DAC-62180

FINAL REPORT

STUDY OF NONPARAMETRIC TECHNIQUES FOR ESTIMATING RELIABILITY AND OTHER LIFE QUALITY PARAMETERS

FEBRUARY 1968

By
V.K. Murthy

N68 18499

Prepared under Contract No. NASw-1367
by Douglas Aircraft Company
Missile and Space Systems Division
Santa Monica, California
for
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

FOREWORD

This report is submitted to the National Aeronautics and Space Administration by the Douglas Aircraft Company under Contract No. NASw-1367. It was prepared by Systems Development Corporation for Douglas under Subcontract No. L. S. -1407-A.

STUDY OF NONPARAMETRIC TECHNIQUES FOR ESTIMATING RELIABILITY AND OTHER LIFE QUALITY PARAMETERS

by

V.K. Murthy

CHAPTER I

SUMMARY

This study was undertaken to carry out a comprehensive investigation of procedures for estimating probability distributions of life lengths, as well as the corresponding probability densities and hazard functions, in situations where little or no information is available on the family of the underlying failure laws. In carrying out this study, nonparametric techniques were developed for obtaining estimates and confidence bands for various kinds of samples of life lengths--i.e., for random, censored, and truncated samples. Large-sample techniques of this kind are derived and discussed in Chapters II-IV. A procedure for optimizing the methods of estimation described in Chapter II is recounted in Chapter VIII. Discontinuities of the life distribution function correspond to times at which the devices considered are exposed to increased "instantaneous hostility." A statistical test is proposed in Chapter V which makes it possible to determine whether such moments of increased instantaneous hostility are present.

Exact small-sample probability distributions as well as well as the corresponding asymptotic (large-sample) distributions to selected Rényi-type statistics are presented in Chapter VI. Extensive numerical tabulations of these distributions have been performed, and results are summarized in compact tables ready for practical use.

A family of life distributions with a number of interesting properties is studied in Chapter VII. Applications of this family of failure laws are postponed to a future study.

CHAPTER II

ESTIMATION AND CONFIDENCE BANDS FOR RANDOM SAMPLES

Concepts and Notations

The concepts and definitions introduced pertain to the performance of a component. It is assumed that failure of the component is well defined and that the time from inception to failure is observable and measured. Evidently these concepts and definitions apply to any system of components in which failure of the system is well defined and the time to failure is measured.

To establish a base for the principal results derived in a later section, the remainder of this section will be devoted to a derivation, accompanied with the explicit definition of key notions and terminology of a component's distribution function of time to failure and the associated hazard function. Let the nonnegative real number T denote the observed time to failure of a component. Other conditions remaining the same, if identical experiments are conducted to determine the respective times to failure on identical specimens of a component type, the actual observed times to failure need not be the same, even for identical experiments. In this sense, the observed time to failure T is a random (stochastic or chance) variable. For any $t > 0$, the event $T \leq t$ is the event that the observed time to failure is less than or equal to a designated time instant t, or, equivalently, the event that the item has failed by time t. Let

$$F(t) = P(T \leq t) \ .$$

F(t), which is the probability of the event that the item has failed by time t, is the distribution function of the random variable T. The complementary event $T > t$, $t \geq 0$, is the event that the observed time to failure is greater than t or, equivalently, the event that the item survived time instant t. The probability of this event is denoted by $R(t)$, where

$$R(t) = P(T > t) = 1 - F(t) \ . \tag{2}$$

$R(t)$ is customarily called reliability of the item at time t. Assuming that the singular part of $F(t)$ is identically zero, $F(t)$ can be uniquely decomposed into (for example, Cramer [4, pp. 52-53])

$$F(t) = F_1(t) + F_2(t) \ , \tag{3}$$

where $F_1(t)$ is an everywhere absolutely continuous function and $F_2(t)$ is a pure step function with steps of magnitude, say, S_ν, $S_\nu > 0$, at the points $t = t_\nu$, $\nu = 0, 1, 2, \ldots$, and both $F_1(t)$ and $F_2(t)$ are nondecreasing and uniquely determined.

Let

$$dF_1(t) = f(t)\ dt \ , \tag{4}$$

where $f(t)$, which is the derivative of the absolutely continuous part of the distribution function $F(t)$, is called the probability density function, and the symbol, dt, refers to an infinitesimal time increment.

Let

$$\begin{aligned} A &: t < T \leq t + dt \\ B &: T > t \qquad . \end{aligned} \tag{5}$$

A stands for the event that the component fails during an instantaneous neighborhood of the time instant t, while B stands for the event that the component survived time t. Clearly

$$P(A/B) = \frac{P(AB)}{P(B)} = \frac{P(A)}{P(B)}, \tag{6}$$

because the event A is contained in the event B; symbolically

$$A \subset B .$$

Now, at a point of continuity t of the distribution F(t)

$$\left.\begin{aligned} P(A) &= f(t)\,dt \\ \text{and} \quad P(B) &= 1-F(t) \\ \text{for all } t \geq 0. & \end{aligned}\right\} \tag{7}$$

The left-hand side of equation (6) denotes the conditional probability that the item, having survived time t, fails between t and t + dt, t being a point of continuity of the distribution F(t). Let z(t)dt denote the left-hand side of equation (6), where, in view of equation (7),

$$z(t)dt = \frac{f(t)dt}{1-F(t)} \tag{8}$$

at a point of continuity t of the underlying law of failures F(t).
The function z(t) is called the hazard rate or the conditional failure rate, or sometimes, simply the failure rate.

A Basic Decomposition of the Cumulative Hazard Function

The function, Z(t), is called the cumulative hazard function, where

$$Z(t) = -\log\left(1 - F(t)\right) \tag{9}$$

and $F(t)$ is the underlying distribution function of time to failure given by equation (3). Assuming the singular part of the distribution to be identically zero, $F(t)$ has the representation given in (3), namely,

$$F(t) = F_1(t) + F_2(t) \ , \tag{10}$$

where $F_1(t)$ is the absolutely continuous part, and $F_2(t)$ is the pure discrete part (step function) of the distribution $F(t)$. Since $F(t)$ is a distribution function,

$$\left.\begin{aligned} &F(0) = F_1(0) + F_2(0) = 0, \\ &\text{which implies} \\ &F_1(0) = 0, \ F_2(0) = 0 \end{aligned}\right\} \tag{11}$$

since F_1 and F_2 are nonnegative. Also

$$F(\infty) = F_1(\infty) + F_2(\infty) = 1. \tag{12}$$

Substituting equations (10) and (12) in (9), one obtains

$$\begin{aligned} Z(t) &= -\log\Big(F_1(\infty) + F_2(\infty) - F_1(t) - F_2(t)\Big) \\ &= -\log\left(\Big(F_1(\infty) - F_1(t)\Big)\left(1 + \frac{F_2(\infty) - F_2(t)}{F_1(\infty) - F_1(t)}\right)\right) \\ &= -\log\Big(F_1(\infty) - F_1(t)\Big) - \log\left(1 + \frac{F_2(\infty) - F_2(t)}{F_1(\infty) - F_1(t)}\right) , \end{aligned} \tag{13}$$

for $t < \tau$, where τ is the smallest integer such that $F_1(\infty) = F_1(t)$ for $t \geq \tau$; i.e., F is strictly discrete for $t \geq \tau$. This τ can be 0.

For $t \geq \tau$, we have

$$Z(t) = -\log (F_2(\infty) - F_2(t)).$$

To decompose $Z(t)$ into its absolutely continuous part and pure step function or discrete part, consider for $t < \tau$:

$$dZ(t) = \frac{f(t)\,dt}{F_1(\infty) - F_1(t)} - d\left(\log\left(1 + \frac{F_2(\infty) - F_2(t)}{F_1(\infty) - F_1(t)}\right)\right) \tag{14}$$

where $f(t) = \frac{d}{dt} F_1(t)$.

At this stage, it is interesting to notice that while

$$\int_0^t \frac{f(\tau)d\tau}{F_1(\infty) - F_1(\tau)}$$

is the cumulative hazard function associated with $F_1(t)$, the absolutely continuous part of the distribution function $F(t)$, $\frac{f(t)}{F_1(\infty) - F_1(t)}$

being the hazard rate corresponding to $F_1(t)$, the function

$$\int_0^t \frac{f(\tau)d\tau}{F_1(\infty) - F_1(\tau)}$$

does not constitute the absolutely continuous part of the cumulative hazard function $Z(t)$ associated with the entire distribution $F(t)$. The reason for this is that

$$d\left(\log\left(1 + \frac{F_2(\infty) - F_2(t)}{F_1(\infty) - F_1(t)}\right)\right)$$

contains an absolutely continuous component which must be added to

$$\frac{f(t)dt}{F_1(\infty) - F_1(t)}$$

to yield the absolutely continuous part of $Z(t)$, the cumulative hazard function corresponding to $F(t)$.

Now

$$d\left(\log\left(1 + \frac{F_2(\infty) - F_2(t)}{F_1(\infty) - F_1(t)}\right)\right) = \left(\frac{F_1(\infty) - F_1(t)}{1 - F(t)}\right)$$

$$\left(\frac{\left(F_2(\infty) - F_2(t)\right) f(t)dt - \left(F_1(\infty) - F_1(t)\right) dF_2(t)}{\left(F_1(\infty) - F_1(t)\right)^2}\right)$$

$$= \frac{f(t)dt}{1 - F(t)} \frac{F_2(\infty) - F_2(t)}{F_1(\infty) - F_1(t)} - \frac{dF_2(t)}{1 - F(t)} \qquad (15)$$

Substituting (15) in (14), it is deduced that

$$dZ(t) = \frac{f(t)dt}{F_1(\infty) - F_1(t)}\left(1 - \frac{F_2(\infty) - F_2(t)}{1 - F(t)}\right) + \frac{dF_2(t)}{1 - F(t)}$$

$$= \frac{f(t)dt}{1 - F(t)} + \frac{dF_2(t)}{1 - F(t)} \quad . \tag{16}$$

Hence,

$$Z(t) = Z_1(t) + Z_2(t), \tag{17}$$

where

$$Z_1(t) = \int_o^t \frac{f(\tau)d\tau}{1 - F(\tau)} \tag{18}$$

and

$$Z_2(t) = \int_o^t \frac{dF_2(\tau)}{1 - F(\tau)} = \sum_{j=o}^{\infty} \frac{U(t - t_j)S_{t_j}}{1 - F(t_j)} \ , \tag{19}$$

$U(x)$ being the Heaviside Unit Function with

$$U(x) \begin{cases} = 1 \text{ for } x \geq o \\ = o, \text{ otherwise.} \end{cases}$$

Clearly, $Z_1(t)$ and $Z_2(t)$ are, respectively, the absolutely continuous part and pure discrete part of the cumulative hazard function $Z(t)$ associated with the distribution function of time to failure $F(t)$.

Since $Z(t)$ exists and is equal to $-\log(1 - F(t))$, the existence of either $Z_1(t)$ or $Z_2(t)$ must be proved before the representation (17) is a valid decomposition of $Z(t)$.

Clearly, $Z(t) = -\log(1 - F(t))$ is non-decreasing, $Z(0) = 0$, $Z(+\infty) = +\infty$, and $Z(t) = +\infty$ for $t \to F(t) = 1$.

To this end, consider

$$Z_1(t) = \int_0^t \frac{f(\tau)d\tau}{1 - F(\tau)}$$

Write, for $t < \tau$,

$$\frac{f(t)}{1 - F(t)} = \left(\frac{f(t)}{F_1(\infty) - F_1(t)}\right)\left(\frac{F_1(\infty) - F_1(t)}{F_1(\infty) - F_1(t) + F_2(\infty) - F_2(t)}\right) .$$

Now

$$\frac{f(t)}{F_1(\infty) - F_1(t)}$$

is absolutely integrable, since it is positive, and for $t < \tau$ (see p.6),

$$\int_0^t \frac{f(\tau)}{F_1(\infty) - F_1(\tau)} d\tau = -\log\left(F_1(\infty) - F_1(t)\right) ,$$

and

$$\frac{F_1(\infty) - F_1(t)}{F_1(\infty) - F_1(t) + F_2(\infty) - F_2(t)} \leq 1 .$$

Therefore,

$$\frac{f(t)}{1 - F(t)} \leq \frac{f(t)}{F_1(\infty) - F_1(t)} ,$$

which is absolutely integrable.

Hence,

$$\frac{f(t)}{1 - F(t)}$$

is also integrable and therefore $Z_1(t)$ exists for $t < \tau$.

For $t \geq \tau$: $f(t) \equiv 0$, hence $Z_1(t) = \int_0^t \frac{f(t)}{1 - F(t)}\, dt = Z_1(\tau) = Z_1(+\infty)$.

The proof of the basic decomposition (17) of $Z(t)$, the cumulative hazard function of $F(t)$, is thus complete.

Estimation of the Density of the Underlying Law of Failures at a Point of Continuity

Let $F(t)$, the distribution function of time to failure T, be given by

$$F(t) = F_1(t) + F_2(t), \tag{20}$$

where $F_1(t)$ is the absolutely continuous part, and $F_2(t)$ is the pure step function with steps of magnitude, S_ν at the points $t = t_\nu$, $\nu = 0, 1, 2, \ldots \infty$. Now let the random variable T denote the observed time to failure of an item. Let $T_1, T_2, \ldots T_n$ denote the actual observed times to failure of n identical items put to a life testing experiment. In other words, $T_1, T_2, \ldots T_n$ are the observed values of n independently identically distributed random variables with

$$P(T_i \leq t) = F(t),\ i = 1, 2, \ldots n.$$

Since $F_1(t)$ is absolutely continuous,

$$F_1(t) = \int_0^t f(\tau)\, d\tau \tag{21}$$

where $f(t)$ is the probability density function at a point of continuity t of the distribution function $F(t)$.

Let

$$F_n(t) = \frac{1}{n}\,[\text{number of observations} \le t \text{ among } T_1, T_2, \dots T_n]\ . \tag{22}$$

Clearly, the random variable $F_n(t)$, which is the empirical distribution function based on the observed sample, is a binomially distributed random variable with expectation and variance given by

$$E\left(F_n(t)\right) = F(t)$$

$$\operatorname{Var}\left(F_n(t)\right) = \frac{F(t)\left(1 - F(t)\right)}{n}. \tag{23}$$

A weight function $K(\omega)$ is called a window if it satisfies the following conditions:

$$K(\omega) \geq 0\ ,$$

$$K(\omega) = K(-\omega)\ ,$$

$$\lim_{|\omega| \to \infty} \omega K(\omega) = 0\ ,$$

$$\int_{-\infty}^{\infty} K(\omega)\, d\omega = 1\ . \tag{24}$$

Propose

$$f_n(t_o) = \int_o^\infty B_n K\left(B_n(t - t_o)\right) dF_n(t)$$

$$= \frac{B_n}{n} \sum_{j=1}^{n} K\left(B_n(T_j - t_o)\right) , \qquad (25)$$

as an estimate of the density $f(t_o)$ of the underlying law of failures at a point of continuity t_o of the distribution function of time to failure $F(t)$, where $\{B_n\}$ is a sequence of nonnegative constants depending on the sample size n such that

$$\lim_{n\to\infty} B_n = \infty \quad . \qquad (26)$$

Asymptotic Unbiasedness of the Estimate $f_n(t_o)$

Taking expectation on both sides of (25) one obtains (since the observed times to failure $T_1, T_2, \ldots T_n$ are independently identically distributed with the common distribution $F(t)$)

$$E\left(f_n(t_o)\right) = E\left(\frac{B_n}{n} \sum_{j=1}^{n} K\left(B_n(T_j - t_o)\right)\right)$$

$$= \int_o^\infty B_n K\left(B_n(t - t_o)\right) dF(t) . \qquad (27)$$

It will now be proven that

$$\lim_{n\to\infty} E\left(f_n(t_o)\right) = f(t_o) , \qquad (28)$$

at a point of continuity t_o of the distribution $F(t)$ where the density $f(t)$ is also continuous. The meaning of (28) is that as the sample size n increases indefinitely the mean value of the estimate $f_n(t_o)$ converges to its true value $f(t_o)$, which makes the estimate asymptotically (for large samples) unbiased. To prove (28), the following lemma is needed.

Lemma 1

Let $K(t)$ be a window satisfying (24). Let $t_i (i = o, 1, 2, \ldots)$ be the points of discontinuity of the distribution $F(t)$, and let S_i be the magnitude of the jump in $F(t)$ at $t = t_i$. Further, let $A_n(t) = B_n \, K\left(B_n(t - t_o')\right)$ where the density $f(t)$ is also continuous at t_o'.

Then

$$\lim_{n \to \infty} J(A_n) = \lim_{n \to \infty} \int_o^{\infty} A_n(t) \, dF(t) = f(t_o') , \tag{29}$$

provided the series

$$\sum_i \frac{S_i}{|t_i - t_o'|} \quad \text{converges.} \tag{30}$$

(Note that this assumption is used only for (36) and (38).)

Proof

Now

$$J(A_n) = \int_0^\infty B_n K\bigl(B_n(t - t'_o)\bigr)\, dF(t)$$

$$= \int_0^\infty B_n K\bigl(B_n(t - t'_o)\bigr)\, dF_1(t)$$

$$+ \int_0^\infty B_n K\bigl(B_n(t - t'_o)\bigr)\, dF_2(t), \tag{31}$$

where $F_1(t)$ and $F_2(t)$ are, respectively, the absolutely continuous part and the discrete part of the failure distribution $F(t)$.

Now,

$$\int_0^\infty B_n K\bigl(B_n(t - t'_o)\bigr)\, dF_1(t) = \int_0^\infty B_n K\bigl(B_n(t - t'_o)\bigr)\, f(t)dt\,.$$

Put

$$x = B_n(t - t'_o)\,.$$

Then,

$$\int_0^\infty B_n K\bigl(B_n(t - t'_o)\bigr) f(t)dt = \int_{-B_n t'_o}^\infty K(x)\, f\left(t'_o + \frac{x}{B_n}\right)dx. \tag{32}$$

Taking limit as $n \to \infty$ on both sides of (32) we have, since $f(t)$ is assumed continuous at t_o', that

$$\lim_{n\to\infty} \int_0^\infty B_n K\left(B_n(t - t_o')\right) \quad dF_1(t) = \int_{-\infty}^{\infty} K(x)\, f(t_o')dx$$

$$= f(t_o') \int_{-\infty}^{\infty} K(x)\, dx$$

$$= f(t_o'), \tag{33}$$

at a point of continuity t_o' of $F(t)$ and also of $f(t)$.

