SURVIVAL MODELS

AND THEIR ESTIMATION

Third Edition

Dick London, FSA

SURVIVAL MODELS AND THEIR ESTIMATION

Third Edition

Dick London, FSA

ACTEX Publications
Winsted, Connecticut

Request for permission should be addressed to:
ACTEX Publications
PO Box 715
New Hartford CT 06057
support@actexmadriver.com

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN: 978-1-63588-654-2

CONTENTS

PART I
THE NATURE AND PROPERTIES OF SURVIVAL MODELS

PART II
ESTIMATION OF SURVIVAL MODELS FROM SAMPLE DATA

PART III
APPLICATIONS AND EXTENSIONS

PREFACE TO THE SECOND EDITION

Survival Models and Their Estimation is a general textbook describing the properties and characteristics of survival models, and statistical procedures for estimating such models from sample data. Although it is written primarily for actuaries, it is also intended to be of interest to a broader mathematical and statistical audience. Academically, the text is aimed at the fourth year undergraduate or the first year graduate level.

Actuaries and other applied mathematicians work with models which predict the survival pattern of humans or other entities (animate or inanimate), and frequently use these models as the basis for calculations of considerable financial importance. Specifically, actuaries use such models to calculate the financial values associated with individual life insurance policies, pension plans, and income loss coverages. Demographers and other social scientists use survival models for making predictive statements about the future make-up of a population to which the model is deemed to apply.

This text is not primarily concerned with the *uses* of survival models, but rather with the question of how such models are established. This exercise is sometimes referred to as survival model *development* or survival model *construction*; in this text, however, we prefer the more descriptive phrase ***survival model estimation***.

It cannot be noted too strongly that the "real" survival distribution (or survival probabilities) which apply to a group of persons is unknown, and probably will forever be so. What we, therefore, attempt to do is *estimate* that distribution, based on the data of a sample and a chosen estimation procedure. It is vitally important that this be clearly understood. Since the name of the game is estimation, there are no "right" answers. There are only sound (or unsound) procedures.

Because the result of our exercise is an estimate of the theoretical, underlying, operative survival distribution, based on the particular experience of a sample, we recognize that the estimate is a realization of a random variable, called an ***estimator***. In turn, this random variable has properties such as expected value and variance, and these properties tell us something about the quality of the estimator. Note that we do not judge the "accuracy" of the resulting estimate, but rather the quality (or validity) of the procedure which produced the estimate. Properties of estimator random variables are defined in Appendix A. Readers who are not entirely familiar with these properties may wish to review Appendix A before studying the specific estimators developed in the text.

Frequently the estimated survival model produced directly from a study is not entirely suitable for practical use, and is, therefore, systematically revised before such use. The process of revising the initial estimates into revised estimates is called ***graduation***. This step in the development of a useable estimated survival model is the topic of a companion text to this one entitled *Graduation: The Revision of Estimates*.

Survival Models and Their Estimation is said to be a general text in that it treats survival model estimation from the viewpoint of several different practitioners, including the actuary, the demographer, and the biostatistician, without attempting to be an exhaustive treatment of any one of these traditions.

A more thorough treatment of the actuarial tradition, from a different perspective, can be found in texts by Gershenson [32], Benjamin and Pollard [11], and Batten [8]; demographic approaches are the main theme of the works by Keyfitz and Beekman [46], Spiegelman [71], and Chiang [19]; the medical, or biostatistical, tradition is more deeply pursued by Elandt-Johnson and Johnson [25]. Additional texts, which deal with the statistical analysis of survival data at the graduate level, include those by Lawless [50], Lee [51], Miller [56], and Kalbfleisch and Prentice [42].

How is an initial estimated survival model determined from sample data? There are many approaches to this. A survival model estimation problem will generally have three basic components: (1) the form and nature of the sample data (which might also be called the ***study design***); (2) the chosen ***estimation procedure***; and (3) any ***simplifying assumptions*** made along the way. All of these concepts will be further developed in this text. The traditional actuarial approach, for example, is characterized by a cross-sectional study design using the transactional data of an insurance company or pension fund operation, a method-of-moments estimation procedure, and the Balducci distribution assumption. In this text we will consider, as well, other study designs, primarily those encountered by the clinical statistician or the reliability engineer. In addition, we will consider other estimation procedures, especially the maximum likelihood and product-limit methods. Finally, we will consider other simplifying assumptions, such as the uniform and exponential distributions.

The text presumes a basic familiarity with probability and statistics, including the topics of estimation and hypothesis testing. The application of these ideas specifically to the estimation of survival models is then developed throughout the text. An effort has been made to keep the mathematics and the pedagogy at a level which does not require a prior familiarity with the topic. Whenever a choice between mathematical rigor and pedagogic effectiveness appeared to be necessary, we opted for the latter. As a result, the level of mathematical rigor in the text may be somewhat less than that desired by the precise mathematician, but the increased clarity which results from sacrificing some rigor will hopefully be welcomed by the student reader.

The first edition of this text, published in 1986, included the subject matter contained in the first eight chapters of the new edition. Chapters 5 and 6 have been completely rewritten from the first edition, Chapter 7 has been substantially revised, and three entirely new chapters have been added. The first edition was adopted by the Society of Actuaries as a reference for its examination program in 1987, and many valuable suggestions for improvement were contributed by students and educators.

Drafts of the material in both editions of the text were submitted to a review team, whose many valuable comments are reflected in the final version. The indispensable assistance of this group is hereby gratefully acknowledged.

Warren R. Luckner, FSA, of the Society of Actuaries, coordinated the efforts of the review team and made many valuable comments himself.

Stuart A. Klugman, FSA, Ph.D., of the University of Iowa, was particularly adept at detecting mathematical errors in the drafts, and much of the precision that the text has attained is due to his careful efforts.

Stanley Slater, ASA, of Metropolitan Life Insurance Company, did a remarkable job of editing the drafts for improvements in writing style and clarity, especially for the benefit of the student reader.

