STUDIES ON HAZARD RATIO FUNCTIONS IN SURVIVAL ANALYSIS
WITH SPECIAL REFERENCE TO PROPORTIONAL
HAZARD FUNCTIONS MODELS

by

Brenda Kay Edwards
Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1018
June 1975
The proportional hazard function (PHF) model, $\mu_1(t) = \theta \mu_2(t)$, where $\mu_g(t)$ is the hazard function for population $\pi_g$, $g = 1, 2$, is used as a model for comparing survivorship functions of two (or more) populations. In Chapter II we discuss the estimation of the ratio $\theta = \mu_1(t)/\mu_2(t)$ and $\Delta = \theta/(1 + \theta)$ in the PHF model. We derive three new estimators and the maximum likelihood (ML) estimator of $\Delta$. The properties and distributions of the estimators for small sample sizes in two sampling schemes are presented. We consider estimation conditional on the observed numbers of individuals from $\pi_1$ and $\pi_2$ at risk of dying at each of the times of death (SSI). We also consider estimation when the numbers of individuals in the risk sets are taken as random variables which reflect the natural decrement occurring in the two populations (SSII). In SSI all four estimators are biased, the third estimator proposed was usually found to have the smallest mean square error (MSE) and was close to the ML estimator. In SSII the first estimator proposed was shown to be a function of the (unbiased) Mann-Whitney U-statistic, and the third estimator was usually found to have the smallest MSE.

The PHF is a special case when the hazard ratio function (HRF) is assumed constant for all $t$, i.e. $\theta(t) = \mu_1(t)/\mu_2(t) = \theta$ for all $t$. 
In Chapter III we investigate the more general situation in which the HRF is a function of \( t \). For many empirical data cases we found that the HRF is not constant, except perhaps over a restricted range of values of \( t \). The use of experimental evidence to draw conclusions about the form of the HRF may be difficult, especially for small samples. If, however, we specify the parametric family of the failure distributions, the HRF may be a well-behaved function of \( t \), and is studied analytically. The results presented in Chapter III suggest that the PHF assumption may be valid only under certain conditions on the parameter values or on the range of values of \( t \).

In Chapter IV we present the results of a Monte Carlo simulation study designed to compare some "average" of the function

\[
\Delta(t) = \frac{\theta(t)}{1 + \theta(t)}
\]

at the observed times of death with the estimates of "\( \Delta \)" obtained under the PHF model. The failure distributions of \( \pi_1 \) and \( \pi_2 \) are specified by a parametric model. We also suggest some alternatives to the PHF model and outline methods for assessing the adequacy of the PHF model with respect to these alternatives which are not based on any a priori knowledge of the failure distributions in \( \pi_1 \) and \( \pi_2 \).

An example is given in Chapter V to illustrate the use of the estimators of \( \Delta \), proposed in Chapter II, that allow for both tied and censored observations. We conclude the dissertation with some suggestions for future research.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................... v

LIST OF TABLES .............................................. vi

LIST OF FIGURES ........................................... viii

I. INTRODUCTION AND REVIEW OF LITERATURE

1.1 Introduction ............................................ 1

1.2 Definitions .............................................. 4

1.3 Parametric Models ....................................... 5

1.4 Nonparametric Models ................................... 5

1.4.1 Linear Approximation ................................ 8

1.4.2 Piece-wise Exponential Approximation ............... 9

1.4.3 Step-function Approximation ......................... 10

1.5 Survival Models with Concomitant Variables ............ 11

1.5.1 Parametric Models with Concomitant Variables ..... 12

1.6 "Quasi Parametric" Models .............................. 16

1.7 Comparison of Survival Curves ....................... 19

II. ESTIMATION IN THE PHF MODEL

2.1 Estimators of the Ratio $\theta = \mu_1(t)/\mu_2(t)$ .......... 21

2.2 Estimators of $\Delta$ and Their Properties Under Sample

   Scheme I .................................................. 25

   2.2.1 Approximation to First and Second Moments of $\Delta$ 25

   2.2.2 Choice of Weights $\{\alpha_i\}$ ................. 27

   2.2.3 Three Estimators of $\Delta$ .................... 28
2.2.4 (Conditional) Maximum Likelihood Estimator of $\Delta$ ... 30
2.2.5 Nearly Unbiased (NU) Estimators of $\Delta$ ............. 32

2.3 Estimators of $\Delta$ and Their Properties Under Sample

Scheme II .................................................................. 32
2.3.1 Relationship of $\hat{\Delta}_1$ to Mann-Whitney U-statistic .. 33
2.3.2 Other Estimators of $\Delta$ ....................... 35

2.4 Extensions to Ties and Censoring ......................... 36

2.4.1 Ties ......................................................... 36
2.4.2 Censoring ............................................. 40

2.5 Distributions of $\hat{\Delta}_\ell$ ($\ell = 1, 2, 3$) and $\hat{\Delta}_4$ and Some of Their Properties .............................. 40

2.5.1 SSI (No Censoring) ..................................... 41
2.5.2 SSII (No Censoring) .................................... 52

III. SOME ANALYTICAL AND EMPIRICAL INVESTIGATIONS OF THE HAZARD RATIO FUNCTION (HRF)

3.1 The HRF for Some Population Mortality Data ........ 63

3.2 The HRF for Some Empirical Data ...................... 64

3.3 The HRF for Some Parametric Models ................. 75

3.3.1 The HRF for Some Parametric Models from the Same Family of Distributions .................................. 77

3.3.1.1 HRF for Two Weibull Distributions ............... 79
3.3.1.2 HRF for Two Makeham Gompertz Distributions .. 79
3.3.1.3 HRF for Two Linear Hazard Distributions ....... 81
3.3.1.4 HRF for Two Lognormal Distributions ............ 82
3.3.1.5 HRF for Two (Left) Truncated Logistic Distributions ......................... 83

3.3.1.6 HRF for Two Gamma Distributions ............... 86
ACKNOWLEDGEMENTS

The contributions of Professor Regina C. Elandt-Johnson as adviser to both my academic endeavors and avocational pursuits are gratefully acknowledged and kindly remembered. To her and to Professor Norman L. Johnson I express a sincere appreciation for their guidance, encouragement, availability and friendship, both jointly and separately, during the course of this research. I also wish to thank my other committee members, Dr. Joan Cornoni, Dr. C. Ed Davis, Dr. David G. Hoel and in particular Dr. Dana E. A. Quade, for their helpful comments.

The financial support for this research and graduate study in the Department of Biostatistics was through Training Grant GM00038-21 from the National Institute of General Medical Sciences. I appreciate the opportunities which this support has afforded.

I wish to thank Dr. Norman E. Breslow for making data from the Children's Cancer Study Group A available, the Occupational Health Studies Group of the University of North Carolina for the data which they provided, and J. O. Kitchen for helping me adapt to my own needs. His program for evaluating $S_B$ parameters.

For the skill, equanimity and good-humored cooperation which Ann Hackney brought to the task of typing this manuscript under many changing deadlines I am most grateful. I also appreciate Gay Hinnant for her skillful typing of most of the tables in Chapters II and III.

For the sundry "aids" which were extended to me by family and friends during graduate school I express a warm appreciation.
LIST OF TABLES

2.1 SSI, Exact Distribution, k = 6, \( r'_{1} = [4 \ 3 \ 2 \ 1 \ 1 \ 1] \), 
\( r'_{2} = [4 \ 4 \ 4 \ 3 \ 2] \) ........................................ 44

2.2 SSI, Exact Distribution, k = 7, \( r'_{1} = [4 \ 4 \ 4 \ 3 \ 2 \ 2 \ 1] \), 
\( r'_{2} = [4 \ 3 \ 2 \ 2 \ 2 \ 1 \ 1] \) ........................................ 45

2.3 SSI, Exact Distribution, k = 7, \( r'_{1} = [4 \ 4 \ 4 \ 4 \ 3 \ 2] \), 
\( r'_{2} = [5 \ 4 \ 3 \ 2 \ 1 \ 1 \ 1] \) ........................................ 46

2.4 SSI, Exact Distribution, k = 8, \( r'_{1} = [4 \ 3 \ 3 \ 3 \ 3 \ 2 \ 2 \ 1] \), 
\( r'_{2} = [5 \ 5 \ 4 \ 3 \ 2 \ 2 \ 1 \ 1] \) ........................................ 47

2.5 SSI, Exact Distribution, k = 7, \( r'_{1} = [5 \ 4 \ 3 \ 3 \ 2 \ 2 \ 1] \), 
\( r'_{2} = [5 \ 5 \ 5 \ 4 \ 4 \ 3 \ 3] \) ........................................ 48

2.6 SSI, Sample Distribution, k = 18, Sample Size = 220 ........... 50

2.7 SSII, Exact Distribution, k = 7, \( r'_{11} = 3 \), \( r'_{21} = 4 \) and 
\( k = 8 \), \( r'_{11} = 4 \), \( r'_{21} = 4 \) ........................................ 54

2.8 SSII, Exact Distribution, k = 9, \( r'_{11} = 4 \), \( r'_{21} = 5 \) and 
\( k = 10 \), \( r'_{11} = 5 \), \( r'_{21} = 5 \) ........................................ 55

2.9 Chi-square Goodness of Fit Test Statistics, Exact Distributions, 
SSII ............................................................... 57

2.10 SSII, Sample Distribution, \( k = 20 \), \( r'_{11} = 10 \), \( r'_{21} = 10 \), 
Sample Size = 186 ........................................... 59
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.11</td>
<td>Kolmogorov-Smirnov One Sample Test Statistics $\sqrt{n} D_n$ SSII, Sample Distribution, Sample Size = 186</td>
<td>61</td>
</tr>
<tr>
<td>3.1</td>
<td>Example 1, Sample Estimates of HRF</td>
<td>67</td>
</tr>
<tr>
<td>3.2</td>
<td>Example 2, Sample Estimates of HRF</td>
<td>69</td>
</tr>
<tr>
<td>3.3</td>
<td>Example 3, Sample Estimates of HRF's</td>
<td>71</td>
</tr>
<tr>
<td>3.4</td>
<td>Some Parametric Families of Distributions</td>
<td>76</td>
</tr>
<tr>
<td>3.5</td>
<td>Summary of the Behavior of the HRF for Some Parametric Models from the Same Family of Distributions</td>
<td>78</td>
</tr>
<tr>
<td>3.6</td>
<td>Parameter Values of the HRF for Pairs of Lognormal Distributions</td>
<td>84</td>
</tr>
<tr>
<td>3.7</td>
<td>Parameter Values of the HRF for Pairs of Gamma Distributions</td>
<td>87</td>
</tr>
<tr>
<td>4.1</td>
<td>Summary of 200 Sample Estimates of &quot;Δ&quot; Under Some Specified Parametric Models, SS1</td>
<td>97</td>
</tr>
<tr>
<td>4.2</td>
<td>Summary of 200 Sample Estimates of &quot;Δ&quot; Under Some Specified Parametric Models, SSII</td>
<td>103</td>
</tr>
<tr>
<td>5.1</td>
<td>Summary of Data on Six Populations of Workers in a Major Tire Manufacturing Company, Akron, Ohio, 1964-1973</td>
<td>118</td>
</tr>
<tr>
<td>5.2</td>
<td>Estimates of Δ for Three Comparisons of Populations of Workers in a Major Tire Manufacturing Company, Akron, Ohio, 1964-1973</td>
<td>119</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

3.1 Males and Females, United States 1966 Population ............ 65
3.2 Nonwhite and White Males, United States 1966 Population .... 65
3.3 United States (Males) and England and Wales (Males), 1967
Population ........................................ 66
3.4 United States (Males) and West German (Males), 1967 Population . 66
3.5 Example 1, Plot of Estimated HRF ............................ 68
3.6 Example 2, Plot of Estimated HRF ............................ 70
3.7 Example 3, Plots of Estimated HRF's .......................... 73
3.8 HRF for Pairs of Weibull Distributions ...................... 80
3.9 HRF for Two Makeham Compertz Distributions ................ 81
3.10 HRF for Two Linear Hazard Distributions ................... 82
3.11 HRF for Two Lognormal Distributions ....................... 85
3.12 HRF for Two Truncated Logistic Distributions ............... 85
3.13 HRF for Two Gamma Distributions ........................... 87
3.14 HRF for Two Burr Distributions ............................. 88
4.1 Histograms of 200 Sample Estimates of $\Delta$ When $\Delta(t)$ is
Specified by Two Weibull Distributions, SSI ..................... 98
4.2 Histograms of 200 Sample Estimates of $\Delta$ When $\Delta(t)$ is
Specified by Two Compertz Distributions, SSI ................... 99
4.3 Histograms of 200 Sample Estimates of $\Delta$ When $\Delta(t)$ is
Specified by Two Weibull Distributions, SSI ................... 100
4.4 Histograms of 200 Sample Estimates of $\Delta$ When $\Delta(t)$ is
Specified by Two Weibull Distributions, SSII ................... 104
4.5 Histograms of 200 Sample Estimates of $\Delta$ When $\Delta(t)$ is Specified by Two Gompertz Distributions, SSII .............. 105

4.6 Histograms of 200 Sample Estimates of $\Delta$ When $\Delta(t)$ is Specified by Two Weibull Distributions, SSII .............. 106
CHAPTER I
INTRODUCTION AND REVIEW OF LITERATURE

1.1. Introduction.

The random variable of interest, T, say, in many epidemiological studies is the length of time elapsing until an individual experiences a well-defined event, often termed "death" or "failure". The distribution function of this random variable, i.e. the cumulative failure time distribution, its complement (often called the survival function), and the instantaneous rate of failure (generally termed the hazard function), are of both practical and theoretical importance. The survival function and the hazard function are frequently used to describe and summarize the results of follow-up studies in epidemiological investigations. These longitudinal or prospective studies (cohort studies) are basic designs in biomedical research. Implicit in the design is the definition of the population (cohort) to be followed for a specified period of time. Also, the outcome event which may occur in that population during the follow-up period must be well-defined. Incidence studies, clinical trials and intervention trials are important areas of application of follow-up studies.

Statistical methods for estimating the functions that describe survivorship depend on the manner in which measurements on the random variable T are obtained. Problems in the analysis of the data are briefly summarized into three groups. These are associated with

(i) censored observations,
(ii) competing risks,

and (iii) concomitant variables (or covariables).

Before describing these problems it is first necessary to distinguish between two periods of time: the study period and the follow-up period. The study period is the interval of time measured from the beginning until the termination of the study. It may be expressed as the actual calendar time or as the number of units of time, usually in months or years, of the study period.

The follow-up period is the length of time each individual is observed during the study period; the time he enters the study period is the beginning of his follow-up period, i.e. $t_0 = 0$. The survival information contributed by each individual during his follow-up period is used to estimate the survival function. The follow-up period for each individual is of variable length and is limited by the closing date of the study. In this case the survival time observation is (right) censored.

Intervening events, i.e. deaths from other causes (competing risks), may also remove the individual from observation before the study is completed or before the event of interest has occurred. Thus, censoring and competing risks produce partial information on survivorship for each individual.

On the other hand, incorporation of measurements on variables in addition to survival time poses different problems. It is usually assumed that survival time is some function of this additional information, or a function of certain concomitant variables.

Let $t_1 < t_2 < \cdots < t_k$ represent the observed (ordered) times of failure during the follow-up period. Without loss of generality we
will refer to them as the times of "death". We may define the risk set \( R_i \) (more precisely \( R_{t_i} \)) as the set of individuals alive and at risk of dying at \( t_i \), with \( r_i \) denoting the number of individuals in \( R_i \) (\( i = 1, \ldots, k \)). In general, \( R_i \) and \( R_{i+1} \) need not contain any members in common, e.g. when the distinct times of death represent data from different independent experiments. Estimation of the survivorship functions could be conditional on the underlying risk sets \( \{ R_i \} \). Such a situation occurs at the conclusion of a study when estimation is conditional on the observed risk set for each time of death. We will refer to estimation of the survivorship functions which is conditional on the risk sets as Sample Scheme I (SSI).

In most situations, however, such as follow-up studies, the set \( R_{i+1} \) represents the survivors out of \( R_i \) who were not censored or lost in \( [t_i, t_{i+1}) \). In these particular circumstances \( R_i \) and \( R_{i+1} \) do contain members in common. Estimation of the survivorship where the risk sets are considered as random variables will be called Sample Scheme II (SSII).

The remaining parts of this chapter will give a brief review of literature on methods available for estimating and comparing survival functions.

Special models based on the assumption of a proportional relationship between two hazard functions (the PHF model) which are frequently utilized in the estimation and comparison of survival functions are considered in Chapter II. Three new estimators of the parameter in the PHF model, their distributions, and their properties for Sample Scheme I and Sample Scheme II are given.
A discussion of some empirical evidence and theoretical conditions which illustrate the inappropriateness of the PHF model in certain circumstances is presented in Chapter III.

In Chapter IV some consequences of estimating the parameter in the PHF model when an alternate model is valid are given. Two indirect "tests" of the PHF model are suggested. Chapter V contains an example of the application of the methods of analysis of survival data investigated in this paper to ten years of follow-up data from a population of workers in the rubber industry.

1.2. Definitions.

We briefly review the basic definitions useful in survivorship analysis. Let $T$ be a positive continuous random variable representing the lifetime of an individual.

Cumulative Distribution Function (cdf)

$$F(t) = \Pr[ T \leq t ] = \Pr\{ \text{individual dies at or before time } t \}.\]

$$F(0) = 0; \quad F(\infty) = 1. \quad (1.2.1)$$

Survival Function

$$P(t) = \Pr[ T > t ] = \Pr\{ \text{individual survives longer than time } t \}.\]

$$P(0) = 1; \quad P(\infty) = 0. \quad (1.2.2)$$

Hazard Function

$$\mu(t) = \lim_{\Delta t \to 0} \frac{\Pr[ t < T \leq t + \Delta t | T > t ]}{\Delta t} = \frac{f(t)}{P(t)}, \quad (1.2.3)$$

$$= - \frac{d}{dt} \ln[P(t)], \quad (1.2.4)$$

where $f(t) = \frac{d}{dt} F(t) = - \frac{d}{dt} P(t)$ is the probability density function
(pdf) of the distribution function of times of death. The hazard function is also called the force of mortality and is, in fact, the instantaneous death rate.

From (1.2.4) we have the relations

$$P(t) = \exp[-\int_0^t \mu(\tau)d\tau].$$  \hspace{1cm} (1.2.5)

$$f(t) = \mu(t)\exp[-\int_0^\infty \mu(\tau)d\tau].$$  \hspace{1cm} (1.2.6)

It is only necessary to specify the form of any one of the functions to obtain the form of the other two.

1.3. Parametric Models.

Parametric models (Buckland, 1964; Johnson and Kotz, 1970) frequently applicable to the description of lifetime distributions will be discussed in Chapter III. A standard reference point is a constant hazard function, or \( \mu(t) = \mu \) for all \( t \); it is associated with the Poisson Process. A constant hazard function is equivalent to assuming deaths occur at random. The distribution of times of death is described by a (negative) exponential distribution. Monotone hazard functions correspond to always positive or negative wearing mechanisms. The linear hazard function model, Weibull, Makeham Gompertz, gamma, Burr, and (truncated) logistic families of distributions all have monotone (increasing or decreasing) hazard functions. For some distributions, e.g. the lognormal distribution, the hazard function is not monotone — it is increasing for some values of the variable and decreasing for others.


The classical approach to estimating survivorship functions has
been through the use of life table methods developed by actuaries. Life table functions which are estimated from information contributed by the cohort of individuals in a follow-up study are defined below.

Divide the follow-up time into \( \omega \) non-overlapping intervals, \([t_i, t_{i+1})\), each of width \( h_i \) (\( i = 0, \ldots, \omega - 1 \)), \( t_0 = 0 \).

\( \{\lambda_i\} \) is the nonincreasing sequence of values representing the expected numbers (or proportions) of individuals surviving to at least \( t_i \) out of \( \lambda_0 \) individuals alive at \( t_0 \) (when \( \lambda_0 = 1 \)).

More precisely, this should be \( \lambda_{t_i} \), but for convenience we will simply write the subscript \( i \) for all the life table functions that depend on \( t_i \) unless ambiguity arises, (e.g. \( \lambda_{t_i} \equiv \lambda_i \)). Notice that the set \( \{t_i\} \) now represents interval endpoints of failure rather than observed failure times.

\[
P_i = \frac{\lambda_i}{\lambda_0} = \Pr(T \geq t_i) = \Pr(\text{surviving to at least } t_i)
\]

\[
= \exp\left[-\int_0^{t_1} \mu(\tau) d\tau\right]
\]

\[
= \exp\left[-\int_0^{t_1} \mu(\tau) d\tau - \int_t^{t_2} \mu(\tau) d\tau - \ldots - \int_t^{t_{i-1}} \mu(\tau) d\tau\right]
\]

\[
= p_0 p_1 \cdots p_{i-1} = \prod_{j=0}^{i-1} p_j \tag{1.4.1}
\]

where
\[ p_i = \exp\left[-\int_{t_i}^{t_{i+1}} \mu(\tau) d\tau\right] \]  
\[ = \Pr\{T > t_{i+1} | T > t_i\} \]  
\[ = \frac{\ell_{i+1}}{\ell_i} = \frac{p_{i+1}}{p_i}, \]  

(1.4.2) (1.4.3)

that is, \( p_i \) is the conditional probability of surviving at least \( h_i \) additional years, given alive at \( t_i \). Its complement

\[ q_i = 1 - p_i = 1 - \exp\left[-\int_{t_i}^{t_{i+1}} \mu(\tau) d\tau\right], \]  

(1.4.4)

is the conditional probability of dying in the interval \([t_i, t_{i+1})\), given alive at \( t_i \). Expression (1.4.1) illustrates that the survival function can be represented as the product of the conditional probabilities of surviving successive intervals of time.

\[ d_i = \int_{t_i}^{t_{i+1}} \mu(\tau) \ell_i d\tau = \ell_i - \ell_{i+1}, \]  

(1.4.5)

where \( d_i \) is the expected number of deaths occurring in the interval \([t_i, t_{i+1})\).

\[ m_i = \frac{\int_{t_i}^{t_{i+1}} \mu(\tau) \ell_i d\tau}{\int_{t_i}^{t_{i+1}} \ell_i d\tau} = \frac{d_i}{\int_{t_i}^{t_{i+1}} \ell_i d\tau}, \]  

(1.4.6)

where \( m_i \) is the central death rate. The denominator of the function is the expected number of years lived between \( t_i \) and \( t_{i+1} \) by the \( \ell_i \) individuals alive at \( t_i \). Thus \( m_i \) is the expected number of deaths per person per unit of time. Precise actuarial notation would be \( p_{t_i} \) and \( m_{t_i} \). For convenience we will use \( p_i \) and \( m_i \).
Life table methods are primarily counting techniques which estimate the denominator of (1.4.3) or (1.4.6) using partial information on survival. These methods have been used in their simpler forms by epidemiologists since 1910 (Elderton and Perry, 1910, 1913; Weinberg, 1913).

Two approximations to the survival function which have been used with grouped data, and a method used with ungrouped data will be discussed. Their application to the analysis of follow-up data from epidemiological studies is cited.

1.4.1. Linear Approximation.

In analyzing a set of follow-up data grouped into intervals it may be reasonable to make the following assumptions.

(i) Times of death are uniformly distributed over each interval.

(ii) Each person leaving (or entering) the experience contributes, on the average, to one-half of the interval length of exposure.

(iii) Events in different intervals are independent.

A uniform distribution of deaths in each interval (i) is equivalent to assuming the survival function is linear in the interval. It may be necessary to group the data into intervals of unequal lengths in which this assumption holds (Merrell and Shulman, 1955). For example, infant mortality or post-operative mortality is approximately linear over small intervals of time, whereas later follow-up mortality is effectively linear over much larger intervals of time.

In 1933 Frost used the actuarial method and assumptions (i)-(iii) to obtain an estimator of \( m_1 \), the age-specific annual death rate, for a
study population exposed to pulmonary tuberculosis. The denominator of the estimator was defined in terms of person-years, or "risk time units" (one person exposed to the risk of dying for one year). In the 1950's the same methods were applied to estimate the conditional probability of dying, \( q_i \), using partial information (Berkson and Gage, 1950; Cutler and Ederer, 1958).

**Actuarial estimators of \( q_i \),**

\[ \tilde{q}_i = \frac{\text{Number of deaths in } [t_i, t_{i+1})}{\text{Number initially exposed to risk in } [t_i, t_{i+1})}, \]

and their corresponding large sample variances, have also been obtained. The earliest known derivation of the approximate standard error of the estimated survival function was given in 1926 by Greenwood (see also Irwin, 1949; Cutler and Ederer, 1958; Kaplan and Meier, 1958). For a review of actuarial methods used to estimate survivorship functions see Gehan (1969).

1.4.2. Piece-wise Exponential Approximation.

We now consider the situation where assumption (i) is replaced by the assumption that the hazard function \( \mu(t) \) is a constant \( \mu_i \) over the interval \([t_i, t_{i+1})\), \((i = 0, \ldots, \omega - 1)\). The number of deaths in an interval has a binomial distribution with the probability of death given by (1.4.4). Using this assumption the probability of death simplifies to \( q_i = 1 - \exp(-\mu_i h_i) \). A maximum likelihood (ML) estimator of \( \mu_i \) based on this assumption can be derived in terms of the MLE of \( p_i = 1 - q_i \) (Gehan, 1969). It is worth remarking at this point that other "MLE's" of \( \mu_i \) can also be derived under other assumptions, if (i) and (iii) hold (see Elveback (1958)).

Where there are no losses Chiang (1968) used the assumption of a
constant hazard function over each interval of follow-up time to estimate the survival function. He also divided the risk sets \( \{R_i\} \) into two independent populations: those "due" for censoring in the interval and those not "due" for censoring in the interval. When the number of deaths among those due to be censored is zero, his estimator of the conditional probability of survival, \( p_i \), is equivalent to the actuarial estimator. Also, when the number due to be censored and/or the hazard rate function in each interval is small, it has been shown that there is no practical difference between actuarial and Chiang's estimation of \( q_i \) (Kuzma, (1967)).

When the hazard function \( \mu_i \) is small there is essentially no difference between the estimators of the survival function obtained by linear approximation and by piece-wise exponential approximation. This is obvious, for when \( \mu_i \) is small the conditional probability of death in each interval is approximately uniform, i.e.

\[
q_i = 1 - \exp\left(-\mu_i h_i\right)
\]

\[
\approx \mu_i h_i .
\]

(1.4.7)

In this case the survival function is approximately linear over the interval.

1.4.3. Step-function Approximation.

If the times of each single death are available the survival function can also be approximated by the "empirical" distribution, i.e. a step-function with jumps at the times of death. Since the estimated conditional probability of survival is equal to one over intervals in which no deaths occur, it is only necessary to consider the intervals determined by the times of death. Kaplan and Meier (1958) called this
the product-limit (PL) estimator of the survival function. Turnbull
(1973) extended the PL estimator to doubly censored data, i.e. data
that is both left and right censored.

It should be noted that the three methods - life table with
linear approximation, life table with exponential approximation, and the
PL step-function approximation, have been developed for analysis of
survival data at the close of a study, conditionally on the risk sets
\( \{R_i\} \). These methods are in the frame of Sample Scheme I.

1.5. Survival Models with Concomitant Variables.

It is usually reasonable to assume that the probability of sur-
vival in the presence of a morbid or lethal condition is a function of
one or more variables in addition to time or age. These have been
termed explanatory or concomitant variables. The experimental design
of a follow-up study usually incorporates some such variables as control
factors. However, the analysis of survival data may need to use infor-
mation contributed by these variables, i.e.

(i) randomization procedures may fail to achieve a balance
    of some factors among populations,
(ii) it may be impractical to control for a large number of
    factors, or
(iii) the importance of the variables appears after the study
    design is implemented.

The use of models of survival which include concomitant variables
(also sometimes called regression models) or covariance analysis has
four main purposes.

(i) To inquire whether the explanatory variables are signifi-
    cant in the model (McGilchrist, 1974; Greenberg et al.,
1974).

(ii) To correct for lack of comparability with respect to the concomitant variables among groups whose survival is being compared (Feigl and Zelen, 1965; Glasser, 1967).

(iii) To select optimum treatment assignments for patients in a clinical trial (Aitchison, 1970; Byar, 1974).

(iv) To reduce the variance in the survival time (or response variable) (Armitage and Gehan, 1974).

1.5.1. Parametric Models with Concomitant Variables.

Adjustments of observed survival times to account for concomitant variables have primarily been developed by modeling the functional dependence of the hazard function (and therefore of the survival distribution) on the covariates. The simplest formulations have been for the exponential model. It has often been shown that this model describes survival following diagnosis of various types of cancer (Myers et al., 1966). Other models, notably the multivariate logistic, the Weibull, the log gamma and the Gompertz have been considered. The exponential is a special case of all these models except the logistic.

In 1965 Feigl and Zelen suggested an exponential model to describe survival of leukemia patients. Let \( z \) be the observed value of the random variable corresponding to white blood cell (WBC) count, and \( \beta_0 \) be the parameter for effect due to membership in a specific population. The expected survival time of each individual in that population is then assumed to be a linear function of the concomitant variable \( z \), i.e.

\[
E[T] = \mu^{-1} = \beta_0 + \beta_1 z
\]  

(1.5.1)
or

\[ \mu = (\beta_0 + \beta_1 z)^{-1}, \quad \mu > 0. \]  \hspace{1cm} (1.5.2)

Inclusion of the parameter \( \beta_1 \) allows the survival function to vary among individuals according to the value of each person's WBC count measured at some point, usually \( t_0 \). The likelihood function to be maximized is the product of exponential density functions where the parameter \( \mu = g(z; \beta_0, \beta_1) \) is a function of the covariable and parameters in (1.5.2). This model was extended by Zippen and Armitage (1966) to include censored observations. Byar et al. (1974) utilized the exponential model where \( \mu = \beta_0 + \beta_1 z \).

Concurrently Glasser (1967) suggested an exponential model to adjust censored survival data following surgery for lung cancer in two populations which differed with respect to their age distributions. Let \( z \) be the deviation of an individual's age from the average age of the total sample populations, and \( \mu_0 \) be the hazard function for an individual whose age is equal to the average age of the total populations. The mean survival of each individual is then assumed to be defined by

\[ E[T] = \mu^{-1}, \]  \hspace{1cm} (1.5.3)

where

\[ \mu = \mu_0 \exp(\beta z), \quad \mu, \mu_0 > 0. \]  \hspace{1cm} (1.5.4)

This type of model is usually termed a log-linear model. Maximum likelihood estimation of the parameters in this model, generalized to more than one covariable, has recently been given (Breslow, 1974a). A log-linear
model for each of a number of subintervals of the follow-up period has been suggested (Holford, 1972). This results in a piece-wise exponential model for the overall survival function of each individual.

The multivariate logistic function has been proposed (Truett et al., 1967; Walker and Duncan, 1967) as a model for fitting a dichotomous response variable (such as death or survival) to a number of independent covariables. In this model it is assumed that the unconditional probability of dying depends on \( m \) covariables \( z \) according to the relation

\[
Q(z) = [1 + \exp(-\beta'z)]^{-1}
= \exp(\beta'z)/[1 + \exp(\beta'z)],
\]

(1.5.5)

where \( \beta \) is the corresponding vector of parameters to be estimated. A generalization of the model to fit a categorical response variable with more than two outcomes has also been proposed (Mantel, 1966a). This model has the property that the logarithm of the ratio of the probability of dying to that of surviving is a linear function of the covariables, i.e.

\[
\log\left[\frac{Q(z)}{1 - Q(z)}\right] = \sum_{j=1}^{m} \beta_j z_j
\]

(1.5.6)

This model describes the response over varying units of time. It is possible to include the actual length of follow-up time in the model as one of the covariables (see Biometric Society Committee's Report on Trials of Hypoglycemic Agents, 1975); however, this implies that response time is a covariable.

Extensions of the multivariate logistic risk function to take into account individual response time information have also been
suggested (Myers et al., 1973). Let the follow-up period be divided into intervals of equal length \( h \). Assume each individual has the same constant force of mortality \( \mu \) per unit length of such intervals. The probability of survival over any unit is \( p = \exp(-\mu h) \). The probability of surviving each of the units preceding the unit in which an individual's first death occurs is given by a geometric distribution with parameter \( q = 1 - p = 1 - \exp(-\mu h) \). Further, it is assumed that the log odds of survival to death for each interval satisfies the **linear logistic function**

\[
\log \left( \frac{1 - p^h}{p^h} \right) = \sum_{j=1}^{m} \beta_j z_j,
\]

(1.5.7)

where \( z \) is a \( m \times 1 \) vector of covariates and \( \beta \) the corresponding vector of parameters. Rearrangement of (1.5.7) gives the hazard function of the exponential distribution in terms of the covariates, i.e.

\[
\mu = \frac{1}{h} \log [1 + \exp(\sum_{j=1}^{m} \beta_j z_j)].
\]

(1.5.8)

This model has been called the **logistic-exponential** regression model.

It can be shown that the limiting case when \( h \) is zero is a log-linear model; when \( h \) is infinite the hazard function is linear in the regressor variables (i.e. the exponential model where \( \mu = \sum_{j=1}^{m} \beta_j z_j \)). A suggestion for extending the model to permit individual variation in the hazard function was also given in this paper.

Other parametric models have been proposed which include concomitant variables in various functional relations to the survivorship functions. Theoretical results for fitting a normal distribution by weighted linear regression to the logarithm of a Gompertz hazard function have been described (Prentice and El Shaarawi, 1973).
For small class intervals and short follow-up periods age-specific death rates can be considered as estimating the hazard function. It was suggested that covariables related to age class could be included in an exponentially multiplicative manner in this model. Prentice (1973) has also suggested a Weibull model. This is one of several parametric forms which allows the hazard function to depend on time and on a set of covariables which are independent of time. Inclusion of covariables in a generalized log-gamma model were presented elsewhere (Prentice, 1974). Estimation and tests of hypotheses become rather complicated in these last two models.