In view of (33), the proof of Lemma 1 is complete if it can be shown that

$$\lim_{n\to\infty} \int_0^\infty B_n\, K\left(B_n(t - t_o')\right) dF_2(t) = 0, \tag{34}$$

at a point of continuity t_o' .

To show (34), consider

$$\int_0^\infty B_n K\left(B_n(t - t_o')\right) dF_2(t) = \sum_i B_n K\left(B_n(t_i - t_o')\right) S_i$$

$$= \sum_{i \le m} B_n K\left(B_n(t_i - t_o')\right) S_i$$

$$+ \sum_{i > m} B_n K\left(B_n(t_i - t_o')\right) S_i$$

$$= \sum\nolimits_1 + \sum\nolimits_2 \text{, say.} \tag{35}$$

Since t_o' is a point of continuity, $t_i \neq t_o'$ for all i. Because $|t\,K(t)| \to 0$ as $t \to \pm\infty$, an $N_o > 0$ can be chosen such that

$$\left| B_n (t_i - t_o') K\left(B_n(t_i - t_o')\right) \right| < \epsilon \text{ for } n > N_o \text{ and all } i < m \tag{35'}$$

where ϵ is positive and arbitrary.

Hence,

$$\left|\sum\nolimits_1\right| < \epsilon \sum_{i<m} \frac{S_i}{|t_i - t_o'|} \le \epsilon A \quad , \tag{36}$$

where

$$A = \sum_i \frac{S_i}{|t_i - t_o|} < \infty \text{ , by assumption.}$$

Also, since $t\,K(t) \to 0)$ as $|t| \to \infty$, it follows that $|t\,K(t)|$ is bounded. Hence $|t\,K(t)| \le K_o$ (finite) for all t. Therefore,

$$\left|\sum\nolimits_2\right| \le K_o \sum_{i>m} \frac{S_i}{|t_i - t_o'|} \tag{37}$$

Since $\sum_i \frac{S_i}{|t_i - t_o'|}$ converges by assumption, an integer m can be chosen such that

$$\sum_{i>m} \frac{S_i}{|t_i - t_o'|} < \epsilon, \tag{38}$$

where ϵ is positive and arbitrary.

Note that one has to choose m first, so that (38); then N, so that (35').

Therefore, combining (35), (36), (37), and (38) we discover that

$$\lim_{n\to\infty} \sum_i B_n K\left(B_n(t_i - t_o')\right) S_i = 0, \tag{39}$$

which proves (34).

This completes the proof of Lemma 1.

<u>Remark 1</u>

Note that if the points of discontinuity of the distribution function are isolated points the condition imposed in the lemma, namely,

$$\sum_i \frac{S_i}{|t_i - t_o'|} < \infty \quad ,$$

is automatically satisfied.

For, in this case,

$$\inf_i \mid t_i - t_o' \mid > 0 \tag{39}$$

for every point of continuity t_o' and consequently,

$$\sum_i \frac{S_i}{|t_i - t_o'|} \le \frac{1}{t''} \sum_i S_i \le \frac{1}{t''}$$

where

$$t'' = \inf_i \mid t_i - t_o' \mid .$$

Since in practice only isolated discontinuities are encountered in the law of failures, the assumption

$$\sum_i \frac{S_i}{|t_i - t_o'|} < \infty$$

is always satisfied for applications.

<u>Remark 2</u>

If as assumed in Lemma 1 K(t) does not satisfy (24), $\int_{-\infty}^{\infty} K(t)\,dt \neq 1$ but is finite, i.e., $\int_{-\infty}^{\infty} K(t)dt < \infty$, then the limit in (29) will be

$$\lim_{n\to\infty} J\,(A_n) = f\,(t_o) \int_{-\infty}^{\infty} K\,(t)\,dt. \tag{40}$$

Now applying (29) of Lemma 1 to (28) it is at once clear that

$$\lim_{n\to\infty} E\,(f_n(t_o)) = f(t_o)\int_{-\infty}^{+\infty} K(t)\,dt\ , \tag{41}$$

at a point of continuity t_o of $F(t)$ and $f(t)$.

The Consistency of the Estimate $f_n(t_o)$ for Estimating $f(t_o)$

The consistency of $f_n(t_o)$ will be now established by showing that the variance of $f_n(t_o)$ goes to zero as the sample size n tends to infinity. This, together with the property of asymptotic unbiasedness proved earlier, will establish the consistency.

Taking variance on both sides of (25), we obtain

$$\mathrm{Var}\left[f_n(t_o)\right] = \frac{B_n^2}{n}\left[E\left(K^2\left(B_n(T - t_o)\right)\right) - E^2\left(K\left(B_n(T - t_o)\right)\right)\right] \tag{42}$$

Taking limit $n \to \infty$ on both sides of (42), one obtains, in view of (28),

$$\lim_{n\to\infty} \operatorname{Var}\left[f_n(t_o)\right] = \lim_{n\to\infty} \frac{B_n^2}{n} E\left(K^2\left(B_n(T - t_o)\right)\right)$$

$$= \lim_{n\to\infty} \frac{B_n^2}{n} \int_o^\infty K^2\left(B_n(t - t_o)\right) dF(t). \tag{43}$$

We now observe that the function $K^2(t)$ has all but one of the properties of $K(t)$, namely $\int_{-\infty}^{\infty} K^2(t)dt \neq 1$, but that it is finite. Lemma 1, therefore, holds for $K^2(t)$, with limit as given by (40).

Therefore,

$$\lim_{n\to\infty} B_n \int_o^\infty K^2(B_n(t - t_o))\, dF(t) = f(t_o) \int_{-\infty}^{\infty} K^2(t)\, dt, \tag{44}$$

at a point of continuity t_o of the distribution $F(t)$ and also of the density $f(t)$. Combining (43) and (44), we discover that

$$\lim_{n\to\infty} \left(\frac{n}{B_n}\right) \operatorname{Var}\left[f_n(t_o)\right] = f(t_o) \int_{-\infty}^{\infty} K^2(t)\, dt, \tag{45}$$

at a point of continuity t_o of $F(t)$ and $f(t)$.

Assuming now that $B_n \to \infty$ as the sample size $n \to \infty$ more slowly than n in such a way that

$$\lim_{n\to\infty} \left(\frac{B_n}{n}\right) = 0, \tag{46}$$

a sufficient condition for consistency, one obtains in view of (45) that

$$\lim_{n \to \infty} \text{Var} \left[f_n(t_o) \right] = 0 \quad . \tag{47}$$

(28) and (47) together establish the consistency of the estimator $f_n(t_o)$ for estimating the density $f(t_o)$ at a continuity point t_o of $F(t)$ and $f(t)$.

Estimation of the Reliability Function R(t)

The empirical reliability function $R_n(t)$ based on observations is defined by

$$R_n(t) = 1 \; F_n(t) , \tag{48}$$

where $F_n(t)$ is the empirical distribution function given by (22). Therefore,

$$R_n(t) = \frac{1}{n} \left[\text{number of observations} > t \text{ among } T_1, T_2, \ldots T_n \right], \tag{49}$$

where $T_1, T_2, \ldots T_n$ are the observed times to failure of n identical items subjected to a life testing experiment.

Evidently $R_n(t)$ is a binominally distributed random variable with mean and variance given by

$$E(R_n(t)) = R(t),$$
$$\text{Var} \left| R_n(t) \right| = \frac{R(t)(1 - R(t))}{n} \tag{50}$$

at all points t, whether they are points of continuity or not of the underlying law of failures $F(t)$. The meaning of (50) is that the empirical reliability function $R_n(t)$ based on the observations $T_1, T_2, \ldots T_n$ is unbiased and consistent for estimating the true reliability $R(t)$ at all points of time t.

Another class of estimators $R_n^*(t)$ for estimating the reliability function $R(t)$ is now proposed and its properties examined. Let us start with the estimate $f_n(t)$ given by (25), which was earlier shown to be asymptotically unbiased and consistent for estimating the density $f(t)$ at all points of continuity t of $F(t)$ and $f(t)$ where

$$f_n(t) = \frac{B_n}{n} \sum_{j=1}^{n} K(B_n(T_j - t)). \tag{51}$$

Now define the class of estimators

$$\begin{aligned} R_n^*(t) &= \int_t^{\infty} f_n(t)\,dt \\ &= \frac{B_n}{n} \sum_{j=1}^{n} \int_t^{\infty} K(B_n(T_j - t))\,dt \quad . \end{aligned} \tag{52}$$

It can now be proven that at a point of continuity t of the distribution $F(t)$

$$\lim_{n \to \infty} E\,(R_n^*(t)) = R(t) \tag{53}$$

$$\lim_{n \to \infty} \left| n \operatorname{Var}(R_n^*(t)) \right| = R(t)\,(1 - R(t)). \tag{54}$$

Let

$$G(t) = \int_{-\infty}^{t} K(t)\,dt\,. \tag{55}$$

In terms of $G(t)$, $R_n^*(t)$ can be written as

$$R_n^*(t) = \frac{1}{n} \sum_{j=1}^{n} G(B_n(T_j - t)). \tag{56}$$

Taking expectation on both sides of (56), the following is obtained:

$$E\left(R_n^*(t)\right) = \int_0^\infty G(B_n(\tau - t)\, dF(\tau)$$

$$= 1 - \int_0^\infty B_n K\left(B_n(\tau - t)\right) F(\tau)\, d\tau. \tag{57}$$

Now

$$\int_0^\infty B_n K(B_n(\tau - t)) F(\tau)\, d\tau = \int_{-B_n t}^\infty K(\lambda) F\left(t + \frac{\lambda}{B_n}\right) d\lambda. \tag{58}$$

If t is a point of continuity of the distribution function of time to failure F(t), one obtains, taking limit as the sample size $n \to \infty$ on both sides of (58),

$$\lim_{n\to\infty} \int_0^\infty B_n K\left(B_n(\tau - t)\right) F(\tau)\, d\tau = \int_{-\infty}^\infty F(t)\, K(\lambda)\, d\lambda = F(t). \tag{59}$$

Combining (57) and (59), we discover that

$$\lim_{n\to\infty} E\left(R_n^*(t)\right) = 1 - F(t) = R(t), \tag{60}$$

at a point of continuity t of the underlying law of failures F(t). Equation (60) establishes the asymptotic unbiasedness of $R_n^*(t)$ for estimating R(t) at every point of continuity.

Taking variance on both sides of (56), we find that

$$\mathrm{Var}\left[R_n^*(t)\right] = \frac{1}{n}\,\mathrm{Var}\left[G\Big(B_n(T - t)\Big)\right]$$

$$= \frac{1}{n}\left[E\left(G^2\Big(B_n(T-t)\Big)\right) - E^2\left(G\Big(B_n(T-t)\Big)\right)\right] . \qquad (61)$$

Now

$$E\left(G^2\Big(B_n(T-t)\Big)\right) = \int_0^\infty G^2\Big(B_n(\tau-t)\Big)\, dF(\tau)$$

$$= 1 - 2\int_0^\infty G\Big(B_n(\tau-t)\Big) B_n K\Big(B_n(\tau-t)\Big)\, F(\tau) d\tau, \qquad (62)$$

after integration by parts. Substituting $B_n(\tau - t) = \lambda$, (62) can be written as

$$E\left(G^2\Big(B_n(T - t)\Big)\right) = 1 - 2\int_{-B_n t}^\infty G(\lambda)\, K(\lambda)\, F\left(t + \frac{\lambda}{B_n}\right) d\lambda. \qquad (63)$$

Taking limit as $n \to \infty$ on both sides of (63) gives at a point of continuity t of the distribution F(t)

$$\lim_{n\to\infty} E\left(G^2\Big(B_n(T - t)\Big)\right) = 1 - 2\,F(t)\int_{-\infty}^\infty G(\lambda)\, K(\lambda)\, d\lambda$$

$$= 1 - F(t) = R(t), \qquad (64)$$

since $\int_{-\infty}^\infty G(\lambda)\, K(\lambda)\, d\lambda = \frac{1}{2}$.

Combining (57), (60), (61) and (64), we discover that

$$\lim_{n\to\infty}\left[n \text{ Var } \left(R_n^{*}(t)\right)\right] = R(t) - R^2(t) = R(t)\left(1 - R(t)\right) \tag{65}$$

at every point of continuity t of the underlying distribution F(t). Equations (60) and (65) together establish the consistency of the class of estimators $R_n^*(t)$ for estimating the reliability R(t) at every point of continuity t of the underlying law of failures F(t).

Also, at a point of continuity t, the estimate $R_n(t)$ (which is the empirical reliability function) and the class of estimators $\left\{R_n^*(t)\right\}$ have the same asymptotic variance and order of consistency. In this sense, both $R_n(t)$ and $R_n^*(t)$ are asymptotically equivalent. But, for any given sample (finite), for a given window K(t), the corresponding $R_n^*(t)$ may be more efficient than the empirical reliability function $R_n(t)$ for estimating the reliability R(t) at time t.

Having thus established the equivalence of the estimate $R_n(t)$ and the class of estimators $R_n^*(t)$ at a point of continuity t, we will examine in a subsequent section of this report (Chapter V) what happens to these estimators at a point of discontinuity of the underlying law of failures. In this case, it is shown that the estimators are not asymptotically equivalent and, indeed, provide a method of estimating the probability of failure of the item due to undergoing instantaneous hostility at any such time instant.

Estimation of the Hazard Rate at a Point of Continuity t of the Law of Failures, F(t)

The basic decomposition theorem of the cumulative hazard function for any arbitrary law of failure, $F(t)$, establishes that at any point of continuity t the unique derivative of the absolutely continuous part of the cumulative hazard function (also called the hazard rate) is given by

$$z(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{R(t)}. \tag{66}$$

The interpretation of $z(t)\,dt$ is, as usual, the probability that the item having survived time t fails between t and $t + dt$.

Propose $z_n(t)$ as an estimate of the hazard rate $z(t)$ at a point of continuity t of the distribution $F(t)$, where

$$z_n(t) = \frac{f_n(t)}{R_n(t)}\ , \tag{67}$$

$f_n(t)$ and $R_n(t)$ being, respectively, given by (25) and (49). It has been shown earlier that $f_n(t)$ is consistent for estimating the density $f(t)$ at all continuity points, i.e., $f_n(t)$ converges in probability to $f(t)$. In symbols,

$$\operatorname*{Plim}_{n\to\infty} f_n(t) = f(t), \tag{68}$$

the symbol "Plim" standing for probability limit in the sense of convergence in probability. Also from (50), note that $R_n(t)$ is consistent for estimating $R(t)$ at all points t, i.e.,

$$\operatorname*{Plim}_{n\to\infty} R_n(t) = R(t) \tag{69}$$

Since $z_n(t)$ is a rational function of $f_n(t)$ and $R_n(t)$ and since the probability limit of the denominator in $z_n(t)$ does not go to zero except at $t = \infty$, we have (using a well known convergence theorem of Cramer [4, p. 254])

$$\operatorname*{Plim}_{n\to\infty} z_n(t) = \frac{f(t)}{R(t)} = z(t). \tag{70}$$

The meaning of (70) is that the estimator $z_n(t)$ is consistent for estimating the hazard rate $z(t)$ at time t.

Also proposed is the class of estimators $z_n^*(t)$ for estimating $z(t)$ where

$$z_n^*(t) = \frac{f_n(t)}{R_n^*(t)}, \tag{71}$$

$f_n(t)$ and $R_n^*(t)$ being respectively given by (25) and (52). Since $R_n^*(t)$ and $R_n(t)$ are asymptotically equivalent for estimating the reliability $R(t)$, where t is a point of continuity of the distribution $F(t)$, it follows that $z_n^*(t)$ is also consistent for estimating $z(t)$ and that both $z_n(t)$ and $z_n^*(t)$ are asymptotically equivalent for estimating the hazard rate $z(t)$ at time t, t being a point of continuity of $F(t)$.

Asymptotic Normality and Confidence Bands for f(t), R(t), and z(t)

It is now possible to investigate the reliability of any time t, as well as the question of what happens to the distributions of the estimates for the density of the underlying law of failures at a point of continuity, and finally the hazard or failure rate at a point of continuity of the underlying law of failures. In particular, it will be shown that these distributions are asymptotically Gaussian and thus provide the basis for large-sample confidence bands for these life quality paramenters at any desired level of confidence.

In order to prove the asymptotic normality, the following lemma is needed.

Lemma 2

Let $V_1, V_2, \ldots V_n, \ldots$ be a sequence of independently and identically distributed random variables. Define the sequence $\{S_n\}$ where

$$S_n = \frac{1}{n} \sum_{j=1}^{n} V_j .$$

Then a sufficient condition for the sequence $\{S_n\}$ to be asymptotically normally distributed is that for some $\delta > 0$

$$\lim_{n\to\infty} \frac{E|V_n - E(V_n)|^{2+\delta}}{n^{\delta/2}\left(\mathrm{Var}\,(V_n)\right)^{1+\delta/2}} = 0. \tag{72}$$

For proof of this well-known lemma, refer to Parzen [9, p. 1019].