Other members of the review team who made valuable contributions to the final text include Robert L. Brown, FSA, and Frank G. Reynolds, FSA, both of the University of Waterloo, Cecil J. Nesbitt, FSA, Ph.D., of the University of Michigan, Geoffrey Crofts, FSA, of the University of Hartford, and Robert Hupf, FSA, of United of Omaha Life Insurance Company.

Much of the research and writing time invested in this project was supported by a grant from the Actuarial Education and Research Fund. The author would like to express his appreciation to the directors of AERF for this support.

Special thanks and appreciation are expressed to Marilyn J. Baleshiski of ACTEX Publications who did the electronic typesetting for the entire text, through what must have appeared to be an endless series of revisions.

Despite the efforts of the review team and the author to attain pedagogic clarity and mathematical accuracy, errors and imperfections are undoubtedly still present in the text. For this the author and the publisher take full responsibility and sincerely apologize to the reader. We respectfully request that you report these errors to the author at ACTEX Publications, P.O. Box 974, Winsted, CT 06098.

Winsted, Connecticut
June, 1988

Dick London, FSA

PREFACE TO THE THIRD EDITION

It has been nearly ten years since the publication of the second edition of *Survival Models and Their Estimation*, and its adoption by the Society of Actuaries as the principal references for its Course 160 examination. It appears that the text has proved satisfactory in that role from the point of view of exam candidate and examiner alike.

The mathematics of survival models themselves, and how they might be estimated from sample data, has been a fairly stable topic, so that a revision of the theory presented in the first eight chapters of the textbook does not seem to be required at this time. Consequently, the reader familiar with the second edition will note various clarifications and improvements in presentation in these chapters, but no substantive change in the overall content. Why, then, is a new edition appearing at this time?

Beginning in 1994, a Society of Actuaries Board Task Force on Education has been working toward a new model of actuarial education for the twenty-first century. Among many other important principles, the Task Force has established that actuarial education in the future should include guidance for the application of standard actuarial techniques in disciplines beyond the traditional actuarial areas of insurance and pensions. Since *Survival Models and Their Estimation* plays its small part in the actuarial education arena, as the reference text for Course 160, it naturally follows that a revision of the text, guided by the Task Force principle of broadened application, is now appropriate. The result is the appearance of new Chapters 10, 11, and 12, which present applications of the general theory of survival models in such fields as epidemiology, facilities planning, economics, investments, reliability engineering, and others.

Two other issues have affected certain changes from the prior to the current edition of this text as well.

The first is that Chapter 9 in the prior edition, which described the demographer's process of estimating survival models from general population data, has been deleted from the text. The material in the prior edition was based on out-of-date studies, namely the Canadian census of 1981 and the U.S. census of 1980, and is not included in the course of reading for the Course 160 exam. Furthermore, the content of that chapter is also included, in up-dated form, in the new (third) edition of Robert L. Brown's *Introduction to the Mathematics of Demography* [16], and interested readers are directed to that reference.

The second important change reflected in the new edition of this text is a recognition that the traditional theory and data processing mechanics for large-scale actuarial studies, developed in the pre-computer age, should no longer receive the emphasis that it has in the past. Accordingly, the description of the theory of the traditional actuarial approach and a critique of that theory, as presented in Section 6.4, has been appropriately reduced. In addition, Chapter 11 of the prior edition, which dealt with the now out-of-date practice of calculating actuarial exposue as a by-product of life insurance liability valuation, has been deleted. On the other hand, after considerable deliberation it was decided that the prior Chapter 10 should be retained as Chapter 9 in the new edition. Although dated in some respects, a description of the actual data processing mechanics involved in the actuarial technique of estimating survival models from insurance company and pension fund data, was deemed to still be an important component of actuarial education.

The author would like to acknowledge the contributions of several colleagues to the new edition of the text.

A review of Chapter 10 was provided by Bruce Leonard Jones, FSA, FCIA, Ph.D., of the University of Western Ontario. Frank G. Bensics, FSA, Ph.D., of the College of Insurance suggested much of the content of Chapter 11, and reviewed the drafts of that chapter as well. An important contribution to the development of Chapter 11 was also made by Matthew J. Hassett, ASA, Ph.D., of Arizona State University. A similar role for Chapter 12 was played by Rohan J. Dalpatadu, ASA, Ph.D., of the University of Nevada at Las Vegas. Overall guidance for the content of the new edition was provided by Robert A. Conover, FSA, the Education Actuary for Course 160 at the Society of Actuaries.

The text layout design and typesetting was again handled by Marilyn J. Baleshiski at ACTEX Publications. The author is very appreciative of her considerable skills, professionalism, and patience.

We hope that readers familiar with the prior edition of this text will agree that the changes reflected in the new edition make a valuable contribution to actuarial education as we enter the twenty-first century. As always, corrections and suggestions for improvements are welcome.

Winsted, Connecticut
April, 1997

Dick London, FSA

SURVIVAL MODELS AND THEIR ESTIMATION

Third Edition

PART I

THE NATURE AND PROPERTIES OF SURVIVAL MODELS

The main topic of this text is the statistical estimation of survival models and the analysis of those estimated models.

Before we tackle the estimation idea, however, we must first develop a considerable familiarity with survival models themselves, and that is the purpose of the first three chapters of the text.

Chapter 1 introduces the general idea of survival models in a conceptual manner, and gives an overview of the entire text.

Chapter 2 presents a symbolic analysis of the survival model, and gives several examples of distributions that might be used as parametric survival models.

Chapter 3 describes the nature and properties of the traditional tabular survival model, the life table. A strong effort is made in this chapter to show that life tables, assisted by mortality distribution assumptions, have the same capabilities as the parametric models of Chapter 2.

CHAPTER 2

THE MATHEMATICS OF SURVIVAL MODELS

2.1 INTRODUCTION

Before we begin our exploration of the topic of estimating a survival model, we need to develop a complete understanding of the nature of survival models themselves.

Since a survival model is a special kind of probability distribution, most of the material in this chapter will be familiar to those with a good knowledge of probability. Furthermore, the survival model is discussed in many standard textbooks on actuarial mathematics. (See, for example, Bowers, et al. [12].)

