\begin{array}{c|c} X_A & 0 \\ \hline 0 & X_B \end{array}\right]\beta,\ \left[\begin{array}{c|c} h_1^{-1}I_A & 0 \\ \hline 0 & h_2^{-1}I_B \end{array}\right]\right),$$

and, given λ (and I^*, Z_S and h),

$$\beta \sim N\left(\lambda,\ \left[\begin{array}{c|c} h_1^{-1}\tau_1^{-1} & 0 \\ \hline 0 & h_2^{-1}\tau_2^{-1} \end{array}\right]\right).$$

Thus, by Lemma 2.2.1, the marginal (conditional) distribution of $\binom{Y_A}{Y_B}$ is normal with mean

$$\mu_{AB} = \left[\begin{array}{c|c} X_A & 0 \\ \hline 0 & X_B \end{array}\right]\lambda = \begin{bmatrix} X_A\lambda_1 \\ X_B\lambda_2 \end{bmatrix}$$

and with variance

$$V_{AB} = \left[\begin{array}{c|c} h_1^{-1}I_A & 0 \\ \hline 0 & h_2^{-1}I_B \end{array}\right] + \left[\begin{array}{c|c} X_A & 0 \\ \hline 0 & X_B \end{array}\right]\left[\begin{array}{c|c} h_1^{-1}\tau_1^{-1} & 0 \\ \hline 0 & h_2^{-1}\tau_2^{-1} \end{array}\right]\left[\begin{array}{c|c} X'_A & 0 \\ \hline 0 & X'_B \end{array}\right]$$

$$= \left[\begin{array}{c|c} h_1^{-1}(I_A+X_A\tau_1^{-1}X'_A) & 0 \\ \hline 0 & h_2^{-1}(I_B+X_B\tau_2^{-1}X'_B) \end{array}\right].$$

This means that

$$f(\underset{\sim}{Y}_A, \underset{\sim}{Y}_B \mid I^*, Z_S, h)$$

$$\propto |V_{AB}|^{-1/2}\exp\left\{-\frac{1}{2}\left(\binom{Y_A}{Y_B}-\mu_{AB}\right)' V_{AB}^{-1}\left(\binom{Y_A}{Y_B}-\mu_{AB}\right)\right\}.$$

Now h and Z_S are independent random vectors (conditional on I^*) so that $f(\underset{\sim}{h}|Z_S, I^*) = f(\underset{\sim}{h}|I^*)$. We thus have, from (3.2.3),

$$f(\underset{\sim}{h}|II^*) \propto |V_{AB}|^{-1/2} \exp\left\{-\frac{1}{2}\left(\left(\begin{matrix}Y_A\\Y_B\end{matrix}\right)-\mu_{AB}\right)' V_{AB}^{-1}\left(\left(\begin{matrix}Y_A\\Y_B\end{matrix}\right)-\mu_{AB}\right)\right\}$$

$$\times [h_1^{-e_1/2} \exp\{-\frac{e_1h_1}{2\eta_1}\}][h_2^{-e_2'/2} \exp\{-\frac{e_2h_2}{2\eta_2}\}]$$

$$= (|h_1^{-1}(I_A+X_A\tau_1^{-1}X_A')| \cdot |h_2^{-1}(I_B+X_B\tau_2^{-1}X_B')|)^{-1/2}$$

$$\times \exp\{-\frac{1}{2}[h_1(Y_A-X_A\lambda_1)'(I_A+X_A\tau_1^{-1}X_A')^{-1}(Y_A-X_A\lambda_1)]\}$$

$$\times \exp\{-\frac{1}{2}[h_2(Y_B-X_B\lambda_2)'(I_B+X_B\tau_2^{-1}X_B')^{-1}(Y_B-X_B\lambda_2)]\}$$

$$\times [h_1^{(e_1/2)-1} \exp\{-\frac{e_1h_1}{2\eta_1}\}][h_2^{(e_2/2)-1} \exp\{-\frac{e_2h_2}{2\eta_2}\}] .$$

Note that

$$|h_1^{-1}(I_A+X_A\tau_1^{-1}X_A')| = h_1^{-r}|I_A+X_A\tau_2^{-1}X_A'|$$

and that

$$|h_2^{-1}(I_B+X_B\tau_2^{-1}X_B')| = h_2^{-m}|I_B+X_B\tau_2^{-1}X_B'| .$$

Then

$$f(\underset{\sim}{h}|\text{II}^*) \propto h_1^{(r+e_1)/2-1} \exp\{-\frac{1}{2}[\frac{e_1}{\eta_1}+(Y_A-X_A\lambda_1)'(I_A+X_A\tau_1^{-1}X_A')^{-1}(Y_A-X_A\lambda_1)]h_1\}$$

$$\times h_2^{(m+e_2)/2-1} \exp\{-\frac{1}{2}[\frac{e_2}{\eta_2}+(Y_B-X_B\lambda_2)'(I_B+X_B\tau_2^{-1}X_B')^{-1}(Y_B-X_B\lambda_2)]h_2\}$$

$$= h_1^{(r+e_1)/2-1} \exp\{-\frac{1}{2}(\frac{e_1}{\eta_1})Q_1h_1\}$$

$$\times h_2^{(m+e_2)/2-1} \exp\{-\frac{1}{2}(\frac{e_2}{\eta_2})Q_2h_2\} \tag{3.2.4}$$

where

$$Q_1 := 1 + e_1^{-1}(Y_A-X_A\lambda_1)'(\eta_1)(I_A+X_A\tau_1^{-1}X_A')^{-1}(Y_A-X_A\lambda_1)$$

and

$$Q_2 := 1 + e_2^{-1}(Y_B-X_B\lambda_2)'(\eta_2)(I_B+X_B\tau_2^{-1}X_B')^{-1}(Y_B-X_B\lambda_2) .$$

We now use (2.2.12), (3.2.2) and (3.2.4) in (3.2.1) to obtain

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi}, \underset{\sim}{h}|\text{II}^*)$$

$$\propto (|V_{CD}|)^{-1/2}\exp\left\{-\frac{1}{2}[(\begin{matrix}Y_C\\Y_D\end{matrix})-\mu_{CD}]'V_{CD}^{-1}[(\begin{matrix}Y_C\\Y_D\end{matrix})-\mu_{CD}]\right\}$$

$$\times \prod_{ij\in D}[\pi_i^{Z_{ij}}(1-\pi_i)^{1-Z_{ij}}](|B_\beta|)^{-1/2}\exp\{-\frac{1}{2}(\beta-\hat{\beta})'B_\beta^{-1}(\beta-\hat{\beta})\}$$

$$\times \prod_{i=1}^{g}\left\{\frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')}\pi_i^{p_i'-1}(1-\pi_i)^{q_i'-1}\right\} h_1^{(r+e_1)/2-1}\exp\{-\frac{1}{2}(\frac{e_1}{\eta_1})Q_1h_1\}$$

$$\times h_2^{(m+e_2)/2-1}\exp\{-\frac{1}{2}(\frac{e_2}{\eta_2})Q_2h_2\} . \tag{3.2.5}$$

In (3.2.5) Z_D, β, π and h are nuisance parameters. We eliminate them as we did in Section 2.2 by finding the marginal density, $f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D | II^*)$.

Before doing so, it is interesting to note that

$$f(\underset{\sim}{\beta}, \underset{\sim}{h} | II^*) = f(\underset{\sim}{\beta} | h, II^*) f(\underset{\sim}{h} | II^*)$$

where

$$f(\underset{\sim}{\beta} | h, II^*) = f(\underset{\sim}{\beta} | II, \eta_1, e_1, \eta_2, e_2) = f(\underset{\sim}{\beta} | II) .$$

Thus, using (2.2.9) and (3.2.4), we have

$$\begin{aligned} f(\underset{\sim}{\beta}, \underset{\sim}{h} | II^*) \propto (|B_\beta|)^{-1/2} &\exp\{-\tfrac{1}{2}(\beta-\hat{\beta})'B_\beta^{-1}(\beta-\hat{\beta})\} \\ \times\, h_1^{(r+e_1)/2-1} &\exp\{-\tfrac{1}{2}(\tfrac{e_1}{\eta_1})Q_1 h_1\} \\ \times\, h_2^{(m+e_2)/2-1} &\exp\{-\tfrac{1}{2}(\tfrac{e_2}{\eta_2})Q_2 h_2\} \end{aligned} \tag{3.2.6}$$

where

$$\hat{\beta} := \begin{Bmatrix} (X_A'X_A+\tau_1)^{-1}(X_A'Y_A+\tau_1\lambda_1) \\ (X_B'X_B+\tau_2)^{-1}(X_B'Y_B+\tau_2\lambda_2) \end{Bmatrix}.$$

Note that

$$|B_\beta| = h_1^{-x}|(X_A'X_A+\tau_1)^{-1}| \cdot h_2^{-x}|(X_B'X_B+\tau_2)^{-1}| \propto h_1^{-x}h_2^{-x}$$

for II^* known. Also note that

$$(\beta-\hat{\beta})'B_{\beta}^{-1}(\beta-\hat{\beta}) = h_1B_1 + h_2B_2$$

where

$$B_1 := (\beta_1-\hat{\beta}_1)'(X_A'X_A+\tau_1)(\beta_1-\hat{\beta}_1)$$

and where

$$B_2 := (\beta_2-\hat{\beta}_2)'(X_B'X_B+\tau_2)(\beta_2-\hat{\beta}_2) .$$

Thus, (3.2.6) becomes

$$\begin{aligned} f(\underset{\sim}{\beta}, \underset{\sim}{h} \mid \Pi^*) &\propto h_1^{x/2}h_2^{x/2}\exp\{-\tfrac{1}{2}B_1h_1\}\exp\{-\tfrac{1}{2}B_2h_2\} \\ &\quad \times h_1^{(r+e_1)/2-1}\exp\{-\tfrac{1}{2}(\frac{e_1}{\eta_1})Q_1h_1\} \\ &\quad \times h_2^{(m+e_2)/2-1}\exp\{-\tfrac{1}{2}(\frac{e_2}{\eta_2})Q_2h_2\} \\ &= h_1^{\frac{1}{2}(x+r+e_1)-1}\exp\{-\tfrac{1}{2}[\frac{e_1}{\eta_1}Q_1+B_1]h_1\} \\ &\quad \times h_2^{\frac{1}{2}(x+m+e_2)-1}\exp\{-\tfrac{1}{2}[\frac{e_2}{\eta_2}Q_2+B_2]h_2\} . \end{aligned} \tag{3.2.7}$$

We can now find the posterior density of β in this revised model. Using (3.2.7) we obtain

$$f(\underset{\sim}{\beta}|\mathrm{II}^*) = \int_0^\infty \int_0^\infty f(\underset{\sim}{\beta}, \underset{\sim}{h}|\mathrm{II}^*)dh_1dh_2$$

$$\propto [\frac{e_1}{\eta_1} Q_1+B_1]^{-\frac{1}{2}(x+r+e_1)} \Gamma(\frac{1}{2}(x+r+e_1))$$

$$\times [\frac{e_2}{\eta_2} Q_2+B_2]^{-\frac{1}{2}(x+m+e_2)} \Gamma(\frac{1}{2}(x+m+e_2))$$

$$\propto [\frac{e_1}{\eta_1} Q_1+B_1]^{-\frac{1}{2}(x+r+e_1)} \cdot [\frac{e_2}{\eta_2} Q_2+B_2]^{-\frac{1}{2}(x+m+e_2)} . \tag{3.2.8}$$

But, if we define

$$B_1^* := [(\frac{e_1}{r+e_1})(\frac{Q_1}{\eta_1})]^{-1} B_1 \quad \text{and} \quad B_2^* := [(\frac{e_2}{m+e_2})(\frac{Q_2}{\eta_2})]^{-1} B_2 ,$$

then

$$[\frac{e_1}{\eta_1} Q_1+B_1]^{-\frac{1}{2}(x+r+e_1)} = \underbrace{(\frac{e_1}{\eta_1} Q_1)^{-\frac{1}{2}(x+r+e_1)}}_{\text{known scalar}} [1+(\frac{e_1}{\eta_1} Q_1)^{-1} B_1]^{-\frac{1}{2}(x+r+e_1)}$$

$$\propto [1+(\frac{e_1}{\eta_1} Q_1)^{-1} B_1]^{-\frac{1}{2}(x+r+e_1)}$$

$$= [1+(r+e_1)^{-1} B_1^*]^{-\frac{1}{2}(x+r+e_1)}$$

and, similarly,

$$[\frac{e_2}{\eta_2} Q_2+B_2]^{-\frac{1}{2}(x+m+e_2)} \propto [1+(m+e_2)^{-1} B_2^*]^{-\frac{1}{2}(x+m+e_2)} .$$

This means that (3.2.8) may be re-written as

$$f(\underset{\sim}{\beta}|\Pi^*) \propto [1+(r+e_1)^{-1}B_1^*]^{-\frac{1}{2}(x+r+e_1)} \cdot [1+(m+e_2)^{-1}B_2^*]^{-\frac{1}{2}(x+m+e_2)} .$$

Thus, it is clear that $f(\underset{\sim}{\beta}|\Pi^*) = f(\underset{\sim}{\beta}_1|\Pi^*)f(\underset{\sim}{\beta}_2|\Pi^*)$ where $f(\underset{\sim}{\beta}_1|\Pi^*)$ and $f(\underset{\sim}{\beta}_2|\Pi^*)$ are multivariate-t densities having means and covariance matrices

$$(\hat{\beta}_1, (\frac{r+e_1}{r+e_1-2})(\frac{e_1}{r+e_1})(\frac{Q_1}{\eta_1})(X_A'X_A+\tau_1)^{-1})$$

and

$$(\hat{\beta}_2, (\frac{m+e_2}{m+e_2-2})(\frac{e_2}{m+e_2})(\frac{Q_2}{\eta_2})(X_B'X_B+\tau_2)^{-1})$$

respectively. These posterior means may be interpreted as Bayes estimates of β_1 and β_2 under squared error loss. Note that they are the same estimates obtained in the model of Chapter II.

We now return to the task of eliminating nuisance parameters from (3.2.5). Recall that $|B_\beta| \propto h_1^{-x}h_2^{-x}$ and that $(\beta-\hat{\beta})'B_\beta(\beta-\hat{\beta}) = h_1B_1 + h_2B_2$. Also note that

$$|V_{CD}| = |h_2^{-1}I_C| \cdot |V_D|$$

$$= h_2^{-(n-m-r)} h_1^{-\Sigma_{ij\epsilon D} Z_{ij}} h_2^{-(N-n-\Sigma_{ij\epsilon D} Z_{ij})}$$

and that, for $\mu_{ij} := Z_{ij}X_i'\beta_1 + (1-Z_{ij})X_i'\beta_2 \quad (ij \in U)$,

$$\left(\begin{pmatrix} Y_C \\ Y_D \end{pmatrix} - \mu_{CD}\right)' V_{CD}^{-1} \left(\begin{pmatrix} Y_C \\ Y_D \end{pmatrix} - \mu_{CD}\right)$$

$$= h_2(Y_C - X_C\beta_2)'(Y_C - X_C\beta_2) + \sum_{ij \in D} h_1^{Z_{ij}} h_2^{(1-Z_{ij})} (Y_{ij} - \mu_{ij})^2$$

$$= h_2 \sum_{ij \in C} (Y_{ij} - X_i'\beta_2)^2 + \sum_{ij \in D} [Z_{ij}h_1 + (1-Z_{ij})h_2](Y_{ij} - \mu_{ij})^2$$

$$= h_1 \sum_{ij \in D} Z_{ij}(Y_{ij} - \mu_{ij})^2 + h_2 \left\{ \sum_{ij \in C} (Y_{ij} - X_i'\beta_2)^2 + \sum_{ij \in D} (1-Z_{ij})(Y_{ij} - \mu_{ij})^2 \right\}.$$

We can rewrite (3.2.5) as

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi}, \underset{\sim}{h} \,|\, II^*)$$

$$\propto h_1^{\frac{1}{2}\Sigma_{ij \in D} Z_{ij}} \; h_2^{\frac{1}{2}((N-n)+(n-m-r)-\Sigma_{ij \in D} Z_{ij})}$$

$$\times \exp\left\{ -\frac{1}{2} \left[\sum_{ij \in D} Z_{ij}(Y_{ij} - \mu_{ij})^2 \right] h_1 \right\}$$

$$\times \exp\left\{ -\frac{1}{2} \left[\sum_{ij \in C} (Y_{ij} - X_i'\beta_2)^2 + \sum_{ij \in D} (1-Z_{ij})(Y_{ij} - \mu_{ij})^2 \right] h_2 \right\}$$

$$\times \left\{ \prod_{ij \in D} [\pi_i^{Z_{ij}} (1-\pi_i)^{1-Z_{ij}}] \right\} \times h_1^{x/2} \exp\{-\tfrac{1}{2} B_1 h_1\}$$

$$\times h_2^{x/2} \exp\{-\tfrac{1}{2} B_2 h_2\} \times \prod_{i=1}^{g} \left\{ \frac{\Gamma(p_i' + q_i')}{\Gamma(p_i')\Gamma(q_i')} \pi_i^{p_i'-1} (1-\pi_i)^{q_i'-1} \right\} \quad \times$$

$$\times\ [h_1^{(r+e_1)/2-1} \exp\{-\tfrac{1}{2}(\tfrac{e_1}{\eta_1})Q_1 h_1\}][h_2^{(m+e_2)/2-1} \exp\{-\tfrac{1}{2}(\tfrac{e_2}{\eta_2})Q_2 h_2\}]\ . \tag{3.2.9}$$

We will first eliminate Z_D. We will then eliminate β, h and finally π. The following lemma will be helpful.