Asymptotic Normality of the Estimate $f_n(t)$ and Associated Confidence Bands for the Density $f(t)$

The estimator $f_n(t_o)$ for estimating the density of the underlying law of failures $F(t)$ at a point of continuity t_o is given by

$$f_n(t_o) = \int_o^\infty B_n\, K\left(B_n(t - t_o)\right) dF_n(t)$$

$$= \frac{B_n}{n} \sum_{j=1}^{n} K\left(B_n(T_j - t_o)\right). \tag{73}$$

Now (73) can be written as

$$f_n(t_o) = \frac{1}{n} \sum_{j=1}^{n} V_j ,$$

where

$$V_j = B_n K\left(B_n(T_j - t_o)\right),\ j = 1, 2, \ldots n. \tag{74}$$

The sequence $\{V_j\}$ given by (74) is independently and identically distributed as a random variable

$$V(n) = B_n K\left(B_n(T - t_o)\right). \tag{75}$$

Applying Lemma 1 to the random variable V(n) given by (75), we discover that

$$E\left|V(n)\right|^{2+\delta} \sim B_n^{1+\delta/2}\ f(t_o) \int_{-\infty}^{\infty} \left(K(t)\right)^{2+\delta} dt ,$$

$$\operatorname{Var}\left(V(n)\right) \sim B_n\ f(t_o) \int_{-\infty}^{\infty} K^2(t)\, dt , \tag{76}$$

at every point of continuity t_o of the underlying law of failures F(t) and the density f(t).

In view of $\int_{-\infty}^{\infty} K(t)\, dt = 1$,

$$\int_{-\infty}^{\infty} \left(K(t)\right)^{2+\delta} dt < \infty, \text{ for all } \delta \geq 0. \tag{77}$$

Taking (76) and (77) and the condition that $\frac{B_n}{n} \to 0$ as $n \to \infty$, it is at once clear that V(n) given by (75) satisfies the condition (72) of Lemma 2. It is thus proved that the estimator $f_n(t_o)$ given by (73) is asymptotically normal for estimating $f(t_o)$ at every point of continuity t_o of F(t) and f(t); that is

$$\lim_{n\to\infty} P\left\{ \left(\frac{n}{B_n}\right)^{1/2} \left(\frac{f_n(t) - f(t)}{\left[f(t) \int_{-\infty}^{\infty} K^2(\tau) d\tau\right]^{1/2}} \right) \leq x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2} y^2} dy, \tag{78}$$

at every continuity point t of F(t) and f(t).

Now let t_α be the upper α percentage point of a normal distribution with zero mean and unit variance. Then, the confidence interval with confidence coefficient α for the density f(t) at every point of continuity t is given by the expression in the parenthesis of the following equation:

$$\lim_{n\to\infty} P\left(f_n(t) - t_{\frac{\alpha}{2}} \sqrt{\frac{B_n}{n} \, f_n(t) \int_{-\infty}^{\infty} K^2(\tau) d\tau} \leq f(t) \leq f_n(t) + t_{\frac{\alpha}{2}} \sqrt{\frac{B_n}{n} \, f_n(t) \int_{-\infty}^{\infty} K^2(\tau) d\tau} \right) = \alpha \tag{79}$$

Asymptotic Normality of the Estimate $R_n(t)$ and the Class of Estimators $R_n^*(t)$ for Estimating the Reliability Function $R(t)$

The empirical reliability function $R_n(t)$ given by (49), where

$$R_n(t) = \frac{1}{n}\,[\text{number of observations} > t \text{ among } T_1, T_2, \ldots T_n], \tag{80}$$

is binominally distributed with

$$E\big(R_n(t)\big) = R(t)$$

$$\operatorname{Var}\big(R_n(t)\big) = \frac{1}{n}\,R(t)\,\big(1 - R(t)\big)\,, \tag{81}$$

at all time points t.

From the normal approximation to the binominal distribution, it follows that

$$\lim_{n\to\infty} P\left\{\sqrt{n}\,\frac{R_n(t) - R(t)}{\sqrt{R(t)\big(1 - R(t)\big)}} \le x\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}y^2}\,dy. \tag{82}$$

Hence, if t_α is the upper α percentage point of a normal distribution with zero mean and unit variance, then

$$\lim_{n\to\infty} P\left(R_n(t) - t_{\frac{\alpha}{2}}\left(\frac{R_n(t)F_n(t)}{n}\right)^{1/2} \le R(t) \le R_n(t) + t_{\frac{\alpha}{2}}\left(\frac{R_n(t)F_n(t)}{n}\right)^{1/2}\right) = \alpha. \tag{83}$$

The expression within parentheses in the above equation is the α percentage large-sample confidence interval for the reliability function R(t) based on the empirical reliability function $R_n(t)$. Since the exact distribution of $R_n(t)$ is known, exact confidence intervals based on the binominal distribution can also be obtained for reliability function $R(t)$.

Now consider the class of estimators $R_n^*(t)$ given by (52), where

$$R_n^*(t) = \int_t^\infty f_n(\tau)\, d\tau = \frac{1}{n} \sum_{j=1}^{n} G\Big(B_n(T_j - t)\Big) . \tag{84}$$

The class of estimators $R_n^*(t)$ given by (84) has been shown earlier to be consistent for estimating the reliability function $R(t)$ at every point of continuity t of the underlying law of failures $F(t)$.

Now (84) can be written as

$$R_n^*(t) = \frac{1}{n} \sum_{j=1}^{n} V_j \quad , \tag{85}$$

where

$$V_j = G\Big(B_n(T_j - t)\Big)$$

and the sequence of random variables $V_1, V_2, \ldots$ is independently identically distributed as the random variable

$$V(n) = G\Big(B_n(T - t)\Big) . \tag{86}$$

Applying Lemma 1 to the random variable $V(n)$ given by (86), we discover that

$$\lim_{n \to \infty} E|V(n)|^{2+\delta} < \infty, \text{ for } \delta > o, \text{ and } \lim_{n\to\infty} \operatorname{Var}\Big(V(n)\Big) < \infty , \tag{87}$$

at every point of continuity t of the distribution $F(t)$.

In view of (87), the condition (72) of Lemma 2 is satisfied for the random variable $V(n)$ given by (86). Hence, the class of estimators $R_n^*(t)$ is asymptotically normal for estimating the reliability function $R(t)$ at every point of continuity t of the underlying law of failures $F(t)$, i.e.,

$$\lim_{n \to \infty} P \left\{ \sqrt{n} \; \frac{R_n^*(t) - R(t)}{\sqrt{R_n^*(t) \; F_n^*(t)}} < x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2} y^2} \, dy \, . \tag{88}$$

Comparing (82) and (88), we find that the α percentage confidence interval for $R(t)$ is the same whether it is obtained from $R_n(t)$ or $R_n^*(t)$, since both $R_n(t)$ and $R_n^*(t)$ are asymptotically equivalent for estimating $R(t)$ at every point of continuity t of the failure distribution $F(t)$.

Asymptotic Normality of the Class of Estimators $z_n(t)$ for Estimating the Hazard Rate $z(t)$

The estimator $z_n(t)$ for estimating the hazard rate $z(t)$ at the point of continuity t is given by

$$z_n(t) = \frac{f_n(t)}{R_n(t)} , \tag{89}$$

where $f_n(t)$ and $R_n(t)$ are, respectively, given by (73) and (80).
From (50) one finds that $R_n(t)$ is consistent for estimating $R(t)$, i.e.,

$$\operatorname*{Plim}_{n \to \infty} R_n(t) = R(t). \tag{90}$$

Also from (78),

$$\lim_{n\to\infty} P\left\{\left(\frac{n}{B_n}\right)^{1/2}\left(\frac{f_n(t)-f(t)}{\left(f(t)\int_{-\infty}^{\infty}K^2(t)dt\right)^{1/2}}\right)\leq x\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}y^2}dy . \tag{91}$$

Combining (90) and (91) and using the convergence theorem of Cramer [4, p.254], we discover that

$$\lim_{n\to\infty} P\left\{\left(\frac{n}{B_n}\right)^{1/2}\frac{\frac{f_n(t)}{R_n(t)}-\frac{f(t)}{R_n(t)}}{\left(\frac{f(t)}{R^2(t)}\int_{-\infty}^{\infty}K^2(x)dx\right)^{1/2}}<x\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}y^2}dy . \tag{92}$$

Now consider

$$y_n = \left(\frac{n}{B_n}\right)^{1/2}\left(R_n(t)-R(t)\right). \tag{93}$$

In view of (50)

$$E(y_n) = 0$$

$$\operatorname{Var}(y_n) = \left(\frac{1}{B_n}\right)R(t)\left(1-R(t)\right), \tag{94}$$

and hence

$$\operatorname*{Plim}_{n\to\infty} y_n = 0. \tag{95}$$

Combining (92) and (95) and again using Cramer's convergence theorem we obtain

$$\lim_{n\to\infty} P\left\{\left(\frac{n}{B_n}\right)^{1/2} \frac{z_n(t) - z(t)}{\left(\frac{z(t)}{R(t)}\int_{-\infty}^{\infty} K^2(x)dx\right)^{1/2}} \le x\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}y^2} dy. \tag{96}$$

Equation (96) establishes that the class of estimators $z_n(t)$ are asymptotically normal for estimating the hazard rate $z(t)$ at every continuity point t of the underlhing law of failures $F(t)$.

Now, consider the class of estimators $z_n^*(t)$ for estimating the hazard rate $z(t)$, where

$$z_n^*(t) = \frac{f_n(t)}{R_n^*(t)}, \tag{97}$$

$f_n(t)$ and $R_n^*(t)$ are, respectively, given by (51) and (52).

Since $R_n(t)$ and $R_n^*(t)$ are asymptotically equivalent at every continuity point t of $F(t)$ using a similar argument as in the case of $z_n(t)$, it is evident that the class of estimators $z_n^*(t)$ is also asymptotically normal for estimating the hazard rate $z(t)$ at every continuity point t of $F(t)$.

Now let t_α be the upper α percentage point of a normal distribution with zero mean and unit variance. Then the confidence interval with confidence coefficient α for the hazard rate $z(t)$ at every point of continuity t of the underlying law of failures $F(t)$ is given by the expression in parentheses in the following equation:

$$\lim_{n\to\infty} P\left(z_n(t) - t_{\alpha/2}\sqrt{\frac{B_n}{n}\frac{z_n(t)}{R_n(t)}\int_{-\infty}^{\infty} K^2(x)dx} \le z(t) \le z_n(t) + t_{\alpha/2}\sqrt{\frac{B_n}{n}\frac{z_n(t)}{R_n(t)}\int_{-\infty}^{\infty} K^2(x)dx}\right) = \alpha \tag{98}$$

CHAPTER III

ESTIMATION AND CONFIDENCE BANDS FOR CENSORED SAMPLES

In this chapter we consider the case of censored sampling, i.e., the situation when N items are put to test and the test is terminated as soon as $M = \alpha N$ items $0 < \alpha < 1$ have failed.

Let $\tau_1 \leq \tau_2 \leq \cdots \leq \tau_M$ be the observed times to failure of the M items.

The empirical distribution function is given by

$$F_{M,N}(t) = \frac{1}{M}\left\{\text{Min}\begin{Bmatrix}\text{number observations} \leq t \text{ among}\\ \text{the sample of size } N\end{Bmatrix}, M\right\} \tag{99}$$

$$= \frac{\ddot{m}}{M} \text{ , say.}$$

Now

$$p\{M = i\} = \frac{\binom{N}{i}F^i(t)(1-F(t))^{N-i}}{\sum_{i=0}^{M}\binom{N}{i}F^i(t)(1-F(t))^{N-i}} \tag{100}$$

$$= \frac{\binom{N}{i}\left(\frac{F(t)}{1-F(t)}\right)^i}{\sum_{i=0}^{M}\binom{N}{i}\left(\frac{F(t)}{1-F(t)}\right)^i} \text{ ,}$$

$$i = 0,1,2,\cdots M \text{ .}$$

Let $\quad (C)^i = \left(\frac{F(t)}{1-F(t)}\right)^i \quad ;$

$$\text{Prob}\{M = i\} = \frac{\binom{N}{i}C^i}{\sum_{j=0}^{M}\binom{N}{j}(C)^j} \text{ .}$$

Hence

$$E(M) = \frac{\sum_{i=0}^{M} i\binom{N}{i}C^i}{\sum_{i=0}^{M}\binom{N}{i}C^i} \text{ .}$$

Now $$(1+C)^N = \sum_{i=0}^{N} \binom{N}{i} C^i$$

$$\sum_{i=0}^{N} i \binom{N}{i} C^i = \binom{N}{i} C + 2\binom{N}{2} C^2 + \cdots + (N-1)\binom{N}{N-1} C^{N-1} + N\binom{N}{N} C^N$$

$$= NC + 2\,\frac{N(N-1)}{1\cdot 2}\;C^2 + 3\,\frac{N(N-1)(N-2)}{1\cdot 2\cdot 3}\,C^3 + \cdots N\binom{N}{N} C^N$$

$$= NC\left[1 + (N-1)C + \frac{(N-1)(N-2)}{1\cdot 2}\,C + \binom{N-1}{N-2} C^{N-2} + \binom{N-1}{N-1} C^{N-1}\right]$$

$$= NC \sum_{j=0}^{N-1} \binom{N-1}{j} C^j$$

$$= NC\,(1+C)^{N-1}$$

$$E(M) = \frac{\sum_{i=0}^{N} i \binom{N}{i} C^i - \sum_{i=M+1}^{N} i \binom{N}{i} C^i}{\sum_{i=0}^{N} \binom{N}{i} C^i - \sum_{i=M+1}^{N} \binom{N}{i} C^i}$$

$$= \frac{\dfrac{\sum_{i=0}^{N} i \binom{N}{i} C^i}{\sum_{i=0}^{N} \binom{N}{i} C^i} - \dfrac{\sum_{i=M+1}^{N} i \binom{N}{i} C^i}{\sum_{i=0}^{N} \binom{N}{i} C^i}}{1 - \dfrac{\sum_{i=M+1}^{N} \binom{N}{i} C^i}{\sum_{i=0}^{N} \binom{N}{i} C^i}}$$

$$= \frac{\dfrac{NC}{1+C} - \dfrac{\sum_{i=M+1}^{N} i\binom{N}{i}C^i}{\sum_{i=0}^{N}\binom{N}{i}C^i}}{1 - \dfrac{\sum_{i=M+1}^{N}\binom{N}{i}C^i}{\sum_{i=0}^{N}\binom{N}{i}C^i}}$$

$$E\left(\frac{M}{N} F_{M,N}(t)\right) = \frac{\dfrac{C}{1+C} - \dfrac{\sum_{i=M+1}^{N} i\binom{N}{i}C^i}{N\sum_{i=0}^{N}\binom{N}{i}C^i}}{1 - \dfrac{\sum_{i=M+1}^{N}\binom{N}{i}C^i}{\sum_{i=0}^{N}\binom{N}{i}C^i}}$$

Consider now

$$\frac{\sum_{i=M+1}^{N}\binom{N}{i}C^i}{\sum_{i=0}^{N}\binom{N}{i}C^i} = \frac{\sum_{i=M+1}^{N}\binom{N}{i}C^i}{(1+C)^N} = \frac{\sum_{i=M+1}^{N}\binom{N}{i}C^i}{\left(1+\dfrac{F(t)}{1-F(t)}\right)^N}$$

$$= (1-F(t))^N \sum_{i=M+1}^{N}\binom{N}{i}C^i$$

Suppose $M = \theta N$ where $0 < \theta < 1$. Then the above

$$\doteq (1-F(t))^N \sum_{i=\gamma N}^{N} \binom{N}{i} C^i$$

where $\gamma N = \theta N+1$

Hence

$$\lim_{N\to\infty} \frac{\sum_{i=M+1}^{N} \binom{N}{i} C^i}{\sum_{i=0}^{N} \binom{N}{i} C^i} = \lim_{N\to\infty} (1-F(t))^N \sum_{i=\gamma N}^{N} \binom{N}{i} C^i$$

$$= 0 .$$

Now
$$\sum_{i=M+1}^{N} i \binom{N}{i} C^i \leq N \sum_{i=M+1}^{N} \binom{N}{i} C^i .$$

Hence
$$\frac{\sum_{i=M+1}^{N} i \binom{N}{i} C^i}{N \sum_{i=0}^{N} \binom{N}{i} C^i} \leq \frac{\sum_{i=M+1}^{N} \binom{N}{i} C^i}{\sum_{i=0}^{N} \binom{N}{i} C^i} .$$

Hence
$$\lim_{N\to\infty} \frac{\sum_{i=M+1}^{N} i \binom{N}{i} C^i}{N \sum_{i=0}^{N} \binom{N}{i} C^i} = 0 .$$

Thus

$$\lim_{N\to\infty} E(\theta F_{M,N}(t)) = F(t). \tag{101}$$

We will now prove that $\theta F_{M,N}(t)$ is consistent for estimating $F(t)$ for any t. In view of (101), since $\theta F_{M,N}(t)$ is asymptotically unbiased for estimating $F(t)$, it is sufficient to show that the variance of $\theta F_{M,N}(t)$ goes to zero as the sample size $N\to\infty$.

Now

$$\mathrm{Var}\ (\theta F_{M,N}(t)) = \mathrm{Var}\ (\tfrac{M}{N} F_{M,N}(t))$$

$$= \frac{1}{N^2}\ \mathrm{Var}\ (m) = \frac{E(m^2)}{N^2} - \left[\frac{E(m)}{N}\right]^2 \quad . \tag{102}$$

In a straightforward manner, as in the case of $E(\theta F_{M,N}(t))$, it can be shown easily that

$$\lim_{N\to\infty} \frac{E(m^2)}{N} = F^2(t) \quad , \tag{103}$$

and

$$\lim_{N\to\infty} \frac{E(m^2)}{N^2} = F^2(t) \quad . \tag{104}$$

combining (102), (103) and (104) we discover that

$$\lim_{N\to\infty} \mathrm{Var}\ (\theta F_{M,N}(t)) = 0 \quad . \tag{105}$$

Equations (101) and (105) together establish the consistency of $\theta F_{M,N}(t)$ for estimating the underlying law of failures F(t).