Choice among models, particularly regarding the manner in which concomitant variables affect the survival function as expressed through the hazard function, has only recently received much attention. Fisher and Kanarek (1974) have proposed a metric to measure the "distance" between the survival curves of different models which are functions of only one covariable. Their primary concern is the size of the "difference" between models, not methods of discriminating between models. As yet they have published only preliminary results of their investigation. The choice between models is not only a question of which model fits the data best; it is also a question of which set of assumptions is the most reasonable representation of the physical process being studied.


In other recent work the comparison of survival functions has been approached using essentially "nonparametric" methods, but assuming the survival functions are in a specific (parametric) relationship to each other. In particular, the proportional hazard function (PHF) model has been assumed. For two populations, this specifies
\[ \mu_1(t) = \theta \mu_2(t), \quad \theta > 0, \quad (1.6.1) \]

where \( \mu_g(t) \) is the hazard function of population \( g = 1, 2 \) at time \( t \).

A "generalized proportional" relationship among hazard functions was assumed by Cox (1972),

\[ \mu(t; z) = \mu_0(t) \exp(z' \beta), \quad t > 0, \quad (1.6.2) \]

where \( z' = (z_1, z_2, \ldots, z_m) \) is a \( 1 \times m \) vector of concomitant variables for an individual, \( \beta \) is the corresponding vector of parameters, and \( \mu_0(t) \) is the (unknown) hazard function for the condition \( z = 0 \) or \( \beta = 0 \). The vector \( z \) may be zero-one indicator variables; in this model they may also be functions of time. If \( z \) is a function of time \( z(t) \) it must not involve either of the two hazard functions \( \mu(t; z) \) or \( \mu_0(t) \).

Cox obtained equations for the ML estimators of \( \beta \) based on a conditional likelihood function conditioned on the actual times of death. Consider the probability

\[
\Pr\{\text{individual (i) dies in } [t_i, t_i + dt_i]|\text{an individual} \}
\]

\[
(\ell) \epsilon R_i \text{ dies in } [t_i, t_i + dt_i)]
\]

\[
\left. \left. \left[ \mu_1(t_i) dt_i \right] \sum_{\ell \epsilon (R_i - i)} \left[ (1 - \mu_\ell(t_i) dt_i) \right] \right|_{[t_i, t_i + dt_i]} \right.
\]

\[
= \lim_{dt_i \to 0} \sum_{j \epsilon R_i} \left( \left[ \mu_j(t_i) dt_i \right] \left[ (1 - \mu_\ell(t_i) dt_i) \right] \right)
\]

\[
= \frac{\mu_1(t_i)}{\sum_{j \epsilon R_i} \mu_j(t_i)}
\quad (1.6.3)
\]

where the instantaneous probability of individual (i) dying in
\([t_i, t_i + dt_i]\), conditional on being alive at \(t_i\), is very nearly
\(\mu_i(t_i)dt_i, i = 1, \ldots, k\). Elimination of the nuisance parameters in
\(\mu_0(t_i)\) makes the likelihood depend only on \(\beta\) and \(z_i\).

In practice "ties" may result from measurement errors or slight

grouping of the data. The continuous model (1.6.2) becomes complicated
to apply, regardless of how one approaches estimating \(\beta\). Cox (1970)
himself defines a different model, called the "logistic" model, when
the information is recorded with more than one death occurring at \(t_i\).
This logistic model specifies the log odds of the instantaneous probabil-
ities of death relative to survival in the interval \([t, t + dt)\) as

\[
\frac{\mu(t; z)dt}{1 - \mu(t; z)dt} = \frac{\mu_0(t)dt}{1 - \mu_0(t)dt} \exp(z'\beta). 
\]  

The results obtained for estimating \(\beta\) using the conditional
likelihood are analogous to those appropriate for sampling without
replacement from the fixed finite populations of the risk sets \(\{R_i\}\),
in which the selection of an individual from \(R_i\) at each observed
(fixed) death time is an independent experiment. As Cox points out,
however, the risk set depends on what happened at previous time points
and on any intervening censoring.

Kalbfleisch and Prentice (1973) considered the same functional
relationship for the hazard function as that proposed by Cox (1.6.2).
Marginal likelihood estimators of \(\beta\) were obtained which depend only
on the information contained in the rank statistics of ordered times
of death. This approach, however, does not allow the covariables to be
functions of time. For both censored and uncensored data, expressions
for the estimators of \(\beta\) are equivalent to those given by Cox's con-
ditional likelihood when the covariables in Cox's model are those
permitted by Kalbfleisch and Prentice's underlying assumption, i.e. covariables which are not functions of time. However, generalization of the continuous model (1.6.2) to include ties results in estimators of $\beta$ which are not equivalent to those in Cox's logistic model.

When it is unreasonable to assume a covariable (or factor) has a multiplicative effect on the hazard function, Kalbfleisch (1974b) proposes that one assumes Cox's model is appropriate for each of the levels (or strata) of that covariable. The likelihood function is thus the product of the likelihood for each of the levels of the factor. For data including censored observations and covariables which are not functions of time, the rank order statistics of the times of death are also sufficient for the ML estimators of the parameters in this model.

1.7. Comparison of Survival Curves.

Comparison of survival curves is a very broad topic. In general comparisons have been made either in terms of tests on parameters in various parametric models or tests based on rank statistics. A few of the publications on this topic will be briefly mentioned. Results which are relevant to the discussion in Chapter IV will be given in more detail at that point.

Consider the model (1.6.2) where $z$ is a single indicator variable equal to 1 if an individual is from $\pi_1$ and equal to 0 if an individual is from $\pi_2$. Then

$$\mu(t; z) = \mu_0(t)e^{\beta z}. \quad (1.7.1)$$

This is a special case of the PHF model given by (1.6.1). This model, and methods which are concerned with estimating $\theta$ or $\beta$, have the advantage that data from each population contribute information to the
estimate of the hazard function in the other population. A dis-
advantage lies in the restrictions on the estimates of the parameters
in the model, i.e. they must be proportional. It should be noted that
the PHF model is equivalent to assuming the relationship of the two
survival functions is

\[ P_1(t) = [P_2(t)]^\theta. \]  \hspace{1cm} (1.7.2)

In hypothesis testing, for \( \theta \neq 1 \), this is a form of the well-known
Lehmann alternative.

Numerous standard nonparametric tests for comparison of survival
(distribution) functions have been extended to account for both censored
data and more than two groups (Halperin, 1960; Gehan, 1965a, 1965b;
Thomas, 1974). An evaluation of the power of several parametric and
nonparametric two-sample tests for small samples from exponential and
Weibull distributions, with and without censoring, has recently been
presented (Lee et al., 1974). The efficiency of Cox's model (1.6.2)
for the special case of the PHF model has been considered (Kalbfleisch,
1974a) for special cases: (1) assuming an underlying exponential
model and (2) assuming an underlying Weibull model with common shape
parameter, for the form of \( \mu_0(t) \). McRae and Thomas (1972) presented
both point estimation and tests of fit for two populations with the PHF
model and arbitrarily right censored data. Breslow (1974b) discusses
the PHF model in terms of its relationship to the analysis of survival
data using contingency tables. The literature cited reviews briefly
the main methods currently available for comparison of survival functions
using nonparametric tests.
CHAPTER II

ESTIMATION IN THE PHF MODEL

As indicated in Chapter I, the proportional hazard function (PHF) model, \( \mu_1(t) = \theta \mu_2(t) \), where \( \mu_g(t) \) is the hazard function for population \( g \) (\( g = 1,2 \)) at time \( t \), has received much attention as a model in the analysis of survival data. It has been used as a model to estimate and to compare survivorship functions, and to describe the relationship of non-time dependent covariables to survivorship functions. In this chapter, three new estimators and the MLE for a function of the parameter \( \theta \) are given. The properties and distributions of these estimators under Sample Scheme I and Sample Scheme II are also presented.

2.1. Estimators of the Ratio \( \theta = \mu_1(t)/\mu_2(t) \).

Consider the case of two populations denoted by \( \pi_1 \) and \( \pi_2 \). Let

\[ R_g(t) \text{ denote the risk set of } \pi_g \text{, or those members of population } \]

\[ (g) \text{ who survive until at least time } t, \ g = 1,2. \]

\[ r_g(t) \text{ denote the number of individuals in } R_g(t), \ g = 1,2. \]

When there is no confusion we will drop the \( t \) and simply write \( R_g \) and \( r_g \). Recall, by definition (1.2.5)

\[ P_g(t) = \exp(-\int_0^t \mu_g(\tau)d\tau), \quad g = 1,2. \]

For any model
\[ p = \Pr\{\text{member of } \pi_1 \text{ dies first} \mid R_1 \text{ and } R_2\} \quad (2.1.1) \]

\[ = \int_0^\infty \sum_{\ell \in R_1} \left\{ \left[ \prod_{n \in (R_1 - \ell)} P_{1n}(t) \prod_{m \in R_2} P_{2m}(t) \right] f_\ell(t) dt \right\}, \quad (2.1.2) \]

where \( P_{gj}(t) \) is the probability that the \( j^\text{th} \) individual in \( \pi_g \) survives to at least time \( t \), \( g = 1,2 \). This simplifies, if all members of any one population have the same survival function, to

\[ p = r_1 \int_0^\infty [P_1(t)]^{r_1-1} [P_2(t)]^{r_2} f_1(t) dt. \quad (2.1.3) \]

Using the relationships defined in (1.2.3) and (1.2.4) we obtain alternate expressions for (2.1.3)

\[ p = r_1 \int_0^\infty [P_1(t)]^{r_1} [P_2(t)]^{r_2} \mu_1(t) dt \quad (2.1.4) \]

\[ = r_1 \int_0^\infty [P_1(t)]^{r_1} [P_2(t)]^{r_2} \left\{- \frac{1}{P_1(t)} \frac{d}{dt} P_1(t) \right\} dt. \quad (2.1.5) \]

And so for the PHF model substituting \( P_1(t) = [P_2(t)]^{\theta} \), i.e., equation (1.7.2),

\[ p = r_1 \int_0^\infty [P_1(t)]^{r_1} [P_1(t)]^{r_2/\theta} \mu_1(t) dt \]

\[ = r_1 \int_0^\infty \left\{ \exp\left( -\int_0^t \mu_1(\tau) d\tau \right) (r_1 + r_2/\theta) \right\} \mu_1(t) dt \]

\[ = r_1 \int_0^\infty (-\theta (r_1 \theta + r_2))^{-1} \left[ \frac{d}{dt} \exp\left( -\int_0^t \mu_1(\tau) d\tau \right) (r_1 + r_2/\theta) \right] dt \]

\[ = -(r_1 \theta) (r_1 \theta + r_2)^{-1} \left[ \exp\left( -\int_0^t \mu_1(\tau) d\tau \right) (r_1 + r_2/\theta) \right]_0^\infty \]

\[ = \frac{r_1 \theta}{r_1 \theta + r_2}, \quad (2.1.6) \]

since

\[ \exp\left( -\int_0^\infty \mu_g(\tau) d\tau \right) = P_g(\infty) = 0, \]

\[ \exp\left( -\int_0^0 \mu_g(\tau) d\tau \right) = P_g(0) = 1, \quad g = 1,2. \]
Let

\( t_i \) denote the observed distinct (ordered) time of the \( i^{th} \)

independent death in the two populations, \( i = 1, \ldots, k, \)

\( t_1 < t_2 < \ldots < t_k. \)

\( R_{gi} \) denote the risk set of \( \pi_g \), or those members of population

\( (g) \) who survive up to time \( t_i, \ g = 1, 2, \ i = 1, \ldots, k. \)

\( r_{gi} \) denote the number of individuals in \( R_{gi}, \ g = 1, 2, \)

\( i = 1, \ldots, k. \)

Then by a similar argument to that leading to (2.1.6), under the PHF model

\[
p_i = \Pr\{\text{member of } \pi_1 \text{ dies first} \mid R_{1i}, R_{2i}\} = \frac{r_{1i} \theta}{r_{1i} \theta + r_{2i}}.
\]

(2.1.7)

A natural unbiased estimator of \( p_i \) is

\[
\hat{p}_i = \begin{cases} 
1 & \text{if a member of } \pi_1 \text{ dies first, given } r_{1i}, r_{2i}, \\
0 & \text{if a member of } \pi_2 \text{ dies first, given } r_{1i}, r_{2i}.
\end{cases}
\]

The estimator \( \hat{p}_i \) has a Bernoulli distribution with

\[
E[\hat{p}_i] = p_i = \frac{r_{1i} \theta}{r_{1i} \theta + r_{2i}}.
\]

\[
\text{Var}[\hat{p}_i] = p_i(1 - p_i) = \frac{\theta r_{1i} r_{2i}}{(r_{1i} \theta + r_{2i})^2}.
\]

\[
\text{Cov}[\hat{p}_i, \hat{p}_j] = 0 \text{ if the times of death are independent.}
\]
From equation (2.1.7)
\[ \theta = \left[ \frac{r_{2i}}{r_{1i}} \right] \frac{p_i}{1 - p_i}, \quad i = 1, \ldots, k. \] (2.1.8)

At each of the failure times \( \theta \) is defined by the corresponding parameter \( p_i \) and the numbers at risk \( r_{1i} \) and \( r_{2i} \); it is the weighted odds ratio of the probability an individual from population (1) dies first to the probability an individual from population (2) dies first. Substitution of the estimator \( \hat{p}_i \) in (2.1.8) would in fact give values of zero or infinity for \( \theta \). This problem can be avoided by combining the estimators of \( \theta \) at each of the failure points in the following way.

Expression (2.1.8) may be written as
\[ (r_{1i} \theta + r_{2i}) p_i = r_{1i} \theta. \] (2.1.9)

Let \( \{ \alpha_i \} \) be a set of arbitrary weights, usually positive, and let \( \sum \) denote \( \sum_{i=1}^{k} \). The sum of weighted values of (2.1.9) gives
\[ \sum_{i=1}^{k} \alpha_i (r_{1i} \theta + r_{2i}) p_i = \sum_{i=1}^{k} \alpha_i r_{1i} \theta, \] (2.1.10)

from which
\[ \theta = \frac{\sum \alpha_i r_{2i} p_i}{\sum \alpha_i r_{1i} (1 - p_i)}. \] (2.1.11)

Substituting \( \hat{p}_i \) in (2.1.11) gives an estimator of \( \theta \)
\[ \hat{\theta} = \frac{\sum \alpha_i r_{2i} \hat{p}_i}{\sum \alpha_i r_{1i} (1 - \hat{p}_i)}. \] (2.1.12)

\[ = \frac{\sum (1) \alpha_i r_{2i}}{\sum (2) \alpha_i r_{1i}}, \] (2.1.13)
where \( \sum_{(g)} \) is the sum over all time points among \( t_1, \ldots, t_k \) at which an individual from \( \pi_g \) dies, \( g = 1,2 \).

To avoid an estimator with possibly infinite moments, which would occur in the case of all deaths in \( \pi_1 \) preceding all deaths in \( \pi_2 \) (giving \( \tilde{\theta} = \infty \)), an arbitrary monotonic increasing function of \( \theta \) can be defined by

\[
\Delta = \frac{\theta}{1 + \theta}.
\]  

(2.1.14)

Notice that \( \Delta \) is the probability that, given one person chosen at random from \( \pi_1 \) and one from \( \pi_2 \), the person from \( \pi_1 \) dies first. It is in the bounded interval \([0,1]\).

We shall estimate \( \Delta \) by statistics of the form

\[
\tilde{\Delta} = \frac{\tilde{\theta}}{1 + \tilde{\theta}} = \frac{\sum_{1 \leq i < j < k} \alpha_{ij} r_{ij} \hat{p}_i}{\sum_{1 \leq i < j < k} \alpha_{ij} r_{ij} \hat{p}_i + \sum_{1 \leq i < j < k} \alpha_{ij} r_{ij} (1 - \hat{p}_i)}.
\]  

(2.1.15)

\[
= \frac{\sum_{1 \leq i < j < k} \alpha_{ij} r_{ij} \hat{p}_i}{\sum_{1 \leq i < j < k} \alpha_{ij} r_{ij} \hat{p}_i + \sum_{1 \leq i < j < k} \alpha_{ij} r_{ij} (1 - \hat{p}_i)}.
\]  

(2.1.16)

The choice of the set of weights \( \{\alpha_i\} \) will be considered in Section 2.2.2.

2.2. Estimators of \( \Delta \) and Their Properties Under Sample Scheme I.

Recall that Sample Scheme I (SSI) denotes estimation of \( \Delta \) (or \( \theta \)) which is conditional on the observed risk sets at every time of death, i.e., the numbers at risk are fixed values. We will now consider estimation of \( \Delta \) under SSI.

2.2.1. Approximation to First and Second Moments of \( \tilde{\Delta} \).

Under certain conditions, in SSI we can get a good approximation to the expected value and variance of \( \tilde{\Delta} \) by using statistical
differentials.

Let

\[ X = \sum \alpha_i r_{2i} \hat{p}_i, \]
\[ Y = \sum \alpha_i r_{1i} (1 - \hat{p}_i), \]
\[ Z = X + Y = \sum \alpha_i r_{2i} \hat{p}_i + \sum \alpha_i r_{1i} (1 - \hat{p}_i), \]
\[ \Delta = \frac{X}{Z}, \]
\[ \xi = \text{E}[X] = \sum \alpha_i r_{2i} p_i, \]
\[ \eta = \text{E}[Y] = \sum \alpha_i r_{1i} (1 - p_i), \]
\[ \sigma_X^2 = \text{Var}[X] = \sum \alpha_i^2 r_{2i}^2 p_i (1 - p_i), \]
\[ \sigma_Z^2 = \text{Var}[Z] = \sum \alpha_i^2 (r_{2i} - r_{1i})^2 p_i (1 - p_i), \]
\[ \sigma_{XZ} = \text{Cov}[X,Z] = \sum \alpha_i^2 r_{2i} (r_{2i} - r_{1i}) p_i (1 - p_i), \]
\[ \theta = \frac{\xi}{\eta}, \]
\[ \Delta = \frac{\theta}{1 + \theta} = \frac{\xi}{\xi + \eta}, \]

and

\[ \psi = \left[ \sum \alpha_i^2 (r_{2i} - r_{1i}) r_{1i} r_{2i} (r_{1i} \theta + r_{2i})^{-1} \right] \left[ \sum \alpha_i r_{1i} r_{2i} (r_{1i} \theta + r_{2i})^{-1} \right]^{-2}. \]

Then

\[ \text{E}[\Delta] = \Delta \left( 1 + \frac{\sigma_Z^2}{(\xi + \eta)^2} - \frac{\sigma_{XZ}}{\xi (\xi + \eta)} \right). \quad (2.2.1) \]

\[ \text{Var}[\Delta] = \left\{ \frac{\xi}{\xi + \eta} \right\} ^2 \left\{ \frac{\sigma_X^2}{\xi^2} + \frac{\sigma_Z^2}{(\xi + \eta)^2} - \frac{2 \sigma_{XZ}}{\xi (\xi + \eta)} \right\} \]
\[ = \left\{ \frac{\xi}{\xi + \eta} \right\} ^2 \left\{ \sum \alpha_i^2 p_i (1 - p_i) \left[ \frac{r_{2i}}{\xi} - \frac{(r_{2i} - r_{1i})}{(\xi + \eta)} \right] ^2 \right\} \]
\[
\hat{\Delta} = \left( \frac{\xi}{\xi + \eta} \right)^2 \left( \sum \alpha_i^2 r_{1i} r_{2i} (r_{1i}^2 + r_{2i}^2) \right)^{-2} \left[ \frac{\eta}{\xi (\xi + \eta)} \right]^2 (r_{1i}^2 + r_{2i}^2)^2
\]

\[
\approx \left( \frac{n \theta}{(\xi + \eta)^4} \right) \left( \sum \alpha_i^2 r_{1i} r_{2i} \right), \quad \text{(using } \xi = \theta \eta \text{).} \quad (2.2.2)
\]

\[
\text{Bias}[^\Delta] = E[^\Delta] - \Delta
\]

\[
\approx \Delta \left( \frac{\sigma^2}{(\xi + \eta)^2} - \frac{\sigma_{xz}}{\xi (\xi + \eta)} \right)
\]

\[
\approx \Delta \left( \sum \alpha_i^2 (r_{2i} - r_{1i}) p_i (1 - p_i) \left( \frac{r_{2i} - r_{1i}}{\xi (\xi + \eta)} \right)^2 - \frac{r_{2i}}{\xi (\xi + \eta)} \right)
\]

\[
\approx \Delta \left( - (\xi + \eta)^{-2} \sum \alpha_i^2 (r_{2i} - r_{1i}) r_{1i} r_{2i} (r_{1i}^2 + r_{2i})^{-1} \right)
\]

\[
\approx - \Delta (1 - \Delta)^2 \psi. \quad (2.2.3)
\]

2.2.2. Choice of Weights \{\alpha_i\}.

It is reasonable to choose the set of weights \{\alpha_i\} so that some optimal condition is achieved. One way of choosing \{\alpha_i\} is to try to minimize the (approximate) variance of \hat{\Delta} as given by (2.2.2) with respect to \alpha_i for fixed \{r_{1i}\}, \{r_{2i}\}. Although \hat{\Delta} is not necessarily unbiased (2.2.3), minimizing the (approximate) mean square error of \hat{\Delta} would yield no better results than minimizing the (approximate) variance since \text{[Bias(\hat{\Delta})]}^2 contains terms of order four. The approximation for the variance ignores terms of order greater than two.

Let

\[
A = \sum \alpha_i^2 r_{1i} r_{2i}.
\]
Thus

$$\text{Var}[\Delta] \doteq \frac{\eta \theta}{(\xi + \eta)^4} \Lambda$$

$$\frac{d}{d\alpha_i}[\text{Var}(\Delta)] \doteq \left\{ \frac{n\xi}{(\xi + \eta)^4} \right\} \left\{ 2\alpha_i r_{11} r_{2i} \right\} + \left\{ \frac{n\Lambda}{(\xi + \eta)^4} \right\} \left\{ r_{2i} p_i \right\}$$

$$+ \left\{ \frac{\xi A}{(\xi + \eta)^4} \right\} \left\{ r_{11} (1 - p_i) \right\} - \left\{ \frac{4n\xi}{(\xi + \eta)^5} \right\} \left\{ r_{11} (1 - p_i) + r_{2i} p_i \right\} \right.$$

$$\left. i = 1, \ldots, k. \right.$$ 

Set the above expression equal to zero and solve for $\alpha_i$.

$$\alpha_i \doteq \left\{ \frac{(\xi + \eta)^4}{(\xi + \eta)^4} \right\} \left\{ \frac{4n\xi}{(\xi + \eta)^5} \right\} \left\{ r_{11} (1 - p_i) + r_{2i} p_i \right\}$$

$$- \frac{n}{(\xi + \eta)^4} (r_{2i} p_i) - \frac{\xi}{(\xi + \eta)^4} \left\{ r_{11} (1 - p_i) \right\} \right.$$ 

$$\doteq \left\{ \frac{A}{2n\xi(\xi + \eta)} \right\} \left\{ 4n\xi r_{11} (1 - p_i) + 4n\xi r_{2i} p_i - (\xi + \eta) n r_{2i} p_i \right.$$ 

$$- (\xi + \eta) \xi r_{11} (1 - p_i) \right\}$$

$$\doteq \left\{ \frac{A}{2n\xi(\xi + \eta)} \right\} \left\{ 2\eta^2 r_{11} r_{2i} + 2\eta^2 \xi^2 r_{11} r_{2i} \right\} \left\{ r_{11} \theta + r_{2i} \right\}^{-1}$$

$$\text{(using } \xi = 0n),$$

$$\doteq \left\{ \frac{A(1 + \theta)}{(\xi + \eta)} \right\} \left\{ r_{11} \theta + r_{2i} \right\}^{-1}$$

$$\doteq \left\{ \frac{A}{\eta} \right\} \left\{ r_{11} \theta + r_{2i} \right\}^{-1}.$$ 

Thus we take

$$\alpha_i \propto (r_{11} \theta + r_{2i})^{-1} \quad (2.2.4)$$

2.2.3. Three Estimators of $\Delta$.

Optimal choice of the set of weights (2.2.4) would depend on the unknown parameter $\theta$, which is the ratio of the two hazard functions in
the PHF model. Three estimators are proposed based on different assumptions regarding $\theta$.

(i) Take $\{a_i\}$ constant, i.e., $a_i = a$ for $i = 1, \ldots, k$.

This assumption, which is equivalent to ignoring the question of weighting the estimators of $\theta$ at each of the times of death, gives

$$\hat{\theta}_1 = \frac{\sum r_{2i} \hat{p}_i}{\sum r_{2i} \hat{p}_i + \sum r_{1i} (1 - \hat{p}_i)}.$$  \hfill (2.2.5)

(ii) Take $\theta = 1$ in $\{a_i\}$, $i = 1, \ldots, k$.

This assumption is equivalent to supposing the two hazard functions are equal at each of the times of death. It is reasonable to suppose that we will often encounter situations in which $\theta$ is not too far from one, or else the difference in the two survivorship functions would be so clear as not to need experimentation. Estimators of $\theta$ and $\Delta$ are

$$\hat{\theta}_2 = \frac{\sum r_{2i} \hat{p}_i (r_{1i} + r_{2i})^{-1}}{\sum r_{1i} (1 - \hat{p}_i) (r_{1i} + r_{2i})^{-1}},$$  \hfill (2.2.6)

and

$$\hat{\Delta}_2 = \frac{\sum r_{2i} \hat{p}_i (r_{1i} + r_{2i})^{-1}}{\sum r_{2i} \hat{p}_i (r_{1i} + r_{2i})^{-1} + \sum r_{1i} (1 - \hat{p}_i) (r_{1i} + r_{2i})^{-1}},$$  \hfill (2.2.7)

respectively.

(iii) Take $\theta = \hat{\theta}_2$ in $\{a_i\}$, $i = 1, \ldots, k$.

This assumption allows us to "improve" our guess of $\Delta$ starting from a specified value of $\theta$. Of course, one can continue to "improve" the estimators of $\theta$ and $\Delta$ (see Section 2.2.4), but this iterative approach detracts from the simplicity of the proposed estimators. It will be seen that iteration is often unnecessary because $\hat{\theta}_2$ differs little in value from successive values in the iteration. With a second iteration we have
\[ \Delta_3 = \frac{\sum_{i=1}^{r_{21}} \hat{\theta}_i (r_{1i} \hat{\theta}_2 + r_{2i})^{-1}}{\sum_{i=1}^{r_{2i}} \hat{\theta}_i (r_{1i} \theta_2 + r_{2i})^{-1} + \sum_{i=1}^{r_{11}} (1 - \hat{\theta}_i)(r_{1i} \hat{\theta}_2 + r_{2i})^{-1}}, \]

(2.2.8)

where \( \hat{\theta}_2 \) is given by (2.2.6).

2.2.4. (Conditional) Maximum Likelihood Estimator of \( \Delta \).

A fourth estimator of \( \Delta \) to be considered is

\[ \hat{\Delta}_4 = \frac{\hat{\theta}}{1 + \hat{\theta}}, \]

(2.2.9)

where \( \hat{\theta} \) is the ML estimator which maximizes the conditional likelihood for the two population PHF model given by Cox (1972) (see Section 1.6, Chapter I).

Rewrite the conditional likelihood for all times of death under model (1.6.3) (see also (1.7.1)),

\[ L(\theta; z, r_1, r_2) = \prod_{i=1}^{k} \left\{ \frac{\theta^{z_i}}{(r_{1i} \theta + r_{2i})} \right\}, \]

(2.2.10)

where

\[ z_i = \begin{cases} 
1 & \text{if an individual from } \pi_1 \text{ dies first at } t_i, \\
0 & \text{if an individual from } \pi_2 \text{ dies first at } t_i. 
\end{cases} \]

Notice that the single indicator covariable \( z_i \) is equivalent to the estimator \( \hat{p}_i \) described in Section 1.1, i.e., \( z_i = \hat{p}_i \). Notice also that the conditional likelihood depends on the hazard function only at the times of death; the unconditional likelihood of a model describing times of death depends on the hazard function over the entire interval \([0, \infty)\), and hence implies assumptions about \( \mu(t) \) over the entire interval.

In the notation of Section 1.1, the likelihood conditional on all individuals in the two populations who are at risk of dying at \( t_i \) is
\[
L(\theta; \hat{p}_i, r_1, r_2) = \prod_{i=1}^{k} \frac{r_{1i}^{\hat{p}_i}}{r_{1i}^\theta + r_{2i}} \left\{ \frac{r_{2i}}{r_{1i}^\theta + r_{2i}} \right\}^{1-\hat{p}_i}
= \theta \sum_{i=1}^{k} \frac{\hat{p}_i}{(r_{1i}^\theta + r_{2i})^{-1}} \left\{ \frac{r_{1i}^\hat{p}_i}{r_{1i}^\theta + r_{2i}} \right\}^{(1-\hat{p}_i)}. \quad (2.2.11)
\]

Except for a constant the log likelihood of (2.2.11) is equivalent to the log of the conditional likelihood given by Cox (2.2.10). The MLE of \( \theta, \hat{\theta} \), is the value of \( \theta \) that satisfies the equation
\[
0 = \frac{\sum_{i=1}^{k} \hat{p}_i}{\theta} - \sum \left\{ \frac{r_{1i}}{r_{1i}^\theta + r_{2i}} \right\}. \quad (2.2.12)
\]

For SSI, from (2.2.11) we see that \( \sum \hat{p}_i \) is a sufficient statistic for \( \theta \). For SSII, however, this is not true since the vectors \( r_1 \) and \( r_2 \) are functions of the \( \hat{p}_i \)'s. Although not functions of a sufficient statistic, our estimators \( \hat{\theta}_k \) and \( \hat{\Delta}_k \) (\( k = 1, 2, 3 \)) of \( \theta \) and \( \Delta \), respectively, seem to be useful approximations to the MLE's of these parameters. In fact, repeated iterations give estimators tending to the MLE's. This is obvious, since the iterative equation given in (2.1.12) is easily shown to be equivalent to that for the ML estimator (2.2.12), i.e.,
\[
\theta = \frac{\sum \left\{ \frac{r_{2i}^\hat{p}_i}{r_{1i}^\theta + r_{2i}} \right\}}{\sum \left\{ \frac{r_{1i}(1 - \hat{p}_i)}{r_{1i}^\theta + r_{2i}} \right\}}, \quad \text{using} \quad \alpha_1 = (r_{1i}^\theta + r_{2i})^{-1}. \quad (2.2.13)
\]

\[
\theta \sum \left\{ \frac{r_{1i}}{r_{1i}^\theta + r_{2i}} \right\} - \theta \sum \left\{ \frac{r_{1i}^\hat{p}_i}{r_{1i}^\theta + r_{2i}} \right\} = \sum \left\{ \frac{r_{2i}^\hat{p}_i}{r_{1i}^\theta + r_{2i}} \right\}.
\]

\[
\theta \sum \left\{ \frac{r_{1i}}{r_{1i}^\theta + r_{2i}} \right\} = \sum \hat{p}_i,
\]

which gives
\[ 0 = \frac{\sum \hat{p}_i}{\theta} - \sum \left( \frac{r_{li}}{r_{li} \theta + r_{2i}} \right) \]

Also, the form of our estimators gives a rather clear picture of the relation of the estimators to the data.

2.2.5. Nearly Unbiased (NU) Estimators of \( \Delta \).

Nearly unbiased (NU) estimators of \( \Delta \) were also obtained.

Recall that

\[ E[\hat{\Delta}] = \Delta + \text{Bias}[\hat{\Delta}] \]

\[ \approx \Delta - \Delta(1 - \Delta)^2 \psi, \quad \text{using (2.2.3)} \]

\[ \approx \Delta[1 - (1 - \Delta)^2 \psi]. \]

Thus

\[ E\left[ \frac{\hat{\Delta}}{1 - (1 - \Delta)^2 \psi} \right] \approx \Delta. \]

Each estimator \( \hat{\Delta}_\ell \) (\( \ell = 1, 2, 3 \)) was divided by

\[ (1 - (1 - \Delta)^2 \psi), \quad (2.2.14) \]

with the corresponding estimators of \( \theta \) and \( \Delta \) substituted into (2.2.14), to obtain the NU estimators \( \hat{\Delta}'_\ell \) (\( \ell = 1, 2, 3 \)). The MLE \( \hat{\theta}_4 \) was also divided by (2.2.14), substituting \( \hat{\theta}_4 \) and \( \hat{\Delta}_4 \) for \( \theta \) and \( \Delta \) respectively, giving \( \hat{\Delta}'_4 \) since this estimator is unbiased only asymptotically. The properties of these NU estimators will be compared to the properties of the original estimators. A summary of these comparisons is given in Section 2.5.

2.3. Estimators of \( \Delta \) and Their Properties Under Sample Scheme II.

Recall that Sample Scheme II (SSII) denotes estimation (of \( \theta \) and \( \Delta \)) in which the risk sets \( R_{li} \) and \( R_{2i} \) are no longer fixed, but are subject to the observed natural decrement reflected by \( \{\hat{p}_i\} \),
In this sampling situation $\sum \hat{p}_i$ is not a sufficient statistic for $\theta$. At each time of death, with no censoring, the numbers of individuals in the risk sets $R_{1i}$ and $R_{2i}$, i.e., $r_{1i}$ and $r_{2i}$ respectively, are related to the $\hat{p}_i$'s by

$$r_{1i} = r_{1i} - \sum_{j<i} \hat{p}_j.$$  \hspace{1cm} (2.3.1)

$$r_{2i} = r_{2i} - \sum_{j<i} (1 - \hat{p}_j).$$  \hspace{1cm} (2.3.2)

We will now consider estimation of $\Delta$ under SSII.

2.3.1. Relationship of $\Delta$ to Mann-Whitney U-statistic.

Recall that $\Delta$ is the probability that, given one person chosen at random from $\pi_1$ and one from $\pi_2$, the person from $\pi_1$ dies first. We can show that the estimator $\hat{\Delta}_1$ of $\Delta$, under SSII, is in fact the standard Mann-Whitney (1947) U-statistic for estimating $\Delta$.

From (2.2.5), we have

$$\tilde{\Delta}_1 = \frac{\sum_i r_{2i} \hat{p}_i}{\sum_i r_{2i} \hat{p}_i + \sum_i r_{1i} (1 - \hat{p}_i)},$$

where $k = r_{1i} + r_{2i}$, and $\sum_i$ represents $\sum_{i=1}^{k}$.

