CONTRIBUTIONS TO STATISTICAL THEORY OF LIFE TESTING AND RELIABILITY

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

Satya Deva Dubey

1960

This is to certify that the

thesis entitled

CONTRIBUTIONS TO STATISTICAL THEORY

OF LIFE TESTING AND RELIABILITY

presented by

Satya Deva Dubey

has been accepted towards fulfillment
of the requirements for

Ph. D. degree in Statistics

Kenneth Arnold

Major professor

Date May 19, 1960

O-169

CONTRIBUTIONS TO STATISTICAL THEORY OF LIFE TESTING AND RELIABILITY

By

SATYA DEVA DUBEY

A THESIS

Submitted to the School of Graduate Studies of
Michigan State University of Agriculture and
Applied Science in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Statistics

1960

Satya Deva Dubey

Candidate for the degree of

Doctor of Philosophy

Final examination, May 11, 1960, 2:00 P.M., Physics-Mathematics Building

Dissertation: Contributions to Statistical Theory of Life Testing and Reliability

Outline of Studies

Major subjects: Probability, Mathematical and Applied Statistics
Minor subjects: Mathematics

Biographical Items

Born, February 10, 1930, Sakara Bazid, India

Undergraduate Studies: Patna University, 1947-51

Graduate Studies: Indian Statistical Institute, Calcutta 1951-53; Carnegie Institute of Technology, Pittsburgh, Pa., Fall, 1956; Michigan State University, 1957-60.

Experience: Technical Assistant, Indian Institute of Technology, Kharagpur, India, 1953-56; Graduate Assistant, Carnegie Institute of Technology, Fall 1956; Temporary Instructor (Feb. to June 1957), Full-time Research Worker (Summer 1957) and Teaching and Special Graduate Research Assistant (1957-60) Michigan State University, Editorial Collaborator of the Journal of American Statistical Association (Dec. 1957 & March 1958).

Founder Member of Indian Society of Theoretical and Applied Mechanics, 1955; full member of Society of Sigma Xi, member of Institute of Mathematical Statistics.

ACKNOWLEDGEMENTS

I wish to express my deep gratitude to Professor Kenneth J. Arnold for his expert supervision, valuable criticisms and active interest throughout this investigation. I am also grateful to Professor Leo Katz for his encouragement at an early stage of this work and to Professor Fritz Herzog of the Department of Mathematics for helping me in the proof of Theorem 2 of Chapter II. My sincere thanks are due to Mrs. Jane Joyaux for the excellent typing of the manuscript. Finally, I am indebted to the Office of National Science Foundation for financial support.

D E D I C A T E D

T O

My Beloved Parents

CONTRIBUTIONS TO STATISTICAL THEORY OF LIFE TESTING AND RELIABILITY

By

SATYA DEVA DUBEY

AN ABSTRACT

Submitted to the School of Graduate Studies of
Michigan State University of Agriculture and
Applied Science in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Statistics

1960

Approved ______________________

Errata Sheet

Page	Line	Printed	Read
1	8	location	scale
24	6	δ^*	δ^*_n
54	7	Student's t test	Student's t
55	3 (from below)	t_i and t_j	t_1 and t_r
77	6	Involing	Invoking
91	8	$\Pi B\left(\frac{i-1}{m}, \frac{1}{m}\right)$	$\prod_{i=1}^{j-1} B\left(\frac{i-1}{m}, \frac{1}{m}\right)$
92	3	$\frac{1}{m}\Gamma\left(\frac{1}{m}\right)^j$	$\left[\frac{1}{m}\Gamma\left(\frac{1}{m}\right)\right]^j$
114	16 (4th from below)	with	which

SATYA DEVA DUBEY

ABSTRACT

Several problems in testing hypotheses about and in estimating parameters of Weibull distributions, particularly the exponential, are considered. Some attention is given to the possibility that simple assumptions about the intensity function will lead to classes of distributions of wide applicability in describing distributions of length of life.

In the case of the exponential failure law with known location parameter, the minimum-variance-single-observation-unbiased estimator of the location parameter is investigated. It is found that if the rth observation in order of increasing time is the single observation on which this estimate is based and if we write $r = n\delta_n$, then

$$\lim_{n \to \infty} \delta_n = \delta_0 \doteq 0.8 .$$

Several tests of parameters are developed for the case in which no observations beyond the rth in order of magnitude are used. When the scale parameter is known, the likelihood ratio test that the location parameter is a given value is uniformly most powerful against all alternatives. When the scale parameter is unknown, the likelihood ratio test that the location parameter is a given value is uniformly most powerful unbiased. For the latter situation a simple test function based on the first and rth observation is proposed. This test function is shown to be unbiased and for the left-sided alternatives the power of the likelihood ratio test and of the simple test function is shown to be identical.

SATYA DEVA DUBEY

ABSTRACT

For a simple hypothesis on the location and the scale parameters the test function derived by means of the Neyman-Pearson lemma is shown to be uniformly most powerful against alternatives confined to the south-west quadrant. A uniformly most powerful unbiased test for the scale parameter is derived for the case in which the location parameter is unknown and for the case in which it is known a similar test function of first r observations is suggested. The power functions for these tests are expressed in terms of the Incomplete Gamma Function. Two simple test functions for testing the hypothesis on the scale parameter when the location parameter is known and unknown respectively are proposed and their power functions are derived.

Some results are extended to two sample problems. For the likelihood ratio test based on the first r_1 $(\leq n_1)$ and r_2 $(\leq n_2)$ observations to test the hypothesis on the equality of two location parameters assuming the same but unknown scale parameter, the power function is derived and it is shown that the test is unbiased.

The percentile estimators for the parameters of the exponential laws are derived for various situations. The choices of the cumulative probabilities are made so that we have minimum variance unbiased percentile estimators for the estimators. The asymptotic results are given for the sampling distributions, the means, the variances and the coveriance of the unbiased percentile estimators.

SATYA DEVA DUBEY

ABSTRACT

The moment-recurrence formulas for the Weibull laws are established and the moment estimators of Weibull parameters are derived through them. The percentile and the modified percentile estimator for these parameters are derived explicitly and by using the reliability and the intensity functions other estimators are obtained.

Starting from the intensity function, a large number of potentially useful failure laws are generated and the estimation of the parameters is considered. Finally the applications of some of these failure laws are pointed out.

CONTENTS

I. INTRODUCTION

Epstein [23]* has derived simple estimators of the parameters of exponential distributions whose probability density functions (p.d.f.) are

$$f(x;\theta)=\begin{cases}\frac{1}{\theta}e^{-\frac{x}{\theta}}, & x>0,\ \theta>0\\ 0, & \text{otherwise}\end{cases} \tag{1}$$

and

$$f(x;\theta,G)=\begin{cases}\frac{1}{\theta}e^{-\frac{x-G}{\theta}}, & 0<G<x<\infty,\ \theta>0\\ 0, & \text{otherwise}\end{cases} \tag{2}$$

when samples are censored. There he is led to investigate the properties of an unbiased estimator for θ which involves only the $r(\leq n)$th observation of the sample of size n drawn from (1). Denoting the rth observation by $x_{r,n}$, his unbiased estimator of θ is given by

$$\theta^*_{r,n}=\beta_{r,n}x_{r,n} \quad \text{where} \quad \beta_{r,n}=\frac{1}{\sum_{j=1}^{r}\frac{1}{n-j+1}}$$

From [23] we find that in [14] he has shown that $\theta^*_{r,n}$ is of very high efficiency ($\geq$ 96 percent) if $\frac{r}{n}\leq\frac{1}{2}$ [$\geq$ 90 percent if $\frac{r}{n}\leq\frac{2}{3}$] when compared with the best estimator $\hat{\theta}_{r,n}$ where

---

* Numbers in parentheses refer to the bibliography.

$$\hat{\theta}_{r,n} = \frac{\sum_{i=1}^{r} x_{in} + (r-r)x_{rn}}{r} .$$

(The reference [14] was not available to this author.)

Here we have considered the problem of finding the most efficient single-observation estimator of θ. The results concerning this have been investigated in Chapter II. whose summary we give below.

In the second chapter we show that the smallest sample observation in case of the 1-parameter exponential failure law provides a worse estimator (in the sense of minimum variance unbiased estimator) among single-observation estimators for average life than any one of the (n-1) remaining sample observations in a sample of size n. It follows from [23] that the largest sample observation, up to the sample size five, provides the best estimator for average life in the same sense. Here we show that a single-observation unbiased estimator for average life based on the $r(\leq n)$ th statistic where $r = n\delta_n$ with

$$\lim_{n} \delta_n = \delta_0 \doteq 0.8$$

possesses minimum variance. It is about 66 percent efficient in comparison with the minimum variance unbiased estimator based on all observations in the sample. The sample median has only 48 percent efficiency. The smallest sample observation is $\frac{100}{n}$ (n, sample size) percent efficient and the largest sample observation has $\frac{600}{n\pi^2} \ln^2(n)$ asymptotic efficiency. Since the life test data are naturally ordered we have <u>dependent</u> random variables to work with. In this connection we have found the product moment correlation coefficient to order $(n+2)^{-1}$ between any ith and jth $(i < j)$ ordered sample observations from exponential population. This correlation is

$$\rho(t_i, t_j) = \sqrt{\frac{i(n-j+1)}{j(n-i+1)}}\left\{1 - \frac{1}{4(n+2)}\left(\frac{1}{n-j+1} - \frac{1}{n-i+1}\right)\right\}.$$

It is asymptotically equal to $\sqrt{\frac{i(n-j+1)}{j(n-i+1)}}$.

Epstein and Sobel have derived the best test for the H: $\theta = \theta_1$ against A: $\theta = \theta_2$ $(< \theta_1)$ based on the first $r(\leq n)$ observations out of a sample of n drawn from (1) in [1]. In the first part of Chapter III, we have considered tests for the hypotheses:

i) H: $\theta = \theta_0$ against A: $\theta \neq \theta_0$, G known.

ii) H: $\theta = \theta_0$ against A: $\theta \neq \theta_0$, G unknown.

iii) H: $G = G_0$ against A: $G \neq G_0$, θ known,

and

iv) H: $G = G_0$ against A: $G \neq G_0$, θ unknown.

Paulson [4] and Lehmann [2] have considered these hypotheses under the assumption that all n sample observations are available. Here we have extended the results of Paulson and Lehman for all the cases. The extension consists in the fact that our tests are based on only the first $r(\leq n)$ ordered observations from a sample of size n . Furthermore we have also considered v) H: $G_1 = G_2$ against A: $G_1 \neq G_2$ and assuming the same but unknown scale parameter. Paulson [4] has considered this hypothesis under the assumption that all the sample observations are available. Epstein and Tsao [7] have derived the reduced likelihood ratio test when samples are censored from the right. Here following Paulson [4] we have derived the power function of this test and have shown that the test is unbiased. For the tests concerning the hypotheses i), ii), iii) and iv) we have derived the power functions and have investigated some of their properties. Following Lehmann [2] we have obtained the uniformly most powerful (UMP) unbiased test for the hypothesis ii) when sample is censored from the right. In the case of the hypothesis iii) we have shown that the likelihood ratio test is UMP against all alternatives and for the hypothesis iv) we have shown, by following Lehmann [2], that the corresponding likelihood ratio test is UMP unbiased test. Furthermore we have pointed out that the best test for the H: $\theta = \theta_1$ against A: $\theta = \theta_2$ $(< \theta_1)$, considered by Epstein and Sobel [1], is UMP for the H: $G = G_0$ and $\theta = \theta_0$ against A: $G < G_0$ and $\theta < \theta_0$. This test is also UMP for the H: $\theta \geq \theta_0$ against A: $\theta < \theta_0$ with known G .

Epstein and Sobel have considered a test based on the rth observation only to test H: $\theta = \theta_1$ against A: $\theta = \theta_2$ $(< \theta_1)$ in [1]. This has led us to consider simple test functions for the hypotheses i), ii), and iv) in the second part of this chapter. Since the test function for the hypothesis iii) is based on the first observation alone and is UMP we have considered simple test functions for the remaining three cases. When guarantee time (location parameter) is known, the proposed simple test function to test the hypothesis about θ , the scale parameter, is based on only the rth observation of the sample. When guarantee time is unknown, a corresponding simple test function is based on the first and the rth observations. For both cases the power functions have been derived and have been reduced to the Incomplete Beta Function. When the scale parameter is unknown a simple test function based on the first and the $r(\leq n)$ th observations has been suggested to test hypotheses on guarantee time. Its power function has been derived which shows that the test is unbiased and for the left-sided alternatives its power function and the power function of the corresponding likelihood ratio test is identical. Two moment-recurrence formulas have been established to compute higher moments of this simple statistic.

In Chapter IV. we have made an extensive study of the percentile estimators for both the location and the scale parameters of the exponential failure law in three different cases. Some of the results of this report are extensions of the results of Chapter II. We have derived the sampling distributions of the percentile estimators; have derived their asymptotic distributions, have given expressions for their kth moments and have made choices for cumulative probabilities such that the corresponding percentiles in first two cases insure minimum variance single-observation unbiased percentile estimators provided the sample size is large. For the third case we have given the asymptotic form of the covariance matrix.

Kao [12] has derived the maximum likelihood estimators (m.l.e.) of θ and m for the 2-parameter Weibull law when a random sample is censored from the right. His likelihood equations demand the use of successive approximations. Duggan [13] has obtained moment estimators for all the three parameters of the Weibull law which requires use of a table especially prepared for this purpose. Representing the p.d.f. of the 3-parameter Weibull law by

$$f(t) = \begin{cases} \dfrac{m(t-G)^{m-1}}{\theta} e^{-\frac{1}{\theta}(t-G)^m}, & t \geq G \in (-\infty, \infty) \;\&\; \theta, m \in (0,\infty). \\ 0, & \text{otherwise,} \end{cases}$$

we have presented in Chapter V. the results listed in the following paragraph.

First we show that β_1 (measure of skewness) and β_2 (measure of kurtosis) for the Weibull laws are functions of m, the shape parameter, only. Then we establish a lemma which reveals the relationship between the rth moment and the rth power of the first moment of the Weibull law. This lemma is used for deriving moment estimators for the Weibull parameters. We obtain percentile and modified percentile estimators for the Weibull parameters in the form of formulas. By means of the reliability function and the intensity function we have derived some other estimators for the Weibull parameters as well.

In the sixth chapter, a large number of failure laws have been generated by various reasonable assumptions about the form of the intensity function. The applications of some failure laws, generated in this manner, have been pointed out and the estimation of parameters of such failure laws has been considered.

II. SOME RESULTS RELEVANT TO EXPONENTIAL FAILURE LAW

In this chapter we shall derive some results of interest in life testing problems where the random variable has the following exponential probability density function

$$f(X) = \begin{cases} \frac{1}{\theta} e^{-\frac{1}{\theta}(X-G)} & , X > G \in (-\infty,\infty) \ \& \ \theta \in (0,\infty) \\ 0 , & \text{otherwise.} \end{cases}$$

We shall assume G to be known throughout this chapter and since we are concerned with life test data in time units we shall write t instead of X - G and reduce the above exponential probability law to the one-parameter exponential probability density function (p.d.f.) whose form is given by

$$f(t) = \begin{cases} \frac{1}{\theta} e^{-\frac{1}{\theta}t} & , t > 0 \\ 0 , & \text{otherwise.} \end{cases}$$

Now we proceed to prove the following simple theorem.

<u>Theorem 1</u>: For the above one-parameter exponential failure law, the maximum of sample observations provides a more efficient estimator of average life than its minimum.

<u>Proof</u>: Let $(t_1, t_2, \ldots, t_n)$ be a sample of size n. It does not matter here whether they are ordered or not. However, it is clear that in the life testing situations our observations will always appear ordered. Let $\xi = \max(t_1, t_2, \ldots, t_n)$ and $\eta = \min(t_1, t_2, \ldots, t_n)$. Now the p.d.f. of ξ is given by

$$f(\xi) = \begin{cases} \frac{n}{\theta} e^{-\frac{\xi}{\theta}} \left(1 - e^{-\frac{\xi}{\theta}}\right)^{n-1}, & \xi > 0 \\ 0, & \text{otherwise.} \end{cases}$$

This gives

$$E\xi^k = n\theta^k \Gamma(k+1) \sum_{j=0}^{n-1} \binom{n-1}{j} (-1)^j \frac{1}{(j+1)^{k+1}}$$

Hence,

$$E\xi = n\theta \sum_{j=0}^{n-1} \binom{n-1}{j} (-1)^j \frac{1}{(j+1)^2} = \theta \sum_{j=0}^{n-1} \binom{n}{j+1} (-1)^j \frac{1}{j+1}$$

$$= \theta \sum_{k=1}^{n} \binom{n}{k} (-1)^{k-1} \frac{1}{k} = \theta \sum_{k=1}^{n} \frac{1}{k}$$, by virtue of the following lemma.