Lemma 3.2.10. Let $\theta(Z_D)$ be a real-valued function of Z_D, the subvector of Z corresponding to units in D. Then

$$\sum_{\substack{Z_{ij} \in \{0,1\} \\ ij \in D}} \theta(Z_D) = \sum_{\omega \in W} \theta(Z_\omega = 1_\omega, Z_{\omega*} = 0_{\omega*})$$

where $W := 2^D$, the set of all subsets of D,

$\omega* := D \setminus \omega$, $n_\omega :=$ number of elements in ω,

$n_{\omega*} :=$ number of elements in $\omega* = N-n-n_\omega$,

$Z_\omega := (n_\omega \times 1)$ subvector of Z_D corresponding to units in ω,

$Z_{\omega*} := (n_{\omega*} \times 1)$ subvector of Z_D corresponding to units in $\omega*$,

$1_\omega := (n_\omega \times 1)$ vector of ones and where

$0_{\omega*} := (n_{\omega*} \times 1)$ vector of zeros.

The notation $\theta(Z_\omega = 1_\omega,\ Z_{\omega*} = 0_{\omega*})$ means that the function $\theta(\cdot)$ is evaluated for $Z_{ij} = 1$, $ij \in \omega$, and $Z_{ij} = 0$, $ij \in \omega*$.

Proof. Note that W has 2^{N-n} members so that the sum on the left hand side (l. h. s.) and the sum on the right hand side (r. h. s.) both have 2^{N-n} terms. Consider the sum on the l. h. s. In each term, Z_{ij} is either 0 or 1 for every $ij \in D$. In fact, the 2^{N-n} terms in the sum have a 1-1 correspondence with the 2^{N-n} possible sequences, $(Z_{ij} : ij \in D)$, where Z_{ij} is either 0 or 1 for $ij \in D$. For any one of these sequences, let ω be the set of subscripts corresponding to those Z_{ij}'s taking the value 1. Then ω^*, as defined in the lemma, consists of those subscripts for which the corresponding Z_{ij}'s take the value 0. Now ω is clearly a subset of D. Further, there is a unique ω for each possible sequence, and the collection of these ω's is obviously W. But this establishes a 1-1 correspondence between terms in the sum on the l. h. s. and terms in the sum on the r. h. s. The lemma follows immediately. $\square$

We have, by Lemma 3.2.10,

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi}, \underset{\sim}{h} | \Pi^*)$$

$$= \sum_{\substack{Z_{ij} \in \{0, 1\} \\ ij \in D}} f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{Z}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi}, \underset{\sim}{h} | \Pi^*) \quad \propto$$

$$\propto \sum_{\omega \in W} \Bigg\{ h_1^{\frac{1}{2} n_\omega} \, h_2^{\frac{1}{2}(n_{\omega*}+n-m-r)}$$

$$\times \exp\left\{ -\frac{1}{2}\left[\sum_{ij \in \omega} (Y_{ij} - X_i'\beta_1)^2 \right] h_1 \right\}$$

$$\times \exp\left\{ -\frac{1}{2}\left[\sum_{ij \in C \cup \omega*} (Y_{ij} - X_i'\beta_2)^2 \right] h_2 \right\}$$

$$\times \left\{ \prod_{i=1}^{g} [\pi_i^{\omega_i}(1-\pi_i)^{\omega_i^*}] \right\} \cdot h_1^{x/2} \exp\{-\tfrac{1}{2} B_1 h_1\} \cdot h_2^{x/2} \exp\{-\tfrac{1}{2} B_2 h_2\}$$

$$\times \prod_{i=1}^{g} \left\{ \frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \pi_i^{p_i'-1} (1-\pi_i)^{q_i'-1} \right\}$$

$$\times [h_1^{(r+e_1)/2-1} \exp\{-\tfrac{1}{2}(\tfrac{e_1}{\eta_1}) Q_1 h_1\}][h_2^{(m+e_2)/2-1} \exp\{-\tfrac{1}{2}(\tfrac{e_2}{\eta_2}) Q_2 h_2\}] \Bigg\}$$

(3.2.11)

where ω_i := number of elements from group i in ω and ω_i^* := number of elements from group i in ω^*, $i = 1, 2, \ldots, g$.

We next integrate (3.2.11) with respect to β and obtain

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{\pi}, \underset{\sim}{h} \mid \Pi^*)$$

$$= \int_{\mathbb{R}^x} \int_{\mathbb{R}^x} f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{\beta}, \underset{\sim}{\pi}, \underset{\sim}{h} \mid \Pi^*) d\beta_1 d\beta_2$$

$$\propto \sum_{\omega \in W} \Bigg\{ \{ \prod_{i=1}^{g} [\pi_i^{\omega_i}(1-\pi_i)^{\omega_i^*}]\} \times [h_1^{(r+e_1)/2-1} \exp\{-\tfrac{1}{2}(\tfrac{e_1}{\eta_1}) Q_1 h_1\}] \quad \times$$

$$\times \left[h_2^{(m+e_2)/2-1} \exp\{-\tfrac{1}{2}(\tfrac{e_2}{\eta_2})Q_2h_2\}\right]$$

$$\times \prod_{i=1}^{g}\left\{\frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')}\pi_i^{p_i'-1}(1-\pi_i)^{q_i'-1}\right\}\times \mathrm{Int}(\omega)\Bigg\} \qquad (3.2.12)$$

where

$$\mathrm{Int}(\omega) := \int_{\mathbb{R}^x}\int_{\mathbb{R}^x}\Bigg\{ h_1^{\frac{1}{2}n_\omega} h_2^{\frac{1}{2}(n_{\omega*}+n-m-r)}$$

$$\times \exp\left\{-\frac{1}{2}\left[\sum_{ij\in\omega}(Y_{ij}-X_i'\beta_1)^2\right]h_1\right\}$$

$$\times \exp\left\{-\frac{1}{2}\left[\sum_{ij\in C\cup\omega*}(Y_{ij}-X_i'\beta_2)^2\right]h_2\right\}$$

$$\times h_1^{x/2}\exp\{-\tfrac{1}{2}B_1h_1\}h_2^{x/2}\exp\{-\tfrac{1}{2}B_2h_2\}\Bigg\}d\beta_1 d\beta_2 .$$

Recall that $B_1 := (\beta_1-\hat{\beta}_1)'(X_A'X_A+\tau_1)(\beta_1-\hat{\beta}_1)$ and $B_2 := (\beta_2-\hat{\beta}_2)'(X_B'X_B+\tau_2)(\beta_2-\hat{\beta}_2)$.

To find $\mathrm{Int}(\omega)$ for any $\omega \in W$, we will use Lemma 2.2.1 and argue conditionally on h_1, h_2, ω and Π^*. Let X_ω and $X_{\omega*}$ be submatrices of X corresponding to units in ω and $\omega*$ respectively. Let I_ω and $I_{\omega*}$ be identity matrices of sizes n_ω and $n_{\omega*}$ respectively, and let Y_ω and $Y_{\omega*}$ be those subvectors of Y_D corresponding to units in ω and $\omega*$ respectively. Given β,

$$\begin{bmatrix} Y_\omega \\ Y_C \\ Y_{\omega*} \end{bmatrix} \sim N\left(\left[\begin{array}{c|c} X_\omega & 0 \\ \hline 0 & X_C \\ \hline 0 & X_{\omega*} \end{array}\right]\beta,\ \left[\begin{array}{c|c|c} h_1^{-1} I_\omega & 0 & 0 \\ \hline 0 & h_2^{-1} I_C & 0 \\ \hline 0 & 0 & h_2^{-1} I_{\omega*} \end{array}\right] \right)$$

while, given II,

$$\beta = \binom{\beta_1}{\beta_2} \sim N\left(\hat{\beta},\ \left[\begin{array}{c|c} h_1^{-1}(X_A'X_A+\tau_1)^{-1} & 0 \\ \hline 0 & h_2^{-1}(X_B'X_B+\tau_2)^{-1} \end{array}\right]\right).$$

Lemma 2.2.1 shows that the marginal distribution of

$$\begin{bmatrix} Y_\omega \\ Y_C \\ Y_{\omega*} \end{bmatrix} \quad \text{is normal with mean} \quad \begin{bmatrix} X_\omega \hat{\beta}_1 \\ X_C \hat{\beta}_2 \\ X_{\omega*} \hat{\beta}_2 \end{bmatrix}$$

and with covariance matrix given by

$$\left[\begin{array}{c|c|c} h_1^{-1} I_\omega & 0 & 0 \\ \hline 0 & h_2^{-1} I_C & 0 \\ \hline 0 & 0 & h_2^{-1} I_{\omega*} \end{array}\right] + \left[\begin{array}{c|c} X_\omega & 0 \\ \hline 0 & X_C \\ \hline 0 & X_{\omega*} \end{array}\right] \left[\begin{array}{c|c} h_1^{-1}(X_A'X_A+\tau_1)^{-1} & 0 \\ \hline 0 & h_2^{-1}(X_B'X_B+\tau_2)^{-1} \end{array}\right]$$

$$\times \left[\begin{array}{c|c} X_\omega & 0 \\ \hline 0 & X_C \\ \hline 0 & X_{\omega*} \end{array}\right]' = \left[\begin{array}{c|c} h_1^{-1} H_\omega^{-1} & 0 \\ \hline 0 & h_2^{-1} H_{\omega*}^{-1} \end{array}\right]$$

where $H_\omega^{-1} := I_\omega + X_\omega(X_A'X_A+\tau_1)^{-1}X_\omega'$ and

$$H_{\omega*}^{-1} := \left[\begin{array}{c|c} I_C + X_C(X_B'X_B+\tau_2)^{-1}X_C' & X_C(X_B'X_B+\tau_2)^{-1}X_{\omega*}' \\ \hline X_{\omega*}(X_B'X_B+\tau_2)^{-1}X_C' & I_{\omega*}+X_{\omega*}(X_B'X_B+\tau_2)^{-1}X_{\omega*}' \end{array}\right]$$

This means that

$$\begin{aligned} \mathrm{Int}(\omega) \propto\ & (|h_1^{-1}H_\omega^{-1}|)^{-1/2}(|h_2^{-1}H_{\omega*}|)^{-1/2} \\ & \times \exp\left\{-\frac{1}{2}\left[(Y_\omega - X_\omega\hat{\beta}_1)', \left(\begin{pmatrix} Y_C \\ Y_{\omega*}\end{pmatrix} - \begin{pmatrix} X_C \\ X_{\omega*}\end{pmatrix}\hat{\beta}_2\right)'\right]\right. \\ & \left. \times \left[\begin{array}{c|c} h_1 H_\omega & 0 \\ \hline 0 & h_2 H_{\omega*}\end{array}\right] \begin{bmatrix} Y_\omega - X_\omega\hat{\beta}_1 \\ Y_C - X_C\hat{\beta}_2 \\ Y_{\omega*} - X_{\omega*}\hat{\beta}_2 \end{bmatrix}\right\} \\ =\ & h_1^{\frac{1}{2}n_\omega} |H_\omega|^{1/2} \exp\{-\frac{h_1}{2} Q_\omega^*\} \\ & \times h_2^{\frac{1}{2}(n-m-r+n_{\omega*})} |H_{\omega*}|^{1/2} \exp\{-\frac{h_2}{2} Q_{\omega*}^*\} \end{aligned}$$

where

$$Q_\omega^* := (Y_\omega - X_\omega\hat{\beta}_1)'H_\omega(Y_\omega - X_\omega\hat{\beta}_1)$$

and

$$Q_{\omega*}^* := \left[\begin{pmatrix} Y_C \\ Y_{\omega*}\end{pmatrix} - \begin{pmatrix} X_C \\ X_{\omega*}\end{pmatrix}\hat{\beta}_2\right]' H_{\omega*} \left[\begin{pmatrix} Y_C \\ Y_{\omega*}\end{pmatrix} - \begin{pmatrix} X_C \\ X_{\omega*}\end{pmatrix}\hat{\beta}_2\right].$$

The constant of proportionality is clearly

$$(2\pi)^{-\frac{1}{2}n_\omega}(2\pi)^{-\frac{1}{2}(n-m-r+n_{\omega*})}.$$

Thus,

$$\mathrm{Int}(\omega) = (2\pi)^{-\frac{1}{2}(N-n+n-m-r)} h_1^{\frac{1}{2}n_\omega} |H_\omega|^{1/2} \exp\{-\frac{h_1}{2} Q_\omega^*\}$$

$$\times\ h_2^{\frac{1}{2}(n-m-r+n_{\omega*})} |H_{\omega*}|^{1/2} \exp\{-\frac{h_2}{2} Q_{\omega*}^*\}\ .$$

Expression (3.2.12) can now be written as

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{\pi}, \underset{\sim}{h} | II^*)$$

$$\begin{aligned}\propto \sum_{\omega \in W} \Bigg\{ & \prod_{i=1}^{g} \left\{ \frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \pi_i^{(p_i'+\omega_i)-1} (1-\pi_i)^{(q_i'+\omega_i^*)-1} \right\} \\ & \times \left[(2\pi)^{-\frac{1}{2}(N-n+n-m-r)} |H_\omega|^{1/2} |H_{\omega*}|^{1/2} \right] \\ & \times \left[h_1^{\frac{1}{2}(n_\omega+r+e_1)-1} \exp\{-\frac{1}{2}[(\frac{e_1}{\eta_1})Q_1+Q_\omega^*]h_1\} \right] \\ & \times \left[h_2^{\frac{1}{2}(n-m-r+n_{\omega*}+m+e_2)-1} \exp\{-\frac{1}{2}[(\frac{e_2}{\eta_2})Q_2+Q_{\omega*}^*]h_2\} \right] \Bigg\}\ . \end{aligned} \tag{3.2.13}$$

Next we integrate (3.2.13) with respect to h_1 and h_2 and obtain

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{\pi} | II^*)$$

$$\begin{aligned}\propto \sum_{\omega \in W} \Bigg\{ & \prod_{i=1}^{g} \left\{ \frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \pi_i^{(p_i'+\omega_i)-1} (1-\pi_i)^{(q_i'+\omega_i^*)-1} \right\} \\ & \times (2\pi)^{-\frac{1}{2}(N-n+n-m-r)} |H_\omega|^{1/2} |H_{\omega*}|^{1/2} (I_{h_1})(I_{h_2}) \Bigg\} \end{aligned} \tag{3.2.14}$$

where

$$I_{h_1} := \int_0^\infty h_1^{\frac{1}{2}(n_\omega+r+e_1)-1} \exp\{-\frac{1}{2}[(\frac{e_1}{\eta_1})Q_1+Q_\omega^*]h_1\}d_{h_1}$$

$$= [\frac{1}{2}(\frac{e_1}{\eta_1}Q_1+Q_\omega^*)]^{-\frac{1}{2}(n_\omega+r+e_1)}\Gamma(\frac{1}{2}(n_\omega+r+e_1))$$

and

$$I_{h_2} := \int_0^\infty h_1^{\frac{1}{2}(n_{\omega*}+n-m-r+m+e_2)-1} \exp\{-\frac{1}{2}[(\frac{e_2}{\eta_2}Q_2+Q_{\omega*}^*]h_2\}d_{h_2}$$

$$= [\frac{1}{2}(\frac{e_2}{\eta_2}Q_2+Q_{\omega*}^*)]^{-\frac{1}{2}(n_{\omega*}+n-m-r+m+e_2)}$$

$$\times\ \Gamma(\frac{1}{2}(n_{\omega*}+n-m-r+m+e_2))\ .$$

If we define

$$Q_\omega := (Y_\omega-X_\omega\hat{\beta}_1)'[(\frac{e_1}{r+e_1})(\frac{Q_1}{\eta_1})H_\omega^{-1}]^{-1}(Y_\omega-X_\omega\hat{\beta}_1)$$

and

$$Q_{\omega*} := [\binom{Y_C}{Y_{\omega*}}-\binom{X_C}{X_{\omega*}}\hat{\beta}_2]'[(\frac{e_2}{m+e_2})(\frac{Q_2}{\eta_2})H_{\omega*}^{-1}]^{-1}[\binom{Y_C}{Y_{\omega*}}-\binom{X_C}{X_{\omega*}}\hat{\beta}_2]\ ,$$