In order to obtain consistent estimates for the density f(t) of the underlying law of failures and the hazard function $Z(t) = f(t)/R(t)$, we proceed as follows:

Let K(x) be the usual weight-function or window as defined in Chapter II above, and let $\{B_N\}$ be a sequence of nonnegative constants depending on the sample size N such that $\frac{B_N}{N} \to 0$ as $N\to\infty$

Consider

$$F^*_{M,N}(\tau) = \int_0^\infty B_N K(B_N(t-\tau))\ F_{M,N}(t)\ dt \quad , \tag{106}$$

where from (99)

$$F_{M,N}(t) = \frac{1}{M} \left\{ \mathrm{Min} \left[\begin{matrix} \text{Number of observations} \le t \\ \text{among } T_1, T_2, \cdots T_N \end{matrix} \quad , M \right] \right\} \tag{107}$$

$$= \frac{1}{M} \sum_{j=1}^{M} U(t-\tau_j) \quad ,$$

and $U(x) = 1$ for $x \geq 0$

$= 0$ otherwise, $\tau_1 \leq \tau_2 \leq \cdots \leq \tau_M$ are the observed times to failure of the M items.

Combining (106) and (107) we obtain

$$F^*_{M,N}(t) = \frac{1}{M} \sum_{j=1}^{M} B_N \int_0^{\infty} K(B_N(t-\tau))\ U(t-\tau_j)\ dt \tag{108}$$

$$= \frac{B_N}{M} \sum_{j=1}^{M} \int_{\tau_j}^{\infty} K(B_N(t-\tau))\ dt\ .$$

Let

$$1 - \int_0^t K(x)dx = \int_t^{\infty} K(x)dx = G(t)\ . \tag{109}$$

Making the substitution $B_N(t-\tau) = x$ in (108) we obtain

$$F^*_{M,N}(\tau) = \frac{1}{N} \sum_{j=1}^{N} \int_{B_N(\tau_j-\tau)}^{\infty} K(x)\ dx \tag{110}$$

$$= \frac{1}{N} \sum_{j=1}^{N} G(B_N(\tau_j-\tau))\ .$$

In the form (110) the statistic $F^*_{M,N}(\tau)$ is straightforward and easy to calculate.

The consistency of $F^*_{M,N}(\tau)$ for estimating $F(\tau)$ can be established and the details of the calculation will not be given here.

To obtain the estimate for the density $f(\tau)$ we differentiate both sides of (110) with respect to τ and obtain

$$f^*_{M,N}(\tau) = \frac{1}{N} \sum_{j=1}^{N} B_N K(B_N(\tau_j-\tau))\ . \tag{111}$$

From the consistency of $F^*_{M,N}(\tau)$ for estimating $F(\tau)$, the consistency of $f^*_{M,N}(\tau)$ $\{$the derivative of $F^*_{M,N}(\tau)\}$ for estimating $f(\tau)$ $\{$the derivative of $F(\tau)\}$ follows.

To estimate the hazard function $Z(\tau)$ at time τ, we propose

$$Z^*_{M,N}(\tau) = \frac{f^*_{M,N}(\tau)}{1 - F_{M,N}(\tau)} \tag{112}$$

The asymptotic variance and the consistency of $Z^*_{M,N}(\tau)$ can be obtained by simple but tedious calculations. Also the asymptotic normality and the associated confidence bands for all the estimates considered can be established by methods similar to those in Chapter II.

CHAPTER IV

ESTIMATION AND CONFIDENCE BANDS FOR THE TRUNCATED SAMPLING SCHEME

This section addresses the situation where the life testing scheme is truncated. Let N identical items be put to a life testing experiment and let the experiment be terminated after time τ.

Let $T_1, T_2, \ldots, T_M$ denote the observed times to failure obtained in the above experiment. Here each T_i, $i = 1, 2, \ldots, M$ is a random variable and each is less than or equal to τ. Also, the sample size M itself is a random variable, as we do not know before the experiment is performed how many of the sample items fail by the fixed test period τ.

Let

$$P\{T_i \leq t\} = F(t), \; i = 1, 2, \ldots, M.$$

F(t) is the underlying distribution function of time to failure T or, equivalently, the so-called underlying law of failures.

Our object now is to estimate the underlying law of failures, the reliability function, and the hazard rate, based on the above truncated sampling scheme, without assuming anything about the form of the underlying law of failures, e.g., Weibull, lognormal, gamma, etc.

The Empirical Distribution for the Truncated Scheme

Let

$$F_M(t) = \frac{1}{M}\begin{bmatrix}\text{Number of observations among} \\ T_1,\ T_2,\ \ldots,\ T_M \leq t\end{bmatrix} \tag{113}$$

$$\text{for } t \leq \tau\ ;$$

$$= 1, \text{ otherwise,}$$

since all T_1, T_2, ..., T_M are less than or equal to τ.
From (113) it is easy to see that we cannot estimate the underlying law of failures $F(t)$ or, equivalently, the reliability function $R(t)$ for values of time exceeding the test period τ.
We will now write down the empirical reliability function as follows:

$$R_M(t) = 1 - F_M(t), \qquad \text{where} \tag{114}$$

$F_M(t)$ is given by (113).

Let

$$U(x) = 1 \text{ for } x \geq 0 \tag{115}$$

$$= 0 \text{ otherwise.}$$

In terms of (115), (113) can be written as

$$F_M(t) = \frac{1}{M} \sum_{j=1}^{M} U(t - T_j) \ , \tag{116}$$

$$\text{for } t \leq \tau \ ,$$

$$= 1, \text{ otherwise.}$$

Sampling Properties of $F_M(t)$ for the Truncated Case

Clearly, the sample size M is a binomially distributed random variable with parameters N and probability p given by

$$p = P(T \leq \tau) = F(\tau) \ . \tag{117}$$

Hence,

$$P\{M = m\} = \binom{N}{m} p^m(1 - p)^{N-m}, \tag{118}$$

$$E(M) = Np = NF(\tau), \tag{119}$$

$$\mathrm{Var}(M) = Np(1 - p) = NF(\tau)R(\tau) \ . \tag{120}$$

Now for any fixed M (say, M = m) we have, from (116),

$$E\left(F_m(t)\right) = E\left(U(t - T_j)\right) \ . \tag{121}$$

Since all $T_j \le \tau$, we have for the distribution function of T_j

$$P(T_j \le t) = \frac{F(t)}{F(\tau)}, \quad 0 \le t \le \tau. \tag{122}$$

Using (121) and (122) we obtain

$$E\left(F_m(t)\right) = E \quad U(t - T_j) \tag{123}$$

$$= \int_0^t d\left(\frac{F(t)}{F(\tau)}\right)$$

$$= \frac{F(t)}{F(\tau)} .$$

Since (123) is independent of m, we have from (113)

$$E\left(F_M(t)\right) = \frac{F(t)}{F(\tau)} . \tag{124}$$

$F_M(t)$ is thus not an unbiased estimate of the underlying failure distribution $F(t)$.

Equation (124) thus suggests the following exactly unbiased estimate which unfortunately involves $F(\tau)$:

$$F(\tau)F_M(t) = \frac{F(\tau)}{M} \sum_{j=1}^{M} U(t - T_j), \quad \text{for } t \leq \tau , \tag{125}$$

$$= 1 \text{ otherwise} .$$

Now since $F(\tau)$ is a binomial probability, the best estimate for $F(\tau)$ is given by M/N. Substituting for $F(\tau)$ its best estimate M/N in (125), we obtain

$$\hat{F}_M(t) = \frac{1}{N} \sum_{j=1}^{M} U(t - T_j) , \quad t \leq \tau . \tag{126}$$

In (125) and (126) we assumed that the binomial random variable $M \geq 1$. From (118) we obtain that the conditional distribution of M, given $M \geq 1$, is

$$P\{M = m/M \geq 1\} = \frac{\binom{N}{m} p^m (1-p)^{N-m}}{1-(1-p)^N} , \tag{127}$$

$$m = 1, 2, 3, \ldots, N .$$

Hence

$$E(M/M \geq 1) = \sum_{m=1}^{N} \frac{m\binom{N}{m} p^m (1-p)^{N-m}}{1-(1-p)^N} . \tag{128}$$

Thus

$$E(M/M \geq 1) = \frac{NF(\tau)}{1-(1-F(\tau))^N} \tag{129}$$

Now the expected value of (126) for fixed M, say $M = m$, is given by

$$E\left(\hat{F}_m(t)\right) = \frac{m}{N}\ \frac{F(t)}{F(\tau)} \tag{130}$$

Now taking expectation with respect to m in (130) we discover, in view of (129), that

$$E\left(\hat{F}_M(t)\right) = \frac{NF(\tau)}{N\left(1-(1-F(\tau))^N\right)}\ \frac{F(t)}{F(\tau)} \tag{131}$$

Taking limit as $N \to \infty$ in (131), we discover that

$$\lim_{N \to \infty} E\left(\hat{F}_M(t)\right) = F(t)\ ,\ t \le \tau. \tag{132}$$

Equation (132) establishes that $\hat{F}_M(t)$ given by (126) is asymptotically unbiased for estimating $F(t)$ for $t \le \tau$.

In order to show that $\hat{F}_M(t)$ is consistent for estimating $F(t)$ for $t \le \tau$, we will first compute the variance of $\hat{F}_M(t)$ and then show that it goes to zero as the sample size $N \to \infty$.

We have

$$\mathrm{Var}\left(\hat{F}_M(t)\right) = \frac{E(M)}{N^2} \quad \mathrm{Var}\left(U(t - T_j)\right) \tag{133}$$

$$= \frac{NF(\tau)}{N^2[1-(1-F(\tau))^N]} \; \frac{F(t)}{F(\tau)} \left(1 - \frac{F(t)}{F(\tau)}\right)$$

$$= \frac{1}{N}\left[\frac{F(t)\left(1 - \frac{F(t)}{F(\tau)}\right)}{1-(1-F(\tau))^N}\right] .$$

Thus

$$\lim_{N \to \infty} N \,\mathrm{Var}\left(\hat{F}_M(t)\right) = F(t)\left(1 - \frac{F(t)}{F(\tau)}\right), \quad t \leq \tau . \tag{134}$$

It follows from (134) that

$$\lim_{N \to \infty} \mathrm{Var}\left(\hat{F}_M(t)\right) = 0 . \tag{135}$$

Combining (132) and (135), we discover that $\hat{F}_M(t)$ given by (126) is consistent for estimating $F(t)$, $t \leq \tau$.

Following exactly the same procedure as in Chapter II, we also discover that the sequence of statistics given by

$$\sqrt{N} \quad \frac{\hat{F}_M(t) - F(t)}{\left[\hat{F}_M(t) \left(1 - \frac{N\hat{F}_M(t)}{M} \right) \right]^{1/2}} \tag{136}$$

converges in distribution to a normal distribution with zero mean and unit variance as the sample size $N \to \infty$, for $t \le \tau$.

It now follows from (114) and (136) that the sequence of statistics

$$\sqrt{N} \quad \frac{\hat{R}_M(t) - R(t)}{\left[\hat{F}_M(t) \left(1 - \frac{N\hat{F}_M(t)}{M} \right) \right]^{1/2}}$$

converges in distribution to a normal distribution with zero mean and unit variance as the sample size $N \to \infty$, where the empirical reliability statistic is given by

$$\hat{R}_M(t) = 1 - \hat{F}_M(t) \ , \quad \text{for } t \le \tau \ .$$

The estimation procedure for estimating the hazard rate $z(t) = f(t)/R(t)$ in this case is very similar to the simple random sampling situation given in Chapter II. Also, the asymptotic normality follows in a similar manner.

Thus confidence bands at any desired level of confidence for the reliability function and the hazard rate follow from the above results and remarks.

CHAPTER V

THE MEANING OF JUMPS OF F(t); ESTIMATION, AND TESTS OF HYPOTHESES

Interpretation of a Discontinuity and Jump in the Underlying Law of Failures

If the distribution function F(t) of time to failure T is absolutely continuous, then the pure step function $F_2(t)$ is identically zero for all t and

$$F(t) = F_1(t) = \int_0^t f(\tau)d\tau \ , \tag{137}$$

where f(t), the derivative of the absolutely continuous part, is the probability density function. In this case,

$$P(T = t_o) = 0 \quad , \tag{138}$$

where t_o is any specified time instant. In other words, the probability of the event that the item fails at time t_o is identically zero. On the other hand, if the distribution function F(t) is not absolutely continuous and if the time instant t_o corresponds to a point of discontinuity in the distribution F(t), we have

$$P(T = t_\nu) = S_\nu \ , \ \nu = 0, 1, 2, \ . \ . \ \infty \ , \tag{139}$$

where S_ν is the magnitude of the jump of $F_2(t)$ or $F(t)$ at $t = t_\nu$. The meaning of (139) is that there is a strictly positive probability, equal to the size of the jump, that the item fails at time instants corresponding to the points of discontinuity in the underlying law of failures. This may happen if the item is subjected to instantaneous hostile atmosphere at these time points. A vehicular system traversing through space and being impinged upon by failure-causing meteorites provides an example of such a situation.

Estimation of the Jump S_i at the Discontinuity t_i of the Underlying Law of Failures $F(t)$

Assuming the singular part to be identically zero, the distribution $F(t)$ can be decomposed into (see Cramer [4, pp. 52, 53]):

$$F(t) = F_1(t) + F_2(t) \quad , \tag{140}$$

where $F_1(t)$ is an everywhere continuous function, and $F_2(t)$ is a pure step function with steps of magnitude, say, S_ν at the points $t = t_\nu$, $\nu = 1, 2, \ldots$; $F_1(t)$ and $F_2(t)$ are nondecreasing and uniquely determined. Substituting (140) in equation (57), Chapter II, we obtain

$$\begin{aligned} E\left(R_n^*(t)\right) &= \int_0^\infty G\left(B_n(\tau-t)\right) dF_1(\tau) + \int_0^\omega G\left(B_n(\tau-t)\right) dF_2(\tau) \\ &= I_1 + I_2 \quad , \quad \text{say}. \end{aligned} \tag{141}$$

Using arguments similar to those following equation (57) (Chapter II), we readily obtain

$$\lim_{n\to\infty} I_1 = F_1(\infty) - F_1(t) \quad .$$

Since $F_1(t)$ is continuous at $t = t_i$,

$$\lim_{n\to\infty} I_1 = F_1(\infty) - F_1(t_i) \tag{142}$$

at the discontinuity $t = t_i$ of the underlying law of failures $F(t)$.

Now

$$I_2 = \int_0^\infty G\Big(B_n(\tau - t)\Big)\, dF_2(\tau) = \sum_{\nu=1}^{\infty} S_\nu G\Big(B_n(t_\nu - t)\Big) \quad . \tag{143}$$

Denoting by $\sum\limits_{t_\nu > t_i}$ summation over all ν such that $t_\nu > t_i$ and by $\sum\limits_{t_\nu < t_i}$ summation over all ν such that $t_\nu < t_i$, at the discontinuity $t = t_i$ of the distribution $F(t)$, I_2 can be written as

$$I_2 = I_{21} + I_{22} + I_{23} \; ,$$

where

$$I_{21} = \sum_{t_\nu < t_i} S_\nu G\Big(B_n(t_\nu - t_i)\Big) \, ,$$

$$I_{22} = S_i G(0) = 1/2 \; S_i \; ,$$

and

$$I_{23} = \sum_{t_\nu > t_i} S_\nu G\Big(B_n(t_\nu - t_i)\Big) \, .$$

Now

$$I_{21} = \sum\nolimits_1 + \sum\nolimits_2$$

where

$$\sum_1 = \sum_{\substack{t_\nu < t_i, \\ |\nu| \le m}} S_\nu G\left(B_n(t_\nu - t_i)\right),$$

and

$$\sum_2 = \sum_{\substack{t_\nu < t_i, \\ |\nu| > m}} S_\nu G\left(B_n(t_\nu - t_i)\right).$$

It can be argued, as in the proof of Lemma 1 (Chapter II) that Σ_2 can be made arbitrarily small by choosing m sufficiently large (no matter what n is); and Σ_1, for fixed m, can be made arbitrarily small by choosing n sufficiently large, i.e.,

$$\lim_{n\to\infty} I_{21} = 0 .$$

From

$$I_{23} = \sum_{t_\nu > t_i} S_\nu - \sum_{t_\nu > t_i} S_\nu \left[1 - G\left(B_n(t_\nu - t_i)\right)\right] ,$$

it is discovered that

$$\lim_{n\to\infty} I_{23} = \sum_{t_\nu > t_i} S_\nu .$$

Of course, it should be noted that in proving the above statement it is assumed that

$$\sum_{\nu \ne i} \frac{S_\nu}{|t_\nu - t_i|} < \infty .$$

This proves that, at the discontinuity $t = t_i$ of the underlying law of failures F(t)

$$\lim_{n\to\infty} I_2 = 1/2\ S_i + \sum_{t_\nu > t_i} S_\nu . \tag{144}$$

Combining equations (141), (142), and (144) we obtain

$$\lim_{n\to\infty} E\left(R_n^*(t_i)\right) = F_1(\infty) - F_1(t_i) + 1/2\, S_i + \sum_{t_\nu > t_i} S_\nu \tag{145}$$

at the discontinuity t_i of the underlying law of failures $F(t)$.

Now

$$F(t_i) = \int_0^{t_i} d\left(F_1(t) + F_2(t)\right) = F_1(t_i) + \sum_{t_\nu \le t_i} S_\nu ,$$

and therefore

$$\begin{aligned} R(t_i) &= 1 - F(t_i) = F_1(\infty) + F_2(\infty) - F_1(t_i) - \sum_{t_\nu \le t_i} S_\nu \\ &= F_1(\infty) - F_1(t_i) + \sum_{t_\nu > t_i} S_\nu . \end{aligned} \tag{146}$$

Substituting equation (146) in equation (145) we discover that

$$\lim_{n\to\infty} E\left(R_n^*(t_i)\right) = R(t_i) + 1/2\, S_i. \tag{147}$$

From equation (50) of Chapter II we have that

$$E\left(R_n(t_i)\right) = R(t_i) . \tag{148}$$

Write

$$H_n(t_i) = 2\left[R_n^*(t_i) - R_n(t_i)\right]. \tag{149}$$

In view of equations (147) and (148), we obtain

$$\lim_{n\to\infty} E\left(H_n(t_i)\right) = S_i \tag{150}$$

at the discontinuity t_i of the underlying law of failures $F(t)$.