<u>Lemma 1</u>: $$\sum_{j=1}^{n} \binom{n}{j} (-1)^{j-1} \frac{1}{j} = \sum_{j=1}^{n} \frac{1}{j}$$

Proof: We write $$(1-x)^n = \sum_{j=0}^{n} \binom{n}{j} (-1)^j x^j$$

$$= 1 + \sum_{j=1}^{n} \binom{n}{j} (-1)^j x^j$$

$$= 1 - x \sum_{j=1}^{n} \binom{n}{j} (-1)^{j-1} x^{j-1}$$

Therefore,

$$\frac{1-(1-x)^n}{x} = \sum_{j=1}^{n} \binom{n}{j} (-1)^{j-1} x^{j-1} \quad \text{for} \quad x \neq 0.$$

Now

$$\frac{1-(1-x)^n}{x} = \sum_{j=0}^{n-1} (1-x)^i \quad \text{for} \quad x \neq 0$$

$$\int_0^1 \sum_{i=0}^{n-1} (1-x)^i \, dx = \sum_{i=0}^{n-1} \int_0^1 (1-x)^i \, dx = \sum_{j=1}^{n} \frac{1}{j}$$

and

$$\int_0^1 \sum_{j=1}^{n} \binom{n}{j} (-1)^{j-1} x^{j-1} dx = \sum_{j=1}^{n} \binom{n}{j} (-1)^{j-1} \int_0^1 x^{j-1} dx$$

$$= \sum_{j=1}^{n} \binom{n}{j} (-1)^{j-1} \frac{1}{j} .$$

Hence the lemma is proved.

Returning to $E\xi = \theta \sum_{k=1}^{n} \frac{1}{k}$, we see that $\xi \left[\sum_{k=1}^{n} \frac{1}{k} \right]^{-1}$ is an unbiased estimator of θ. Its variance is found to be

$\left[\sum_{k=1}^{n} \frac{1}{k} \right]^{-2} Var(\xi)$. An expression for $Var(\xi)$ is given in the following lemma.

<u>Lemma 2</u>: If t_r is the rth ordered sample observation out of a random sample of size $n (1 \leq r \leq n)$ drawn from the above one-parameter exponetial law then the

$$Var\, t_r = \theta^2 \sum_{j=1}^{r} \frac{1}{(n-r+j)^2} .$$

<u>Proof</u>: The p.d.f. of t_r is given by

$$f(t_r) = \begin{cases} \frac{n!}{(r-1)!(n-r)!\theta} e^{-(n-r+1)\frac{t_r}{\theta}} (1-e^{-\frac{t_r}{\theta}})^{r-1}, & t_r > 0 \\ 0, & \text{otherwise.} \end{cases}$$

For the sake of convenience we drop subscript r from t. The characteristic function of the above p.d.f. is given by

$$\phi_1(u) = E e^{iut} = \int_0^\infty e^{iut} f(t)\, dt$$

$$= \frac{n!}{(r-1)!(n-r)!\theta} \int_0^\infty e^{-(n-r-iu+1)\frac{t}{\theta}} (1-e^{-\frac{t}{\theta}})^{r-1} dt .$$

To integrate out the above integral, write $\frac{t}{\theta} = y$ then

$$E e^{iuy} = \frac{n!}{(r-1)!(n-r)!} \int_0^\infty e^{-(n-r+1-iu)y} (1-e^{-y})^{r-1} dy .$$

Let $e^{-y} = z$ then

$$E e^{iuy} = \frac{n!}{(r-1)!(n-r)!} \int_0^1 z^{n-r-iu} (1-z)^{r-1} dz .$$

Since the real parts of (n-r+1-iu) and r are positive, the above integral is a Beta function whose arguments are n-r+1-iu and r respectively ([22], p. 212). Thus

$$\phi_2(u) = E e^{iuy} = \frac{n!}{(r-1)!(n-r)!} B(n-r+1-iu; r)$$

$$= \frac{\prod_{j=1}^{r} (n-r+j)}{\prod_{j=1}^{r} (n-r+j-iu)} .$$

Hence

$$\ln \phi_2(u) = \sum_{j=1}^{r} \ln(n-r+j) - \sum_{j=1}^{r} \ln(n-r+j-iu),$$

where any branch can be chosen.

The rth cumulant is given by $k_r = (-1)^r \frac{d^r \ln \phi_2(u)}{du^r}\Big|_{u=0}$ and we have $k_2 = -\frac{d^2 \ln \phi_2(u)}{du^2}\Big|_{u=0} = \sum_{j=1}^{r} \frac{1}{(n-r+j)^2} = \mathrm{Var}\, y = \frac{1}{\theta^2}\mathrm{Var}\, t_r$.

Therefore, $\mathrm{Var}\, t_r = \theta^2 \sum_{j=1}^{r} \frac{1}{(n-r+j)^2}$.

Q.E.D.

Incidentally $k_1 = (-1)\frac{d \ln \phi_2(u)}{du}\Big|_{u=0} = \sum_{j=1}^{r} \frac{1}{n-r+j}$

$= E y$ which gives $E t_r = \theta \sum_{j=1}^{r} \frac{1}{n-r+j}$.

ξ, the maximum of sample observations is t_n when sample observations are ordered, therefore, $\mathrm{Var}\, \xi = \mathrm{Var}\, t_n = \theta^2 \sum_{j=1}^{n} \frac{1}{j^2}$ and again incidentally, $E\xi = E t_n = \theta \sum_{j=1}^{n} \frac{1}{j}$ which checks with the direct derivation of the expectation of the maximum of sample observations.

Thus $\mathrm{Var}\left(\frac{\xi}{\sum_1^n \frac{1}{k}}\right) = \theta^2 \frac{\sum_1^n \frac{1}{k^2}}{\left(\sum_1^n \frac{1}{k}\right)^2} \leq \theta^2$.

The p.d.f. of η, the minimum of sample observations is given by

$$f(\eta) = \begin{cases} \frac{n}{\theta} e^{-\frac{n}{\theta}\eta}, & \eta > 0 \\ 0, & \text{otherwise.} \end{cases}$$

This gives $E\eta^k = \left(\frac{\theta}{n}\right)^k (k!)$

Hence $E\eta = \frac{\theta}{n}$ and $\operatorname{Var}\eta = \frac{\theta^2}{n^2}$.

The unbiased estimator of θ based on the minimum sample observation is $n\eta$ whose variance is θ^2 .

Thus we see that $\operatorname{Var}\left(\frac{\xi}{\sum_1^n \frac{1}{k}}\right) \leq \operatorname{Var}(n\eta)$ which immediately proves the theorem.

The following lemma is useful in deriving the p.d.f. of η from the joint p.d.f. of ξ and η .

__Lemma 3__:

$$\sum_{j=0}^{n} \binom{n}{j} (-1)^j \frac{1}{j+1} = \frac{1}{n+1}$$

__Proof__: We write $(1-x)^n = \sum_{j=0}^{n} \binom{n}{j} (-1)^j x^j, \quad 0 < x < 1.$

Now $\int_0^1 (1-x)^n dx = \frac{1}{n+1}$ and $\sum_{j=0}^{n} \binom{n}{j} (-1)^j \int_0^1 x^j dx = \sum_{j=0}^{n} \binom{n}{j} (-1)^j \frac{1}{j+1}$,

hence the lemma is proved.

Noting that $\sum_{1}^{n} \frac{1}{k} = \ln(n+\frac{1}{2}) + \gamma + O(\frac{1}{n^2})$, where γ is Euler's constant, we see that the unbiased estimator of θ based on the maximum of sample observations has asymptotic variance zero. However the variance of the unbiased estimator of θ based on the minimum sample observation is θ^2, independent of sample size n. Thus its asymptotic variance is also θ^2.

It is shown now that in fact the unbiased estimator of θ, based on the minimum sample observation provides the worst estimator of θ in the sense of minimum variance unbiased estimator in a class of single-observation estimators.

__Proof__: For the $r (1 \leq r \leq n)$ th ordered sample observation we have established that

$$E t_r = \theta \sum_{j=1}^{r} \frac{1}{n-r+j} \quad \& \quad Var\, t_r = \theta^2 \sum_{j=1}^{r} \frac{1}{(n-r+j)^2} .$$

Here $r = 1$ gives $E t_1 = E\eta = \frac{\theta}{n}$ and $Var\, t_1 = Var\, \eta = \frac{\theta^2}{n^2}$;

$r = n$ gives $E t_n = E\xi = \theta \sum_{1}^{n} \frac{1}{j}$ and $Var\, t_n = Var\, \xi = \theta^2 \sum_{1}^{n} \frac{1}{j^2}$;

and for $1 < r < n$ we have the expressions for Et_r and $Var\, t_r$ given above.

Now $Var(n\eta) = \theta^2 > \theta^2 \dfrac{\sum_{j=1}^{r} \frac{1}{(n-r+j)^2}}{\left(\sum_{j=1}^{r} \frac{1}{n-r+j}\right)^2} = Var\left(\dfrac{t_r}{\sum_{j=1}^{r} \frac{1}{n-r+j}}\right)$,

for $r > 1$,

hence it is proved that the minimum sample observation provides the worst estimator of θ in a class of single-observation estimators.

Since

$$Var\left(\frac{\xi}{\sum_{1}^{n} \frac{1}{j}}\right) = \theta^2 \frac{\sum_{1}^{n} \frac{1}{j^2}}{\left(\sum_{1}^{n} \frac{1}{j}\right)^2} \leq \theta^2$$

and

$$Var\left(\frac{t_r}{\sum_{j=1}^{r} \frac{1}{n-r+j}}\right) = \theta^2 \frac{\sum_{j=1}^{r} \frac{1}{(n-r+j)^2}}{\left(\sum_{j=1}^{r} \frac{1}{n-r+j}\right)^2} \leq \theta^2,$$

it is not clear at this stage whether the maximum sample observation provides the best estimator of θ among estimators based on a single observation. Numerical computation ([23], Table II) shows that the maximum sample observation provides the best single-observation estimator of θ so long as the sample size does not exceed five. Beyond five the maximum of sample observations no longer provides the best single-observation estimator of θ. We would like to know for which rth ordered sample observation we obtain the best estimator of θ.

In other words we want to find an integer r, say r_n (depending on n) such that $F(n, r_n)$ is minimum of $F(n, r)$ for fixed n where

$$F(n,r) = \frac{\sum_{j=1}^{r} \frac{1}{(n-r+j)^2}}{\left(\sum_{j=1}^{r} \frac{1}{(n-r+j)}\right)^2} = \frac{\sum_{K=n-r+1}^{n} \frac{1}{K^2}}{\left(\sum_{K=n-r+1}^{n} \frac{1}{K}\right)^2}.$$

We shall answer this question for large n in this chapter. It is clear that $0 < F(n, r) \leq 1$. Let $\delta_n = \frac{r_n}{n}$. Now we shall show that the limit inferior of δ_n is larger than zero. First we prove the following lemma.

<u>Lemma 4</u>: $F(n, r) \geq \frac{1}{r}$.

<u>Proof</u>: Let us define a random variable X with the following probability distribution.

$$P[X = n-r+j] = \begin{cases} \dfrac{\frac{1}{(n-r+j)^2}}{\sum_{j=1}^{r} \frac{1}{(n-r+j)^2}} & \text{for } j = 1, 2, \cdots r. \\ 0, & \text{otherwise.} \end{cases}$$

This gives

$$EX = \frac{\sum_{j=1}^{r} \frac{1}{n-r+j}}{\sum_{j=1}^{r} \frac{1}{(n-r+j)^2}} \quad \text{and} \quad EX^2 = \frac{r}{\sum_{j=1}^{r} \frac{1}{(n-r+j)^2}}$$

Now $\operatorname{Var} X \geq 0$ implies $F(n, r) \geq \frac{1}{r}$. Q.E.D.

A very simple upper estimate of $F(n, r)$ is computed as

$$F(n,r) = \frac{\sum_{n-r+1}^{n} \frac{1}{k^2}}{\left(\sum_{n-r+1}^{n} \frac{1}{k}\right)^2} \leq \frac{\frac{r}{(n-r+1)^2}}{\left(\frac{r}{n}\right)^2} = \frac{n^2}{r(n-r+1)^2} .$$

From $F(n, r) \geq \frac{1}{r}$ and $\min_r F(n, r) \leq \min_r \frac{n^2}{r(n-r+1)^2}$, we have

$\frac{1}{r_n} \leq F(n, r_n) \leq \min_r \frac{n^2}{r(n-r+1)^2}$ which gives $\frac{1}{r_n} \leq \min_r \frac{n^2}{r(n-r+1)^2}$.

Therefore $\delta_n \geq \max_r \frac{r(n-r+1)^2}{n^3} = \frac{4}{27}\left(1 + \frac{1}{n}\right)^3$ or

$\frac{1}{27}\left(1 + \frac{2}{n}\right)\left(2 + \frac{1}{n}\right)^2$ or $\frac{1}{27}\left(2 + \frac{3}{n}\right)^2$ according as

$n = 2 \bmod 3$, $1 \bmod 3$ and $\bmod 3$ respectively. Hence

$\liminf \delta_n \geq \frac{4}{27} \doteq 0.148$.

A better upper estimate of $\sum_{n-r+1}^{n} \frac{1}{k^2}$ given by $\frac{\pi^2}{6(n-r+1)}$

follows from the result of Blom ([25], p. 80 equations 7.3.11 & 7.3.13). Further using the fact that $\frac{1}{k}$ is a strictly monotone nonincreasing function, an upper estimate of $F(n, r)$ is computed as

$$F(n,r) < \frac{\pi^2}{6(n-r+1)} \left(\int_{n-r+1}^{n+1} \frac{dx}{x^2} \right)^{-2} = \frac{\pi^2}{6(n-r+1)} \ln^{-2}\left(\frac{n+1}{n-r+1}\right)$$

which gives $\liminf \delta_n \geq \frac{24}{(\pi e)^2} \doteq 0.329$.

Now we proceed to prove the following Theorem.

<u>Theorem 2</u>: Let $\min\limits_{1 \leq r \leq n} F(n,r) = F(n,r_n)$ and $\delta_n = \frac{r_n}{n}$ where

$$F(n,r) = \frac{\sum\limits_{n-r+1}^{n} \frac{1}{k^2}}{\left(\sum\limits_{n-r+1}^{n} \frac{1}{k}\right)^2} .$$

Then i) $\delta_n \leq \alpha < 1$ for some $\alpha < 1$ and sufficiently large n; in fact

ii) δ_n tends to the limit δ_0 as n tends to plus infinity where δ_0 is the positive root of the equation, $\ln(1-\delta)+2\delta=0$;

iii) $\frac{nF(n,r)}{\psi(\delta)}$ converges to unity on the open δ interval (0, 1) as n tends to plus infinity where $\psi(\delta) = \frac{\delta}{(1-\delta)\ln^2(1-\delta)}$.

In other words, F(n, r) is asymptotically equal to $\frac{1}{n}\psi(\delta)$, and in particular $F(n, r_n)$ is asymptotically equal to $\frac{1}{n}\psi(\delta_0)$; and

iv) $\frac{nF(n,r)}{\psi(\delta)}$ converges uniformly in $\delta\ (0 < \delta = \frac{r}{n} \leq \alpha)$ to unity as n tends to plus infinity, provided α is any constant satisfying $0 < \alpha < 1$.

Proof: i) First we show that for large n, $F(n, n) > F(n, n-1)$.

Let $\sum_1^n \frac{1}{K^2} = s_n$ and $\sum_1^n \frac{1}{K} = t_n$.

Then $F(n,n) = \frac{s_n}{t_n^2}$, $F(n, n-1) = \frac{s_n - 1}{(t_n - 1)^2}$

$$F(n,n) - F(n,n-1) = \frac{1}{t_n^2(t_n-1)^2}\left[s_n(t_n-1)^2 - (s_n-1)t_n^2\right]$$

$$> \frac{t_n(t_n - 2s_n)}{t_n^4} > \frac{t_n - \frac{\pi^2}{3}}{t_n^3} > 0 \text{ for large } n, \text{ since}$$

$t_n \longrightarrow +\infty$. From now on we assume that $1 \leq r \leq n-1$. Put $\frac{r}{n} = \delta$ so that $\frac{1}{n} \leq \delta \leq 1 - \frac{1}{n}$.

Let $\psi(\delta) = \frac{\delta}{(1-\delta)\ln^2(1-\delta)}$, $0 < \delta < 1$.