2.2 THE DISTRIBUTION OF T

2.2.1 The Survival Distribution Function

In Chapter 1 we chose to define and describe a survival model in terms of the function $S(t)$, which represents $Pr(T > t)$, where T is the failure time random variable. This function of the random variable T is called the Survival Distribution Function (SDF). We recall that it gives the probability that failure (death) will occur *after* time t, which is the same as the probability that the entity, known to exist at time $t = 0$, will survive to *at least* time t. We also recall that $S(0) = 1$ and $S(\infty) = 0$.

2.2.2 The Cumulative Distribution Function

The Cumulative Distribution Function (CDF) of T is $F(t)$. The CDF gives the probability that the random variable will assume a value less than or equal to t. That is,

$$F(t) = Pr(T \leq t). \tag{2.1}$$

In the special case of our failure time random variable, $F(t)$ gives the probability that failure (death) will occur *not later than* time t. It should be clear that

$$F(t) = 1 - S(t), \tag{2.2}$$

and that $F(0) = 0$ and $F(\infty) = 1$.

In most probability textbooks, the CDF, $F(t)$, is given greater emphasis than is the SDF, $S(t)$. But for our special kind of random variable, $S(t)$ will receive greater attention.

2.2.3 The Probability Density Function

For the special case of a continuous random variable, the Probability Density Function (PDF), $f(t)$, is defined as the derivative of $F(t)$. Thus

$$f(t) = \frac{d}{dt}F(t) = -\frac{d}{dt}S(t), \quad t \geq 0. \tag{2.3}$$

Consequently, it is easy to see that

$$F(t) = \int_0^t f(y)\,dy, \tag{2.4}$$

and

$$S(t) = \int_t^\infty f(y)\,dy. \tag{2.5}$$

Of course it must be true that

$$\int_0^\infty f(y)\,dy = 1. \tag{2.6}$$

Although we have given mathematical definitions of $f(t)$, it will be useful to describe $f(t)$ more fully in the context of the failure time random variable. Whereas $F(t)$ and $S(t)$ are probabilities which relate to certain time *intervals*, $f(t)$ relates to a point of time, and is not a probability, per se. We prefer to refer to $f(t)$ by its conventional description as "probability density." It is the density of failure *at* time t, and is an *instantaneous* measure, as opposed to an interval measure.

It is important to recognize that $f(t)$ is the *unconditional* density of failure at time t. By this we mean that it is the density of failure at time t given *only* that the entity existed at $t = 0$. The significance of this point will become clearer in the next subsection.

2.2.4 The Hazard Rate Function

We have just established that the PDF of $T, f(t)$, is the unconditional density of failure at time t. We now define a *conditional* density of failure at time t, such density to be conditional on survival to time t. This conditional instantaneous measure of failure at time t, given survival to time t, will be called the ***hazard rate*** at time t, or the Hazard Rate Function (HRF) when viewed as a function of t. It will be denoted by $\lambda(t)$.

In general, if a conditional measure is multiplied by the probability of obtaining the condition, then the corresponding unconditional measure will result. Specifically,

(Conditional density of failure at time t, given survival to time t)
$\times$ (Probability of survival to time t)
$=$ (Unconditional density of failure at time t).

Symbolically this states that

$$\lambda(t) \cdot S(t) = f(t), \tag{2.7}$$

or

$$\lambda(t) = \frac{f(t)}{S(t)}. \tag{2.8}$$

Mathematically, Equations (2.8) and (2.3) define the HRF and the PDF of the failure time random variable, and these mathematical definitions are, of course, very important. However, it is equally important to have a clear understanding of the *descriptive* meanings of $\lambda(t)$ and $f(t)$. They are both instantaneous measures of the density of failure at time t; they differ from each other in that $\lambda(t)$ is conditional on survival to time t, whereas $f(t)$ is unconditional (*i.e.*, given only existence at time $t = 0$).

In the actuarial context of human survival models, failure means death, or mortality, and the hazard rate is normally called the ***force of mortality***. We will discuss the actuarial context further in this chapter and in Chapter 3.

Some important mathematical consequences follow directly from Equation (2.8). Since $f(t) = -\frac{d}{dt}S(t)$, it follows that

$$\lambda(t) = \frac{-\frac{d}{dt}S(t)}{S(t)} = -\frac{d}{dt}\ln S(t). \tag{2.9}$$

Integrating, we have

$$\int_0^t \lambda(y)\,dy = -\ln S(t), \tag{2.10}$$

or

$$S(t) = exp\left[-\int_0^t \lambda(y)\,dy\right]. \tag{2.11}$$

The Cumulative Hazard Function (CHF) is defined to be

$$\Lambda(t) = \int_0^t \lambda(y)\,dy = -\ln S(t), \tag{2.12}$$

so that

$$S(t) = e^{-\Lambda(t)}. \tag{2.13}$$

2.2.5 The Moments of the Random Variable T

The first moment of a continuous random variable defined on $[0, \infty)$ is given by

$$E[T] = \int_0^\infty t \cdot f(t)\,dt, \tag{2.14}$$

if the integral exists, and otherwise the first moment is undefined. Integration by parts yields the alternative formula

$$E[T] = \int_0^\infty S(t)\,dt, \tag{2.15}$$

a form which is frequently used to find the first moment of a failure time random variable.

The second moment of T is given by

$$E[T^2] = \int_0^\infty t^2 \cdot f(t)\,dt, \tag{2.16}$$

if the integral exists, so the variance of T can be found from

$$Var(T) = E[T^2] - \{E[T]\}^2. \tag{2.17}$$

Specific expressions can be developed for the moments of T for specific forms of $f(t)$. This will be pursued in the following section.

Another property of the future lifetime random variable that is of interest is its median value. We recall that the median of a random variable is the value for which there is a 50% chance that T will exceed (and thus also not exceed) that value. Mathematically, y is the median of T if

$$Pr(T \geq y) = Pr(T \leq y) = \tfrac{1}{2}, \tag{2.18}$$

so that $S(y) = F(y) = \frac{1}{2}$.

2.2.6 Actuarial Survival Models

Thus far in this section we have considered only the random variable T, and have looked at various quantities related to that random variable and the interrelationships among those quantities. Exactly the same quantities and relationships exist for the actuarial survival model represented by the SDF $S(x)$, $x \geq 0$.