To obtain the variance of the statistic $H_n(t_i)$, we have

$$\mathrm{Var}\Big(H_n(t_i)\Big) = 4\Big[\mathrm{Var}\Big(R_n^*(t_i)\Big) + \mathrm{Var}\Big(R_n(t_i)\Big) - 2\,\mathrm{Cov}\Big[R_n^*(t_i),\ R_n(t_i)\Big]\Big] \tag{151}$$

Since we know Var $(R_n(t_i))$ as given by equation (50) of Chapter II, we only have to obtain Var $(R_n(t_i))$ and Cov $\left[R_n^*(t_i),\ R_n(t_i)\right]$.

From equation (61) (Chapter II) we have

$$n\ \mathrm{Var}\Big(R_n^*(t)\Big) = E\Big(G^2\Big(B_n(T-t)\Big)\Big) - E^2\Big(G\Big(B_n(T-t)\Big)\Big) \tag{152}$$

Now

$$\begin{aligned} E\Big(G^2\Big(B_n(T-t)\Big)\Big) &= \int_0^\infty G^2\Big(B_n(\tau-t)\Big)\, dF_1(\tau) \\ &\quad + \int_0^\infty G^2\Big(B_n(\tau-t)\Big)\, dF_2(\tau), \\ &= J_1 + J_2, \quad \text{say.} \end{aligned}$$

It is easily seen that

$$\begin{aligned} \lim_{n\to\infty} J_1 &= F_1(\infty) - F_1(t_i), \\ \lim_{n\to\infty} J_2 &= 1/4\ Si + \sum_{t_\nu > t_i} S_\nu, \end{aligned} \tag{153}$$

at the discontinuity t_i of $F(t)$.

Therefore,

$$\begin{aligned} \lim_{n\to\infty} E\Big[G^2\Big(B_n(T-t_i)\Big)\Big] &= F_1(\infty) - F_1(t_i) + 1/4\ S_i + \sum_{t_\nu > t_i} S_\nu \\ &= R(t_i) + 1/4\ S_i\ . \end{aligned} \tag{154}$$

Combining equations (147), (152), and (154) we obtain

$$\lim_{n\to\infty}\left[n\ \mathrm{Var}\left(R_n^*(t_i)\right)\right] = R(t_i) + 1/4\ S_i - \left(R(t_i) + 1/2\ S_i\right)^2, \tag{155}$$

To find the covariance between $R_n(t)$ and $R_n^*(t)$, let us recall that

$$R_n^*(t) = 1/n \sum_{\gamma=1}^{n} G\left(B_n(T_\gamma - t)\right) \tag{156}$$

and

$$\begin{aligned} R_n(t) &= 1/n \begin{bmatrix} \text{number of observations} > t \\ \text{among } T_1,\ T_2, \ldots T_n \end{bmatrix} \\ &= 1/n \sum_{\gamma=1}^{n} U\,(T_\gamma - t), \end{aligned} \tag{157}$$

where

$$\begin{aligned} U(x) &= 1 \text{ for } x > 0 \\ &= 0 \text{ for } x \le 0 \quad . \end{aligned}$$

Now

$$\begin{aligned} \mathrm{Cov}\ \left|R_n^*(t),\ R_n(t)\right| &= 1/n^2 \sum_{\gamma=1}^{n} \mathrm{Cov}\ G\left[\left(B_n(T_\gamma - t),\ U\,(T_\gamma - t)\right)\right] \\ &= 1/n\ \ \mathrm{Cov}\left[G\left(B_n(T-t)\right),\ U(T-t)\right] . \end{aligned} \tag{158}$$

We have

$$\mathrm{Cov}\left[G\left(B_n(T-t)\right),\ U(T-t)\right] = M_1 - M_2 \tag{159}$$

where

$$M_1 = \int_0^\infty U(\tau-t)\, G\Big(B_n(\tau-t)\Big)\, d\, \Big(F_1(\tau) + F_2(\tau)\Big)$$

$$= M_{11} + M_{12}, \text{ say,} \tag{160}$$

and

$$M_2 = E\Big[U(T-t)\Big]\, E\Big[G\Big(B_n(T-t)\Big)\Big]. \tag{161}$$

It can be easily verified that

$$\lim_{n\to\infty} M_{11} = \lim_{n\to\infty} \int_0^\infty U(\tau-t)\, G\Big(B_n(\tau-t)\Big)\, dF_1(\tau) = F_1(\infty) - F_1(t_i), \tag{162}$$

at the discontinuity $t = t_i$ of $F(t)$.

Also,

$$M_{12} = \int_0^\infty U(\tau-t)\, G\Big(B_n(\tau-t)\Big)\, dF_2(\tau)$$

$$= \sum_{t_\nu > t_i} S_\nu\, G\Big(B_n(t_\nu - t_i)\Big),$$

Hence,

$$\lim_{n\to\infty} M_{12} = \sum_{t_\nu > t_i} S_\nu, \tag{163}$$

at the discontinuity $t = t_i$.

Summing up,

$$\lim_{n\to\infty} M_1 = F_1(\infty) - F_1(t_i) + \sum_{t_\nu > t_i} S_\nu = R(t_i) \tag{164}$$

at the discontinuity t_i of $F(t)$.

We have

$$\begin{aligned} E\Big(U(T-t_i)\Big) &= \int_0^\infty U(\tau - t_i)\, d\Big(F_1(\tau) + F_2(\tau)\Big) \\ &= \int_{t_i}^\infty d\, F_1(\tau) + \sum_{t_\nu > t_i} S_\nu \\ &= F_1(\infty) - F_1(t_i) + \sum_{t_\nu > t_i} S_\nu \\ &= R(t_i). \end{aligned} \tag{165}$$

Combining equations (147), (161), and (165), we have

$$\lim_{n\to\infty} M_2 = R(t_i)\Big(R(t_i) + 1/2\, S_i\Big), \tag{166}$$

at the discontinuity t_i.

Finally, combining equations (158), (164), and (166) we discover that

$$\lim_{n\to\infty} \mathrm{Cov}\Big[G\Big(B_n(T-t)\Big),\ U(T-t)\Big] = R(t_i)\Big[1 - R(t_i) - 1/2\, S_i\Big], \tag{167}$$

at the discontinuity t_i.

Combining all the above, we obtain

$$\lim_{n\to\infty} \left[n \operatorname{Var}\left(H_n(t_i)\right)\right] = S_i(1-S_i) \tag{168}$$

at the discontinuity $t = t_i$ of the distribution $F(t)$.

Writing the estimator $H_n(t_i)$ as

$$H_n(t_i) = 1/n \sum_{\gamma=1}^{n} \xi_\gamma,$$

where

$$\xi_\gamma = 2\left[G\left(B_n(T_\gamma - t_i)\right) - U(T_\gamma - t_i)\right], \tag{169}$$

one can easily verify that the sufficient condition equation (72) of Chapter II is satisfied by the sequence $\{\xi_\gamma\}$ of independently and identically distributed random variables. Thus we have proved

<u>Theorem 3</u>

The class of estimators $\{H_n(t_i)\}$ are consistent and asymptotically normal for estimating the Jump S_i corresponding to the discontinuity $t = t_i$ of the underlying law of failures $F(t)$.

Consider now the estimator

$$f_n^*(t_i) = 1/B_n f_n(t_i) \tag{170}$$

and $f_n(t_i)$ is given by equation (25) of Chapter II at the discontinuity t_i of $F(t)$.

Since

$$f_n^*(t_i) = 1/n \sum_{\gamma=1}^{n} K\left(B_n(T_\gamma - t)\right), \tag{171}$$

a straightforward calculation yields that

$$\lim_{n\to\infty} E\left(f_n^*(t_i)\right) = K(0)\, S_i ,$$

$$\lim_{n\to\infty} \left[n \operatorname{Var}\left(f_n^*(t_i)\right)\right] = K^2(0)\, S_i(1-S_i) \tag{172}$$

at the discontinuity $t = t_i$ of the underlying law of failures $F(t)$; and finally, the estimate $f_n^*(t_i)$ is asymptotically normal. Thus, the classes of estimators $H_n(t_i)$ and $\left[\frac{1}{K(0)} f_n^*(t_i)\right]$ are asymptotically equivalent for estimating the jump S_i at the discontinuity t_i of $F(t)$.

Testing the Hypothesis that the Underlying Law of Failures Has no Jump

Let T_1, T_2, .. T_n be the observed times to failures of n identical items put to life testing experiment. Let K(t) be the usual weight function or window as defined by equation (24) of Chapter II. Let B_n be a sequence of nonnegative constants tending to infinity as $n \to \infty$.

Consider the statistic

$$S_n = \frac{1}{\binom{n}{2}} \sum_{i \neq j}^{n} K\Big(B_n(T_i - T_j)\Big) \quad .$$

It can be easily shown that

$$\lim_{n\to\infty} E(S_n) = K(0) \sum_{\nu=1}^{\infty} S_\nu^2 \quad ,$$

i.e., the estimator $\frac{S_n}{K(0)}$ is asymptotically unbiased for estimating the sum of squares of Jumps in the underlying law of failures.

In a similar manner, it can be shown that

$$\lim_{n\to\infty} \Big[n \operatorname{Var}(S_n)\Big] = K^2(0)\, S(1 - S)$$

where

$$S = \sum_{\nu=1}^{\infty} S_\nu^2 \quad .$$

It can also be shown that the sequence of estimators $\frac{S_n}{K(0)}$ is asymptotically normal for estimating S.

Thus we have

<u>Theorem 4</u>

The sequence of estimators $\frac{S_n}{K(0)}$ is consistent and asymptotically normal for estimating the sum of squares of Jumps S in the underlying law of failures.

The above theorem at once yields a large sample test of significance for testing the hypothesis that

$$S = \sum_{\nu=1}^{\infty} S_\nu^2 = 0 ,$$

i.e., whether the underlying law of failures F(t) has Jumps, or not.

CHAPTER VI

ESTIMATION OF THE FAILURE LAW BY RÉNYI-TYPE STATISTICS

In this chapter appear tables of the exact distribution of some Rényi-type statistics. Expressions for the exact and limiting distributions are given. Also discussed in this chapter are the accuracy of the tables and their use for obtaining upper confidence conditions for the unknown distribution function of failures. We also discuss the size of the sample necessary to use the limiting distribution in place of the exact distribution.

Let $T \leq \ldots \leq T_n$ be an ordered sample from a random variable T (time to failure) with continuous distribution function, F. Let F_n be the empirical distribution function of this sample.

If little is known about the distribution function, F, then one seeks an upper confidence contour for F, i.e., a function $G_n(s)$, depending on the sample, such that the assertion $F(s) \leq G_n(s)$ can be made on a preassigned confidence level, for every $s \geq 0$ or at least within some meaningful range. The two statistics that will be considered are

$$D_1 = \sup_{F_n(t) \geq a} \{F_n(t) - F(t)\}$$

and

$$D_2 = \sup_{F_n(t) \geq a} \frac{F_n(t) - F(t)}{F_n(t)} .$$

The exact distribution of each of these statistics is known from [3] and the limiting distributions were developed in [7] and [10].

The exact distributions of the statistics, D_1 and D_2, are given by

$$P_1(n,a,c) = P[\sup_{F_n(t)\geq a} \{F_n(t)-F(t)\} < c] = \tag{173}$$

$$= 1-c \sum_{j=0}^{k} \binom{n}{j} \left(\frac{j}{n} + c\right)^{j-1} (1-c-j/n)^{n-j} \quad ,$$

$$\text{where } k = \Big([n(1-a)] - 1\Big)^{+}$$

and by

$$P_2(n,a,c) = P[\sup_{F_n(t)\geq a} \frac{F_n(t)-F(t)}{F_n(t)} < c] = \tag{174}$$

$$= 1-c \sum_{j=0}^{k} \binom{n}{j} \left(\frac{j}{n}(1-c)+c\right)^{j-1} \Big((1-c)(1-j/n)\Big)^{n-j} \quad ,$$

$$\text{where } k = \Big([n(1-a)] - 1\Big)^{+} \quad .$$

The notation $(d)^+$ is defined by $(d)^+ = \begin{cases} d & \text{if } d \geq 0 \\ 0 & \text{if } d < 0 \end{cases}.$

The limiting distributions of the statistics are given by

$$\lim_{n\to\infty} P[\sqrt{n} \sup_{F_n(t)\geq a} \{F_n(t)-F(t)\} < c] = \tag{175}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c_1} e^{-t^2/2} dt - \frac{e^{-2c^2}}{\sqrt{2\pi}} \int_{-\infty}^{c_3} e^{-t^2/2} dt$$

where $c_1 = \frac{c}{\sqrt{a(1-a)}}$ and $c_3 = \frac{c-2ac}{\sqrt{a(1-a)}}$

and by

$$\lim_{n\to\infty} P\left[\sqrt{n} \sup_{F_n(t)\geq a} \frac{F_n(t)-F(t)}{F_n(t)} < c\right] = \tag{176}$$

$$= \frac{2}{\sqrt{2\pi}} \int_0^{c\,(a/1-a)^{1/2}} e^{-t^2/2}\, dt.$$

The following is a brief description of the computations leading to the tables. The computations for both tables were done by an IBM 360 computer system using the expressions (173) and (174) for the respective statistics. The programs were written in FORTRAN IV language in double precision. Computations of the exact distributions of both statistics were done for the following values of n,a, and c:

$n=5(5)50$, $a=.10(.10).80$, and $c=.05(.05).90$.

Due to double precision, the probabilities are accurate to five decimal places. Since the limiting distributions are much easier to calculate than the exact distributions, it clearly is useful for practical reasons to know at what level of sample size the exact and limiting distributions differ by less than, say, .001. Additional computations using expressions (175) and (176) have indicated that this is the case for n greater than 40 for the statistic D_1 and for n greater than 50 for the statistic D_2.

If the sample $T_1,\ldots T_n$ is censored from above to m observations $(m<n)$, then F is known only for those values of t for which $F_n(t) \le \frac{m}{n}$. While the two statistics, D_1 and D_2, correspond to censoring from below, the following simple replacement in the statistics leads to the case of censoring from above: In both statistics let F(t) be replaced by 1-F(-t), $F_n(t)$ by $1-F_n(-t)$, and -t by t. Then one has:

$$P\left[D_1\le c\right] = P\left[\sup_{F_n(t)\le 1-a} \{F(t)-F_n(t)\} \le c\right] \tag{177}$$

$$P\left[D_2\le c\right] = P\left[\sup_{F_n(t)\le 1-a} \frac{F(t)-F_n(t)}{1-F_n(t)} \le c\right]. \tag{178}$$

Simplifying both (177) and (178) leads to

$$P\left[D_1\le c\right] = P\left[F(t)\le F_n(t) + c \quad \text{for } F_n(t)\le 1-a\right] = \tag{179}$$

$$= P\left[F(t)\le F_n(t) + c \quad \text{for } t\le T_k,\ k=\left(\left[n(1-a)\right]-1\right)^+\right]$$

and

$$P\left[D_2\le c\right] = P\left[F(t)\le F_n(t)(1-c) + c \text{ for } F_n(t)\le 1-a\right] \tag{180}$$

$$= P\left[F(t)\le F_n(t)(1-c) + c \text{ for } t\le T_k,\ k=\left(\left[n(1-a)\right] -1\right)^+\right].$$

Let D denote either of the statistics, D_1 or D_2. In practical situations one is given a sample size n, a censoring level a, $0<a<1$, and a confidence level $1-\alpha$; and for both statistics one seeks the minimum c, denoted c_α, in the range of c given above, so that

$$P\left[D\le c_\alpha\right]\ge 1-\alpha\ . \tag{181}$$

For each statistic this will give an upper confidence contour at the $1-\alpha$ level.

For each statistic, for each n and a in the range given above, and for $\alpha=.10$, .05, .02, and .01, the value of c_α in the range .05(.05).90 appears in the tables along with $P\left[D\le c_\alpha\right]$ and $P\left[D\le c_\alpha - .05\right]$.

Thus, for a given n,a, and α one obtains from Tables I and II the critical values, $c_{\alpha,1}$ and $c_{\alpha,2}$ so that the following inequalities hold:

$$P\left[D_1\le c_{\alpha,1}\right]= P\left[F(t)\le F_n(t)+c_{\alpha,1}\quad \text{for} \quad t\le T_k,\ k=\left(\left[n(1-a)\right]-1\right)^+\right]\ge 1-\alpha \tag{182}$$

and

$$P\left[D_2\le c_{\alpha,2}\right]= P\left[F(t)\le F_n(t)(1-c_{\alpha,2})+c_{\alpha,2}\ \text{for} \quad t\le T_k,\ k=\left(\left[n(1-a)\right]-1\right)^+\right]\ge 1-\alpha\ . \tag{183}$$

For example, for n=10, a=.20, $\alpha=.05$ for the statistic D_1 one finds in Table I the value $c_\alpha=.40$. The corresponding probability, obtained from Table I is $P_1(10, .20, .40) = .970505$. Also, from the same row, one has $P_1(10,.20,.35) =$ $= .933015$. The confidence contour obtained is given by

$$P\left[F(t)\le F_{10}(t)+.40\quad \text{for}\ t\le T_7\right] = .970505\ > .9500 = 1-\alpha\ .$$

For the statistic D_2 one obtains from Table II for n=10, a=.20, and α=.05 the critical value c_α=.75 and the corresponding probabilities $P_2(10,.20,.75) = .962044$, $P_2(10,.20,.70) = .938123$.

