# Springer Series in Statistics

# Springer Series in Statistics

*Andersen/Borgan/Gill/Keiding:* Statistical Models Based on Counting Processes.
*Atkinson/Riani:* Robust Diagnostic Regression Analysis.
*Atkinson/Riani/Cerioli:* Exploring Multivariate Data with the Forward Search.
*Berger:* Statistical Decision Theory and Bayesian Analysis, 2nd edition.
*Borg/Groenen:* Modern Multidimensional Scaling: Theory and Applications.
*Brockwell/Davis:* Time Series: Theory and Methods, 2nd edition.
*Chan/Tong:* Chaos: A Statistical Perspective.
*Chen/Shao/Ibrahim:* Monte Carlo Methods in Bayesian Computation.
*Coles:* An Introduction to Statistical Modeling of Extreme Values.
*David/Edwards:* Annotated Readings in the History of Statistics.
*Devroye/Lugosi:* Combinatorial Methods in Density Estimation.
*Efromovich:* Nonparametric Curve Estimation: Methods, Theory, and Applications.
*Eggermont/LaRiccia:* Maximum Penalized Likelihood Estimation, Volume I: Density Estimation.
*Fahrmeir/Tutz:* Multivariate Statistical Modelling Based on Generalized Linear Models, 2nd edition.
*Fan/Yao:* Nonlinear Time Series: Nonparametric and Parametric Methods.
*Farebrother:* Fitting Linear Relationships: A History of the Calculus of Observations 1750–1900.
*Federer:* Statistical Design and Analysis for Intercropping Experiments, Volume I: Two Crops.
*Federer:* Statistical Design and Analysis for Intercropping Experiments, Volume II: Three or More Crops.
*Ghosh/Ramamoorthi:* Bayesian Nonparametrics.
*Glaz/Naus/Wallenstein:* Scan Statistics.
*Good:* Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses, 2nd edition.
*Gouriéroux:* ARCH Models and Financial Applications.
*Gu:* Smoothings Spline ANOVA Models.
*Györfi/Kohler/Krzyżak/Walk:* A Distribution-Free Theory of Nonparametric Regression.
*Haberman:* Advanced Statistics, Volume I: Description of Populations.
*Hall:* The Bootstrap and Edgeworth Expansion.
*Härdle:* Smoothing Techniques: With Implementation in S.
*Harrell:* Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis.
*Hart:* Nonparametric Smoothing and Lack-of-Fit Tests.
*Hastie/Tibshirani/Friedman:* The Elements of Statistical Learning: Data Mining, Inference, and Prediction.
*Hedayat/Sloane/Stufken:* Orthogonal Arrays: Theory and Applications.
*Heyde:* Quasi-Likelihood and its Application: A General Approach to Optimal Parameter Estimation.
*Huet/Bouvier/Poursat/Jolivet:* Statistical Tools for Nonlinear Regression: A Practical Guide with S-PLUS and R Examples, 2nd edition.
*Ibrahim/Chen/Sinha:* Bayesian Survival Analysis.

*(continued after index)*

Sam C. Saunders

# Reliability, Life Testing and the Prediction of Service Lives

## For Engineers and Scientists

Springer

Sam C. Saunders
Kirkland, WA 98033
USA
smru@comcast.net

Library of Congress Control Number: 2006921753

ISBN-10: 0-387-32522-0 e-ISBN-10: 0-387-48538-4
ISBN-13: 978-0-387-32522-4 e-ISBN-13: 978-0-387-48538-6

Printed on acid-free paper.



9 8 7 6 5 4 3 2 1

springer.com

# Preface

The prerequisite for reading this text is a calculus-based course in Probability and Mathematical Statistics, along with the usual curricular mathematical requirements for every science major. For graduate students from disciplines other than mathematical sciences much advantage, viz., both insight and mathematical maturity, is gained by having had experience quantifying the assurance for safety of structures, operability of systems or health of persons. It is presumed that each student will have some familiarity with Mathematica or Maple or better yet also have available some survival-analysis software such as S-Plus or R, to handle the computations with the data sets.

This material has been selected under the conviction that the most practical aid any investigator can have is a good theory. The course is intended for persons who will, during their professional life, be concerned with the 'theoretical' aspects of applied science. This implies consulting with industrial mathematicians/statisticians, lead engineers in various fields, physcists, chemists, material scientists and other technical specialists who are collaborating to solve some difficult technological/scientific problem. Accordingly, there are sections devoted to the deportment of applied mathematicians during consulting. This corresponds to the 'bedside manner' of physicians and is a important aspect of professionalism.

While Henri Poincaré lectured successively in: capillarity, elasticity, thermodynamics, optics, electricity, telegraphy, cosmogeny, not to name all; very few of us can be such *universalists*. But he was an expert in each of these fields because he could understand the mathematical problems at the foundations of each. That is what we hope, in small measure, to foster here: To present the basic methods for application of probability and statistics to the ubiquitous task of calculating the reliability, or its equivalent, for some of the engineered systems in modern civilization.

Remembering the sense of satisfaction I obtained as a student when I discovered an oversight in a textbook, I have not sought, exhaustively, to deprive the readers of this text from experiencing that same private exhilaration.

> The beginner... should not be discouraged if... he finds he does not have the prerequisites for reading the prerequisites.
>
> Paul Halmos

> Science is not a collection of facts anymore than a heap of stones is a house,
>
> Henri Poincaré

# Acknowledgements

An acknowledgement is owed to influential teachers and exemplars; the former category includes Professors Ralph Badgely, Ivan Niven, and Z.W. Birnbaum and in the latter are Carl Allendoerfer and, Edwin Hewitt. The Mathematical Analysis group, headed by Burton Colvin, at the erstwhile Boeing Scientific Research Laboratories contained notable colleagues Frank Proschan, George Marsaglia, Albert Marshall, Gordon Crawford and James Esary. Z.W.(Bill) Birnbaum said this group rivaled the Analysis Group led by Lev Sierspinski at Lwöw, Poland, when he studied there with Stephan Banach. Students have helped in organizing material, in correcting my errors and suggesting clarifications. In particular Prof. Juhn-Hsiong Wong, Prof. Jung Soo Woo and Dr. Jonathan Martin are owed a debt of gratitude.

Of course we are all influenced by our genealogy: ··· Ferdinand Lindemann begat David Hilbert who begat Hugo Steinhaus who begat Z. W. Birnbaum who begat my mathematical siblings Ron Pyke and Albert Marshall, who have both remained life-long colleagues and friends.

> Every man who rises above the common level has received two educations: the first from his teachers; the second, more personal and important, from himself.
>
> Edward Gibbon

> If I have seen farther than others it was because I was standing on the shoulders of giants.
>
> Sir Isaac Newton

# Vörtrekkers

1. *Statistical Theory of Reliability and Life Testing;* Richard Barlow and Frank Proschan, Holt, Rinehart & Winston, 1981, reprinted SIAM 1996.
2. *Probabilistic Models of Cumulative Damage;* J.L. Bogdanoff and F. Kozin, John Wiley & Sons Inc., 1985.