Special symbols are used in the actuarial context for some of the concepts defined in this section. The hazard rate, called the force of mortality, is denoted by μ_x, rather than $\lambda(x)$. Thus

$$\mu_x = \frac{-\frac{d}{dx}S(x)}{S(x)} = -\frac{d}{dx}\ln S(x). \tag{2.9a}$$

It is also customary to denote the first moment of X by $\mathring{e}_0$. Thus

$$\mathring{e}_0 = E[X] = \int_0^\infty x \cdot f(x)\,dx. \tag{2.19}$$

Since $\mathring{e}_0$ is the unconditional expected value of X, given only alive at $x = 0$, it is called the ***complete expectation of life at birth***.

For the select model $S(t;x)$, recall that t is a value of the random variable T, and x is the age at which the person to whom $S(t;x)$ refers was selected. The expected value of T, $E[T;x]$, gives the expected future lifetime (or expectation of life) for a person selected at age x, and is denoted by $\mathring{e}_{[x]}$. The HRF is denoted by $\mu_{[x]+t}$, and is given by

$$\mu_{[x]+t} = \frac{-\frac{d}{dt}S(t;x)}{S(t;x)} = -\frac{d}{dt}\ln S(t;x). \tag{2.9b}$$

We recognize that the moments of X or T given above are all unconditional. Conditional moments, and other conditional measures, are defined conceptually in Section 2.4, and the standard actuarial notation for them is reviewed in Chapter 3.

2.3 EXAMPLES OF PARAMETRIC SURVIVAL MODELS

In this section we explore several non-negative continuous probability distributions which are candidates for serving as survival models. In practice, some distributions fit better than others to the empirical evidence of the shape of a failure time distribution, so we will comment on each distribution we present regarding its suitability as a survival model.

2.3.1 The Uniform Distribution

The uniform distribution is a simple two-parameter distribution, with a constant PDF. The parameters of the distribution are the limits of the interval on the real number axis over which it is defined, and its PDF is the reciprocal of that interval length. Thus if the random variable is defined over the interval $[a,b]$, then $f(t) = \frac{1}{b-a}$ for $a \leq t \leq b$, and $f(t) = 0$ elsewhere.

For the special case of the future lifetime random variable, $a = 0$. Therefore, b is the length of the interval, as well as the greatest value of t for which $f(t) > 0$. When the uniform distribution is used as a survival model, the Greek ω is frequently used for this parameter, so the distribution is defined by

$$f(t) = \frac{1}{\omega}, \; 0 \leq t \leq \omega. \tag{2.20}$$

The following properties of the uniform distribution easily follow, and should be verified by the reader:

$$F(t) = \int_0^t f(y)\,dy = \frac{t}{\omega} \tag{2.21}$$

$$S(t) = 1 - F(t) = \int_t^\omega f(y)\,dy = \frac{\omega - t}{\omega} \tag{2.22}$$

$$\lambda(t) = \frac{f(t)}{S(t)} = \frac{1}{\omega - t} \tag{2.23}$$

$$E[T] \;=\; \int_0^{\omega} t \cdot f(t)\,dt \;=\; \frac{\omega}{2} \tag{2.24}$$

$$Var(T) \;=\; E[T^2] - \{E[T]\}^2 \;=\; \frac{\omega^2}{12} \tag{2.25}$$

The uniform distribution, as a survival model, is not appropriate over a broad range of time, at least as a model for *human* survival. It is of historical interest, however, to note that it was the first continuous probability distribution to be suggested for that purpose, in 1724, by Abraham de Moivre.

The major use of this distribution is over short ranges of time (or age). We will explore this use of the uniform distribution quite thoroughly in Section 3.5.1.

2.3.2 The Exponential Distribution

This very popular one-parameter distribution is defined by its SDF to be

$$S(t) \;=\; e^{-\lambda t},\; t \geq 0,\; \lambda \geq 0. \tag{2.26}$$

It then follows that the PDF is

$$f(t) \;=\; -\frac{d}{dt}\,S(t) \;=\; \lambda e^{-\lambda t}, \tag{2.27}$$

so that the HRF is

$$\lambda(t) \;=\; \frac{f(t)}{S(t)} \;=\; \lambda, \tag{2.28}$$

a constant. In the actuarial context, where the hazard rate is generally called the force of mortality, the exponential distribution is referred to as the ***constant force distribution***.

EXAMPLE 2.1 Show that, for the exponential distribution,

$$E[T] \;=\; \frac{1}{\lambda} \tag{2.29}$$

and

$$Var(T) \;=\; \frac{1}{\lambda^2}. \tag{2.30}$$

$\boxed{\textbf{SOLUTION}}$ $E[T] = \int_0^\infty t \cdot f(t)\,dt = \int_0^\infty t \cdot \lambda e^{-\lambda t}\,dt.$ Integration by parts produces $\int_0^\infty e^{-\lambda t}\,dt$, whence $E[T] = -\frac{1}{\lambda}e^{-\lambda t}\Big|_0^\infty = \frac{1}{\lambda}$. We also have

$$E[T^2] = \int_0^\infty t^2 \cdot \lambda e^{-\lambda t}\,dt = 2\int_0^\infty t \cdot e^{-\lambda t}\,dt = \frac{2}{\lambda}\int_0^\infty e^{-\lambda t}\,dt = \frac{2}{\lambda^2}.$$

Then

$$Var(T) = \frac{2}{\lambda^2} - \left\{\frac{1}{\lambda}\right\}^2 = \frac{1}{\lambda^2}. \qquad \square$$

The exponential distribution, with its property of a constant hazard rate, is frequently used in reliability engineering as a survival model for inanimate objects such as machine parts (see Chapter 12). Like the uniform distribution, however, it is not appropriate as a model for human survival over a broad range, but is used extensively over short intervals, such as one year, due to its mathematical simplicity. This will be explored in Section 3.5.2.

Since we do not contemplate using the uniform or exponential as a model for human survival, we use T, rather than X for our failure time random variable. For the next three distributions, we use X to suggest that they are more useful as models of human survival.