This yields an upper confidence contour given by

$$P\left[F(t)\leq F_{10}(t)(.25) + .75 \text{ for } x\leq T_7\right] = .962044 > .9500 = 1-\alpha .$$

TABLE I

	a	c_α	$P_1(n,a,c_\alpha)$	$P_1(n,a,c_\alpha-.05)$
α=.10, n=5				
	.10	.45	.902998	.845440
	.20	.45	.902998	.845440
	.30	.45	.902998	.845440
	.40	.45	.902998	.845440
	.50	.45	.915907	.871040
	.60	.45	.915907	.871040
	.70	.40	.922224	.883971
	.80	.40	.922224	.883971
α=.10, n=10				
	.10	.35	.933013	.864536
	.20	.35	.933013	.864536
	.30	.35	.933013	.864536
	.40	.35	.933369	.868256
	.50	.35	.936865	.878165
	.60	.35	.944435	.893918
	.70	.30	.915152	.849516
	.80	.30	.941519	.891907
α=.10, n=15				
	.10	.30	.946009	.870227
	.20	.30	.946009	.870227
	.30	.30	.946036	.871628
	.40	.30	.946682	.875512
	.50	.30	.952051	.891090
	.60	.25	.902533	.806772
	.70	.25	.932071	.858129
	.80	.25	.949720	.889823
α=.10, n=20				
	.10	.25	.931173	.822814
	.20	.25	.931173	.822814
	.30	.25	.931250	.825257
	.40	.25	.933149	.834565
	.50	.25	.938765	.850587
	.60	.25	.948270	.872754
	.70	.20	.900856	.782405
	.80	.20	.934560	.842035

TABLE I (Continued)

	a	c_α	$P_1(n,a,c_\alpha)$	$P_1(n,a,c_\alpha-.05)$
α=.10, n=25				
	.10	.25	.963496	.881683
	.20	.25	.963496	.881683
	.30	.25	.963546	.883444
	.40	.25	.964240	.888442
	.50	.20	.902474	.765900
	.60	.20	.915859	.793050
	.70	.20	.941222	.843814
	.80	.20	.960763	.885483
α=.10, n=30				
	.10	.20	.920985	.764733
	.20	.20	.920985	.765179
	.30	.20	.921449	.771860
	.40	.20	.924870	.786787
	.50	.20	.943524	.808842
	.60	.20	.944213	.837623
	.70	.20	.959268	.873403
	.80	.15	.916464	.761349
α=.10, n=35				
	.10	.20	.947231	.812358
	.20	.20	.947231	.812569
	.30	.20	.947567	.818474
	.40	.20	.949477	.829111
	.50	.20	.955642	.850795
	.60	.20	.962942	.872285
	.70	.15	.907688	.734762
	.80	.15	.938773	.797235
α=.10, n=40				
	.10	.20	.964758	.850342
	.20	.20	.964758	.850441
	.30	.20	.964849	.853648
	.40	.20	.966056	.863043
	.50	.20	.969509	.878461
	.60	.20	.975350	.899362
	.70	.15	.925244	.759626
	.80	.15	.954956	.827198
α=.10, n=45				
	.10	.20	.976463	.880635
	.20	.20	.976463	.880682
	.30	.20	.976529	.883434
	.40	.20	.977214	.890252
	.50	.15	.904711	.702222
	.60	.15	.920582	.738592
	.70	.15	.944967	.795265
	.80	.15	.966760	.852366

TABLE I (Continued)

	a	c_α	$P_1(n,a,c_\alpha)$	$P_1(n,a,c_\alpha-.05)$
α=.05, n=5				
	.10	.55	.970152	.944000
	.20	.55	.970152	.944000
	.30	.55	.970152	.944000
	.40	.55	.970152	.944000
	.50	.55	.970805	.948500
	.60	.55	.970805	.948500
	.70	.50	.968750	.949671
	.80	.50	.968750	.949671
α=.05, n=10				
	.10	.40	.970505	.933013
	.20	.40	.970505	.933013
	.30	.40	.970505	.933013
	.40	.40	.970505	.933013
	.50	.40	.971167	.936865
	.60	.40	.973919	.944435
	.70	.35	.955852	.915152
	.80	.35	.970418	.941519
α=.05, n=15				
	.10	.35	.981021	.946009
	.20	.35	.981021	.946009
	.30	.35	.981020	.946036
	.40	.35	.981040	.946682
	.50	.30	.952051	.891090
	.60	.30	.957019	.902533
	.70	.30	.970924	.932071
	.80	.30	.979255	.949720
α=.05, n=20				
	.10	.30	.978466	.931173
	.20	.30	.978466	.931173
	.30	.30	.978466	.931251
	.40	.30	.978595	.933149
	.50	.30	.979740	.938765
	.60	.30	.982595	.948270
	.70	.25	.961147	.900856
	.80	.25	.976280	.934560
α=.05, n=25				
	.10	.25	.963496	.881683
	.20	.25	.963496	.881683
	.30	.25	.963546	.883444
	.40	.25	.964240	.888442
	.50	.25	.967970	.902474
	.60	.25	.972484	.915859
	.70	.25	.981731	.941222
	.80	.20	.960763	.885483

TABLE I (Continued)

	a	c_α	$P_1(n,a,c_\alpha)$	$P_1(n,a,c_\alpha-.05)$
α=.05, n=30				
	.10	.25	.980638	.920985
	.20	.25	.980638	.920985
	.30	.25	.980638	.920985
	.40	.25	.980920	.924880
	.50	.25	.982342	.932524
	.60	.25	.985348	.944213
	.70	.20	.959268	.873403
	.80	.20	.976313	.916464
α=.05, n=35				
	.10	.25	.989730	.947231
	.20	.25	.989730	.947231
	.30	.25	.989732	.947567
	.40	.25	.989838	.949477
	.50	.20	.955642	.850795
	.60	.20	.962942	.872285
	.70	.20	.975411	.907688
	.80	.20	.985626	.938773
α-.05, n=40				
	.10	.20	.964758	.850342
	.20	.20	.964758	.850342
	.30	.20	.974849	.853648
	.40	.20	.966056	.863043
	.50	.20	.969509	.878461
	.60	.20	.975350	.899362
	.70	.20	.983014	.925244
	.80	.15	.954956	.827198
α=.05, n=45				
	.10	.20	.976463	.880635
	.20	.20	.976463	.880682
	.30	.20	.976529	.883434
	.40	.20	.977214	.890252
	.50	.20	.979876	.904712
	.60	.20	.983587	.920582
	.70	.20	.989634	.944967
	.80	.15	.966760	.852366

TABLE I (Continued)

	a	c_α	$P_1(n,a,c_\alpha)$	$P_1(n,a,c_\alpha -.05)$
α=.02, n=5				
	.10	.60	.984960	.970152
	.20	.60	.984960	.970152
	.30	.60	.984960	.970152
	.40	.60	.984960	.970152
	.50	.60	.984960	.970805
	.60	.60	.984960	.970805
	.70	.55	.981547	.968750
	.80	.55	.981547	.968750
α=.02, n=10				
	.10	.45	.988554	.970505
	.20	.45	.988554	.970505
	.30	.45	.988554	.970505
	.40	.45	.988554	.970505
	.50	.45	.988554	.971167
	.60	.45	.989244	.973919
	.70	.45	.991098	.979063
	.80	.40	.986141	.970418
α=.02, n=15				
	.10	.35	.981021	.946009
	.20	.35	.981021	.946009
	.30	.35	.981021	.946036
	.40	.35	.981034	.946682
	.50	.35	.982170	.952051
	.60	.35	.983756	.957019
	.70	.35	.989022	.970924
	.80	.35	.992344	.979255
α=.02, n=20				
	.10	.35	.994622	.978466
	.20	.35	.994622	.978466
	.30	.35	.994622	.978466
	.40	.35	.994623	.978595
	.50	.35	.994730	.979740
	.60	.30	.982595	.948270
	.70	.30	.987096	.961147
	.80	.30	.992543	.976279
α=.02, n=25				
	.10	.30	.991411	.963496
	.20	.30	.991411	.963496
	.30	.30	.991411	.963546
	.40	.30	.991436	.964240
	.50	.30	.991966	.967970
	.60	.30	.992958	.972485
	.70	.25	.981731	.941222
	.80	.25	.988701	.960763

TABLE I (Continued)

	a	c_α	$P_1(n,a,c_\alpha)$	$P_1(n,a,c_\alpha-.05)$
α=.02, n=30				
	.10	.25	.980638	.920985
	.20	.25	.980638	.920985
	.30	.25	.980640	.921449
	.40	.25	.980920	.924880
	.50	.25	.982342	.932524
	.60	.25	.985348	.944213
	.70	.25	.989701	.959268
	.80	.25	.994580	.976313
α=.02, n=35				
	.10	.25	.989730	.947231
	.20	.25	.989730	.947231
	.30	.25	.989732	.947567
	.40	.25	.989838	.949477
	.50	.25	.990714	.955642
	.60	.25	.992192	.962942
	.70	.25	.995063	.975411
	.80	.20	.985626	.938773
α=.02, n=40				
	.10	.25	.994552	.964758
	.20	.25	.994552	.964758
	.30	.25	.994552	.964849
	.40	.25	.994594	.966056
	.50	.25	.994937	.969509
	.60	.25	.995838	.975350
	.70	.20	.983014	.925244
	.80	.20	.991241	.954956
α=.02, n=45				
	.10	.25	.997110	.976463
	.20	.25	.997110	.976463
	.30	.25	.997110	.976529
	.40	.25	.997126	.977214
	.50	.25	.997330	.979876
	.60	.20	.983587	.920582
	.70	.20	.989634	.944967
	.80	.20	.994646	.966760

PRECEDING PAGE BLANK NOT FILMED.

TABLE I (Continued)

	a	c_α	$P_1(n,a,c_\alpha)$	$P_1(n,a,c_\alpha-.05)$
α=.01, n=30				
	.10	.30	.996574	.980638
	.20	.30	.996574	.980638
	.30	.30	.996574	.980640
	.40	.30	.996579	.980920
	.50	.30	.996697	.982342
	.60	.30	.997154	.985348
	.70	.30	.998001	.989701
	.80	.25	.994580	.976313
α=.01, n=35				
	.10	.30	.998633	.989730
	.20	.30	.998633	.989730
	.30	.30	.998633	.989732
	.40	.30	.998634	.989838
	.50	.25	.990714	.955642
	.60	.25	.992192	.962942
	.70	.25	.995063	.975411
	.80	.25	.997387	.985626
α=.01, n=40				
	.10	.25	.994552	.964758
	.20	.25	.994552	.964758
	.30	.25	.994552	.964849
	.40	.25	.994594	.966056
	.50	.25	.994937	.969509
	.60	.25	.995838	.975350
	.70	.25	.997235	.983014
	.80	.20	.991241	.954956
α=.01, n=45				
	.10	.25	.997110	.976463
	.20	.25	.997110	.976463
	.30	.25	.997110	.976529
	.40	.25	.997126	.977214
	.50	.25	.997330	.979876
	.60	.25	.997781	.983587
	.70	.25	.998659	.989634
	.80	.20	.994646	.966760

TABLE II

	a	c_α	$P_2(n,a,c_\alpha)$	$P_2(n,a,c_\alpha-.05)$
α=.10, n=5				
	.10	.75	.910758	.874423
	.20	.75	.910758	.874423
	.30	.65	.927216	.895265
	.40	.65	.927216	.895265
	.50	.50	.904750	.865339
	.60	.50	.904750	.865339
	.70	.40	.922240	.883971
	.80	.40	.922240	.883971
α=.10, n=10				
	.10	.75	.903122	.864586
	.20	.65	.907319	.869529
	.30	.60	.927309	.892857
	.40	.50	.904894	.861918
	.50	.45	.911194	.866802
	.60	.40	.914735	.867673
	.70	.35	.918037	.867401
	.80	.30	.924851	.870963
α=.10, n=15				
	.10	.65	.901136	.861726
	.20	.60	.917586	.880166
	.30	.50	.923233	.883215
	.40	.45	.918203	.873833
	.50	.40	.936651	.894112
	.60	.30	.924993	.873326
	.70	.30	.939817	.887618
	.80	.25	.927235	.859496
α=.10, n=20				
	.10	.70	.931302	.898124
	.20	.55	.916571	.876754
	.30	.45	.903903	.855207
	.40	.40	.915308	.864764
	.50	.35	.918447	.862796
	.60	.30	.916341	.851632
	.70	.25	.910792	.831626
	.80	.20	.904768	.803238
α=.10, n=25				
	.10	.60	.910059	.870518
	.20	.50	.907360	.862590
	.30	.40	.902838	.848362
	.40	.40	.942353	.899795
	.50	.30	.909015	.840222
	.60	.30	.943948	.889383
	.70	.25	.948012	.885179
	.80	.20	.934598	.845578

TABLE II (Continued)

	a	c_α	$P_2(n,a,c_\alpha)$	$P_2(n,a,c_\alpha-.05)$
α=.10, n=30				
	.10	.60	.908212	.868174
	.20	.50	.929810	.890122
	.30	.40	.916158	.864584
	.40	.35	.925087	.869812
	.50	.30	.924601	.860314
	.60	.25	.916786	.835993
	.70	.20	.901457	.791727
	.80	.20	.954656	.877856
α=.10, n=35				
	.10	.55	.907848	.865579
	.20	.45	.911647	.863566
	.30	.40	.942496	.899052
	.40	.35	.943617	.894928
	.50	.30	.948847	.894838
	.60	.25	.936977	.864543
	.70	.20	.931165	.834699
	.80	.15	.902808	.748228
α=.10, n=40				
	.10	.55	.906418	.863767
	.20	.45	.928641	.884791
	.30	.35	.909502	.848473
	.40	.30	.914798	.845559
	.50	.25	.907631	.821147
	.60	.25	.952016	.887615
	.70	.20	.940771	.848939
	.80	.15	.922296	.776908
α=.10, n=45				
	.10	.55	.933717	.897158
	.20	.40	.902394	.846514
	.30	.35	.931718	.877530
	.40	.30	.930646	.866920
	.50	.25	.929615	.851692
	.60	.20	.906416	.794210
	.70	.20	.958029	.878686
	.80	.15	.937637	.801746
α=.10, n=50				
	.10	.55	.932785	.895906
	.20	.40	.917080	.864728
	.30	.35	.938614	.886819
	.40	.30	.943368	.885009
	.50	.25	.937969	.863701
	.60	.20	.921834	.816069
	.70	.20	.963801	.888918
	.80	.15	.949790	.823390

TABLE II (Continued)

	a	c_α	$P_2(n,a,c_\alpha)$	$P_2(n,a,c_\alpha-.05)$
α=.05, n=5				
	.10	.80	.952246	.910758
	.20	.80	.952246	.910758
	.30	.70	.952482	.927216
	.40	.70	.952482	.927216
	.50	.60	.958303	.935358
	.60	.60	.958303	.935358
	.70	.50	.968750	.949672
	.80	.50	.968750	.949672
α=.05, n=10				
	.10	.80	.973956	.945877
	.20	.75	.962044	.938123
	.30	.65	.953728	.927309
	.40	.60	.961843	.937808
	.50	.55	.966641	.943828
	.60	.50	.970083	.947800
	.70	.40	.951960	.918037
	.80	.35	.958452	.924851
α=.05, n=15				
	.10	.75	.958954	.933549
	.20	.70	.968381	.946795
	.30	.55	.952572	.923233
	.40	.50	.950103	.918203
	.50	.45	.964616	.936651
	.60	.40	.958483	.924993
	.70	.35	.970060	.939817
	.80	.30	.964953	.927235
α=.05, n=20				
	.10	.75	.957422	.931302
	.20	.65	.968828	.946945
	.30	.55	.965060	.939926
	.40	.45	.950392	.915308
	.50	.40	.954930	.918447
	.60	.35	.956414	.916341
	.70	.30	.956626	.910792
	.80	.25	.958021	.904768
α=.05, n=25				
	.10	.70	.964749	.941327
	.20	.60	.965405	.941224
	.30	.50	.965622	.941692
	.40	.45	.969453	.942353
	.50	.35	.952327	.909015
	.60	.35	.974264	.943948
	.70	.30	.978949	.911801
	.80	.25	.975411	.934598

TABLE II (Continued)

	a	c_α	$P_2(n,a,c_\alpha)$	$P_2(n,a,c_\alpha-.05)$
α=.05, n=30				
	.10	.70	.963840	.939972
	.20	.55	.958233	.929810
	.30	.45	.951725	.916158
	.40	.40	.960372	.925087
	.50	.35	.962992	.924601
	.60	.30	.962061	.916786
	.70	.25	.958727	.901457
	.80	.20	.954656	.877856
α=.05, n=35				
	.10	.65	.964611	.940609
	.20	.55	.970125	.946495
	.30	.45	.969998	.942496
	.40	.40	.972557	.943617
	.50	.35	.977809	.948847
	.60	.30	.974124	.936977
	.70	.25	.975268	.931165
	.80	.20	.968336	.902808
α=.05, n=40				
	.10	.65	.963905	.939559
	.20	.50	.959016	.928641
	.30	.40	.950168	.909502
	.40	.35	.957339	.914798
	.50	.30	.957385	.907631
	.60	.25	.952016	.887615
	.70	.25	.980308	.940771
	.80	.20	.977764	.922296
α=.05, n=45				
	.10	.60	.960332	.933717
	.20	.50	.968484	.942157
	.30	.40	.965362	.931718
	.40	.35	.967584	.930646
	.50	.25	.970689	.929615
	.60	.25	.963310	.906416
	.70	.20	.958029	.878686
	.80	.20	.984314	.937637
α=.05, n=50				
	.10	.60	.959693	.932785
	.20	.45	.952971	.917080
	.30	.40	.969891	.938614
	.40	.35	.975283	.943368
	.50	.30	.975488	.937969
	.60	.25	.971847	.921834
	.70	.20	.963801	.888918
	.80	.20	.988893	.949790