Then $$\psi'(\delta) = \frac{\ln(1-\delta) + 2\delta}{(1-\delta)^2 \ln^3(1-\delta)}.$$

Here the denominator, $(1-\delta)^2 \ln^3(1-\delta)$, is < 0 for $0 < \delta < 1$, and the numerator, $\ln(1-\delta) + 2\delta$, is > 0 for $0 < \delta < \delta_0$ where δ_0 is the positive root of the equation, $2\delta + \ln(1-\delta) = 0$, and it is < 0 for $\delta_0 < \delta < 1$. [The function $g(\delta) = 2\delta + \ln(1-\delta)$, $0 < \delta < 1$ has the following properties.

i) $\lim_{\delta \downarrow 0} g(\delta) = 0$; ii) $\lim_{\delta \uparrow 1} g(\delta) = -\infty$, and

iii) $\max_{\delta} g(\delta) = g\left(\frac{1}{2}\right) > 0$. Therefore, δ_0, a positive root of $g(\delta)$ lies between $\frac{1}{2}$ and 1 . It is obvious that there is no real root of $g(\delta)$ greater than one]. This means that $\psi(\delta)$ decreases from $+\infty$ to $\psi(\delta_0)$ as δ increases from 0 to δ_0 and $\psi(\delta)$ increases from $\psi(\delta_0)$ to $+\infty$ as δ increases from δ_0 to 1 . (1)

[$\delta_0 \doteq .797$, which gives $\psi(\delta_0) = 1.545$].

An upper estimate of F(n, r) is found as

$$F(n,r) < \frac{\int_{n-r}^{n} \frac{dx}{x^2}}{\left(\int_{n-r+1}^{n+1} \frac{dx}{x}\right)^2} = \frac{\frac{1}{n-r} - \frac{1}{n}}{\ln^2\left(\frac{n+1}{n-r+1}\right)}$$

$$= \frac{r}{n(n-r)} \ln^{-2}\left(1 - \frac{r}{n+1}\right) = \frac{\psi(\delta)}{n}\left[\frac{\ln(1-\delta)}{\ln\left(1 - \frac{\delta}{1+\frac{1}{n}}\right)}\right]^2.$$

Now we prove the following lemma.

Lemma 5: If β is constant, $0 < \beta < 1$, and $0 < u < 1$ then

$$\Phi(u) = \frac{\ln(1-u)}{\ln(1-\beta u)}$$ increases with u .

Proof:

$$\phi'(u) = \frac{\beta(1-u)\ln(1-u) - (1-\beta u)\ln(1-\beta u)}{(1-u)(1-\beta u)\ln^2(1-\beta u)}$$

The denominator, $(1-u)(1-\beta u)\ \ln^2(1-\beta u)$, is >0 .

The numerator

$$=-\beta(1-u)\left(u+\frac{1}{2}u^2+\cdots\right)+(1-\beta u)\left(\beta u+\frac{1}{2}\beta^2u^2+\cdots\right)$$

$$=\frac{1}{1.2}(\beta-\beta^2)u^2+\frac{1}{2.3}(\beta-\beta^3)u^3+\cdots+\frac{1}{(n-1)n}(\beta-\beta^n)u^n+$$
$$+\cdots\cdots>0.$$

Q.E.D.

Hence, by virtue of the lemma 5,

$$F(n,r)<\frac{\Psi(\delta)}{n}\left\{\frac{\ln[1-(1-\frac{1}{n})]}{\ln\left[1-\frac{1-\frac{1}{n}}{1+\frac{1}{n}}\right]}\right\}^2=\frac{\Psi(\delta)}{n}\left[\frac{\ln(n)}{\ln(\frac{n+1}{2})}\right]^2 \quad (2)$$

We may note that

$$\frac{\ln(n)}{\ln(\frac{n+1}{2})}=\frac{\ln(n)}{\ln(n)+\ln(1+\frac{1}{n})-\ln 2}\to 1 \text{ as } n\to+\infty.$$

Similarly we find a lower estimate of $F(n, r)$.

$$F(n,r)>\frac{\int_{n-r+1}^{n+1}\frac{dx}{x^2}}{\left(\int_{n-r}^{n}\frac{dx}{x}\right)^2}=\frac{\frac{1}{n-r+1}-\frac{1}{n+1}}{\ln^2\left(\frac{n}{n-r}\right)}$$

$$=\frac{r}{(n+1)(n+1-r)}\ln^{-2}\left(1-\frac{r}{n}\right)=\frac{\Psi(\delta)}{n}\cdot\frac{1}{1+\frac{1}{n}}\cdot\frac{1-\delta}{1+\frac{1}{n}-\delta}$$

Now choose α and β such that $\delta_0 < \beta < \alpha < 1$ and $\psi(\alpha) = 3\,\psi(\beta)$.
We consider the following cases.

Case I: $0 < \delta \leq \alpha < 1$.

The following lower estimate of $F(n, r)$ actually holds for any fixed δ.
Since $\dfrac{1-\delta}{1+\frac{1}{n}-\delta} = \dfrac{1}{1+\frac{1}{n(1-\delta)}}$ decreases with δ,

$$F(n,r) > \frac{\psi(\delta)}{n}\cdot\frac{1}{1+\frac{1}{n}}\cdot\frac{1-\alpha}{1+\frac{1}{n}-\alpha} \qquad (3)$$

where $\dfrac{1}{1+\frac{1}{n}}\cdot\dfrac{1-\alpha}{1+\frac{1}{n}-\alpha} \longrightarrow 1$ as $n \longrightarrow +\infty$.

Case II: $\alpha < \delta < 1$.

$$F(n,r) > \frac{\psi(\delta)}{n}\cdot\frac{1}{1+\frac{1}{n}}\cdot\frac{1-(1-\frac{1}{n})}{1+\frac{1}{n}-(1-\frac{1}{n})} \geq \frac{\psi(\alpha)}{n}\frac{1}{1+\frac{1}{n}}\cdot\frac{1}{2}$$

For large n, $\dfrac{1}{2}\dfrac{1}{1+\frac{1}{n}} > \dfrac{1}{3}\cdot\dfrac{\ln(n)}{\ln\left(\frac{n+1}{2}\right)}$, therefore,

$$\frac{\psi(\delta)}{n}\cdot\frac{1}{2}\cdot\frac{1}{1+\frac{1}{n}} > \frac{1}{3}\frac{\psi(\alpha)}{n}\frac{\ln(n)}{\ln\left(\frac{n+1}{2}\right)} = \frac{\psi(\beta)}{n}\frac{\ln(n)}{\ln\left(\frac{n+1}{2}\right)}$$

$$\geq \frac{\psi(\delta')}{n}\cdot\frac{\ln(n)}{\ln\left(\frac{n+1}{2}\right)} \quad \text{for } \delta_0 \leq \delta' \leq \beta .$$

Thus we see that for large n,

$$F(n,r) > \frac{\psi(\delta')}{n}\frac{\ln(n)}{\ln\left(\frac{n+1}{2}\right)} \quad \text{for } \delta' \in [\delta_0, \beta] . \qquad (4)$$

Consequently, $F(n, r) > F(n, r')$ when $\delta = \frac{r}{n} > \alpha$ and $\delta' = \frac{r'}{n} \in [\delta_0, \beta]$ and n large. There is at least one such value of r' when n is large. Hence $F(n, r)$ does not assume its minimum when $\frac{r}{n} = \delta > \alpha$. That is, $F(n, r)$ assumes its minimum, $F(n, r_n)$, when $\delta_n = \frac{r_n}{n} \leq \alpha$.

Q.E.D.

ii) We have here $\delta_n = \frac{r_n}{n} \leq \alpha$ and

$\frac{\psi(\delta)}{n}(1-\epsilon'_n) < F(n,r) < \frac{\psi(\delta)}{n}(1+\epsilon''_n)$ for $0 < \delta \leq \alpha$ where $\epsilon'_n\,(>0)$ and $\epsilon''_n\,(>0)$ approach zero as n tends to plus infinity.

$$\left(\text{Since } F(n,r) < \frac{\psi(\delta)}{n}\left[\frac{\ln(n)}{\ln\left(\frac{n+1}{2}\right)}\right]^2 = \frac{\psi(\delta)}{n}\left[\left(1 - \left\{\frac{\ln\left(\frac{2n}{n+1}\right)}{\ln(n)}\right\}\right)^{-1}\right]^2\right.$$

and,

$$F(n,r) > \frac{\psi(\delta)}{n}\,\frac{1}{1+\frac{1}{n}}\,\frac{1-\alpha}{1+\frac{1}{n}-\alpha} = \frac{\psi(\delta)}{n}\left(1+\frac{1}{n}\right)^{-1}\left(1+\frac{1}{n(1-\alpha)}\right)^{-1}$$

$$\left. \text{for } 0 < \delta \leq \alpha \right).$$

We proceed to show that δ_n tends to δ_0 as $n \longrightarrow +\infty$. Let $\epsilon > 0$ be given. We choose δ^* in $(\delta_0 - \epsilon, \delta_0 + \epsilon)$ such that $\delta^* \neq \delta_0$ and $\psi(\delta_0) < \psi(\delta^*) < \min[\psi(\delta_0 - \epsilon),\ \psi(\delta_0 + \epsilon)] = \psi(\delta_0 \pm \epsilon)$, say. (That this is possible follows from properties enumerated under (i)). Now we choose N so large that for $n > N$,

$$\psi(\delta^*)(1+\epsilon''_n) < \psi(\delta_0 \pm \epsilon)(1-\epsilon'_n). \qquad (6)$$

(That this is possible follows from properties enumerated under (i).

ii) $\epsilon'_n\ (>0) \longrightarrow 0$, $\epsilon''_n\ (>0) \longrightarrow 0$ and

iii) $\delta^* \in (\delta_0 - \epsilon, \delta_0 + \epsilon)$.) For $|\delta - \delta_0| \geq \epsilon$, $\psi(\delta) \geq \psi(\delta_0 \pm \epsilon)$

hence

$$\frac{\psi(\delta_0 \pm \epsilon)}{n}(1-\epsilon'_n) \leq \frac{\psi(\delta)}{n}(1-\epsilon'_n) < F(n,r). \quad (7)$$

And, for $\delta^* = \frac{r_n^*}{n} \in (\delta_0, \delta^*)$, $\psi(\delta_n^*) < \psi(\delta^*)$,

so,

$$\frac{\psi(\delta^*)}{n}(1+\epsilon''_n) > \frac{\psi(\delta_n^*)}{n}(1+\epsilon''_n) > F(n, r_n^*). \quad (8)$$

Hence from (6), (7) and (8) it immediately follows that for $|\delta - \delta_0| \geq \epsilon$, $F(n, r) > F(n, r_n^*)$.

Thus we have $F(n, r) > F(n, r_n^*) > F(n, r_n)$. This implies that $\delta_n = \frac{r_n}{n}$ must lie in $(\delta_0 - \epsilon, \delta_0 + \epsilon)$. That is, $|\delta_n - \delta_0| < \epsilon$ for $n > N$.

Q.E.D.

iii) From (2) we have for $0 < \delta < 1$,

$$F(n,r) < \frac{\psi(\delta)}{n}\left[\frac{\ln(n)}{\ln\left(\frac{n+1}{2}\right)}\right]^2$$

This gives

$$\overline{\lim_{n}}\ \frac{n F(n,r)}{\psi(\delta)} \leq 1 .$$

From (3), we have for $0 < \delta \leq \alpha < 1$,

$$F(n,r) > \frac{\psi(\delta)}{n} \cdot \frac{1}{1+\frac{1}{n}} \cdot \frac{1-\alpha}{1+\frac{1}{n}-\alpha} .$$

This gives

$$\underline{\lim_{n}} \frac{nF(n,r)}{\psi(\delta)} \geq 1 .$$

Hence for every δ in $0 < \delta \leq \alpha$, $\lim_{n} \frac{nF(n,r)}{\psi(\delta)} = 1$.

Since we can choose α as close to unity as we wish, it immediately follows that $\frac{nF(n,r)}{\psi(\delta)}$ converges to unity on open interval $(0, 1)$ as $n \longrightarrow +\infty$. Furthermore, from the definition of asymptotic equivalence ([28], p. 10) it follows that $F(n, r)$ is asymptotically equal to $\frac{1}{n}\psi(\delta)$ and in particular $F(n, r_n)$ is asymptotically equal to $\frac{1}{n}\psi(\delta_0)$.

Q.E.D.

iv) To prove the statement iv) true we show that for every $\epsilon > 0$, there exists an N(depending on ϵ and α) such that

$$\left| \frac{nF(n,r)}{\psi(\delta)} - 1 \right| < \epsilon$$

for $n \geq N$ and all δ in $(0, \alpha]$. That is, $\lim_{n} \frac{nF(n,r)}{\psi(\delta)} = 1$ for all δ satisfying $0 < \delta \leq \alpha < 1$. From 2) we get, for,

$$0 < \delta \leq \alpha < 1 \quad , \frac{nF(n,r)}{\psi(\delta)} < \left[\frac{\ln(n)}{\ln(\frac{n+1}{2})} \right]^2 = 1 + \epsilon_n'' \text{ where } \epsilon_n'' (>0)$$

tends to zero as n tends to plus infinity. That is, for every $\epsilon > 0$, $\exists$ an $N(\epsilon)$ such that $|\epsilon_n''| < \epsilon$ for $n \geq N$.

From 3), we get for $0 < \delta \leq \alpha$,

$$\frac{nF(n,r)}{\psi(\delta)} > \frac{1}{1+\frac{1}{n}} \cdot \frac{1-\alpha}{1+\frac{1}{n}-\alpha} = 1-\epsilon'_n ,$$

where ϵ'_n (> 0) tends to zero as n tends to plus infinity. That is, for every $\epsilon > 0$, $\exists$ $N(\epsilon, \alpha)$ such that for $n \geq N$, $|\epsilon'_n| < \epsilon$. Now take N = largest of $N(\epsilon)$ and $N(\epsilon, \alpha)$. Then for $n \geq N$,

$\frac{nF(n,r)}{\psi(\delta)} - 1 < |\epsilon''_n| < \epsilon$ gives $\lim_n \frac{nF(n,r)}{\psi(\delta)} \leq 1$, for every δ in $(0, \alpha]$; and, $\frac{nF(n,r)}{\psi(\delta)} - 1 > -\epsilon'_n$ gives

$\lim_n \left[\frac{nF(n,r)}{\psi(\delta)} - 1\right] > -\lim_n \epsilon'_n = 0$ which means

$\lim_n \frac{nF(n,r)}{\psi(\delta)} \geq 1$, for every δ in $(0, \alpha]$. Hence

$\lim_n \frac{nF(n,r)}{\psi(\delta)} = 1$ for every δ in $(0, \alpha]$. By virtue of the existence of a single N, the uniform convergence is established.

Q.E.D.

We claim that $\frac{nF(n,r)}{\psi(\delta)}$ does not converge uniformly to unity on the (0, 1).

<u>Proof</u>: Suppose that $\frac{nF(n,r)}{\psi(\delta)}$ converges uniformly to unity on the (0, 1). Then, for any sequence $\{\rho_n\}$ with $0 < \rho_n < 1$, it would be true that

$$\lim_n \frac{nF(n, n\rho_n)}{\psi(\rho_n)} = 1 .$$

Now choose $\rho_n = \frac{n-1}{n}$. Then $\psi\left(\frac{n-1}{n}\right) = \frac{n-1}{\ln^2(n)}$, hence

$$\frac{nF(n,n-1)}{\psi\left(\frac{n-1}{n}\right)} = \frac{n}{n-1}\,\frac{\ln^2(n)}{\left(\sum_{k=2}^{n}\frac{1}{k}\right)^2}\cdot\sum_{2}^{n}\frac{1}{k^2}$$ which approaches

$$\sum_{2}^{\infty}\frac{1}{k^2} = \frac{\pi^2}{6} - 1 < 1.$$ This is a contradiction.

Q.E.D.

It seems interesting to point out, beyond the result of Theorem 2, the following facts. If we approximate

$$\sum_{n-r+1}^{n}\frac{1}{k^2} \text{ and } \sum_{n-r+1}^{n}\frac{1}{k}$$ by

$$\sum_{n-r+1}^{n}\frac{1}{k^2} \doteq \int_{n-r}^{n}\frac{dx}{x^2} = \frac{1}{n-r} - \frac{1}{n},$$ and

$$\sum_{n-r+1}^{n}\frac{1}{k} \doteq \int_{n-r}^{n}\frac{dx}{x} = \ln\left(\frac{n}{n-r}\right)$$

respectively then we can write

$$F(n,r) \doteq \frac{\frac{1}{n-r} - \frac{1}{n}}{\ln^2\left(\frac{n}{n-r}\right)}.$$

Let $\delta = \frac{r}{n}$. Then this approximation for $F(n, r)$ reduces to $\frac{\psi(\delta)}{n}$. The minimum of $\frac{\psi(\delta)}{n}$ gives the same δ_0 which has been shown in Theorem 2 to be the limit of δ_n. This result can also be derived by means of the Euler's Summation Formula.