2.3.3 The Gompertz Distribution

This distribution was suggested as a model for human survival by Gompertz [33] in 1825. The distribution is usually defined by its hazard rate as

$$\lambda(x) = Bc^x, \ x \geq 0, \ B > 0, \ c > 1. \tag{2.31}$$

Then the SDF is given by

$$S(x) = exp\left[-\int_0^x \lambda(y)\,dy\right] = exp\left[\frac{B}{\ln c}(1 - c^x)\right]. \tag{2.32}$$

The PDF is given by $\lambda(x) \cdot S(x)$, and is clearly not a very convenient mathematical form. In particular, the mean of the distribution, $E[X]$, is not easily found.

2.3.4 The Makeham Distribution

In 1860 Makeham [53] modified the Gompertz distribution by taking the HRF to be

$$\lambda(x) \;=\; A + Bc^x,\; x \geq 0,\; B > 0,\; c > 1,\; A > -B. \qquad (2.33)$$

Makeham was suggesting that part of the hazard at any age is independent of the age itself, so a constant was added to the Gompertz hazard rate.

The SDF for this distribution is given by

$$S(x) \;=\; exp\left[-\int_0^x (A+Bc^y)\,dy\right] \;=\; exp\left[\frac{B}{\ln c}\,(1-c^x) - Ax\right]. \qquad (2.34)$$

Again it is clear that the PDF for this distribution is not mathematically tractable, so the calculation of probabilities, moments, or other quantities is somewhat difficult.

2.3.5 The Weibull Distribution

This distribution is defined by

$$\lambda(x) \;=\; k \cdot x^n\,,\; x \geq 0,\; k > 0,\; n > -1. \qquad (2.35)$$

Its SDF is given by

$$S(x) \;=\; exp\left[-\int_0^x k \cdot y^n\,dy\right] \;=\; exp\left[-\;\frac{k \cdot x^{n+1}}{n+1}\right]. \qquad (2.36)$$

2.3.6 Other Distributions

Other probability distributions are very useful as models for other random variables, such as the ***amount of claim random variable*** in non-life insurance applications (see, for example, Hogg and Klugman [37]). These distributions, which include the gamma, the chi-square (a special case of the gamma), the normal, the lognormal, the Pareto, and others, are not appropriate for the failure time random variable which we are considering in this text.

The chi-square distribution, however, is useful in testing the fit of empirical data to a hypothesized parametric distribution (see Chapter 8).

2.3.7 Summary of Parametric Models

We have briefly explored five distributions here: two (uniform and exponential) which are mathematically simple, and three (Gompertz, Makeham and Weibull) which are not.

For our actuarial survival model, denoted by $S(x)$, the last three will receive further consideration in Chapter 8. For many illustrations, where we wish to avoid mathematical complexity, we will use the uniform or the exponential for illustrative purposes only, not necessarily suggesting that they are applicable in practice. The exponential has been assumed to be applicable in many situations not involving healthy human lives, and has been widely used in those situations.

2.4 CONDITIONAL MEASURES AND TRUNCATED DISTRIBUTIONS

Thus far we have only considered probabilities measured from age $x = 0$, denoting such probabilities by $S(x)$ or $F(x)$. Specifically, such probabilities were unconditional, since we knew only that the person was alive at $x = 0$. Now we consider the case of a person known to be alive at age $x > 0$, and we seek probabilities (and densities) of survival (or failure) measured from age x.

2.4.1 Conditional Probabilities and Densities

What is the probability that a person, known to be alive at age x, will still be alive n years later (*i.e.*, at age $x+n$)? We seek

$$Pr(\text{survival to } x+n, \text{ given survival to } x).$$

If we multiply this conditional probability by the probability of obtaining the condition, which is $S(x)$, we obtain the unconditional probability of survival to age $x+n$, which is $S(x+n)$. Thus the desired probability, which we denote by ${}_np_x$, is given by

$${}_np_x = \frac{S(x+n)}{S(x)}. \tag{2.37}$$

The companion conditional probability for death prior to age $x+n$, given alive at x, is given by

$${}_nq_x = 1 - {}_np_x = \frac{S(x) - S(x+n)}{S(x)}. \tag{2.38}$$

It is important to distinguish ${}_np_x$, a conditional probability, from the unconditional probability represented by $S(n;x)$. In each case we seek the probability that a person age x will survive to age $x+n$. When we determine this probability in accordance with the model $S(x)$, it is conditional, it is denoted by ${}_np_x$, and it is given by $\frac{S(x+n)}{S(x)}$. If the desired probability is determined from $S(t;x)$, then it is unconditional, it is given directly by $S(n;x)$, and it is denoted by ${}_np_{[x]}$, to distinguish it from ${}_np_x$.

Similar remarks hold for the companion probability of death prior to age $x+n$. If it is determined from $S(x)$, it is conditional (on survival to x), it is given by $\frac{S(x)-S(x+n)}{S(x)}$, and it is denoted by ${}_nq_x$. But if this probability is determined from $S(t;x)$, then it is unconditional, it is given directly by $F(n;x)$, and it is denoted by ${}_nq_{[x]}$.

This is not to suggest that we cannot have conditional probabilities in terms of $S(t;x)$, as shown by the following example.

EXAMPLE 2.2 Find, in terms of $S(t;x)$, the probability that a person selected at age x, but known to be alive at $x+10$, will die prior to age $x+20$.

SOLUTION We seek the probability of death prior to age $x+20$, given alive at age $x+10$. We denote this probability by ${}_{10}q_{[x]+10}$. It is equal to $1-Pr(\text{Survival to } x+20, \text{ given survival to } x+10) = 1-{}_{10}p_{[x]+10}$. Now if the conditional probability ${}_{10}p_{[x]+10}$ is multiplied by the probability of obtaining the condition, which is $S(10;x)$, the result is the unconditional probability of survival to $x+20$, namely $S(20;x)$. Thus

$${}_{10}q_{[x]+10} = 1 - {}_{10}p_{[x]+10} = 1 - \frac{S(20;x)}{S(10;x)}. \qquad \square$$

Consider next the PDF for death at age y, given alive at age x, for $y > x$. If this conditional density is multiplied by the probability of obtaining the condition, which is $S(x)$, then the unconditional density, which is $f(y)$, results. Thus the conditional density is

$$\frac{f(y)}{S(x)}. \tag{2.39}$$

We will derive this conditional density more formally in the next subsection.