Consider first the denominator $D$, say, of $\tilde{\Delta}_1$.

$$D = \sum_i r_{2i} \hat{p}_i + \sum_i r_{1i} (1 - \hat{p}_i)$$  \hspace{1cm} (2.3.3)

$$= \sum_i [r_{2i} - \sum_{j<i} (1 - \hat{p}_j)] \hat{p}_i + \sum_i [r_{1i} - \sum_{j<i} \hat{p}_j] (1 - \hat{p}_i).$$

Using

$$\sum_i \sum_{j<i} \hat{p}_j = \sum_i (r_{1i} + r_{2i} - i) \hat{p}_i,$$

and
\[
\sum_{i \neq j} \hat{p}_i \hat{p}_j \leq \left( \frac{r_{11}}{2} \right) = \frac{1}{4} r_{11} (r_{11} - 1),
\]

we have

\[
D = r_{11} r_{21} - \sum_{i} (i - 1) \hat{p}_i + \frac{1}{2} r_{11} (r_{11} - 1) + r_{11} r_{21}
- \sum_{i} (r_{11} + r_{21} - i) \hat{p}_i + \frac{1}{2} r_{11} (r_{11} - 1),
\]

so that

\[
D = r_{11} r_{21} .
\]

Thus,

\[
\Delta_1 = \frac{r_{11} r_{21} + \frac{1}{2} r_{11} (r_{11} - 1) - \sum_{i} (i - 1) \hat{p}_i}{r_{11} r_{21}} .
\]

The numerator of \( \Delta_1 \) can be shown to be a U-statistic.

Let

\( U_2 = \) the number of pairs for which \( t^{(1)} < t^{(2)} \), i.e., the
number of pairs of observations from \( \pi_1 \) and \( \pi_2 \) for
which the time of death of an individual in \( \pi_1 \), \( t^{(1)} \),
is less than the time of death of an individual in \( \pi_2 \), \( t^{(2)} \).

\( U_1 \) = the number of pairs for which \( t^{(1)} > t^{(2)} \),

\( W_g = \) the sum of the ranks of the times of death of members
of \( \pi_g \). \( W_g \) is the Wilcoxon rank sum statistic, \( g = 1, 2 \).

In the notation of the estimators for \( \Delta \), we notice that
\[
\sum_{i=1}^{n} \hat{p}_i = W_1 = U_1 + \frac{1}{2} r_{11} (r_{11} + 1) \\
= r_{11} r_{21} - U_2 + \frac{1}{2} r_{11} (r_{11} + 1),
\]
(2.3.6)

where  \( U_1 + U_2 = r_{11} r_{21} \).

Substituting the relationship (2.3.6) in (2.3.5) and simplifying, we obtain

\[
\hat{\Delta}_1 = \frac{U_2}{r_{11} r_{21}}.
\]
(2.3.7)

Thus, \( \frac{U_2}{r_{11} r_{21}} \) is the standard U-statistic for estimating \( \Delta \).

The statistic was proposed by Pitman (1948) and by Birnbaum (1956) (see also Hollander and Wolfe, 1973, pp. 68-77). Lehmann (1951) showed \( \hat{\Delta}_1 \) is the minimum variance unbiased estimator of \( \Delta \) over the class of continuous distributions. In addition, \( \hat{\Delta}_1 \) has all the properties of the Wilcoxon-Mann-Whitney U-statistics, one of which is asymptotic normality as \( \min(r_{11}, r_{21}) \to \infty \), i.e., that

\[
U_2 \sim N \left[ \frac{r_{11} r_{21} (r_{11} + r_{21} + 1)}{12} \right].
\]

The distribution of \( \hat{\Delta}_1 \) is asymptotically normal with mean \( \Delta \) and variance \( \frac{r_{11} + r_{21} + 1}{12 r_{11} r_{21}} \).

2.3.2. Other Estimators of \( \Delta \).

Although the estimators of \( \Delta \) proposed in Section 2.2 were developed for an optimal choice of the set of weights \( \{a_k\} \) under SSI, there was nothing in the discussion of \( \Delta \) in Section 2.1 which imposed any restrictions on the choice of \( \{a_k\} \). We have just seen that the unweighted estimator \( \hat{\Delta}_1 \) under SSII is an unbiased estimator of \( \Delta \). The estimators \( \hat{\Delta}_2, \hat{\Delta}_3 \) and \( \hat{\Delta}_4 \) given by (2.2.7), (2.2.8) and (2.2.9), respectively,
under SSII will also be considered. The distribution properties of all four estimators of \( \Delta \) will be presented in Section 2.5.

2.4. Extensions to Ties and Censoring.

2.4.1. Ties.

The estimators proposed in Section 2.1 can be extended to handle apparent "multiple deaths" (Cox's "multiplicities"). In statistical terms these represent ties. In practice they result from using a unit of time to record times of death which is larger than the smallest time between any two successive deaths.

Two estimators of \( \theta \) and \( \Delta \) are proposed which are reasonable when there are ties among the recorded times of death.

Let

\[
\begin{align*}
  r_{gi} &= \text{number of individuals from } \pi_g \text{ at risk of dying at } t_i, \ g = 1, 2. \\
  d_{gi} &= \text{number of individuals from } \pi_g \text{ who die in } [t_i, t_{i+1}), \\
           &\qquad g = 1, 2, i = 1, \ldots, k. \\
  d_i &= d_{1i} + d_{2i} = \text{total number of deaths observed in } [t_i, t_{i+1}), \ i = 1, \ldots, k; \ d_i \geq 0.
\end{align*}
\]

Consider the \( d_i \) subintervals in \( [t_i, t_{i+1}) \) formed by the times of each of the \( d_i \) deaths. Now let

\[
  d_i^{(n)} = \text{number of deaths from } \pi_g \text{ in } [t_i, t_{i+1}) \text{ which have occurred as of the time of the } n^{th} \text{ death, } n = 1, \ldots, d_i.
\]

Actually the superscript is \( n_i \), but for convenience we will write \( n \).
Notice that $d_{1i}^{(n)} + d_{2i}^{(n)} = n$.

For each of the $n$ unknown deaths in $[t_i, t_{i+1})$,

$$p_i^{(n)} = \Pr\{\text{member of } \pi_1 \text{ dies first as the } n^{th} \text{ death} \mid r_{1i} - d_{1i}^{(n-1)}, r_{2i} - d_{2i}^{(n-1)}\}$$

$$= \frac{[r_{1i} - d_{1i}^{(n-1)}] \theta}{[r_{1i} - d_{1i}^{(n-1)}] \theta + [r_{2i} - d_{2i}^{(n-1)}]}.$$

(2.4.1)

Linearizing (2.4.1) we have $n$ equations of the form

$$\theta[r_{1i} - d_{1i}^{(n-1)}] [1 - p_i^{(n)}] = [r_{2i} - d_{2i}^{(n-1)}] p_i^{(n)}.$$

(2.4.2)

We may sum the equations with respect to $n$ and define $\hat{\theta}$ as the solution of

$$\hat{\theta} \sum_{n=1}^{d_i} \{[r_{1i} - d_{1i}^{(n-1)}] [1 - \hat{p}_i^{(n)}]\} = \sum_{n=1}^{d_i} [r_{2i} - d_{2i}^{(n-1)}] \hat{p}_i^{(n)},$$

(2.4.3)

where

$$\hat{p}_i^{(n)} = \begin{cases} 1 & \text{if member of } \pi_1 \text{ dies first as the } n^{th} \text{ death.} \\ 0 & \text{if member of } \pi_2 \text{ dies first as the } n^{th} \text{ death.} \end{cases}$$

Simplifying (2.4.3) we have

$$\hat{\theta} r_{1i} d_{2i} - \hat{\theta}_1^{(2)} d_{1i}^{(n-1)} = r_{2i} d_{1i} - \sum_{n}^{(1)} d_{2i}^{(n-1)},$$

(2.4.4)

where $\sum_{n}^{(g)}$ is the sum over the $n = 1, \ldots, d_i$ subintervals in which a death from $\pi_g$ occurs, $g = 1, 2$.

Notice that $\sum_{n}^{(1)} d_{2i}^{(n-1)}$ is the sum over all times of death from $\pi_1$ of the number of previous deaths from $\pi_2$. This is in fact $U_2$, the Mann-Whitney U-statistic defined in Section 2.3.1, which is the sum
of all pairs of death times, one from \( \pi_1 \) and one from \( \pi_2 \), in which a member of \( \pi_1 \) dies before a member of \( \pi_2 \). Similarly, \( \sum_{n}^{(2)} d_{11}^{(n-1)} \) is \( U_1 \) defined in Section 2.3.1.

There are \( d_{1i}d_{2i} \) pairs of failure times; for each pair

\[
\Pr(\text{member of } \pi_1 \text{ of the pair dies first}) = \frac{\theta}{1 + \theta} = \Lambda. \quad (2.4.5)
\]

If we replace the summations in (2.4.4) by their respective expected values, i.e., put

\[
\sum_{n}^{(2)} d_{11}^{(n-1)} = E\left[ \sum_{n}^{(2)} d_{11}^{(n-1)} \right] = d_{1i}d_{2i} \frac{\theta}{1 + \theta},
\]

and

\[
\sum_{n}^{(1)} d_{21}^{(n-1)} = E\left[ \sum_{n}^{(1)} d_{21}^{(n-1)} \right] = d_{1i}d_{2i} \frac{1}{1 + \theta},
\]

we obtain

\[
\tilde{\theta}^2 d_{21} (r_{1i} - d_{1i}) + \tilde{\theta} (r_{1i} d_{2i} - r_{2i} d_{1i}) - d_{1i} (r_{2i} - d_{2i}) = 0. \quad (2.4.6)
\]

The solutions to (2.4.6) involve a quadratic equation in \( \tilde{\theta} \) for each \( i = 1, \ldots, k \).

Combining the \( k \) estimates of \( \tilde{\theta} \), using some arbitrary set of weights \( \{a'_i\} \), we have

\[
\tilde{\theta}^2 \sum_{i} a'_i d_{2i} (r_{1i} - d_{1i}) + \tilde{\theta} \sum_{i} a'_i (r_{1i} d_{2i} - r_{2i} d_{1i}) - \sum_{i} a'_i d_{1i} (r_{2i} - d_{2i}) = 0.
\]

(2.4.7)

The solutions to (2.4.7) are \( \tilde{\theta} = -1 \), which is obviously not possible (and, in fact was introduced by multiplying through by \( 1 + \tilde{\theta} \) to obtain (2.4.6)), and
\[ 0 = \frac{\sum_{i} \alpha_i d_{1i} (r_{2i} - d_{2i})}{\sum_{i} \alpha_i d_{2i} (r_{1i} - d_{1i})}. \]  

(2.4.8)

An estimator for \( \Delta \) is

\[ \hat{\Delta} = \frac{\sum_{i} \alpha_i d_{1i} (r_{2i} - d_{2i})}{\sum_{i} \alpha_i d_{1i} (r_{2i} - d_{2i}) + \sum_{i} \alpha_i d_{2i} (r_{1i} - d_{1i})}. \]  

(2.4.9)

For the special case of \( d_i = 1 \) for all \( i = 1, \ldots, k \), both (2.4.8) and (2.4.9) reduce to the respective estimators (2.1.12) and (2.1.15), where \( d_{1i} = \hat{p}_i \), \( d_{2i} = (1 - \hat{p}_i) \), and \( \alpha_i = \alpha \) for all \( i \). The choice of the appropriate set of weights when \( d_i > 1 \) for some \( i \) is left arbitrary at this time.

Within each interval \( [t_i, t_{i+1}) \) we are in a SSII situation. However, the above analysis applies whether we have a SSI or a SSII situation from one interval to another.

A second, more heuristic approach, for an estimator of \( \theta \) and of \( \Delta \) which allows for ties is now suggested. The form of the respective estimators is the same as that given for the single death case. The risk sets are adjusted to some "average" number at risk \( r_{1i}^* \) and \( r_{2i}^* \) over each interval, and \( \hat{p}_i = p_i = \frac{d_{1i}}{d_{1i} + d_{2i}} \). The estimators suggested are

\[ \hat{\theta}^* = \frac{\sum_{i} \alpha_i r_{2i}^* \hat{p}_i}{\sum_{i} \alpha_i r_{1i}^* (1 - \hat{p}_i)} \]  

(2.4.10)

and

\[ \hat{\Delta}^* = \frac{\sum_{i} \alpha_i r_{2i}^* \hat{p}_i}{\sum_{i} \alpha_i r_{1i}^* (1 - \hat{p}_i)}. \]  

(2.4.11)
Again, the choice of the weights \( \{\alpha^*_i\} \) is left arbitrary, and the "average" risk sets are not specified. These estimators do not reduce to the single death estimators given by (2.1.12) and (2.1.15), respectively.

2.4.2. Censoring.

In most follow-up studies there will be censored observations. The exact time of censoring may be known, or the interval in which censoring occurs may be all that is known (recorded). Ties may also occur among both the censored and the uncensored observations.

The expressions (2.4.8) and (2.4.9) may be modified to account for censored observations. We may subtract a proportion of the number of individuals censored in each population during each interval from the respective numbers initially at risk in \( \pi_1 \) and \( \pi_2 \) at \( t_i, i = 1, \ldots, k \).

It is also possible to find "interval" estimators of \( \theta \) and \( \Delta \) in the sense of lower and upper bounds which correspond to extreme assumptions about the times of censoring. That is, for an upper bound estimator we may assume the time of censoring of individuals in \( \pi_1 \) is at the beginning of each interval, and the time of censoring of individuals in \( \pi_2 \) is at the end of each interval. Reversing the assumptions in each population would give a lower bound estimator.

2.5. Distributions of \( \hat{\Lambda}_2 \) (\( \ell = 1, 2, 3 \)) and \( \hat{\Lambda}_4 \) and Some of Their Properties.

The discrete distributions and properties of the estimators of \( \Delta \) given in Sections 2.2 and 2.3 were investigated when the true value of \( \theta \) is assumed known. Nine values of \( \theta \) were considered,

\[ \theta' = [0.25 \ 0.50 \ 0.75 \ 0.90 \ 1.0 \ 1.1 \ 1.3 \ 2.0 \ 4.0], \] with
corresponding values of $\Delta$,

$$\Delta' = [0.20000, 0.33333, 0.42857, 0.47368, 0.50000, 0.52381, 0.56522, 0.66667, 0.80000].$$

Recall that $\Delta = \frac{\theta}{1 + \theta}$, or $\theta = \frac{\Delta}{1 - \Delta}$. Occasionally we may refer to $\theta$ and $\Delta$ interchangeably.

The exact distributions of the estimators of $\Delta$ proposed under SSI and SSII were generated for several cases. When the number of outcomes in the sample space became large a sample of the possible outcomes was drawn. A Monte Carlo simulation is presented in Chapter IV which investigated the large sample properties of the estimators in an approach somewhat different than the sampling study described in this section.

2.5.1. SSI (No Censoring).

Suppose we observe $k$ successive deaths at times $t_1 < t_2 < \ldots < t_k$ either from $\pi_1$ or from $\pi_2$, given that $r_1$ and $r_2$ are the $1 \times k$ vectors of values representing the numbers of individuals at risk of dying in $\pi_1$ and $\pi_2$, respectively, at each of the times of death. The only restriction on the vectors is that $r_{gi} \geq 1$ for all $g = 1, 2$, $i = 1, \ldots, k$.

Let the vector $Z' = [Z_1, \ldots, Z_k]$ be defined such that $Z_i = g$ if the $i^{th}$ death is from $\pi_g$, $g = 1, 2$, $i = 1, \ldots, k$. There are $2^k$ "possible vectors $Z$" which represent the sequences of populations in which deaths can occur. For illustration, let $k = 4$, and let the numbers at risk be given by $r_1' = [r_{11}, r_{12}, r_{13}, r_{14}]$ and $r_2' = [r_{21}, r_{22}, r_{23}, r_{24}]$, where each $r_{gi}$ is a specified value, $g = 1, 2$, $i = 1, \ldots, k$. The $2^4 = 16$ possible vectors are
For each of the $2^k$ possible vectors each of the four estimators of $\Delta$ described in Section 2.2, $\tilde{\Delta}_\ell$ ($\ell = 1,2,3$) and $\hat{\Delta}_4$, can be calculated. The corresponding nearly unbiased (NU) estimators $\Delta'_\ell$ ($\ell = 1,2,3$) and $\hat{\Delta}'_4$ can also be determined (2.2.14). Under the PHF model the probability of each of the $j = 1, \ldots, 2^k$ possible sequences is calculated from

$$v_j(\hat{p}; \theta) = \frac{k!}{\prod_{i=1}^{k} \left[ \frac{r_{1i} \theta}{r_{1i} \theta + r_{2i}} \right]^{\hat{p}_i} \left[ \frac{r_{2i}}{r_{1i} \theta + r_{2i}} \right]^{1-\hat{p}_i}}, \quad (2.5.1)$$

where

$$\hat{p}_i = \begin{cases} 1 & \text{if } i^{th} \text{ death from } \pi_1, \\ 0 & \text{if } i^{th} \text{ death from } \pi_2, \end{cases} \quad i = 1, \ldots, k.$$ 

For any specified value of $\theta$ we can evaluate (2.5.1) and find the moments of the distribution of the estimators of $\Delta$. The $\gamma^{th}$ moment about zero of any estimator $\hat{\Delta}$, say, is

$$u'_\gamma(\hat{\Delta}; \theta) = E[\hat{\Delta}^\gamma] = \sum_{j=1}^{N} \Delta_j v_j(\hat{p}; \theta), \quad (2.5.2)$$

where $N = 2^k$, $\Delta_j$ is the estimate of $\Delta$ for the $j^{th}$ sequence, and

$$\Delta\in\{\tilde{\Delta}_\ell, \hat{\Delta}_\ell, \Delta'_\ell \ (\ell = 1,2,3), \hat{\Delta}_4, \hat{\Delta}'_4\}.$$

**Small Samples**

When $k$ is small ($k \leq 8$, say) it is easy to determine the exact distribution of each of the eight estimators proposed for $\Delta$ under SSI.
Five examples are given in Tables 2.1–2.5 which illustrate the small sample properties of each of the eight estimators of \( \Delta \) for nine values of \( \theta \). The vectors of the numbers at risk in the two populations, \( r_1 \) and \( r_2 \), are given in Tables 2.1–2.5.

From these examples we can reasonably conclude:

1. All eight estimators are biased.
2. There is no clear pattern as to whether each estimator overestimates or underestimates \( \Delta \). This property seems to depend on the relative magnitudes of the numbers at risk, \( r_{11} \) and \( r_{21} \), at each time of death. All the original estimators seem to overestimate \( \Delta \) if \( r_1 > r_2 \) for all \( i = 1, \ldots, k \), and otherwise to underestimate \( \Delta \) for at least some values of \( \theta \).
3. \( \tilde{\Delta}_1 \) usually has the smallest bias.
4. \( \tilde{\Delta}_3 \) almost always has the smallest mean square error, MSE, where \( \text{MSE}[\hat{\Delta}] = \text{Var}[\hat{\Delta}] + \text{Bias}^2[\hat{\Delta}] \). \( \tilde{\Delta}_1 \) always has the largest MSE.
5. The MSE of the original estimators is usually smaller than the MSE of the corresponding NU estimators. This may not be true for all values of \( \theta \), but in general there is little reason to prefer an estimator which has a reduction in bias that is frequently accompanied by a much larger increase in the variance. If one had more information on the influence of the relationship of the numbers at risk it might be possible to decide when it is worthwhile using nearly unbiased estimators.
<table>
<thead>
<tr>
<th>θ</th>
<th>Δ</th>
<th>MEAN</th>
<th>SD</th>
<th>MSE</th>
<th>MEAN</th>
<th>SD</th>
<th>MSE</th>
<th>MEAN</th>
<th>SD</th>
<th>MSE</th>
<th>MEAN</th>
<th>SD</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>.25</td>
<td>.20000</td>
<td>.18316</td>
<td>.18271</td>
<td>.18447</td>
<td>.18389</td>
<td>.20023</td>
<td>.19585</td>
<td>.19774</td>
<td>.19713</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.20062</td>
<td>.19136</td>
<td>.19157</td>
<td>.19137</td>
<td>.21123</td>
<td>.20311</td>
<td>.20342</td>
<td>.20319</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.00401</td>
<td>.03692</td>
<td>.03694</td>
<td>.03688</td>
<td>.00408</td>
<td>.00412</td>
<td>.00413</td>
<td>.00412</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.50</td>
<td>.33333</td>
<td>.31519</td>
<td>.30902</td>
<td>.31047</td>
<td>.31001</td>
<td>.33439</td>
<td>.32846</td>
<td>.32906</td>
<td>.32955</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.23088</td>
<td>.21999</td>
<td>.21843</td>
<td>.21882</td>
<td>.23944</td>
<td>.22960</td>
<td>.22812</td>
<td>.22849</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.05327</td>
<td>.04899</td>
<td>.04823</td>
<td>.04843</td>
<td>.05734</td>
<td>.05274</td>
<td>.05205</td>
<td>.05222</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.75</td>
<td>.42857</td>
<td>.40836</td>
<td>.40161</td>
<td>.40222</td>
<td>.40205</td>
<td>.43042</td>
<td>.42406</td>
<td>.42483</td>
<td>.42461</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.23586</td>
<td>.22574</td>
<td>.22303</td>
<td>.22382</td>
<td>.24211</td>
<td>.23236</td>
<td>.22974</td>
<td>.23050</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.05604</td>
<td>.05169</td>
<td>.05044</td>
<td>.05080</td>
<td>.05862</td>
<td>.05401</td>
<td>.05279</td>
<td>.05315</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.90</td>
<td>.47368</td>
<td>.45302</td>
<td>.44619</td>
<td>.44626</td>
<td>.44627</td>
<td>.47594</td>
<td>.46956</td>
<td>.46980</td>
<td>.46977</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.23494</td>
<td>.22489</td>
<td>.22177</td>
<td>.22270</td>
<td>.23946</td>
<td>.22983</td>
<td>.22680</td>
<td>.22770</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.05562</td>
<td>.05133</td>
<td>.04993</td>
<td>.05035</td>
<td>.05734</td>
<td>.05284</td>
<td>.05145</td>
<td>.05186</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>.50000</td>
<td>.47254</td>
<td>.47261</td>
<td>.47212</td>
<td>.47227</td>
<td>.50230</td>
<td>.49616</td>
<td>.49605</td>
<td>.49614</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.23338</td>
<td>.22340</td>
<td>.21207</td>
<td>.21208</td>
<td>.23684</td>
<td>.22729</td>
<td>.22407</td>
<td>.22504</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.05490</td>
<td>.05067</td>
<td>.04921</td>
<td>.04964</td>
<td>.05610</td>
<td>.05168</td>
<td>.05022</td>
<td>.05066</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.10</td>
<td>.52381</td>
<td>.50305</td>
<td>.49627</td>
<td>.49565</td>
<td>.49591</td>
<td>.52652</td>
<td>.52020</td>
<td>.51983</td>
<td>.52003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.23133</td>
<td>.22142</td>
<td>.21796</td>
<td>.21800</td>
<td>.23381</td>
<td>.22435</td>
<td>.22099</td>
<td>.22201</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.05391</td>
<td>.04979</td>
<td>.04830</td>
<td>.04874</td>
<td>.05468</td>
<td>.05035</td>
<td>.04885</td>
<td>.04930</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.30</td>
<td>.56522</td>
<td>.54670</td>
<td>.53808</td>
<td>.53685</td>
<td>.53733</td>
<td>.56928</td>
<td>.56224</td>
<td>.56124</td>
<td>.56162</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.22636</td>
<td>.21659</td>
<td>.21296</td>
<td>.21406</td>
<td>.22710</td>
<td>.21779</td>
<td>.21429</td>
<td>.21555</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.05266</td>
<td>.04765</td>
<td>.04616</td>
<td>.04660</td>
<td>.05150</td>
<td>.04744</td>
<td>.04634</td>
<td>.04639</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.00</td>
<td>.66667</td>
<td>.64803</td>
<td>.64222</td>
<td>.63948</td>
<td>.64050</td>
<td>.67043</td>
<td>.66520</td>
<td>.66405</td>
<td>.66774</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.20657</td>
<td>.19722</td>
<td>.19374</td>
<td>.19475</td>
<td>.20313</td>
<td>.19423</td>
<td>.19093</td>
<td>.19189</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.04302</td>
<td>.03949</td>
<td>.03828</td>
<td>.03861</td>
<td>.04127</td>
<td>.03773</td>
<td>.03667</td>
<td>.03683</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.00</td>
<td>.80000</td>
<td>.78660</td>
<td>.78271</td>
<td>.77864</td>
<td>.78014</td>
<td>.80387</td>
<td>.80060</td>
<td>.79688</td>
<td>.79827</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.16298</td>
<td>.15448</td>
<td>.15251</td>
<td>.15283</td>
<td>.15523</td>
<td>.14709</td>
<td>.14527</td>
<td>.14558</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.02674</td>
<td>.02416</td>
<td>.02371</td>
<td>.02375</td>
<td>.02541</td>
<td>.02164</td>
<td>.02111</td>
<td>.02120</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$</td>
<td>$\Delta$</td>
<td>$\tilde{\Delta}_1$</td>
<td>$\tilde{\Delta}_2$</td>
<td>$\tilde{\Delta}_3$</td>
<td>$\tilde{\Delta}_4$</td>
<td>$\tilde{\Delta}_1'$</td>
<td>$\tilde{\Delta}_2'$</td>
<td>$\tilde{\Delta}_3'$</td>
<td>$\tilde{\Delta}_4'$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.25</td>
<td>.20000</td>
<td>MEAN .20723</td>
<td>.20679</td>
<td>.20826</td>
<td>.20774</td>
<td>.20097</td>
<td>.20033</td>
<td>.20178</td>
<td>.20128</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.15576</td>
<td>.14408</td>
<td>.14353</td>
<td>.14356</td>
<td>.15295</td>
<td>.14131</td>
<td>.14078</td>
<td>.14081</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.02431</td>
<td>.02080</td>
<td>.02067</td>
<td>.02067</td>
<td>.02339</td>
<td>.01997</td>
<td>.01982</td>
<td>.01983</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.50</td>
<td>.33333</td>
<td>MEAN .34224</td>
<td>.34258</td>
<td>.34366</td>
<td>.34328</td>
<td>.33516</td>
<td>.33428</td>
<td>.33533</td>
<td>.33496</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.18988</td>
<td>.17587</td>
<td>.17469</td>
<td>.17500</td>
<td>.18865</td>
<td>.17644</td>
<td>.17348</td>
<td>.17378</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.03615</td>
<td>.03102</td>
<td>.03062</td>
<td>.03072</td>
<td>.03559</td>
<td>.03050</td>
<td>.03001</td>
<td>.03020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.75</td>
<td>.42857</td>
<td>MEAN .43936</td>
<td>.43861</td>
<td>.43816</td>
<td>.43896</td>
<td>.43091</td>
<td>.42992</td>
<td>.43044</td>
<td>.43025</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.20344</td>
<td>.18857</td>
<td>.18722</td>
<td>.18760</td>
<td>.20353</td>
<td>.18883</td>
<td>.18728</td>
<td>.18767</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.04151</td>
<td>.03566</td>
<td>.03516</td>
<td>.03550</td>
<td>.04143</td>
<td>.03556</td>
<td>.03508</td>
<td>.03522</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.90</td>
<td>.47668</td>
<td>MEAN .48461</td>
<td>.48383</td>
<td>.48410</td>
<td>.48460</td>
<td>.47620</td>
<td>.47519</td>
<td>.47513</td>
<td>.47534</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.20707</td>
<td>.19200</td>
<td>.19065</td>
<td>.19104</td>
<td>.20777</td>
<td>.19265</td>
<td>.19130</td>
<td>.19170</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.04300</td>
<td>.03697</td>
<td>.03645</td>
<td>.03660</td>
<td>.04317</td>
<td>.03712</td>
<td>.03660</td>
<td>.03675</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>.50000</td>
<td>MEAN .51092</td>
<td>.51013</td>
<td>.51024</td>
<td>.51019</td>
<td>.50260</td>
<td>.50158</td>
<td>.50167</td>
<td>.50162</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.20837</td>
<td>.19233</td>
<td>.19191</td>
<td>.19230</td>
<td>.20941</td>
<td>.19422</td>
<td>.19290</td>
<td>.19329</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.04356</td>
<td>.03744</td>
<td>.03694</td>
<td>.03708</td>
<td>.04386</td>
<td>.03772</td>
<td>.03722</td>
<td>.03736</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.10</td>
<td>.52381</td>
<td>MEAN .53667</td>
<td>.53388</td>
<td>.53385</td>
<td>.53384</td>
<td>.52647</td>
<td>.52545</td>
<td>.52540</td>
<td>.52540</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.20902</td>
<td>.19387</td>
<td>.19259</td>
<td>.19296</td>
<td>.21037</td>
<td>.19515</td>
<td>.19387</td>
<td>.19424</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.04381</td>
<td>.03769</td>
<td>.03719</td>
<td>.03733</td>
<td>.04426</td>
<td>.03808</td>
<td>.03759</td>
<td>.03773</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.30</td>
<td>.56522</td>
<td>MEAN .57586</td>
<td>.57507</td>
<td>.57481</td>
<td>.57488</td>
<td>.56795</td>
<td>.56694</td>
<td>.56666</td>
<td>.56674</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.20895</td>
<td>.19385</td>
<td>.19267</td>
<td>.19301</td>
<td>.21079</td>
<td>.19561</td>
<td>.19464</td>
<td>.19477</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.04377</td>
<td>.03768</td>
<td>.03722</td>
<td>.03734</td>
<td>.04444</td>
<td>.03827</td>
<td>.03781</td>
<td>.03794</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.00</td>
<td>.66667</td>
<td>MEAN .67516</td>
<td>.67542</td>
<td>.67572</td>
<td>.67494</td>
<td>.66936</td>
<td>.66843</td>
<td>.66771</td>
<td>.66794</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.20189</td>
<td>.18741</td>
<td>.18662</td>
<td>.18681</td>
<td>.20473</td>
<td>.19014</td>
<td>.18935</td>
<td>.18954</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.04085</td>
<td>.03520</td>
<td>.03489</td>
<td>.03497</td>
<td>.04192</td>
<td>.03616</td>
<td>.03585</td>
<td>.03593</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.00</td>
<td>.80000</td>
<td>MEAN .80670</td>
<td>.80616</td>
<td>.80528</td>
<td>.80556</td>
<td>.80211</td>
<td>.80144</td>
<td>.80053</td>
<td>.80082</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.17447</td>
<td>.16205</td>
<td>.16194</td>
<td>.16189</td>
<td>.17794</td>
<td>.16540</td>
<td>.16529</td>
<td>.16524</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.03048</td>
<td>.02630</td>
<td>.02625</td>
<td>.02626</td>
<td>.03167</td>
<td>.02736</td>
<td>.02732</td>
<td>.02730</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>θ</td>
<td>A</td>
<td>Δ</td>
<td>Δ₁</td>
<td>Δ₂</td>
<td>Δ₃</td>
<td>Δ₄</td>
<td>Δ₁̂</td>
<td>Δ₂̂</td>
<td>Δ₃̂</td>
<td>Δ₄̂</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.25</td>
<td>.20000</td>
<td>MEAN</td>
<td>.20714</td>
<td>.12133</td>
<td>.21708</td>
<td>.21534</td>
<td>.13528</td>
<td>.19993</td>
<td>.20460</td>
<td>.20274</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.15252</td>
<td>.14304</td>
<td>.14091</td>
<td>.14128</td>
<td>.14792</td>
<td>.13876</td>
<td>.13678</td>
<td>.13703</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.02531</td>
<td>.02061</td>
<td>.02015</td>
<td>.02017</td>
<td>.02390</td>
<td>.01925</td>
<td>.01873</td>
<td>.01878</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.50</td>
<td>.33333</td>
<td>MEAN</td>
<td>.34241</td>
<td>.35000</td>
<td>.35326</td>
<td>.35206</td>
<td>.32769</td>
<td>.33480</td>
<td>.33768</td>
<td>.33674</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.19041</td>
<td>.18040</td>
<td>.17649</td>
<td>.17759</td>
<td>.18896</td>
<td>.17936</td>
<td>.17558</td>
<td>.17666</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.03634</td>
<td>.03282</td>
<td>.03155</td>
<td>.03189</td>
<td>.03574</td>
<td>.03217</td>
<td>.03085</td>
<td>.03122</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.75</td>
<td>.42857</td>
<td>MEAN</td>
<td>.43790</td>
<td>.44645</td>
<td>.44799</td>
<td>.44740</td>
<td>.42289</td>
<td>.43109</td>
<td>.43250</td>
<td>.43125</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.20627</td>
<td>.19618</td>
<td>.19189</td>
<td>.19217</td>
<td>.20327</td>
<td>.19758</td>
<td>.19336</td>
<td>.19463</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.04264</td>
<td>.03881</td>
<td>.03720</td>
<td>.03767</td>
<td>.04299</td>
<td>.03905</td>
<td>.03740</td>
<td>.03789</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.90</td>
<td>.47368</td>
<td>MEAN</td>
<td>.48286</td>
<td>.49163</td>
<td>.49232</td>
<td>.49203</td>
<td>.46811</td>
<td>.47661</td>
<td>.47718</td>
<td>.47693</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.21082</td>
<td>.20074</td>
<td>.19653</td>
<td>.19779</td>
<td>.21290</td>
<td>.20320</td>
<td>.19903</td>
<td>.20029</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.04453</td>
<td>.04062</td>
<td>.03897</td>
<td>.03946</td>
<td>.04536</td>
<td>.04130</td>
<td>.03963</td>
<td>.04013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>.50000</td>
<td>MEAN</td>
<td>.50902</td>
<td>.51784</td>
<td>.51805</td>
<td>.51793</td>
<td>.49453</td>
<td>.50312</td>
<td>.50322</td>
<td>.50314</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.21360</td>
<td>.20254</td>
<td>.19845</td>
<td>.19967</td>
<td>.21528</td>
<td>.20557</td>
<td>.20151</td>
<td>.20273</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.04528</td>
<td>.04134</td>
<td>.03971</td>
<td>.04019</td>
<td>.04638</td>
<td>.04227</td>
<td>.04062</td>
<td>.04111</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.10</td>
<td>.52381</td>
<td>MEAN</td>
<td>.53264</td>
<td>.54147</td>
<td>.54126</td>
<td>.54128</td>
<td>.51844</td>
<td>.52709</td>
<td>.52677</td>
<td>.52683</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.21363</td>
<td>.20362</td>
<td>.19968</td>
<td>.20085</td>
<td>.21684</td>
<td>.20714</td>
<td>.20322</td>
<td>.20338</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.04572</td>
<td>.04177</td>
<td>.04048</td>
<td>.04065</td>
<td>.04705</td>
<td>.04292</td>
<td>.04131</td>
<td>.04178</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.30</td>
<td>.56522</td>
<td>MEAN</td>
<td>.57362</td>
<td>.58236</td>
<td>.58148</td>
<td>.58174</td>
<td>.56008</td>
<td>.56870</td>
<td>.56771</td>
<td>.56800</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.21419</td>
<td>.20423</td>
<td>.20066</td>
<td>.20169</td>
<td>.21820</td>
<td>.20852</td>
<td>.20494</td>
<td>.20597</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.04595</td>
<td>.04200</td>
<td>.04053</td>
<td>.04095</td>
<td>.04764</td>
<td>.04349</td>
<td>.04201</td>
<td>.04243</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.00</td>
<td>.66667</td>
<td>MEAN</td>
<td>.67363</td>
<td>.68159</td>
<td>.67944</td>
<td>.68014</td>
<td>.66229</td>
<td>.67028</td>
<td>.66801</td>
<td>.66874</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.20817</td>
<td>.19856</td>
<td>.19627</td>
<td>.19684</td>
<td>.21358</td>
<td>.20420</td>
<td>.20190</td>
<td>.20247</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.04338</td>
<td>.03965</td>
<td>.03869</td>
<td>.03993</td>
<td>.04568</td>
<td>.04171</td>
<td>.04077</td>
<td>.04100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.00</td>
<td>.80000</td>
<td>MEAN</td>
<td>.80442</td>
<td>.81018</td>
<td>.80753</td>
<td>.80840</td>
<td>.79703</td>
<td>.80290</td>
<td>.80614</td>
<td>.80104</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>.18095</td>
<td>.17246</td>
<td>.17232</td>
<td>.17213</td>
<td>.19712</td>
<td>.17855</td>
<td>.17843</td>
<td>.17822</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>.03276</td>
<td>.02985</td>
<td>.02975</td>
<td>.02970</td>
<td>.03502</td>
<td>.03189</td>
<td>.03184</td>
<td>.03176</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 2.4