Following ([24], [26] and [27]9 we derive the expressions for the expectation and the variance of the rth ordered sample observation from one-parameter exponential law up to order $(n + 2)^{-1}$. They are:

$$E t_r = \theta \left[\ln\left(\frac{n+1}{n-r+1}\right) + \frac{r}{2(n-r+1)(n+2)} \right]$$

and,

$$\operatorname{Var} t_r = \theta^2 \frac{r}{(n-r+1)(n+2)} \quad .$$

As n increases with $\frac{r}{n}$ fixed, the asymptotic distribution of t_r has the mean, $\theta \ln\left(\frac{n+1}{n-r+1}\right)$, and variance $\theta^2 \frac{r}{n(n-r+1)}$ respectively. This follows from the results in [24]. Hence the asymptotic variance of the unbiased estimator, based on the t_r th statistic, for θ works out to be

$$\frac{\theta^2}{n} \frac{r}{n-r+1} \ln^{-2}\left(\frac{n+1}{n-r+1}\right) \quad .$$

Now we want to determine r so that this variance is minimum. Treating the expression for the variance as a function of r , taking its first derivative with respect to r and setting it equal to zero, we get, after simplification,

$$\ln\left(1 - \frac{r}{n+1}\right) + \frac{2r}{n+1} = 0$$

Writing $\frac{r}{n+1} = x$, we reduce the above equation to $\ln(1-x) + 2x = 0$, $0 < x < 1$. If x_0 is the solution of the above equation then

$r_n \doteq (n+1)\, x_0$ provides minimum variance for large n. We get $x_0 \doteq 0.797$ which gives asymptotic minimum variance $\doteq \frac{1.545}{n} \theta^2$. This result also follows from the fact that $F(n, r_n)$ is asymptotically equal to $\frac{1}{n} \phi(\delta_0)$ where $\phi(\delta_0) \doteq 1.545$. We know that in the present situation the sample mean is the unique minimum variance unbiased estimator for θ [19]. The variance of this estimator is $\frac{\theta^2}{n}$. Hence the asymptotic efficiency of the estimator based on the r_n th statistic is found to be about 66 percent. The subject matter for the Case I in Chapter IV. is clearly related to the present discussion. From the results derived there, it follows that the sample median has approximately 48 percent asymptotic efficiency.

As a point of interest, the product-moment correlation coefficient between two order statistics, say t_i and $t_j (i < j)$ in a random sample of size n drawn from one-parameter exponential law to order $(n+2)^{-1}$ is obtained by means of the formula given in [24] and [26]. This works out to be

$$\rho(t_i, t_j) = \sqrt{\frac{i(n-j+1)}{j(n-i+1)}} \left\{ 1 - \frac{1}{2(n+2)} \left(\frac{j}{n-j+1} - \frac{i}{n-i+1} \right) \right\}$$

$$= \sqrt{\frac{p_1(1-p_2)}{p_2(1-p_1)}} \left\{ 1 - \frac{1}{2(n+2)} \left(\frac{p_2}{1-p_2} - \frac{p_1}{1-p_1} \right) \right\},$$

where $p_1 = \frac{i}{n+1}$, and $p_2 = \frac{j}{n+1}$. Its asymptotic expression is clearly $\sqrt{\frac{p_1(1-p_2)}{p_2(1-p_1)}}$ or $\sqrt{\frac{i(n-j+1)}{j(n-i+1)}}$. This result also follows from Cramér ([10], p. 369).

III. SOME TESTS ON PARAMETERS OF EXPONENTIAL FAILURE LAW

Part 1.

If a random variable (r.v.) X has the probability density function (p.d.f.)

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}}, & x > 0 \text{ and } \theta > 0 \\ 0, & \text{otherwise}, \end{cases}$$

the likelihood ratio test of H: $\theta = \theta_1$ against A: $\theta = \theta_2 (< \theta_1)$ given the first $r (\leq n)$ ordered observations from a random sample of n yields, as shown by Epstein and Sobel [1], the critical region $\sum_{i=1}^{r} x_i + (n-r) x_r < C$. Epstein and Sobel [1] show that $\frac{2}{\theta}\left[\sum_1^r x_i + (n-r) x_r\right]$ has a chi-square distribution with 2r degrees of freedom. The parameter θ is the expected value of X and is, in life testing, called average life.

Let $\phi(x)$ be the probability of accepting the alternative when the observation vector is x. Now following Fraser [3] it is easy to see that $E_\theta\{\phi(X)\} \leq \alpha$ for $\theta \geq \theta_1$ and furthermore the test does not depend on θ_2 so long as $\theta_2 < \theta_1$. Hence it at once follows that the test function, $\phi(x)$ derived by means of Neyman-Pearson lemma is a uniformly most powerful (UMP) test for the H: $\theta \geq \theta_1$ versus A: $\theta < \theta_1$. Similarly we can show that the corresponding test function, namely,

$$\phi(x) = \begin{cases} 0 & \text{if } \sum_{1}^{r} x_i + (n-r)x_r \leq c \\ 1, & \text{otherwise}, \end{cases}$$

for the H: $\theta = \theta_1$ against A: $\theta = \theta_2$ $(> \theta_1)$ is UMP test for the modified H: $\theta \leq \theta_1$ against A: $\theta > \theta_1$, although in life testing problem, H: $\theta \geq \theta_1$ seems in general to be of practical interest. However, there does not exist a UMP test for the H: $\theta = \theta_1$ against A: $\theta \neq \theta_1$. But if we restrict ourselves to a class of unbiased tests there does exist a UMP unbiased test for the H: $\theta = \theta_1$ against A: $\theta \neq \theta_1$.

Here we propose to consider testing statistical hypotheses connected with the two-parameter exponential law whose probability density function (p.d.f.) is given by

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{1}{\theta}(x-G)}, & x > G \in (-\infty,\infty) \;\&\; \theta \in (0,\infty) \\ 0, & \text{otherwise}, \end{cases}$$

where θ is known as the scale parameter and G the location parameter. G is identified as guarantee time or minimum life in life testing situations.

Since our life test data refer to measurement of time it seems appropriate to denote the observation vector by t instead of x . In the sequel we shall use T and t in the sense of random and observation vector respectively.

Now, in failure analysis we generate data by destructive tests and so from an economic point of view we consider a censoring scheme in which we use only the first $r(\leq n)$ ordered observations: $G \leq t_1 < t_2 < \ldots < t_r < \infty$.

To test the $H:\ \theta = \theta_1$ against $A:\ \theta = \theta_2 (< \theta_1)$, assuming G to be known, say G_0 (not necessarily zero), the best α-level test function based on the first r out of n ordered observations can of course be derived directly from the test function given above for $G = 0$. Obviously, this test function would provide UMP test for modified $H:\ \theta \geq \theta_1$ against $A:\ \theta < \theta_1$ and would possess all the other properties which have been pointed out regarding the hypothesis considered by Epstein and Sobel.

When G is unknown, the best α - level test function for simple $H:\ \theta = \theta_1$ against simple $A:\ \theta = \theta_2\ (< \theta_1)$ is found to be

$$\phi(t) = \begin{cases} 0 & \text{if } \sum_{i=2}^{r} (t_i - t_1) + (n-r)(t_r - t_1) \geq c \\ 1 & \text{if } \qquad\qquad\qquad\qquad < c \end{cases}$$

where $\frac{2}{\theta}\left[\sum_{i=2}^{r}(t_i - t_1) + (n-r)(t_r - t_1)\right]$ has chi-square distribution with $2r - 2$ degrees of freedom. For modified $H:\ \theta = \theta_0$ against $A:\ \theta \neq \theta_0$ and G unknown, the UMP <u>unbiased</u> α - level test function is found to be

$$\phi(t) = \begin{cases} 0 \text{ if } c_1 \leq \sum_2^r (t_i - t_1) + (n-r)(t_r - t_1) = y \leq c_2 \\ 1, \text{ otherwise}, \end{cases}$$

where c_1 and c_2 are determined from

$$\frac{1}{(r-2)!\,\theta_0^{r-1}} \int_{c_1}^{c_2} y^{r-2} e^{-\frac{y}{\theta_0}} dy = 1 - \alpha \qquad (1)$$

in conjunction with

$$\left[\frac{\partial}{\partial \theta} \int_{c_1}^{c_2} f(y)\, dy \right]_{\theta = \theta_0} = 0 \qquad (2)$$

where

$$f(y) = \begin{cases} \frac{1}{(r-2)!\,\theta^{r-1}} y^{r-2} e^{-\frac{y}{\theta}}, & y > 0 \\ 0, & \text{otherwise.} \end{cases}$$

The relation (2) yields,

$$c_1^{r-1} e^{-\frac{c_1}{\theta_0}} = c_2^{r-1} e^{-\frac{c_2}{\theta_0}}$$

The numerical solution for c_1 and c_2 can be carried out by successive approximations or graphically. This test function is derived by following the hint given in Lehmann ([2], Problem 12 i), p. 202)

For the H: $\theta = \theta_0$ against A: $\theta \neq \theta_0$ assuming G to be known, say G_0, an α - level test function is defined by

$$\phi(t) = \begin{cases} 0 & \text{if } c_1 \leq \sum_{1}^{r} (t_i - G_0) + (n-r)(t_r - G_0) \leq c_2 \\ 1, & \text{otherwise.} \end{cases}$$

Noting that

$$\frac{2}{\theta}\left[\sum_{1}^{r} (t_i - G_0) + (n-r)(t_r - G_0)\right]$$

has chi-square distribution with 2r degrees of freedom we obtain the p.d.f. of

$$u = \sum_{1}^{r} (t_i - G_0) + (n-r)(t_r - G_0) .$$

This is

$$f(u) = \begin{cases} \frac{1}{(r-1)!\,\theta^r} u^{r-1} e^{-\frac{u}{\theta}}, & u > 0 \\ 0, & \text{otherwise.} \end{cases}$$

We determine c_1 and c_2 from the equation

$$\frac{1}{(r-1)!}\,\frac{1}{\theta_0^{r}}\int_{c_1}^{c_2} u^{r-1} e^{-\frac{u}{\theta_0}}\,du = 1-\alpha, \tag{3}$$

in conjunction with

$$\left[\frac{\partial}{\partial\theta}\int_{c_1}^{c_2} f(u)\,du\right]_{\theta=\theta_0} = 0 \tag{4}$$

The relation (4) yields

$$c_1^{r} e^{-\frac{c_1}{\theta_0}} = c_2^{r} e^{-\frac{c_2}{\theta_0}}.$$

The equations (1) and (3) can be expressed in terms of the Incomplete Gamma Function as

$$\alpha = 1 - I\left[r-1;\frac{c_2}{\theta_0}\right] + I\left[r-1;\frac{c_1}{\theta_0}\right],$$

and

$$\alpha = 1 - I\left[r;\frac{c_2}{\theta_0}\right] + I\left[r;\frac{c_1}{\theta_0}\right]$$

respectively, where $I[p; q]$ is the Incomplete Gamma Function whose values are tabulated in [5].

And finally the power functions corresponding to $H: \theta = \theta_0$ against

A: $\theta \neq \theta_0$ assuming G unknown and H: $\theta = \theta_0$ against A: $\theta \neq \theta_0$ assuming G known are expressed as

$$P(\theta) = 1 - I\left[r-1, \frac{c_2}{\theta}\right] + I\left[r-1, \frac{c_1}{\theta}\right]$$

and

$$P(\theta) = 1 - I\left[r, \frac{c_2}{\theta}\right] + I\left[r, \frac{c_1}{\theta}\right]$$

respectively.

Earlier we have seen that the test function,

$$\phi(t) = \begin{cases} 0 & \text{if } \sum_{1}^{r} (t_i - G_0) + (n-r)(t_r - G_0) > c \\ 1, & \text{otherwise}, \end{cases}$$

is UMP for the H: $\theta \geq \theta_0$ against A: $\theta < \theta_0$, assuming G to be known, say G_0. We shall show later on that this test function is again UMP for the H: $G = G_0$ and $\theta = \theta_0$ against A: $G < G_0$ and $\theta < \theta_0$.

For testing the H: $G = G_0$ against A: $G \neq G_0$ and θ known, say θ_0 the UMP test based on the first $r(\leq n)$ ordered observations out of a random sample of size n from the two-parameter exponential population is obtained by means of likelihood ratio. The likelihood ratio,

$$\lambda = \frac{\sup_{\omega} f(t)}{\sup_{\Omega} f(t)} ,$$

where $\omega = \{G_0 , \theta_0\}$ and $\Omega = \{G, \theta_0\}$ gives in the present situation

$$\lambda = \begin{cases} e^{-\frac{n}{\theta_0}(t_1 - G_0)} & \text{if } t_1 > G_0 \\ 0 & \text{if } t_1 < G_0 . \end{cases}$$

This implies that the α - level likelihood ratio test is given by

$$\phi(t) = \begin{cases} 0 \text{ if } e^{-\frac{n}{\theta_0}(t_1 - G_0)} > \lambda_\alpha \\ 1 \text{ if } \qquad\qquad \leq \lambda_\alpha \text{ or } t_1 < G_0 , \end{cases}$$

where λ_α is determined from the integral, $\int_0^{\lambda_\alpha} f(\lambda)\, d\lambda = \alpha$. $f(\lambda)$, the p.d.f. of λ under null hypothesis is a uniform distribution over unit interval which immediately gives $\lambda_\alpha = \alpha$. An equivalent α - level test function is then given by

$$\phi(t) = \begin{cases} 0 \text{ if } G_0 \leq t_1 \leq G_0 - \frac{1}{n}\theta_0 \ln \alpha \\ 1 \text{ if } t_1 < G_0 \quad \text{or} \quad t_1 \geq G_0 - \frac{1}{n}\theta_0 \ln \alpha . \end{cases}$$

For $r = n$ where n is the size of a random sample from the same exponential law, Paulson [4] has considered, among other things, $H: G = 0$ against $A: G \neq 0$. Here we shall reformulate Paulson's problems and obtain more general results besides considering other problems as well.

We may note that whenever a UMP test exists it is unbiased since its power cannot fall below that of the test $\phi(t) \equiv \alpha$, and in addition Neyman and Pearson have shown that if a UMP test exists, it is the likelihood ratio test.

Now we shall derive power function, $P(G)$ of the likelihood ratio test for $H: G = G_0$ versus $A: G \neq G_0$, assuming $\theta = \theta_0$ known and show that the likelihood ratio test is unbiased and UMP.

From the equivalent α - level test function given above, we write the power function as

$$P(G) = \int_{G_1}^{G_0} f(t_1)\,dt_1 + \int_{a}^{\infty} f(t_1)\,dt_1 ,$$

where

i) $$f(t_1) = \begin{cases} \frac{n}{\theta_0} e^{-\frac{n}{\theta_0}(t_1 - G)} , & t_1 > G \\ 0 , & \text{otherwise} \end{cases}$$

and

ii) $$a = G_0 - \frac{1}{n}\theta_0 \ln\alpha .$$

Now we consider three cases.

Case I: $G \leq G_0$.

$$P(G) = 1 - \int_{G_0}^{a} f(t_1)\,dt_1 = 1-(1-\alpha)e^{\frac{n}{\theta_0}(G-G_0)},$$

which gives $P(G_0) = \alpha$, as expected.

Case II: $G_0 \leq G \leq a$.

$$P(G) = \int_{a}^{\infty} f(t_1)\,dt_1 = 1 - \int_{G}^{a} f(t_1)\,dt_1 = \alpha e^{\frac{n}{\theta_0}(G-G_0)}$$

which gives $P(G_0) = \alpha$ as expected.

Case III: $G \geq a$.

$$P(G) = \int_{a}^{\infty} f(t_1)\,dt_1 = \int_{G}^{\infty} f(t_1)\,dt_1 = 1.$$

For $G \leq G_0$,

$$e^{\frac{n}{\theta_0}(G-G_0)} \leq 1,$$ which implies

$$P(G) = 1-(1-\alpha)e^{\frac{n}{\theta_0}(G-G_0)} \geq \alpha.$$

For $G_0 \leq G \leq a$,

$$e^{\frac{n}{\theta_0}(G-G_0)} \geq 1$$ which implies

$$P(G) = \alpha e^{\frac{n}{\theta_0}(G - G_0)} \geq \alpha .$$

And, for $G \geq a$,

$$P(G) = 1 \geq \alpha \quad .$$

Hence $P(G) \geq \alpha$ for every $G \in (-\infty, \infty)$ and so the likelihood ratio test is unbiased.