then

$$I_{h_1} = (\frac{1}{2}\frac{e_1}{\eta_1}Q_1)^{-\frac{1}{2}(n_\omega+r+e_1)}[1+(r+e_1)^{-1}Q_\omega]^{-\frac{1}{2}(n_\omega+r+e_1)}$$

$$\times\ \Gamma(\frac{1}{2}(n_\omega+r+e_1))$$

and

$$I_{h_2} = \left(\frac{1}{2} \frac{e_2}{\eta_2} Q_2 \right)^{-\frac{1}{2}(n_{\omega*}+n-m-r+m+e_2)}$$

$$\times \left[1+(m+e_2)^{-1} Q_{\omega*} \right]^{-\frac{1}{2}(n_{\omega*}+n-m-r+m+e_2)}$$

$$\times \Gamma\left(\frac{1}{2} (n_{\omega*}+n-m-r+m+e_2) \right) .$$

It is now clear that (3.2.14) can be written as

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{\pi} | II^*)$$

$$\propto \sum_{\omega \in W} \left\{ \prod_{i=1}^{g} \left\{ \frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \pi_i^{(p_i'+\omega_i)-1} (1-\pi_i)^{(q_i'+\omega_i^*)-1} \right\} \right.$$

$$\times \pi^{-\frac{1}{2}n_\omega} e_1^{-\frac{1}{2}n_\omega} [(\frac{Q_1}{\eta_1})^{n_\omega} \cdot |H_\omega^{-1}|]^{-1/2} (\frac{1}{2} \frac{e_1}{\eta_1} Q_1)^{-\frac{1}{2}(r+e_1)}$$

$$\times \left[\pi^{-\frac{1}{2}(n-m-r+n_{\omega*})} e_2^{-\frac{1}{2}(n-m-r+n_{\omega*})} \right.$$

$$\left. \times [(\frac{Q_2}{\eta_2})^{n-m-r+n_{\omega*}} \cdot |H_{\omega*}^{-1}|]^{-1/2} (\frac{1}{2} \frac{e_2}{\eta_2} Q_2)^{-\frac{1}{2}(m+e_2)} \right]$$

$$\times \left[1+(r+e_1)^{-1} Q_\omega \right]^{-\frac{1}{2}(n_\omega+r+e_1)} \Gamma\left(\frac{1}{2} (n_\omega+r+e_1) \right)$$

$$\times \left[1+(m+e_2)^{-1} Q_{\omega*} \right]^{-\frac{1}{2}(n-m-r+n_{\omega*}+m+e_2)}$$

$$\left. \times \Gamma\left(\frac{1}{2} (n-m-r+n_{\omega*}+m+e_2) \right) \right\} \propto$$

$$\propto \sum_{\omega \in W} \left\{ \prod_{i=1}^{g} \left\{ \frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \pi_i^{(p_i'+\omega_i)-1} (1-\pi_i)^{(q_i'+\omega_i^*)-1} \right\} \right.$$

$$\times \left[(\pi e_1)^{-\frac{1}{2}n_\omega} \left| \frac{Q_1}{\eta_1} H_\omega^{-1} \right|^{-1/2} \Gamma\left(\frac{1}{2}(n_\omega+r+e_1)\right) \right.$$

$$\left. \times \left[1+(r+e_1)^{-1} Q_\omega\right]^{-\frac{1}{2}(n_\omega+r+e_1)} \right]$$

$$\times \left[(\pi e_2)^{-\frac{1}{2}(n-m-r+n_{\omega*})} \left| \frac{Q_2}{\eta_2} H_{\omega*}^{-1} \right|^{-1/2} \Gamma\left(\frac{1}{2}(n-m-r+n_{\omega*}+m+e_2)\right) \right]$$

$$\left. \times \left[1+(m+e_2)^{-1} Q_{\omega*}\right]^{-\frac{1}{2}(n-m-r+n_{\omega*}+m+e_2)} \right\}. \qquad (3.2.15)$$

Let $\nu_1 := r + e_1$, $\nu_2 := m + e_2$, $V_\omega := \frac{e_1}{\nu_1} \frac{Q_1}{\eta_1} H_\omega^{-1}$ and $V_{\omega*} := \frac{e_2}{\nu_2} \frac{Q_2}{\eta_2} H_{\omega*}^{-1}$. Then we can write (3.2.15) as

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D, \underset{\sim}{\pi} | \Pi^*)$$

$$\propto \sum_{\omega \in W} \left\{ \prod_{i=1}^{g} \left\{ \frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \pi_i^{(p_i'+\omega_i)-1} (1-\pi_i)^{(q_i'+\omega_i^*)-1} \right\} \right.$$

$$\left. \times f_\omega(\underset{\sim}{Y}_\omega) f_{\omega*}(\underset{\sim}{Y}_C, \underset{\sim}{Y}_{\omega*}) \right\} \qquad (3.2.16)$$

where $f_\omega(\underset{\sim}{Y}_\omega)$ is a $t(n_\omega, \nu_1, X_\omega \hat{\beta}_1, V_\omega^{-1})$ density and $f_{\omega*}(\underset{\sim}{Y}_C, \underset{\sim}{Y}_{\omega*})$ is a $t(n-m-r+n_{\omega*}, \nu_2, \binom{X_C}{X_{\omega*}} \hat{\beta}_2, V_{\omega*}^{-1})$ density. (See A.3.)

Finally, let us find $f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D | \Pi^*)$ by integrating (3.2.16) with respect to the vector π and finding the proportionality constant. Recall that

$$\int_0^\infty t^{\alpha_1 - 1} (1-t)^{\alpha_2 - 1} \, dt = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)}$$

for $\alpha_1, \alpha_2 > 0$. Also note that $f_\omega(\underset{\sim}{Y}_\omega)$ and $f_{\omega*}(\underset{\sim}{Y}_C, \underset{\sim}{Y}_{\omega*})$ do not involve the vector π. Thus,

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D | \Pi^*) \propto \sum_{\omega \in W} \left\{ \prod_{i=1}^{g} \left\{ \frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \cdot \frac{\Gamma(p_i'+\omega_i)\Gamma(q_i'+\omega_i^*)}{\Gamma(p_i'+q_i'+d_i)} \right\} \right.$$
$$\left. \times f_\omega(\underset{\sim}{Y}_\omega) f_{\omega*}(\underset{\sim}{Y}_C, \underset{\sim}{Y}_{\omega*}) \right\}. \qquad (3.2.17)$$

The constant of proportionality is one, since

$$\sum_{\omega \in W} \left\{ \prod_{i=1}^{g} \left\{ \frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \frac{\Gamma(p_i'+\omega_i)\Gamma(q_i'+\omega_i^*)}{\Gamma(p_i'+q_i'+d_i)} \right\} \right.$$
$$\left. \times \int_{\mathbb{R}^{n_\omega}} f_\omega(\underset{\sim}{Y}_\omega) dY_\omega \int_{\mathbb{R}^{n-m-r+n_{\omega*}}} f_{\omega*}(\underset{\sim}{Y}_C, \underset{\sim}{Y}_{\omega*}) dY_C \, dY_{\omega*} \right\}$$
$$= \sum_{\omega \in W} \left\{ \prod_{i=1}^{g} \left\{ \frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \frac{\Gamma(p_i'+\omega_i)\Gamma(q_i'+\omega_i^*)}{\Gamma(p_i'+q_i'+d_i)} \right\} \right\}$$
$$= 1$$

by Lemma (3.2.19), below.

The proportionality sign in (3.2.17) can thus be replaced by a sign of equality so that

$$f(\underset{\sim}{Y}_C, \underset{\sim}{Y}_D \mid \Pi^*) = \sum \left\{ \prod_{i=1}^{g} \left\{ \frac{\Gamma(p_i'+q_i')}{\Gamma(p_i')\Gamma(q_i')} \; \frac{\Gamma(p_i'+\omega_i)\Gamma(q_i'+\omega_i^*)}{\Gamma(p_i'+q_i'+d_i)} \right\} \right.$$

$$\left. \times \; f_\omega(\underset{\sim}{Y}_\omega) f_{\omega*}(\underset{\sim}{Y}_C, \underset{\sim}{Y}_{\omega*}) \right\} \quad . \qquad (3.2.18)$$

Lemma 3.2.19. Let $\rho_1, \rho_2, \ldots, \rho_g, \; \sigma_1, \sigma_2, \ldots, \sigma_g$ be any positive constants. Then, using the notation of Lemma 3.2.10,

$$\sum_{\omega \in W} \left\{ \prod_{i=1}^{g} \frac{\Gamma(\rho_i+\sigma_i)}{\Gamma(\rho_i)\Gamma(\sigma_i)} \; \frac{\Gamma(\rho_i+\omega_i)\Gamma(\sigma_i+\omega_i^*)}{\Gamma(\rho_i+\sigma_i+d_i)} \right\} = 1 .$$

Proof. Using Lemma 3.2.10, we have

$$\sum_{\omega \in W} \left\{ \prod_{i=1}^{g} \frac{\Gamma(\rho_i+\sigma_i)}{\Gamma(\rho_i)\Gamma(\sigma_i)} \; \frac{\Gamma(\rho_i+\omega_i)\Gamma(\sigma_i+\omega_i^*)}{\Gamma(\rho_i+\sigma_i+d_i)} \right\}$$

$$= \sum_{\substack{Z_{k\ell} \in \{0,1\} \\ k\ell \in D}} \left\{ \prod_{i=1}^{g} \frac{\Gamma(\rho_i+\sigma_i)}{\Gamma(\rho_i)\Gamma(\sigma_i)} \; \frac{\Gamma(\rho_i+\Sigma_{j:ij\in D} Z_{ij})\Gamma(\sigma_i+d_i-\Sigma_{j:ij\in D} Z_{ij})}{\Gamma(\rho_i+\sigma_i+d_i)} \right\}$$

$$= \prod_{i=1}^{g} \left\{ \sum_{\substack{Z_k \in \{0,1\} \\ k \in \{1,2,\ldots,d_i\}}} \frac{\Gamma(\rho_i+\sigma_i)}{\Gamma(\rho_i)\Gamma(\sigma_i)} \; \frac{\Gamma(\rho_i+\Sigma_{k=1}^{d_i} Z_k)\Gamma(\sigma_i+d_i-\Sigma_{k=1}^{d_i} Z_k)}{\Gamma(\rho_i+\sigma_i+d_i)} \right\} =$$

$$= \prod_{i=1}^{g} \left\{ \sum_{\mu_i=0}^{d_i} \binom{d_i}{\mu_i} \frac{\Gamma(\rho_i+\sigma_i)}{\Gamma(\rho_i)\Gamma(\sigma_i)} \frac{\Gamma(\rho_i+\mu_i)\Gamma(\sigma_i+d_i-\mu_i)}{\Gamma(\rho_i+\sigma_i+d_i)} \right\}$$

$$= 1$$

by Lemma 3.2.20, following. $\square$

<u>Lemma 3.2.20</u>. For any $b, c > 0$ and any positive integer k,

$$\sum_{z=0}^{k} \binom{k}{z} \frac{\Gamma(b+c)}{\Gamma(b)\Gamma(c)} \frac{\Gamma(b+z)\Gamma(c+k-z)}{\Gamma(b+c+k)} = 1 .$$

<u>Proof</u>.

$$\sum_{z=0}^{k} \binom{k}{z} \frac{\Gamma(b+c)}{\Gamma(b)\Gamma(c)} \frac{\Gamma(b+z)\Gamma(c+k-z)}{\Gamma(b+c+k)}$$

$$= \sum_{z=0}^{k} \left[\binom{k}{z} \frac{\Gamma(b+c)}{\Gamma(b)\Gamma(c)} \int_0^1 t^{(b+z)-1}(1-t)^{(c+k-z)-1} dt \right]$$

$$= \frac{\Gamma(b+c)}{\Gamma(b)\Gamma(c)} \int_0^1 t^{b-1}(1-t)^{c-1} \left[\sum_{z=0}^{k} \binom{k}{z} t^z (1-t)^{k-z} \right] dt$$

$$= \frac{\Gamma(b+c)}{\Gamma(b)\Gamma(c)} \int_0^1 t^{b-1}(1-t)^{c-1} dt \qquad \text{by the Binomial Formula}$$

$$= 1 . \square$$

3.3. Posterior Prediction

Section 1.2 showed that the Bayes predictor of the population total, T, under squared error loss is

$$\hat{T} := \sum_{ij \in A \cup B} Y_{ij} + E(R|II^*)$$

where

$$R := \sum_{ij \in C \cup D} Y_{ij} .$$

In this section we will find $E(R|II^*)$ as well as $Var(R|II^*)$, the posterior variance of R.

From (2.3.11) we have

$$E(R|II) = \sum_{i=1}^{g} c_i X_i'\hat{\beta}_2 + \sum_{i=1}^{g} d_i[\hat{\pi}_i X_i'\hat{\beta}_1 + (1-\hat{\pi}_i)X_i'\hat{\beta}_2] .$$

Note that this expression is independent of $h = (h_1, h_2)'$. Also note that

$$E(R|h, II^*) = E(R|II, e_1, \eta_1, e_2, \eta_2) = E(R|II) ,$$

since the constants e_1, η_1 e_2 and η_2 do not appear in $E(R|II)$.

We now use (2.3.8) to see that

$$
\begin{aligned}
E(R|\Pi^*) &= E_h[E(R|h, \Pi^*)] \\
&= E_h[E(R|\Pi)] \\
&= E(R|\Pi) \\
&= \sum_{i=1}^{g} c_i X_i'\hat{\beta}_2 + \sum_{i=1}^{g} d_i[\hat{\pi}_i X_i'\hat{\beta}_1 + (1-\hat{\pi}_i)X_i'\hat{\beta}_2] \,, \qquad (3.3.1)
\end{aligned}
$$

since $E(R|\Pi)$ does not involve h_1 or h_2. By comparing (3.3.1) with (2.3.11) we see that, for predicting T under squared error loss, it makes no difference whether the precision vector, h, is fixed as in Chapter II or random as in this chapter.

Before finding $Var(R|\Pi^*)$, we note that, given Π^*, h_1 and h_2 are independent random variables having gamma distributions. In fact, using (3.2.4), we see that, given Π^*, h_1 and h_2 are independent random variables such that

$$
f(\underset{\sim}{h}_1|\Pi^*) = \frac{[\frac{1}{2}(\frac{e_1}{\eta_1})Q_1]^{\frac{1}{2}(r+e_1)}}{\Gamma[\frac{1}{2}(r+e_1)]} h_1^{\frac{1}{2}(r+e_1)-1} \exp\{-\frac{1}{2}(\frac{e_1}{\eta_1})Q_1 h_1\}
$$

and

$$
f(\underset{\sim}{h}_2|\Pi^*) = \frac{[\frac{1}{2}(\frac{e_2}{\eta_2})Q_2]^{\frac{1}{2}(m+e_2)}}{\Gamma[\frac{1}{2}(m+e_2)]} h_2^{\frac{1}{2}(m+e_2)-1} \exp\{-\frac{1}{2}(\frac{e_2}{\eta_2})Q_2 h_2\} .
$$

Thus,

$$E[h_1^{-1}|\Pi^*] = \int_0^\infty h_1^{-1} f(\underset{\sim}{h}_1|\Pi^*)dh_1$$

$$= \frac{[\frac{1}{2}(\frac{e_1}{\eta_1})Q_1]^{\frac{1}{2}(r+e_1)}}{\Gamma[\frac{1}{2}(r+e_1)]} \times \frac{\Gamma[\frac{1}{2}(r+e_1)-1]}{[\frac{1}{2}(\frac{e_1}{\eta_1})Q_1]^{\frac{1}{2}(r+e_1)-1}}$$

$$\times \int_0^\infty \frac{[\frac{1}{2}(\frac{e_1}{\eta_1})Q_1]^{\frac{1}{2}(r+e_1)-1}}{\Gamma[\frac{1}{2}(r+e_1)-1]} h_1^{[\frac{1}{2}(r+e_1)-2]} \exp\{-\frac{1}{2}(\frac{e_1}{\eta_1})Q_1 h_1\}dh_1$$

$$= \frac{[\frac{1}{2}(\frac{e_1}{\eta_1})Q_1]^{\frac{1}{2}(r+e_1)}\Gamma[\frac{1}{2}(r+e_1)-1]}{\Gamma[\frac{1}{2}(r+e_1)][\frac{1}{2}(\frac{e_1}{\eta_1})Q_1]^{\frac{1}{2}(r+e_1)-1}} .$$

Now, for any $a > 0$, $\Gamma(a+1) = a\Gamma(a)$. Thus,

$$\Gamma[\frac{1}{2}(r+e_1)] = \Gamma[\frac{1}{2}(r+e_1)-1+1] = [\frac{1}{2}(r+e_1)-1]\Gamma[\frac{1}{2}(r+e_1)-1]$$

provided $\frac{1}{2}(r+e_1) > 1$. In any practical situation, the number of respondents, r, is larger than 2 so that $\frac{1}{2}(r+e_1) > 1$. We then have

$$E[h_1^{-1}|II^*] = \frac{1}{2}\left(\frac{e_1}{\eta_1}\right)Q_1\left[\frac{1}{2}(r+e_1)-1\right]^{-1}$$

$$= \frac{e_1 Q_1}{\eta_1(r+e_1-2)} . \quad (3.3.2)$$

Similarly,

$$E[h_2^{-1}|II^*] = \frac{e_2 Q_2}{\eta_2(m+e_2-2)} . \quad (3.3.3)$$

The results of Chapter IV suggest that, for the purpose of interval estimation of R, the posterior density $f(\underset{\sim}{R}|II)$ can be approximated by a normal density having mean $E(R|II)$ and variance $Var(R|II)$ in the case where the precision vector, h, is fixed. Fact A.7 shows that, as $\nu_1 \to +\infty$ and $\nu_2 \to +\infty$, the densities $f_\omega(\underset{\sim}{Y}_\omega)$ and $f_{\omega*}(\underset{\sim}{Y}_C, \underset{\sim}{Y}_{\omega*})$, $\omega \in W$, may be approximated by multivariate normal densities. This in turn suggests that $f(\underset{\sim}{R}|II^*)$ may be approximated by a mixture of normal densities similar to $f(\underset{\sim}{R}|II)$. It is reasonable then to think that the distribution of R given II^* may, even for samples of moderate size, be well approximated by the normal distribution having mean $E(R|II^*)$ and variance $Var(R|II^*)$.