TABLE II (Continued)

	a	c_α	$P_2(n,a,c_\alpha)$	$P_2(n,a,c_\alpha-.05)$
α=.02, n=5				
	.10	.85	.981385	.952246
	.20	.85	.981385	.952246
	.30	.80	.984893	.971408
	.40	.80	.984893	.971508
	.50	.70	.985958	.974771
	.60	.70	.985958	.974771
	.70	.55	.981547	.968750
	.80	.55	.981547	.968750
α=.02, n=10				
	.10	.85	.980235	.971879
	.20	.85	.982748	.973956
	.30	.75	.985828	.972898
	.40	.70	.988984	.978396
	.50	.60	.981661	.966641
	.60	.55	.984048	.970083
	.70	.50	.986393	.973507
	.80	.45	.985482	.978336
α=.02, n=15				
	.10	.80	.984328	.958954
	.20	.75	.983217	.968381
	.30	.65	.985883	.972849
	.40	.60	.985338	.971711
	.50	.50	.981762	.964616
	.60	.50	.990041	.978728
	.70	.40	.986281	.970060
	.80	.35	.984402	.964953
α=.02, n=20				
	.10	.80	.982607	.957422
	.20	.70	.983486	.968828
	.30	.60	.981405	.965060
	.40	.55	.986774	.973145
	.50	.50	.989461	.977107
	.60	.45	.991055	.979225
	.70	.35	.980803	.956626
	.80	.30	.983230	.958021
α=.02, n=25				
	.10	.75	.981075	.964749
	.20	.65	.981485	.965405
	.30	.55	.983624	.967622
	.40	.50	.985307	.969453
	.50	.45	.990277	.977271
	.60	.40	.989422	.974264
	.70	.35	.992451	.978949
	.80	.30	.991839	.975411

TABLE II (Continued)

	a	c_α	$P_2(n,a,c_\alpha)$	$P_2(n,a,c_\alpha-.05)$
α=.02, n=30				
	.10	.75	.982949	.970413
	.20	.65	.988893	.977233
	.30	.55	.987852	.974495
	.40	.45	.980987	.960372
	.50	.40	.983692	.962992
	.60	.35	.984632	.962061
	.70	.30	.984826	.958727
	.80	.25	.985441	.954656
α=.02, n=35				
	.10	.70	.980974	.964611
	.20	.60	.984911	.970125
	.30	.50	.985892	.969998
	.40	.45	.988072	.972557
	.50	.40	.991543	.977809
	.60	.35	.990746	.974124
	.70	.30	.992404	.975268
	.80	.25	.991309	.968336
α=.02, n=40				
	.10	.70	.980548	.963905
	.20	.60	.989944	.978519
	.30	.50	.988832	.975046
	.40	.40	.980885	.957339
	.50	.35	.982673	.957385
	.60	.30	.982247	.952016
	.70	.25	.980308	.940771
	.80	.25	.994779	.977764
α=.02, n=45				
	.10	.70	.989597	.978393
	.20	.50	.987732	.972966
	.30	.45	.984264	.965362
	.40	.40	.986624	.967584
	.50	.35	.989452	.970689
	.60	.30	.987762	.963310
	.70	.25	.988005	.958029
	.80	.20	.984314	.937637
α=.02, n=50				
	.10	.65	.989383	.977999
	.20	.55	.988762	.975686
	.30	.45	.986868	.969891
	.40	.40	.990604	.975283
	.50	.35	.991730	.975488
	.60	.30	.991531	.971847
	.70	.25	.990424	.963801
	.80	.20	.988893	.949790

TABLE II (Continued)

	a	c_α	$P_2(n,a,c_\alpha)$	$P_2(n,a,c_\alpha-.05)$
α=.01, n=5				
	.10	.90	.990083	.981385
	.20	.90	.990083	.981385
	.30	.85	.991471	.984893
	.40	.85	.991471	.984893
	.50	.75	.993023	.985958
	.60	.75	.993023	.985958
	.70	.65	.994748	.989760
	.80	.65	.994748	.989760
α=.01, n=10				
	.10	.90	.990927	.980235
	.20	.90	.990927	.980235
	.30	.80	.993711	.985828
	.40	.75	.995128	.988984
	.50	.65	.990850	.981661
	.60	.60	.992230	.984048
	.70	.55	.993579	.986393
	.80	.50	.995240	.989438
α=.01, n=15				
	.10	.90	.995244	.984328
	.20	.80	.992436	.983217
	.30	.70	.993525	.985883
	.40	.65	.993213	.985338
	.50	.55	.991460	.981762
	.60	.50	.990041	.978728
	.70	.45	.994274	.986281
	.80	.40	.993642	.984402
α=.01, n=20				
	.10	.85	.995824	.982607
	.20	.75	.992405	.983486
	.30	.65	.991535	.981405
	.40	.60	.994199	.986774
	.50	.55	.995687	.989461
	.60	.45	.991055	.979225
	.70	.40	.992347	.980803
	.80	.35	.993977	.983230
α=.01, n=25				
	.10	.80	.991370	.981075
	.20	.70	.991249	.981485
	.30	.60	.992616	.983624
	.40	.55	.993708	.985307
	.50	.45	.990277	.977271
	.60	.45	.996168	.989422
	.70	.35	.992451	.978949
	.80	.30	.991839	.975411

TABLE II (Continued)

	a	c_α	$P_2(n,a,c_\alpha)$	$P_2(n,a,c_\alpha-.05)$
α=.01, n=30				
	.10	.80	.991278	.982949
	.20	.70	.995314	.988893
	.30	.60	.994906	.987852
	.40	.50	.991869	.980987
	.50	.45	.993652	.983692
	.60	.40	.994548	.984632
	.70	.35	.995159	.984826
	.80	.30	.995981	.985441
α=.01, n=35				
	.10	.75	.990114	.980975
	.20	.65	.993288	.984911
	.30	.55	.994146	.985892
	.40	.50	.995461	.988072
	.50	.40	.991543	.977809
	.60	.35	.990746	.974124
	.70	.30	.992404	.975268
	.80	.25	.991309	.968336
α=.01, n=40				
	.10	.75	.990892	.980548
	.20	.65	.995920	.989944
	.30	.55	.995635	.988832
	.40	.45	.992470	.980885
	.50	.40	.993895	.982673
	.60	.35	.994392	.982247
	.70	.30	.994507	.980308
	.80	.25	.994779	.977764
α=.01, n=45				
	.10	.75	.995770	.989597
	.20	.60	.993268	.984490
	.30	.50	.993725	.984264
	.40	.45	.995222	.986624
	.50	.40	.996782	.989452
	.60	.35	.996584	.987762
	.70	.30	.997200	.988005
	.80	.25	.996848	.984314
α=.01, n=50				
	.10	.75	.995673	.989384
	.20	.60	.995477	.988762
	.30	.50	.995016	.98686
	.40	.40	.990604	.975283
	.50	.35	.991730	.975488
	.60	.30	.991531	.971847
	.70	.25	.990424	.963801
	.80	.25	.998089	.988893

CHAPTER VII

FAILURE DISTRIBUTIONS WITH DECREASING MEAN RESIDUAL LIFE

Summary

The study of distributions with decreasing mean residual life (DMR) has received little attention in the literature (Barlow, Marshall and Proschan [2], Watson and Wells [11]). It is well known that this class of distributions contains the class with increasing hazard rate (IHR), which is studied in the literature in considerable detail. In this chapter, starting with a DMR distribution, a sequence of distributions is constructed that preserves the DMR property. It is further shown that this sequence of distributions converges to a stable limit that has very interesting properties. It also turns out that the only distribution which exactly reproduces itself in this sequence is the exponential distribution, which thus may be looked upon as the boundary distribution between DMR and IMR (increasing mean residual life) distributions. It is believed that several inequalities derived under the IHR assumption could be derived under the weaker assumption of the DMR property. As an illustration, an inequality which was derived previously under the IHR assumption is shown to be true for any arbitrary failure distribution.

Introduction and Notation

Let $T_0 \geq 0$ be a nonnegative random variable with $P[T_0 \leq t] = F_0(t) = \int_0^t f_0(t)dt$ where $F_0(t)$ and $f_0(t)$ are respectively the distribution function and

Clearly

$$F_1(t) = 1-R_1(t) \tag{190}$$

is a distribution function of a random variable, say, T_1 which is induced by T_0. Let us write

$$F_1(t) = 1-e^{-\int_0^t \eta_1(x)dx} \tag{191}$$

where, in accordance with (184), $\eta_1(t)$ is the hazard rate associated with T_1. Now

$$\frac{\mu_0(t)}{\mu_0(0)} = e^{-\int_0^t \eta_1(x)dx} . \tag{192}$$

Therefore

$$\log \frac{\mu_0(t)}{\mu_0(0)} = -\int_0^t \eta_1(x)dx ,$$

Differentiating with respect to t both sides, we get

$$\frac{\mu_0(0)}{\mu_0(t)} \frac{\mu_0'(t)}{\mu_0(0)} = \frac{\mu_0'(t)}{\mu_0(t)} = -\eta_1(t) ,$$

where differentiation with respect to t is denoted by a prime.

$$\eta_1(t) = \frac{R_0(t)}{\mu_0(t)} = \frac{1}{\nu_1(t)} . \tag{193}$$

Clearly

$$F_1(t) = 1-R_1(t) \tag{190}$$

is a distribution function of a random variable, say, T_1 which is induced by T_0. Let us write

$$F_1(t) = 1-e^{-\int_0^t \eta_1(x)dx} \tag{191}$$

where, in accordance with (184), $\eta_1(t)$ is the hazard rate associated with T_1. Now

$$\frac{\mu_0(t)}{\mu_0(0)} = e^{-\int_0^t \eta_1(x)dx} . \tag{192}$$

Therefore

$$\log \frac{\mu_0(t)}{\mu_0(0)} = -\int_0^t \eta_1(x)dx ,$$

Differentiating with respect to t both sides, we get

$$\frac{\mu_0(0)}{\mu_0(t)} \frac{\mu_0'(t)}{\mu_0(0)} = \frac{\mu_0'(t)}{\mu_0(t)} = -\eta_1(t) ,$$

where differentiation with respect to t is denoted by a prime.

$$\eta_1(t) = \frac{R_0(t)}{\mu_0(t)} = \frac{1}{\nu_1(t)} . \tag{193}$$

Thus the hazard rate $\eta_1(t)$ of the random variable T_1 is the reciprocal of the mean residual life of the given random variable T_0. For consistency of concept we define

$$\nu_0(t) = \frac{1}{Z_0(t)} \text{ , which implies}$$

$$\eta_0(t) = Z_0(t) \; . \tag{194}$$

Generalizing (189) we define the distribution function $F_k(t)$ of the random variable T_K (induced by the given random variable T_0), by

$$F_K(t) = 1-R_K(t) = 1- \frac{\mu_{K-1}(t)}{\mu_{K-1}(0)} \tag{195}$$

where

$$\mu_{K-1}(t) = \int_t^\infty R_{K-1}(x)dx \; , \tag{196}$$

$$K = 1,2,3,\ldots\ldots \; ,$$

and

$$\mu_{K-1}(0) = E(T_{K-1}) \; . \tag{197}$$

Writing

$$F_K(t) = 1-e^{-\int_0^t \eta_K(x)dx} \; , \tag{198}$$

one obtains for the hazard rate $\eta_K(t)$ associated with T_K

$$\eta_K(t) = \frac{-\mu'_{K-1}(t)}{\mu_{K-1}(t)} = \frac{R_{K-1}(t)}{\mu_{K-1}(t)} , \tag{199}$$

which shows that the mean residual life $\nu_K(t)$ of the random variable T_{K-1} is related to $\eta_K(t)$ by

$$\nu_K(t) = \frac{\mu_{K-1}(t)}{R_{K-1}(t)} = \frac{\int_t^\infty R_{K-1}(x)dx}{R_{K-1}(t)} = \frac{1}{\eta_K(t)} . \tag{200}$$

Assuming that the density exists, differentiating (195) with respect to t, one one obtains for $f_K(t)$ the probability density function of T_K that

$$f_K(t) = \frac{R_{K-1}(t)}{\mu_{K-1}(0)} , \qquad K=1,2,\ldots . \tag{201}$$

Equation (201) shows that $f_k(t)$ is a decreasing function of t for all $K = 1,2,3,\ldots$. Starting from the given random variable T_0 we have thus generated a sequence $\{T_K\}$ $K=1,2,\ldots$ of random variables, whose distribution functions and associated properties have the representations discussed above.

<u>Properties of the Class of Distributions $F_K(t)$</u>

In this section we will prove that if the given random variable T_0 has a distribution with DMR property, then the induced sequence of random variables T_K, $K=1,2,\ldots,\infty$, will all have distributions with the DMR property. To prove this we need the following Lemma.

Lemma 5

$$\nu_K'(t)-\nu_K(t)\eta_{K-1}(t)+1 = 0, \qquad K=1,2,3,\ldots . \tag{202}$$

Proof

By definition from (200),

$$\nu_K(t)R_{K-1}(t) = \mu_{K-1}(t) = \int_t^\infty R_{K-1}(x)dx .$$

Differentiating both sides with respect to t, we obtain

$$\nu_K'(t)R_{K-1}(t)-\nu_K(t)f_{K-1}(t) = -R_{K-1}(t),$$

which implies

$$\nu_K'(t)-\nu_K(t)\eta_{K-1}(t)+1 = 0.$$

Putting K=1 in the lemma, one obtains

$$\nu_1'(t)-\nu_1(t)Z_0(t)+1 = 0, \text{ where} \tag{203}$$

$Z_0(t)$ and $\nu_1(t)$ are respectively the hazard rate and mean residual life of the given random variable T_0. Equation (203) was obtained by W. R. Knight [6] . We will now prove the following:

Theorem 6

$F_0(t)$ is DMR implies $F_K(t)$ is DMR, $K = 1,2,\ldots,\infty$.

Proof

We will prove the theorem by showing that $\gamma_{K-1}(t)$ is decreasing in t will imply that $\nu_K(t)$ is decreasing in t which, combined with the assumption $\nu_1(t)$ is decreasing in t, proves the theorem.

We have from (200) that

$$\nu_K(t) = \int_t^\infty R_{K-1}(x)dx/R_{K-1}(t) \ . \tag{204}$$

Multiplying both sides by $R_{K-2}(t)$,

$$\nu_K(t)R_{K-2}(t) = \frac{R_{K-2}(t)}{R_{K-1}(t)} \int_t^\infty \frac{R_{K-1}(x)}{R_{K-2}(x)} R_{K-2}(x)dx \ . \tag{205}$$

By definition from (199) we have

$$\eta_{K-1}(t) = \frac{R_{K-2}(t)}{\int_t^\infty R_{K-2}(x)dx} = \frac{R_{K-2}(t)}{\mu_{K-2}(0)R_{K-1}(t)} , \tag{206}$$

since

$$\mu_{K-2}(t) = \int_t^\infty R_{K-2}(x)dx = \int_t^\infty \mu_{K-2}(0)f_{K-1}(t)dt$$

from equation (201).

Substituting from (206) in (205) we have

$$\nu_K(t)R_{K-2}(t) = \mu_{K-2}(0)\eta_{K-1}(t)\int_t^\infty \frac{1}{\mu_{K-2}(0)\eta_{K-1}(x)} R_{K-2}(x)dx \tag{207}$$

$$= \eta_{K-1}(t)\int_t^\infty \frac{1}{\eta_{K-1}(x)} R_{K-2}(x)dx$$

$$= \frac{1}{\nu_{K-1}(t)}\int_t^\infty \nu_{K-1}(x)R_{K-2}(x)dx \quad .$$

Since $\nu_{K-1}(t)$ is assumed to be decreasing in t, equation (207) becomes

$$\nu_K(t)R_{K-2} \leq \int_t^\infty R_{K-2}(x)dx \ ,$$

which implies from (199) that

$$\nu_K(t)\eta_{K-1}(t) \leq 1 \ . \tag{208}$$

Now from Lemma 5,

$$\nu_K(t)\eta_{K-1}(t) = 1+\nu_K'(t) \ . \tag{209}$$

Combining (208) and (209, we discover that

$$\nu_K'(t) \leq 0 \ , \tag{210}$$

which implies that $\nu_K(t)$ is decreasing in t.

Convergence of the Sequence of Distribution $\{F_K(t)\}$

We will first prove that the sequence of mean residual times $\{\gamma_K(t)\}$ associated with the sequence of distributions $\{F_K(t)\}$ forms a monotonic decreasing sequence when the given distribution $F_0(t)$ has the DMR property. From equation (202 we obtain

$$\nu_K'(t)-\nu_K(t)\,\frac{1}{\nu_{K-1}}\,(t)+1 = 0\ ,$$

which reduces to

$$\nu_K'(t) = \frac{\nu_K(t)-\nu_{K-1}(t)}{\nu_{K-1}(t)}\ . \tag{211}$$

Since by inequality (211) $\gamma_K'(t) \leq 0$, in view of equation (211) we discover that

$$\nu_K(t) \leq \nu_{K-1}(t)\ , \qquad K = 1,2,\ldots,\infty\ . \tag{212}$$

Now from equation (198),

$$F_K(t) = 1-e^{-\int_0 \frac{1}{K(x)}dx}\ . \tag{213}$$

From (212) we have

$$\int_0^t \frac{1}{\nu_K(x)}\,dx \geq \int_0^t \frac{1}{\nu_{(K-1}(x)}\,dx\ ,$$

which implies that

$$F_K(t) \geq F_{K-1}(t), \qquad K = 1,2,\ldots,\infty . \tag{214}$$

Thus the sequence of distributions $\{F_K(t)\}$ is a monotonic increasing sequence, being uniformly bounded by unity; the sequence therefore converges to a limiting function $F_\infty(t)$ which is a distribution function by a well-known theorem of Helly.

Defining the characteristic function of the random variable T_K by

$$\Phi_K(u) = E(e^{iT_K u}) = \int_0^\infty e^{ix}u\; f_K(x)dx \tag{215}$$

and using the relation (201) one can deduce the following recursive relation satisfied by the sequence of characteristic functions $\{\Phi_{K-1}(u)\}$:

$$iu\; \mu_{K-1}(0)\Phi_K(u) = \Phi_{K-1}(u)-1 . \tag{216}$$

From (214) it follows that

$$R_K(t) \leq R_{K-1}(t) \quad ,$$

and therefore

$$0 \leq \mu_K(0) = \int_0^\infty R_K(x)dx \leq \mu_{K-1}(0) . \tag{217}$$

Equation (217) implies that the sequence $\{\mu_K(0)\}$ of means of the random variables $\{T_K\}$ converges to a limit denoted by μ_∞.