To show that this test is UMP, we need to establish that the critical region provided by the above α - level equivalent test function gives maximum power of the test for each and every alternative to $H;\ G = G_0$. To see this, we consider two simple alternatives, namely, $G = G_1\ (> G_0)$ and $G = G_2\ (< G_0)$ for the null hypothesis $G = G_0$. For simple hypothesis against simple alternative we find the α - level best test by the application of the Neyman-Pearson lemma. Here Neyman-Pearson lemma gives the following α - level best test function for the $H:\ G = G_0$ against $A:\ G = G_1\ (> G_0)$ and also against $A:\ G = G_2 (< G_0)$

$$\phi(t) = \begin{cases} 0 & \text{if } G_0 \leq t_1 \leq G_0 - \frac{1}{n}\theta_0 \ln \alpha \\ 1 & \text{if } t_1 < G_0 \quad \text{or} \quad t_1 > G_0 - \frac{1}{n}\theta_0 \ln \alpha . \end{cases}$$

This test function is identical with the equivalent α - level test function based on the likelihood ratio test and furthermore it is independent of G_1 and G_2 and hence of every alternative for the hypothesis $G = G_0$; so clearly the likelihood ratio test is UMP. This property is also the subject of Lehmann's problem 13.2 (i) of page 110. [2].

For testing the H: $G = G_0$ against A: $G \neq G_0$ when the scale parameter θ is unknown, the likelihood ratio λ based on the first $r(\leq n)$ ordered observations out of a random sample of size n is given by

$$\lambda = \left[\frac{\sum_{2}^{r} (t_i - t_1) + (n-r)(t_r - t_1)}{\sum_{1}^{r} (t_i - G_0) + (n-r)(t_r - G_0)} \right]^{r}$$

which can be written as

$$\lambda^{\frac{1}{r}} = \left[1 + \frac{n(t_1 - G_0)}{\sum_{2}^{r} (t_i - t_1) + (n-r)(t_r - t_1)} \right]^{-1}$$

From [1] and [19] it follows that $\frac{2}{\theta} n(t_1 - G_0)$ and $\frac{2}{\theta}\left[\sum_{2}^{r} (t_i - t_1) + (n-r)(t_r - t_1)\right]$ are independently distributed as chi-squares with 2 and $2r-2$ degrees of freedom respectively and hence the quantity,

$$z = \frac{n(r-1)(t_1 - G_0)}{\sum_{i=2}^{r} (t_i - t_1) + (n-r)(t_r - t_1)}$$

has the well-known F distribution with 2 and 2r-2 degrees of freedom. The statistic Z is clearly equivalent to the above likelihood ratio test. Thus the α - level test function for the aforementioned hypothesis is:

$$\phi(t) = \begin{cases} 0 & \text{if } 0 \le Z \le b \\ 1, & \text{otherwise}, \end{cases}$$

where b is determined from the F - table. The use of this test function is equivalent to the decision rule: accept H when $G_0 < t_1 < G_0 + \frac{bu}{n}$, where

$$u = \frac{\sum_{2}^{r}(t_i - t_1) + (n-r)(t_r - t_1)}{r-1}, \text{ whose p.d.f. is given by}$$

$$f(u) = \begin{cases} \dfrac{(r-1)^{r-1} u^{r-2} e^{-\frac{1}{\theta}(r-1)u}}{\theta^{r-1}(r-2)!}, & u > 0 \\ 0, & \text{elsewhere} \end{cases}$$

and p.d.f. of t_1 is given by

$$f(t_1) = \begin{cases} \frac{n}{\theta} e^{-\frac{n}{\theta}(t_1 - G)}, & t_1 > G \\ 0, & \text{elsewhere.} \end{cases}$$

We may note here that t_1 and u are independently distributed so that $f(t_1, u) = f(t_1)\, f(u)$. The power function is derived now.

Case I: $G \le G_0$.

The relations $G \leq G_0 < t_1 < G_0 + \frac{bu}{n} < \infty$ gives

$$P(G) = 1 - \int_0^\infty f(u) \left[\int_{G_0}^{G_0 + \frac{bu}{n}} f(t_1)\, dt_1 \right] du = 1 - (1-\alpha) e^{\frac{n}{\theta}(G - G_0)};$$

where $\alpha = \left(1 + \frac{b}{r-1}\right)^{-(r-1)}$ which follows from the fact that

$$\int_b^\infty f(z)\, dz = \alpha .$$

Case II: $G \geq G_0$.

Here the relations $G_0 \leq G < t_1 < G_0 + \frac{bu}{n} < \infty$ gives

$$P(G) = \int_{\frac{n}{b}(G - G_0)}^{\infty} f(u) \left[\int_G^{G_0 + \frac{bu}{n}} f(t_1)\, dt_1 \right] du$$

$$= \alpha e^{\frac{n}{\theta}(G - G_0)} + I\left[r-1; \frac{n(r-1)(G-G_0)}{\theta b} \right]$$

$$- \alpha e^{\frac{n}{\theta}(G - G_0)} I\left[r-1; \frac{n(r-1+b)(G-G_0)}{\theta b} \right],$$

where $I[p; x] = \frac{\Gamma_x(p)}{\Gamma(p)} = \frac{\int_0^x x^{p-1} e^{-x} dx}{\int_0^\infty x^{p-1} e^{-x} dx}$

which is the form in which the Incomplete Gamma Function has been tabulated [5].

Having derived the power function we show now that the likelihood ratio test is <u>unbiased</u>.

For $G \leq G_0$, $e^{\frac{n}{\theta}(G-G_0)} \leq 1$ which means

$$P(G) = 1-(1-\alpha)e^{\frac{n}{\theta}(G-G_0)} \geq \alpha .$$

For $G \geq G_0$, we write

$$P(G) = 1 - \int_{\frac{n}{b}(G-G_0)}^{\infty} f(u)\,du \int_{G}^{G_0+\frac{bu}{n}} f(t_1)\,dt_1$$

$$= 1 + \int_{\frac{n}{b}(G-G_0)}^{\infty} \left[e^{\frac{1}{\theta}[n(G-G_0)-bu]} - 1 \right] f(u)\,du$$

Differentiating $P(G)$ with respect to G,

$$P'(G) = \frac{n}{\theta} \int_{\frac{n}{b}(G-G_0)}^{\infty} e^{\frac{1}{\theta}[n(G-G_0)-bu]} f(u)\,du$$

Recalling that $G \geq G_0$, the integral expression for $P'(G)$ is clearly positive which implies that $P(G)$ is a monotone non-decreasing function of G for $G \geq G_0$ and we have shown earlier that for $G \leq G_0$, $P(G) \geq \alpha$ hence it follows that for any G, $P(G) \geq \alpha$ and so the likelihood ratio test is <u>unbiased</u>.

The likelihood ratio test for the alternative $G \neq G_0$, θ unknown, is not UMP. But when we restrict ourselves to a class of unbiased tests there does exist a UMP unbiased test. Such UMP unbiased test function is derived by making use of the hint given in Lehmann ([2], Problem 12 ii), page 202). This is

$$\phi(t) = \begin{cases} 0 & \text{if } 0 \leq \dfrac{n(r-1)(t_1 - G_0)}{\sum_{2}^{r}(t_i - t_1) + (n-r)(t_r - t_1)} = Z \leq c \\ 1, & \text{otherwise}, \end{cases}$$

where as shown earlier Z has F distribution with 2 and $2r-2$ degrees of freedom.

It may be noted that our UMP unbiased test function and our unbiased likelihood ratio test function for the H: $G = G_0$ versus A: $G \neq G_0$ are identical.

For testing the H: $G = G_0$ and $\theta = \theta_0$ versus A: $G < G_0$ and $\theta < \theta_0$, the UMP test based on the first $r (\leq n)$ ordered observations is obtained by means of the Neyman-Pearson lemma. The test function is given by

$$\phi(t) = \begin{cases} 0 \text{ if } \sum_1^r (t_i - G_0) + (n-r)(t_r - G_0) \geq c \\ 1, \quad \text{otherwise} \end{cases},$$

where under the null hypothesis $\frac{2}{\theta_0}\left[\sum_1^r (t_i - G_0) + (n-r)(t_r - G_0)\right]$

has chi-square distribution with $2r$ degrees of freedom. The proof that the above test function indeed provides the UMP test for the H: $G = G_0$ and $\theta = \theta_0$ versus A: $G < G_0$ and $\theta < \theta_0$ follows directly from the similar arguments of Neyman and Pearson [6]. It may be remarked at this stage that we have obtained earlier the exactly same test function for testing the H: $\theta = \theta_1$ against A: $\theta = \theta_2$ $(< \theta_1)$ assuming G to be known, say G_0.

A likelihood ratio test based on the first $r_1 (\leq n_1)$ and $r_2 (\leq n_2)$ observations out of two ordered samples of sizes n_1 and n_2 respectively drawn randomly from exponential failure laws,

$$f_1(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{1}{\theta}(x - G_1)}, & x > G_1 \\ 0, & \text{otherwise} \end{cases}$$

and,

$$f_2(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{1}{\theta}(x - G_2)}, & x > G_2 \\ 0, & \text{otherwise}, \end{cases}$$

for testing H: $G_1 = G_2$ assuming θ, the common scale parameter unknown has been derived by Epstein and Tsao [7]. Noting

$t_{11} < t_{12} < \dots < t_{1r_1}$ and $t_{21} < t_{22} < \dots < t_{2r_2}$ are the first r_1 $(\leq n_1)$ and r_2 $(\leq n_2)$ observations of random samples of sizes n_1 and n_2 respectively, an equivalent likelihood ratio test function is given by

$$\phi(t) = \begin{cases} 0 & \text{if } 0 < w < c \\ 1, & \text{otherwise} \end{cases}$$

where

$$w = \begin{cases} \dfrac{(r_1+r_2-2)n_1(t_{11}-t_{21})}{u} & \text{if } t_{11} > t_{21} \\[2ex] \dfrac{(r_1+r_2-2)n_2(t_{21}-t_{11})}{u} & \text{if } t_{21} > t_{11} \end{cases}$$

with

$$u = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}(t_{ij}-t_{i1}) + (n_i-r_i)(t_{ir_i}-t_{i1})\right]$$

The statistic w has an F distribution with 2 and $2r_1 + 2r_2 - 4$ degrees of freedom. Paulson [4] has considered the same hypothesis and has selected another equivalent (differing by constant only) likelihood ratio test function for this hypothesis when $r_1 = n_1$ and $r_2 = n_2$. Furthermore he has shown that the likelihood ratio test for this hypothesis is unbiased and has expressed its power function in terms of the Incomplete Gamma Function. For $r_1 \leq n_1$ and $r_2 \leq n_2$, the Paulson's form of the test function is given by

$$\phi(t) = \begin{cases} 0 \text{ if } 0 < \dfrac{w}{r_1 + r_2 - 2} < c_1 \\ 1, \quad \text{otherwise}, \end{cases}$$

where $c_1 = \dfrac{c}{r_1 + r_2 - 2}$ and w is the same as given above. Let $r = r_1 + r_2$. For the α - level test function we have $\alpha = (1+c_1)^{-(r-2)}$ which follows from the relation,

$$\int_{c_1(r-2)}^{\infty} f(w)\,dw = \alpha$$

Now we proceed to derive power function for the above test function and show that it is <u>unbiased</u> by following Paulson [4].

Writing
$$Z = \begin{cases} n_1(t_{11} - t_{21}) & \text{if} \quad t_{11} > t_{21} \\ n_2(t_{21} - t_{11}) & \text{if} \quad t_{21} > t_{11} \end{cases},$$

we have $\dfrac{w}{r-2} = \dfrac{Z}{u}$ which gives the acceptance region for the null hypothesis as: $0 \leq Z \leq c_1 u$, where the p.d.f. of u is given by

$$f(u) = \begin{cases} \dfrac{u^{r-3} e^{-\frac{u}{\theta}}}{\theta^{r-2}(r-3)!}, & u > 0. \\ 0, \quad \text{otherwise.} \end{cases}$$

The p.d.f. of Z is derived by observing that the probability that Z lies in any interval is the sum of the probabilities that $n_2(t_{21} - t_{11})$ and $n_1(t_{11} - t_{21})$ lie in that interval and by then using standard methods for finding the distribution of the difference of two random variables. For the case $H = G_2 - G_1 \geq 0$, the p.d.f. of Z is

$$f(Z) = \begin{cases} f_1(Z) = \dfrac{e^{-\frac{n_1}{\theta}H}}{(n_1+n_2)\theta}\left[n_1 e^{\frac{n_1 Z}{n_2 \theta}} + n_2 e^{-\frac{Z}{\theta}}\right], & 0 \leq Z \leq n_2 H \\[2ex] f_2(Z) = \dfrac{\left[n_1 e^{\frac{n_2 H}{\theta}} + n_2 e^{-\frac{n_1 H}{\theta}}\right]}{(n_1+n_2)\theta} e^{-\frac{Z}{\theta}}, & n_2 H \leq Z \leq \infty . \end{cases}$$

Likewise for the case $H \leq 0$, the p.d.f. of Z is

$$f^*(z) = \begin{cases} f_1^*(z) = \dfrac{e^{\frac{n_2}{\theta}H}}{(n_1+n_2)\theta}\left[n_2 e^{\frac{n_2 z}{n_1\theta}} + n_1 e^{-\frac{z}{\theta}}\right], & 0 \le Z \le -n_1 H \\[2ex] f_2^*(z) = \dfrac{\left[n_2 e^{-\frac{n_1 H}{\theta}} + n_1 e^{\frac{n_2 H}{\theta}}\right]}{(n_1+n_2)\theta} e^{-\frac{z}{\theta}}, & -n_1 H \le Z \le \infty \end{cases}$$

The power function, $P(H)$, for the case $H \geq 0$, is

$$P(H) = 1 - \left\{ \int_0^u du \int_0^{n_2 H} f_1(z) f(u)\, dz - \int_0^{\frac{n_2 H}{c_1}} du \int_{c_1 u}^{n_2 H} f_1(z) f(u)\, dz + \int_{\frac{n_2 H}{c_1}}^{\infty} du \int_{n_2 H}^{c_1 u} f_2(z) f(u)\, dz \right\}$$

Upon integrating out and simplifying, the power function becomes

$$P(H) = \alpha\left(\frac{n_2 e^{-\frac{n_1 H}{\theta}}}{n_1+n_2}\right) I\left[r-2; \frac{n_2 H}{c_1\theta}\right]$$

$$+\alpha\left(\frac{n_1 e^{\frac{n_2 H}{\theta}}}{n_1+n_2}\right)\left\{1 - I\left[r-2; \frac{n_2 H(1+c_1)}{c_1\theta}\right]\right\}$$

$$-\frac{n_2}{n_1+n_2} e^{-\frac{n_1 H}{\theta}} \left(\frac{n_2}{n_2 - n_1 c_1}\right)^{r-2} I\left[r-2; \frac{H(n_2 - n_1 c_1)}{c_1\theta}\right]$$

The power function for the case $H \leq 0$ is

$$P(H) = 1 - \left\{ \int_0^{\infty} du \int_0^{-n_1 H} f_1^*(z) f(u)\, du\, dz - \int_0^{-\frac{n_1 H}{e_1}} du \int_{e_1 u}^{-n_1 H} f_1^*(z) f(u)\, dz \right.$$

$$\left. + \int_{-\frac{n_1 H}{e_1}}^{\infty} du \int_{-n_1 H}^{e_1 u} f_2^*(z) f(u)\, du \right\}$$

Again on integrating out and simplifying, we get

$$P(H) = \alpha\left(\frac{n_1 e^{\frac{n_2 H}{\theta}}}{n_1 + n_2} \right) + I\left[r-2 ; -\frac{n_1 H}{e_1 \theta} \right]$$

$$+ \alpha\left(\frac{n_2 e^{-\frac{n_1 H}{\theta}}}{n_1 + n_2} \right) \left\{ 1 - I\left[r-2 ; \frac{-n_1 H (1+e_1)}{e_1 \theta} \right] \right\}$$

$$- \frac{n_1}{n_1 + n_2} e^{\frac{n_2 H}{\theta}} \left(\frac{n_1}{n_1 - n_2 e_1} \right)^{r-2} I\left[r-2 ; \frac{-H(n_1 - n_2 e_1)}{e_1 \theta} \right].$$

To show that $P(H) > \alpha$ when $H \neq 0$, it is sufficient to show that the derivative $p'(H) > 0$ when $H > 0$ and $P'(H) < 0$ when $H < 0$. Of course $P(H = 0) = \alpha$. For $H > 0$, we write $P(H)$ after integrating w.r.t. Z as

$$P(H) = 1 - \frac{n_2}{n_1+n_2}\left[1 - e^{-\frac{H}{\theta}(n_1+n_2)}\right]$$

$$+ \int_0^{\frac{n_2H}{e_1}} \frac{n_2 e^{-\frac{n_1H}{\theta}}}{n_1+n_2}\left[-e^{\frac{n_1H}{n_2\theta}} - e^{-\frac{Z}{\theta}}\right]_{e_1u}^{n_2H} f(u)\,du$$

$$- \int_{\frac{n_2H}{e_1}}^{\infty} \frac{\left(n_1 e^{\frac{n_2H}{\theta}} + n_2 e^{-\frac{n_1H}{\theta}}\right)}{n_1+n_2}\left[-e^{-\frac{Z}{\theta}}\right]_{-n_2H}^{e_1u} f(u)\,du$$

where $[f(x)]_a^b = f(b) - f(a)$. Upon differentiating and simplifying we get

$$P'(H) = \frac{n_1n_2}{(n_1+n_2)\theta}\int_0^{\frac{n_2H}{e_1}} -e^{-\frac{n_1H}{\theta}}\left[e^{\frac{n_1e_1u}{n_2\theta}} - e^{-\frac{e_1u}{\theta}}\right] f(u)\,du$$

$$+ \frac{n_1n_2}{(n_1+n_2)\theta}\int_{\frac{n_2H}{\theta}}^{\infty} -e^{-\frac{e_1u}{\theta}}\left[e^{\frac{n_2H}{\theta}} - e^{-\frac{n_1H}{\theta}}\right] f(u)\,du.$$

Both integrals are clearly positive, so $P'(H) > 0$ when $H > 0$. Similarly we can show that $P'(H) < 0$ when $H < 0$. Therefore, it follows that the test is unbiased.