Finally, consider the conditional HRF (or force of mortality) for death at age y, given alive at age x ($y > x$). Recall that the HRF is itself always conditional on survival to the age at which it applies. (There is no such thing as an unconditional HRF.) Thus, since $y > x$ and μ_y itself is conditional on survival to y, then the statement "given survival to x" is redundant. Therefore, this "conditonal" HRF to which we allude is clearly the same as μ_y itself. This intuitive result will be shown more formally in the following subsection.

2.4.2 Lower Truncation of the Distribution of X

When we speak of probabilities (or densities) conditional on survival to age x, we are dealing with the distribution of a subset of the sample space of the random variable X, namely those values of X which fall in excess of x. This distribution is called the ***distribution of X truncated below at x***.

Our conditional survival probability ${}_np_x$ can now be stated formally as

$$ {}_np_x \;=\; Pr(X > x{+}n \,|\, X > x) \;=\; S(x{+}n \,|\, X > x). \tag{2.40}$$

In words, this asks for the probability that the age at death will exceed $x{+}n$, given that it does exceed x. It is easy to see that this is the same concept as "probability of survival to $x{+}n$, given survival to x." Thus, from Equations (2.37) and (2.40) we find that

$$ S(x{+}n \,|\, X > x) \;=\; \frac{S(x{+}n)}{S(x)}. \tag{2.41}$$

Similarly,

$$ \begin{aligned} {}_nq_x &= Pr(X \le x{+}n \,|\, X > x) \\ &= Pr(x < X \le x{+}n \,|\, X > x) \;=\; F(x{+}n \,|\, X > x). \end{aligned} \tag{2.42}$$

Comparison of Equations (2.38) and (2.42) shows that

$$ F(x{+}n \,|\, X > x) \;=\; \frac{S(x) - S(x{+}n)}{S(x)} \;=\; \frac{F(x{+}n) - F(x)}{1 - F(x)}, \tag{2.43}$$

since $S(x) = 1 - F(x)$. Note that both (2.41) and (2.43) result from the general probability relationship $P(A|B) \cdot P(B) = P(A \cap B)$.

Next, the conditional density function for death at age y, given alive at age x $(y > x)$, is denoted by $f(y \,|\, X > x)$. We have

$$ f(y \,|\, X > x) \;=\; \frac{d}{dy}\, F(y \,|\, X > x) \;=\; \frac{d}{dy}\, \frac{F(y) - F(x)}{1 - F(x)} \;=\; \frac{f(y)}{1 - F(x)}, $$

since $\frac{d}{dy}F(x) = 0$. Thus we have

$$f(y \mid X > x) = \frac{f(y)}{S(x)}, \tag{2.39a}$$

as already established intuitively by (2.39).

Finally, the alleged "conditional HRF at y, given alive at x ($y > x$)," was shown intuitively to be the same as the basic μ_y (or $\lambda(y)$). This result can now be mathematically verified. Denoting this "conditional HRF" by $\lambda(y \mid X > x)$, we have

$$\lambda(y \mid X > x) = \frac{f(y \mid X > x)}{S(y \mid X > x)} = \frac{f(y)}{S(x)} \div \frac{S(y)}{S(x)} = \frac{f(y)}{S(y)} = \lambda(y). \tag{2.44}$$

In summary, the functions $S(y \mid X > x)$, $F(y \mid X > x)$ and $f(y \mid X > x)$ are the functions for the truncated distribution of X, truncated below at x. The HRF for this truncated distribution, denoted by $\lambda(y \mid X > x)$, is identical to the untruncated $\lambda(y)$.

2.4.3 Upper and Lower Truncation of the Distribution of X

A more general view of truncated distributions is to consider the distribution of the subset of the sample space of X which falls between y and z. Still using X for the age at death random variable, the truncated SDF is given by

$$\begin{aligned} S(x \mid y < X \leq z) &= Pr(X > x \mid y < X \leq z) \\ &= Pr(x < X \leq z \mid y < X \leq z), \end{aligned} \tag{2.45}$$

for $y < x \leq z$. In words, we speak of the probability that death will occur after age x, given that it does occur between y and z. Since it must occur prior to z, then we are really talking about death between x and z. If this conditional probability is multiplied by the probability of obtaining the condition, which is $S(y) - S(z)$, then the unconditional probability for death between x and z, which is $S(x) - S(z)$, results. Thus

$$S(x \mid y < X \leq z) = \frac{S(x) - S(z)}{S(y) - S(z)}. \tag{2.46}$$

The corresponding truncated CDF is given by

$$F(x \mid y < X \leq z) = Pr(y < X \leq x \mid y < X \leq z). \tag{2.47}$$

Since $F(x \mid y < X \leq z) = 1 - S(x \mid y < X \leq z)$, we have

$$F(x \mid y < X \leq z) = \frac{S(y) - S(x)}{S(y) - S(z)} = \frac{F(x) - F(y)}{F(z) - F(y)}, \tag{2.48}$$

directly from Equation (2.46).

Next, the doubly-truncated PDF is given by

$$f(x \mid y < X \leq z) = -\frac{d}{dx} S(x \mid y < X \leq z) = -\frac{d}{dx} \frac{S(x) - S(z)}{S(y) - S(z)},$$

producing

$$f(x \mid y < X \leq z) = \frac{f(x)}{S(y) - S(z)}, \tag{2.49}$$

since $-\frac{d}{dx} S(x) = f(x)$, and $-\frac{d}{dx} S(z) = 0$.