SSI, Exact Distribution, \( k = 8 \), \( \mathbf{r}_1 = [4 \ 3 \ 3 \ 3 \ 2 \ 2 \ 1] \), \( \mathbf{r}_2 = [5 \ 5 \ 4 \ 3 \ 2 \ 2 \ 1] \)

<table>
<thead>
<tr>
<th>Exact Values</th>
<th>Original Estimators</th>
<th>Nearly Unbiased Estimators</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta )</td>
<td>( \hat{\theta}_1 )</td>
<td>( \hat{\theta}_2 )</td>
</tr>
<tr>
<td>0.25</td>
<td>0.20000</td>
<td>0.19682</td>
</tr>
<tr>
<td>SD</td>
<td>0.10479</td>
<td>0.14185</td>
</tr>
<tr>
<td>MSE</td>
<td>0.03979</td>
<td>0.02012</td>
</tr>
<tr>
<td>0.50</td>
<td>0.33333</td>
<td>0.32835</td>
</tr>
<tr>
<td>SD</td>
<td>0.18302</td>
<td>0.16826</td>
</tr>
<tr>
<td>MSE</td>
<td>0.03352</td>
<td>0.02831</td>
</tr>
<tr>
<td>0.75</td>
<td>0.42857</td>
<td>0.42266</td>
</tr>
<tr>
<td>SD</td>
<td>0.19218</td>
<td>0.17705</td>
</tr>
<tr>
<td>MSE</td>
<td>0.03697</td>
<td>0.03135</td>
</tr>
<tr>
<td>0.90</td>
<td>0.47368</td>
<td>0.46747</td>
</tr>
<tr>
<td>SD</td>
<td>0.19382</td>
<td>0.17873</td>
</tr>
<tr>
<td>MSE</td>
<td>0.03761</td>
<td>0.03194</td>
</tr>
<tr>
<td>1.00</td>
<td>0.50000</td>
<td>0.49366</td>
</tr>
<tr>
<td>SD</td>
<td>0.19601</td>
<td>0.17899</td>
</tr>
<tr>
<td>MSE</td>
<td>0.03768</td>
<td>0.03204</td>
</tr>
<tr>
<td>1.10</td>
<td>0.52381</td>
<td>0.51738</td>
</tr>
<tr>
<td>SD</td>
<td>0.19709</td>
<td>0.17979</td>
</tr>
<tr>
<td>MSE</td>
<td>0.03756</td>
<td>0.03197</td>
</tr>
<tr>
<td>1.30</td>
<td>0.56522</td>
<td>0.55872</td>
</tr>
<tr>
<td>SD</td>
<td>0.19205</td>
<td>0.17741</td>
</tr>
<tr>
<td>MSE</td>
<td>0.03693</td>
<td>0.03148</td>
</tr>
<tr>
<td>2.00</td>
<td>0.66667</td>
<td>0.66042</td>
</tr>
<tr>
<td>SD</td>
<td>0.18199</td>
<td>0.16838</td>
</tr>
<tr>
<td>MSE</td>
<td>0.03312</td>
<td>0.02835</td>
</tr>
<tr>
<td>4.00</td>
<td>0.80000</td>
<td>0.79518</td>
</tr>
<tr>
<td>SD</td>
<td>0.15304</td>
<td>0.14212</td>
</tr>
<tr>
<td>MSE</td>
<td>0.02345</td>
<td>0.02020</td>
</tr>
</tbody>
</table>
**TABLE 2.5**

SSI, EXACT DISTRIBUTION, \( k = 7, \epsilon_2 = \{5 4 3 2 2 1\}, \epsilon_2 = \{5 5 5 4 3 3\} \)

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>( \Delta )</th>
<th>( \hat{\alpha}_1 )</th>
<th>( \hat{\alpha}_2 )</th>
<th>( \hat{\alpha}_3 )</th>
<th>( \hat{\alpha}_4 )</th>
<th>( \hat{\alpha}_5 )</th>
<th>( \hat{\alpha}_6 )</th>
<th>( \hat{\alpha}_7 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>.25</td>
<td>.20000</td>
<td>.19220</td>
<td>.18988</td>
<td>.19076</td>
<td>.19048</td>
<td>.20027</td>
<td>.19797</td>
<td>.19889</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.17435</td>
<td>.16882</td>
<td>.16876</td>
<td>.16870</td>
<td>.18025</td>
<td>.17472</td>
<td>.17467</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>.03046</td>
<td>.02860</td>
<td>.02857</td>
<td>.02855</td>
<td>.03249</td>
<td>.03053</td>
<td>.03051</td>
</tr>
<tr>
<td>.50</td>
<td>.33333</td>
<td>.32252</td>
<td>.31927</td>
<td>.31997</td>
<td>.31975</td>
<td>.33415</td>
<td>.33099</td>
<td>.33172</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.20033</td>
<td>.19386</td>
<td>.19302</td>
<td>.19322</td>
<td>.20475</td>
<td>.19832</td>
<td>.19751</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>.04025</td>
<td>.03777</td>
<td>.03744</td>
<td>.03752</td>
<td>.04192</td>
<td>.03934</td>
<td>.03901</td>
</tr>
<tr>
<td>.75</td>
<td>.42857</td>
<td>.41669</td>
<td>.41310</td>
<td>.41336</td>
<td>.41329</td>
<td>.42987</td>
<td>.42642</td>
<td>.42671</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.20598</td>
<td>.19908</td>
<td>.19779</td>
<td>.19816</td>
<td>.20865</td>
<td>.20187</td>
<td>.20059</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>.04257</td>
<td>.03987</td>
<td>.03933</td>
<td>.03950</td>
<td>.04354</td>
<td>.04075</td>
<td>.04024</td>
</tr>
<tr>
<td>.90</td>
<td>.47368</td>
<td>.46161</td>
<td>.45795</td>
<td>.45793</td>
<td>.45795</td>
<td>.47322</td>
<td>.47172</td>
<td>.47174</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.20561</td>
<td>.19856</td>
<td>.19711</td>
<td>.19754</td>
<td>.20735</td>
<td>.20044</td>
<td>.19900</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>.04242</td>
<td>.03967</td>
<td>.03910</td>
<td>.03927</td>
<td>.04300</td>
<td>.04018</td>
<td>.03960</td>
</tr>
<tr>
<td>1.00</td>
<td>.50000</td>
<td>.48791</td>
<td>.48423</td>
<td>.48403</td>
<td>.48412</td>
<td>.50167</td>
<td>.49817</td>
<td>.49801</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.20455</td>
<td>.19743</td>
<td>.19590</td>
<td>.19635</td>
<td>.20573</td>
<td>.19876</td>
<td>.19724</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>.04199</td>
<td>.03923</td>
<td>.03863</td>
<td>.03881</td>
<td>.04233</td>
<td>.03951</td>
<td>.03891</td>
</tr>
<tr>
<td>1.10</td>
<td>.52381</td>
<td>.51176</td>
<td>.50809</td>
<td>.50772</td>
<td>.50787</td>
<td>.52560</td>
<td>.52211</td>
<td>.52179</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.20307</td>
<td>.19589</td>
<td>.19430</td>
<td>.19478</td>
<td>.20773</td>
<td>.19671</td>
<td>.19515</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>.04138</td>
<td>.03862</td>
<td>.03801</td>
<td>.03819</td>
<td>.04151</td>
<td>.03870</td>
<td>.03809</td>
</tr>
<tr>
<td>1.30</td>
<td>.56522</td>
<td>.55317</td>
<td>.54975</td>
<td>.54908</td>
<td>.54933</td>
<td>.56720</td>
<td>.56378</td>
<td>.56315</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.19913</td>
<td>.19208</td>
<td>.19042</td>
<td>.19091</td>
<td>.19909</td>
<td>.19201</td>
<td>.19038</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>.03987</td>
<td>.03713</td>
<td>.03652</td>
<td>.03670</td>
<td>.03964</td>
<td>.03687</td>
<td>.03625</td>
</tr>
<tr>
<td>2.00</td>
<td>.66667</td>
<td>.65664</td>
<td>.65375</td>
<td>.65332</td>
<td>.65184</td>
<td>.66899</td>
<td>.66593</td>
<td>.66538</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.18190</td>
<td>.17658</td>
<td>.17498</td>
<td>.17562</td>
<td>.18154</td>
<td>.17461</td>
<td>.17287</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>.03393</td>
<td>.03137</td>
<td>.03085</td>
<td>.03099</td>
<td>.03296</td>
<td>.03042</td>
<td>.02989</td>
</tr>
<tr>
<td>4.00</td>
<td>.80000</td>
<td>.79246</td>
<td>.79008</td>
<td>.78799</td>
<td>.78787</td>
<td>.80229</td>
<td>.80015</td>
<td>.79816</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.14877</td>
<td>.14180</td>
<td>.14082</td>
<td>.14096</td>
<td>.14430</td>
<td>.13752</td>
<td>.13660</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>.02219</td>
<td>.02020</td>
<td>.01998</td>
<td>.02000</td>
<td>.02083</td>
<td>.01891</td>
<td>.01866</td>
</tr>
</tbody>
</table>
Large Samples

As the observed number of successive deaths from \( \pi_1 \) and \( \pi_2 \) increases \((k > 8, \text{say})\), generation of the exact distribution of the eight estimators of \( \Delta \) for the \( 2^k \) possible sequence of populations in which the deaths occur soon becomes economically prohibitive. Some results are presented which come from a preliminary sampling study of the properties of the eight estimators of \( \Delta \) using the following example.

Example: Suppose we have observed \( k = 18 \) deaths from \( \pi_1 \) and \( \pi_2 \), where the numbers at risk in the two populations at each time a death occurred are

\[
\begin{align*}
\mathbf{r}_1' & = [10 \ 9 \ 8 \ 7 \ 6 \ 6 \ 5 \ 5 \ 4 \ 4 \ 3 \ 3 \ 3 \ 3 \ 3 \ 2 \ 2], \quad (2.5.3) \\
\mathbf{r}_2' & = [10 \ 10 \ 10 \ 10 \ 9 \ 8 \ 8 \ 7 \ 7 \ 6 \ 6 \ 5 \ 5 \ 4 \ 3 \ 2 \ 2 \ 1]. \quad (2.5.4)
\end{align*}
\]

For each of the \( j = 1, \ldots, 2^{18} \) sequences a random number \( \epsilon \) was generated from the uniform distribution over the interval \([0,1]\). If \( \epsilon \leq 0.00085 \) that sequence would be selected for inclusion in the sample. The sampling fraction \( 0.00085 \) was chosen so that the sample size \( (n) \) from the 262,144 sequences was approximately 200.

Eight estimates of \( \Delta \) corresponding to each of the \( n \) vectors of populations selected for the sample were determined. The \( \gamma \)th moments about zero were also found for \( \gamma = 1,2, \text{i.e.} \)

\[
u'_\gamma(\Delta; \theta) = E[\Delta^\gamma] = \sum_{j=1}^{n} \Delta_j v_j(\hat{p}_j; \theta)/\sum_{j=1}^{n} v_j(\hat{p}_j; \theta). \quad (2.5.5)
\]

Five samples (of sizes 220, 224, 210, 219, and 243) were drawn for the risk sets given in (2.5.3) and (2.5.4). Table 2.6 presents the summary statistics for one of these samples. There is no evidence that the NU estimators are to be preferred to the original estimators, and \( \Delta_3 \)
<table>
<thead>
<tr>
<th>Exact Values</th>
<th>Original Estimates</th>
<th>Nearly Unbiased Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta )</td>
<td>( \Delta )</td>
<td>( \hat{\Delta}_1 )</td>
</tr>
<tr>
<td>.25</td>
<td>.20000</td>
<td>( \text{MEAN} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{SD} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{MSE} )</td>
</tr>
<tr>
<td>.50</td>
<td>.33333</td>
<td>( \text{MEAN} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{SD} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{MSE} )</td>
</tr>
<tr>
<td>.75</td>
<td>.42857</td>
<td>( \text{MEAN} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{SD} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{MSE} )</td>
</tr>
<tr>
<td>.90</td>
<td>.47368</td>
<td>( \text{MEAN} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{SD} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{MSE} )</td>
</tr>
<tr>
<td>1.00</td>
<td>.50000</td>
<td>( \text{MEAN} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{SD} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{MSE} )</td>
</tr>
<tr>
<td>1.10</td>
<td>.52381</td>
<td>( \text{MEAN} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{SD} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{MSE} )</td>
</tr>
<tr>
<td>1.30</td>
<td>.56522</td>
<td>( \text{MEAN} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{SD} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{MSE} )</td>
</tr>
<tr>
<td>2.00</td>
<td>.66667</td>
<td>( \text{MEAN} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{SD} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{MSE} )</td>
</tr>
<tr>
<td>4.00</td>
<td>.80000</td>
<td>( \text{MEAN} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{SD} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{MSE} )</td>
</tr>
</tbody>
</table>
usually has the smallest MSE. It seems reasonable that the estimators are biased for large samples as they were shown to be for small samples. For these five examples, however, we are unable to distinguish true bias from bias introduced by sampling variability.

It was of interest to test two null hypotheses

\[ H_0^{(1)}: \text{A beta distribution describes the distribution of the estimator } \hat{\Lambda}, \]

\[ H_0^{(2)}: \text{A Johnson } S_B \text{ distribution describes the distribution of the estimator } \hat{\Lambda}, \]

(2.5.6)

(2.5.7)

where \( \Delta \in \{ \hat{\Lambda}_x, \hat{\Lambda}'_x, \hat{\Lambda}_4, \hat{\Lambda}'_4 \} \).

Two test statistics were calculated. Using the method of moments, the parameters of the beta (Elderton and Johnson, 1969) and the \( S_B \) distributions (Johnson and Kitchen, 1971) were estimated. These estimates are available in manuscript.

A Kolmogorov-Smirnov (KS) one-sample test statistic was calculated for each estimator at each of the nine values of \( \theta \). Critical values were used which are appropriate for large samples. The test is conservative when the parameters are unknown and must be estimated from the sample data. There is evidence to reject both \( H_0^{(1)} \) and \( H_0^{(2)} \).

Chi-square goodness of fit tests using ten intervals of width .1 were also calculated. These test statistics, each with seven degrees of freedom, also rejected both \( H_0^{(1)} \) and \( H_0^{(2)} \).

It is not necessarily surprising to find these two continuous distributions fail to closely represent the discrete distributions of the estimators of \( \Lambda \). Although the sample sizes on which the test statistics were calculated are reasonably large (about 200 observations
per sample), the number of distinct values the estimators can assume is much smaller. Under SSI the number of deaths which occur in \( \pi_1 \), \( \sum_{i=1}^{k} \hat{p}_i \) is a sufficient statistic. There are only \((k + 1)\) distinct values which the ML estimator \( \hat{\Delta}_4 \) (and hence \( \Delta_4' \)) may assume. The other estimators \( \{\hat{\Delta}_k, \Delta_k' \) \( (k = 1, 2, 3)\) also assume less than \( 2^k \) (or \( n \)) distinct values. Until \( k \) becomes very large no simple continuous distribution is likely to closely approximate the distributions of the estimators.

2.5.2. SSII (No Censoring).

Suppose that from some initial time \( t_0 \) we follow \( r_{11} \) and \( r_{21} \) individuals from \( \pi_1 \) and \( \pi_2 \), respectively, until all of the \( k = r_{11} + r_{21} \) individuals die. Let \( t_1 < t_2 < \ldots < t_k \) represent the ordered times at which the \( k \) deaths are observed during the follow-up period. The numbers at risk of dying in each population at the time a death is observed are random variables, i.e.,

\[
\begin{align*}
r_1' &= [r_{11} \ r_{12} \ \ldots \ r_{1k}], \\
r_2' &= [r_{21} \ r_{22} \ \ldots \ r_{2k}],
\end{align*}
\]

where

\[
\begin{align*}
r_{1i} &= r_{1i} - \sum_{j<i} \hat{p}_j, \\
r_{2i} &= r_{2i} - \sum_{j<i} (1 - \hat{p}_j).
\end{align*}
\]

The vectors \( r_1 \) and \( r_2 \) depend on the number initially at risk, \( r_{gi} \), \( g = 1, 2 \), and on the populations from which deaths have occurred previously. As in Section 2.5.1, let the vector \( Z \) be defined such that each element \( Z_i = g \) if the \( i \)th death is from \( \pi_g \), \( g = 1, 2 \),
\[ i = 1, \ldots, k. \] There are \( \frac{r_{11} + r_{21}}{r_{11}} = NC \) possible orders in which the deaths from the two populations could occur. As an illustration, let \( r_{11} = r_{21} = 2. \) The \( \binom{4}{2} = 6 \) possible orders of the populations from which the four deaths can occur, and the corresponding risk sets, are

1. \( [1 1 2 2] \) \( r'_{\sim 1} = [2 1 0 0] \) \( r'_{\sim 2} = [2 2 2 1], \)
2. \( [1 2 1 2] \) \( r'_{\sim 1} = [2 1 1 0] \) \( r'_{\sim 2} = [2 2 1 1], \)
3. \( [1 2 2 1] \) \( r'_{\sim 1} = [2 1 1 1] \) \( r'_{\sim 2} = [2 2 1 0], \)
4. \( [2 1 1 2] \) \( r'_{\sim 1} = [2 2 1 0] \) \( r'_{\sim 2} = [2 1 1 1], \)
5. \( [2 1 2 1] \) \( r'_{\sim 1} = [2 2 1 1] \) \( r'_{\sim 2} = [2 1 1 0], \)
6. \( [2 2 1 1] \) \( r'_{\sim 1} = [2 2 2 1] \) \( r'_{\sim 2} = [2 1 0 0], \)

Notice that the sequences of risk sets used for examples in SSI are in fact possible sequences for SSII. However, in general this is not true for all SSI sequences.

Estimates of \( \hat{\Delta} \) for all possible NC sequences for any value of \( r_{11} \) and \( r_{21} \) can be found using the four estimators described in Section 2.3, \( \hat{\Delta}_{\ell} (\ell = 1, 2, 3) \) and \( \hat{\Delta}_{4}. \) It has been shown (Section 2.3.1) that \( \hat{\Delta}_{\sim 1} \) is an unbiased estimator of \( \hat{\Delta}. \) It was decided not to consider nearly unbiased estimators of \( \hat{\Delta} \) corresponding to \( \hat{\Delta}_{\sim 2}, \hat{\Delta}_{\sim 3} \) and \( \hat{\Delta}_{\sim 4} \) because one of our estimators, \( \hat{\Delta}_{\sim 1}, \) was unbiased.

Small Samples

The small sample properties of the four estimators under SSII were investigated by generation of their exact distributions which depend on the possible orders of death in the two populations, for

\[ k = r_{11} + r_{21} \leq 10. \] Four examples are given in Tables 2.7 and 2.8 to illustrate their properties. From these examples we can reasonably
<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$\Delta$</th>
<th>$\hat{\delta}_1$</th>
<th>$\hat{\delta}_2$</th>
<th>$\hat{\delta}_3$</th>
<th>$\hat{\delta}_4$</th>
<th>$\hat{\delta}_1$</th>
<th>$\hat{\delta}_2$</th>
<th>$\hat{\delta}_3$</th>
<th>$\hat{\delta}_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>.25</td>
<td>.200000</td>
<td>.200000</td>
<td>.19267</td>
<td>.20235</td>
<td>.20991</td>
<td>.20000</td>
<td>.18997</td>
<td>.19907</td>
<td>.19763</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.188560</td>
<td>.17812</td>
<td>.17694</td>
<td>.17701</td>
<td>.16812</td>
<td>.15573</td>
<td>.15496</td>
<td>.15507</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.035560</td>
<td>.03178</td>
<td>.03131</td>
<td>.03133</td>
<td>.02833</td>
<td>.02435</td>
<td>.02402</td>
<td>.02405</td>
</tr>
<tr>
<td>.50</td>
<td>.333333</td>
<td>.333333</td>
<td>.33033</td>
<td>.33737</td>
<td>.33633</td>
<td>.33333</td>
<td>.32486</td>
<td>.33159</td>
<td>.33058</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.223610</td>
<td>.22138</td>
<td>.21452</td>
<td>.21546</td>
<td>.20242</td>
<td>.19704</td>
<td>.19132</td>
<td>.19221</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.030000</td>
<td>.04002</td>
<td>.04604</td>
<td>.04644</td>
<td>.04097</td>
<td>.03890</td>
<td>.03661</td>
<td>.03685</td>
</tr>
<tr>
<td>.75</td>
<td>.428571</td>
<td>.428571</td>
<td>.43106</td>
<td>.43630</td>
<td>.43382</td>
<td>.42857</td>
<td>.42445</td>
<td>.42712</td>
<td>.42712</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.234340</td>
<td>.235220</td>
<td>.22632</td>
<td>.22758</td>
<td>.21396</td>
<td>.21248</td>
<td>.20465</td>
<td>.20582</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.054492</td>
<td>.05534</td>
<td>.05125</td>
<td>.05187</td>
<td>.04578</td>
<td>.04517</td>
<td>.04188</td>
<td>.04236</td>
</tr>
<tr>
<td>.90</td>
<td>.473684</td>
<td>.473684</td>
<td>.47896</td>
<td>.48019</td>
<td>.47999</td>
<td>.47368</td>
<td>.47213</td>
<td>.47330</td>
<td>.47313</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.235659</td>
<td>.23697</td>
<td>.22778</td>
<td>.22908</td>
<td>.21616</td>
<td>.21552</td>
<td>.20716</td>
<td>.20848</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.053562</td>
<td>.05618</td>
<td>.05393</td>
<td>.05252</td>
<td>.04673</td>
<td>.04645</td>
<td>.04286</td>
<td>.04346</td>
</tr>
<tr>
<td>1.00</td>
<td>.500000</td>
<td>.500000</td>
<td>.50866</td>
<td>.50690</td>
<td>.50687</td>
<td>.50000</td>
<td>.50000</td>
<td>.50000</td>
<td>.50000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.235700</td>
<td>.23658</td>
<td>.22744</td>
<td>.22873</td>
<td>.21651</td>
<td>.21600</td>
<td>.20767</td>
<td>.20890</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.055560</td>
<td>.056020</td>
<td>.05177</td>
<td>.05236</td>
<td>.04688</td>
<td>.04665</td>
<td>.04313</td>
<td>.04364</td>
</tr>
<tr>
<td>1.10</td>
<td>.52381</td>
<td>.52381</td>
<td>.53205</td>
<td>.53102</td>
<td>.53114</td>
<td>.52381</td>
<td>.52522</td>
<td>.52416</td>
<td>.52631</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.234930</td>
<td>.23336</td>
<td>.22637</td>
<td>.22764</td>
<td>.21622</td>
<td>.21561</td>
<td>.20733</td>
<td>.20855</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.053539</td>
<td>.05566</td>
<td>.05330</td>
<td>.05187</td>
<td>.04675</td>
<td>.04645</td>
<td>.04299</td>
<td>.04350</td>
</tr>
<tr>
<td>1.30</td>
<td>.56522</td>
<td>.56522</td>
<td>.57565</td>
<td>.57282</td>
<td>.57320</td>
<td>.56522</td>
<td>.56899</td>
<td>.56613</td>
<td>.56655</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.232130</td>
<td>.23128</td>
<td>.22285</td>
<td>.22404</td>
<td>.21439</td>
<td>.21307</td>
<td>.20515</td>
<td>.20633</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.053890</td>
<td>.05360</td>
<td>.04972</td>
<td>.05026</td>
<td>.04596</td>
<td>.04541</td>
<td>.04209</td>
<td>.04257</td>
</tr>
<tr>
<td>2.00</td>
<td>.66667</td>
<td>.66667</td>
<td>.68077</td>
<td>.67421</td>
<td>.67512</td>
<td>.66667</td>
<td>.67514</td>
<td>.66841</td>
<td>.66942</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.217310</td>
<td>.21222</td>
<td>.20547</td>
<td>.20628</td>
<td>.20242</td>
<td>.20164</td>
<td>.19704</td>
<td>.19921</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.047220</td>
<td>.04481</td>
<td>.04223</td>
<td>.04262</td>
<td>.04097</td>
<td>.03890</td>
<td>.03661</td>
<td>.03695</td>
</tr>
<tr>
<td>4.00</td>
<td>.80000</td>
<td>.80000</td>
<td>.81345</td>
<td>.80476</td>
<td>.80560</td>
<td>.80000</td>
<td>.81003</td>
<td>.80093</td>
<td>.80237</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.178470</td>
<td>.16640</td>
<td>.16439</td>
<td>.16432</td>
<td>.16832</td>
<td>.15573</td>
<td>.15496</td>
<td>.15507</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.031850</td>
<td>.02727</td>
<td>.02702</td>
<td>.02704</td>
<td>.02833</td>
<td>.02435</td>
<td>.02402</td>
<td>.02405</td>
</tr>
<tr>
<td>Exact Values</td>
<td>Estimators, $k = 9$, $r_{11} = 4$, $r_{21} = 5$</td>
<td>Estimators, $k = 10$, $r_{11} = 5$, $r_{21} = 5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------------------------------</td>
<td>-----------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$</td>
<td>$\delta_1$ $\delta_2$ $\delta_3$ $\delta_4$</td>
<td>$\delta_1$ $\delta_2$ $\delta_3$ $\delta_4$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.25</td>
<td>.20000</td>
<td>.19233</td>
<td>.2158</td>
<td>.20003</td>
<td>.20000</td>
<td>.19110</td>
<td>.19956</td>
<td>.19807</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.16193</td>
<td>.14968</td>
<td>.14860</td>
<td>.14878</td>
<td>.14847</td>
<td>.13580</td>
<td>.13496</td>
<td>.13538</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.02622</td>
<td>.02246</td>
<td>.02208</td>
<td>.02213</td>
<td>.02204</td>
<td>.01821</td>
<td>.01821</td>
<td>.01828</td>
<td></td>
</tr>
<tr>
<td>.50</td>
<td>.33333</td>
<td>.32863</td>
<td>.33537</td>
<td>.33432</td>
<td>.33333</td>
<td>.32551</td>
<td>.33175</td>
<td>.33079</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.19293</td>
<td>.18746</td>
<td>.18818</td>
<td>.18279</td>
<td>.17888</td>
<td>.17160</td>
<td>.16713</td>
<td>.16795</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.03722</td>
<td>.03516</td>
<td>.03309</td>
<td>.03341</td>
<td>.03200</td>
<td>.02958</td>
<td>.02793</td>
<td>.02821</td>
<td></td>
</tr>
<tr>
<td>.75</td>
<td>.42857</td>
<td>.42875</td>
<td>.43187</td>
<td>.43140</td>
<td>.42857</td>
<td>.42472</td>
<td>.42782</td>
<td>.42720</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.20204</td>
<td>.20040</td>
<td>.19309</td>
<td>.19420</td>
<td>.19217</td>
<td>.18552</td>
<td>.17910</td>
<td>.18009</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.04108</td>
<td>.04016</td>
<td>.03730</td>
<td>.03772</td>
<td>.03580</td>
<td>.03443</td>
<td>.03208</td>
<td>.03243</td>
<td></td>
</tr>
<tr>
<td>.90</td>
<td>.47368</td>
<td>.47667</td>
<td>.47768</td>
<td>.47749</td>
<td>.47368</td>
<td>.47223</td>
<td>.47311</td>
<td>.47316</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.20416</td>
<td>.20235</td>
<td>.19478</td>
<td>.19591</td>
<td>.19118</td>
<td>.18823</td>
<td>.18156</td>
<td>.18128</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.04146</td>
<td>.04095</td>
<td>.03795</td>
<td>.03840</td>
<td>.03655</td>
<td>.03543</td>
<td>.03293</td>
<td>.03330</td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>.50000</td>
<td>.50430</td>
<td>.50437</td>
<td>.50435</td>
<td>.50000</td>
<td>.50000</td>
<td>.50000</td>
<td>.50000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.20412</td>
<td>.20225</td>
<td>.19469</td>
<td>.19582</td>
<td>.19168</td>
<td>.18666</td>
<td>.18183</td>
<td>.18284</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.04167</td>
<td>.04092</td>
<td>.03792</td>
<td>.03836</td>
<td>.03667</td>
<td>.03559</td>
<td>.03306</td>
<td>.03343</td>
<td></td>
</tr>
<tr>
<td>1.10</td>
<td>.52381</td>
<td>.52944</td>
<td>.52849</td>
<td>.52862</td>
<td>.52381</td>
<td>.52513</td>
<td>.52414</td>
<td>.52429</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.20353</td>
<td>.20138</td>
<td>.19394</td>
<td>.19505</td>
<td>.19123</td>
<td>.18831</td>
<td>.18152</td>
<td>.18254</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.04143</td>
<td>.04059</td>
<td>.03763</td>
<td>.03807</td>
<td>.03657</td>
<td>.03546</td>
<td>.03293</td>
<td>.03332</td>
<td></td>
</tr>
<tr>
<td>1.30</td>
<td>.56252</td>
<td>.57298</td>
<td>.57032</td>
<td>.57070</td>
<td>.56252</td>
<td>.56876</td>
<td>.56609</td>
<td>.56648</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.20123</td>
<td>.19813</td>
<td>.19113</td>
<td>.19219</td>
<td>.18959</td>
<td>.18604</td>
<td>.17955</td>
<td>.18024</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.04049</td>
<td>.03932</td>
<td>.03658</td>
<td>.03697</td>
<td>.03594</td>
<td>.03462</td>
<td>.03224</td>
<td>.03260</td>
<td></td>
</tr>
<tr>
<td>2.00</td>
<td>.66667</td>
<td>.67809</td>
<td>.67184</td>
<td>.67278</td>
<td>.66667</td>
<td>.67449</td>
<td>.66825</td>
<td>.66921</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.18856</td>
<td>.18127</td>
<td>.17646</td>
<td>.17726</td>
<td>.17888</td>
<td>.17160</td>
<td>.16713</td>
<td>.16795</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.03556</td>
<td>.03299</td>
<td>.03116</td>
<td>.03146</td>
<td>.03200</td>
<td>.02958</td>
<td>.02793</td>
<td>.02821</td>
<td></td>
</tr>
<tr>
<td>4.00</td>
<td>.80000</td>
<td>.81106</td>
<td>.80272</td>
<td>.80410</td>
<td>.80000</td>
<td>.80890</td>
<td>.80444</td>
<td>.80193</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.15402</td>
<td>.14177</td>
<td>.14123</td>
<td>.14139</td>
<td>.14847</td>
<td>.13580</td>
<td>.13496</td>
<td>.13513</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.02400</td>
<td>.02022</td>
<td>.01995</td>
<td>.02001</td>
<td>.02204</td>
<td>.01852</td>
<td>.01821</td>
<td>.01828</td>
<td></td>
</tr>
</tbody>
</table>
conclude, in addition to

(1) $\hat{\Delta}_1$ is an unbiased estimator of $\Delta$.

that

(2) In the special case $r_{11} = r_{21}$

(i) All estimators are unbiased for $\theta = 1$.

(ii) $\hat{\Delta}_2$ underestimates $\Delta$ for $\theta < 1$ and overestimates $\Delta$ for $\theta > 1$. The same appears to be true for $\hat{\Delta}_3$ and $\hat{\Delta}_4$, though for these estimators the bias is slight.

(iii) The MSE for estimators is at a maximum when $\theta = 1$.

As $k = r_{11} + r_{21}$ increases the MSE decreases.

(3) For the case $r_{11} = r_{21} + 1$ it appears that $\hat{\Delta}_3$ and $\hat{\Delta}_4$ always overestimate $\Delta$. Whether $\hat{\Delta}_2$ overestimates or underestimates $\Delta$ seems to depend on the value of $\theta$ and the relationship of $r_{11}$ and $r_{21}$.