Part 2.

In Part 1. we have shown that for testing the $H: G = G_0$ against $A: G \neq G_0$ assuming θ unknown, an equivalent likelihood ratio test function is:

$$\phi(t) = \begin{cases} 0 & \text{if } 0 \leq Z \leq b \\ 1, & \text{otherwise} \end{cases}$$

where

$$Z = \frac{n(r-1)(t_1 - G_0)}{\sum_{2}^{r}(t_i - t_1) + (n-r)(t_r - t_1)}$$

has an F distribution with 2 and $2r-2$ degrees of freedom. This test is UMP unbiased. Now we propose a new statistic denoted by $s_{r,n}$ which is analogous to Carlson's statistic h_n^* [8] to test the above hypothesis. This statistic is defined by

$$s_{r,n} = \frac{t_{1n} - G_0}{t_{rn} - t_{1n}}$$, or for convenience in writing $s_r = \frac{t_1 - G_0}{t_r - t_1}$.

A great advantage in choosing this statistic stems from the fact that it requires only two observations namely, the first and the rth one to test the hypothesis. However, recommendation regarding the use of this new statistic depends mainly on its possessing satisfactory properties of a good test function. Superficially, this statistic has properties similar to Student's t test in that it is homogeneous of degree zero in the variable $(t_1 - G_0)$, and the numerator and denominator are independently distributed. In the present discussion we hope to derive i) the p.d.f. and c.d.f. of $s_{r,n}$, ii) the expectation of $s_{r,n}$, iii) Two Moment-Recurrence formulas to compute variance etc. iv) the power function for this test function, and v) some properties of this test function.

Wherever we shall consider it necessary to use $s_{r,n}$ to avoid ambiguity we shall use it, otherwise we shall write s for $s_{r,n}$. Incidentally, it is easy to see that if we were to replace $t_2, t_3, \ldots, t_{r-1}$ by t_r in Z , it would reduce to

$$Z^* = \frac{n(r-1)(t_1 - G_0)}{(n-1)(t_r - t_1)} = \frac{n(r-1)}{n-1} s_{r,n}.$$

On the basis of our new statistic s , an α - level test function for the H: $G = G_0$ versus A: $G \neq G_0$, assuming θ unknown is defined as

$$\phi(t) = \begin{cases} 0 & \text{if } 0 \leq s \leq c \\ 1 & \text{if } s > c \end{cases},$$

where c is determined from the relation $\int_c^{\infty} f(s)\,ds = \alpha$; $f(s)$ being the p.d.f. of s under null hypothesis which we proceed to derive now.

The joint p.d.f. of x_i and x_j $(i < j)$, the ith and the jth ordered sample observations out of an ordered sample of size n drawn from any continuous p.d.f., say $f(X)$ is given by

$$f(x_i, x_j) = \begin{cases} \dfrac{n!}{(i-1)!\,(j-i-1)!\,(n-j)!}\left[F(x_i)\right]^{i-1} \cdot \left[F(x_j) - F(x_i)\right]^{j-i-1} \cdot \\ \quad \cdot \left[1 - F(x_j)\right]^{n-j} \cdot f(x_i) f(x_j), \\ \qquad -\infty < x_i < x_j < \infty . \\ 0, \quad \text{otherwise} \end{cases}$$

where $F(x)$ is the c.d.f. of X whose p.d.f. is $f(x)$. In the present case

$$f(X) = \begin{cases} \dfrac{1}{\theta} e^{-\frac{1}{\theta}(X - G)}, & X > G \\ 0, & \text{otherwise.} \end{cases}$$

Replacing x by t as we are dealing with time measurement and writing $i = 1$, $j = r$ we get the following joint distribution of t_1 and t_r under the null hypothesis.

$$f(t_1, t_r) = \begin{cases} \dfrac{n!}{(r-2)!\,(n-r)!\,\theta^2}\, e^{-\frac{1}{\theta}\left[(t_1 - G_0) + (n-r+1)(t_r - G_0)\right]} \cdot \\ \quad \cdot \left[e^{-\frac{1}{\theta}(t_1 - G_0)} - e^{-\frac{1}{\theta}(t_r - G_0)}\right]^{r-2}, \\ \qquad G_0 < t_1 < t_r < \infty . \\ 0, \quad \text{otherwise.} \end{cases}$$

Making the transformations, $u = t_1 - G_0$ and $v = t_r - t_1$, we get the following joint p.d.f. of u and v.

$$f(u,v) = \begin{cases} \dfrac{n!}{(r-2)!\,(n-r)!\,\theta^2}\, e^{-\frac{1}{\theta}\left[nu + (n-r+1)v\right]} \cdot \left(1 - e^{-\frac{v}{\theta}}\right)^{r-2}, & u, v > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Clearly we have here $f(u, v) = f(u)\, f(v)$ with ranges for u and v independent of each other, hence it immediately follows that u and v are independently distributed, an observation made in the beginning of the 2nd part of this chapter.

Now $s = \dfrac{t_1 - G_0}{t_r - t_1} = \dfrac{u}{v}$ gives

$$f(s,v) = \begin{cases} \dfrac{n!\, v}{(r-2)!\,(n-r)!\,\theta^2}\, e^{-\frac{1}{\theta}\left[(ns + n - r + 1)v\right]} \cdot \left(1 - e^{-\frac{v}{\theta}}\right)^{r-2}, & s, v > 0 \\ 0, & \text{otherwise.} \end{cases}$$

This finally gives the p.d.f. of s under null hypothesis as

$$f(s) = \frac{n!}{(r-2)!\,(n-r)!\,\theta^2} \int_0^\infty v\, e^{-\frac{1}{\theta}\left[ns + n - r + 1\right]v} \cdot \left(1 - e^{-\frac{v}{\theta}}\right)^{r-2} dv,$$

which after integration can be written as

$$f(s)=\begin{cases}\dfrac{n!}{(r-2)!\,(n-r)!}\displaystyle\sum_{j=0}^{r-2}\binom{r-2}{j}(-1)^j\frac{1}{(ns+n-r+1+j)^2} & \text{for } s \geq 0\\[2ex] 0\,, & \text{otherwise.}\end{cases}$$

Incidentally, $\int_0^\infty f(s)\,ds = 1$ proves the following interesting lemma.

<u>Lemma 1</u>:
$$(n-1)\binom{n-2}{r-2}\sum_{j=0}^{r-2}\binom{r-2}{j}(-1)^j\frac{1}{n-r+1+j} = 1 .$$

When $r = n$, $f(s)$ reduces to

$$f(s)=\frac{n!}{(n-2)!\,\theta^2}\int_0^\infty v\, e^{-\frac{1}{\theta}(ns+1)v}\left(1-e^{-\frac{v}{\theta}}\right)^{n-2} dv$$

which agrees with Carlson's [8] expression (2.5) where $\theta = \frac{1}{\alpha}$, $v = \hat{\omega}_n$, and $f(s) = g_4\,(h_n^*)$.

To compute α percentage points of $s_{r,n}$ we derive an expression for $F(s)$, the c.d.f. of s, so that for given α_0 we have the relation, $1 - \alpha_0 = F(s_0)$ from which we determine s_0 .

$$F(s) = \int_0^s f(s)\,ds$$

$$= \frac{n!}{(r-2)!\,(n-r)!\,\theta^2} \int_0^s \int_0^\infty v\,e^{-\frac{1}{\theta}(ns+n-r+1)v} \cdot \left(1 - e^{-\frac{v}{\theta}}\right)^{r-2} dv\,ds$$

$$= \frac{(n-1)!}{(r-2)!\,(n-r)!\,\theta} \int_0^\infty \left(1 - e^{-\frac{v}{\theta}}\right)^{r-2} \left[- \int_0^s -\frac{nv}{\theta} e^{-\frac{v}{\theta}(ns+n-r+1)} ds \right] dv$$

$$= \frac{(n-1)!}{(r-2)!\,(n-r)!\,\theta} \int_0^\infty \left(1 - e^{-\frac{v}{\theta}}\right)^{r-2} e^{-(n-r+1)\frac{v}{\theta}} \left(1 - e^{-ns\frac{v}{\theta}}\right) dv$$

$$= 1 - \frac{B(n+ns-r+1,\, r-1)}{B(n-r+1,\, r-1)} ,$$

where $B(p, q) = \int_0^1 x^{p-1}\,(1-x)^{q-1}\,dx$ is known as a Beta-function. Again $r = n$, gives $F(s) = 1 - (n-1)\,B(ns+1,\, n-1)$, which agrees with Carlson's expression (2.6) where $s = h_n^*$ and $F(s) \equiv G_4\,(h_n^*)$.

Now we derive the expression for expected value of s . This is obtained as

$$E\delta_{r,n} = \frac{n!}{(r-2)!(n-r)!\,\theta^2}\int_0^\infty\int_0^\infty s\,e^{-\frac{1}{\theta}(ns+n-r+1)v}\cdot v\left(1-e^{-\frac{v}{\theta}}\right)^{r-2} dv\,ds$$

$$= \frac{(n-1)!}{n(r-2)!(n-r)!}\int_0^\infty \frac{1}{v}e^{-\frac{1}{\theta}(n-r+1)v}\cdot\left(1-e^{-\frac{v}{\theta}}\right)^{r-2} dv.$$

By making the transformation $y = e^{-\frac{v}{\theta}}$, we write the above integral as

$$E\delta_{r,n} = -\frac{(n-1)!}{n(r-2)!(n-r)!}\int_0^1 \frac{y^{n-r}(1-y)^{r-2}}{\ln y}\,dy$$

Using the standard integral given in [9], namely,

$$\int_0^1 \frac{x^m - x^n}{\ln x}\,dx = \ln\left(\frac{m+1}{n+1}\right), \text{ for } m+1>0 \ \& \ n+1>0$$

and
writing $(1-y)^{r-2} = (1-y)(1-y)^{r-3}$

$$= (1-y)\sum_{t=0}^{r-3}\binom{r-3}{t}(-1)^t y^t$$, we get

$$E\delta_{r,n} = \frac{(n-1)!}{n(r-2)!(n-r)!}\left[\sum_{t=0}^{r-3}\binom{r-3}{t}(-1)^t\ln(n-r+t+2) - \sum_{t=0}^{r-3}\binom{r-3}{t}(-1)^t\ln(n-r+t+1)\right],$$

provided $r \geq 3$.

Letting $k = t + 1$ in the first expression occurring in the above parentheses,

$$E\,\delta_{r,n} = \frac{(n-1)!}{n(r-2)!\,(n-r)!}\left[\sum_{k=1}^{r-2}\binom{r-3}{k-1}(-1)^{k-1}\ln(n-r+k+1)\right.$$

$$\left.-\sum_{t=0}^{r-3}\binom{r-3}{t}(-1)^{t}\ln(n-r+t+1)\right].$$

Now renaming the index k as t and combining the sums,

$$E\,\delta_{r,n} = \frac{(n-1)!}{n(r-2)!\,(n-r)!}\left[\sum_{t=1}^{r-3}\left\{\binom{r-3}{t-1}+\binom{r-3}{t}\right\}(-1)^{t-1}\ln(n-r+t+1)\right.$$

$$\left.+\,(-1)^{r-3}\ln(n-1)-\ln(n-r+1)\right].$$

Using the famous identity, $\binom{r-3}{t-1}+\binom{r-3}{t}=\binom{r-2}{t}$,

$$E\,\delta_{r,n} = \frac{(n-1)!}{n(r-2)!\,(n-r)!}\left[(-1)^{r-3}\ln(n-1)-\ln(n-r+1)\right.$$

$$\left.+\sum_{t=1}^{r-3}\binom{r-2}{t}(-1)^{t-1}\ln(n-r+t+1)\right]$$

$$=\frac{(n-1)!}{n(r-2)!\,(n-r)!}\sum_{t=0}^{r-2}\binom{r-2}{t}(-1)^{t-1}\ln(n-r+t+1),$$

provided $r \geq 3$.

This is a simple direct form for calculating the expectation of $s_{r,n}$ Carlson's [8] expression (4.3) is a special case of the present expression for $Es_{r,n}$ when $r = n$. The variance and the higher moments of $s_{r,n}$ can be computed by means of either of the two recurrence formulas established below.

<u>Lemma 2</u>: For $k > 1$ and $k + 1 < r \leq n$, the kth moment of $s_{r,n}$ can be expressed as follows:

$$E s_{r,n}^{k} = \frac{k(n-r+1)}{n(k-1)}\left[E s_{r-1,n}^{k-1} - E s_{r,n}^{k-1}\right], \text{ and also}$$

$$= \frac{k(n-1)}{n(k-1)}\left[\left(\frac{n-1}{n}\right)^{k-1} E s_{r-1,n-1}^{k-1} - E s_{r,n}^{k-1}\right].$$

<u>Proof</u>:

$$E s_{r,n}^{k} = \frac{n!}{(r-2)!\,(n-r)!\,\theta^2}\int_0^\infty\int_0^\infty s^k e^{-\frac{1}{\theta}(rs+n-r+1)v}\cdot v\left(1-e^{-\frac{v}{\theta}}\right)^{r-2} dv\,ds$$

$$= \frac{(n-1)!\,k!\,\theta^{k-1}}{(r-2)!\,(n-r)!\,n^k}\int_0^\infty \frac{e^{-\frac{1}{\theta}(n-r+1)v}}{v^k}\left(1-e^{-\frac{v}{\theta}}\right)^{r-2} dv.$$

Wherever convenient we shall denote integral of the following type

$$\int_0^\infty \frac{e^{-\frac{1}{\theta}nv}}{v^k}\left(1-e^{-\frac{v}{\theta}}\right)^{r} dv \quad \text{by } S(n,r;k)$$

Now integrating by parts the integral $S(n-r+1,\ r-2;\ k)$ under the assumption of $k > 1$ and $k+1 < r \leq n$, we get

$$E\delta_{r,n}^{k} = \frac{(n-1)!\,k!\,\theta^{k-1}}{(r-2)!\,(n-r)!\,n^{k}}\left[\frac{(r-2)}{\theta(k-1)}S(n-r+2,r-3;k-1) - \frac{(n-r+1)}{\theta(k-1)}S(n-r+1,r-2;k-1)\right].$$

Noting that

$$S(n-r+2,r-3;k-1) = \frac{(r-3)!\,(n-r+1)!\,n^{k-1}}{(n-1)!\,(k-1)!\,\theta^{k-2}}\,E\delta_{r-1,n}^{k-1}$$

and,

$$S(n-r+1,r-2;k-1) = \frac{(r-2)!\,(n-r)!\,n^{k-1}}{(n-1)!\,(k-1)!\,\theta^{k-2}}\,E\delta_{r,n}^{k-1},$$

we finally get after simplification,

$$E\delta_{r,n}^{k} = \frac{k(n-r+1)}{n(k-1)}\left[E\delta_{r-1,n}^{k-1} - E\delta_{r,n}^{k-1}\right].$$

To show another relationship between (k-1)th and kth moments of s,

we write $e^{-\frac{1}{\theta}(n-r+2)v} = e^{-\frac{1}{\theta}(n-r+1)v}\left[1-\left(1-e^{-\frac{v}{\theta}}\right)\right]$

in the integrand of S(n-r+2, r-3; k-1) which gives

$$E\delta_{r,n}^{k} = \frac{(n-1)!\,k!\,\theta^{k-2}}{(k-1)(r-2)!\,(n-r)!\,n^{k}}\left[(r-2)S(n-r+1,r-3;k-1) - (n-1)\,S(n-r+1,r-2;k-1)\right]$$

Noting that

$$S(n-r+1,r-3;k-1) = \frac{(r-3)!\,(n-r)!\,(n-1)^{k-1}}{(k-1)!\,(n-2)!\,\theta^{k-2}}\,E\delta_{r-1,n-1}^{k-1}$$

and

$$S(n-r+1, r-2; k-1) = \frac{(r-2)!\,(n-r)!\,n^{k-1}}{(k-1)!\,(n-1)!\,\theta^{k-2}} \, E s_{r,n}^{k-1} ,$$

we finally get after simplification,

$$E s_{r,n}^{k} = \frac{k(n-1)}{n(k-1)} \left[\left(\frac{n-1}{n}\right)^{k-1} E s_{r-1,n-1}^{k-1} - E s_{r,n}^{k-1} \right] .$$

Q.E.D.