$E(R|II^*)$ is given by (3.3.1). To find $Var(R|II^*)$, we use (2.3.9) as follows:

$$Var(R|II^*) = E_h[Var(R|h, II^*)] + Var_h[E(R|h, II^*)] .$$

We have already shown that $E(R|h, \Pi^*) = E(R|\Pi)$ does not involve h. Thus $\mathrm{Var}_h[E(R|h, \Pi^*)] = 0$. Further, $\mathrm{Var}(R|h, \Pi^*) = \mathrm{Var}(R|\Pi, e_1, \eta_1, e_2, \eta_2) = \mathrm{Var}(R|\Pi)$, since the constants e_1, η_1 e_2 and η_2 do not appear in $\mathrm{Var}(R|\Pi)$. (See (2.3.19).) Thus,

$$\mathrm{Var}(R|\Pi^*) = E_h[\mathrm{Var}(R|\Pi)],$$

and $\mathrm{Var}(R|\Pi^*)$ is calculated by replacing h_1^{-1} by $\dfrac{e_1 Q_1}{\eta_1(r+e_1-2)}$ and h_2^{-1} by $\dfrac{e_2 Q_2}{\eta_2(m+e_2-2)}$ in (2.3.19). Note that

$$[E(h_1|\Pi^*)]^{-1} = \frac{e_1 Q_1}{\eta_1(r+e_1)}$$

and that

$$[E(h_1|\Pi^*)]^{-1} = \frac{e_2 Q_2}{\eta_2(m+e_2)} .$$

IV. EXAMPLE: REGRESSION THROUGH THE ORIGIN

Previous chapters have presented a Bayesian approach to the nonresponse problem of Section 1.3 for the case where the characteristic of interest is linearly related to one or more auxiliary variables. In this chapter we will confine our attention to the case where there is one auxiliary variable and where the expected value of the characteristic of interest is proportional to that auxiliary variable. We first summarize the results of Chapters II and III for this special case then present a hypothetical example to illustrate these results and to suggest a computationally convenient approximation.

4.1. Summary of Results

We now assume that there is one auxiliary variable available for each unit in U. We assume that $x = 1$ and that X_i $(i = 1, 2, \ldots, g)$, β_1, β_2, λ_1, λ_2, τ_1 and τ_2 are scalars. The assumptions of Chapters II and III are modified accordingly. In particular, given X, h, β and Z, we assume that $Y_{11}, \ldots, Y_{gn_g}$ are independent random variables such that

$$Y_{ij} \sim N(Z_{ij}X_i\beta_1 + (1-Z_{ij})X_i\beta_2, (\frac{1}{h_1})^{Z_{ij}}(\frac{1}{h_2})^{1-Z_{ij}}) \tag{4.1.1}$$

for $i = 1, 2, \ldots, g$ and $j = 1, 2, \ldots, n_i$.

The posterior means for β_1 and β_2 found in (2.2.9) become

$$\hat{\beta}_1 := (X_A'X_A+\tau_1)^{-1}(X_A'Y_A+\tau_1\lambda_1)$$

$$= \frac{(\Sigma_{ij\in A} X_iY_{ij})+\tau_1\lambda_1}{(\Sigma_{i=1}^{g} a_iX_i^2)+\tau_1} \tag{4.1.2}$$

and

$$\hat{\beta}_2 := (X_B'X_B+\tau_2)^{-1}(X_B'Y_B+\tau_2\lambda_2)$$

$$= \frac{(\Sigma_{ij\in B} X_iY_{ij})+\tau_2\lambda_2}{(\Sigma_{i=1}^{g} b_iX_i^2)+\tau_2} \; . \tag{4.1.3}$$

The posterior density, $f(\underset{\sim}{R}|II)$, in (2.2.28) remains unchanged except that, from (2.2.23) and (2.2.24), we now have

$$\mu_R^* = \sum_{i=1}^{g} [S_iX_i\hat{\beta}_1+(c_i+d_i-S_i)X_i\hat{\beta}_2] \tag{4.1.4}$$

for $\hat{\beta}_1$ and $\hat{\beta}_2$ given by (4.1.2) and (4.1.3) and

$$V_R^* = (n-m-r)h_2^{-1} + \sum_{i=1}^{g} [h_1^{-1}S_i+h_2^{-1}(d_i-S_i)]$$

$$+ h_1^{-1}\left[\sum_{i=1}^{g} S_i^2X_i^2\right] \Bigg/ \left[\left(\sum_{i=1}^{g} a_iX_i^2\right)+\tau_1\right] \quad +$$

$$+ h_1^{-1} \sum_{i=1}^{g} \sum_{\substack{k=1 \\ k \neq i}}^{g} \left[S_i S_k X_i X_k \Big/ \left(\sum_{\ell=1}^{g} a_\ell X_\ell^2 + \tau_1 \right) \right]$$

$$+ h_2^{-1} \left[\sum_{i=1}^{g} (c_i + d_i - S_i)^2 X_i^2 \right] \Big/ \left[\left(\sum_{i=1}^{g} b_i X_i^2 \right) + \tau_2 \right] \tag{4.1.5}$$

$$+ h_2^{-1} \sum_{i=1}^{g} \sum_{\substack{k=1 \\ k \neq i}}^{g} \left[(c_i + d_i - S_i)(c_k + d_k - S_k) X_i X_k \Big/ \left(\sum_{\ell=1}^{g} b_\ell X_\ell^2 + \tau_2 \right) \right] .$$

The posterior means of $\pi_1, \ldots \pi_g$ are again given by (2.3.2). The expression for $\hat{R}$ in (2.3.11) is now

$$\hat{R} = \hat{\beta}_1 \sum_{i=1}^{g} d_i \hat{\pi}_i X_i + \hat{\beta}_2 \sum_{i=1}^{g} [c_i + d_i(1 - \hat{\pi}_i)] X_i . \tag{4.1.6}$$

The posterior variance, $V(R|\Pi)$, given by (2.3.19) remains the same with

$$\delta_{ik} := X_i X_k / \left[\left(\sum_{i=1}^{g} a_i X_i^2 \right) + \tau_1 \right] , \quad i, k = 1, 2, \ldots, g,$$

and

$$\epsilon_{ik} := X_i X_k / \left[\left(\sum_{i=1}^{g} b_i X_i^2 \right) + \tau_2 \right] , \quad i, k = 1, 2, \ldots, g$$

and $(X_i' = X_i,\ i = 1, 2, \ldots, g)$.

4.2. Simulation Example

In order to illustrate the ideas summarized in Section 4.1, a population of size $N = 120$ was generated. It was assumed that $\lambda_1 = 5$, $\lambda_2 = 2$, $h_1^{-1} = h_2^{-1} = 25$ and $\tau_1 = \tau_2 = 125$. The prior precisions of β_1 and β_2 were thus $h_1\tau_1 = 5$ and $h_2\tau_2 = 5$ respectively. This produced $\beta_1 = 5.3214$ and $\beta_2 = 1.5686$. The population had $g = 3$ groups with $X_1 = 35$, $X_2 = 40$ and $X_3 = 50$. It was assumed that π_1, π_2 and π_3 had prior means 0.6, 0.7 and 0.8 and prior variances 0.04, 0.04 and 0.04 respectively.

A systematic sample of size $n = 105$ was chosen, and it was assumed that $\pi_1 = 0.6$, $\pi_2 = 0.7$ and $\pi_3 = 0.8$. As will be seen later, n/N was chosen large for ease of calculation of $f(\underset{\sim}{R}|\Pi)$. For this sample, we have the following relevant information:

i	X_i	n_i	s_i	a_i	b_i	c_i	d_i
1	35	40	35	23	6	6	5
2	40	40	35	27	4	4	5
3	50	40	35	31	2	2	5
		$N=120$	$n=105$	$r=81$	$m=12$	12	15

$$\sum_{ij \in A} X_i Y_{ij} = 789,567 \qquad \sum_{ij \in B} X_i Y_{ij} = 28,800$$

$$\sum_{i=1}^{g} a_i X_i^2 = 148,875 \qquad \sum_{i=1}^{g} b_i X_i^2 = 18,750$$

$$\hat{\beta}_1 = 5.3033 \qquad \hat{\beta}_2 = 1.5391$$

$$\sum_{ij \in A} Y_{ij} = 18,221.6 \qquad \sum_{ij \in B} Y_{ij} = 720.5$$

$$\hat{\pi}_1 = 0.650, \quad \hat{\pi}_2 = 0.764, \quad \hat{\pi}_3 = 0.879$$

Thus, from (4.1.6), we have $\hat{R} = 3,515.84$ so that $\hat{T} = 22,457.94$. Note that $T = 21,963.5$.

Figure 1 shows the posterior density, $f(\underset{\sim}{R}|II)$, for this example as well as a normal density function having the same mean and variance. The corresponding cumulative distribution functions are plotted in Figure 2.

The Bayes point estimate of R in this example is $\hat{R} = 3,515.84$ as found by using (4.1.6). A corresponding interval estimate is made by finding points R_1 and R_2 such that the posterior probability of the event "$R_1 \leq R \leq R_2$" is, say, 0.95. Typically, R_1 and R_2 are chosen with the further restriction that $R_2 - R_1$ be as small as possible. A careful examination of Figure 2 reveals that nearly the same intervals will be obtained by approximating the actual cumulative distribution function by a normal cumulative distribution function having the mean and variance found

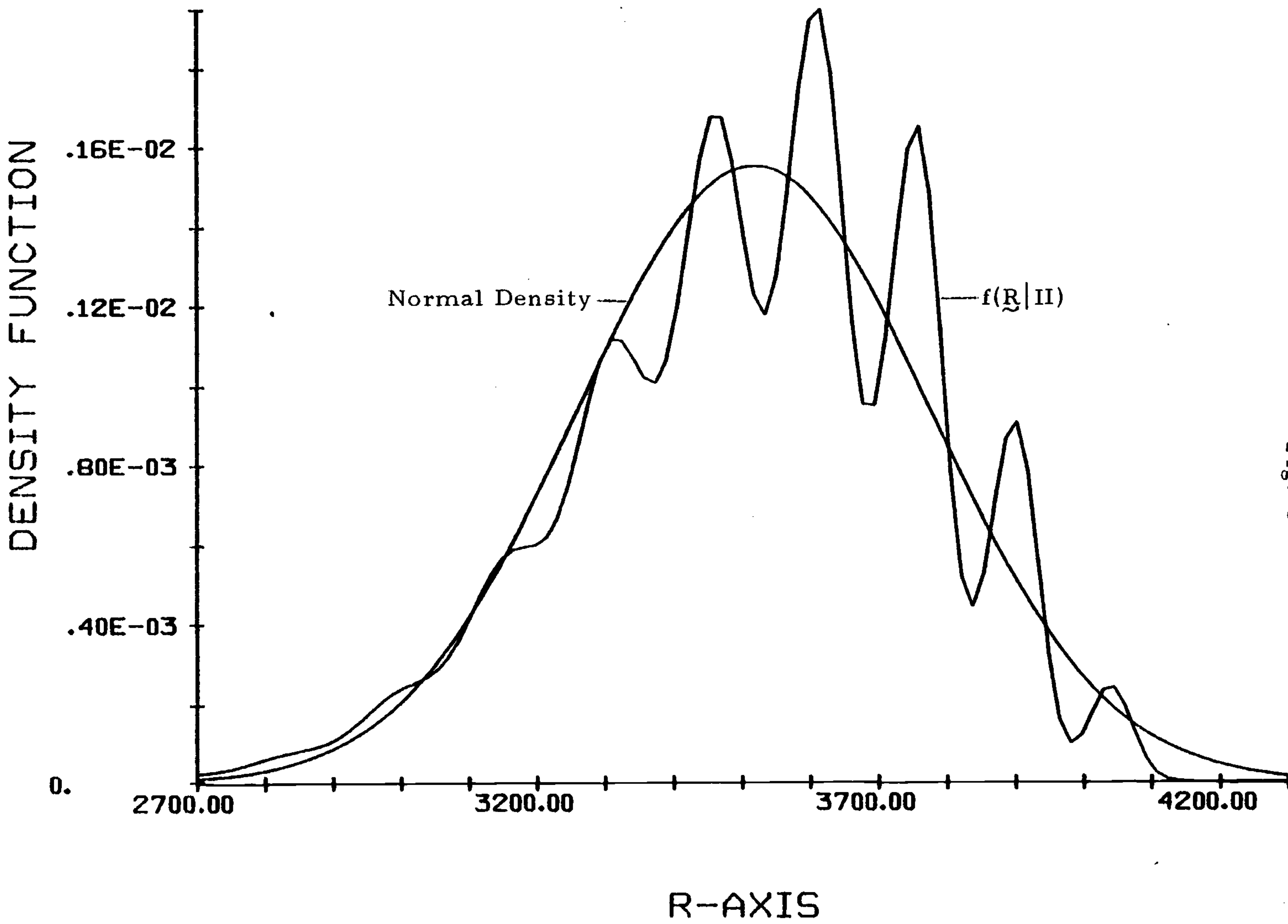
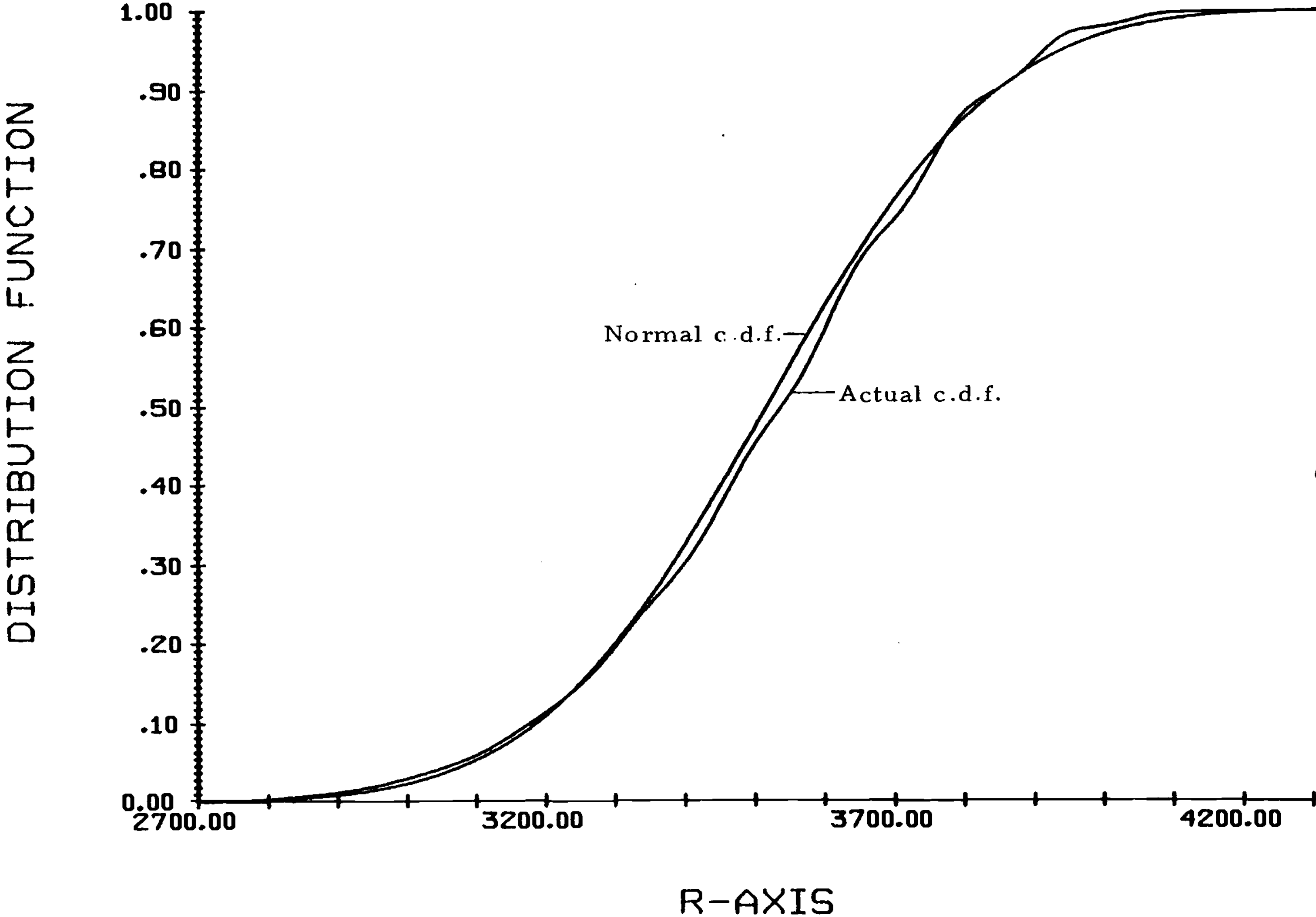

DENSITY FUNCTION
.16E-02
.12E-02
.80E-03
.40E-03
0.
2700.00
3200.00
3700.00
4200.00
R-AXIS
Normal Density
f(R|II)

Fig. 1



DISTRIBUTION FUNCTION
1.00
.90
.80
.70
.60
.50
.40
.30
.20
.10
0.00
2700.00
3200.00
3700.00
4200.00
R-AXIS
Normal c.d.f.
Actual c.d.f.