# Glossary

- as *means* "almost surely or with probability one"
- arv *means* " associated random variable or vector"
- asas *means* "after some algebraic simplification"
- cdf *means* "cumulative distribution function"
- pdf *means* "probability density function"
- sdf *means* "survival distribution function"
- edf *means* the same as ecdf or "empirical (cumulative) distribution function"
- esf *means* the same as "empirical survival distribution function"
- iff *means* "if and only if"
- iid *means* "independent and identically distributed"
- K-M *means* "Kaplan - Meier" e.g. as an affix to edf
- mle *means* "maximum likelihood estimator"
- nasc *means* " necessary and sufficient condition"
- NB *means* Nota Bene, Latin for "It should be well noted that"
- rhs (or lhs) *means* "right-hand side" (left-hand side)
- rwt *means* "random waiting time"
- rv or rv's *means* "random variable or random vector and its plural"
- sp *means* "stochastic process"
- tidpat *means* "Thus it doth plainly appear that" (Lagrange's phrase) but it is often paraphrased as "This is difficult, paradoxical and tedious."
- wrt *means* " with respect to"
- wlog *means* " without loss of generality"
- wp *means* " with probability"
- $:=$ *means* "is defined to be equal to"
- $\doteq$ *means* "is closely approximated by"
- $\asymp$ *means* "is asymptotically equal to"
- $\ll$ *means* "is much less than "
- $\preceq$, *means* "is stochastically less than"
- $\downarrow$ ($\uparrow$) *means* " non-increasing" (non-decreasing); so $F \in\uparrow$ means $F$ is non-decreasing.
- $\perp$ *means* "mutually, stochastically independent"
- $\sim$ *means* " has the distribution or is distributed by"
- $\square$ *means* the same as *quod erat demonstrandum* and marks the end of a proof.
- $\imath = \sqrt{-1}$ is the unit of imaginary numbers
- $\Re$ denotes the real line, viz., $\{x : -\infty < x < \infty\}$
- $\mathrm{I}(x \pi y)$ is the indicator of the relation $x \pi y$ taking the value 1 if true, 0 otherwise

# Admonitions

··· Mathematical ideas originate in empirics, although the genealogy is sometimes long and obscure. But, once they are so conceived, the subject begins to live a peculiar life of its own and is better compared to a creative one, governed by almost entirely aesthetical motivations, than to anything else, in particular, to an empirical science. There is, however, a further point which, I believe, needs stressing. As a mathematical discipline travels far from its empirical source, or still more, if it is a second and third generation only indirectly inspired by ideas coming from 'reality', it is beset with very grave dangers.It becomes more and more purely aestheticising, more and more purely *l'art pour l'art.* This need not be bad, if the field is surrounded by correlated subjects, which still have closer empirical connections, or if the discipline is under the influence of men with an exceptionally well-developed taste. But there is a grave danger that the subject will develop along the line of least resistance, that the stream, so far from its source, will separate into a multitude of insignificant branches and that the discipline will become a disorganized mass of details and complexities. In other words, at a great distance from its empirical source, or after much 'abstract' inbreeding a mathematical subject is in danger of degeneration.

John von Neumann

An explanation is satisfactory only if we are able to reconstruct it logically from our previous knowledge and apply that understanding to circumstances different from those in which it was originally offered. That is why science teachers, to the chagrin of many students in the humanities, put heavy emphasis on problem solving. In order to demonstrate that (s)he has understood a scientific principle, a student is expected to be able to apply this understanding to situations different from the ones in which it was first learned. Similarly, a mathematics student is deficient who knows a theorem, in general, but cannot apply it in an unfamiliar context. Neither memorizing nor reproducing what one has seen or heard in a lecture ensures understanding.

Roger O. Newton

Let no one who is ignorant of Geometry (mathematics) enter here (proceed farther).

Written at the entrance to Plato's Academy

Mathematicians are like Frenchmen: whatever you say to them they translate into their own jargon and thenceforth it becomes something entirely different.

Johann W. von Goethe

We have therefore the equation of condition

$$F(x) = \int dq\, Q(q) \cos(qx).$$

If we substitute for $Q$ any function of $q$ and conduct the integration from $q = 0$ to $q = \infty$, we should find a function of $x$; it is required to solve the inverse problem, that is to say, to ascertain what function of $q$, after being substituted for $Q$, gives as a result the specified function $F(x)$, a remarkable problem whose solution demands attentive examination.

Joseph Fourier

# Contents

CHAPTER 1

# Requisites

## 1.1. Why Reliability Is Important

All artifacts of mankind will eventually fail in service or be discarded because of wear or obsolescence. This was as true for the roads and aquaducts of Rome as it is today for the infrastructure of America. All constructs bearing the hallmark of civilization, from the tomahawk to the cruise missle of the same name, suffer from material weakness or imperfection. This applies *a fortiori* to our electronic systems, computers, and video communication as well as the military's smart-bombs and the Concorde's avionics. Palliative efforts include preventative maintenance for machines and medicines for humankind. But always "an ounce of prevention is better than a pound of cure" and not only because it is earlier and cheaper. The origins of ubiquitious failure are manifold; the designer has neglected or been unaware of the severity of some of the factors of the environment in which the system/structure must operate; the owner-operator of the system has needfully operated it outside its design envelope; the manufacturer failed to eliminate minor defects from the system either during construction or inspection; the supplier has substituted inferior material, causing inherent weaknesses in a component. Such imperfections can cause the early and unexpected failure of the system, or just the incapability of the system to perform its function during its warranteed life. Such practice may lead to hazards both to the operators and to the public weal.

The primary source of failure is the gradual impairment of structural components caused through repetition of their designed duty cycle. These include friction or abrasive wear, metallic fatigue, stress corrosion or chemical degradation. Failure is caused by mistakes (accidents) aggravated by operation or insufficient maintenance, more often than the confluence of unhappy circumstance, labeled "acts of God."

Most often, system or structural failure is the result of many coincident factors. The failure of a dynamically loaded structure may, for example, be the result of a small defect in a critical component undetected because of the insufficient quality control during production. This leads to crack initiation, and a growing fatigue crack accelerated by a corrosive environment; thus, ultimately an extreme random load exceeds the residual strength of the component. Who was at fault? Were the designers, the users, or the nature or all three?

Unfortunately, all failures in service have undesirable economic consequences and not always to those responsible. For example, the destruction by bomb of a Boeing 747 over Lockerbie, Scotland, resulted in the bankruptcy of Pan-American

Airlines; the fading of a new environmentally safe paint from UV-radiation caused the bankruptcy of the Studebaker Avanti automobile corporation. Fortunately, most system failures necessitate only repair or component replacement with its entailed disruption of service. Sometimes such failures may cause an interruption in manufacturing among secondary users. Sometimes dramatic failures force a system redesign or even concept abandonment.

As a rule of thumb, the expenditure of funds for the maintenance of systems, structures, machinery, or equipment amounts to about half the initial investment cost before obsolescence forces replacement. For well-designed, long-life items maintenance expenses may be much more than the initial cost. It is estimated that about 6–8% of the gross national product (GNP) is spent annually for maintenace. This may be small in comparison with the loss of production attributable to unwarranted in-service failures.

Today many purchasers of equipment are aware of the cost of subsequent maintenance. Consequently, they consider not only purchase price but the total life-cycle costs, including maintenance and repair. This is now routinely done for the evaluation of system proposals by industry for the military and certain governmental divisions, such as the federal highways. For political reasons these costs are often underestimated. In the 18th century the costs for the construction and operation of the frigate *Constitution*, "Old Ironsides," were underestimated by about the same percentage as they were in the 20th century for the Stealth Bomber; and with the same perceived political reasons for such mendacity.