(4) In all cases of $r_{11} = r_{21}$ and $r_{11} = r_{21} + 1$, $\hat{\Delta}_3$ has the smallest MSE (although often only slightly less than that of $\hat{\Delta}_4$). When $r_{11}$ and $r_{21}$ are very different from each other this was not always found to be true.

The distributions of the estimators under SSII are also discrete.

However, the NC estimates corresponding to each possible point in the sample space are usually distinct. Chi-square goodness of fit test statistics were calculated to test $H_0^{(1)}$ (2.5.6) and $H_0^{(2)}$ (2.5.7) for $\Delta \in \{\hat{\Delta}_\ell | \ell = 1,2,3\}$. As before, the test statistics were calculated using ten equal intervals, each of width .1. Parameters for the beta and $S_B$ curves were determined from the first two moments of the exact distributions. From Table 2.9 there is evidence to suggest that in this case the beta distribution and the $S_B$ distribution are
| $k = 7 \ (3,4)$ |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|
| $\hat{\alpha}_1$ | $S_B$ | 133.60 | 28.91 | 8.08 | 4.70 | 4.20 | 4.67 | 7.75 | 34.08 | 197.76 |
| Beta | 204.26 | 29.53 | 6.80 | 3.46 | 2.96† | 3.60† | 6.40† | 35.59 | 373.00 |
| $\hat{\alpha}_2$ | $S_B$ | 214.11 | 33.94 | 7.21 | 2.77† | 1.69‡+ | 1.58‡++ | 3.41† | 24.35 | 193.27 |
| Beta | 403.63 | 36.30 | 6.57 | 2.20† | 1.17‡++ | 1.09‡++ | 2.92† | 25.83 | 391.76 |
| $\hat{\alpha}_3$ | $S_B$ | 153.03 | 26.43 | 5.71 | 2.51† | 2.01‡++ | 2.41‡ | 5.23 | 30.95 | 225.85 |
| Beta | 235.89 | 27.57 | 5.42 | 2.24† | 1.75‡++ | 2.16‡ | 4.93 | 33.86 | 477.52 |
| $\hat{\alpha}_4$ | $S_B$ | 152.86 | 26.26 | 5.72 | 2.54† | 2.04‡++ | 2.43‡ | 5.20 | 30.56 | 224.57 |
| Beta | 234.69 | 27.32 | 5.43 | 2.28† | 1.79‡++ | 2.16‡ | 4.91 | 33.43 | 473.69 |

| $k = 8 \ (4,4)$ |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|
| $\hat{\alpha}_1$ | $S_B$ | 376.33 | 64.21 | 13.06 | 5.43 | 4.29 | 5.23 | 11.34 | 64.21 | 376.35 |
| Beta | 693.31 | 68.56 | 11.86 | 4.33 | 3.22† | 4.13† | 10.34 | 68.56 | 693.40 |
| $\hat{\alpha}_2$ | $S_B$ | 681.02 | 79.29 | 12.44 | 3.60 | 2.31† | 3.37† | 10.66 | 79.29 | 681.02 |
| Beta | 1839.68 | 88.37 | 11.39 | 2.61† | 1.36‡++ | 2.38‡ | 9.59 | 88.37 | 1840.80 |
| $\hat{\alpha}_3$ | $S_B$ | 461.16 | 62.47 | 10.68 | 3.43 | 2.35‡ | 3.23‡ | 9.22 | 62.47 | 461.16 |
| Beta | 943.72 | 66.64 | 9.94 | 2.81† | 1.77‡++ | 2.62‡ | 8.50 | 66.64 | 944.21 |
| $\hat{\alpha}_4$ | $S_B$ | 495.85 | 65.18 | 10.55 | 2.96† | 1.84‡++ | 2.76‡ | 9.02 | 65.18 | 495.85 |
| Beta | 1044.61 | 71.14 | 10.15 | 2.58† | 1.48‡++ | 2.37‡ | 8.61 | 71.14 | 1044.88 |

| $k = 9 \ (5,5)$ |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|
| $\hat{\alpha}_1$ | $S_B$ | 748.77 | 126.43 | 22.50** | 6.96 | 4.66 | 6.70 | 20.54* | 148.08 | 1094.23 |
| Beta | 1510.22 | 136.00 | 19.55* | 4.31 | 2.07‡++ | 3.98† | 17.35* | 161.77 | 2812.40 |
| $\hat{\alpha}_2$ | $S_B$ | 1503.45 | 169.49 | 26.36** | 7.02 | 3.55 | 4.85† | 18.01* | 160.21 | 1727.00 |
| Beta | 4689.09 | 187.98 | 23.41** | 4.76 | 1.59‡++ | 2.98† | 16.07 | 181.61 | 6474.90 |
| $\hat{\alpha}_3$ | $S_B$ | 57.21 | 148.11 | 22.98** | 6.24 | 3.86 | 6.09 | 21.22* | 177.85 | 1860.34 |
| Beta | 3242.01 | 160.68 | 20.30** | 4.17 | 1.82‡++ | 3.87 | 18.34* | 199.34 | 7278.40 |
| $\hat{\alpha}_4$ | $S_B$ | 57.57 | 146.20 | 22.33** | 5.83 | 3.38 | 5.48 | 20.17* | 172.29 | 1806.46 |
| Beta | 3181.93 | 159.00 | 20.32** | 3.95 | 1.52‡++ | 3.44 | 17.51* | 193.09 | 7012.39 |

| $k = 10 \ (5,5)$ |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|
| $\hat{\alpha}_1$ | $S_B$ | 2322.08 | 322.32 | 50.87 | 13.95 | 8.53 | 12.96 | 43.42 | 322.32 | 113.03 |
| Beta | 6335.88 | 354.47 | 44.48 | 8.72 | 3.59 | 7.78 | 37.16 | 354.48 | 6338.67 |
| $\hat{\alpha}_2$ | $S_B$ | 4349.10 | 399.02 | 52.79 | 11.94 | 6.09 | 10.87 | 44.42 | 399.02 | 4540.10 |
| Beta | 20345.34 | 457.34 | 48.21 | 8.27 | 2.67† | 7.24† | 39.90 | 457.35 | 20521.33 |
| $\hat{\alpha}_3$ | $S_B$ | 5135.67 | 450.14 | 60.95 | 16.11 | 9.72 | 14.94 | 51.73 | 450.14 | 5135.67 |
| Beta | 24408.24 | 508.93 | 52.93 | 10.18 | 4.22 | 9.08 | 43.99 | 508.97 | 24630.48 |
| $\hat{\alpha}_4$ | $S_B$ | 4926.88 | 435.05 | 59.20 | 15.48 | 9.23 | 14.33 | 50.23 | 435.05 | 4926.88 |
| Beta | 23200.52 | 491.82 | 51.72 | 9.90 | 4.06 | 8.83 | 43.00 | 491.86 | 23327.32 |

$\dagger^\dagger F(x_9^2 < x_{10}^2) = .01$  
$F(x_9^2 < x_{10}^2) = .05$  
$F(x_9^2 < x_{10}^2) = .95$  
$F(x_9^2 < x_{10}^2) = .99$  

When the symbol applies to all values in a column, it is given at the bottom of that column.
reasonable forms to describe the distributions of the four estimators of $\Delta$ when $\theta$ is close to 1. In each case the beta distribution gives a better fit than the $S_8$ distribution (as judged by the $X^2_9$ values).

Large Samples

As $k$, the total number of deaths from $\pi_1$ and $\pi_2$ increases, generation of the exact distributions of the estimators soon becomes economically prohibitive. Some results from a preliminary sampling study, using as an example $r_{11} = r_{21} = 10$, of the large sample properties of the four estimators of $\Delta$ will be reported.

Example: Suppose $r_{11} = r_{21} = 10$ individuals from $\pi_1$ and $\pi_2$, respectively, are followed until all 20 die. For each of the $NC = \binom{20}{10}$ possible orders in which the deaths can occur, a random number $\varepsilon$ was generated from the unit uniform distribution. If $\varepsilon \leq .0011$ that order would be selected for inclusion in the sample, and the risk sets $r_g, g = 1, 2$, corresponding to that order were evaluated. The sampling fraction $.0011$ was chosen so that the sample size $n$ was approximately 200. Due to the symmetry of the distribution of the possible orders of death when $r_{11} = r_{21}$ only half of the set of combinations was generated; the other half was obtained from the complements of the orders of death of the first half.

For each of the $n$ orders selected for a sample, four estimates of $\Delta$ corresponding to that order were determined. The $\gamma^{th}$ moments about zero for $\gamma=1,2$ were found using (2.5.5). Five samples (of sizes 186, 170, 192, 184 and 222) were drawn. Table 2.10 presents the summary statistics for one of these samples. The conclusions generally confirm those described for small samples ($\Delta_1$ is naturally no longer
<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$\Delta$</th>
<th>Exact Values</th>
<th>Estimators</th>
<th>Estimators</th>
<th>Estimators</th>
<th>Estimators</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MEAN</td>
<td>$\hat{\theta}_1$</td>
<td>$\hat{\theta}_2$</td>
<td>$\hat{\theta}_3$</td>
<td>$\hat{\theta}_4$</td>
</tr>
<tr>
<td>0.25</td>
<td>0.20000</td>
<td>MEAN</td>
<td>0.1921</td>
<td>0.1838</td>
<td>0.1879</td>
<td>0.1875</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>0.0809</td>
<td>0.0735</td>
<td>0.0733</td>
<td>0.0730</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MSE</td>
<td>0.0066</td>
<td>0.0057</td>
<td>0.0055</td>
<td>0.0055</td>
</tr>
<tr>
<td>0.50</td>
<td>0.33333</td>
<td>MEAN</td>
<td>0.3354</td>
<td>0.3282</td>
<td>0.3315</td>
<td>0.3311</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>0.1335</td>
<td>0.1275</td>
<td>0.1259</td>
<td>0.1261</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MSE</td>
<td>0.0178</td>
<td>0.0163</td>
<td>0.0159</td>
<td>0.0159</td>
</tr>
<tr>
<td>0.75</td>
<td>0.42857</td>
<td>MEAN</td>
<td>0.4186</td>
<td>0.4333</td>
<td>0.4368</td>
<td>0.4366</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>0.1245</td>
<td>0.1196</td>
<td>0.1170</td>
<td>0.1172</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MSE</td>
<td>0.0156</td>
<td>0.0143</td>
<td>0.0138</td>
<td>0.0138</td>
</tr>
<tr>
<td>0.90</td>
<td>0.47368</td>
<td>MEAN</td>
<td>0.4781</td>
<td>0.4769</td>
<td>0.4774</td>
<td>0.4774</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>0.1209</td>
<td>0.1157</td>
<td>0.1132</td>
<td>0.1135</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MSE</td>
<td>0.0146</td>
<td>0.0134</td>
<td>0.0128</td>
<td>0.0129</td>
</tr>
<tr>
<td>1.00</td>
<td>0.50000</td>
<td>MEAN</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.5000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>0.1203</td>
<td>0.1151</td>
<td>0.1125</td>
<td>0.1128</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MSE</td>
<td>0.0145</td>
<td>0.0132</td>
<td>0.0127</td>
<td>0.0127</td>
</tr>
<tr>
<td>1.10</td>
<td>0.52381</td>
<td>MEAN</td>
<td>0.5198</td>
<td>0.5209</td>
<td>0.5204</td>
<td>0.5205</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>0.1208</td>
<td>0.1156</td>
<td>0.1130</td>
<td>0.1134</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MSE</td>
<td>0.0146</td>
<td>0.0134</td>
<td>0.0128</td>
<td>0.0129</td>
</tr>
<tr>
<td>1.30</td>
<td>0.56522</td>
<td>MEAN</td>
<td>0.5557</td>
<td>0.5587</td>
<td>0.5574</td>
<td>0.5575</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>0.1219</td>
<td>0.1187</td>
<td>0.1163</td>
<td>0.1166</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MSE</td>
<td>0.0154</td>
<td>0.0141</td>
<td>0.0136</td>
<td>0.0137</td>
</tr>
<tr>
<td>2.00</td>
<td>0.66667</td>
<td>MEAN</td>
<td>0.6644</td>
<td>0.6718</td>
<td>0.6685</td>
<td>0.6689</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>0.1335</td>
<td>0.1275</td>
<td>0.1259</td>
<td>0.1261</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MSE</td>
<td>0.0178</td>
<td>0.0163</td>
<td>0.0159</td>
<td>0.0159</td>
</tr>
<tr>
<td>4.00</td>
<td>0.80000</td>
<td>MEAN</td>
<td>0.8079</td>
<td>0.8162</td>
<td>0.8121</td>
<td>0.8125</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>0.0809</td>
<td>0.0735</td>
<td>0.0733</td>
<td>0.0730</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MSE</td>
<td>0.0066</td>
<td>0.0057</td>
<td>0.0055</td>
<td>0.0055</td>
</tr>
</tbody>
</table>
unbiased in this restricted set, although it is actually unbiased).

The null hypotheses $H_0^{(1)}$ (2.5.6) and $H_0^{(2)}$ (2.5.7) were tested for $\hat{A}_\ell[\hat{A}_\ell (\ell = 1, 2, 3), \hat{A}_4]$. Table 2.11 gives the one-sample Kolmogorov-Smirnov test statistics $D_n$ for each estimator at each of the nine values of $\theta$. The parameters were estimated from the sample using the method of moments. These are available in manuscript form.

The KS test statistics are conservative when the parameters are unknown and must be estimated from the sample. They are also conservative when the true distribution is discrete (see Gibbons, 1971, pp. 85-87). There is evidence to reject both $H_0^{(1)}$ and $H_0^{(2)}$ for all values of $\theta$, except for $\theta = 1$, or $A = \frac{1}{2}$. It remains uncertain whether a test statistic more appropriate for the beta and the $S_B$ distributions with parameters unknown would lead to the same conclusions for $\theta = 1$. Modifications of the significance limits of the KS statistic along the lines of the research of Stephens (1974) or of Durbin (1975) should be possible when the underlying distributions are either a beta or a $S_B$ distribution. As yet, however, such results are not available.
<table>
<thead>
<tr>
<th>Exact Values</th>
<th>$\theta$</th>
<th>$.25$</th>
<th>$.50$</th>
<th>$.75$</th>
<th>$.90$</th>
<th>$1.0$</th>
<th>$1.1$</th>
<th>$1.3$</th>
<th>$2.0$</th>
<th>$4.0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta_1$</td>
<td>$S_B$</td>
<td>12.0306</td>
<td>7.0131</td>
<td>3.2319</td>
<td>1.4601†</td>
<td>.5269</td>
<td>1.3672†</td>
<td>2.9807</td>
<td>7.0131</td>
<td>12.0306</td>
</tr>
<tr>
<td>Beta</td>
<td>12.0673</td>
<td>6.9389</td>
<td>3.1811</td>
<td>1.4389†</td>
<td>.4985</td>
<td>1.3515</td>
<td>2.9318</td>
<td>6.9389</td>
<td>12.0673</td>
<td></td>
</tr>
<tr>
<td>$\Delta_2$</td>
<td>$S_B$</td>
<td>12.3335</td>
<td>7.6430</td>
<td>3.5397</td>
<td>1.6073†</td>
<td>.6419</td>
<td>1.5091†</td>
<td>3.2601</td>
<td>7.6430</td>
<td>12.3335</td>
</tr>
<tr>
<td>Beta</td>
<td>12.3904</td>
<td>7.6000</td>
<td>3.4813</td>
<td>1.5692†</td>
<td>.6123</td>
<td>1.4720†</td>
<td>3.2045</td>
<td>7.6000</td>
<td>12.3904</td>
<td></td>
</tr>
<tr>
<td>$\Delta_3$</td>
<td>$S_B$</td>
<td>12.3666</td>
<td>7.6924</td>
<td>3.5880</td>
<td>1.6581††</td>
<td>.6961</td>
<td>1.5602†</td>
<td>3.3087</td>
<td>7.6924</td>
<td>12.3666</td>
</tr>
<tr>
<td>Beta</td>
<td>12.4303</td>
<td>7.6508</td>
<td>3.5318</td>
<td>1.6216†</td>
<td>.6676</td>
<td>1.5246†</td>
<td>3.2551</td>
<td>7.6508</td>
<td>12.4303</td>
<td></td>
</tr>
<tr>
<td>$\Delta_4$</td>
<td>$S_B$</td>
<td>12.3604</td>
<td>7.6944</td>
<td>3.5883</td>
<td>1.6583††</td>
<td>.6962</td>
<td>1.5604†</td>
<td>3.3089</td>
<td>7.6944</td>
<td>12.3604</td>
</tr>
<tr>
<td>Beta</td>
<td>12.4239</td>
<td>7.6526</td>
<td>3.5317</td>
<td>1.6216†</td>
<td>.6676</td>
<td>1.5246†</td>
<td>3.2551</td>
<td>7.6526</td>
<td>12.4239</td>
<td></td>
</tr>
</tbody>
</table>

$\dagger$Pr($\sqrt{D_n} > z$) = .01, $\dagger\dagger$Pr($\sqrt{D_n} > z$) = .001

When the symbol applies to all values in a column, it is given at the bottom of that column.
CHAPTER III

SOME ANALYTICAL AND EMPIRICAL INVESTIGATIONS OF

THE HAZARD RATIO FUNCTION (HRF)

In Chapter II we considered the special case when the ratio of two
hazard functions is constant in the form

\[ \frac{\mu_1(t)}{\mu_2(t)} = \theta, \]

or more conveniently in the form

\[ \frac{\mu_1(t)}{\mu_1(t) + \mu_2(t)} = \Delta. \]

We called this model the proportional hazard function (PHF) model.

In this chapter we wish to consider a more general situation
when the ratio of two hazard functions is a function of \( t \). Thus we
introduce the hazard ratio function (HRF)

\[ \theta(t) = \frac{\mu_1(t)}{\mu_2(t)} \quad \text{for all } t, \quad (3.0.1) \]

and the hazard delta function

\[ \Delta(t) = \frac{\mu_1(t)}{\mu_1(t) + \mu_2(t)} \quad \text{for all } t. \quad (3.0.2) \]

Sometimes we may write (3.0.1) as

\[ \theta(t) = \frac{\mu_1(t; \xi)}{\mu_2(t; \xi)}, \quad (3.0.3) \]

where \( \xi = [\xi_1, \ldots, \xi_m] \) is the \( 1 \times m \) vector of parameters for the
hazard function of $\pi_g$, $g = 1, 2$.

We will discuss the behavior of the HRF for some situations when

(i) population mortality data,

(ii) empirical data, and

(iii) parametric models with known forms of the hazard functions $\mu_g(t; \xi_g)$, $g = 1, 2$,

are used.

The special case of the HRF defined by the proportional hazard function model has frequently been assumed by both actuaries and demographers to describe the relationship between mortality experience in two populations. In particular, the PHF model is assumed to hold in many statistical tests for comparing survival curves (for example see Peto and Peto (1972) and Crowley (1974)).

The brief overview assessment of the HRF in this chapter is motivated by a problem to be considered in Chapter IV, that is, the effect of $\theta(t)$ not being constant for all $t$, on the estimation of $'\theta'$ and $'\Delta'$ under the assumption that the PHF model is valid.

3.1. The HRF for Some Population Mortality Data.

Four examples of the HRF for some population mortality data are presented in Figures 3.1 to 3.4. The life table data are taken from Keyfitz and Flieger (1971), pp. 354-372.

Let $t_0 < t_1 < \ldots < t_{\omega-1}$ denote a sequence of non-overlapping intervals of width $h_i = t_{i+1} - t_i$ and midpoint $t_i' = \frac{1}{2}(t_i + t_{i+1})$. The usual approximation to the hazard function at $t_i'$ is

$$\mu(t_i') = -\frac{1}{h_i} \ln(1 - q_i) = -\frac{1}{h_i} \frac{\#n_i}{\#p_i}, \quad (3.1.1)$$
where $q_i$ is the conditional probability of dying in the interval $[t_i, t_{i+1})$ given alive at $t_i$, provided $q_i$ is small (less than .3, say).

The values of

$$\theta(t_i) = \frac{\ln p_{1i}}{\ln p_{2i}}, \quad i = 0, \ldots, \omega - 1$$  \hspace{1cm} (3.1.2)

were plotted in Figures 3.1 to 3.4 for all age intervals up to age 85.

The four examples are

(i) Males and females, United States 1966 population,  
(Figure 3.1).

(ii) Nonwhite and white males, United States 1966 population 
(Figure 3.2).

(iii) United States (males) and England and Wales (males), 1967 
population (Figure 3.3).

(iv) United States (males) and West German (males), 1967 
population (Figure 3.4).

From these illustrations we can see that the HRF's are not simple 
functions. In particular, they cannot be regarded as effectively con-
stant, except over very restricted ranges of $t$.

3.2. The HRF for Some Empirical Data.

Three examples are presented of the HRF estimated from some 
empirical data reported in previously cited references. A brief descrip-
tion of these examples is given.

Example 1: The data are one year results of a clinical trial reported 
by Freireich et al. (1963) which compared a treatment against a placebo 
in the maintenance of remissions in acute leukemia patients. There are
FIGURE 3.1
MALES AND FEMALES, UNITED STATES 1966 POPULATION
(Keyfitz and Flieger (1971), p. 354)

FIGURE 3.2
NONWHITE AND WHITE MALES, UNITED STATES 1966 POPULATION
(Keyfitz and Flieger (1971), pp. 356, 358)
FIGURE 3.3

UNITED STATES (MALES) AND ENGLAND AND WALES (MALES), 1967 POPULATION

(Keyfitz and Flieger (1971), pp. 360, 472)

FIGURE 3.4

UNITED STATES (MALES) AND WEST GERMANY (MALES), 1967 POPULATION

(Keyfitz and Flieger (1971), pp. 360, 438)
21 individuals in each population. The data have been reanalyzed by Gehan (1965a) and Cox (1972). Cox concluded (graphically) that an exponential distribution was a reasonable model for the lifetime data in each population. This naturally implies that the PHF model is reasonable for comparing survival experience. Table 3.1 presents a life table analysis of the data and Figure 3.5 is a plot of the estimate of \( \theta(t) \).

### Table 3.1

**Example 1, Sample Estimates of HRF**

(Data taken from Freireich et al. (1963))

<table>
<thead>
<tr>
<th>Interval</th>
<th>( \pi_1 ): Treatment Std. Error</th>
<th>( \pi_2 ): Placebo Std. Error</th>
<th>Est. HRF</th>
<th>Std. Error ( \theta(t') ) Est.</th>
</tr>
</thead>
<tbody>
<tr>
<td>([t_i, t_{i+1})) weeks</td>
<td>( -\frac{1}{h_i} \ln p_{1i} ) (Approx. Est.)</td>
<td>( -\frac{1}{h_i} \ln p_{2i} ) (Approx. Est.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[0,4)</td>
<td>.00000</td>
<td>.00000</td>
<td>.06798</td>
<td>.00000</td>
</tr>
<tr>
<td>[4,8)</td>
<td>.05427</td>
<td>.02719</td>
<td>.07192</td>
<td>.03608</td>
</tr>
<tr>
<td>[8,12)</td>
<td>.01786</td>
<td>.01787</td>
<td>.17329</td>
<td>.07217</td>
</tr>
<tr>
<td>[12,16)</td>
<td>.02175</td>
<td>.02176</td>
<td>.17329</td>
<td>.10206</td>
</tr>
<tr>
<td>[16,20)</td>
<td>.02634</td>
<td>.02635</td>
<td>.10137</td>
<td>.10206</td>
</tr>
<tr>
<td>[20,24)</td>
<td>.08412</td>
<td>.05976</td>
<td>.00000</td>
<td>.00000</td>
</tr>
</tbody>
</table>
Example 2: The data are some preliminary results from a cooperative clinical trial which compared six maintenance regimens for remission of acute leukemia. The data were made available by Breslow (1974a) with permission of the Children's Cancer Study Group A. He noted that an exponential distribution is a reasonable approximation to the estimated remission curve. Table 3.2 presents a life table analysis of two treatment populations formed by pooling regimens 1, 2 and 4 and regimens 3 and 5, with 152 and 116 individuals, respectively. A plot of the estimated HRF is given in Figure 3.6.
### TABLE 3.2

**EXAMPLE 2, SAMPLE ESTIMATES OF HRF**

(Data from Breslow (1974a), the Children's Cancer Study Group A)

<table>
<thead>
<tr>
<th>Interval $[t_i, t_{i+1})$</th>
<th>$\pi_1$: Regimens 1,2,4</th>
<th>$\pi_2$: Regimens 3,5</th>
<th>(~\frac{1}{h_i}\ln p_{1i}) (Approx. Est.)</th>
<th>(~\frac{1}{h_i}\ln p_{2i}) (Approx. Est.)</th>
<th>(\tilde{\theta}(t'_i)) (Approx. Est.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0, 120)</td>
<td>.00069</td>
<td>.00020</td>
<td>.00125</td>
<td>.00031</td>
<td>0.5502</td>
</tr>
<tr>
<td>[120, 240)</td>
<td>.00186</td>
<td>.00035</td>
<td>.00301</td>
<td>.00055</td>
<td>0.6188</td>
</tr>
<tr>
<td>[240, 360)</td>
<td>.00176</td>
<td>.00039</td>
<td>.00208</td>
<td>.00054</td>
<td>0.8482</td>
</tr>
<tr>
<td>[360, 480)</td>
<td>.00229</td>
<td>.00051</td>
<td>.00155</td>
<td>.00055</td>
<td>1.4759</td>
</tr>
<tr>
<td>[480, 600)</td>
<td>.00214</td>
<td>.00062</td>
<td>.00292</td>
<td>.00104</td>
<td>0.7320</td>
</tr>
<tr>
<td>[600, 720)</td>
<td>.00125</td>
<td>.00063</td>
<td>.00205</td>
<td>.00119</td>
<td>0.6101</td>
</tr>
<tr>
<td>[720, 840)</td>
<td>.00056</td>
<td>.00056</td>
<td>.00140</td>
<td>.00141</td>
<td>0.3963</td>
</tr>
<tr>
<td>[840, 960)</td>
<td>.00250</td>
<td>.00177</td>
<td>.00000</td>
<td>.00000</td>
<td>-</td>
</tr>
</tbody>
</table>
Example 3: The data are some preliminary results from a clinical trial reported by Glasser (1967) of survival following surgery for lung cancer. The two populations are those with "low" ratios of vital capacity to predicted vital capacity and those with "high" ratios. There are 36 and 95 individuals respectively in the two populations. Glasser also states that an exponential distribution provides a reasonable fit to the survival curve. Two life table analyses of the data are given in Tables 3.3a and 3.3b with the respective plots of the estimates of $\theta(t)$ given in Figures 3.7a and 3.7b. The two analyses are presented to indicate the effect of grouping on the estimate of the HRF.
<table>
<thead>
<tr>
<th>Interval</th>
<th>$\pi_1$: &quot;Low&quot; ratios</th>
<th>Std. Error</th>
<th>$\pi_2$: &quot;High&quot; ratios</th>
<th>Std. Error</th>
<th>E.HRF</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[t_i, t_{i+1})$</td>
<td>$\frac{1}{h_i} \ell \ln p_{1i}$ (Approx. Est.)</td>
<td>$\frac{1}{h_i} \ell \ln p_{2i}$ (Approx. Est.)</td>
<td>$\tilde{\theta}(t_i)$ (Approx. Est.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$[0, 30)$</td>
<td>.00838</td>
<td>.00297</td>
<td>.00332</td>
<td>.00111</td>
<td>2.5250</td>
<td>.0416</td>
</tr>
<tr>
<td>$[30, 60)$</td>
<td>.00251</td>
<td>.00178</td>
<td>.00241</td>
<td>.00098</td>
<td>1.0414</td>
<td>.0294</td>
</tr>
<tr>
<td>$[60, 90)$</td>
<td>.00600</td>
<td>.00301</td>
<td>.00088</td>
<td>.00062</td>
<td>6.8414</td>
<td>.1906</td>
</tr>
<tr>
<td>$[90, 120)$</td>
<td>.00000</td>
<td>.00000</td>
<td>.00206</td>
<td>.00103</td>
<td>0.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>$[120, 150)$</td>
<td>.00202</td>
<td>.00202</td>
<td>.00292</td>
<td>.00131</td>
<td>0.6927</td>
<td>.0265</td>
</tr>
<tr>
<td>$[150, 180)$</td>
<td>.00468</td>
<td>.00331</td>
<td>.00196</td>
<td>.00113</td>
<td>2.3926</td>
<td>.0738</td>
</tr>
<tr>
<td>$[180, 210)$</td>
<td>.00000</td>
<td>.00000</td>
<td>.00320</td>
<td>.00160</td>
<td>0.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>$[210, 240)$</td>
<td>.00445</td>
<td>.00445</td>
<td>.00361</td>
<td>.00180</td>
<td>1.2340</td>
<td>.0484</td>
</tr>
<tr>
<td>$[240, 270)$</td>
<td>.00545</td>
<td>.00545</td>
<td>.00099</td>
<td>.00099</td>
<td>5.5173</td>
<td>.2559</td>
</tr>
<tr>
<td>$[270, 300)$</td>
<td>.01352</td>
<td>.01361</td>
<td>.00222</td>
<td>.00157</td>
<td>6.0797</td>
<td>.2456</td>
</tr>
<tr>
<td>$[300, 330)$</td>
<td>.00000</td>
<td>.00000</td>
<td>.00238</td>
<td>.00168</td>
<td>0.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>$[330, 360)$</td>
<td>-</td>
<td>-</td>
<td>.00283</td>
<td>.00200</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$[360, 390)$</td>
<td>-</td>
<td>-</td>
<td>.00287</td>
<td>.00287</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
**TABLE 3.3b**

**EXAMPLE 3, SAMPLE ESTIMATES OF HRF's**

*(Data from Glasser (1967))*

<table>
<thead>
<tr>
<th>Interval</th>
<th>$\pi_1$: &quot;Low&quot; ratios</th>
<th>Std. Error</th>
<th>$\pi_2$: &quot;High&quot; ratios</th>
<th>Std. Error</th>
<th>Est. HRF</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[t_i, t_{i+1})$</td>
<td>$\frac{1}{h_1} \ln \rho_{1i}$ (Approx. Est.)</td>
<td>$\frac{1}{h_1} \ln \rho_{2i}$ (Approx. Est.)</td>
<td>$\hat{\theta}(t_i')$ (Approx. Est.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[0, 50)</td>
<td>0.00504</td>
<td>0.00179</td>
<td>0.00319</td>
<td>0.00085</td>
<td>1.5793</td>
<td>0.0149</td>
</tr>
<tr>
<td>[50, 100)</td>
<td>0.00530</td>
<td>0.00217</td>
<td>0.00162</td>
<td>0.00066</td>
<td>3.2736</td>
<td>0.0378</td>
</tr>
<tr>
<td>[100, 150)</td>
<td>0.00121</td>
<td>0.00121</td>
<td>0.00205</td>
<td>0.00084</td>
<td>0.5916</td>
<td>0.0135</td>
</tr>
<tr>
<td>[150, 200)</td>
<td>0.00317</td>
<td>0.00224</td>
<td>0.00253</td>
<td>0.00104</td>
<td>1.2495</td>
<td>0.0217</td>
</tr>
<tr>
<td>[200, 250)</td>
<td>0.00575</td>
<td>0.00408</td>
<td>0.00267</td>
<td>0.00120</td>
<td>2.1544</td>
<td>0.0378</td>
</tr>
<tr>
<td>[250, 300)</td>
<td>0.00579</td>
<td>0.00581</td>
<td>0.00192</td>
<td>0.00111</td>
<td>3.0128</td>
<td>0.0716</td>
</tr>
<tr>
<td>[300, 350)</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00145</td>
<td>0.00103</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>[350, 400)</td>
<td>-</td>
<td>-</td>
<td>0.00444</td>
<td>0.00257</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
FIGURE 3.7

EXAMPLE 3, PLOTS OF ESTIMATED HRF's
(Data from Glasser (1967))

(a)

(b)
These three examples each include censored observations (in which the censoring is not necessarily the same in both populations being compared). An actuarial estimator of \( p_{g1} \), \( g = 1,2 \), which uses the concept of "effective" number initially exposed to the risk of dying in an interval was employed (Benjamin and Haycocks, 1970). Estimates of the hazard functions \( \tilde{\mu}_g(t'_i) = -\frac{1}{h_i} \tilde{\lambda}_{n_{g1}} \) and the HRF \( \tilde{\theta}(t'_i) = \frac{\tilde{\lambda}_{n_{p11}}}{\tilde{\lambda}_{n_{p21}}} \), \( g = 1,2 \); \( i = 0, \ldots, \omega - 1 \) were then calculated. An approximation to the variance of the estimated hazard functions is given by

\[
\text{Var}[\tilde{\mu}_g(t'_i)] = \frac{q_{g1}}{h_i^2 \tilde{\lambda}_{n_{g1}} \tilde{E}_{g1}},
\]

(3.2.1)

where \( \tilde{E}_{g1} \) is the "effective" number of individuals in \( \pi_g \) exposed to the risk of dying in \( [t'_i, t'_{i+1}] \). It was taken to be the number of individuals in \( \pi_g \) who survived the interval, plus the number of individuals in \( \pi_g \) who died during the interval and the proportion of the interval in which the individuals censored were observed. An approximation to the variance of the estimated \( \tilde{\theta}(t'_i) \), using the first order approximation to the variance of a ratio of random variables, is given by

\[
\text{Var}[\tilde{\theta}(t'_i)] = \frac{q_{11}^2}{q_{21}^2 h_i^2 \tilde{E}_{11} \tilde{p}_{11} q_{11} + \tilde{E}_{21} \tilde{p}_{21} q_{21}}.
\]

(3.2.2)

Since the hazard function is a rate (Elandt-Johnson, 1975), usually described in terms of events per person time units, we must group the data in order to estimate it. As can be seen pictorially from Figures 3.7a and 3.7b, the interval lengths which determine the groups are an important factor in the resulting estimate of the HRF. Also, the sample size and observed number of deaths affect the precision of the estimate of the HRF. This suggests that we will have a difficult task in using empirical results to make conclusions about the form of
the HRF. If, however, the data fit parametric distributions, then the functional forms of the hazard rates are known, and the behavior of their ratios can be investigated analytically.