Fig. 2

by using (4. 1. 6) and (2. 3. 19) respectively. This approximation is explored further in Section 4. 3.

4. 3. Empirical Suggestion of Normal Approximation

The example of Section 4. 2 suggests that, for the purpose of making interval estimates of R, the actual posterior distribution function can be approximated by a normal distribution function having mean $E(R|II)$ and variance $Var(R|II)$. While this mean and variance are easily calculated, the actual distribution function is expensive to calculate. The normal approximation is easy to use, since tables of normal distribution function values are readily available.

The example of Section 4. 2 is unusual, since $N - n = 15$. In most surveys, the difference between N and n will be at least 500 and more often 1000 or more. In order to see the effect of increasing $N - n$, it was assumed that the sample of Section 4. 2 was actually drawn from a larger population. Figures 3 through 16 show the resulting density function and distribution function plots. In each case, the odd numbered figure corresponds to Figure 1, while the even numbered figure corresponds to Figure 2. The numbers of nonsampled units in groups 1, 2 and 3 were assumed to be as follows:

Figures	d_1	d_2	d_3	N-n
3-4	8	5	5	18
5-6	10	8	7	25
7-8	10	10	10	30
9-10	0	0	30	30
11-12	0	30	0	30
13-14	15	0	15	30
15-16	0	50	0	50

The calculation of $f(\underset{\sim}{R} \mid II)$ and the corresponding distribution function is prohibitively expensive for N-n larger than 50.

Figures 3 through 16 suggest that the use of this normal approximation is appropriate whenever the nonsampled units are not clustered in one group (Figures 9 and 10). In most real situations, this will not be a problem.

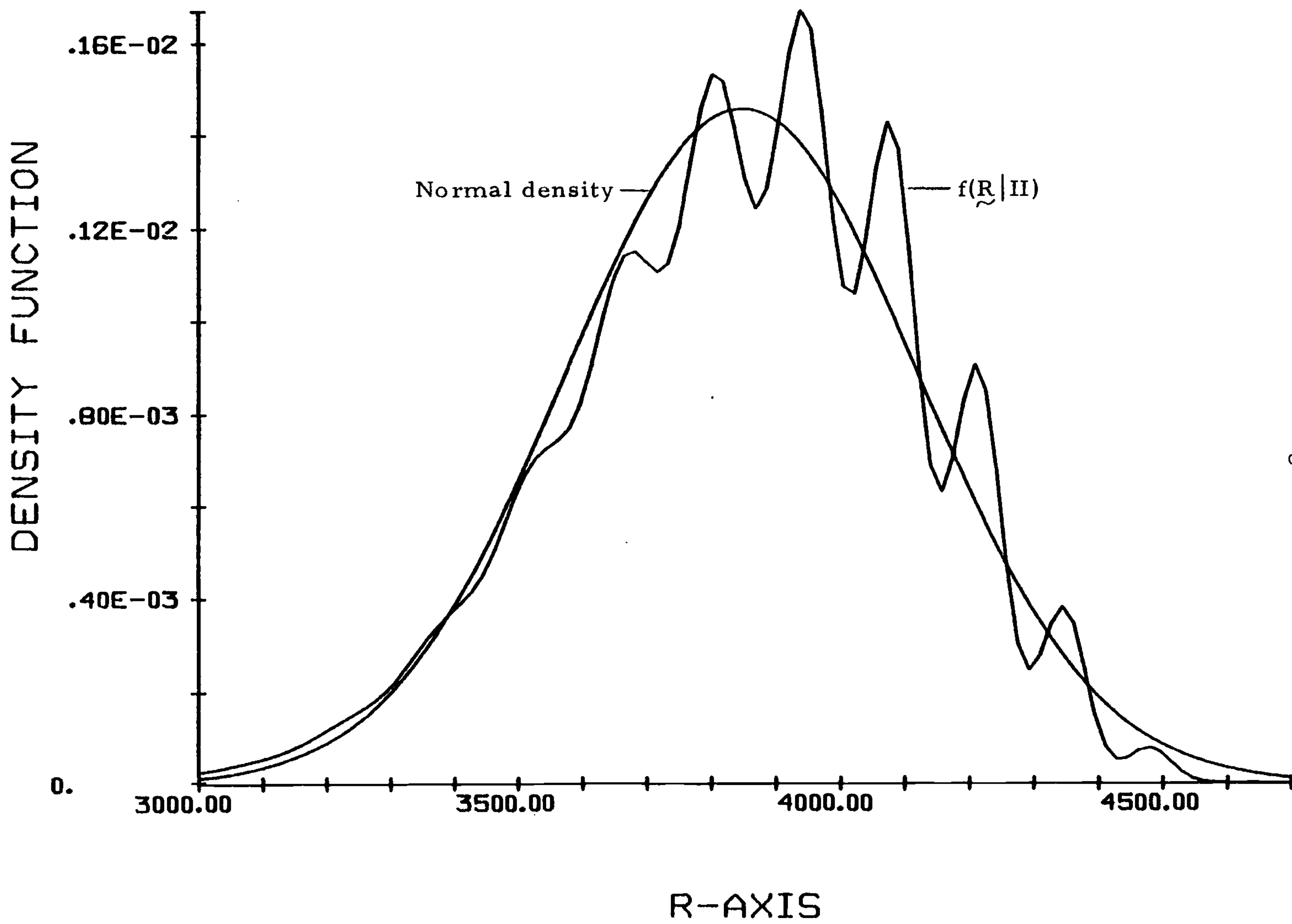

DENSITY FUNCTION
.16E-02
.12E-02
.80E-03
.40E-03
0.
3000.00
3500.00
4000.00
4500.00
R-AXIS
Normal density
f(R|II)

Fig. 3

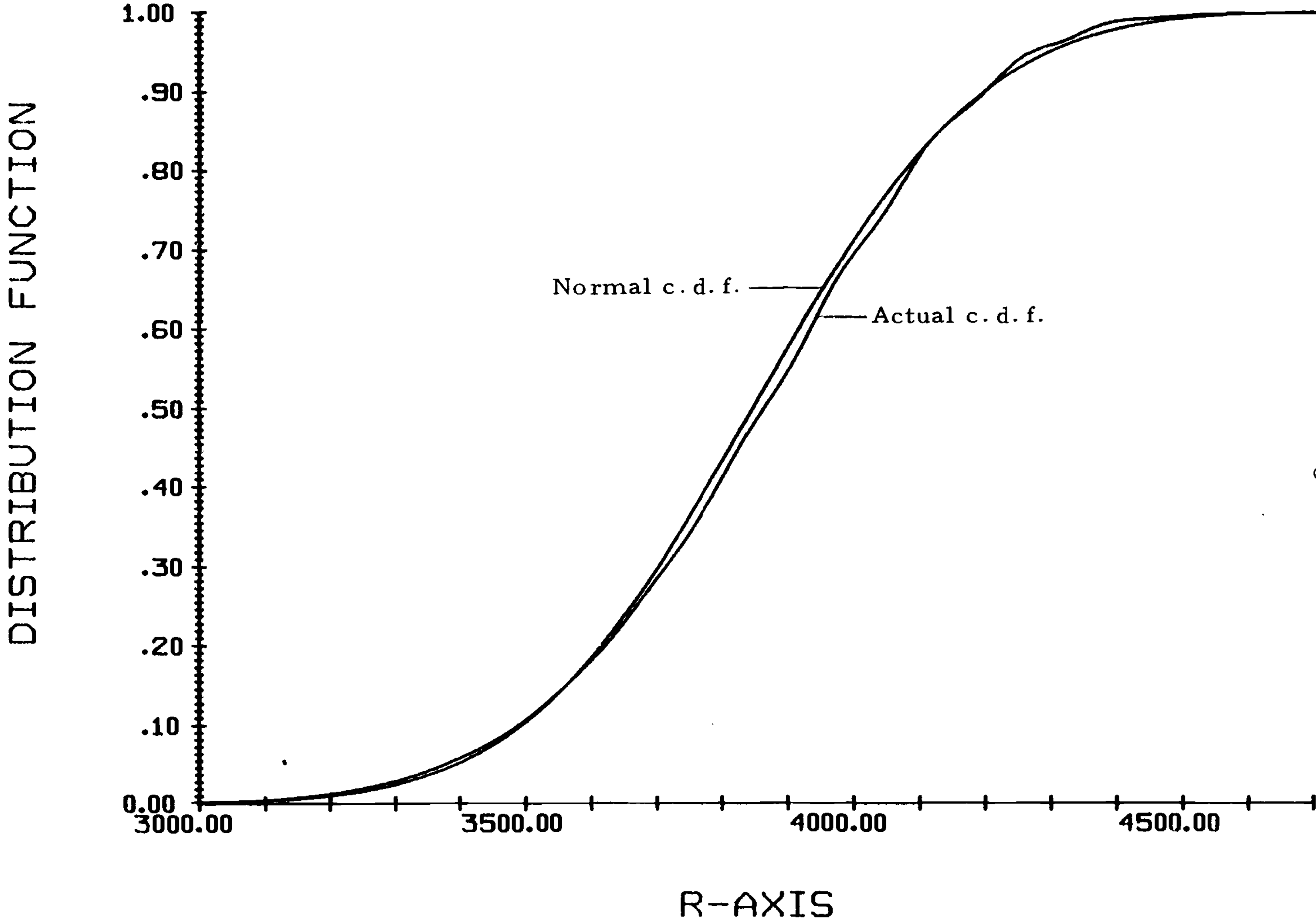

DISTRIBUTION FUNCTION
1.00
.90
.80
.70
.60
.50
.40
.30
.20
.10
0.00
3000.00
3500.00
4000.00
4500.00
R-AXIS
Normal c.d.f.
Actual c.d.f.

Fig. 4

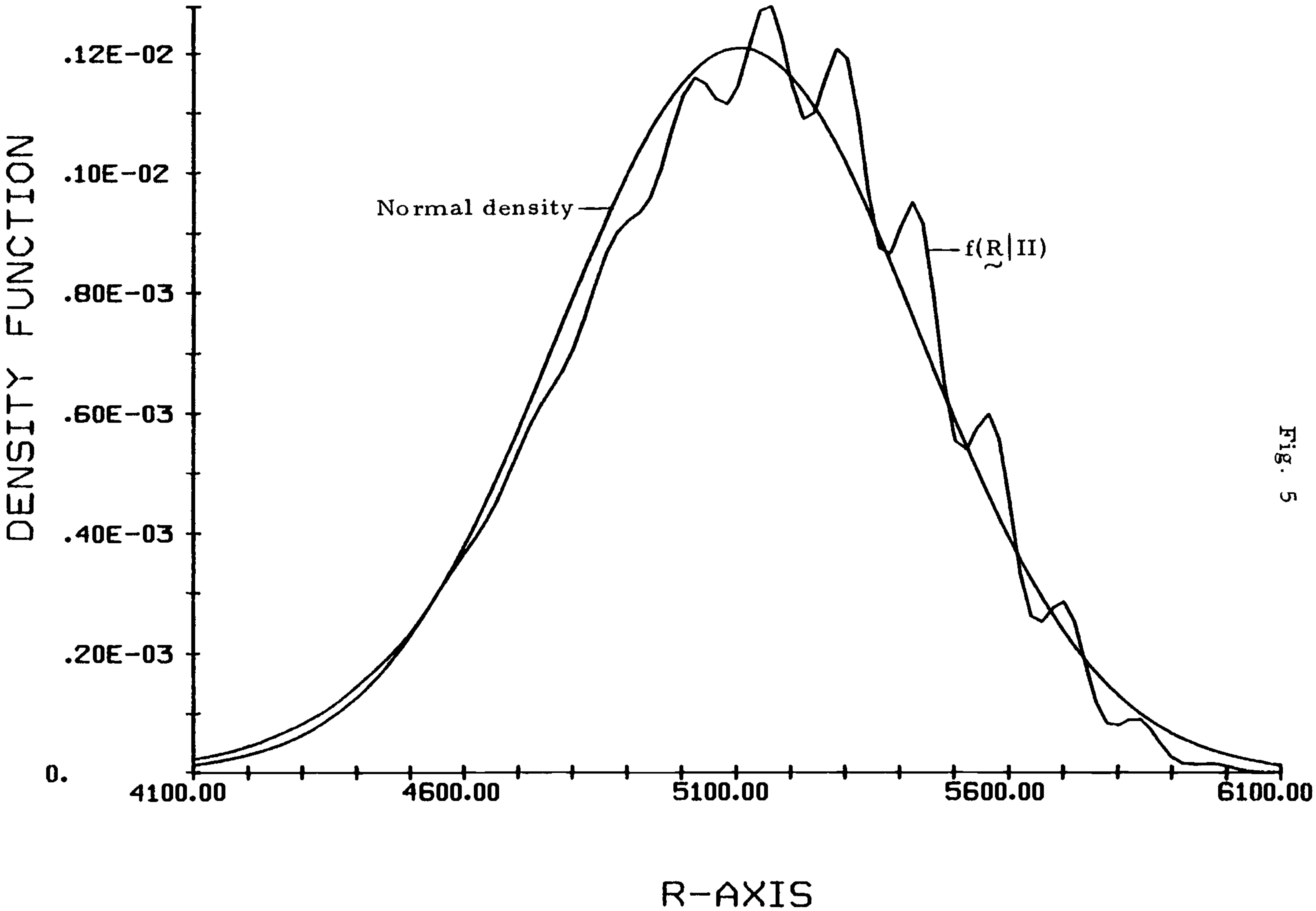

DENSITY FUNCTION
.12E-02
.10E-02
.80E-03
.60E-03
.40E-03
.20E-03
0.
4100.00
4600.00
5100.00
5600.00
6100.00
R-AXIS
Normal density
f(R|II)

Fig. 5

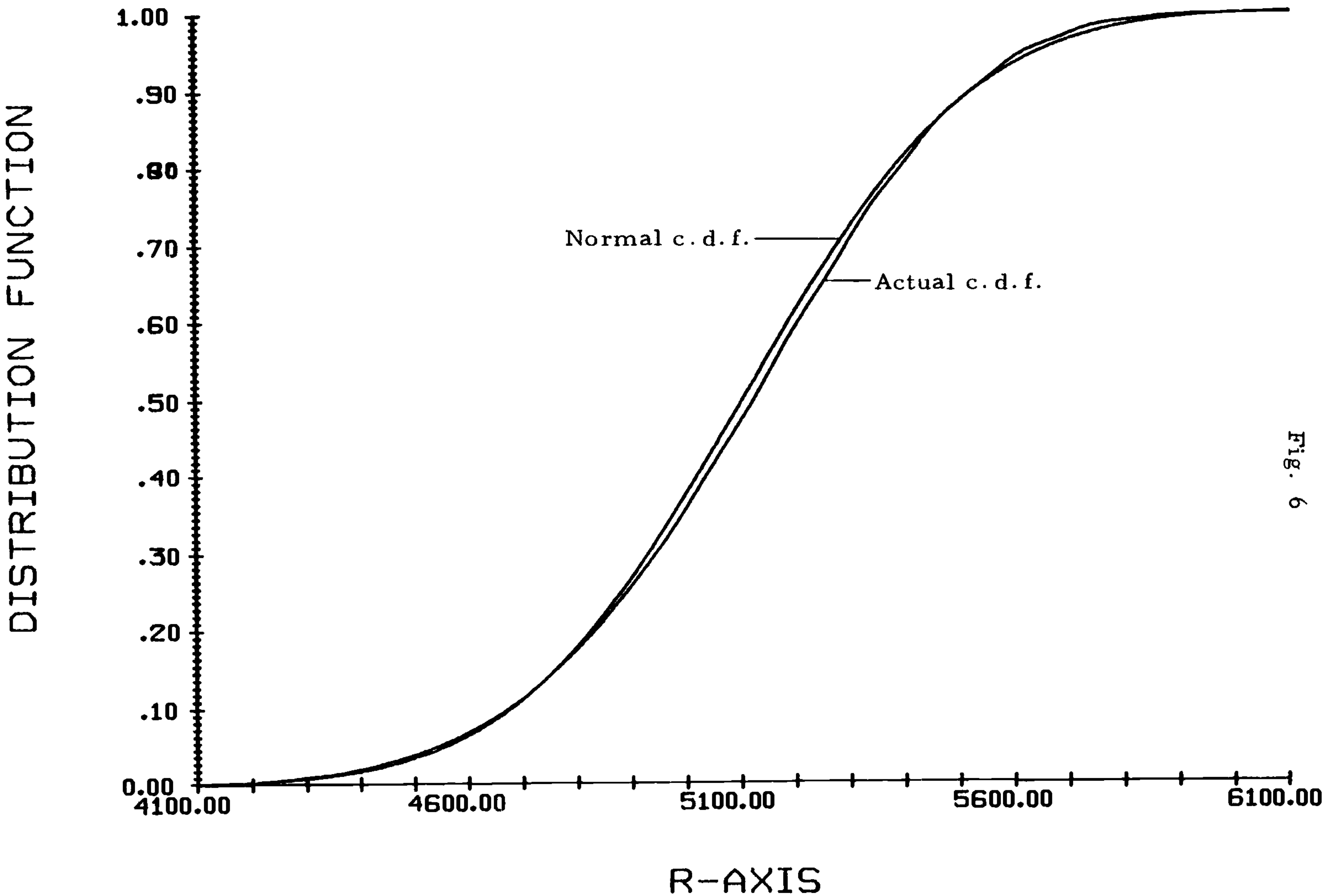

DISTRIBUTION FUNCTION
1.00
.90
.80
.70
.60
.50
.40
.30
.20
.10
0.00
4100.00
4600.00
5100.00
5600.00
6100.00
R-AXIS
Normal c.d.f.
Actual c.d.f.