For many large complex systems/structures, such as high-rise buildings, nuclear power plants, off-shore oil structures, aircraft, and life-support medical equipment, the consequences of unreliability involve public welfare and safety. Of course, the failure of comparatively minor products can, besides just being an annoyance, have serious consequences for the public, for example, the failure of an electric razor or a battery in a fire-alarm system.

Issues concerning safety have, during the last few decades, come increasingly to public attention and hence they become more important to the design engineer. One of the most controversial is nuclear power (since Chernobyl and Three-Mile Island), one of the most dramatic is in-flight safety, while the much higher frequency of death from automobile accidents is, relatively, of small concern. Governmental requirements for safety analysis of systems is increasing, as it should. Legislation aportioning responsibility for product liability is proceeding, as it should, to help avoid the excessively large compensation claims awarded by the American tort system, whenever insurance for the manufacturer is available or the industry is large enough to have "deep pockets."

All of this will increase the need for industry to perform systematic studies for the identification and reduction of causes of failures (with hope for their virtual elimination). These studies must be performed by persons who (i) can identify and quantify the modes of failure, (ii) know how to obtain and analyze the statistics of failure occurrences, and (iii) can construct mathematical models of failure that depend upon, for example, the parameters of material strength or design quality, fatigue or wear resistance, and the stochastic nature of the anticipated duty cycle.

Only then can procedures for optimal design be implemented in parallel with a study of the economic consequences for each failure mode so as to reduce, in an optimal way, the probability of their occurrence.

The purpose of this book is to help supply information about the mathematical and statistical aspects of calculating the reliability in order to make a valid service-life prediction for the use of cognizant persons and to help provide them with the capability to utilize or develop analytical techniques specific to their usage and the needs of their employers.

> The beginner ... should not be discouraged if ... he finds that he does not have the prerequisites for reading the prerequisites.
>
> Paul Halmos

# 1.2. Valuable Concepts

## 1.2.1. Concepts from Probability

Consider the outcome of a well-defined experiment, the result of which can not be exactly anticipated, except it will terminate in some measured quantity, denoted by $X$, within a known set of possible outcomes, labeled $\mathcal{X}$ and called the *sample space*. This measured quantity $X$ is called a *random* or *stochastic variable* (which we abbreviate by rv). We will use $X$, $Y$, $Z$, with or without affixes, to denote rv's. We presume the sample space of each rv is a subset of the real line $\Re$, or of the product of real lines, $\Re^n$ for some $n$. Associated with each experiment is the relative frequency of the different outcomes, which would occur if the experiment could be replicated indefinitely. This probabilistic behavior of the repeated determinations of $X$ is summarized in a mathematical function, called the *cumulative distribution function*, abbreviated cdf and usually denoted by $F_X$ (without affix when no ambiguity results), which maps $\Re$, or $\Re^n$, respectively, onto the unit interval. This function may often be classified into one of the two cases, viz., *discrete* or *continuous*, according to whether as $F_X$ is a step function or is absolutely continuous with density $F'_X := f_X$. In the latter case, that "$X$ is continuous," by which we mean the the cdf is absolutely continuous, it is given by

$$\Pr[X \leq x] := F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt \qquad \text{for all } x \in \Re. \tag{1.1}$$

The *support* of a continuous rv $X$ is the set $\mathcal{X} = \{x \epsilon \Re \mid f_X(x) > 0\}$. When $X$ is discrete, the support is some countable set, say $\mathcal{X} = \{x_1, x_2, \ldots\}$, where

$$\Pr[X = x] = \begin{cases} p_i > 0 & \text{if } x = x_i \text{ for some } i = 1, 2, \ldots, \\ 0 & \text{otherwise.} \end{cases} \tag{1.2}$$

Here $p_i$ is the height of the $i$th saltus of the cdf $F_X$ and

$$\Pr[X \leq x] := F_X(x) = \sum_{x_i \leq x} p_i \qquad \text{for all } x \in \Re. \tag{1.3}$$

If $Y$ be a constant rv, i.e., the event $[Y = c]$ occurs with probability one for some $c \in \Re$, then we write the cdf $F_Y$ in terms of the indicator function of an event, viz., $\mathrm{I}(x \in A) = 0$ iff $x \in A$ and $\mathrm{I}(x \in A) =)$ iff $x \notin A$. When the interval is $[c, \infty)$, we write

$$F_Y(y) = \mathrm{I}(y \geq c) = \begin{cases} 1 \text{ for } y \geq c, \\ 0 \text{ for } y < c. \end{cases} \quad \text{for all } y \in \Re. \tag{1.4}$$

In physics, $\varepsilon(y) := \mathrm{I}(y \geq 0)$ is called the *Heaviside* function. The "derivative" of $\varepsilon(y)$ is the *Dirac delta-function*, which is often used in heuristic arguments, especially when the correctness of the result can be verified empirically.

Thus, the distribution of any discrete rv, as defined in eqn (1.3), can be written

$$F_X(x) = \sum_{i=1}^{\infty} p_i \mathrm{I}(x \geq x_i) \qquad \text{for all } x \in \Re.$$

We say that $X$ is a *mixed* rv iff for some $\gamma \in (0, 1)$ we have, for every $x \in \Re$, and some denumerable subset $\{x_1, x_2, \ldots\}$,

$$F_X(x) = \gamma \sum_{i=1}^{\infty} p_i \mathrm{I}(x \geq x_i) + (1 - \gamma) \int_{-\infty}^{x} f_X(t)\, dt. \tag{1.5}$$

*Familiar Discrete Densities and Distributions*

A *Bernoulli* rv, $X$, is a binary rv that takes the values 0 or 1. Thus, its pdf, i.e., probability distribution function, is defined by

$$\mathrm{P}[X = x] = f(x; p) = p^x q^{1-x} \qquad \text{for} \quad x = 0, 1; 0 < p < 1, q = 1 - p.$$

A *De Moivre* rv, say $X$, usually called "Binomial," has pdf defined by

$$\mathrm{P}[X = x] = f(x; n, p) = \tbinom{n}{x} p^x q^{n-x} \qquad \text{for} \quad x = 0, 1, \ldots, n; 0 < p < 1, \; q = 1 - p.$$

A *Poisson* rv, $X$, has pdf defined by

$$\mathrm{P}[X = x] = f(x; \lambda) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad \text{for} \quad x = 0, 1, \cdots; 0 < \lambda.$$

This distribution had its origin as the "law of small numbers" since it can be derived from the De Moivre distribution as the limit when $n \to \infty$, $p \to 0$ with $np = \lambda$, a constant. Consequently, we have

$$\binom{n}{x} p^x q^{n-x} \doteq \frac{e^{-np}(np)^x}{x!} \qquad \text{when } n \text{ is large and } p \text{ is small.}$$

A *Pascal* rv, say $X$, has pdf defined, again letting $q = 1 - p$, by

$$\mathrm{P}[X = x] = f(x; r, p) = \binom{r + x - 1}{x} p^r q^x \quad \text{for} \quad x = 0, 1, \cdots, n; 0 < p < 1,$$

which is the probability that the $r$th success in a sequence of independent Bernoulli trials occurs at the $(r+x)$th trial. It is often called a 'Negative Binomial' since $f(x;r,p) = \binom{-r}{x} p^r (-q)^x$.