This investigation points out a slip in Carlson's lemma ([8], p. 52). The correct expression of his lemma is found to be

$$E h_n^{*k} = \frac{k(n-1)}{n(k-1)} \left[\left(\frac{n-1}{n}\right)^{k-1} E h_{n-1}^{*k-1} - E h_n^{*k-1} \right] .$$

In addition to this, his Table 2. does not seem to record correct numerical values. In particular, the correct value of $Eh_4^{*2} = 0.196$ against the recorded value 0.131 . Further work shows that numerical error is not due to the typographical error in the formula.

<u>Illustration</u>: Compute $Es_{4,5}^2$ by two formulas established in the above lemma and check them by direct evaluation of $Es_{4,5}^2$ from integration. The first formula gives

$$E s_{4,5}^2 = \frac{4}{5}\left[E s_{3,5} - E s_{4,5} \right]$$

$$= \frac{4}{5}\left[\frac{12}{5}(2\ln 2 - \ln 3) - \frac{12}{5}(2\ln 3 - 3\ln 2) \right]$$

$$= \frac{48}{25}\left[5\ln 2 - 3\ln 3\right].$$

The second formula gives

$$E\delta^2_{4,5} = \frac{8}{5}\left[\frac{4}{5}E\delta_{3,4} - E\delta_{4,5}\right]$$

$$= \frac{8}{5}\left[\frac{4}{5}\left\{\frac{3}{2}(\ln 3 - \ln 2)\right\} - \frac{12}{5}(2\ln 3 - 3\ln 2)\right]$$

$$= \frac{48}{25}\left[5\ln 2 - 3\ln 3\right].$$

The direct computation gives

$$E\delta^2_{4,5} = \frac{24}{25}\,\theta\, S(2,2;2)$$

$$= \frac{48}{25}\left[S(2,1;1) - 2S(2,2;1)\right]$$

$$= \frac{48}{25}\left[5\ln 2 - 3\ln 3\right].$$

Hence we have a perfect check.

Now immediately we get

$$\mathrm{Var}(\delta_{4,5}) = \frac{48}{25}(5\ln 2 - 3\ln 3) - \frac{144}{25}(2\ln 3 - 3\ln 2)^2 = 0.2460.$$

We now proceed to derive the power function for the test function

$$\phi(t) = \begin{cases} 0 & \text{if } 0 \le \dfrac{t_1 - \sigma_0}{t_r - t_1} \le c \\ 1, & \text{otherwise} \end{cases}$$

which has been suggested earlier for testing the H: $G = G_0$ against A: $G \neq G_0$ assuming θ to be unknown. Writing u for $t_r - t_1$, the above test function is equivalent to

$$\phi(t) = \begin{cases} 0 & \text{if } G_0 \leq t_1 \leq cu + G_0 \\ 1, & \text{otherwise.} \end{cases}$$

Here we recall that t_1 and u are independently distributed. The p.d.f. of t_1 and u are

$$f(t_1) = \begin{cases} \frac{n}{\theta} e^{-\frac{n}{\theta}(t_1 - G)}, & t_1 > G \\ 0, & \text{otherwise} \end{cases}$$

and,

$$f(u) = \begin{cases} \frac{(n-1)!}{(r-2)!(n-r)!\,\theta} e^{-(n-r+1)\frac{u}{\theta}} \left(1 - e^{-\frac{u}{\theta}}\right)^{r-2}, & u > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Now we are ready to derive expressions for power function by considering two cases.

Case I: $G \leq G_0$.

$$P(G) = 1 - \int_0^\infty f(u) \left[\int_{G_0}^{cu+G_0} f(t_1)\,dt_1 \right] du$$

$$= 1 - e^{\frac{n}{\theta}(G - G_0)} \int_0^\infty \left(1 - e^{-\frac{n}{\theta}uc}\right) f(u)\,du$$

$$= 1 - \frac{(n-1)!\, e^{\frac{n}{\theta}(G-G_0)}}{(r-2)!\,(n-r)!} \left[B(n-r+1, r-1) - B(n+nc+1-r, r-1) \right]$$

$$= 1 - (1-\alpha) e^{\frac{n}{\theta}(G - G_0)} \quad , \text{ where}$$

$$\alpha = \frac{B(n + nc + 1 - r, r-1)}{B(n - r + 1, r-1)}$$

which follows from the expression of cumulative distribution function given earlier. It may be noted that the power function for this simple test function is the same under Case I. as for the likelihood ratio test. Furthermore for $G \leq G_0$ it is clear that $P(G) \geq \alpha$ and $P(G_0) = \alpha$.

Case II: $G \geq G_0$.

$$P(G) = 1 - \int_{\frac{G-G_0}{c}}^{\infty} f(u) \left[\int_{G}^{cu+G_0} f(t_1)\, dt_1 \right] du$$

$$= 1 - \int_{\frac{G-G_0}{c}}^{\infty} f(u) \left[1 - e^{-\frac{n}{\theta}(cu + G_0 - G)} \right] du$$

$$= 1 - I_a(n - r + 1; r - 1) + \alpha e^{\frac{n}{\theta}(G - G_0)} I_a(n + nc - r + 1; r - 1)$$

where i) $a = e^{\frac{G_0 - G}{c\theta}}$ and

$$I_a(p;q) = \frac{\int_0^a x^{p-1}(1-x)^{q-1}\, dx}{\int_0^1 x^{p-1}(1-x)^{q-1}\, dx}$$

is the Incomplete Beta function whose values are tabulated in [20]. This gives $P(G_0) = \alpha$ as expected. Here for the Case II., the power function of the simple test function is different from the power function of the corresponding likelihood ratio test.

Writing $$P(G) = 1 + \int_{\frac{G-G_0}{c}}^{\infty} \left[e^{-\frac{n}{\theta}(cu+G_0-G)} - 1 \right] f(u)\,du$$

and taking its first derivative w.r.t. G we get

$$P'(G) = \frac{n}{\theta} \int_{\frac{G-G_0}{c}}^{\infty} e^{-\frac{n}{\theta}(cu+G_0-G)} f(u)\,du$$

which is positive for $G \geq G_0$ and $P(G_0) = \alpha$, hence $P(G) \geq \alpha$ for $G \geq G_0$ and earlier we have seen that for $G \leq G_0$, $P(G) \geq \alpha$ so, it immediately follows that the simple test is unbiased. It may be mentioned that the Incomplete Beta function can be expressed as cumulative binomial probabilities.

Now that we have shown that the simple test function is unbiased; it may be interesting to compute suitable power function tables and graphs to point out differences between the simple test function and the likelihood ratio test. In life testing situations, H: $G \geq G_0$ against A: $G < G_0$ is of interest and for this situation it is clear that both the likelihood ratio test and the simple test function are equally good in terms of the power and the unbiasedness of the test. Because the likelihood ratio provides UMP unbiased test it follows that in the class of unbiased tests the power of the likelihood ratio test is uniformly better than that of the simple test.

For computing the moments and especially the variance of this simple test function we have established two recurrence formulas earlier. We need not establish such recurrence formulas for computing moments of the likelihood ratio statistic, Z, since it has a well-known F distribution with 2 and 2r-2 degrees of freedom. Noting that

$$f(Z) = \begin{cases} \dfrac{1}{\left(1 + \frac{Z}{r-1}\right)^{r}}, & Z > 0 \\ 0, & \text{otherwise} \end{cases}$$

we have, $EZ^k = (r-1)^{k+1} B(r-k-1, k+1)$, valid for $r > k+2$, which can be further simplified to

$$EZ^k = \frac{(r-1)^k}{\binom{r-2}{k}}$$, where $\binom{r-2}{k}$ is a Binomial coefficient.

This gives $EZ = \frac{r-1}{r-2}$; $EZ^2 = \frac{2(r-1)^2}{(r-2)(r-3)}$ and

$$\operatorname{Var}(Z) = \frac{(r-1)^3}{(r-2)^2(r-3)} = \frac{1}{\left(1 - \frac{1}{r-1}\right)^2\left(1 - \frac{2}{r-1}\right)}$$

which is independent of sample size n . From the moment recurrence formulas for the simple test function it appears that its variance is not independent of sample size, n .

Now we propose two simple test functions for testing the $H: \theta = \theta_0$ against $A: \theta \neq \theta_0$ assuming G to be known and unknown respectively and derive their power functions. For G known, say G_0, we define a simple test function as

$$\phi(t) = \begin{cases} 0 & \text{if } c_1 \leq t_r - G_0 \leq c_2 \\ 1, & \text{otherwise;} \end{cases}$$

and for G unknown we define another simple test function as

$$\phi(t) = \begin{cases} 0 & \text{if } c_1 \leq t_r - t_1 \leq c_2 \\ 1, & \text{otherwise.} \end{cases}$$

The p.d.f. of t_r for the case when G is known, is given by

$$f(t_r) = \begin{cases} \dfrac{n!}{(r-1)!\,(n-r)!\,\theta} e^{-\frac{1}{\theta}(n-r+1)(t_r - G_0)} \times \left(1 - e^{-\frac{1}{\theta}(t_r - G_0)}\right)^{r-1}, & t_r > G_0 \\ 0, & \text{otherwise.} \end{cases}$$

Letting $t_r - G_0 = x$ we get

$$f(x) = \begin{cases} \dfrac{n!}{(r-1)!\,(n-r)!\,\theta} e^{-\frac{1}{\theta}(n-r+1)x} \cdot \left(1 - e^{-\frac{x}{\theta}}\right)^{r-1}, & x > 0. \\ 0, & \text{otherwise.} \end{cases}$$

Under null hypothesis we have

$$f(x) = \begin{cases} \dfrac{n!}{(r-1)!\,(n-r)!\,\theta_0} e^{-\frac{1}{\theta_0}(n-r+1)x} \cdot \left(1 - e^{-\frac{x}{\theta_0}}\right)^{r-1}, & x > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Letting $e^{-\frac{x}{\theta_0}} = y$, we get

$$f(y) = \begin{cases} \dfrac{n!}{(r-1)!\,(n-r)!} y^{n-r} (1-y)^{r-1}, & 0 < y < 1 \\ 0, & \text{otherwise.} \end{cases}$$

We determine c_1 and c_2 from the equation

$$\int_{e^{-\frac{c_2}{\theta_0}}}^{e^{-\frac{c_1}{\theta_0}}} f(y)\,dy = 1 - \alpha \qquad (5)$$

in conjunction with

$$\left[\frac{\partial}{\partial \theta} \int_{e^{-\frac{c_2}{\theta}}}^{e^{-\frac{c_1}{\theta}}} f(y)\,dy \right]_{\theta = \theta_0} = 0 \qquad (6)$$

The relation (6) yields

$$c_1 e^{-(n-r+1)\frac{c_1}{\theta_0}} \left(1 - e^{-\frac{c_1}{\theta_0}}\right)^{r-1} = c_2 e^{-(n-r+1)\frac{c_2}{\theta_0}} \left(1 - e^{-\frac{c_2}{\theta_0}}\right)^{r-1}.$$

The relation (5) can be reduced to the Incomplete Beta Function as

$$\alpha = 1 - I_{e^{-\frac{c_1}{\theta_0}}}[n-r+1; r] + I_{e^{-\frac{c_2}{\theta_0}}}[n-r+1; r] .$$

The numerical solution for c_1 and c_2 can be carried out by successive approximations or graphically. The power function is now given by

$$P(\theta) = 1 - I_{e^{-\frac{c_1}{\theta}}}[n-r+1; r] + I_{e^{-\frac{c_2}{\theta}}}[n-r+1; r]$$

When G is unknown, the p.d.f. of $t_r - t_1 = u$ is given by

$$f(u) = \begin{cases} \dfrac{(n-1)!}{(r-2)!(n-r)!\theta} e^{-(n-r+1)\frac{u}{\theta}} \left(1 - e^{-\frac{u}{\theta}}\right)^{r-2}, & u > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Letting $e^{-\frac{u}{\theta}} = v$ we get

$$f(v) = \begin{cases} \dfrac{(n-1)!}{(r-2)!(n-r)!} v^{n-r} (1-v)^{r-2}, & 0 < v < 1. \\ 0, & \text{otherwise.} \end{cases}$$

We determine c_1 and c_2 from the equation

$$\int_{e^{-\frac{c_2}{\theta_0}}}^{e^{-\frac{c_1}{\theta_0}}} f(v)\,dv = 1-\alpha \tag{7}$$

in conjunction with

$$\left[\frac{\partial}{\partial\theta}\int_{e^{-\frac{c_2}{\theta}}}^{e^{-\frac{c_1}{\theta}}} f(v)\,dv\right]_{\theta=\theta_0} = 0 \tag{8}$$

The relation (8) gives

$$c_1 e^{-(n-r)\frac{c_1}{\theta_0}}\left(1-e^{-\frac{c_1}{\theta_0}}\right)^{r-2} = c_2 e^{-(n-r)\frac{c_2}{\theta_0}}\left(1-e^{-\frac{c_2}{\theta_0}}\right)^{r-2}.$$

The relation (7) can be expressed in terms of the Incomplete Beta Function as

$$\alpha = 1 - I_{e^{-\frac{c_1}{\theta_0}}}\left[n-r+1;\, r-1\right] + I_{e^{-\frac{c_2}{\theta_0}}}\left[n-r+1;\, r-1\right]$$

The power function of this test is given by

$$P(\theta) = 1 - I_{e^{-\frac{c_1}{\theta}}}\left[n-r+1;\, r-1\right] + I_{e^{-\frac{c_2}{\theta}}}\left[n-r+1;\, r-1\right].$$

IV. PERCENTILE ESTIMATORS FOR PARAMETERS OF EXPONENTIAL FAILURE LAW

The problem of obtaining percentile estimators for parameters of exponential failure laws and investigating some of their properties is taken up in this chapter. A percentile estimator for the shape parameter of the Weibull law is derived in the next chapter. Its expression being unwieldy, it has been thought useful to take up somewhat exhaustive investigation on percentile estimators of exponential laws first.

We write p.d.f. of exponential failure law as

$$f(t) = \begin{cases} \frac{1}{\theta} e^{-\frac{1}{\theta}(t-G)}, & t > G \in (-\infty, \infty) \ \& \ \theta \in (0, \infty). \\ 0, & \text{otherwise.} \end{cases}$$

For a given cumulative probability p, the percentile Υ_p is obtained from

$$p = \int_{G}^{\Upsilon_p} f(t)\, dt$$

which gives

$$\Upsilon_p = G - \theta \ln(1-p), \quad p \in (0,1).$$

Corresponding to population percentile Υ_p, we denote sample percentile by t_p and obtain the following percentile estimators for the parameters of the above failure law.

Case I. If $\theta = \theta_0$ is known, a percentile estimator for G, denoted by g, is given by

$$g = t_p + \theta_0 \ln(1-p) \text{ for any } p \in (0,1).$$

Case II. If $G = G_0$ is known, a percentile estimator for θ, denoted by x, is given by

$$x = (G_0 - t_p)[\ln(1-p)]^{-1} \text{ for any } p \in (0,1).$$

Case III. When both G and θ are unknown, percentile estimators for G and θ can be derived from the equations

$$t_{p_1} = G - \theta \ln(1-p_1) ,$$

$$t_{p_2} = G - \theta \ln(1-p_2) ,$$

where p_1 and p_2 both belong to unit open interval and are chosen in such a manner that $\mu_1 = [np_1] < \mu_2 = [np_2]$; n being the size of random sample drawn from the above exponential population and [np], as usual means the largest integer in np. Clearly $\mu_1 < \mu_2$ would mean

$p_1 < p_2$. The above two equations give the following percentile estimators for θ and G .

$$x = a(t_{p_2} - t_{p_1}) \quad \text{where} \quad a = [\ln(1-p_1) - \ln(1-p_2)]^{-1} > 0 \quad \text{and}$$

$$g = bt_{p_1} + (1-b)\, t_{p_2} \quad \text{where} \quad b = -a \ln(1-p_2) > 0 \ .$$

When $G = 0$, the above failure law reduces to the commonly used one-parameter exponential failure law in which case the percentile estimator for θ boils down to

$$x = -t_p[\ln(1-p)]^{-1} \text{ for } p \in (0,1) .$$

We may note that for $0 \leq p < 1$, we can write $-\ln(1-p) = \sum_{i=1}^{\infty} \frac{1}{i} p^i$.