3.3. The HRF for Some Parametric Models.

In this section we derive analytic forms of the HRF for some parametric distributions. The examples selected are models frequently used in the description of biomedical lifetime distributions. Table 3.4 presents a summary of eight of these models (Barlow and Proschan, 1965; Johnson and Kotz, 1970; Ord, 1972; Patel, 1973; Mann, Schafer and Singpurwalla, 1974). For convenience we will use IHR (DHR) to describe hazard functions which are monotonic increasing (decreasing) functions of time. Those hazard functions which are IHR (DHR) for \( t < \tau \) an DHR (IHR) for \( t > \tau \) will be called unimodal. For illustrative purposes we will deal with parameter values of the models for which \( \mu(t) < 1 \) for all \( t \in [0,100] \).

It is useful to take note of the following properties of hazard functions.

Remark 1: Effect of transformations on hazard functions. Let \( T \) be a positive continuous random variable with density function \( f_T(t) \) and hazard function \( \mu_T(t) \). Let \( Y = g(T) \) be a monotonic increasing transformation of \( T \) such that \( T = g^{-1}(Y) \) exists. It is easy to show that the relationship of the hazard functions of \( T \) and \( Y \) is given by

\[
\mu_Y(g(t)) = \mu_T(g^{-1}(y)) \frac{d}{dy} [g^{-1}(y)].
\]

(3.3.1)

In particular, consider the linear transformation of \( T \) with location and scale parameters \( \xi \) and \( \sigma \) respectively, \( T = \sigma Y + \xi \).
<table>
<thead>
<tr>
<th>Family</th>
<th>Hazard Function $h(t) &gt; 0$</th>
<th>Survival Function $P(t) = 1 - F(t)$</th>
<th>Central Moments</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Exponential</td>
<td>$u(t; \lambda) = \lambda$, $\lambda &gt; 0$</td>
<td>$\exp[-\lambda t]$</td>
<td>$u_1 = \frac{1}{\lambda}$, $u_2 = \frac{1}{\lambda^2}$</td>
<td>IHR, DHR</td>
</tr>
<tr>
<td>2. Weibull (Extreme Value, Type III)</td>
<td>$u(t; \alpha, \kappa) = \alpha \kappa t^{\alpha-1} \exp[-(\frac{t}{\alpha})^\kappa]$, $\alpha, \kappa &gt; 0$</td>
<td>$\exp[-(\frac{t}{\alpha})^\kappa]$</td>
<td>$u_1 = \alpha \kappa^\frac{1}{\alpha}$, $u_2 = \alpha \kappa^\frac{2}{\alpha}$</td>
<td>$\kappa$-1 exponential $e$-3, $\kappa$ and $e$-1 is similar in shape to a normal distribution. IHR $\neq 1$, DHR $\neq 0$ ($\kappa &lt; 1$)</td>
</tr>
<tr>
<td>3. Makeham Compertz (A=0 is Compertz or Extreme Value, Type I)</td>
<td>$u(t; A, B, x) = A + Be^{xt}$, $A, B &gt; 0$</td>
<td>$\exp[-(B/e)(e^{xt} - 1) - At]$</td>
<td>See Appendix A1</td>
<td>$A$-1 exponential $e$-3, $B$-1 exponential $e$-3, $B$-1 and $e$-1 is similar in shape to a normal distribution. DHR $\neq 0$ ($\kappa &lt; 1$)</td>
</tr>
<tr>
<td>4. Linear Hazard Distribution [see Kendin (1963)]</td>
<td>$u(t; \alpha, B) = \alpha + Bt$, $\alpha &gt; 0$</td>
<td>$\exp[-(\alpha t + Bt^2)]$</td>
<td>See Appendix A2</td>
<td>IHR $\neq 0$, DHR $\neq 0$ (Unimodal)</td>
</tr>
<tr>
<td>5. Lognormal</td>
<td>$u(t; \xi, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(\frac{-t^2}{2\sigma^2}\right)$, $\sigma &gt; 0$</td>
<td>$1 - \frac{\Phi(t)}{\Phi(\xi)}$</td>
<td>$u_1 = \exp[t + \frac{\sigma^2}{2}]$, $u_2 = \exp\left[2t + \sigma^2\right] \Phi(\xi)$</td>
<td>Not IHR, Not DHR</td>
</tr>
<tr>
<td>6. Truncated Logistic</td>
<td>$u(t; \xi, \beta) = \frac{(1 + \exp(-\frac{t - \xi}{\beta}))^{-1}}{\beta[1 + \exp(-\frac{t - \xi}{\beta})]}$, $\xi, \beta &gt; 0$</td>
<td>$\frac{1}{\exp(1/\beta)} \left(\frac{\exp(-t/\beta)}{\exp(1/\beta)}\right) \frac{\exp(-t/\beta)}{1 + \exp(-t/\beta)}$</td>
<td>See Appendix A3</td>
<td>IHR</td>
</tr>
<tr>
<td>7. Gamma (Pearson Type III)</td>
<td>$u(t; \alpha, \beta) = \frac{\Gamma(\alpha)}{\Gamma(\alpha) \beta^\alpha} \exp(-t/\beta)$, $\alpha, \beta &gt; 0$</td>
<td>$1 - \int \frac{1}{\Gamma(\alpha)} \exp(-x/\beta) dx$</td>
<td>$u_1 = \alpha \beta$, $u_2 = \alpha \beta^2$</td>
<td>$\alpha$-1 exponential $e$-3, $\beta$-1 in chi-square IHR $\neq 1$, DHR $\neq 0$ ($\alpha &lt; 1$)</td>
</tr>
<tr>
<td>8. Burr [see Burr (1942)]</td>
<td>$u(t; \alpha, \beta) = \frac{\alpha \beta t^{\alpha-1}}{1 + t^{\alpha}}$, $\alpha, \beta &gt; 0$</td>
<td>$\left[1 + t^{-\alpha}\right]^{-\beta}$</td>
<td>$u_1 = 2\beta[\Gamma(\frac{1}{\alpha}) + \Gamma(\frac{1}{\alpha})^2]$</td>
<td>DHR $\neq 0$ (Unimodal)</td>
</tr>
</tbody>
</table>

where $\Gamma(t)$ is the gamma function [see Burr and Glesk (1963)].
The hazard functions are related by

$$\mu_Y(y) = \sigma \mu_T(\sigma y + \xi).$$  \hfill (3.3.2)

**Remark 2:** Effect of truncation on hazard functions. Let

$$F_T(t|a \leq T \leq b)$$
be the cdf of a continuous random variable $T$, $-\infty < t < \infty$, over the restricted range $a \leq T \leq b$, i.e.

$$F_T(t|a \leq t \leq b) = \begin{cases} 
0 & \text{for } a < t, \\
\frac{F_T(t) - F_T(a)}{F_T(b) - F_T(a)} & \text{for } a \leq t \leq b, \\
1 & \text{for } t > b.
\end{cases}$$

The hazard function of the truncated random variable is only affected by truncation on the right, i.e.

$$\mu_T(t|a \leq T \leq b) = \frac{f_T(t|a \leq T \leq b)}{1 - F_T(t|a \leq T \leq b)}$$ \hfill (3.3.3)

$$= \frac{f_T(t)}{F_T(b) - F_T(t)}$$

$$= \mu_T(t|T \leq b).$$ \hfill (3.3.4)

3.3.1. The HRF for Some Parametric Models from the Same Family of Distributions.

A brief review of the HRF for pairs of hazard functions from some parametric models will be considered. A summary of this review is presented in Table 3.5. For illustrations we will restrict ourself to selecting pairs of hazard functions which are monotonic increasing whenever possible. Notice that the delta hazard function, $\Delta(t)$, is monotonic increasing (decreasing) when the HRF $\theta(t)$ is monotonic increasing (decreasing).
<table>
<thead>
<tr>
<th>Family</th>
<th>Constant</th>
<th>Monotonic Increasing</th>
<th>Monotonic Decreasing</th>
<th>Inflection</th>
<th>( \lim_{t \to 0+} G(t) )</th>
<th>( \lim_{t \to \infty} G(t) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Exponential</td>
<td>For all values of ( t )</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>( \lambda )</td>
<td>( \lambda )</td>
</tr>
<tr>
<td>2. Weibull</td>
<td>( k_1 = k_2 )</td>
<td>( k_1 &gt; k_2 )</td>
<td>( k_1 &lt; k_2 )</td>
<td>No</td>
<td>( 0, \ k_1 &gt; k_2 )</td>
<td>( 0, \ k_1 &lt; k_2 )</td>
</tr>
<tr>
<td>3. Makeham</td>
<td>( A_1 = A_2 = 0 ) and ( k_1 = k_2 ) or ( k_1 &gt; k_2 ) and ( k_1 &lt; k_2 )</td>
<td>( A_1 = A_2 = 0 ) and ( k_1 &lt; k_2 ) and ( A_1 + A_2 &gt; 0 )</td>
<td>( A_1 + A_2 &gt; 0 )</td>
<td>( \frac{\alpha_1}{\alpha_2} )</td>
<td>( \frac{\alpha_1}{\alpha_2} )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trivial case: ( B_1 = B_2 = 0 )</td>
<td>( \frac{\alpha_1}{\alpha_2} )</td>
<td>( \frac{\alpha_1}{\alpha_2} )</td>
<td>( \frac{\alpha_1}{\alpha_2} )</td>
<td>( \frac{\alpha_1}{\alpha_2} )</td>
<td>( \frac{\alpha_1}{\alpha_2} )</td>
</tr>
<tr>
<td>4. Linear</td>
<td>( a_1 = a_2 = 0 ) or ( \frac{a_1}{a_2} = \frac{b_1}{b_2} )</td>
<td>( \frac{a_1}{a_2} &lt; \frac{b_1}{b_2} )</td>
<td>( \frac{a_1}{a_2} &gt; \frac{b_1}{b_2} )</td>
<td>No</td>
<td>( a_1 )</td>
<td>( b_1 )</td>
</tr>
<tr>
<td>5. Lognormal</td>
<td>Trivial case ( \xi_1 = \xi_2 ) and ( \sigma_1 = \sigma_2 )</td>
<td>Empirical evidence ( \xi_1 = \xi_2 ) and ( \sigma_1 = \sigma_2 )</td>
<td>Empirical evidence ( \xi_1 = \xi_2 ) and ( \sigma_1 = \sigma_2 )</td>
<td>Empirical evidence ( \xi_1 = \xi_2 ) and ( \sigma_1 = \sigma_2 )</td>
<td>( \frac{\sigma_1}{\sigma_2} )</td>
<td>( \frac{\sigma_1}{\sigma_2} )</td>
</tr>
<tr>
<td></td>
<td>( \xi_1 &lt; \xi_2 ) and ( \sigma_1 &lt; \sigma_2 )</td>
<td>( \xi_1 &gt; \xi_2 ) and ( \sigma_1 &gt; \sigma_2 )</td>
<td>( \xi_1 &gt; \xi_2 ) and ( \sigma_1 &gt; \sigma_2 )</td>
<td>( \xi_1 &lt; \xi_2 ) and ( \sigma_1 &lt; \sigma_2 )</td>
<td>( \xi_1 &lt; \xi_2 ) and ( \sigma_1 &lt; \sigma_2 )</td>
<td></td>
</tr>
<tr>
<td>6. Truncated</td>
<td>( \xi_1 = \xi_2 ) and ( \sigma_1 = \sigma_2 )</td>
<td>( \beta_1 = \beta_2 ) and ( \xi_1 &lt; \xi_2 )</td>
<td>( \beta_1 \neq \beta_2 )</td>
<td>( \beta_1 \neq \beta_2 )</td>
<td>( \beta_1 \neq \beta_2 )</td>
<td>( \beta_1 \neq \beta_2 )</td>
</tr>
<tr>
<td>Logistic</td>
<td>( \xi_1 &gt; \xi_2 ) and ( \sigma_1 &lt; \sigma_2 )</td>
<td>( \xi_1 &lt; \xi_2 ) and ( \sigma_1 &lt; \sigma_2 )</td>
<td>( \xi_1 &lt; \xi_2 ) and ( \sigma_1 &lt; \sigma_2 )</td>
<td>( \xi_1 &gt; \xi_2 ) and ( \sigma_1 &gt; \sigma_2 )</td>
<td>( \xi_1 &gt; \xi_2 ) and ( \sigma_1 &gt; \sigma_2 )</td>
<td></td>
</tr>
<tr>
<td>7. Comma</td>
<td>( a_1 = a_2 = 1 )</td>
<td>Empirical evidence ( a_1 = a_2 = 1 ) and ( b_1 &gt; b_2 ) or ( b_1 &lt; b_2 )</td>
<td>Empirical evidence ( a_1 = a_2 = 1 ) and ( b_1 &gt; b_2 ) or ( b_1 &lt; b_2 )</td>
<td>Empirical evidence ( a_1 = a_2 = 1 ) and ( b_1 &gt; b_2 ) or ( b_1 &lt; b_2 )</td>
<td>( \frac{a_1}{a_2} )</td>
<td>( \frac{a_1}{a_2} )</td>
</tr>
<tr>
<td>8. Burr</td>
<td>( a_1 = a_2 )</td>
<td>No</td>
<td>No</td>
<td>Always</td>
<td>( 0, \ a_1 &gt; a_2 )</td>
<td>( 0, \ a_1 &lt; a_2 )</td>
</tr>
</tbody>
</table>
3.3.1.1. HRF for Two Weibull Distributions.

From Table 3.4 we have

\[ \theta(t) = \frac{\mu_1(t)}{\mu_2(t)} = \frac{\mu(t;\alpha_1,\kappa_1)}{\mu(t;\alpha_2,\kappa_2)} = \left( \frac{\kappa_1}{\kappa_2} \right)^{\kappa_2} \left( \frac{\alpha_2^{\kappa_2}}{\kappa_1^{\kappa_1}} \right) t^{\kappa_1 - \kappa_2}, \quad \alpha_1, \kappa_1 > 0, \quad g = 1, 2, \]

\[ = K t^d, \quad (3.3.5) \]

where \( K = \left( \frac{\kappa_1}{\kappa_2} \right)^{\kappa_2} \left( \frac{\alpha_2}{\kappa_1} \right)^{\kappa_1} \) and \( d = \kappa_1 - \kappa_2. \)

The HRF is monotonic increasing (decreasing) for \( \kappa_1 > \kappa_2 \) (\( \kappa_1 < \kappa_2 \)). When \( d \) is zero the HRF is a constant for all \( t \); this includes the special case \( \kappa_1 = \kappa_2 = 1 \), the HRF for two exponential distributions. The three general shapes of the HRF for two Weibull distributions are shown in Figure 3.8 (i.e., where \( d < 0, 0 < d < 1, \) and \( d > 1 \)).

3.3.1.2. HRF for Two Makeham Gompertz Distributions.

From Table 3.4 we have

\[ \theta(t) = \frac{A_1 + B_1 e^{\kappa_1 t}}{A_2 + B_2 e^{\kappa_2 t}}, \quad A_1, B_1, A_2, B_2 > 0, \quad g = 1, 2. \quad (3.3.6) \]

The HRF is an increasing (decreasing) function for \( \kappa_1 > \kappa_2 \) (\( \kappa_1 < \kappa_2 \)) and \( A_1 = A_2 = 0 \). The HRF is constant for all values of \( t \) when \( A_1 = A_2 = 0 \) and \( \kappa_1 = \kappa_2 \). From consideration of the first derivative of \( \theta(t) \) with respect to \( t \), the HRF is unimodal if \( \kappa_2 > \kappa_1 \) and

\[ \{A_2 B_1 \kappa_1 - [A_1 B_2 \kappa_2 + (\kappa_2 - \kappa_1) B_1 B_2]\} > 0, \] (see Appendix B1). A special
case of the Makeham distribution is the Gompertz distribution \((A = 0)\), i.e., an extreme value type I distribution truncated from below at \(t = 0\). A simple relationship exists between a Weibull and a Gompertz distribution. If \(T\) has a Weibull distribution \(Y = \ln T\) has a Gompertz distribution.

The HRF's for two sets of parameter values are given in Figure 3.9.
3.3.1.3. HRF for Two Linear Hazard Distributions.

From Table 3.4 we have

\[ \theta(t) = \frac{\alpha_1 + \beta_1 t}{\alpha_2 + \beta_2 t}, \quad \alpha_1, \alpha_2 > 0. \] \hspace{1cm} (3.3.7)

The HRF is monotonic increasing (decreasing) for \( \frac{\alpha_1}{\alpha_2} < \frac{\beta_1}{\beta_2} \) \( \left( \frac{\alpha_1}{\alpha_2} > \frac{\beta_1}{\beta_2} \right) \).

It is a constant for \( \alpha_1 = \alpha_2 = 0 \) and for \( \frac{\alpha_1}{\alpha_2} = \frac{\beta_1}{\beta_2} \).

The HRF's for two sets of parameter values are given in Figure 3.10.
3.3.1.4. HRF for Two Lognormal Distributions.

From Table 3.4 we have

\[
\theta(t) = \left\{ \frac{\phi(z_1)}{\sigma_1 t [1 - \phi(z_1)]} \right\} \left\{ \frac{\sigma_2 t [1 - \phi(z_2)]}{\phi(z_2)} \right\},
\]

\[
= \frac{\sigma_2 M(z_2)}{\sigma_1 M(z_1)}, \quad \sigma_1, \sigma_2 > 0,
\]

where \( \phi(z_g) \) is the standard normal cdf, \( \phi(z_g) \) is the corresponding pdf with \( z_{g} = \frac{\ln t - \tau_{g}}{\sigma_{g}} \), and \( M(z_g) \) is the reciprocal of the hazard function commonly called the Mills Ratio, \( g = 1, 2 \).
The behavior of the HRF for two lognormal distributions is difficult to assess analytically. From (3.3.9) we see that the HRF for two lognormal distributions (each of which is unimodal) is of the same form as the HRF for two normal distributions (each of which is monotonic increasing). The behavior of the HRF was investigated by some graphical methods. The parameter values for the pairs of hazard functions we considered are given in Table 3.6; Figure 3.11 illustrates the HRF for one of the pairs of parameter values.

It appears that the HRF is monotonic increasing (decreasing) for \( \xi_1 = \xi_2 \) and \( \sigma_1 > \sigma_2 \) (\( \sigma_1 < \sigma_2 \)) and also for \( \xi_1 > \xi_2 \) and \( \sigma_1 < \sigma_2 \) (\( \xi_1 < \xi_2 \) and \( \sigma_1 > \sigma_2 \)). For some function of the parameter values the HRF may be unimodal or have more than one point of inflection.

3.3.1.5. HRF for Two (Left) Truncated Logistic Distributions.

Recall from Remark 2 in Section 3.1 that truncation on the left (i.e., \( t \geq 0 \) in this case) does not affect the hazard function. From Table 3.4 we have

\[
\theta(t) = \frac{\beta_2 [1 + \exp(-y_2)]}{\beta_1 [1 + \exp(-y_1)]},
\]

where \( y_g = \frac{t - \xi_g}{\beta_g}, \xi_g, \beta_g > 0 \) for \( g = 1, 2 \).

The HRF is constant only in the trivial case of \( \xi_1 = \xi_2 \) and \( \beta_1 = \beta_2 \). From consideration of the first derivative of \( \theta(t) \) with respect to \( t \) (or more conveniently, of \( \log \theta(t) \)) we can conclude that the HRF is monotonic increasing (decreasing) for \( \beta_1 = \beta_2 \) and \( \xi_1 > \xi_2 \) (\( \xi_1 < \xi_2 \)) (see Appendix B2). It is unimodal when \( \beta_1 \neq \beta_2 \).

The HRF's for two sets of parameter values are given in Figure 3.12.
<table>
<thead>
<tr>
<th></th>
<th>$\pi_1$</th>
<th>$\zeta_1$</th>
<th>$\sigma_1$</th>
<th>$\pi_2$</th>
<th>$\zeta_2$</th>
<th>$\sigma_2$</th>
<th>Comment on $\theta(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>0.0</td>
<td>1.30</td>
<td>0.0</td>
<td>1.35</td>
<td>monotonic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2)</td>
<td>0.0</td>
<td>1.30</td>
<td>0.0</td>
<td>1.40</td>
<td>monotonic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3)</td>
<td>0.0</td>
<td>1.50</td>
<td>0.0</td>
<td>1.55</td>
<td>monotonic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4)</td>
<td>0.0</td>
<td>1.20</td>
<td>0.0</td>
<td>1.50</td>
<td>monotonic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5)</td>
<td>0.2</td>
<td>1.00</td>
<td>0.2</td>
<td>1.50</td>
<td>monotonic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6)</td>
<td>0.5</td>
<td>1.30</td>
<td>0.5</td>
<td>1.35</td>
<td>monotonic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7)</td>
<td>0.5</td>
<td>1.60</td>
<td>0.0</td>
<td>1.75</td>
<td>monotonic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8)</td>
<td>0.25</td>
<td>1.75</td>
<td>0.0</td>
<td>1.80</td>
<td>monotonic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(9)</td>
<td>0.0</td>
<td>1.00</td>
<td>1.0</td>
<td>1.00</td>
<td>&gt; 1 inflection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(10)</td>
<td>0.3</td>
<td>1.00</td>
<td>0.2</td>
<td>1.00</td>
<td>&gt; 1 inflection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(11)</td>
<td>0.5</td>
<td>1.00</td>
<td>1.0</td>
<td>1.00</td>
<td>&gt; 1 inflection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(12)</td>
<td>1.5</td>
<td>1.00</td>
<td>1.0</td>
<td>1.00</td>
<td>&gt; 1 inflection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(13)</td>
<td>0.0</td>
<td>1.00</td>
<td>-0.5</td>
<td>1.00</td>
<td>&gt; 1 inflection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(14)</td>
<td>0.0</td>
<td>1.20</td>
<td>0.5</td>
<td>1.20</td>
<td>&gt; 1 inflection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(15)</td>
<td>0.0</td>
<td>1.20</td>
<td>1.0</td>
<td>1.20</td>
<td>&gt; 1 inflection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(16)</td>
<td>0.0</td>
<td>1.30</td>
<td>1.0</td>
<td>1.30</td>
<td>&gt; 1 inflection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(17)</td>
<td>0.5</td>
<td>1.30</td>
<td>1.0</td>
<td>1.30</td>
<td>&gt; 1 inflection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(18)</td>
<td>0.0</td>
<td>1.50</td>
<td>0.5</td>
<td>1.50</td>
<td>&gt; 1 inflection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(19)</td>
<td>0.0</td>
<td>1.30</td>
<td>1.0</td>
<td>1.40</td>
<td>&gt; 1 inflection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(20)</td>
<td>0.0</td>
<td>1.00</td>
<td>0.5</td>
<td>1.50</td>
<td>unimodal</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
FIGURE 3.11

HRF FOR TWO LOGNORMAL DISTRIBUTIONS

$\mu_1(t;0,0,1,0)$ and $\mu_2(t;0,5,1,50)$

FIGURE 3.12

HRF FOR TWO TRUNCATED LOGISTIC DISTRIBUTIONS

$\mu_1(t;10,5)$ and $\mu_2(t;15,10)$
3.3.1.6. HRF for Two Gamma Distributions.

From Table 3.4 we have

\[
\theta(t) = \frac{\Gamma(\alpha_2) \beta_2^\alpha_2 1 - F_2(t) \alpha_1 - \alpha_2 - t(1/\beta_1 - 1/\beta_2)}{\Gamma(\alpha_1) \beta_1^\alpha_1 1 - F_1(t) e^t}, \quad \alpha_g, \beta_g > 0 \tag{3.3.11}
\]

where \( \Gamma(\alpha_g) \) is the gamma function and \( F_g(t) = \int_0^t \frac{(\alpha_g - 1) - x/\beta_g}{\Gamma(\alpha_g) \beta_g} e^{-x/\beta_g} dx \), \( g = 1, 2 \). A special case is the exponential distribution \( (\alpha = 1) \), and the chi-square distribution \( (\alpha = 2\nu, \beta = 2) \) with \( \nu \) degrees of freedom.

The HRF is difficult to assess analytically and is investigated using graphical techniques. An example of the HRF for a set of parameter values is given in Figure 3.13. For all cases considered it appears that the HRF can be either monotonic or unimodal. For some parameter values such that \( \alpha_1 < \alpha_2 \) and \( \beta_1 < \beta_2 \) or \( \alpha_1 > \alpha_2 \) and \( \beta_1 > \beta_2 \) it is unimodal. In other cases it is monotonic.

3.3.1.7. HRF for Two Burr Distributions.

From Table 3.4 we have

\[
\theta(t) = \frac{\alpha_1 \beta_1 [1 + t^{-\alpha_2}]}{\alpha_2 \beta_2 [1 + t^{-\alpha_1}]}, \quad \alpha_g, \beta_g > 0, \tag{3.3.12}
\]

The HRF is constant for \( \alpha_1 = \alpha_2 \). From consideration of the first derivative of \( \theta(t) \) with respect to \( t \) it can be shown that the HRF is always unimodal, (see Appendix B3).

The HRF's for two sets of parameter values are given in Figure 3.14.
FIGURE 3.13

HRF FOR TWO GAMMA DISTRIBUTIONS

\[ u_1(t;6,4) \text{ and } u_2(t;8,5) \]

TABLE 3.7

PARAMETER VALUES OF THE HRF FOR PAIRS OF GAMMA DISTRIBUTIONS

<table>
<thead>
<tr>
<th>( \pi_1 )</th>
<th>( \pi_2 )</th>
<th>Comment on ( \theta(t) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_1 )</td>
<td>( \beta_1 )</td>
<td>( \alpha_2 )</td>
</tr>
<tr>
<td>(1)</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>(2)</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>(3)</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>(4)</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>(5)</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>(6)</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>(7)</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>(8)</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>(9)</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>(10)</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>(11)</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>(12)</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>(13)</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>(14)</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>(15)</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>(16)</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>
3.3.2. The HRF for Pairs of Parametric Models from Different Families of Distributions.

One usually finds it necessary to assume that the survivorship functions of the two populations being compared are represented by the same parametric model with different parameters. Of course, it is possible to have situations in which the hazard functions for $\pi_1$ and $\pi_2$ are from two different parametric families of distributions. In particular, when the shapes of the hazard functions are very different we may find it unreasonable to attempt to describe the lifetime distributions by the same parametric model. As we have seen in Section 3.3, the form of the HRF, even for pairs of hazard functions from the same family of distributions, is not necessarily simple to describe. The HRF for pairs of hazard functions from different families of distributions would often be even more difficult to describe. We will not consider any of these possible cases here.
CHAPTER IV

ESTIMATION OF $\Delta$ UNDER SOME NON-PHF MODELS

In Chapter II we considered the special case, called the PHF model, where the hazard ratio function (HRF) $\theta(t) = \mu_1(t)/\mu_2(t)$ is assumed to be constant for all $t$, i.e. $\theta(t) = \theta$ for all $t$. We derived some estimators $\hat{\Delta}_1$, $\hat{\Delta}_2$, $\hat{\Delta}_3$ and $\hat{\Delta}_4$ of the parameter $\Delta = \theta/(1 + \theta)$ in this model (Sections 2.2 and 2.3). For simplicity of notation, throughout this chapter we will use $\hat{\Delta}_4$ as equivalent to $\hat{\Delta}_4$, so that when we drop the subscript there is no confusion. The estimators $\hat{\Delta}$ are functions of the numbers at risk in $\pi_1$ and $\pi_2$ and the populations from which a death occurs. The importance of the PHF model lies in its frequent use for comparing and for estimating survival functions.

In Chapter III we investigated the more general situation in which the HRF is a function of $t$. When we can specify the parametric family of the failure distributions, the HRF may be a well-behaved function of $t$, and in some cases a constant (see Table 3.5). For many life table models we found that the HRF is not constant, except perhaps over a restricted range of values of $t$. The use of empirical evidence to draw conclusions about the form of the HRF may be difficult, especially for small samples. Thus, the results presented in Chapter III suggest that the PHF assumption may be valid only under certain conditions on the parameter values or on the range of values of $t$. We need to ascertain
what is the "effect" of a non-PHF model (i.e., a nonconstant HRF) on the estimation of \( \Delta \) under a PHF assumption.

In most of the chapter we will assume that we can specify the parametric family of the failure distributions. We will therefore know the HRF \( \theta(t) \) and the hazard delta function
\[
\Delta(t) = \frac{\mu_1(t)}{[\mu_1(t) + \mu_2(t)]} = \frac{\theta(t)}{[1 + \theta(t)]}.
\]

In Section 4.2 we will present the results of a Monte Carlo simulation study which was designed to compare the effects of a non-constant HRF on the estimators \( \hat{\Delta}_l \), \( l = 1, \ldots, 4 \). In Section 4.3 we will suggest some alternatives to the PHF model and outline methods for assessing the adequacy of the PHF model with respect to these alternatives.

4.1. The Estimator \( \hat{\Delta} \) as an Estimator of an "Average" of the Delta Function \( \Delta(t) \).

For any experiment let \( t_1 < \ldots < t_k \) denote \( k \) observed (ordered) times of death. Assuming that we know \( \mu_1(t) \) and \( \mu_2(t) \) we then know \( \Delta(t) \) at each of the times of death. If we assume that the PHF model is valid, we obtain an estimator \( \hat{\Delta} \) of a "\( \Delta \)". This is a constant that may be regarded as some "average" of the \( \Delta(t_i) \)'s, \( i = 1, \ldots, k \). Let \( A[\Delta(t_1), \ldots, \Delta(t_k)] = A[\hat{\Delta}(t)] \) denote the functional form of this average. We will consider the function
\[
d(t; \hat{\Delta}) = A[\hat{\Delta}(t)] - \hat{\Delta} \tag{4.1.1}
\]
as a measure of the deviation of this average from the estimator \( \hat{\Delta} \).

It is apparent that the distribution of \( d(t; \hat{\Delta}) \) in repeated samples would be extremely difficult to determine. It depends on the joint
distribution of the random variables $\Delta(t_i), i = 1, \ldots, k$, and the estimator $\tilde{\Delta}$, which is a function of the set $\{\hat{p}_i\}$ under $\Delta(t)$. In theory we can derive the distribution of $d(t;p)$, but in practice it is almost impossible.

However, we suggest an approximation to the expected value of the estimators, $\tilde{\Delta}_\ell, \ell = 1, \ldots, 4$, with regard to the underlying failure distributions, and consider the deviation measured by the difference between the approximation and $\tilde{\Delta}_\ell$.

To be more specific, we take

$$A_1[\Delta(t)] = E \left[ \frac{\sum_i \alpha_i r_{2i} \hat{p}_i}{\sum_i \alpha_i r_{2i} \hat{p}_i + \sum_i \alpha_i r_{1i} (1 - \hat{p}_i)} \right]$$

$$= \frac{\sum_i \alpha_i r_{2i} E[\hat{p}_i]}{\sum_i \alpha_i r_{2i} E[\hat{p}_i] + \sum_i \alpha_i r_{1i} E[1 - \hat{p}_i]}, \quad (4.1.2)$$

where $\sum_i \alpha_i$ represents the set of weights corresponding to the estimator $\Delta$, and

$$E[\hat{p}_i] = \frac{r_{1i} \Delta(t_i)}{r_{1i} \Delta(t_i) + r_{2i}}$$

$$= \frac{r_{1i} \Delta(t_i)}{[r_{2i} + \Delta(t_i)(r_{1i} - r_{2i})]} = \frac{r_{1i} \Delta(t_i)}{D_i}, \quad (4.1.3)$$

$$i = 1, \ldots, k,$$

with $\theta(t_i) = \Delta(t_i) / [1 - \Delta(t_i)]$ and $D_i = [r_{2i} + \Delta(t_i)(r_{1i} - r_{2i})]$.

Explicitly, the four forms of $A_1[\Delta(t)]$ which correspond to the set of weights $\{\alpha_i\}$ for each of the four estimators $\tilde{\Delta}_\ell, \ell = 1, \ldots, 4$ are given below.
\[ A_{11} [\Delta(t)] = \sum_{i} \frac{r_{1i} r_{2i} \Delta(t_i)}{D_i} \sum_{i} \frac{r_{1i} r_{2i}}{D_i} \]  

(4.1.5)

\[ A_{12} [\Delta(t)] = \sum_{i} \frac{r_{1i} r_{2i} \Delta(t_i)}{(r_{1i} + r_{2i}) D_i} \sum_{i} \frac{r_{1i} r_{2i}}{(r_{1i} + r_{2i}) D_i} \]  

(4.1.6)

\[ A_{13} [\Delta(t)] = \sum_{i} \frac{r_{1i} r_{2i} \Delta(t_i)}{(r_{1i} K + r_{2i}) D_i} \sum_{i} \frac{r_{1i} r_{2i}}{(r_{1i} K + r_{2i}) D_i} \]  

(4.1.7)

where

\[ K = \sum_{i} \frac{r_{1i} r_{2i} \Delta(t_i)}{(r_{1i} + r_{2i}) D_i} \sum_{i} \frac{r_{1i} r_{2i}}{(r_{1i} + r_{2i}) D_i} \]  

\[ A_{14} [\Delta(t)] = \sum_{i} \frac{r_{1i} r_{2i} \Delta(t_i) [1 - \Delta(t_i)]}{D_i} \sum_{i} \frac{r_{1i} r_{2i} [1 - \Delta(t_i)]}{D_i} \]  

(4.1.8)

The corresponding deviations are

\[ d_{1\ell} (t, \hat{p}) = A_{1\ell} [\Delta(t)] - \hat{\Delta}_\ell, \quad \ell = 1, \ldots, 4. \]  

(4.1.9)

We also consider some rather simple weighted averages of the form

\[ A[\Delta(t)] = \frac{\sum_{i} w(t_i) \Delta(t_i)}{\sum_{i} w(t_i)}, \]  

(4.1.10)

where \( w(t_i) \) is a weight which is possibly a function of \( t_i \).
The simplest type of weight would be \( w(t_i) = \frac{1}{k} \) for all \( i = 1, \ldots, k \). This gives

\[
A_2(\Delta(t)) = \frac{1}{k} \sum_i \Delta(t_i),
\]

(4.1.11)

with corresponding measure of deviation

\[
d_2(\hat{\lambda}; \hat{\pi}) = A_2(\Delta(t)) - \hat{\Delta}.
\]

(4.1.12)

A second type of weight which does depend on \( t \) is one that is a function of the numbers at risk of dying in \( \pi_1 \) and \( \pi_2 \) at each of the times of death, such as \( w(t_i) = \frac{r_{1i}r_{2i}}{r_{1i} + r_{2i}} \). This gives

\[
A_3(\Delta(t)) = \sum_i \frac{r_{1i}r_{2i}/\Delta(t_i)}{r_{1i} + r_{2i}},
\]

(4.1.13)

with corresponding measure of deviation

\[
d_3(\hat{\lambda}; \hat{\pi}) = A_3(\Delta(t)) - \hat{\Delta}.
\]

(4.1.14)

Weights such as those suggested for \( A_3(\Delta(t)) \) are reasonable if the extreme values of \( \Delta(t) \), occurring at early values of \( t \), are given weights which reflect the larger numbers of individuals at risk of dying during the initial follow-up period.