Fig. 6

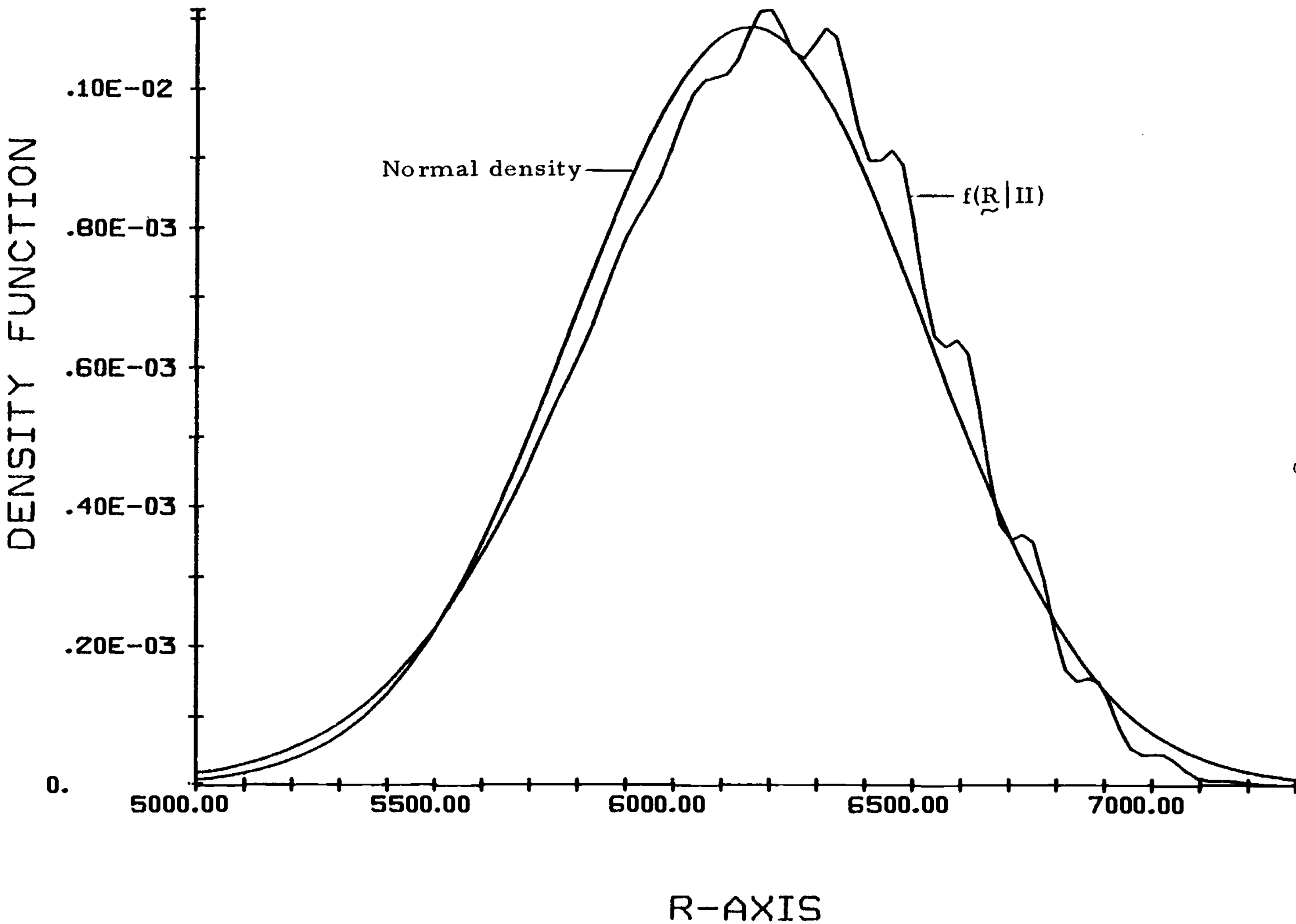

DENSITY FUNCTION
.10E-02
.80E-03
.60E-03
.40E-03
.20E-03
0.
Normal density
f(R|II)
5000.00
5500.00
6000.00
6500.00
7000.00
R-AXIS

Fig. 7

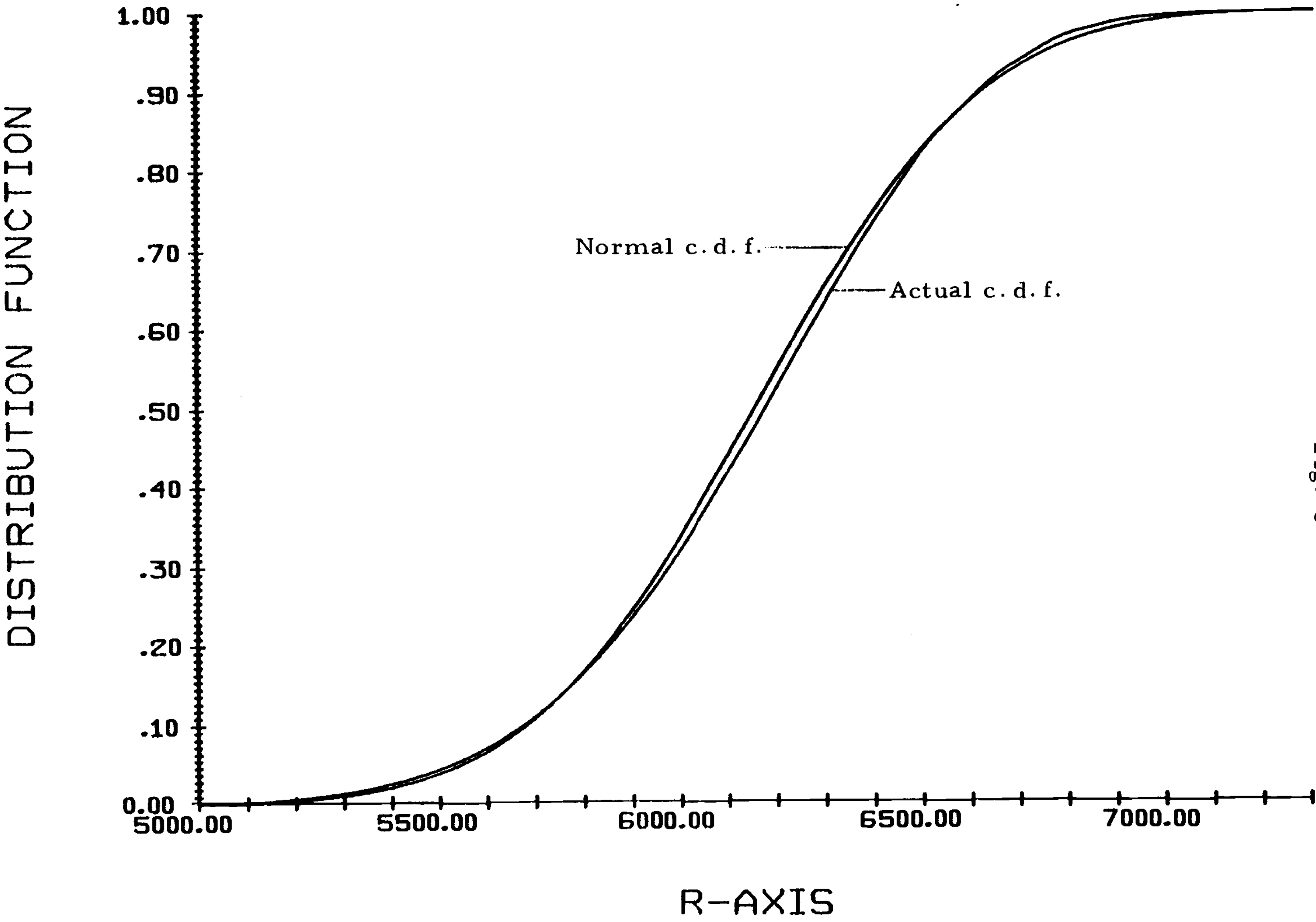

DISTRIBUTION FUNCTION
1.00
.90
.80
.70
.60
.50
.40
.30
.20
.10
0.00
5000.00
5500.00
6000.00
6500.00
7000.00
R-AXIS
Normal c.d.f.
Actual c.d.f.

Fig. 8

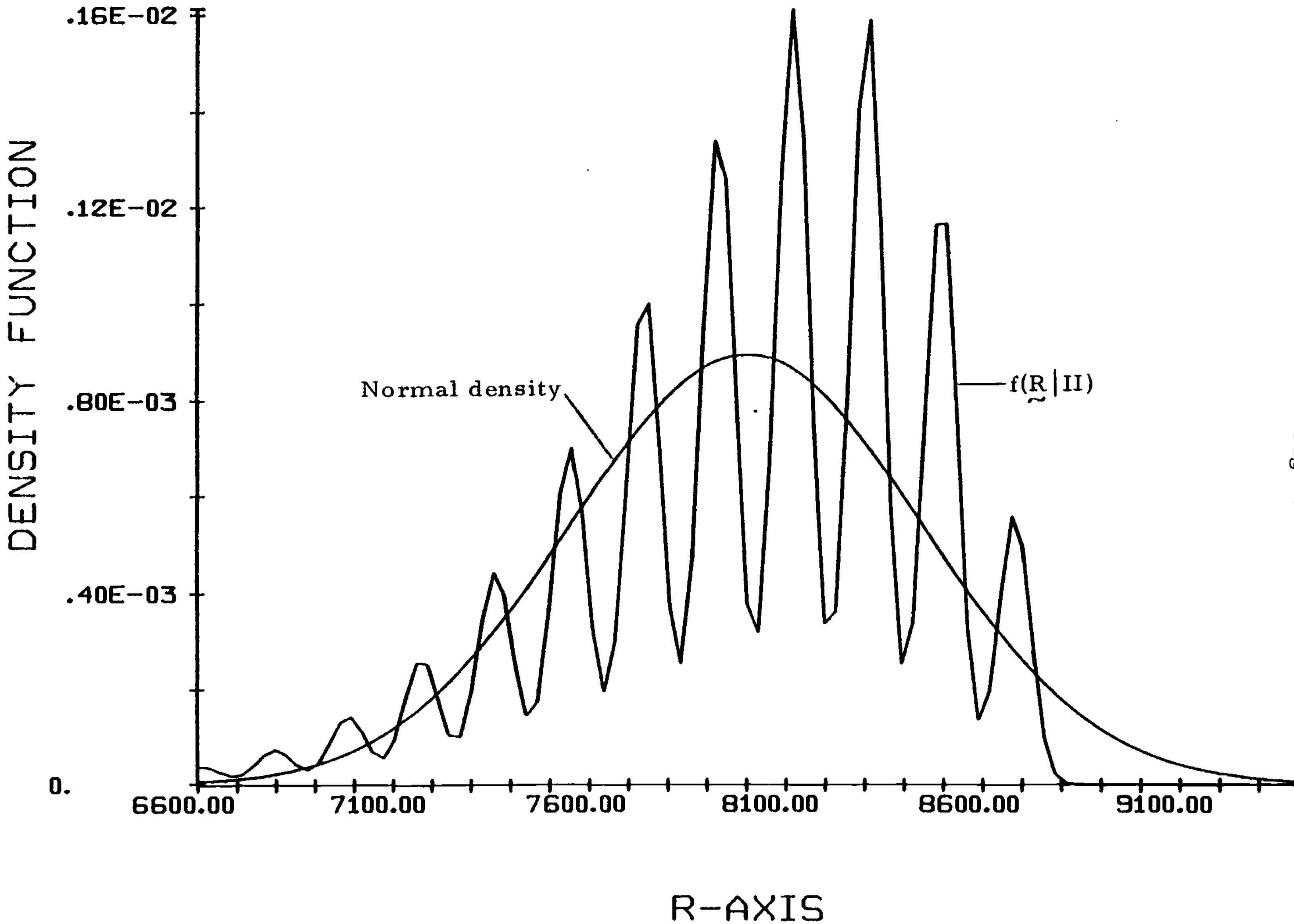

DENSITY FUNCTION
.16E-02
.12E-02
.80E-03
.40E-03
0.
6600.00
7100.00
7600.00
8100.00
8600.00
9100.00
R-AXIS
Normal density
f(R|II)