In view of (216) and the existence of μ_∞ and the continuity theorem on characteristic functions [5,p.54], it follows that the limiting characteristic function $\Phi_\infty(u)$ is given by

$$\Phi_\infty(u) = \frac{1}{1-iu\ \mu_\infty}\ , \tag{218}$$

which is the characteristic function of an exponential random variable with mean μ_∞. This limiting distribution will degenerate to a singular distribution when $\mu_\infty = 0$. Thus we have proved the following:

Theorem 7

The sequence of distributions $\{F_K(t)\}$ $K = 1,2,\ldots\infty$ generated from a given DMR distribution $F_0(t)$ converges either to a singular distribution or to an exponential distribution.

Examples

Example 1

Let $F_0(x)$ be an exponential distribution given by

$$F_0(x) = 1-e^{-\lambda x}\ . \tag{219}$$

From (195) and (196) it follows that

$$F_K(t) = \frac{\int_0^t R_{K-1}(x)dx}{\mu_{K-1}(0)} = \frac{\int_0^t (1-F_{K-1}(x))dx}{\mu_{K-1}(0)}\ . \tag{220}$$

In this case

$$F_1(x) = 1-e^{-\lambda x} = F_0(x)$$

and also

$$F_K(x) = F_0(x) \quad \text{for all } K.$$

We will now characterize the exponential distribution by this property.

Suppose any two successive members of the sequence $\{F_K(t)\}$ are identical, i.e.,

$$F_K(t) = F_{K-1}(t) \text{ for some } K \geq 1 . \tag{221}$$

Equation (221) implies that

$$\mu_K(0) = E(T_K) = \mu_{K-1}(0) = E(T_{K-1}) . \tag{222}$$

Using (220), (221) and (222) we obtain

$$F_K(t) = \int_0^t (1-F_K(x)dx/\mu_K(0) . \tag{223}$$

Differentiating equation (223) we obtain that

$$\frac{f_K(t)}{1-F_K(t)} = \frac{1}{\mu_K(0)} ,$$

which is true if and only if

$$F_K(t) = 1-e^{-\frac{1}{\mu_K(0)}t} . \tag{224}$$

From (224) and the definition of $F_{K+1}(t)$ it follows that

$$F_K(t) = F_{K+1}(t) = F_{K+2}(t) = \ldots .$$

It remains to be shown that $F_{K-2}(t)$ and all previous members of the sequence $\{F_K(t)\}$ are exponential. Writing (K-1) for K in (220) and differentiating, it follows that

$$\begin{aligned} f_{K-1}(t) &= \frac{1}{\mu_{K-2}(0)} (1-F_{K-2}(t)) \\ &= \frac{1}{\mu_{K-1}(0)} e^{-\frac{t}{\mu_{K-1}}} . \end{aligned} \tag{225}$$

Putting t = 0 in (225) we discover that

$$\mu_{K-1}(0) = \mu_{K-2}(0) . \tag{226}$$

Combining (225) and (226) we finally obtain

$$F_{K-2}(t) = 1-e^{-\frac{t}{\mu_{K-1}}} .$$

It is clear that in a similar manner all the distributions have to be exponential. We have thus proved:

Theorem 8

The sequence $\{F_K(t)\}$ is identically exponentially distributed if and only if for some $K \geq 1$, $F_K(t) = F_{K-1}(t)$.

Example 2

$$F_0(x) = x \quad , \quad 0 \leq x \leq 1 \ .$$

It can be easily seen that

$$F_1(t) = 1-(1-t)^2, \ 0 \leq t \leq 1,$$

$$f_1(t) = 2(1-t), \quad 0 \leq t \leq 1 \ ,$$

and more generally,

$$F_n(t) = 1-(1-t)^{n+1}, \quad 0 \leq t \leq 1$$

$$f_n(t) = (n+1)(1-t)^n, \quad 0 \leq t \leq 1$$

$$\mu_n(0) = 1/(n+2)$$

$$Z_n(t) = (n+1)/(1-t)$$

$$\nu_n(t) = (1-t)/(n+2) \ .$$

In this case $F_n(t)$ converges to a degenerate distribution with all the mass concentrated at the origin.

An Inequality for Arbitrary Failure Distribution

One form of the following inequality for the expectation of a random variable is proved by Barlow [1], under the assumption that the distribution function of the random variable has an increasing hazard rate. We shall establish the following inequality in general.

Theorem 9

Let X be a random variable with probability density function f(x), and let $E|X| < \infty$. Then

$$\frac{\int_{-\infty}^{t} t\ f(t)dt}{P[X \leq x]} \leq E(X) \leq \frac{\int_{x}^{\infty} t\ f(t)dt}{P[X > x]} \quad , \tag{227}$$

Proof

$$E(X) = \int_{-\infty}^{\infty} t\ f(t)dt$$

$$= \int_{-\infty}^{x} t\ f(t)dt + \int_{x}^{\infty} t\ f(t)dt \quad . \tag{228}$$

Also,

$$E(X) = \{P[X > x] + P[X \leq x]\} \int_{-\infty}^{\infty} t\ f(t)dt \ . \tag{229}$$

So, from (228) and (229) we have,

$$\int_{-\infty}^{\infty} t\ f(t)dt = \frac{1}{P[X \le x]} \int_{-\infty}^{x} t\ f(t)dt$$

$$+ \frac{1}{P[X \le x]} \left\{ \int_{x}^{\infty} t\ f(t)dt - P[X > x] \int_{-\infty}^{\infty} t\ f(t)dt \right\} . \qquad (230)$$

Thus we see from (230) that

$$\int_{x}^{\infty} t\ f(t)dt < [P\ X > x] \int_{-\infty}^{\infty} t\ f(t)dt$$

$$\Longleftrightarrow \quad \int_{-\infty}^{\infty} t\ f(t)dt < \frac{1}{P[X \le x]} \int_{-\infty}^{x} t\ f(t)dt \quad . \qquad (231)$$

Therefore it suffices to prove only one of the inequalities in (227). Let us break up the range of X into two parts: $E(X) \le x$ and $E(X) > x$. When $E(X) \le x$, we have

$$\int_{x}^{\infty} t\ f(t)dt \ge x \int_{x}^{\infty} f(t)dt = x[P\ X > x] \ ;$$

$$\therefore \qquad \frac{1}{P[X > x]} \int_{x}^{\infty} t\ f(t)dt \ge x \ge E(X) \ . \qquad (232)$$

When $E(X) > x$, we have

$$\int_{-\infty}^{x} t\ f(t)dt \le x \int_{-\infty}^{x} f(t)dt = x\ P[X \le x] \le E(X)\ P[X \le x] \ ;$$

$$\therefore \qquad E(X) - \int_{-\infty}^{x} t\ f(t)dt \ge E(X) - E(X)\ P[X \le x] \ ,$$

or $$\frac{1}{P[X > x]} \int_{x}^{\infty} t\ f(t)dt \ge E(X) \ . \tag{233}$$

(232) and (233) together prove the first part of (231), which implies the second part.

CHAPTER VIII

OPTIMUM ESTIMATION OF THE LAW OF FAILURES

Consider the estimate

$$f_n(t_o) = \int_o^\infty B_n K\left(B_n(\tau - t_o)\right) dF_n(\tau)$$
$$= \frac{B_n}{n} \sum_{j=1}^{n} K\left(B_n(T_j - t_o)\right) \tag{234}$$

for estimating the density f(t) of the underlying law of failures, where $T_1, T_2, \ldots T_n$ are the observed times to failure of the n items put to a life test, and t_o is a point of continuity of the underlying law of failures.

Choose the weight function K(t) such that it vanishes outside a finite interval around $t = t_o$. More specifically, let

$$K(t) = o \text{ for } |t - t_o| > h \tag{235}$$

where h > o is any finite real number.

In view of the results obtained in Chapter II on $f_n(t_o)$, it follows that

$$\lim_{n\to\infty} E(f_n(t_o)) = f(t_o) \int_{t_o-h}^{t_o+h} K(t)\, dt \quad , \tag{236}$$

and

$$\lim_{n\to\infty} \left[\frac{n}{B_n} \text{Var}\,(f_n(t_o))\right] = f(t_o) \int_{t_o-h}^{t_o+h} K^2(t)\, dt \quad . \tag{237}$$

From equation (236) the condition for asymptotic unbiasedness on the weight function K(t) is that

$$\int_{t_o-h}^{t_o+h} K(t)\,dt = 1. \tag{238}$$

Let us now impose a condition on the spread, or equivalently the "bandwidth," of the weight function K(t). One measure of the bandwidth of K(t) is its variance or the second moment about its mean. Denoting this quantity by B, we have

$$B = \int_{t_o-h}^{t_o+h} (t - t_o)^2 K(t)\,dt$$

$$= \int_{t_o-h}^{t_o+h} t^2 K(t)\,dt - t_o^2 \quad . \tag{239}$$

Therefore,

$$\int_{t_o-h}^{t_o+h} t^2 K(t)\,dt = B + t_o^2 \quad . \tag{240}$$

Definition of Optimum Estimate for the Underlying Law of Failures

For a given sample and a given sequence $\{B_n\}$, the estimate for the density $f(t_o)$ at $t = t_o$ depends only on the weight function or window $K(t)$. Choose $K(t)$ such that the corresponding estimate $f_n(t_0)$ has minimum asymptotic variance subject to the conditions of asymptotic unbiasedness and a given bandwidth. Such estimates for the density of the underlying law of failures are called optimum estimates.

The existence of a weight function realizing the above requirement is evident from the fact that the set of all distributions with a bounded second moment is compact.

Optimum weight functions can now be obtained: Namely, the weight function $K(t)$ which is nonnegative, satisfies equations (238) and (240) for a given $B = B_o$, and renders

$$\int_{t_o - h}^{t_o + h} K^2 (t)\, dt \quad \text{a minimum}. \tag{241}$$

The following well-known lemma from the calculus of variations can now be used.

Lemma 10

Let

$$I(K) = \int_a^b F\left(t, K, \frac{dK}{dt}\right) dt \qquad (A), \tag{242}$$

$$W_1(K) = \int_a^b G_1\left(t, K, \frac{dK}{dt}\right) dt - C_1 = o \qquad (B), \tag{243}$$

$$W_2(K) = \int_a^b G_2\left(t, K, \frac{dK}{dt}\right) dt - C_2 = 0 \qquad (C). \tag{244}$$

Then the function K(t), which minimizes the functional (A) subject to the conditions (B) and (C), is given by the Euler-Lagrange differential equation

$$\frac{\partial}{\partial K}\left[F - \lambda_1 G_1 - \lambda_2 G_2\right] - \frac{\partial}{\partial t}\left[\frac{\partial}{\partial K'}\left(F - \lambda_1 G_1 - \lambda_2 G_2\right)\right] = 0 \tag{245}$$

where the symbol prime denotes differentiation with respect to t, and λ_1, λ_2 are the Lagrangian multipliers which are determined by the two conditions (B) and (C). The lemma also assumes that

$$\frac{\partial G_i}{\partial K} - \frac{\partial}{\partial t}\left(\frac{\partial G_i}{\partial K'}\right) \neq 0, \; i = 1, 2. \tag{246}$$

To obtain the optimum weight function K(t), apply the lemma with

$$\begin{aligned} F\ (t, K, K') &= K^2(t), \\ G_1\ (t, K, K') &= K(t), \\ G_2\ (t, K, K') &= t^2 K(t), \end{aligned} \tag{247}$$

and the interval (a, b) is the interval $(t_o - h, t_o + h)$. The Euler-Lagrange differential equation in this case is

$$\frac{\partial}{\partial K}\left[K^2 - \lambda_1 K - \lambda_2 t^2 K\right] = 0 \tag{248}$$

Hence,

$$K(t) = A + B t^2, \tag{249}$$

where we have written A and B for $\lambda_1/2$ and $\lambda_2/2$, respectively. A and B are determined from the conditions of equations (238) and (240). Thus,

$$\int_{t_o - h}^{t_o + h} (A + B t^2)\, dt = 1 ,$$

$$\int_{t_o - h}^{t_o + h} t^2 (A + B t^2) dt = B_o + t_o^{\,2} \tag{250}$$

Simplyfying equation (250), A and B are given by

$$A = \frac{\dfrac{\left(t_o + h\right)^5 - \left(t_o - h\right)^5}{5} - \dfrac{\left(B_o + t_o^{\,2}\right)}{3} \left[\left(t_o + h\right)^3 - \left(t_o - h\right)^3\right]}{\dfrac{2h}{5}\left(\left(t_o + h\right)^5 - \left(t_o - h\right)^5\right) - \dfrac{1}{9}\left[\left(t_o + h\right)^3 - \left(t_o - h\right)^3\right]^2} \tag{251}$$

$$B = \frac{2h\left(B_o + t_o^{\,2}\right) - \dfrac{1}{3}\left[\left(t_o + h\right)^3 - \left(t_o - h\right)^3\right]}{\dfrac{2h}{5}\left[\left(t_o + h\right)^5 - \left(t_o - h\right)^5\right] - \dfrac{1}{9}\left[\left(t_o + h\right)^3 - \left(t_o - h\right)^3\right]^2} . \tag{252}$$

Now suppose that the weight function K(t) is constant in the time interval $(t_o - h,\ t_o + h)$. Then

$$K(t) = \frac{1}{2h} ,$$

since

$$\int_{t_o - h}^{t_o + h} K(t)\, dt = 1 \quad . \tag{253}$$

Also, from equation (239),

$$B_o = \int_{t_o - h}^{t_o + h} t^2 K(t)\, dt - t_o^{\,2}$$

$$= \frac{\left(t_o + h\right)^3 - \left(t_o - h\right)^3}{6h} - t_o^{\,2} \quad . \tag{254}$$

Substituting for B_o from equation (254) in B given by equation (252), it is found that B is identically zero. Also substituting for B_o from equation (254) in A given by equation (251), it is found that

$$A = \frac{1}{2h} \quad .$$

Thus, the rectangular window

$$K(t) = \frac{1}{2h} \quad \text{for} \quad t_o - h \le t \le t_o + h$$

$$= 0 \quad \text{otherwise}$$

is optimum for estimating the underlying law of failures for large samples. Optimum estimates for the underlying law of failures, using other plausible restrictions, can be obtained in a similar manner.

CHAPTER IX

CONCLUDING REMARKS

In the first three chapters of this report a number of asymptotic results have been obtained which deal with estimation and confidence bands for life distributions, their probability densities, and hazard functions, based on random samples, as well as on censored and truncated samples. As is frequently the case with asymptotic procedures, certain pertinent questions still need answering. How large must the sample sizes be under any one of these procedures, to make the asymptotic results practically applicable? Is it possible to replace the asymptotic results by exact small-sample results? While the second question appears very difficult to answer, information leading to answers to the first question is most likely obtainable by the use of Monte Carlo techniques, and studies of this kind should be undertaken in the future.

In Chapter V the meaning and the importance of jumps of the distribution function of life lengths have been discussed, and a test procedure has been proposed which makes it possible to conclude whether a given life distribution has points of discontinuity or not. This procedure, however, does not offer a hint as to the time instant at which those jumps occur. Since these time instants are the times of instantaneous increase of the hazard, it would be of considerable practical importance to be able to estimate their location on the time axis. This, again, constitutes an open problem which should be investigated.

For the Rényi-type statistics discussed in Chapter VI, the exact distributions for finite sample sizes are available, as well as the asymptotic distributions for large samples. Two such statistics have been explored in detail and numerical tables have been computed which make it possible to use them in practical situations. Both of these statistics could be used for the same kinds of problems, and the question has not been answered which of the two statistics is preferable. It appears likely that this question can be answered by analytic methods, by studying the relative asymptotic efficiencies of these two statistics. Should these methods fail, a Monte Carlo study would always be an alternate route to the problem.

The family of DMR life distributions, discussed in Chapter VII, has a number of theoretical properties which suggest that it may be capable of several applications. No such applications, however, have been explored in the present report, and a study of this kind should be undertaken in the future.

In Chapter VIII the question of optimizing the asymptotic estimation procedures discussed in Chapter II is raised. An answer to that question in principle has been obtained in a form which, for a given sequence of constants B_m, determines the weight function K(t). An open problem which should be further explored is that of determining the sequence of the constants B_m in an optimal manner.

REFERENCES

1. Barlow, R.E. Bounds on integrals with applications to reliability problems. Ann. Math. Statist., 1965, 36, 565-574.

2. Barlow, R.E., Marshall, A.W., and Proschan, F. Properties of distributions with monotone hazard rate. Ann. Math. Statist., 1963, 34, 375-389.

3. Birnbaum, Z.W., and Lientz, B.P. Exact distributions of some Rényi-type statistics. To appear in Applicationes Mathematicae.

4. Cramer, H. Mathematical methods of statistics. Princeton University Press, 1946.

5. Gnedenko, B.V. and Kolmogorov, A.N. Limit distributions for sums of independent random variables. Addison-Wesley, 1954.

6. Knight, W.R. Exponential and subexponential distributions in statistical life testing. Ph.D. Thesis. University of Toronto, 1959.

7. Lientz, B.P. Distributions of Rényi and kac-type statistics: power of corresponding tests based on Suzuki-type alternatives. ONR Technical Report No. 51 (1968), contract N-ONR-477(38).

8. Murthy, V.K. Study of nonparametric techniques for estimating reliability and other life quality parameters. Douglas Aircraft Company, Inc., Report DAC-59233, June 1966.

9. Parzen, E. On estimating the probability density function and mode. Ann. Math. Stat. 1962, 33, 1065-1076.

10. Rényi, A. On the theory of order statistics. Acta. Math. Acad. Sci. Hungar. Volume 4, 191-231.

11. Watson, G.S. and Wells, W.T. On the possibility of improving the mean useful life of items by eliminating those with short lives. Technometrics, 1961, 3, 281-298.