A *multivariate De Moivre* rv, say $(X_1, \ldots, X_n)$, has pdf

$$\mathrm{P}[X_1 = x_1, \cdots, X_n = x_n] = n! \prod_{i=1}^{n} \frac{p_i^{x_i}}{x_i!} \quad \text{for} \quad x_i \geq 0;\ 0 < p_i < 1,$$
$$\text{where } \sum_{i=1}^{m} x_i = m,\ \sum_{i=1}^{m} p_i = 1.$$

Another multivariate rv, $(X_1, \ldots, X_n)$, which is continuous, has the *Dirichlet* distribution on the simplex $\mathcal{S} = \{(x_1, x_2, \cdots, x_k) : 0 \leq x_i \leq 1, \sum_1^k x_i \leq 1\}$ when it has density

$$\frac{(n+k)!}{\prod_{j=1}^{k+1}(n_j)!} \left( \prod_{j=1}^{k} x_j^{n_j} \right) \left[ 1 - \sum_1^k x_j \right]^{n_{k+1}},$$

where $n_j \geq 0$ and $n = \sum_1^{k+1} n_j$.

*The Mathematical Expectation*

Let $X \sim F_X$ be a real rv on $\Re$. By the *expectation* of the quantity $\varphi(X)$ we mean the integral

$$\mathrm{E}\varphi(X) = \int_{-\infty}^{\infty} \varphi(x)\, dF_X(x). \tag{1.6}$$

Here we mean the integral is the Stieltjes integral. If this concept is unfamiliar, please read the section in Chapter 15.

By the *moments* of $X$ we mean the quantities

$$\mathrm{E}X^k \quad \text{for } k = 1, 2, \ldots \tag{1.7}$$

for all integral values of $k$ for which the integrals, $\mathrm{E}\{|X|^k\}$ for $k = 1, 2, \ldots$, are finite. Of special interest are the two parameters called the *mean,* often denoted by $\mu$, and the *variance,* often denoted by $\sigma^2$. They are defined, repectively, by

$$\mu := \mathrm{E}[X], \quad \sigma^2 := \mathrm{Var}[X] = \mathrm{E}(X-\mu)^2 = \mathrm{E}[X^2] - \mu^2. \tag{1.8}$$

The first moment, say $\mu$, corresponds to the center of gravity or *centroid,* of the probability "mass" represented by the density. It is a measure of central tendency; the variance corresponds to its moment of inertia about that centroid and so is a measure of dispersion.

*Multivariate Random Variables*

Let $X$, $Y$ be a pair of rv's with *joint distribution* $F_{X,Y}$ and *marginal distributions* $F_X$ and $F_Y$, respectively defined, for all $x, y \in \Re$, by

$$F_{X,Y}(x, y) = \Pr[X \leq x, Y \leq y],$$

with

$$F_X(x) = \Pr[X \leq x] = F_{X,Y}(x, \infty), \quad F_Y(y) = \Pr[Y \leq y] = F_{X,Y}(\infty, y).$$

Moreover the *conditional distribution* of $X$ given that $[Y = y]$ is defined by

$$F_{X|Y}(x|y) = \lim_{h \to 0} \frac{\Pr[X \leq x, y - h < Y \leq y + h]}{F_Y(y + h) - F_Y(y - h)},$$

when it exists. If the conditional distribution $F_{X|Y}$ exists, then we can obtain the cdf of the sum $S = X + Y$ and the product $V = XY$, respectively, as

$$F_S(s) = \int_{-\infty}^{\infty} F_{X|Y}(s - y|y) dF_Y(y) \quad \text{for all } s \in \Re \tag{1.9}$$

and for all $v \in \Re$

$$F_V(v) = \int_0^{\infty} F_{X|Y}(v/y|y) dF_Y(y) + \int_{-\infty}^{0} [1 - F_{X|Y}(v/y|y)] dF_Y(y). \tag{1.10}$$

The *Fourier transform* of the pair $(X, Y)$ is defined, letting $\imath = \sqrt{-1}$, by

$$c_{X,Y}(t, s) := E e^{\imath t X + \imath s Y} \quad \text{for all } (t, s) \in \Re^2.$$

Iff $X$ and $Y$ are *independent*, i.e., $X \perp Y$, does it follow that

$$F_{X,Y}(x, y) = F_X(x) \cdot F_Y(y) \quad \text{for all } (x, y) \in \Re^2,$$

and in this case $F_{X|Y} = F_X$, $F_{Y|X} = F_Y$, and we see $c_{X,Y} = c_X \cdot c_Y$

## 1.2.2. Concepts from Statistics

We now recall some results from the theory of statistics that will be useful in what follows: write $\tilde{\theta}_n = \tilde{\theta}(x_1, \ldots, x_n)$ for any estimator of $\theta \in \Theta$, based on a sample of size $n$.

**Definition 1.** The estimator $\tilde{\theta}_n$ is *unbiased* for $\theta$ iff $\mathrm{E}\tilde{\theta}_n = \theta$ for all $n \in \aleph, \theta \in \Theta$.

**Definition 2.** The estimator $\tilde{\theta}_n$ is *consistent* for $\theta$ iff $\tilde{\theta}_n \xrightarrow{p} \theta$ as $n \to \infty$.

**Definition 3.** The estimator $\tilde{\theta}_n$ is *strongly consistent* for $\theta$ iff $\tilde{\theta}_n \xrightarrow{a.s.} \theta$ as $n \to \infty$.

A *sufficient statistic* for a parameter is one that carries all the information in the sample about that parameter.

**Definition 4.** A statistic $T$ whose value $T(\boldsymbol{x})$ can be computed from data $\boldsymbol{x}$ without knowledge of $\theta$, such that the observed value of $T$ is sufficient to determine the likelihood $\mathcal{L}(\theta|\boldsymbol{x})$, up to a constant of proportionality, is a *sufficient statistic* for $\theta$.

We have the result

**Theorem 1.** *The statistic $T$ is sufficient for $\theta$ iff the density can be factored appropriately*

$$f_{\boldsymbol{X}}(\boldsymbol{x};\theta) = g(\boldsymbol{x}) \cdot h(T(\boldsymbol{x});\theta),$$

*where $g(\cdot)$ is not a function of $\theta$.*

Another formulation serves as an alternate definition.

**Theorem 2.** *The conditional distribution of outcomes $\boldsymbol{X}$, given a sufficient statistic $T$, does not depend upon $\theta$.*

Proof. Let

$$f_T(t;\theta) = \int_{\{\boldsymbol{X}:T(\boldsymbol{X})=t\}} g(\boldsymbol{x})\,d\boldsymbol{x} \cdot h(t;\theta) = \psi(t) \cdot h(t;\theta).$$

Thus, we see

$$f_{X|T}(\boldsymbol{x}|t) = \frac{f_X(\boldsymbol{x};\theta)}{f_T(t;\theta)} = \frac{g(\boldsymbol{x})}{\psi(t)}$$

is independent of the parameter $\theta$. □

Exercise Set 1.A

1. Show that for any rv $X \sim F$, discrete, continuous, or mixed:

$$\mathrm{E}[X] = \int_0^\infty [1 - F(t)]\,dt - \int_{-\infty}^0 F(t)\,dt.$$

2. Much is made in applied mathematics about the nature of the Dirac delta function e.g., even though $\delta(t) = 0$ for $t \neq 0$, it is argued by physical reasoning, involving dimensionality, that $\delta(at+b) \neq \delta(t+\frac{b}{a})$ with $a \neq 0$, since by a change of variable of integration we see, when legitimate,

$$\int \phi(t)\delta(at+b)dt = \int \phi\left(\frac{x-b}{a}\right)\frac{\delta(x)}{|a|}dx = \frac{\phi(-b/a)}{|a|}.$$

So, it is argued $\delta(at+b) = \delta(t+\frac{b}{a})/|a|$. Show that this "anomaly" disappears by using the Heaviside distribution $\varepsilon(t) = \mathrm{I}(t \geq 0)$ in the Stieltjes integral $\int \phi(t)d\,\varepsilon(at+b)$.