Fig. 9

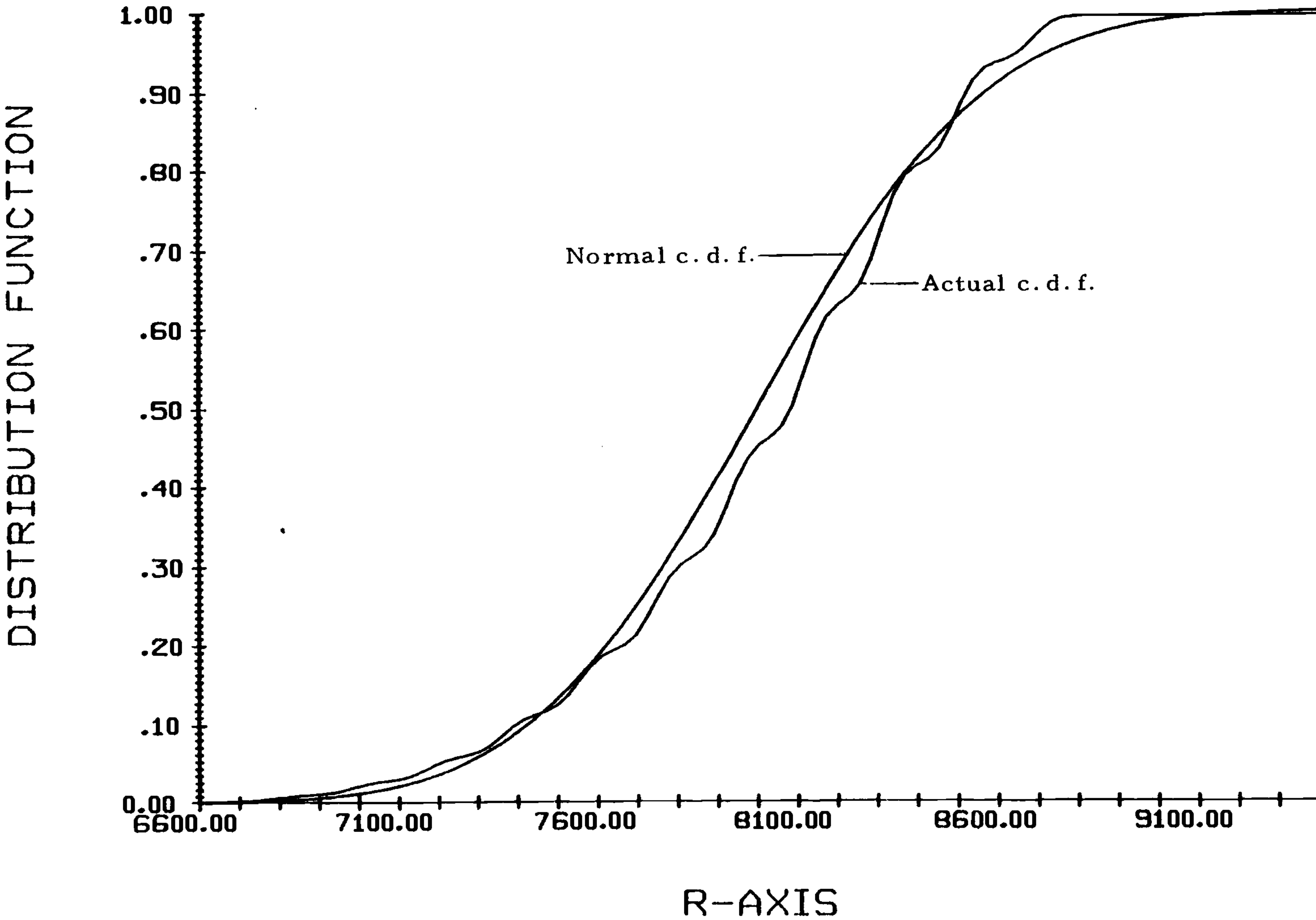
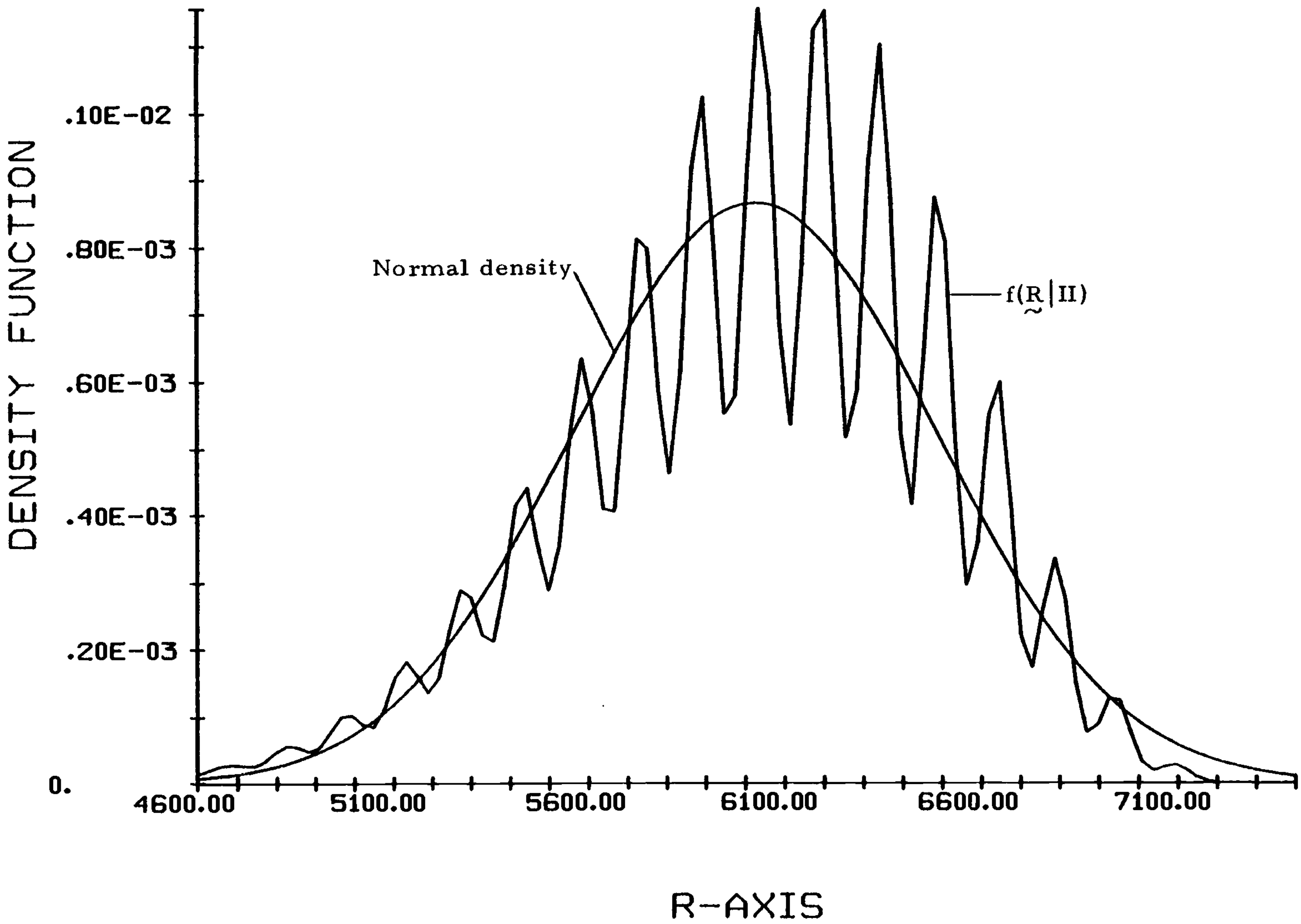

DISTRIBUTION FUNCTION
1.00
.90
.80
.70
.60
.50
.40
.30
.20
.10
0.00
6600.00
7100.00
7600.00
8100.00
8600.00
9100.00
R-AXIS
Normal c. d. f.
Actual c. d. f.

Fig. 10



DENSITY FUNCTION
.10E-02
.80E-03
.60E-03
.40E-03
.20E-03
0.
4600.00
5100.00
5600.00
6100.00
6600.00
7100.00
R-AXIS
Normal density
$f(\underset{\sim}{R}|II)$

Fig. 11

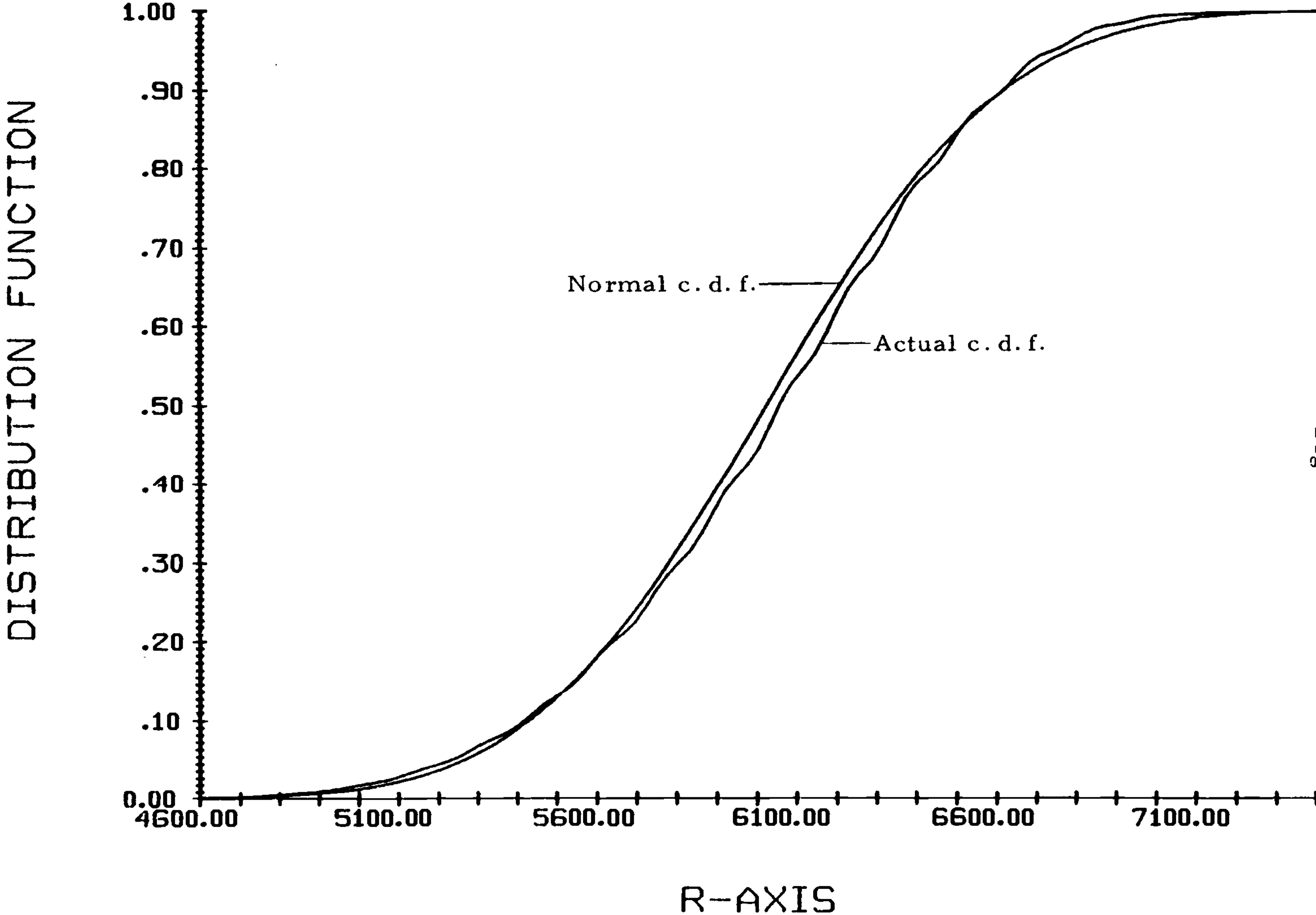

DISTRIBUTION FUNCTION
1.00
.90
.80
.70
.60
.50
.40
.30
.20
.10
0.00
4600.00
5100.00
5600.00
6100.00
6600.00
7100.00
R-AXIS
Normal c.d.f.
Actual c.d.f.

Fig. 12

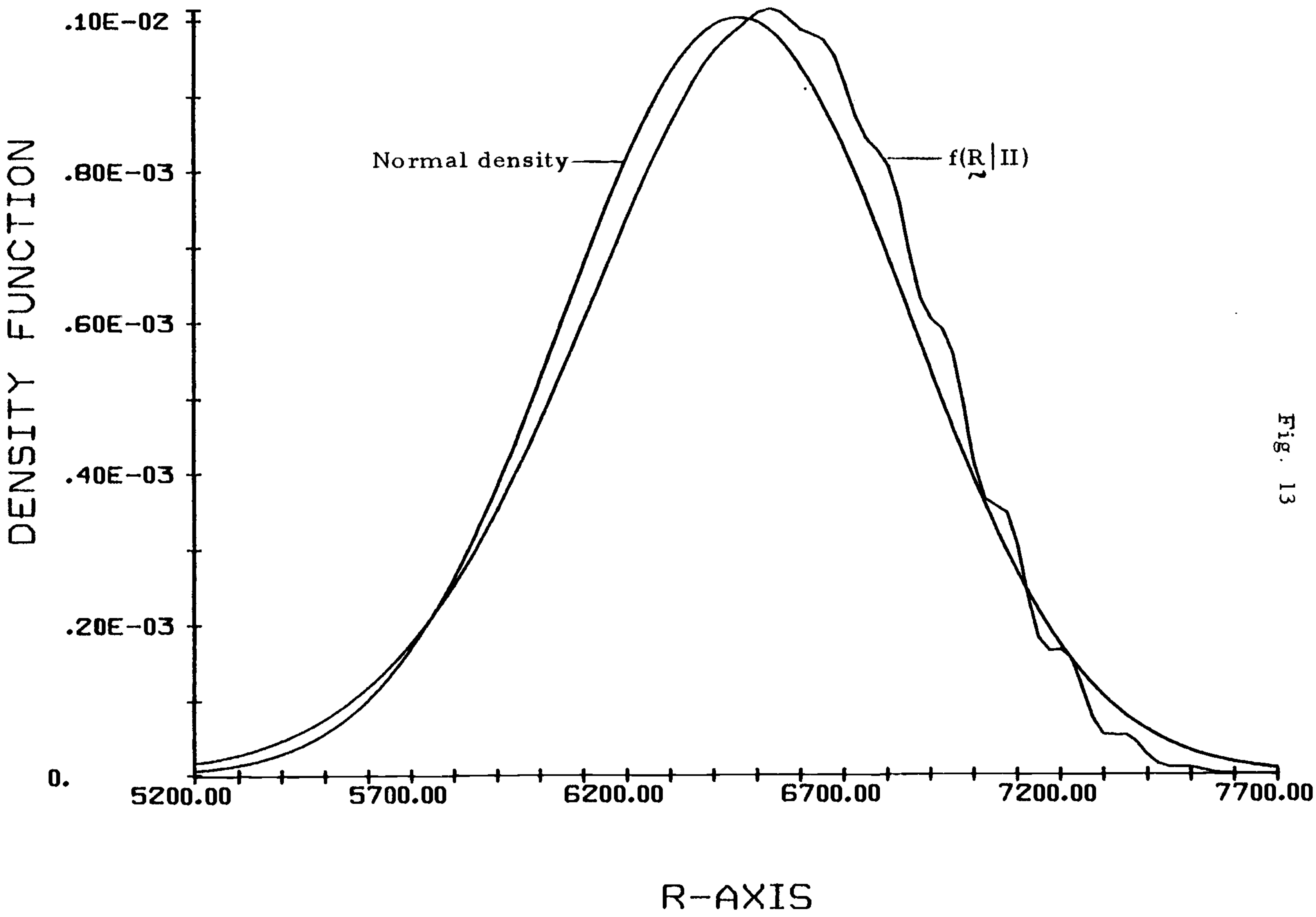

.10E-02
.80E-03
.60E-03
.40E-03
.20E-03
0.
5200.00
5700.00
6200.00
6700.00
7200.00
7700.00
DENSITY FUNCTION
R-AXIS
Normal density
f(R|II)