A Monte Carlo simulation study is described in Section 4.2, which was performed to investigate

(i) comparisons among the estimators \( \hat{\Delta}_\lambda \), \( \lambda = 1, \ldots, 4 \), for some nonconstant HRF models, and

(ii) comparisons among the three "averages" of the \( \Delta(t_i) \)'s
which are assumed to be close to the estimators $\hat{\Delta}_g^x$,
$x = 1, \ldots, 4$.

4.2. Monte Carlo Simulation Study.

Recall the following notation

$r_g = \text{the } k \times 1 \text{ vector of values representing the number of }$
$\Pi_g \text{ individuals from } \Pi_g \text{ who are at risk of dying at each of }$
$\text{the } k \text{ times of death.}$

$\nu_g(t; \xi_{g1}, \xi_{g2}) = \text{the hazard function of } \Pi_g, \text{ indexed by }$
$\text{parameters } \xi_{g1} \text{ and } \xi_{g2}, \ g = 1, 2.$

$t_i = \text{time of the } i^{th} \text{ death, } i = 1, \ldots, k.$

We considered only two parametric models for the hazard functions
of $\Pi_g$, the Weibull and the Gompertz distributions. Using any suitable
random number generator for exponential distributions we can generate
(random) times of death for these two distributions by an appropriate
transformation. In particular, if $X$ is a random variable with a unit
exponential distribution, then

$$T = \frac{1}{\xi_2} X^{1/\xi_2} \quad (4.2.1)$$

has a Weibull distribution and

$$T = \frac{1}{\xi_2} \ln \left[ \frac{\left(\frac{\xi_2}{\xi_1}\right)^X + 1}{\xi_1} \right] \quad (4.2.2)$$

has a Gompertz distribution of the forms given in Table 3.4 where
$\xi_1 = \alpha, \xi_2 = \kappa$ and $\xi_1 = \beta$ and $\xi_2 = \kappa$ respectively. For this study
we considered the HRF for pairs of distributions from the same family of
distributions. We used the random number generator program VARGEN, which
is in the library of the Computer Science facilities at the University of North Carolina.

4.2.1. SSI Simulation Study.

For Sample Scheme I we let \( r_1 \) and \( r_2 \), the vectors of values of the number at risk in \( \pi_1 \) and \( \pi_2 \), respectively, for \( k = 40 \) deaths be

\[
r'_1 = [20 \ 19 \ 19 \ 18 \ 18 \ 17 \ 17 \ 16 \ 16 \ 15 \ 15 \ 14 \ 14 \ 13 \ 13 \ 12 \ 11 \ 10 \ 9 \ 9 \ 8 \ 8 \\
7 \ 7 \ 7 \ 7 \ 7 \ 6 \ 6 \ 6 \ 6 \ 5 \ 5 \ 4 \ 4 \ 3 \ 3 \ 2 \ 1 \ 1],
\]

and

\[
r'_2 = [20 \ 20 \ 19 \ 19 \ 18 \ 18 \ 18 \ 17 \ 17 \ 17 \ 16 \ 16 \ 15 \ 15 \ 15 \ 15 \ 15 \ 15 \ 14 \ 14 \ 14 \ 13 \\
13 \ 12 \ 11 \ 10 \ 9 \ 8 \ 8 \ 7 \ 6 \ 6 \ 5 \ 4 \ 4 \ 3 \ 3 \ 2 \ 2 \ 1 \ 1 \ 1].
\]

At each of the deaths \((i = 1, \ldots, k)\) we generate a time of death \( T_{gij} = t_{gij} \) from the distribution with hazard function \( \mu_g(t; \xi_{g1}, \xi_{g2}) \) for each of the \( r_{gi} \) individuals at risk of dying, \( g = 1, 2, i = 1, \ldots, k, j = 1, \ldots, r_{gi} \). The time of the \( i^{th} \) death is given by

\[
t_i = \min[t_{1i1}, \ldots, t_{1ir_{1i}}, t_{2i1}, \ldots, t_{2ir_{2i}}].
\]

We require that

\[
T_{gij} > t_{i-1},
\]

i.e. each time of death for individuals at risk of dying at the \( i^{th} \) death is greater than the observed time of the \((i - 1)^{th}\) death, with \( t_0 = 0 \). We then determine the observed values of the \( \hat{p}_i \)'s according to the definition used throughout, i.e.
\[ \hat{p}_i = \begin{cases} 1 & \text{if } t_i \text{ is from } \pi_1, \\ 0 & \text{if } t_i \text{ is from } \pi_2. \end{cases} \]

Given the vectors \( r_1, r_2, \) and the resulting vector \( \tilde{\mathbf{p}} \), we obtain the estimates of \( \Delta - \Delta_\delta \), \( \xi = 1, \ldots, 4 \) which were derived in Section 2.2 under the assumption of a constant HRF. We then find \( \Delta(t_i) \), the delta function at each of the \( t_i \) times of death in this sample, \( i = 1, \ldots, k \). Calculation of the "averages" \( A_1[\Delta(t)], A_2[\Delta(t)] \) and \( A_3[\Delta(t)] \) are straightforward from (4.1.5)-(4.1.8), (4.1.11) and (4.1.13), respectively. The values of the deviations \( d_1(t_i, \tilde{\mathbf{p}}), d_2(t_i, \tilde{\mathbf{p}}) \) and \( d_3(t_i, \tilde{\mathbf{p}}) \) follow immediately.

These steps were repeated for 200 samples.

4.2.2. Discussion of SSI Simulation Study.

A summary of the first two sample moments of the four estimators, for \( k = 40 \) deaths and \( r_1 \) and \( r_2 \) given by (4.2.3) and (4.2.4), respectively, and for three choices of the parametric forms of \( \mu (t; \xi, \xi_1, \xi_2) \) is given in Table 4.1. Plots of the respective delta functions are given in Figures 4.1-4.3. Also included in the figures are histograms representing the frequency distributions of the 200 sample estimates given by the four estimators, and the sample means. Note that the histograms have a vertical scale as a base.

From these results we may reasonably conclude the following.

(i) The estimators \( \tilde{\Delta}_2, \tilde{\Delta}_3 \) and \( \tilde{\Delta}_4 \) are similar in regard to their first two sample moments. The distributions of \( \tilde{\Delta}_3 \) and \( \tilde{\Delta}_4 \) are also very similar.

(ii) \( \tilde{\Delta}_1 \) always has the largest variance. When the true delta function is monotonic decreasing (increasing) \( \tilde{\Delta}_1 \) tends to
### TABLE 4.1

**SUMMARY OF 200 SAMPLE ESTIMATES OF "Δ" UNDER SOME SPECIFIED PARAMETRIC MODELS, SSI**

<table>
<thead>
<tr>
<th>(a)</th>
<th>Model specified by two Weibull distributions, $\mu_1(t;50,1.5)$ and $\mu_2(t;60,2.0)$, $k = 40$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>$\hat{\Delta}$</td>
<td>$\hat{\Delta}$</td>
</tr>
<tr>
<td>$\Delta_1$</td>
<td>.5956</td>
</tr>
<tr>
<td>$\Delta_2$</td>
<td>.5384</td>
</tr>
<tr>
<td>$\Delta_3$</td>
<td>.5387</td>
</tr>
<tr>
<td>$\Delta_4$</td>
<td>.5387</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(b)</th>
<th>Model specified by two Compertz distributions, $\mu_1(t;.005,.05)$ and $\mu_1(t;.008,.04)$, $k = 40$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>$\hat{\Delta}$</td>
<td>$\hat{\Delta}$</td>
</tr>
<tr>
<td>$\Delta_1$</td>
<td>.4576</td>
</tr>
<tr>
<td>$\Delta_2$</td>
<td>.4694</td>
</tr>
<tr>
<td>$\Delta_3$</td>
<td>.4689</td>
</tr>
<tr>
<td>$\Delta_4$</td>
<td>.4689</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(c)</th>
<th>PHF model specified by two Weibull distributions, $\mu_1(t;50,1.5)$ and $\mu_2(t;50,1.5)$, $k = 40$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>$\hat{\Delta}$</td>
<td>$\hat{\Delta}$</td>
</tr>
<tr>
<td>$\Delta_1$</td>
<td>.5022</td>
</tr>
<tr>
<td>$\Delta_2$</td>
<td>.4998</td>
</tr>
<tr>
<td>$\Delta_3$</td>
<td>.4994</td>
</tr>
<tr>
<td>$\Delta_4$</td>
<td>.4994</td>
</tr>
</tbody>
</table>
FIGURE 4.1

HISTOGRAMS OF 200 SAMPLE ESTIMATES OF $\Delta$ WHEN $\Delta(t)$ IS SPECIFIED BY TWO WEIBULL DISTRIBUTIONS, $SSI$

$\mu_1(t;50,1.5)$ and $\mu_2(t;60,2.0)$
FIGURE 4.2
HISTOGRAMS OF 200 SAMPLE ESTIMATES OF A MEAN \( \mu \) (\( t \)) IS SPECIFIED BY THE Gompertz DISTRIBUTIONS, SSI

\[ \mu_1 \] (\( t; 0.005, 0.05 \)) and \[ \mu_2 \] (\( t; 0.008, 0.04 \))
FIGURE 4.3

HISTOGRAMS OF 200 SAMPLE ESTIMATES OF $A$ WHEN $A(t)$ IS SPECIFIED BY TWO WEIBULL DISTRIBUTIONS, SSI

$\mu_1(t;50,1.5)$ and $\mu_2(t;50,1.5)$
give a larger (smaller) estimate of $\Delta$ than the other estimators.

(iii) Assuming normality of the mean of the 200 sample estimates $\tilde{\Delta}, \tilde{\Delta}$, with an estimated standard error (SE) given by $SE(\tilde{\Delta}) = SE(\tilde{\Delta})/\sqrt{200}$, when the PHF model is valid (Table 4.1c and Figure 4.3) there is no evidence that the bias in $\tilde{\Delta}$ is significantly different from zero, for all the cases investigated.

In contrast to the conclusions suggested by the results of Section 2.5.1, we find that $\tilde{\Delta}_2$, not $\tilde{\Delta}_3$, has the smallest mean square, and that $\tilde{\Delta}_1$ does not always have the smallest bias.

(iv) A comparison among the absolute values of the deviations defined in Section 4.1, suggests that

(a) $A_2[\tilde{\Delta}(t)]$, the unweighted average of the $\Delta(t_1)$'s, most closely compares to $\tilde{\Delta}_2$ and to $\tilde{\Delta}_3$.

(b) $A_3[\tilde{\Delta}(t)]$, the weighted average of the $\Delta(t_1)$'s, most closely compares to $\tilde{\Delta}_1$.

(c) Only for $\tilde{\Delta}_4$ is the expected average $A_{14}[\tilde{\Delta}(t)]$ closer to $\tilde{\Delta}_4$ than $A_\lambda[\tilde{\Delta}(t)], \lambda = 2,3$.

4.2.3. SSII Simulation Study.

The simulation study for Sample Scheme II is much simpler to design. For each of $r_{gl}$ individuals from $\pi_g$, we generate a time of death $t_{gj}$ from the distribution with hazard function $u_g(t;\xi_{g1},\xi_{g2})$, $g = 1,2$, $j = 1,\ldots,r_{gl}$. We then order the resulting observed times of death $\left\{t_{1j}\right\}$ and $\left\{t_{2j}\right\}$, giving $\xi$, the vector of (ordered) times of death from $\pi_1$ and $\pi_2$. We can then determine the corresponding
vector $\hat{p}$. The numbers at risk of dying at each of the $k = r_{11} + r_{21}$ times of death are given by

$$r_{1i} = r_{11} - \sum_{j<i} \hat{p}_j,$$  \hspace{1cm} (4.2.7)

and

$$r_{2i} = r_{21} - \sum_{j<i} (1 - \hat{p}_j), \hspace{1cm} i = 1, \ldots, k. \hspace{1cm} (4.2.8)$$

Estimation of $\Delta$ and calculation of the "averages" of the delta function at the times of death are as described in Section 4.2.1.

These steps were repeated for 200 samples.

4.2.4. Discussion of SSII Simulation Study.

A summary of the first two sample moments of the four estimators, for $r_{11} = r_{21} = 20$ (or $k = 40$), and for three choices of the parametric forms of $\mu(t; \xi_1, \xi_2)$ is given in Table 4.2. Plots of the respective delta functions are given in Figures 4.4-4.6. Included in the figures are histograms representing the 200 sample estimates of $\Delta$ given by the four estimators; their sample means are also indicated.

From these results we may reasonably conclude the following.

(i) The estimators $\tilde{\Delta}_2$, $\tilde{\Delta}_3$ and $\tilde{\Delta}_4$ are similar in regard to their first two sample moments. The distributions of $\Delta_3$ and $\tilde{\Delta}_4$ are also very similar.

(ii) $\tilde{\Delta}_1$ usually has the largest variance, but for an example not reported (the delta function of Figure 4.4 and $r_{11} = r_{21} = 10$) the variance of $\tilde{\Delta}_2$ was largest.

(iii) The variances of the estimators decrease as $r_{11}$ and $r_{21}$ increase in the same ratio.

(iv) Assuming normality of the mean of the 200 sample estimates
TABLE 4.2
SUMMARY OF 200 SAMPLE ESTIMATES OF "Δ" UNDER SOME
SPECIFIED PARAMETRIC MODELS, SSII

(a) Model specified by two Weibull distributions,
\[ \mu_1(t;50,1.5) \text{ and } \mu_2(t;60,2.0), \quad r_{11} = r_{21} = 20 \]

<table>
<thead>
<tr>
<th>Mean</th>
<th>SD</th>
<th>[A_2[\Delta(t)]]</th>
<th>[A_3[\Delta(t)]]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\tilde{\Delta})</td>
<td>(\tilde{\Delta})</td>
<td>(A_1[\tilde{\Delta}(t)])</td>
<td>(d_1(t,\hat{p}))</td>
</tr>
<tr>
<td>(\Delta_1)</td>
<td>.5926</td>
<td>.0908</td>
<td>.5963</td>
</tr>
<tr>
<td>(\Delta_2)</td>
<td>.5560</td>
<td>.0855</td>
<td>.5853</td>
</tr>
<tr>
<td>(\Delta_3)</td>
<td>.5542</td>
<td>.0837</td>
<td>.5567</td>
</tr>
<tr>
<td>(\Delta_4)</td>
<td>.5543</td>
<td>.0839</td>
<td>.5415</td>
</tr>
</tbody>
</table>

(b) Model specified by two Gompertz distributions
\[ \mu_1(t;.005,.05) \text{ and } \mu_2(t;.008,.04), \quad r_{11} = r_{21} = 20 \]

<table>
<thead>
<tr>
<th>(\tilde{\Delta})</th>
<th>(\tilde{\Delta})</th>
<th>(\hat{\Delta}(t))</th>
<th>(\hat{\Delta}(t))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta_1)</td>
<td>.4547</td>
<td>.0999</td>
<td>.4520</td>
</tr>
<tr>
<td>(\Delta_2)</td>
<td>.4703</td>
<td>.0925</td>
<td>.4489</td>
</tr>
<tr>
<td>(\Delta_3)</td>
<td>.4714</td>
<td>.0906</td>
<td>.4730</td>
</tr>
<tr>
<td>(\Delta_4)</td>
<td>.4713</td>
<td>.0907</td>
<td>.4693</td>
</tr>
</tbody>
</table>

(c) PHF model specified by two Weibull distributions
\[ \mu_1(t;50,1.5) \text{ and } \mu_2(t;50,1.5), \quad r_{11} = r_{21} = 20 \]

<table>
<thead>
<tr>
<th>(\Delta)</th>
<th>(\Delta)</th>
<th>(\hat{\Delta}(t))</th>
<th>(\hat{\Delta}(t))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta_1)</td>
<td>.5140</td>
<td>.0818</td>
<td>.5000</td>
</tr>
<tr>
<td>(\Delta_2)</td>
<td>.5116</td>
<td>.0770</td>
<td>.5000</td>
</tr>
<tr>
<td>(\Delta_3)</td>
<td>.5112</td>
<td>.0763</td>
<td>.5000</td>
</tr>
<tr>
<td>(\Delta_4)</td>
<td>.5112</td>
<td>.0764</td>
<td>.5000</td>
</tr>
</tbody>
</table>
FIGURE 4.4

HISTOGRAMS OF 200 SAMPLE ESTIMATES OF $\Lambda$ WHEN $\Lambda(t)$ IS SPECIFIED BY TWO WEIBULL DISTRIBUTIONS, SSII

$\mu_1(t;50,1.5)$ and $\mu_2(t;50,2.0)$
FIGURE 4.5

HISTOGRAMS OF 200 SAMPLE ESTIMATES OF \( \Delta(t) \) WHEN \( \Delta(t) \) IS SPECIFIED BY TWO COMPETENT DISTRIBUTIONS, SSII

\[ \mu_1(t; 0.005, 0.05) \] and \[ \mu_2(t; 0.008, 0.04) \]
FIGURE 4.6

HISTOGRAMS OF 200 SAMPLE ESTIMATES OF $A$ WHEN $A(t)$ IS SPECIFIED BY TWO WEIBULL DISTRIBUTIONS, SSII

$\mu_1(t;50,1.5)$ and $\mu_2(t;50,1.5)$
When the PHF model is valid there is evidence that the bias is significantly different from zero for \( \bar{\Delta}_1 \). We know, however, that \( \bar{\Delta}_1 \) is in fact, unbiased. The conflicting conclusion must, I think, be attributed to sampling variability for this particular parametric model and numbers initially at risk of dying in \( \pi_1 \) and \( \pi_2 \).

(v) A comparison among the deviations defined in Section 4.1 suggests that

(a) \( \bar{\Delta}_2[\Delta(t)] \) most closely compares to \( \bar{\Delta}_2, \bar{\Delta}_3 \) and \( \bar{\Delta}_4 \), but in some cases \( \bar{\Delta}_l[\Delta(t)] \) is closer to \( \bar{\Delta}_l \) for \( l = 3,4 \).

(b) \( \bar{\Delta}_1 \) most closely approximates to \( \bar{\Delta}_3[\Delta(t)] \).

4.2.5. Some Measures of the Difference Between the \( \Delta(t) \) Function Model and an Estimate \( \bar{\Delta} \) Based on the PHF Model.

There are many ways in which we can define a measure of the difference between the delta function and an estimated constant \( \bar{\Delta} \) based on the PHF model. Simple measures of this difference over some restricted range of \( t, [\tau_1, \tau_2] \), say, are of form

\[
  d = \frac{1}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} \Psi(\Delta(t) - \bar{\Delta}) dt, \quad (4.2.9)
\]

where \( \Psi(\cdot) \) is a well behaved continuous increasing function of the absolute difference \( |\Delta(t) - \bar{\Delta}| \). It is necessary to consider \( \Psi(\cdot) \) over a restricted range of \( t \) because the delta function is frequently zero or one at \( t = 0 \) or \( t = \infty \), and may greatly influence the evaluation of the integral of \( \Psi(\cdot) \). Also, the delta function may approach some limit as \( t \to \infty \) and we may have little interest in the difference.
between $\Delta(t)$ and $\tilde{\Delta}$ as $t \to \infty$, which itself would approach some constant. Restricting the range of $t$ would ignore this part of the $\Psi(\cdot)$ function, while taking $\tau_2 = \infty$ would simply give us the limiting value as $t \to \infty$.

Two possible choices for the form of $\Psi(\cdot)$ are

\begin{align}
(i) \quad & \Psi(\Delta(t) - \tilde{\Delta}) = |\Delta(t) - \tilde{\Delta}| \quad (4.2.10) \\
(ii) \quad & \Psi(\Delta(t) - \tilde{\Delta}) = (\Delta(t) - \tilde{\Delta})^2. \quad (4.2.11)
\end{align}

A serious criticism of any of these measures, however, is the arbitrariness of the $\tau_1$ and $\tau_2$ values. Some preliminary results from the Monte Carlo Studies described in Sections 4.2.1 and 4.2.3 in which $\Psi(\cdot)$ was defined by (4.2.11) indicate that the measure $d$ is very dependent upon the values for $\tau_1$ and $\tau_2$. Until further investigations are made we do not recommend using measures of this type.

4.3. Alternatives to the PHF Model.

In Chapter II we considered estimation of the parameter in the PHF model, but in Chapter III we found that in many situations the HRF was not constant. In Section 4.2 we examined properties of the estimators of $\Delta$ (derived under the PHF model) when the HRF was not constant.

On the basis of the results presented in Chapter III we will suggest the following procedures as practical methods for comparing survivorship functions. These suggestions follow two lines of thought. They are given as a possible basis for practical application and a topic for research and development.
4.3.1. Alternate Models.

We may propose some simple models as alternatives to the PHF model. We can then construct tests of the validity of the PHF model with regard to these simple alternatives.

For example, suppose that the hazards in two populations are not proportional, but have a linear additive relationship such as

$$
\mu_1(t) = \theta \mu_2(t) + \phi, \quad \theta > 0,
$$

(4.3.1)

or more generally,

$$
\mu_1(t) = \theta \mu_2(t) + \phi(t)
$$

(4.3.2)

If either (4.3.1) or (4.3.2) were valid we see that the HRF would be given by

$$
\theta(t) = \frac{\mu_1(t)}{\mu_2(t)} = \theta + \frac{\phi}{\mu_2(t)},
$$

(4.3.3)

or

$$
\theta(t) = \frac{\mu_1(t)}{\mu_2(t)} = \theta + \frac{\phi(t)}{\mu_2(t)},
$$

(4.3.4)

respectively, both of which are functions of time. Notice that the HRF given by (4.3.3) is a monotonic decreasing function if \( \mu_2(t) \) is a monotonic increasing function, with a limiting value of \( \theta \). From the analytic results of some HRF's summarized in Table 3.5, this behavior of the HRF (or delta function) is not unreasonable for some pairs of parametric models (e.g., Makeham Gompertz, linear hazard distribution, truncated logistic, gamma, and Burr).
Let us assume the linear additive model given by (4.3.1) is valid and construct a test for the appropriateness of the PHF model. More specifically, we wish to test $H_0: \phi = 0$ against the alternatives $H_A: \phi \neq 0$. We will briefly describe two tests of this hypothesis which are not based on any \textit{a priori} knowledge of the forms of the separate hazard functions.

Recall from Section 3.2 that we can obtain estimates of the hazard function by several methods. One such nonparametric method for grouped data, using an exponential approximation to the survival function within intervals for $n_g$, is

$$\hat{\mu}_g(t'_i) = -\frac{1}{n_1} \ln \hat{p}_{gi},$$

(4.3.5)

where $\hat{p}_{gi}$ is an estimate of the conditional probability of surviving the interval $[t_{i-1}, t_i)$, given alive at $t_i$, $g = 1, 2$. The midpoint of the interval is $t'_i$. We may then substitute these estimates into (4.3.1) and use a weighted least squares method to estimate $\theta$ and $\phi$. The weights usually chosen are the reciprocals of the variances of the corresponding error terms in the model. Although we can obtain an estimate of the approximate variances of $\hat{\mu}_1(t'_i)$ and $\hat{\mu}_2(t'_i)$ (which are independent), the variance of $\left[\hat{\mu}_1(t'_i) - \theta \hat{\mu}_2(t'_i) - \phi\right]$ involves the unknown parameters $\theta$ and $\phi$ which we wish to estimate. Use of these weights in an iterative procedure is a natural way to proceed.

While this method may prove enlightening with regard to a test of $H_0$, we are overlooking the fact that the estimates $\hat{\mu}_1(t'_i)$ and $\hat{\mu}_2(t'_i)$ are both random variables. This may not be very important in estimation of the parameters $\theta$ and $\phi$, but could be a source of serious error in testing $H_0$. Some of the techniques developed in structural
regression (Kendall and Stuart, Vol. II, Chapter 29, 1972) could possibly be adapted for this problem. In particular, we may have to make some assumptions regarding the variances of the observed $\hat{\mu}_g(t'_1)$, $g = 1, 2$, before we can obtain estimates of the parameters in the model.

4.3.2. Subdivision of the Follow-up Period.

Without going even as far as to specify an alternative to the PHF model we can suggest an "indirect" test for the appropriateness of the PHF model. Sometimes it is possible to divide the follow-up period into a relatively few subperiods within which a PHF model may be appropriate. If this is possible the estimates of $\Delta$ (or $\theta$) for successive subperiods can be compared. For a sufficiently large sample size, we may assume that the estimators $\hat{\Delta}$ (or $\hat{\theta}$) have a normal distribution. An (approximate) estimate of the variance of $\hat{\Delta}$ (or $\hat{\theta}$) can be derived using an approximation to the variance of the ratio of two random variables (e.g. see (2.2.2)). We may then formally test for the equality of the values of $\Delta$ (or $\theta$) over successive periods, using the estimates so obtained.
CHAPTER V

SOME APPLICATIONS OF THE PHF MODEL

The purpose of this chapter is to illustrate the techniques of estimating the parameter $\Delta$ in the PHF model, using an empirical example. We should be aware that this is not a complete analysis of the data, but only one aspect of the analysis.

The estimation of $\Delta$ in the PHF model, $u_1(t) = \theta u_2(t)$, where $\Delta = \theta/(1 + \theta)$, is straightforward. It requires a reasonably simple computing algorithm to obtain the vectors of values $r_1$, $r_2$ and $\hat{p}$ when we are given the length of time of follow-up, population identification, and possibly indication of censoring, for each individual under observation.

Before proceeding with the example we will give an approximation to the variances of the estimators of $\Delta$ derived in Section 2.4 which allow for ties.

5.1. Approximations to the Variances of the Estimators of $\Delta$ Which Allow for Ties.

For convenience, throughout this chapter we will again use $\hat{\Delta}_4$ to denote the ML estimator $\hat{\Delta}_4$.

An approximation to the variances of the estimators of $\Delta$, conditional on the numbers at risk in $\pi_1$ and $\pi_2$ is given by (2.2.2), i.e.,

$$\text{Var}[\hat{\Delta}_4] \approx \frac{\xi_1 \eta_1}{[\xi_1 + \eta_1]^4} \sum_{i=1}^{k} \alpha_i^2 r_{1i} r_{2i},$$

(5.1.1)
\[ \xi_i = \sum_{i=1}^{k} a_{i1}^t r_{2i} p_i, \quad \eta_i = \sum_{i=1}^{k} a_{i1}^t r_{1i}(1 - p_i), \quad \text{and} \quad \{a_{i1}^t\} \text{ is the set of weights corresponding to } \tilde{\lambda}_i, \ i = 1, \ldots, 4. \]

We now consider an approximation to the variance of the estimator of \( \Delta, \tilde{\Delta}' \) say, which allows for ties. Recall from Section 2.4 we derived the estimator

\[
\tilde{\Delta}' = \frac{\sum a_{i1}^t d_{1i}(r_{2i} - d_{2i})}{\sum a_{i1}^t d_{1i}(r_{2i} - d_{2i}) + \sum a_{i1}^t d_{2i}(r_{1i} - d_{1i})}, \tag{5.1.2}
\]

where \( d_{gi} \) is the number of deaths from \( \pi_g \) occurring at \( t_i \), \( g = 1, 2, i = 1, \ldots, k, \{a_{i1}^t\} \) is a set of arbitrary weights, and \( \sum_{i=1}^{k} = \sum_{i=1}^{k} \).

Conditional on \( r_1 \) and \( r_2 \), the random variables are \( d_{1i} \) and \( d_{2i} \). The distribution of \( d_{1i} \) (or \( d_{2i} \)), conditional on \( d_i = d_{1i} + d_{2i} \), the total number of deaths observed at \( t_i \), is difficult to determine.

A very rough approximation to this distribution is to assume \( d_{2i} \) is binomially distributed, i.e., \( d_{2i} \sim \text{b}[d_i, q_i] \), where \( q_i = \frac{r_{2i}}{r_{1i}^2 + r_{2i}^2} \), \( i = 1, \ldots, k. \)

Let

\[
X = \sum a_{i1}^t d_{1i}(r_{2i} - d_{2i}) = \sum a_{i1}^t (d_i - d_{2i})(r_{2i} - d_{2i}). \tag{5.1.3}
\]

\[
Y = \sum a_{i1}^t d_{2i}(r_{1i} - d_{1i}) = \sum a_{i1}^t d_{2i}(r_{1i} - d_i + d_{2i}). \tag{5.1.4}
\]

\[
Z = X + Y = \sum a_{i1}^t[d_i r_{2i} + (r_{1i} - r_{2i} - 2d_i)d_{2i} + 2d_{2i}^2]. \tag{5.1.5}
\]

It can be shown (see Appendix C), using the assumption \( d_{2i} \) is binomially distributed, that

\[
E[X] = \sum a_{i1}^t[d_i r_{2i} + (1 - r_{2i} - d_i)d_i q_i + d_i(d_i - 1)q_i^2]. \tag{5.1.6}
\]
and
\[
E[Y] \doteq \sum_{i=1}^{a_i} \left( (1 + r_{1i} - d_i) d_i q_i + d_i (d_i - 1) q_i^2 \right). \tag{5.1.7}
\]

Also
\[
\text{Var}[X] \doteq \sum_{i=1}^{a_i} \left\{ [d_i - (r_{2i} + d_i)]^2 d_i q_i \\
+ 6d_i - 7 + (r_{2i} + d_i) (6 - 5d_i - r_{2i}) d_i q_i^2 \\
+ 4(2d_i + r_{2i} - 3)d_i (d_i - 1) q_i^3 + 2(3 - 2d_i) d_i (d_i - 1) q_i^4 \right\},
\tag{5.1.8}
\]
\[
\text{Var}[Y] \doteq \sum_{i=1}^{a_i} \left\{ [d_i - (r_{1i} - d_i)]^2 d_i q_i \\
+ 6d_i - 7 + (r_{1i} - d_i) (5d_i - 6 - r_{1i}) d_i q_i^2 \\
+ 4(2d_i - r_{1i} - 3)d_i (d_i - 1) q_i^3 + 2(3 - 2d_i) d_i (d_i - 1) q_i^4 \right\},
\tag{5.1.9}
\]

and
\[
\text{Cov}[X,Y] \doteq \sum_{i=1}^{a_i} \left\{ [d_i - (r_{2i} + d_i)] [1 - (r_{2i} + d_i)] d_i q_i \\
+ \left\{ 6d_i - 7 + (3 - 2d_i) [(r_{2i} + d_i) - (r_{1i} - d_i)] \right\} d_i q_i^2 \\
+ 2[2(d_i - 3) + (r_{2i} - d_i) - (r_{1i} - d_i)] d_i (d_i - 1) q_i^3 \\
+ 2(3 - 2d_i) d_i (d_i - 1) q_i^4 \right\}. \tag{5.1.10}
\]

An approximation to the variance of $\Delta'$ is
\[
\text{Var}[\Delta'] \doteq \left\{ \frac{E[X]}{E[X+Y]} \right\}^2 \left\{ \frac{\text{Var}[X]}{(E[X])^2} + \frac{\text{Var}[X+Y]}{(E[X+Y])^2} - \frac{2 \text{Cov}[X,X+Y]}{E[X]E[X+Y]} \right\}.
\tag{5.1.11}
\]
Substituting (5.1.6)-(5.1.10) into (5.1.11), where

\[ E[X + Y] = E[X] + E[Y], \]

\[ \text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2 \text{Cov}[X, Y], \]

and

\[ \text{Cov}[X, X + Y] = \text{Var}[X] + \text{Cov}[X, Y], \]

gives an expression for the approximate variance of \( \tilde{\Delta}' \). Since the true \( \Delta \) (or \( \theta \)) is unknown, we take a further approximation and substitute the estimate of \( \Delta \) in (5.1.11). We will take the weights \( \{a_i'\} \) to be \( \{a_{k1}'\} \), the weights corresponding to the estimators of \( \Delta \) with single times of death, \( k = 1, \ldots, 4 \).

It should be noted that the approximation given by (5.1.1) for \( \text{Var}[\tilde{\Delta}_4] \) is equal to the asymptotic estimators given by the expected value of the negative reciprocal of the second derivatives of the log-likelihood function (see (2.2.11)), under this assumption about \( d_{2i}' \).

When we have censored observations, we make the assumption that the effective numbers at risk of dying at each of the times of death are

\[ r_{gi}' = r_{gi} - \frac{i}{c_{gi}}, \quad g = 1, 2, \quad i = 1, \ldots, k, \quad (5.1.12) \]

where \( c_{gi} \) is the number of individuals from \( \pi_g \) censored between \( t_i \) and \( t_{i+1} \). Approximations to the estimators of \( \Delta \) will be obtained using \( r_{gi}' \) for \( r_{gi} \) in the respective formulas, including the approximations to the variances of these estimators.
5.2. **Example of Mortality Data Among Hourly Workers in the Rubber Industry.**

The data are from a retrospective study of hourly workers in the rubber industry collected by the Occupational Health Studies Group (OHSG) of the University of North Carolina. A subset of the data file, provided for the author by the OHSG, includes information on length of follow-up, age and number of years of previous work experience in the rubber industry at time of entry into the study, and survival status at the end of the study, i.e., alive (including lost to follow-up and censored) or dead. The subset contained 3907 white male employees between the ages 30 and 55 at entry into the study, from a major tire manufacturing company in Akron, Ohio. The study period was ten years between January 1, 1964 and December 31, 1973. The data are incomplete in a few cases and ties may exist. We will not distinguish between censored observations resulting from termination of the study and voluntary withdrawal, i.e., loss of an individual from the study.