3. Using the Schwarz inequality, viz., $\mathrm{E}^2|XY| \le \mathrm{E}|X|^2 \cdot \mathrm{E}|Y|^2$, show that $\ln \mathrm{E}|X|^r$ is a convex function of $r$. Recall that for all $x$, $y$ in an open interval, we have convexity of $\phi$, viz., $\phi(\frac{x+y}{2}) \ge \frac{1}{2}\phi(x) + \frac{1}{2}\phi(y)$, plus continuity, is equivalent with

$$\phi[tx + (1-t)y] \ge t\phi(x) + (1-t)\phi(y) \quad \text{for all } t \in (0,1).$$

4. Show that $\mathrm{E}^{\frac{1}{r}}|X|^r$ is a nondecreasing function of $r$.

5. The characteristic function of the rv $X$ is given by $c(t) = \exp(-t^2/2)$. Find the density function of this random variable.

6. If $X$ is a random variable with characteristic function $c$, then find, in terms of $c$, the characteristic function of the linear transformation $aX + b$ for any constants $a$ and $b$.

7. Show that the cumulants of $X$ and $Y = X + a$, for any constant $a$, are the same. Cumulants are sometimes called the *semi-invariants.*

8. If the bivariate density of (X,Y) is $f(x, y) = e^{-y}$ for $0 < x < y < \infty$, find the density of $X + Y$.

9. If $X$ and $Y$ are dependent, but $F_{X|Y}$ exists, find formulas for the distributions of $U = \max(X, Y)$, $W = \min(X, Y)$, and $Z = X/Y$.

10. Evaluate E[$X$] and Var[$X$] when $X \sim F$ in the two cases where:
(a)

$$F(x) = \Phi[\xi(x/\beta)/\alpha] \quad \text{for } x > 0; \alpha, \beta > 0,$$

with $\Phi$ defined in eqn (2.25) and $\xi$ defined in eqn (2.35) in Chapter 2.
(b)

$$F'(x) = \frac{\beta}{x^2\Phi(\alpha)}\varphi\left(\frac{\beta}{x} - \alpha\right) \quad \text{for } x > 0 : \alpha, \beta > 0.$$

Remember that $\varphi(x)$ is a transcendental function.

11. * Let the pair of rv's $(X, Y)$ have the density $f$ defined on its support by

$$f(x, y) = [1 + xy(x^2 - y^2)]/4 \quad \text{for all } |x| \le 1 \text{ and } |y| \le 1, \text{ and zero elsewhere.}$$

Show that
(a) $f_X \cdot f_Y \ne f$,
(b) $f_{X+Y}(z) = (2 - |z|)/4$ for $0 \le |z| \le 2$, and zero elsewhere.
(c) Find $c_X, c_Y, c_{X+Y}$.
(d) Can it be true that $c_X(t) \cdot c_Y(t) = c_{X+Y}(t)$ for dependent $X$ and $Y$?

12. The "Law of the Unconscious Statistician" refers to the oversight that occurs when $X \sim F$, with $F$ absolutely continuous, its expectation defined by $\mathrm{E}X := \int_{-\infty}^{\infty} xF'(x)\,dx$, when, without proof, it is assumed, for any (measurable) transformation $g(X)$, its expectation is $\mathrm{E}g(X) = \int_{-\infty}^{\infty} g(x)F'(x)\,dx$. Using the definition $\mathrm{E}X := \int_{-\infty}^{\infty} x\,dF(x)$ show that if $Y = g(X)$ for any monotone $g$, that $\mathrm{E}Y = \int_{-\infty}^{\infty} g(x)\,dF(x)$.

Persons who do not understand mathematics are not truly human; they are, at best, a tolerable subspecies that has learned to wear shoes, bathe and not make messes in the house.

Lazarus Long: A character of Robert Heinlein

I tell them that if they will occupy themselves with the study of Mathematics they will find in it the best remedy aginst the lusts of the flesh.

Thomas Mann: in the *Magic Mountain*

I had a feeling once about Mathematics—that I saw it all. Depth beyond depth was revealed to me—the Byss and the Abyss. I saw—as one might see the transit of Venus or even the Lord Mayor's show—a quantity passing through infinity and changing its sign from plus to minus. I saw exactly how it happened and why the tergiversation was inevitable—but it was just after dinner and I let it all go.

Winston S. Churchill

The supposed advantage of having all those humanities courses taught on campus is more than counterbalanced by the general dopiness of the people who study them.

Richard P. Feynman, on why he left Cornell to go to Cal. Tech.

CHAPTER 2

# Elements of Reliability

## 2.1. Properties of Life Distributions

Reliability studies are concerned with an assessment of the rate of wear, deterioration, or accumulating damage to a structure or system and the entailed distribution of useful service life, i.e., until it can no longer, safely or profitably, perform its operational mission. Since damage during service occurs in a known manner but at unpredictable times, the waiting time until failure occurs is also a random variable which must be nonnegative. Let $T \geq 0$ be an rv denoting *life length;* then we write $T \sim F$ when $F$ is its cdf. NB that $F$ has support on $[0, \infty)$, i.e., $F(t) = 0$ for $t < 0$. The corresponding *survival distribution,* denoted by *sdf*, is defined by

$$\bar{F}(t) := 1 - F(t) = \Pr[T > t] \quad \text{for } t \geq 0. \tag{2.1}$$

It is also called, in many applications, the *reliability* function.

**Remark.** *We know the following statements:*

*If T is* discrete, *then $\bar{F}$ is a decreasing step function.*
*If T is* continuous, *then $\bar{F}$ is continuously decreasing and $F' := f$.*
*If T is* mixed, *then $\bar{F}$ is decreasing with at least one saltus.*

*NB the convention that an rv being 'continuous' means its cdf is absolutely continuous.*

Any function, say $\bar{G}$, is a reliability function for some life-length variate iff it has the following three properties:

$(i)$ $\bar{G}(0) = \delta$ for some $0 < \delta \leq 1$, $(ii)$ $t_1 \leq t_2$ implies $\bar{G}(t_1) \geq \bar{G}(t_2)$,
$(iii)$ $\lim_{t\to\infty} \bar{G}(t) = 0$.