Finally, the doubly-truncated HRF is given by

$$\lambda(x \mid y < X \leq z) = \frac{f(x \mid y < X \leq z)}{S(x \mid y < X \leq z)},$$

producing

$$\lambda(x \mid y < X \leq z) = \frac{f(x)}{S(y) - S(z)} \div \frac{S(x) - S(z)}{S(y) - S(z)} = \frac{f(x)}{S(x) - S(z)}.$$

Since $f(x) = \lambda(x) \cdot S(x)$ in the untruncated distribution, then we have

$$\lambda(x \mid y < X \leq z) = \frac{\lambda(x) \cdot S(x)}{S(x) - S(z)}. \tag{2.50}$$

Equation (2.50) shows that, whereas truncation only from below did not affect the HRF, truncation from above does (since the truncated HRF is a function of z). This result is intuitive. Since the HRF at x is conditional on survival to x, truncation below x is immaterial. However, truncation above x has an effect on the HRF at x, since the time interval remaining for death is shortened. It should be clear that as $z \to x$ from above, $\lambda(x \mid y < X \leq z)$ becomes infinitely large, since the interval for death approaches zero. This result is seen mathematically by taking the limit as $z \to x$ in Equation (2.50).

2.4.4 Moments of Truncated Distributions

The first moment of the doubly-truncated distribution of X is given by

$$E[X \mid y < X \leq z] = \int_y^z x \cdot f(x \mid y < X \leq z)\, dx. \tag{2.51}$$

Of special interest is the distribution of X truncated only below at y. Then

$$E[X \mid X > y] = \int_y^\infty x \cdot f(x \mid X > y)\, dx, \tag{2.52}$$

if the expectation exists. Since X is the age at death of a person known to be alive at y, then Equation (2.52) gives the expected age at death for such a person. If we subtract y from this expected age at death, we obtain the expected future lifetime of such a person. This expected future lifetime is denoted by $\mathring{e}_y$, and is called the ***expectation of life at age y***. Formally,

$$\mathring{e}_y = E[X \mid X > y] - y. \tag{2.53}$$

Since $\int_y^\infty f(x \mid X > y)\, dx = 1$, we can write

$$\mathring{e}_y = \int_y^\infty (x-y) \cdot f(x \mid X > y)\, dx \tag{2.54a}$$

$$= \int_0^\infty t \cdot f(t+y \mid X > y)\, dt, \tag{2.54b}$$

and since $f(t+y \mid X > y)$ is the PDF of $(X-y \mid X > y)$, then "expected future lifetime" is a good name for $\mathring{e}_y$.

Furthermore, if it exists, then

$$E[X^2 \mid X > y] = \int_y^\infty x^2 \cdot f(x \mid X > y)\, dx. \tag{2.55}$$

Then the variance of future lifetime is given by

$$Var(X-y \mid X > y) = Var(X \mid X > y) = E[X^2 \mid X > y] - \left\{E[X \mid X > y]\right\}^2. \tag{2.56}$$

2.4.5 The Central Rate

Another type of conditional measure over the interval from age x to age $x+1$ is called the ***central rate of death***, and is denoted by m_x. It is defined as the weighted average value of the HRF $\lambda(x)$ over the interval, using, as the weight for $\lambda(y)$, the probability of survival to age y. Formally,

$$m_x = \frac{\int_x^{x+1} S(y)\cdot\lambda(y)\,dy}{\int_x^{x+1} S(y)\,dy}, \tag{2.57}$$

where the denominator is the sum of the weights for a continuous case weighted average.

More generally, ${}_nm_x$ is the average hazard, or central rate of death, over the interval from x to $x+n$, and is given by

$${}_nm_x = \frac{\int_x^{x+n} S(y)\cdot\lambda(y)\,dy}{\int_x^{x+n} S(y)\,dy} = \frac{\int_0^n S(x+s)\cdot\lambda(x+s)\,ds}{\int_0^n S(x+s)\,ds}, \tag{2.58}$$

the second expression resulting from the simple change of variable $y = x+s$. If we divide both numerator and denominator of (2.58) by $S(x)$, we obtain

$${}_nm_x = \frac{\int_0^n \frac{S(x+s)}{S(x)}\cdot\lambda(x+s)\,ds}{\int_0^n \frac{S(x+s)}{S(x)}\,ds} = \frac{\int_0^n {}_sp_x\mu_{x+s}\,ds}{\int_0^n {}_sp_x\,ds}, \tag{2.59}$$

since ${}_sp_x$ is the conditional probability $\frac{S(x+s)}{S(x)}$, and μ_{x+s} is the standard actuarial symbol for $\lambda(x+s)$. The second expression in (2.59) is a common actuarial form for ${}_nm_x$. We will return to this function in the next chapter, and make some use of it in estimating the tabular survival model in Part II of the text.

EXAMPLE 2.3 If X has an exponential distribution, show that this implies $m_x = -\ln p_x$.

SOLUTION If X is exponential, the HRF is constant, with $\lambda(y) = \lambda$ for all y. Then, from Equation (2.57), $m_x = \frac{\lambda\cdot\int_x^{x+1} S(y)\,dy}{\int_x^{x+1} S(y)\,dy} = \lambda$. Furthermore, we also have $p_x = \frac{S(x+1)}{S(x)} = \frac{e^{-\lambda(x+1)}}{e^{-\lambda(x)}} = e^{-\lambda}$. Thus $\lambda = -\ln p_x$, and since $m_x = \lambda$, then $m_x = -\ln p_x$. □

2.4.6 Use of Conditional Probabilities in Estimation

We have noted that the main business of this text will be the estimation of an operative survival model, such estimation to be based on the data of a sample.

Suppose we wish to estimate, say, $S(10)$, the probability of survival from $t = 0$ to $t = 10$. In many cases the nature of the study (and the data) will suggest that we consider only the time interval from $t = i$ to $t = i+1$, and estimate the conditional probability of survival over that interval. That is, we will estimate $\frac{S(i+1)}{S(i)}$, the probability of survival to $i+1$, given alive at i. This conditional probability has been called p_i, so the *estimate* of it which we obtain will be called $\hat{p}_i$, for $i = 0, 1, \ldots, 9$. We will then obtain our estimate of $S(10)$ by multiplying these several $\hat{p}_i$. Thus we will obtain $\hat{S}(10) = \hat{p}_0 \cdot \hat{p}_1 \cdot \cdots \cdot \hat{p}_9$, or, in general

$$\hat{S}(t) = \hat{p}_0 \cdot \hat{p}_1 \cdot \cdots \cdot \hat{p}_{t-1}. \tag{2.60}$$

In many cases it will be natural to first estimate q_i, the conditional probability of failure (death) in $(i, i+1]$, given alive at i; then take $\hat{p}_i = 1 - \hat{q}_i$, where $\hat{q}_i$ is the estimate of q_i; and finally obtain $\hat{S}(t)$ by multiplying these conditional $\hat{p}_i$ estimates. This approach to estimating the survival function will be utilized mainly in Chapters 6 and 7.