It is of interest to ascertain whether exposure to the working conditions in a tire manufacturing plant has a deleterious influence on survival of the employees. Age is, of course, an important factor in mortality. We conjecture that exposure (acquired during work experience) is also an important factor.

A description of one approach to comparing the survivorship functions of men with different amounts of exposure, controlling for the effect of age, is now given. Define six populations as follows.

\[ \pi_1: \text{men aged 35-40 with 5-15 years of exposure.} \]

\[ \pi_2: \text{men aged 35-40 with 15-25 years of exposure.} \]
\( \pi_3 \): men aged 40-45 with 10-20 years of exposure.

\( \pi_4 \): men aged 40-45 with 20-30 years of exposure

\( \pi_5 \): men aged 45-50 with 15-25 years of exposure.

\( \pi_6 \): men aged 45-50 with 25-35 years of exposure.

One way of making this comparison is to compare the survivorship functions of \( \pi_1 \) with \( \pi_2 \), \( \pi_3 \) with \( \pi_4 \), and \( \pi_5 \) with \( \pi_6 \). We can do this using any of the estimators of \( \Delta \) derived under the PHF assumption.

A second way emerges if we think of \( \pi_3 \) as being like \( \pi_1 \) after five further years have elapsed and \( \pi_5 \) as being like \( \pi_1 \) after ten years have elapsed, i.e., the cohort of men 35-40 years of age with 5-15 years of work experience will be age 40-45 with 10-20 years of experience five years later and will be age 45-50 with 15-25 years experience ten years later. We may think of \( \pi_4 \) and \( \pi_6 \) with regard to \( \pi_2 \) in a similar way.

The estimators of \( \Delta \) obtained in the three comparisons above provide, in a sense, information on twenty years of survivorship for the two populations of men 35-40 years of age with 5-15 and 15-25 years of exposure. This approach assumes that hazard rates observed for the different age groups between 1964 and 1973 will be those in the next ten year period. It also implies all other factors relating to calendar time, e.g., environmental and health conditions within the plant and associated community, are the same. These assumptions are not likely to be true, but this may have little effect on the comparisons suggested.
A summary of the data for each of the six populations described above is given in Table 5.1; Table 5.2 gives a summary of the estimates of $\Delta$ for the three proposed comparisons of populations.

**TABLE 5.1**

SUMMARY OF DATA ON SIX POPULATIONS OF WORKERS IN A MAJOR TIRE MANUFACTURING COMPANY, AKRON, OHIO, 1964-1973

<table>
<thead>
<tr>
<th>Population</th>
<th>Number Alive at $t_0$</th>
<th>Number of Deaths During Follow-up</th>
<th>Number of Persons Dying During Follow-up</th>
<th>Number of Persons Censored Before Close of Study</th>
<th>Proportion of Pop. at $t_0$</th>
<th>Proportion of Pop. at $t_0$ Censored Before Close of Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_1$</td>
<td>295</td>
<td>9</td>
<td>.0305</td>
<td>67</td>
<td>.2271</td>
<td></td>
</tr>
<tr>
<td>$\pi_2$</td>
<td>541</td>
<td>14</td>
<td>.0259</td>
<td>85</td>
<td>.1571</td>
<td></td>
</tr>
<tr>
<td>$\pi_3$</td>
<td>458</td>
<td>21</td>
<td>.0459</td>
<td>58</td>
<td>.1266</td>
<td></td>
</tr>
<tr>
<td>$\pi_4$</td>
<td>458</td>
<td>22</td>
<td>.0480</td>
<td>50</td>
<td>.1091</td>
<td></td>
</tr>
<tr>
<td>$\pi_5$</td>
<td>681</td>
<td>50</td>
<td>.0734</td>
<td>75</td>
<td>.1101</td>
<td></td>
</tr>
<tr>
<td>$\pi_6$</td>
<td>76</td>
<td>10</td>
<td>.1316</td>
<td>8</td>
<td>.1053</td>
<td></td>
</tr>
</tbody>
</table>
TABLE 5.2
ESTIMATES OF $\Delta$ FOR THREE COMPARISONS OF POPULATIONS OF WORKERS IN A
MAJOR TIRE MANUFACTURING COMPANY, AKRON, OHIO, 1964-1973

<table>
<thead>
<tr>
<th>Comparison</th>
<th>$\hat{\Delta}_1$</th>
<th>$\hat{\Delta}_2$</th>
<th>$\hat{\Delta}_3$</th>
<th>$\hat{\Delta}_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Populations</td>
<td>SE($\hat{\Delta}_1$)</td>
<td>SE($\hat{\Delta}_2$)</td>
<td>SE($\hat{\Delta}_3$)</td>
<td>SE($\hat{\Delta}_4$)</td>
</tr>
<tr>
<td>$C_1: \pi_1$ vs. $\pi_2$</td>
<td>.5509</td>
<td>.5481</td>
<td>.5482</td>
<td>.5481</td>
</tr>
<tr>
<td></td>
<td>.1065</td>
<td>.1059</td>
<td>.1059</td>
<td>.1059</td>
</tr>
<tr>
<td>$C_2: \pi_3$ vs. $\pi_4$</td>
<td>.5000</td>
<td>.4905</td>
<td>.4905</td>
<td>.4905</td>
</tr>
<tr>
<td></td>
<td>.0766</td>
<td>.0763</td>
<td>.0763</td>
<td>.0763</td>
</tr>
<tr>
<td>$C_3: \pi_5$ vs. $\pi_6$</td>
<td>.3504</td>
<td>.3490</td>
<td>.3491</td>
<td>.3491</td>
</tr>
<tr>
<td></td>
<td>.0794</td>
<td>.0787</td>
<td>.0788</td>
<td>.0788</td>
</tr>
</tbody>
</table>

We notice that there are no practical differences among the three estimators $\hat{\Delta}_2$, $\hat{\Delta}_3$ and $\hat{\Delta}_4$ in terms of their estimate of $\Delta$ and standard error of the estimate. Tests of the hypothesis $H_0: \Delta = .5$, which implies there is no difference in the mortality of the two populations, were performed assuming normality of $\hat{\Delta}_L$, $L = 1, \ldots, 4$, with an approximate estimate of the standard error obtained from (5.1.11). None of the tests for the three comparisons was significant.

We suggest that the conclusions of no differences among the mortality of the populations be accepted with caution for the following reasons. The estimates of $\Delta$ for the third comparison seem to differ from the estimates of $\Delta$ for the first two comparisons. We do not know whether this is a real difference which our test procedure failed to detect or whether it reflects the influence of (i) unequal
sample sizes at \( t_0 \), (ii) our method of adjusting for censored observations, particularly at the last time of death, or (iii) the elimination of men who retired early, most of whom would have been in \( \pi_5 \) and \( \pi_6 \).

We are uncertain what influence censoring may have on our estimation. In this example most of the censoring occurred at the close of the study, i.e., at the same fixed point, and we adjusted the risk sets at the last time of death to account for this. In an estimation not presented, we omitted the last time of death from our calculations. There is some indication that this affects the estimates of \( \Delta \) and the standard error of the estimates only in a small way. This point needs more investigation, however, before we can make any definitive statements regarding the influence of censoring on our estimation procedure.

We conclude by suggesting that these results be taken as provisional and proceed with additional methods of analysis. In particular, we may take the results and use them to motivate a regression type analysis similar to that discussed by Cox (1972), or possibly some parametric model which allows for covariables.
CHAPTER VI

SUGGESTIONS FOR FURTHER RESEARCH

We have indicated areas of further research in the comparison of survivorship functions at various points throughout this dissertation. We briefly described (Section 4.3) some methods for detecting departures from the PHF model in Chapter IV. These were regression techniques, and, in particular, methods of structural regression. We pointed out that the typical weighted least squares approach to estimating parameters in an alternate linear additive model has two main limitations. One is the choice of weights since an approximation to the variances of \( \mu_2(t) \) over appropriate intervals depends on the parameters in the model. In general, moderate errors in weights have relatively little influence and fair approximations can be used with some confidence. When the weights depend on the parameters to be measured, natural iterative methods should be considered. A second limitation is that the regressor variables, in addition to the response variables, are random variables. Some of the techniques developed in structural regression could possibly be adapted for this problem. However, application of structural regression requires introduction of reasonable assumptions on the relative magnitude of errors in the variables concerned. This needs detailed study.

The construction and analysis of models alternate to the PHF model appears to be an area offering considerable possibilities of
research. Suitable models should become apparent after sufficiently prolonged study along the lines described in the previous paragraph. Possibly preceding this there may be scope for detailed theoretical analysis of certain simple models similar to the models described in Section 4.3.1.

Indices of the kind described in Sections 4.1 and 4.2.5 (representing variation in the HRF or some function of the HRF) might prove useful in deciding on suitable intervals of time within which a PHF model might be regarded as a sufficiently good approximation to use. One might, for example, have a requirement that for this to be so, the "index" for any such interval should not exceed a specified value. Determination of such a critical value would of course call for a careful assessment of the consequences of its use.

We have suggested a very rough way of allowing for censored data (Section 2.4.1). Some of the calculations in Chapter V indicate that assumptions on the effect of censoring may, in some circumstances, have a noticeable effect on the results obtained. A thorough investigation of these effects seems likely to produce interesting contributions. Of course, even with such studies, it is by no means certain, or even likely, that a universally "best" (or even good) way of allowing for censoring will exist. When we really do have full information on how censoring operates in a particular case we can try to make as appropriate an allowance for it as we can. Very often, however, we can do no more than make a more or less enlightened guess.
BIBLIOGRAPHY


(1974a), Covariance analysis of censored survival data, Biometrics, 30, 89-100.

(1974b), Analysis of Survival Data Under the Proportional Hazards Model, paper presented at the Session on Statistical Aspects of Chronic Disease, Eighth International Biometric Conference, Constanta, Romania, August.


_____ and Cislak, P. J. (1968), On a general system of distributions, I. Its curve-shape characteristics, II. The sample median, Journal of the American Statistical Association, 63, 627-635.


Crowley, J. (1974), A note on some recent likelihoods leading to the log rank test, Biometrika, 61, 533-538.


Durbin, J. (1975), Kolmogorov-Smirnov tests when parameters are estimated with applications to tests of exponentiality and tests on spacings, Biometrika, 62, 5-22.


Freireich, E. J. et al. (1963), The affect of 6-mercaptopurine on the duration of steroid-induced remissions in acute leukemia, Blood, 21, 699-716.


(1965b), A generalized two-sample Wilcoxon test for doubly censored data, Biometrika, 52, 650-653.


Mann, H. B. and Whitney, D. R. (1947), On a test of whether one of two random variables is stochastically larger than the other, Annals of Mathematical Statistics, 18, 50-60.

Mantel, N. (1966a), Models for complex contingency tables and polychotomous dosage response curves, Biometrics, 22, 83-95


McRae, K. B. and Thomas, D. R. (1972), Inference Procedures for Pairs of Distributions with Proportional Failure Rate Functions, paper presented at the Joint Statistical Meetings, Montreal, Canada.


APPENDIX A

MOMENTS ABOUT ZERO OF SOME LIFETIME DISTRIBUTIONS

Let $T$ be a positive continuous random variable with cumulative distribution function $F_T(t)$ such that the probability density function $f_T(t) = \frac{d}{dt}F_T(t)$ exists. It is known that the mean of $T$ is the integral of the survival function, i.e.

$$E[T] = \int_0^\infty [1 - F_T(t)]dt = \int_0^\infty F_T(t)dt,$$  \hspace{1cm} (A.1)

where $P_T(t) = 1 - F_T(t)$ is the survival function.

To show (A.1) is true, by definition we have

$$E[T] = \int_0^\infty t \frac{d}{dt}[1 - F_T(t)]dt.$$

(A.2)

Integration by parts of (A.2) leads to (A.1) if $\lim_{t \to \infty} t[1 - F_T(t)] = 0$.

It can be shown that the $\gamma$th moment about zero can be written as

$$E[T^\gamma] = u'_\gamma = \gamma \int_0^\infty t^{\gamma - 1}[1 - F_T(t)]dt,$$

(A.3)

provided $t^\gamma[1 - F_T(t)] \to 0$ as $t \to \infty$.

A1. **Makeham Gompertz Distribution.**

For a derivation of the first two moments about zero of the Gompertz distribution ($A = 0$) see Hoel (1972); the mean of the Gompertz distribution has been given by Broadbent (1958) and by Flehinger and Lewis (1959). The moments of the Makeham Gompertz distribution are not readily found in the statistical literature and are therefore presented
Using (A.1) the mean is

$$E[T] = u'_1 = \int_0^\infty \{ \exp[-(B/k)(e^{\kappa t} - 1) - At] \} dt \quad (A1.1)$$

To evaluate (A.1) consider the transformation $y = e^{\kappa t}$ which gives

$$u'_1 = \frac{e^{B/k}}{k} \int_1^\infty e^{-(B/k)y} y^{-(1+A/k)} dy. \quad (A1.2)$$

The exponential integral (A.2) may be written in terms of an incomplete gamma function, giving

$$u'_1 = \frac{e^{B/k}}{k} \left( \frac{B}{k} \right) \frac{A/k}{A/k, B/k, \kappa} \quad (A1.3)$$

where $\Gamma(\alpha, x) = \int_x^\infty e^{-t} t^{\alpha-1} dt$ is the form of the incomplete gamma function given by Abramowitz and Stegun (1972) (p. 260, [6.5.3]).

For higher moments we can show that (A.3) is equivalent to

$$E[T^\gamma] = u' = \gamma(-1)\Gamma^{-1} \frac{d\gamma^{-1}u'_1}{d\gamma^{-1}}. \quad (A1.4)$$

The variance is then

$$\text{Var}[T] = u'_2 - [u'_1]^2,$$

where

$$u'_2 = \frac{2e^{B/k}}{k^2} \int_1^\infty e^{-y} y^{-(1+A/k)} dy. \quad (A1.5)$$

Evaluation of both (A.2) and (A.5) requires numerical methods.

A2. Linear Hazard Function Distribution.

For a derivation of the first two moments of the linear hazard function distribution see Kodlin (1967). The expressions given below are equivalent to those given by Kodlin, but are somewhat simpler in
form and may be evaluated using techniques for the standard normal integral.

Using (A.1) the mean is

\[
E[T] = u'_1 = \int_0^\infty \{\exp[-\alpha t - \frac{1}{2} \beta t^2]\} dt.
\]  
(A2.1)

To evaluate the integral in (A2.1), complete the square of the exponential factor and let \( y = \tau \sqrt{\beta} + \alpha / \sqrt{\beta} \). We find

\[
u'_1 = e^{\alpha^2/2\beta} \left( \frac{1}{\sqrt{\beta}} \right) \int_0^\infty e^{-y^2/2} dt
\]
(A2.2)

\[
u'_1 = e^{\alpha^2/2\beta} \left( \frac{2\pi}{\sqrt{\beta}} \right) \left\{ 1 - \phi \left( \frac{\alpha}{\sqrt{\beta}} \right) \right\},
\]
(A2.3)

where \( \phi(z) \) is the standard normal probability integral.

For higher moments we can show that (A.3) is equivalent to

\[
E[T^\gamma] = u'_\gamma = \gamma(-1)^{\gamma-1} \frac{d^{\gamma-1} u'_1}{d\alpha^{\gamma-1}}.
\]  
(A2.4)

The second moment about zero is

\[
u'_2 = -\frac{2\alpha}{\beta} e^{\alpha^2/2\beta} \left( \frac{2\pi}{\sqrt{\beta}} \right) \left\{ 1 - \phi \left( \frac{\alpha}{\sqrt{\beta}} \right) \right\} + 2 e^{\alpha^2/2\beta} \left( \frac{2\pi}{\sqrt{\beta}} \right) \phi \left( \frac{\alpha}{\sqrt{\beta}} \right) \frac{1}{\sqrt{\beta}}
\]
(A2.5)

The variance is

\[
\text{Var}[T] = \frac{2}{\beta} \{1 - \alpha u'_1\} - \{u'_1\}^2.
\]  
(A2.6)


The moments of the truncated logistic distribution do not seem to be published.

Consider the standard logistic density function truncated at \( y = K \),
\[ p(y) = e^{-y} [1 + e^{-y}]^{-2} / [1 - F(K)], \]  
\[ (A3.1) \]

where \( F(K) = [1 + e^{-K}]^{-1} \). The moments about zero are given by

\[ E[Y^\gamma] = C^{-1} \int K y e^{-y} [1 + e^{-y}]^{-2} dy, \]  
\[ (A3.2) \]

where \( C = [1 - (1 + e^{-K})^{-1}] \).

Using the relationship given by

\[ e^{-y} [1 + e^{-y}]^{-2} = \sum_{j=1}^{\infty} (-1)^{j-1} j e^{-jy} \]  
\[ (A3.3) \]

(see Johnson and Kotz, Chapter 22, p. 4), we can write (A3.2), after interchanging the integral and summation, as

\[ C \cdot E[T^\gamma] = \sum_{j=1}^{\infty} (-1)^{j-1} j \int K y e^{-jy} dy. \]  
\[ (A3.4) \]

If \( \gamma \) is an integer we can evaluate (A3.4) by putting \( z = jy \) and then integrating by parts, leading to

\[ u'_\gamma = C^{-1} \left\{ \gamma! \sum_{j=1}^{\infty} (-1)^{j-1} j^{-\gamma} S_{\gamma}(K_j) \right\}, \]  
\[ (A3.5) \]

where

\[ S_{\gamma}(K_j) = e^{-K_j} \left\{ 1 + \frac{(K_j)}{1!} + \frac{(K_j)^2}{2!} + \ldots + \frac{(K_j)^\gamma}{\gamma!} \right\}. \]  
\[ (A3.6) \]

For \( (K_j) > 0 \) we can evaluate \( S_{\gamma}(K_j) \) using the table of Poisson probabilities.

If we want the distribution of \( Y = \frac{T - \xi}{\beta} \), or \( T = \xi + \beta Y \) such that \( T \) is truncated at zero, then \( Y \) is truncated at \( K = -\xi/\beta \). Thus, unless \( \xi < 0 \) we cannot evaluate \( S_{\gamma}(K_j) \) using tables of Poisson probabilities.
APPENDIX B

BEHAVIOR OF THE HRF FOR PAIRS OF DISTRIBUTIONS
FROM THE SAME PARAMETRIC MODEL


The HRF, using the notation in Table 3.4, is

$$\theta(t) = \frac{A_1 + B_1 e^{\kappa_1 t}}{A_2 + B_2 e^{\kappa_2 t}}. \quad (B1.1)$$

Let $y = e^t$. Then

$$\frac{d\theta(y)}{dy}$$

$$= \left[A_2 + B_2 y^{\kappa_2^{-1}}\right]^{-2}\left\{[A_2 + B_2 y^{\kappa_2^{-1}}][\kappa_1 B_1 y^{\kappa_1^{-1}}] - [A_1 + B_1 y^{\kappa_1^{-1}}][\kappa_2 B_2 y^{\kappa_2^{-1}}]\right\}$$

$$= \frac{\kappa_1^{-1}}{y^{\kappa_1^{-1}}} \left[A_2 B_1^{\kappa_1^{-1}} - [A_1 B_2^{\kappa_2^{-1}} + (\kappa_2 - \kappa_1) B_1 B_2 y^{\kappa_2^{-1}}]\right]. \quad (B1.2)$$

Assume $0 < \kappa_1 < \kappa_2$ and consider the behavior of the function $G(y) = \frac{d\theta(y)}{dy}$. Now, $G(y)$ is a decreasing function and has at most one root. There is a root if

$$G(1) = \{A_2 B_1^{\kappa_1} - [A_1 B_2^{\kappa_2} + (\kappa_2 - \kappa_1) B_1 B_2]\} > 0,$$

since $\lim_{y \to \infty} G(y) = -\infty$. If $G(y)$ has a root, then we know there is some value of $y$ (or $t$), $\eta$ say, such that $G(y) > 0$ for $y > \eta$ and
G(y) < 0 for y < \eta. Thus, if

\[
\left( A_2 B_1 \kappa_1 - \left[ A_1 B_1 \kappa_2 + (\kappa_2 - \kappa_1) B_1 B_2 \right] \right) > 0 \quad (B1.3)
\]

for 0 < \kappa_1 < \kappa_2 we conclude that \( \theta(t) \) has a mode.

**B2. Truncated Logistic.**

The HRF, using the notation of Table 3.4, is

\[
\theta(t) = \frac{\beta_2 \left[ 1 + \exp \left( -\frac{t - \xi_2}{\beta_2} \right) \right]}{\beta \left[ 1 + \exp \left( -\frac{t - \xi_1}{\beta_1} \right) \right]}, \quad \beta_1, \beta_2 > 0. \quad (B2.1)
\]

For convenience, we will consider \( \log \theta(t) \). We have

\[
\frac{d \log \theta(t)}{dt} = -\frac{\exp \left( -\frac{t - \xi_2}{\beta_2} \right)}{\beta_2 \left[ 1 + \exp \left( -\frac{t - \xi_2}{\beta_2} \right) \right]} + \frac{\exp \left( -\frac{t - \xi_1}{\beta_1} \right)}{\beta_1 \left[ 1 + \exp \left( -\frac{t - \xi_1}{\beta_1} \right) \right]}. \quad (B2.2)
\]

We see that \( \frac{d \log \theta(t)}{dt} > 0 \) (\(< 0\)) if \( G(t) < \frac{\beta_1}{\beta_2} \) (\( G(t) > \frac{\beta_1}{\beta_2} \)) where

\[
G(t) = \frac{\left[ 1 + \exp \left( \frac{t - \xi_2}{\beta_2} \right) \right]}{\left[ 1 + \exp \left( \frac{t - \xi_1}{\beta_1} \right) \right]}. \quad (B2.3)
\]

To simplify, let \( \alpha_1 = e^{-\xi_1/\beta_1}, \alpha_2 = e^{-\xi_2/\beta_2}, \kappa = \frac{\beta_1}{\beta_2} \) and \( y = e^{t/\beta_1} \).

Then \( H(y) = \frac{1 + \alpha_2 y^\kappa}{1 + \alpha_1 y} \).

Choose \( \beta_1 > \beta_2 > 0 \) and consider the behavior of \( H(y) \). We have

\[
h(y) = \frac{dH(y)}{dy} = \left[ 1 + \alpha_1 y \right]^{-2} \{-\alpha_1 + (\kappa - 1) \alpha_1 \alpha_2 y^\kappa + \alpha_2 ky^{\kappa - 1} \}. \quad (B2.4)
\]
Now, \( h(y) \) clearly has a root, since for \( \kappa > 1 \), \( h(y) \) is an increasing function of \( y \), \( h(0) = -\alpha_1 \) and \( \lim_{y \to \infty} h(y) = \infty \). That is, \( h(y) < 0 \) for some \( y < \eta \), say, and \( h(y) > 0 \) for \( y > \eta \). Thus \( H(y) \) has a mode, i.e., \( H(y) < \frac{\beta_1}{\beta_2} \) for \( y < \eta \) and \( H(y) > \frac{\beta_1}{\beta_2} \) for \( y > \eta \). Hence, \[
\frac{d \log \theta(t)}{dt} > 0 \quad (< 0) \quad \text{for some} \quad t < \eta^* = \frac{1}{\beta_1} \log \eta \quad (t > \eta^*). \] And in general, for \( \beta_1 \neq \beta_2 \) we see that \( \theta(t) \) always is unimodal.

For \( \beta_1 = \beta_2 \), however, we can show that \( \theta(t) \) is always monotonic. If \( \beta_1 = \beta_2 \), or \( \kappa = 1 \), we see that \( H(y) \) is a monotonic increasing (decreasing) function if \( \alpha_1 < \alpha_2 \) (\( \alpha_1 > \alpha_2 \)), or that \[
\frac{d \log \theta(t)}{dt} > 0 \quad (< 0) \quad \text{if} \quad \alpha_1 < \alpha_2 \quad (\alpha_1 > \alpha_2). \]

Thus \( \theta(t) \) is a monotonic increasing (decreasing) function for \( \xi_1 > \xi_2 \), \( \xi_1 < \xi_2 \).

**B3. Burr.**

The HRF, using the notation of Table 3.4, is

\[
\theta(t) = \frac{\alpha_1 \beta_1 [1 + t^{-\alpha_2}]}{\alpha_2 \beta_2 [1 + t^{-\alpha_1}]} \quad g, \beta_g > 0, \ g = 1, 2. \tag{B3.1}
\]

Then

\[
\frac{d \theta(t)}{dt} = \frac{- (\alpha_1 + \alpha_2 + 1)}{\alpha_2 \beta_2 [1 + t^{-\alpha_1}]^2} \left( \alpha_1 t^{\alpha_2} - \alpha_2 t^{\alpha_1} + (\alpha_1 - \alpha_2) \right). \tag{B3.2}
\]

Let \( h(t) = \alpha_1 t^{\alpha_2} - \alpha_2 t^{\alpha_1} + (\alpha_1 - \alpha_2) \). Then \( \frac{d \theta(t)}{dt} > 0 \quad (< 0) \) as \( h(t) > 0 \quad (< 0) \). We see that \( h(0) = \alpha_1 - \alpha_2 \), \( h(t) = 2(\alpha_1 - \alpha_2) \), and \( \lim_{t \to \infty} h(t) = \infty \quad (\rightarrow \infty) \) if \( \alpha_1 < \alpha_2 \) (\( \alpha_1 > \alpha_2 \)). Suppose \( \alpha_1 > \alpha_2 \); we also
observe that \( \frac{dh(t)}{dt} = \alpha_1 \alpha_2 [t^{\alpha_2 - 1} - t^{\alpha_1 - 1}] \) is positive or negative as \( t > 1 \) or \( t < 1 \). So there is some \( t, \eta \) say, such that \( \frac{d\theta(t)}{dt} > 0 \) \( (\leq 0) \) as \( t < \eta \) \( (t > \eta) \) for \( \alpha_1 > \alpha_2 \). Thus \( \theta(t) \) is unimodal. The argument is similar for \( \alpha_1 < \alpha_2 \).
APPENDIX C

APPROXIMATE VARIANCES OF THE ESTIMATORS OF Δ WITH TIES AND CENSORING

Consider the estimators of Δ given in Section 2.4 which includes tied observations, i.e.,

\[ \hat{\Delta}' = \frac{\sum_{i=1}^{k} \alpha_{i}d_{i1}(r_{2i} - d_{2i})}{\sum_{i=1}^{k} \alpha_{i}d_{i1}(r_{2i} - d_{2i}) + \sum_{i=1}^{k} \alpha_{i}d_{i2}(r_{1i} - d_{1i})}, \quad (C.1) \]

with \( \sum_{i=1}^{k} = \sum_{i=1}^{k} \). Assume \( d_{2i} \sim b[d_{1i}, q_{i}] \), where \( d_{i} = d_{1i} + d_{2i} \),

\[ q_{i} = \frac{r_{2i}}{r_{1i} + r_{2i}} \]

and the \( d_{2i} \)'s are independent, \( i = 1, \ldots, k \).

Let

\[ X = \sum_{i=1}^{k} \alpha_{i}d_{i1}(r_{2i} - d_{2i}) = \sum_{i=1}^{k} \alpha_{i}(d_{1i} - d_{2i})(r_{2i} - d_{2i}), \quad (C.2) \]

and

\[ Y = \sum_{i=1}^{k} \alpha_{i}d_{i2}(r_{1i} - d_{1i}) = \sum_{i=1}^{k} \alpha_{i}d_{i2}(r_{1i} - d_{1i} + d_{2i}). \quad (C.3) \]

An approximation to the variance of \( \hat{\Delta}' \) is

\[ \text{Var}[\hat{\Delta}'] = \left[ \frac{E[X]}{E[X + Y]} \right]^{2} \left[ \frac{\text{Var}[X]}{(E[X])^{2}} + \frac{\text{Var}[X + Y]}{(E[X] + Y)^{2}} - \frac{2 \text{Cov}[X, X + Y]}{E[X]E[X + Y]} \right], \quad (C.4) \]

where

\[ E[X + Y] = E[X] + E[Y], \]

\[ \text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2 \text{Cov}[X, Y], \]
and \[ \text{Cov}(X, X + Y) = \text{Var}(X) + \text{Cov}(X, Y). \]

Then

\[
\begin{align*}
E[X] &= \sum_{i} \alpha_i' \mathbb{E}[u_i], \\
E[Y] &= \sum_{i} \alpha_i' \mathbb{E}[v_i], \\
\text{Var}[X] &= \sum_{i} \alpha_i'^2 \text{Var}[u_i], \\
\text{Var}[Y] &= \sum_{i} \alpha_i'^2 \text{Var}[v_i],
\end{align*}
\]

and

\[ \text{Cov}[X, Y] = \sum_{i} \alpha_i'^2 \text{Cov}[u_i, v_i], \]

where \[ u_i = (d_1 - d_{2i})(r_{2i} - d_{2i}) = d_1 r_{2i} - d_{2i}(r_{2i} + d_1) + d_{2i}^2, \]
and

\[ v_i = d_{2i}(r_{1i} - d_i + d_{2i}) = d_{2i}(r_{1i} - d_i) + d_{2i}^2. \] (C.5)

We will drop the subscript \( i \) and proceed to find the first two moments of \( u \) and \( v \) and their covariance. The following relationships between moments about zero and factorial moments of the binomial distribution employed in the derivations are given below.

\[
\begin{align*}
E[d_2^2] &= E[d_2^{(2)}] + E[d_2], \\
E[d_2^3] &= E[d_2^{(3)}] + 3E[d_2^{(2)}] + E[d_2], \\
E[d_2^4] &= E[d_2^{(4)}] + 6E[d_2^{(3)}] + 7E[d_2^{(2)}] + E[d_2],
\end{align*}
\]

where \[ E[d_2^{(Y)}] = d(d - 1) \cdots (d - Y + 1)q^Y. \]

Thus we find

\[
E[u] = d r_2 - (r_2 + d)dq + d(d - 1)q^2 + dq,
= dr_2 + [1 - (r_2 + d)]dq + d(d - 1)q^2. \] (C.7)
\[ E[v] = (r_1 - d)dq + d(d - 1)q^2 + dq, \]
\[ = [1 + r_1 - d]dq + d(d - 1)q^2. \]  
(C.8)

\[ \text{Var}[u] = E[u'^2] - \{E[u']\}^2, \text{ where } u' = d_2^2 - d_2(r_2 + d). \]

We have

\[ E[u'^2] = E[d_2^4 - 2d_2^3(r_2 + d) + d_2^2(r_2 + d)^2] \]
\[ = d(d - 1)(d - 2)(d - 3)q^4 + 3[2 - (r_2 + d)]d(d - 1)(d - 2)q^3 \]
\[ + [7 - 6(r_2 + d) + (r_2 + d)^2]d(d - 1)q^2 + [1 - (r_2 + d)]^2dq, \]  
(C.9)

and

\[ \{E[u']\}^2 = d_2^2(d - 1)^2q^4 + 2d_2^2(d - 1)[1 - (r_2 + d)]q^3 + [1 - (r_2 + d)]^2dq^2, \]  
(C.10)

which gives

\[ \text{Var}[u] = [(d - 2)(d - 3) - d(d - 1)]d(d - 1)q^4 \]
\[ + [(d - 2)[6 - 2(r_2 + d)] - 2d[1 - (r_2 + d)]]d(d - 1)q^3 \]
\[ + [(d - 1)[7 - 6(r_2 + d) + (r_2 + d)^2] - d[1 - (r_2 + d)]^2]dq^2 \]
\[ + [1 - (r_2 + d)]^2dq. \]  
(C.11)

Combining like terms in (C.9) and (C.10), upon simplification we have

\[ \text{Var}[u] = 2(3 - 2d)d(d - 1)q^4 + 4(r_2 + 2d - 3)d(d - 1)q^3 \]
\[ + [6d - 7 + (r_2 + d)(6 - 5d - r_2)]dq^2 + [1 - (r_2 + d)]^2dq. \]  
(C.12)
We can obtain $\text{Var}[v]$ directly from (C.12) since we notice from (C.5) and (C.6) that $u'$ and $v$ are of the same form with $-r_{1i}$ replacing $r_{2i}$. Thus

$$\text{Var}[v] = 2(3 - 2d)d(d - 1)q^4 + 4(2d - r_1 - 3)d(d - 1)q^3$$

$$+ [6d - 7 + (r_1 - d)(5d - 6 - r_{1i})]dq^2 + [1 + r_1 - d]dq.$$  

(C.13)

For the covariance, $\text{Cov}[u,v] = \text{Cov}[u',v]$, we have

$$E[u',v] = E\left\{d^2_2 + [(r_1 - d) - (r_2 + d)]d_2^3 - (r_1 - d)(r_2 + d)d_2^2\right\}$$

$$= d(d - 1)(d - 2)(d - 3)q^4$$

$$+ [6 + (r_1 - d) - (r_2 + d)]d(d - 1)(d - 2)q^3$$

$$+ [7 + 3(r_1 - d) - 3(r_2 + d) - (r_1 - d)(r_2 + d)]d(d - 1)q^2$$

$$+ [1 + (r_1 - d) - (r_2 + d) - (r_1 - d)(r_2 + d)]dq,$$

(C.14)

and

$$E[u']E[v] = d^2(d - 1)^2q^4 + [(1 - (r_2 + d)) + [1 + r_1 - d)]d^2(d - 1)q^3$$

$$+ [1 - (r_2 + d) + (r_1 - d) - (r_2 + d)(r_1 - d)]d^2q^2.$$  \hspace{1cm} (C.15)

We then find, using (C.14) and (C.15), combining like terms, and simplifying that
\[ \text{Cov}[u, v] = 2(3 - 2d) d(d - 1)q^4 \]

\[ + 2[2(d - 3) + (r_2 + d) - (r_1 - d)] d(d - 1)q^3 \]

\[ + \{6d - 7 + (3 - 2d)[(r_2 + d) - (r_1 - d)]\} dq^2 \]

\[ + [1 - (r_2 + d)][1 + r_1 - d] dq. \quad \text{(C.16)} \]