Consider the probability of failure for a life-length $T \sim F$ during the time interval $(t, t + x]$ :

$$\Pr[t < T \leq t + x] = F(t + x) - F(t), \tag{2.2}$$

and the conditional probability of failure, given survival to time $t > 0$:

$$F[t + x|t] := \Pr[t < T \leq t + x|T > t] = \frac{F(t + x) - F(t)}{1 - F(t)}. \tag{2.3}$$

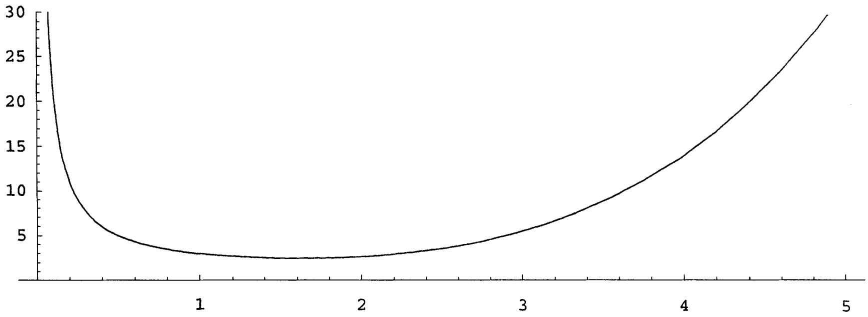


Figure 2.1. The bathtub-shaped hazard rate.

The *hazard rate*, label it $h$ when it exists, is defined as

$$h(t) := \lim_{x \downarrow 0} \frac{F[t+x|t]}{x} = \frac{F'(t)}{1-F(t)}; \tag{2.4}$$

this function is also called the *failure rate*, and in actuarial usage it is called *the force of mortality*.

The mortality tables of both humans and animals exhibit a characteristic behavior; the force of mortality initially decreases and then remains virtually constant for a time and finally increases. See Figure 2.1. Failure data for machines often exhibit the same behavior. This bathtub shape is explained by the operation of three independent failure modes, namely, (i) manufacturing or assembly error causing early failure, (ii) failure due to accidents while in service, and (iii) cumulative damage (wearout or fatigue) failures, which are manifested during late service life.

Hazard rates are sometimes easier to determine from physical considerations than are densities. In fact, the Gompertz–Makeham law of human mortality [38, 1825; 67, 1860], was among the first statistical models ever applied, predating statistical inference. For a given cdf $F$ with density $f$ we have the corresponding hazard rate, say $h$, given by the relation of eqn (2.4) and its integral, the *hazard function* (call it $H$), given by:

$$H(x) = \int_0^x h(t)\,dt = -\ln[1-F(x)] \quad \text{for } x > 0. \tag{2.5}$$

Note that we can define $H := -\ln[1-F]$, which always exists.

Thus three properties determine a hazard function for a life variate, viz., (i) $H(0) = 0$, (ii) $H$ is nondecreasing, and (iii) $H(\infty) = \infty$. Moreover, we can write the sdf as

$$\bar{F}(t) = e^{-H(t)} = \exp\left[-\int_0^t h(x)\,dx\right] \quad \text{for all } t > 0. \tag{2.6}$$

Denote the conditional reliability of a unit of age $t$ during a time $x$ by

$$\bar{F}(x|t) := \frac{\bar{F}(t+x)}{\bar{F}(t)} \qquad \text{if } \bar{F}(t) > 0.$$

Thus, we make the important

**Definition.** A cdf $F$ is IHR, which stands for *increasing hazard rate*, iff $F$ satisfies

$$\bar{F}(x|t) \quad \text{is decreasing in } 0 < t < \infty \text{ for each } x \geq 0. \tag{2.7}$$

NB that a distribution can be IHR without $H'$ existing!

It follows, when a density $f(t)$ exists for an IHR cdf, that

$$h(t) = \lim_{x \downarrow 0} \frac{\left[1 - \bar{F}(x|t)\right]}{x} \qquad \text{is increasing in } t \geq 0.$$

Conversely, when the hazard rate $h(t)$ is increasing, then $\bar{F}(x|t)$ is decreasing in $t \geq 0$ for each $x \geq 0$. Thus, we see that when a density exists, the IHR definition is equivalent with the hazard rate being an increasing function. We also define a DHR distribution as one for which the words "increasing" and "decreasing" are interchanged in the preceeding definition.

There is another concept that has intuitive meaning and can be used in modeling life distributions in reliability. For a given $0 < T \sim F$ which is right-continuous, we consider the *mean residual-life function*, say $m$, defined by

$$m(t) := \mathrm{E}[X - t | X > t], \quad \text{when } \bar{F}(t) > 0, \quad \text{and} \quad = 0 \quad \text{otherwise}.$$

NB that we can write for any $t$ such that $\bar{F}(t) > 0$,

$$m(t) = \int_0^\infty \bar{F}(x|t)dx = \int_t^\infty \frac{\bar{F}(x)}{\bar{F}(t)}dx.$$

If $F' = f$ exists, then we also have

$$m(t) + t = \frac{1}{\bar{F}(t)} \int_t^\infty u f(u)\, du, \tag{2.8}$$

from which the following relationship can be deduced:

$$m'(t) + 1 = m(t) \cdot h(t). \tag{2.9}$$

Let $\mu = \mathrm{E}X$. So if $F(0) = 0$, then $m(0) = \mu$; and if $F(0) > 0$, then $m(0) = \mu / \bar{F}(0)$.

Let us presume that $F(0) = 0$; then we have the representation by setting $F^{-1}(1) = \sup\{t > 0 | F(t) < 1\}$,

$$\bar{F}(t) = \begin{cases} \frac{m(0)}{m(t)} \exp\left\{-\int_0^t \frac{1}{m(u)} du\right\} & \text{for } 0 \leq t < F^{-1}(1), \\ 0 & \text{for } t \geq F^{-1}(1). \end{cases} \tag{2.10}$$

Thus, we have the following classifications:

**Definition.** A distribution $F$ has a *decreasing mean residual life* (DMRL) iff its mean residual-life function is a decreasing function.

**Definition.** A life length rv $T \sim F$ is *new better than used in expectation* (NBUE) iff $m(0) \geq m(t)$ for all $t > 0$.

Exercise Set 2.A

1. The hazard rate for human mortality suggested by Gompertz was $h(t) = \alpha$ and the "Gompertz–Makeham law" was $h(t) = \alpha + \lambda e^{\gamma t}$ with $\alpha, \lambda, \gamma > 0$, for $t > 0$. Why do you think the GM law was more successful in applications?
2. The "logistic distribution," in standard form, is given by

$$f(x) = \frac{e^x}{(1+e^x)^2} \quad \text{for} \quad -\infty < x < \infty,$$

   Add scale and location parameters and discuss its behavior as a model for life length.
3. A component as produced, has a life length, say $T$, which is an rv with hazard rate $h$. Each component is subjected to a *burn-in* of length $\tau$ and passed iff it did not fail during burn-in. Express the hazard rate of the passed component, say $T_\tau$, in terms of $h$.
4. Let the cost for each burn-in test be \$$c$, when the component passes, and \$$C$, when the component fails. The gain is \$$D$ per unit of increase of expected service-life, which is obtained by testing. If the hazard rate $h$ is bathtub-shaped, derive a formula for an optimum burn-in period and the distribution of life obtained.
5. Show that the expectation of the life-length rv $T$ is its accumulated reliability, i.e.,

$$\mu = E(T) = \int_0^\infty \bar{F}(t)\,dt, \quad \text{if} \quad \lim_{t\to\infty} \bar{F}(t)/h(t) = 0.$$