2.5 TRANSFORMATIONS OF RANDOM VARIABLES

Suppose we have a random variable X, with known probability distribution, and we consider a new random variable Y, which is some function of X. That is, let

$$y = g(x) \tag{2.61}$$

be a function of x such that the inverse function $x = g^{-1}(y) = h(y)$ exists. We seek the probability distribution of Y.

Let $y = g(x)$ be a strictly increasing function, as shown in Figure 2.1 on page 30. Since $g(x)$ is increasing, then if X is less than or equal to x, it follows that Y is less than or equal to the unique value of y which corresponds to the given value of x. Thus if $X \leq x$, then $Y \leq g(x)$. Conversely, if $Y \leq y$, then $X \leq h(y)$, and the probabilities of these events are equal. That is,

$$Pr(Y \leq y) = Pr(X \leq h(y)), \tag{2.62}$$

or

$$F(y) = F[h(y)]. \tag{2.63}$$

Equation (2.63) can be confusing since the CDF's on opposite sides of the equation are not the same function. The one on the left is the CDF of the random variable Y, whereas the one on the right is for the random variable X. To clarify this we write

$$F_Y(y) = F_X[h(y)]. \tag{2.63a}$$

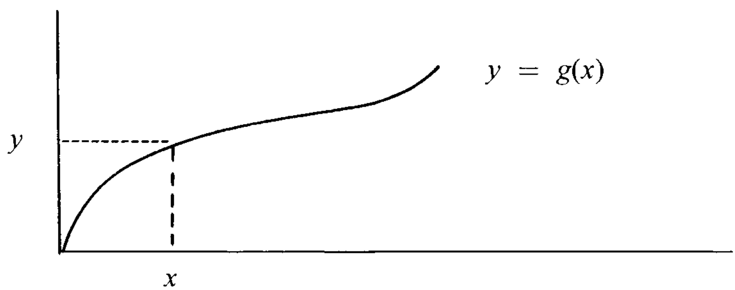

FIGURE 2.1

From (2.63a), which relates the CDF's of the random variables X and Y, we can derive relationships for the SDF's, PDF's and HRF's as well.

Since the SDF is the complement of the CDF, it follows from (2.63a) that $1 - S_Y(y) = 1 - S_X[h(y)]$, or that

$$S_Y(y) = S_X[h(y)]. \tag{2.64}$$

Next, the PDF is the derivative of the CDF, so we differentiate both sides of (2.63a) with respect to y, obtaining

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X[h(y)] = f_X[h(y)] \cdot \frac{d}{dy} h(y),$$

using the chain rule to differentiate $F_X[h(y)]$. Since $h(y)$ is simply x, we can write

$$f_Y(y) = f_X[h(y)] \cdot \frac{dx}{dy}. \tag{2.65}$$

Finally, the HRF is the ratio of the PDF to the SDF. Thus

$$\lambda_Y(y) = \frac{f_Y(y)}{S_Y(y)} = \frac{f_X[h(y)] \cdot \frac{dx}{dy}}{S_X[h(y)]} = \lambda_X[h(y)] \cdot \frac{dx}{dy}. \tag{2.66}$$

EXAMPLE 2.4 Suppose X has an exponential distribution with $\lambda = 1$. Let $y = g(x) = x^{1/2}$. Find the SDF, PDF and HRF of the transformed random variable Y.

SOLUTION Note that $y = g(x)$ is strictly increasing, and $x = h(y) = y^2$. Then $F_Y(y) = F_X(y^2)$. Since X is exponential, then $F_X(x) = 1-e^{-x}$, so we have $F_Y(y) = 1 - e^{-y^2}$, and $S_Y(y) = e^{-y^2}$. Next, $f_Y(y) = \frac{d}{dy} F_Y(y) = 2y \cdot e^{-y^2}$. Finally, $\lambda_Y(y) = \frac{f_Y(y)}{1-F_Y(y)} = \frac{2y \cdot e^{-y^2}}{e^{-y^2}} = 2y$. Alternatively, since we have $\lambda_X(x) = \lambda_X[h(y)] = 1$, then, from (2.66), $\lambda_Y(y)$ is simply $\frac{dx}{dy} = 2y$. Note that Y has a Weibull distribution with $k = 2$ and $n = 1$. □

If $y = g(x)$ is a strictly decreasing function, as shown in Figure 2.2, then our reasoning and our results change a bit.

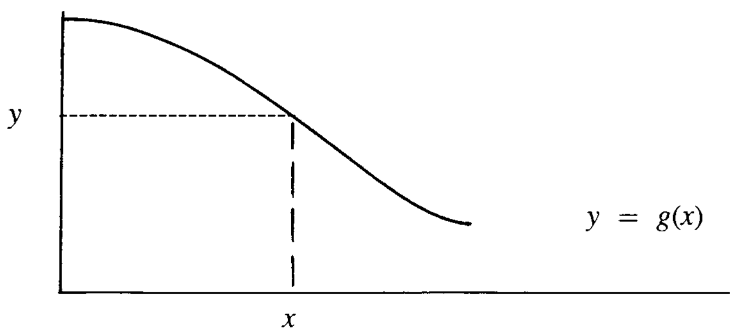

FIGURE 2.2

Here we can see that if X is less than x, then Y *will be greater* than the value of y which corresponds to the given value of x; or, conversely, if $Y > y$, then $X < h(y)$. In terms of probabilities, we have

$$Pr(Y > y) = Pr(X < h(y)) = Pr(X \leq h(y)) \tag{2.67}$$