Fig. 13

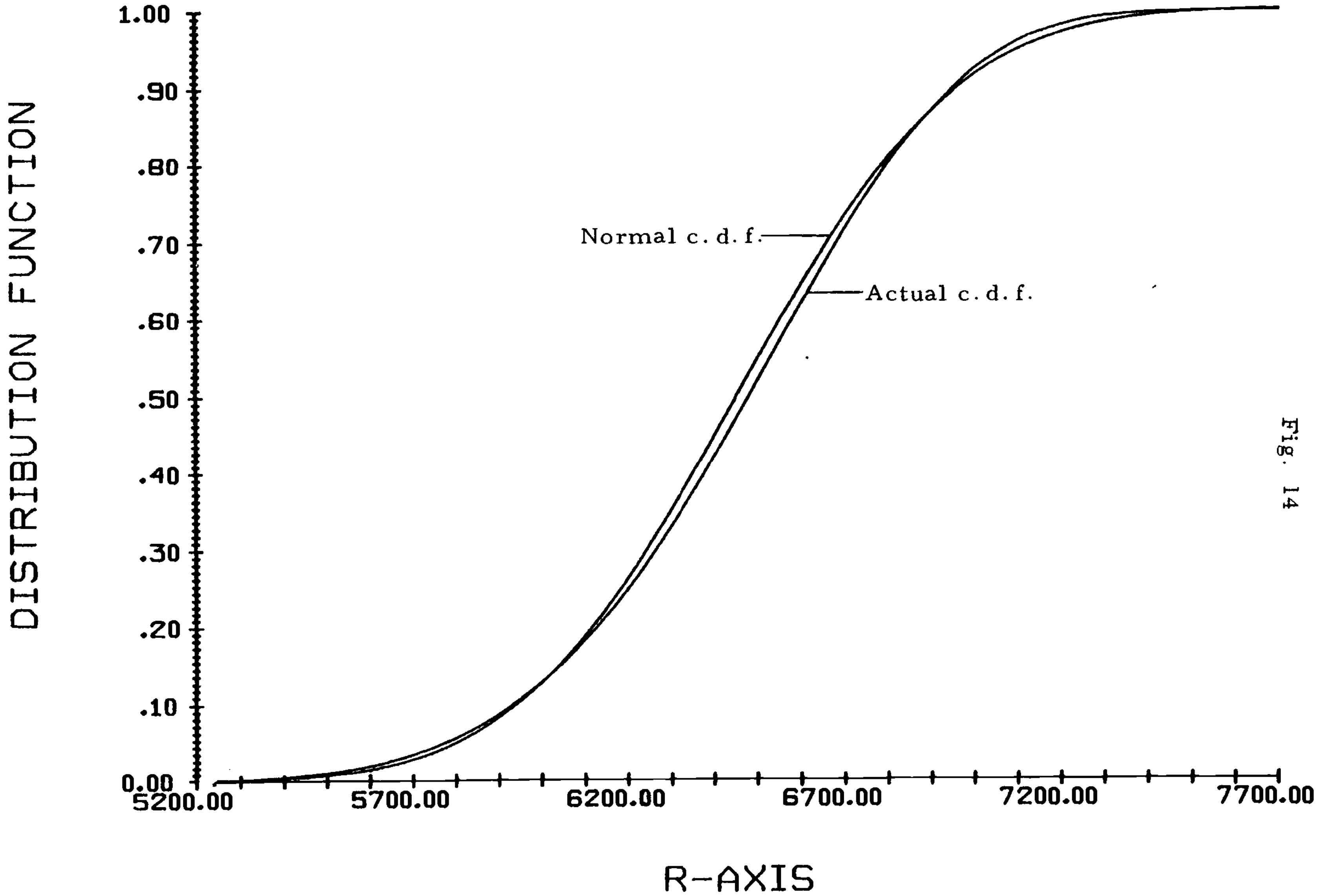
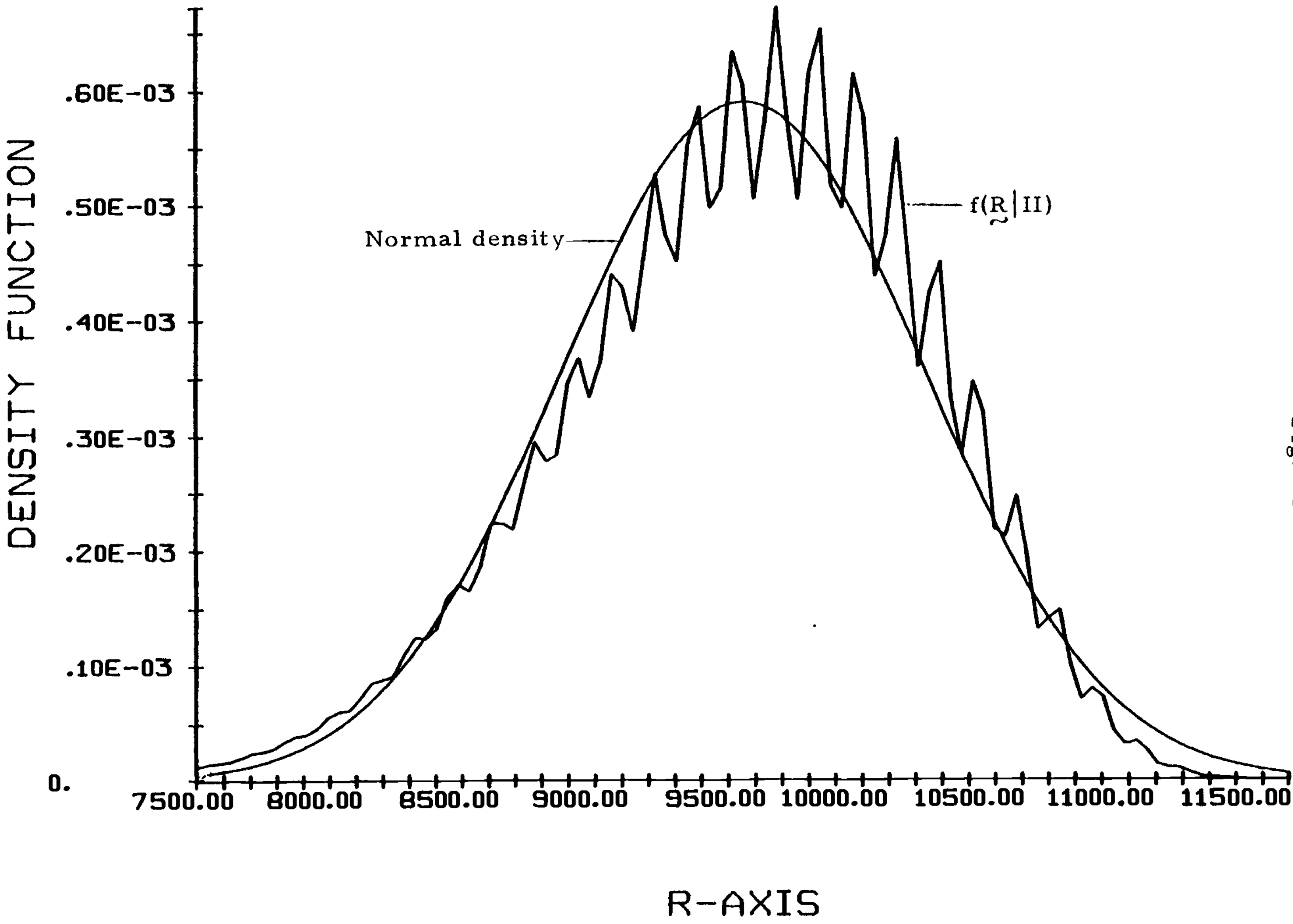

DISTRIBUTION FUNCTION
1.00
.90
.80
.70
.60
.50
.40
.30
.20
.10
0.00
5200.00
5700.00
6200.00
6700.00
7200.00
7700.00
R-AXIS
Normal c. d. f.
Actual c. d. f.

Fig. 14



DENSITY FUNCTION
.60E-03
.50E-03
.40E-03
.30E-03
.20E-03
.10E-03
0.
7500.00
8000.00
8500.00
9000.00
9500.00
10000.00
10500.00
11000.00
11500.00
R-AXIS
Normal density
f(R|II)

Fig. 15

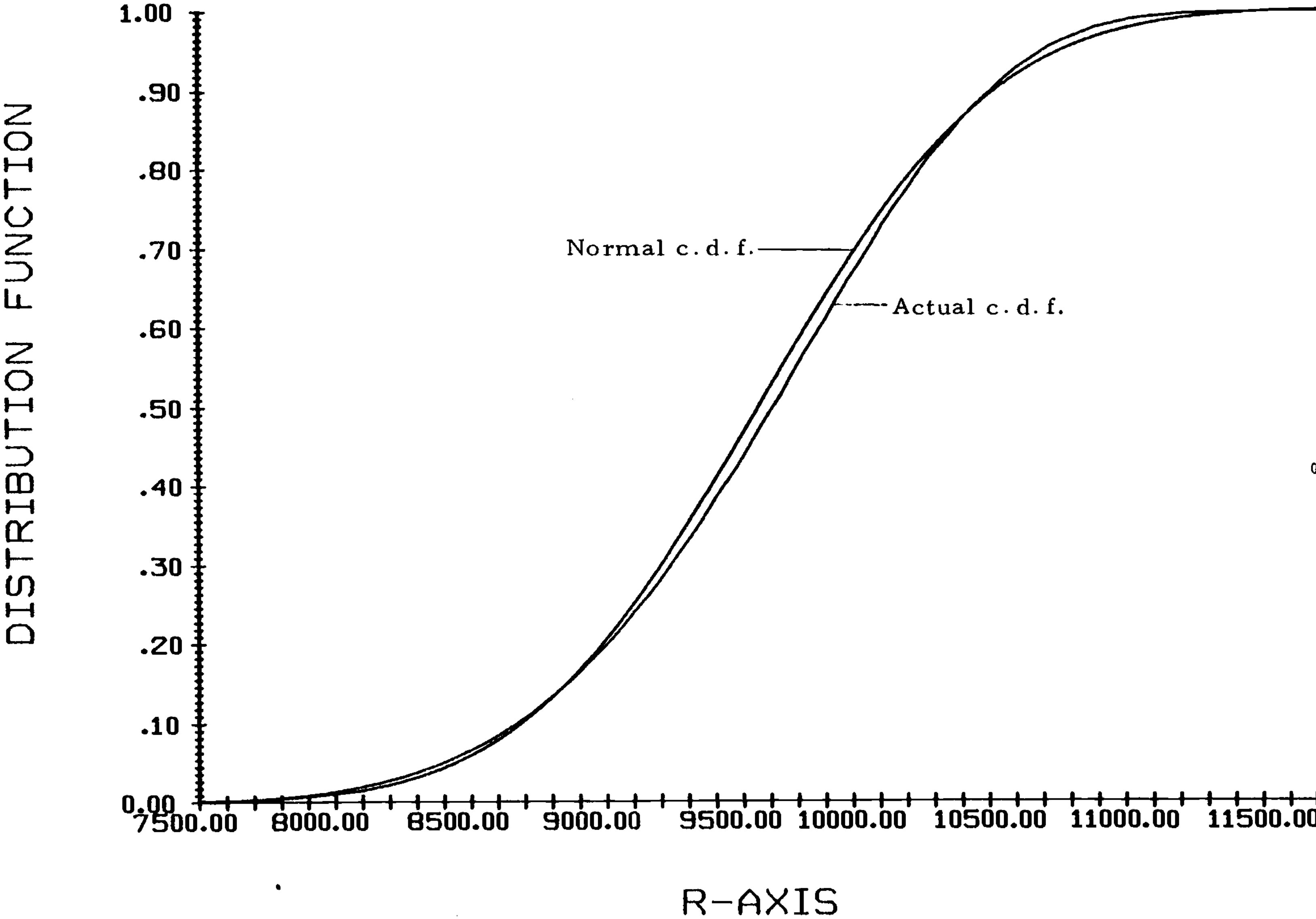

DISTRIBUTION FUNCTION
1.00
.90
.80
.70
.60
.50
.40
.30
.20
.10
0.00
Normal c.d.f.
Actual c.d.f.
7500.00
8000.00
8500.00
9000.00
9500.00
10000.00
10500.00
11000.00
11500.00
R-AXIS

Fig. 16

BIBLIOGRAPHY

Armstrong, J.S., and Overton, T.S. (1977), "Estimating Nonresponse Bias in Mail Surveys," Journal of Marketing Research, XIV, 396-402.

Basu, D. (1971), "An Essay on the Logical Foundations of Survey Sampling, Part One," in Foundations of Statistical Inference, eds. V.P. Godambe and D.A. Sprott, Toronto: Holt, Rinehart & Winston, 203-33.

Cassel, C.M., Särndal, C.E., and Wretman, J.H. (1977), Foundations of Inference in Survey Sampling, New York: John Wiley & Sons.

Cornish, E.A. (1954), "The Multivariate t-Distribution Associated with a Set of Normal Sample Deviates," Australian Journal of Physics, 7, 531-42.

Ericson, W.A. (1967), "Optimal Sample Design with Nonresponse," Journal of the American Statistical Association, 62, 63-78.

Ericson, W.A. (1969), "Subjective Bayesian Models in Sampling Finite Populations," Journal of the Royal Statistical Society, Ser. B, 31, 195-224.

Godambe, V.P. (1955), "A Unified Theory of Sampling from Finite Populations," Journal of the Royal Statistical Society, Ser. B, 17, 269-78.

Godambe, V.P. (1966), "A New Approach to Sampling from Finite Populations I, II," Journal of the Royal Statistical Society, Ser. B, 28, 310-28.

Godambe, V.P. (1969), "Admissibility and Bayes Estimation in Sampling Finite Populations-V," Annals of Mathematical Statistics, 40, 672-6.

Hansen, M.H., and Hurwitz, W.N. (1946), "The Problem of Non-Response in Sample Surveys," Journal of the American Statistical Association, 41, 517-29.

Johnson, N. L., and Kotz, S. (1972), Distributions in Statistics: Continuous Multivariate Distributions, New York: John Wiley & Sons, Chapter 37.

Kanuck, L., and Berenson, C. (1975), "Mail Surveys and Response Rates: A Literature Review," Journal of Marketing Research, XII, 440-53.

Lindley, D.V., and Smith, A. F.M. (1972), "Bayes Estimates for the Linear Model," Journal of the Royal Statistical Society, 34, 1-18.

Neyman, J. (1971), Discussion of paper by R.M. Royall, in Foundations of Statistical Inference, eds. V.P. Godambe and D.A. Sprott, Toronto: Holt, Rinehart & Winston, 276-8.

Rao, J.N.K. (1975), "On the Foundations of Survey Sampling," in A Survey of Statistical Design and Linear Models, ed. J.N. Srivastava, The Hague: North Holland, 489-505.

Royall, R.M. (1971), "Linear Regression Models in Finite Population Samping Theory," in Foundations of Statistical Inference, eds. V.P. Godambe and D.A. Sprott, Toronto: Holt, Rinehart & Winston, 259-74.

Rubin, D.B. (1977), "Formalizing Subjective Notions About the Effect of Nonrespondents in Sample Surveys," Journal of the American Statistical Association, 72, 538-43.

Singh, B., and Sedransk, J. (1975), "Estimation of Regression Coefficients when there is Nonresponse," Technical Report No. 25, Statistical Science Division, State University of New York at Buffalo.

Solomon, H., and Zacks, S. (1970), "Optimal Design of Sampling from Finite Populations: A Critical Review and Indication of New Research Areas," Journal of the American Statistical Association, 65, 653-77.

APPENDICES

Appendix: Multivariate-t Distributions

This appendix summarizes pertinent definitions, notation and facts about multivariate-t distributions used in Chapter III. These facts are adapted from a paper by Cornish (1954) and a book by Johnson and Kotz (1972).

Definition A.1. If $L = (L_1, L_2, \ldots, L_w)'$ is a $w \times 1$ random vector having a probability density function in $\mathbb{R}^w$ given by

$$f(\underset{\sim}{L}) = \frac{\Gamma(\frac{1}{2}(\nu+w))}{(\pi\nu)^{\frac{1}{2}w}\Gamma(\frac{1}{2}\nu)|E|^{1/2}}[1+\nu^{-1}L'E^{-1}L]^{-\frac{1}{2}(\nu+w)}$$

where $\nu > 2$ and where E is a $w \times w$ positive definite matrix, then L is said to have a w-dimensional multivariate-t distribution with ν degrees of freedom and with characterizing matrix E^{-1}.

Fact A.2. If L has the distribution of A.1, then $E(L) = 0$, a $w \times 1$ vector of zero's, and $\mathrm{Cov}(L) := E(LL') = \frac{\nu}{\nu-2}E$.

Definition A.3. Let $K := L + \mu$ where μ is a $w \times 1$ vector of constants and where L has the distribution of A.1. Then K has a density function in $\mathbb{R}^w$ given by

$$f(\underset{\sim}{K}) = \frac{\Gamma(\frac{1}{2}(\nu+w))}{(\pi\nu)^{\frac{1}{2}w}\,\Gamma(\frac{1}{2}\nu)|E|^{1/2}}\,[1+\nu^{-1}(K-\mu)'E^{-1}(K-\mu)]^{-\frac{1}{2}(\nu+w)} .$$

Note that $E(K) = E(L) + \mu = \mu$ and that $\mathrm{Cov}(K) = \mathrm{Cov}(L) = \frac{\nu}{\nu-2}E$. We say that K has a w-dimensional multivariate-t distribution with ν degrees of freedom, with mean μ and with characterizing matrix E^{-1}.

Notation: If K has the distribution of A. 3, we write $K \sim t(w, \nu, \mu, E^{-1})$. Thus, in A. 1, $L \sim t(w, \nu, 0, E^{-1})$.

<u>Fact A. 4</u>. If $L \sim t(w, \nu, 0, E^{-1})$, then the characteristic function of L is given by

$$\phi_L(t) := E(e^{it'L}) = [\Gamma(\tfrac{1}{2}\nu)]^{-1}\int_0^\infty u^{\frac{1}{2}\nu-1}\exp\{-u-\tfrac{1}{4}\nu u^{-1}[t'Et]\}du$$

where $i := \sqrt{-1}$. If $K \sim t(w, \nu, \mu, E^{-1})$ then $K = L + \mu$ where $L \sim t(w, \nu, 0, E^{-1})$. Thus, the characteristic function of K is given by

$$\begin{aligned}\phi_K(t) &:= E(e^{it'K}) = E(e^{it'(L+\mu)}) = e^{it'\mu}\phi_L(t)\\ &= [\Gamma(\tfrac{1}{2}\nu)]^{-1}\int_0^\infty u^{\frac{1}{2}\nu-1}\exp\{-u+it'\mu-\tfrac{1}{4}\nu u^{-1}[t'Et]\}du .\end{aligned}$$

<u>Fact A.5.</u> Let $K \sim t(w, \nu, \mu, E^{-1})$ and let $J := HK$ where H is an $h \times w$ matrix of rank h $(h \leq w)$. Then $J \sim t(h, \nu, H\mu, (HEH')^{-1})$. In particular, the marginal distribution of the first h elements of the random vector K is $t(h, \nu, \mu_h, E_h^{-1})$ where μ_h is an $h \times 1$ vector consisting of the first h elements of μ and where E_h is the leading $h \times h$ submatrix of E.

<u>Fact A.6.</u> If $L \sim t(w, \nu, 0, E^{-1})$, then $L \xrightarrow[\nu \to +\infty]{\mathcal{L}} U$ where $U \sim N(0, E)$. That is, if $\{\nu_1, \nu_2, \ldots\}$ in any sequence of real numbers tending to $+\infty$, and if $\{L_1, L_2, \ldots\}$ is a sequence of random vectors such that $L_i \sim t(w, \nu_i, 0, E^{-1})$ for $i = 1, 2, \ldots,$ then the sequence $\{L_1, L_2, \ldots\}$ converges in law to the random vector U where U has a w-dimensional normal distribution with mean vector 0 and covariance matrix E.

<u>Fact A.7.</u> Slutsky's theorem and Fact A.6 show that, if $K \sim t(w, \nu, \mu, E^{-1})$, then $K \xrightarrow[\nu \to +\infty]{\mathcal{L}} V$ where $V \sim N(\mu, E)$.