   Hint: Evaluate, via L'Hopital's rule, conditions for $\lim_{y\to\infty} y\bar{F}(y) = 0$. Is this a mathematical or practical consideration?
6. If $\nu_r(T) = [E(T)^r]^{1/r}$, for some life-length variate $T$, exists for all $r > 0$, what would be the values of $\lim_{r\to 0} \nu_r(T)$ and $\lim_{r\to\infty} \nu_r(T)$ ?
7. Define $W(t) = \int_t^\infty \bar{F}(x)\,dx$ for $t > 0$, note that $\mu = W(0)$, and show that

$$E[T^2] = 2\int_0^\infty W(t)\,dt, \quad \text{provided that} \quad \lim_{t\to\infty} t^2\bar{F}(t) = \lim_{t\to\infty} tW(t) = 0.$$

   NB that $W(t)/\mu$ is itself a survival distribution.
8. Check the conditions of Exercise 3 for the case when

$$h(t) = \frac{1}{1+at} \quad \text{for } t > 0; \text{ for some } a > 0.$$

# 2.2. Useful Parametric Life Distributions

## 2.2.1. The Epstein (Exponential) Distribution

One of the simplest, yet most useful distributions, was studied in detail by Benjamin Epstein in the 1940s. It is often called the *exponential,* or *negative exponential,* distribution. However, during the 19th century "exponential" was the name given to what we now call the Gaussian (Normal) distribution. In statistics we have "the exponential family" and to avoid confusion we utilize his proper name. We write $T \sim \text{Ep}(\lambda)$ whenever the hazard rate of the rv $T$ is constant, namely,

$$h(t) = \lambda \quad \text{for all } t > 0;\ \lambda > 0.$$

The density and survival distribution are then given, respectively, by

$$f(t) = \lambda e^{-t\lambda}, \quad \bar{F}(t) = e^{-t\lambda} \quad \text{for } t > 0.$$

An alternate parameterization is in terms of the mean $\mu = 1/\lambda$. In this case the notation often used is $T \sim \text{Exp}(\mu)$; so it follows that $E(T) = \mu, \quad \text{Var}(T) = \mu^2$.

The "two-parameter exponential" distribution results when the Epstein origin is changed. It is denoted here as the shifted-Epstein, $T \sim \text{Shep}(\lambda, \nu)$, with the density and sdf given by

$$f(t) = \lambda e^{-\lambda(t-\nu)^+}, \qquad \bar{F}(t) = e^{-\lambda(t-\nu)^+} \quad \text{for } t \in \Re; \lambda > 0, \nu \in \Re. \tag{2.11}$$

There is another distribution related to the exponential, called the two-sided exponential or the *Laplace distribution.* This is denoted by $T \sim \text{Lap}(\lambda, \mu)$ whenever

$$f(t) = \frac{\lambda}{2} e^{-\lambda|t-\mu|} \quad \text{for } t \in \Re; \lambda > 0, \mu \in \Re. \tag{2.12}$$

Exercise Set 2.B

1. Let $T_1, \ldots, T_n$ be iid $\text{Ep}(\eta)$ and set $X_n = \max_{i=1}^n T_i$. Find the distribution of $X_n$ and show its mean and variance satisfy

$$\text{E}X_n = \frac{1}{\eta} \sum_{j=1}^{n} \frac{1}{j}, \qquad \text{Var}X_n = \frac{1}{\eta^2} \sum_{j=1}^{n} \frac{1}{j^2}. \qquad \text{What happens as as } n \to \infty.?$$

2. If $Y_i \sim \text{Ep}(i\eta)$ for $i = 1, \ldots, n$ are independent, what are the mean and variance of $\sum_{i=1}^n Y_i$?

3. Suppose a device consists of $m$ components each of which has an Epstein life with hazard rate proportional to load. If each surviving component shares the total imposed load equally, what is the distribution of life of this device if it fails whenever the critical $k$th element fails for some fixed $1 \le k \le m$.

### 2.2.2. The Gamma Distribution

We say $T$ has a Gamma distribution and write $T \sim \mathcal{G}\text{am}(\alpha, \beta)$ whenever its cdf transform and density are given, respectively, by

$$F^{\dagger}(t) = (1 - t\beta)^{-\alpha}, \quad F'(t) = \frac{t^{\alpha-1}e^{-t/\beta}}{\beta^{\alpha}\Gamma(\alpha)} \quad \text{for } t > 0;\ \alpha, \beta > 0. \tag{2.13}$$

Here $\Gamma(\alpha)$ denotes the *gamma function*, an Eulerian integral of the second kind, which serves as the normalizing factor in the density. There are two special cases: Because of its application in early studies in telephone traffic by Erlang, a Bell Telephone employee, the distribution $\mathcal{G}\text{am}(k, \beta)$ when $k \in \aleph = \{1, 2, \ldots\}$ is called the *Erlang-k* distribution. Also because of its early utilization in classical statistics the distribution $\mathcal{G}\text{am}(n/2, 2)$ is called the *Chi-square distribution*, and this is denoted as

$$\chi_n^2 \sim \mathcal{G}\text{am}(n/2, 2).$$

For $n \geq 30$ the chi-square distribution may be well approximated by the Normal distribution, to be identified subsequently.

We have the reliability of $X \sim \mathcal{G}\text{am}(\alpha, \beta)$ given by

$$\bar{G}(t) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} \int_t^{\infty} u^{\alpha-1}e^{-u/\beta}\, du \tag{2.14}$$

$$= e^{-t/\beta} \sum_{j=0}^{\alpha-1} \frac{(t/\beta)^j}{j!} \quad \text{when } \alpha \in \aleph = \{1, 2, \cdots\}. \tag{2.15}$$

The corresponding hazard rate $h$ is best studied in terms of its reciprocal:

$$\frac{1}{h(t)} = \int_t^{\infty} (x/t)^{\alpha-1}e^{-(x-t)/\beta}dx = \int_0^{\infty} (1 + \frac{u}{t})^{\alpha-1}e^{-u/\beta}du \tag{2.16}$$

$$= \frac{\beta}{\alpha} \sum_{j=1}^{\alpha} (\alpha^{j\downarrow})(\beta/t)^{j-1} \quad \text{whenever } \alpha \in \aleph. \tag{2.17}$$

Note that again we have made use of the factorial power. See Figure 2.2.

### 2.2.3. The Pareto Distribution

A distribution originally derived by an Italian economist, Vilfredo Pareto (1848–1923), to describe the distribution of income within a population has proved to be of considerable interest in other areas of application, including reliability. We write $T \sim \mathcal{P}\text{ar}(\alpha, \beta)$ whenever the survival distribution is given by:

$$\bar{F}(t) = e^{-\alpha \ln(1+\frac{t}{\beta})} = \left(1 + \frac{t}{\beta}\right)^{-\alpha} \quad \text{for all } t > 0; \alpha, \beta > 0, \tag{2.18}$$

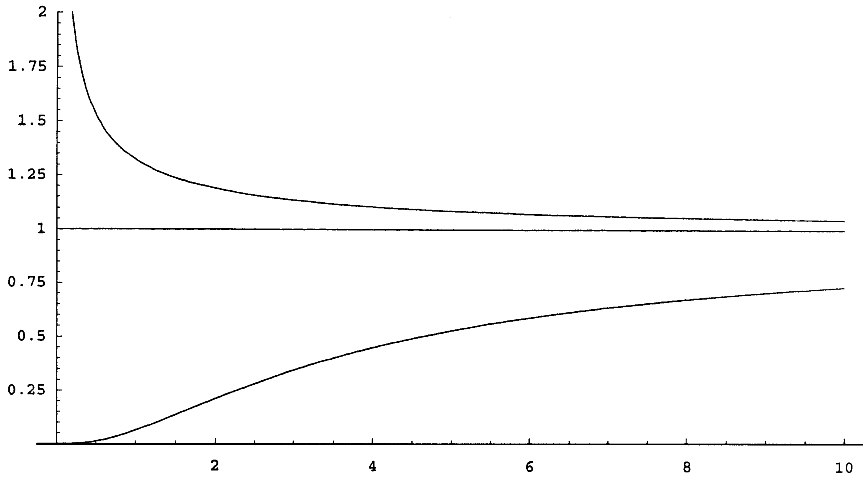


Figure 2.2. Gamma-Hazard-Rate with $\beta = 1$, for $\alpha = 1/2, \alpha = 1, \alpha = 4$.

and the corresponding hazard rate is given by

$$h(t) = \frac{\alpha}{\beta + t} \quad \text{for } t > 0. \tag{2.19}$$

## 2.2.4. The Gaussian or Normal Distribution

The rv $X$ has a Gaussian or Normal distribution with location and scale parameters $\mu$ and $\sigma$, respectively, when the density is given by

$$f_X(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{for } x \in \Re. \tag{2.20}$$

We write $X \sim \mathcal{N}(\mu, \sigma^2)$ for any $\sigma > 0$ and $\mu \in \Re$. In *standardized form* one writes $Z \sim \Phi$ iff $Z \sim \mathcal{N}(0, 1)$. Here we define the standard distribution and density, respectively, by

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt, \qquad \Phi'(x) := \varphi(x) \ \text{ for } x \in \Re. \tag{2.21}$$

Thus, corresponding to the density given in eqn (2.20), $X \sim F_X$ would imply that $E(X) = \mu$ and $E(X - \mu)^2 = \sigma^2$ with the cdf and sdf given, respectively, for any $t \in \Re$ by

$$F_X(t) = \Phi\left(\frac{t-\mu}{\sigma}\right), \quad \text{with} \quad \bar{F}_X(t) = \Phi\left(\frac{\mu - t}{\sigma}\right). \tag{2.22}$$

A classical asymptotic expansion called *Mill's ratio*, see p. 932 of Abramowitz and Stegun [2] , which we here denote by $m(t)$ is

$$m(t) := \frac{t\bar{\Phi}(t)}{\varphi(t)} \asymp 1 - \frac{1}{t^2} + \frac{1\cdot 3}{t^4} - \cdots + \frac{(-1)^n}{t^{2n}}\frac{(2n!)}{2(n!)} + \rho_n(t) \quad \text{for } t \gg 0, \tag{2.23}$$

where the error is always less than the absolute remainder; the remainder is

$$\rho_n(t) = (-1)^n \prod_{j=1}^{n}(2j-1) \times t \int_t^\infty \varphi(x)x^{-2n-2}\,dx. \tag{2.24}$$

The reciprocal of the Gaussian hazard rate, say $h$, is given by

$$\frac{1}{h(x)} = \frac{\bar{\Phi}(x)}{\varphi(x)} = e^{x^2/2}\int_x^\infty e^{-u^2/2}\,du \asymp \frac{1}{x} - \frac{1}{x^3} + \frac{1\cdot 3}{x^5} - \cdots \quad \text{for } x \gg 0. \tag{2.25}$$

Computation of $h(x)$ for $x < 0$ is aided by using the identity

$$m(x) = \frac{x^-}{\varphi(x)} + m(|x|), \quad \text{where } x^- = \min(x, 0). \tag{2.26}$$

## 2.2.5. Transformations to Normality

The Gaussian, or normal, distribution with its unique and useful statistical properties is the distribution assumed most frequently in virtually every field. However, the support of any normal density is $\Re$ and so the probability of a negative value is always positive. Since many physical variables are nonnegative, if one adopts a Gaussian model of, say, life-length which implies there is a nonnegligible probability of being negative, a nonsensical result may occur. Moreover, it is not surprising that there are many practical problems that are "solved" by merely introducing a simple transformation of the data to normality (the logarithm is a popular choice). After finding the answer one transforms it back to the original sample space. Such procedures must be used with caution (avoided?), when they cannot be shown to give approximately correct answers.

### *The Truncated Normal Distribution*

If $X \sim \mathcal{N}(\mu, \sigma^2)$ then the conditional rv $T = [X|X > 0]$ has the $\mathcal{TN}(\mu, \sigma^2)$ distribution with cdf given by

$$F(t) = \frac{\Phi\left(\frac{t-\mu}{\sigma}\right) - \Phi\left(\frac{-\mu}{\sigma}\right)}{\Phi(\frac{\mu}{\sigma})} \quad \text{for } t > 0. \tag{2.27}$$

Another related distribution, that has proven to be of practical utility in tool-life studies, see [51], is the *Alpha* distribution with density given by

$$f(x) = \frac{\beta}{x^2\Phi(\alpha)}\varphi\left(\frac{\beta}{x} - \alpha\right) \quad \text{for } 0 < x < \infty. \tag{2.28}$$

*The Log-Normal Distribution*

This distribution is sometimes called the *Law of Galton*, because of its utilization in the pioneering efforts of Francis Galton (1822–1911) to construct an empirical and conceptual methodology in statistics that was based on probability. An rv $T > 0$ has the *log-normal* distribution whenever its logarithm has a normal distribution, i.e.,

$$T \sim \mathcal{LN}(\mu, \sigma^2) \quad \text{iff} \quad (\ln T) \sim \mathcal{N}(\mu, \sigma^2).$$

Thus, one finds the cdf is given by

$$F_T(t) = \Phi\left(\frac{\ln t - \mu}{\sigma}\right) \quad \text{for } t > 0, \tag{2.29}$$

and we find the mean and variance, respectively, to be

$$\mathrm{E}(T) = e^{\mu + \frac{\sigma^2}{2}} \quad \text{and} \quad \mathrm{Var}(T) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1). \tag{2.30}$$

The assumption that a variate has the log-normal distribution can be easily checked visually by merely taking the logarithm of the observations in a sample and seeing if the cumulative plot of the ordered observations departs significantly from a straight line when plotted on normal-probability paper.

> Every data-set they have is considered Normal unless its scatter covers several orders of magnitude and then they assume it is Log-normal.
>
> Bob McCarty - on engineering practice

Exercise Set 2.C

1. Show if $0 < \alpha \leq 1$ then the gamma distribution is DHR, but for $\alpha \geq 1$ is IHR.
2. Find the density and first two moments of the Weibull distribution.
3. The rv $X$ has the hazard rate $h(x) = (a - x)^{-1}$ for $0 < x < a$; what is the distribution of $X$?
4. Show that if $T \sim \mathrm{Par}(\alpha, \beta)$, that

$$\mathrm{E}(T) = \frac{\beta}{\alpha - 1} \quad \text{if } \alpha > 1: \qquad \mathrm{Var}(T) = \frac{\alpha\beta^2}{(\alpha-1)^2(\alpha-2)} \quad \text{if } \alpha > 2.$$

5. Show that if we have a gamma mixture of Epstein sdfs, namely

$$\bar{F}(t) = \int_0^\infty e^{-\lambda t} dG_\Lambda(\lambda),$$