Now we proceed to derive the sampling distributions of the above percentile estimators of G and θ and investigate some of their properties. Following Cramér ([10], pp. 367 ff), the p.d.f. of t_p , denoting for simplicity by t , in a random sample of size n is found to be

$$f(t) = \begin{cases} \dfrac{n!}{\mu!(n-\mu-1)!\,\theta} e^{-\frac{1}{\theta}(n-\mu)(t-G)} \left(1 - e^{-\frac{1}{\theta}(t-G)}\right)^{\mu}, & t > G. \\ 0 , & \text{otherwise} \end{cases}$$

where $\mu = [np]$ and np is not an integer. If np is an integer we are in the indeterminate case and t_p may be any value in the interval

(t_{np}, t_{np+1}). Since $g = t + \theta_0 \ln(1-p)$, we have

$$Eg^k = E[t + \theta_0 \ln(1-p)]^k = \sum_{j=0}^{k} \binom{k}{j} d^{k-j} Et^j,$$

where $d = \theta_0 \ln(1-p)$ and the expression for Et^j is obtained from $f(t)$ either by direct computation or by means of characteristic functions. From the results of Chapter II., it follows that

$$Eg^k = \frac{n!}{\mu!(n-\mu-1)!} \sum_{i=0}^{\mu} \sum_{j=0}^{k} \binom{\mu}{i} \binom{k}{j} \frac{(-1)^j j! [G + \theta_0 \ln(1-p)]^{k-j} \theta_0^j}{(n-\mu+i)^{j+1}}$$

which gives

$$Eg = G + \theta_0 \ln(1-p) + \theta_0 \sum_{i=0}^{\mu} \frac{1}{n-\mu+i},$$

and,

$$\operatorname{Var} g = \theta_0^2 \sum_{i=0}^{\mu} \frac{1}{(n-\mu+i)^2}.$$

Hence an unbiased estimator of G based on percentile estimator g is

$$t - \theta_0 \sum_{i=0}^{\mu} \frac{1}{n-\mu+i}.$$

Recalling that $\mu = [np] = np-q$ where $0 < q < 1$, it is easy to see that

$$\sum_{i=0}^{\mu} \frac{1}{n-\mu+i} \quad \text{is asymptotically equal to} \quad -\ln(1-p) \quad \text{and}$$

$$\sum_{i=0}^{\mu} \frac{1}{(n-\mu+i)^2} \quad \text{is asymptotically equal to} \quad \frac{p}{n(1-p)} \quad ,$$

which imply that for large n , expectation of g approaches G and variance of g approaches $\frac{\theta_0^2}{n} \frac{p}{1-p}$. Involing Cramér's Theorem ([10], p. 369), the above unbiased estimator of G based on percentile estimator has asymptotically normal distribution with mean G and variance $\frac{\theta_0^2}{n} \frac{p}{1-p}$.

Now in order to obtain minimum variance unbiased percentile estimator of G , we keep n fixed and choose p such that for large n , the variance , $\frac{\theta_0^2}{n} \frac{p}{1-p}$ is minimum. Clearly, any $p \in \left(\frac{1}{n}, \frac{2}{n}\right)$ provides such a percentile estimator of G . This means, $\mu = [np] = 1$, the first (or the smallest) sample observation out of an ordered sample of size n drawn from exponential population yields minimum variance unbiased percentile estimator of G . We note that the maximum likelihood estimator of G is the smallest sample observation.

Since for Case II., $x = (G_0 - t)\, \ln^{-1}(1-p)$, we have

$$Ex^k = E[c(t-G_0)]^k = c^k \sum_{j=0}^{k} \binom{k}{j} (-1)^j G_0^j\, E\,t^{k-j}$$

where $c = -\ln^{-1}(1-p) > 0$. Finally,

$$Ex^k = \frac{n!\,k!}{\mu!\,(n-\mu-1)!} (c\theta)^k \sum_{i=0}^{\mu} \binom{\mu}{i} (-1)^i \frac{1}{(n-\mu+i)^{k+1}} .$$

This gives

$$Ex = c\theta \sum_{i=0}^{\mu} \frac{1}{n-\mu+i} \quad \text{and} \quad \operatorname{Var} x = c^2\theta^2 \sum_{i=0}^{\mu} \frac{1}{(n-\mu+i)^2} .$$

Therefore, it is clear that $x\left[c \sum_{i=0}^{\mu} \frac{1}{n-\mu+i}\right]^{-1}$ is an unbiased estimator of θ with variance equal to

$$\theta^2 \sum_{i=0}^{\mu} \frac{1}{(n-\mu+i)^2} \left[\sum_{i=0}^{\mu} \frac{1}{n-\mu+i}\right]^{-2} .$$

Using asymptotic results relating to $\sum_{i=0}^{\mu} \frac{1}{n-\mu+i}$ and $\sum_{i=0}^{\mu} \frac{1}{(n-\mu+i)^2}$ and recalling that $c = -\ln^{-1}(1-p)$ we see that expectation of x, a percentile estimator of θ, approaches θ and its variance approaches $\frac{\theta^2}{n(1-p)} \frac{p}{\ln^2(1-p)}$. And invoking Cramér's Theorem ([10], p. 369) it follows that unbiased percentile estimator of θ has asymptotically normal distribution with mean θ and variance

$\frac{\theta^2 p}{n(1-p)} \ln^{-2}(1-p)$. We may now attempt to choose p such that the variance of unbiased percentile estimator of θ is minimum. Considering expression for such variance as a function of p, and setting its first derivative with respect to p equal to zero we find the equation, $2p + \ln(1-p) = 0$. Now p_0, the solution of this equation would insure minimum variance of unbiased percentile estimator of θ. By iterative procedure we obtain $p_0 = 0.797$. Hence $\mu = [np_0]$ gives the appropriate ordered sample observation which we should take to form unbiased percentile estimator of θ in order to have an assurance of minimum variance.

For Case III., percentile estimators of θ and G have been derived earlier. They are $x = a(t_{p_2} - t_{p_1})$ with

$a = \ln^{-1}\left(\frac{1-p_1}{1-p_2}\right) > 0$ and $g = bt_{p_1} + (1-b)\, t_{p_2}$ with

$b = \left[1 - \frac{\ln(1-p_1)}{\ln(1-p_2)}\right]^{-1} > 0$ where $p_1 \in (0, 1)$, $p_2 \in (0, 1)$,

$\mu_1 = [np_1] < [np_2] = \mu_2$ (implying $p_1 < p_2$). Noting that μ_1

and μ_2 are integers, for convenience, we replace μ_1 by i and μ_2 by j , satisfying $i < j$. The joint distribution of order statistics t_i and $t_j (i<j)$ in a random sample of size n drawn from exponential population is given by

$$f(t_i, t_j) = \begin{cases} \dfrac{n!\, e^{-\frac{1}{\theta}[(t_i - G) + (n-j+1)(t_j - G)]}}{(i-1)!(j-i-1)!(n-j)!\,\theta^2} \cdot \\ \cdot \left[1 - e^{-\frac{1}{\theta}(t_i - G)}\right]^{i-1} \cdot \left[e^{-\frac{1}{\theta}(t_i - G)} - e^{-\frac{1}{\theta}(t_j - G)}\right]^{j-i-1}, \\ \qquad G < t_i < t_j < \infty. \\ 0, \quad \text{otherwise.} \end{cases}$$

Now the suitable expression for percentile estimator of θ in terms of ith and jth order statistics is $x = a(t_j - t_i)$ where $i < j$ and a as already defined above. The p.d.f. of x is derived from $f(t_i, t_j)$ by integrating out t_j over $t_j > G + \frac{x}{a}$, after expressing t_i in terms of x and t_j . Thus, we get

$$f(x)=\begin{cases} \dfrac{n!}{(i-1)!\,(j-i-1)!\,(n-j)!\,a\theta}\times \\ \quad\times\displaystyle\sum_{k=0}^{i-1}\sum_{m=0}^{j-i-1}\binom{i-1}{k}\binom{j-i-1}{m}\times \\ \quad\times\dfrac{(-1)^{k+m}\,e^{-\frac{x}{a\theta}(n+m-i-j+1)}}{(n+k-i+1)}, & x>0 \\ 0, & \text{otherwise.}\end{cases}$$

This gives rth moment of x as

$$Ex^r=\frac{n!\,r!\,a^r\theta^r}{(i-1)!\,(j-i-1)!\,(n-j)!}\sum_{k=0}^{i-1}\sum_{m=0}^{j-i-1}\binom{i-1}{k}\binom{j-i-1}{m}\cdot$$

$$\times\frac{(-1)^{k+m}\,(n+m-i-j+1)^{-(r+1)}}{(n+k-i+1)}$$

where $a=\ln^{-1}\left(\frac{1-p_1}{1-p_2}\right)$. Since x is a linear function of t_i and t_j, it is clear that we can obtain moments of x in terms of joint moments of t_i and t_j. Thus

$$Ex^k=a^k\,E(t_j-t_i)^k=a^k\sum_{m=0}^{k}\binom{k}{m}(-1)^m\,E\,t_i^m\,t_j^{k-m}$$

By this formula we get

$$Ex = a(Et_j - Et_i) = a\theta\left[\sum_{k=0}^{j}\frac{1}{n-j+k} - \sum_{k=0}^{i}\frac{1}{n-i+k}\right]$$

and

$$\mathrm{Var}\,x = a^2\left[\mathrm{Var}\,t_i + \mathrm{Var}\,t_j - 2\,\mathrm{Cov}(t_i, t_j)\right],$$

where

$$\mathrm{Var}\,t_i = \theta^2 \sum_{k=0}^{i}\frac{1}{(n-i+k)^2},$$ and so Var t_j can be written

by replacing i by j in the expression for variance of t_i ; and

$$\mathrm{Cov}(t_i, t_j) = \theta^2 \sum_{k=0}^{i}\frac{1}{(n-i+k)^2}$$

which also follows from Sarhan [11]. Asymptotically expectation of x approaches θ and its variance approaches

$$\frac{\theta^2}{n}\,\frac{(p_2-p_1)}{(1-p_1)(1-p_2)}\,\ln^{-2}\left(\frac{1-p_1}{1-p_2}\right).$$

The above asymptotic results follow directly from the expressions given for expectation and variance of x . Alternatively these results could have been derived from Cramér's results ([10], pp. 369). Furthermore, Cramér's result would imply that x , percentile estimator of θ , has asymptotically normal distribution with mean θ and variance

$$\frac{\theta^2}{n}\frac{(p_2-p_1)}{(1-p_1)(1-p_2)}\ln^{-2}\left(\frac{1-p_1}{1-p_2}\right).$$

An unbiased percentile estimator of θ is clearly seen to be

$$x\left[a\left(\sum_{k=0}^{j}\frac{1}{n-j+k}-\sum_{k=0}^{i}\frac{1}{n-i+k}\right)\right]^{-1}$$

$$=(t_j-t_i)\left[\sum_{k=0}^{j}\frac{1}{n-j+k}-\sum_{k=0}^{i}\frac{1}{n-i+k}\right]^{-1}$$

whose variance can be written as

$$\theta^2\left[\sum_{k=0}^{j}\frac{1}{(n-j+k)^2}-\sum_{k=0}^{i}\frac{1}{(n-i+k)^2}\right]\times$$

$$\times\left[\sum_{k=0}^{j}\frac{1}{n-j+k}-\sum_{k=0}^{i}\frac{1}{n-i+k}\right]^{-2}.$$

Again invoking Cramér's Theorem ([10], pp. 369 ff) the following is true: This unbiased percentile estimator of θ has asymptotically the same normal distribution as the biased estimator x.

An appropriate expression for percentile estimator of G in terms of ith and jth order statistic is: $g = bt_i + (1-b)\, t_j$ where

$b = \left[1 - \frac{\ln(1-p_1)}{\ln(1-p_2)}\right]^{-1} > 0$ and $i < j$. We derive p.d.f. of g by making the transformation $g = bt_i + (1-b)\, t_j$, $t_j = t_j$ and integrating out t_j over $t_j > \frac{g-bG}{1-b}$. This gives

$$f(g) = \begin{cases} \dfrac{n!}{(i-1)!\,(j-i-1)!\,(n-j)!\, b\theta} \displaystyle\sum_{k=0}^{i-1} \sum_{m=0}^{j-i-1} \binom{i-1}{k} \binom{j-i-1}{m} \times \\ \times \dfrac{(-1)^{k+m}\, e^{-\frac{1}{\theta}\left[\left(\frac{n-i}{1-b} - \frac{1}{b} + j + i - m\right) g - \frac{n-i}{1-b} G\right]}}{n-i-(k+1)\left(\frac{1}{b}-1\right)}, & g > G. \\ 0, \text{ otherwise.} \end{cases}$$

Since g is a linear function of t_i and t_j, the formula for the rth moment of g is given by

$$E g^r = E\left[bt_i + (1-b)t_j\right]^r$$
$$= \sum_{l=0}^{r} \binom{r}{l} b^l (1-b)^{r-l}\, E t_i^r t_j^{r-l},$$

where product moments of t_i and t_j of various orders can be computed. Thus

$$E g = bEt_i + (1-b)\, Et_j$$
$$= G + \theta\left[b \sum_{k=0}^{i} \frac{1}{n-i+k} + (1-b) \sum_{k=0}^{j} \frac{1}{n-j+k}\right]$$

and,

$$\operatorname{Var} g = b^2 \operatorname{Var} t_i + (1-b)^2 \operatorname{Var} t_j + 2b(1-b)\operatorname{Cov}(t_i, t_j)$$

$$= b(2-b)\,\theta^2 \sum_{k=0}^{i} \frac{1}{(n-i+k)^2} + (1-b)^2 \theta^2 \sum_{k=0}^{j} \frac{1}{(n-j+k)^2}$$

Asymptotically expectation of g approaches G and variance of g approaches

$$\frac{\theta^2}{n}\left[(2b-b^2)\frac{p_1}{1-p_1} + (1-b)^2 \frac{p_2}{1-p_2}\right] \quad \text{where } b = \left[1 - \frac{\ln(1-p_1)}{\ln(1-p_2)}\right]^{-1}.$$

On the basis of expectation of g it is clear that

$$g - \theta\left[b \sum_{k=0}^{i} \frac{1}{n-i+k} + (1-b)\sum_{k=0}^{j} \frac{1}{n-j+k}\right]$$

is an unbiased percentile estimator of G , if θ is known. In case θ is unknown we need to replace θ by the unbiased percentile estimator of θ . For θ known, by use of Cramér's Theorem again g has asymptotically normal distribution with mean G and variance

$$\frac{\theta^2}{n}\left[\frac{p_2}{1-p_2} + \frac{p_2-p_1}{(1-p_1)(1-p_2)} \cdot \ln(1-p_2)\ln^{-1}\left(\frac{1-p_1}{1-p_2}\right)\left\{2 + \ln(1-p_2)\ln^{-1}\left(\frac{1-p_1}{1-p_2}\right)\right\}\right]$$

which is the same as given earlier except that b has been expressed in terms of p_1 and p_2 .

Now we proceed to derive joint distribution of x and g from joint distribution of t_i and t_j . The percentile estimators, $x = a(t_j - t_i)$ and $g = bt_i + (1-b)\, t_j$ for θ and G respectively can be conveniently expressed as

$$x = (t_i - t_j)\, \ln^{-1}\left(\frac{1-p_2}{1-p_1}\right), \quad \text{and}$$

$$g = [t_i \ln(1-p_2) - t_j \ln(1-p_1)]\ln^{-1}\left(\frac{1-p_2}{1-p_1}\right) .$$

Making the above transofrmations we get the following joint distribution of x and g from joint distribution of t_i and t_j .

$$f(x,g) = \begin{cases} \dfrac{n!\,\ln\left(\frac{1-p_1}{1-p_2}\right)}{(i-1)!\,(j-i-1)!\,(n-j)!\,\theta^2} \times e^{-\frac{1}{\theta}\left[g - x\ln(1-p_1) - G + (n-j+1)\{g - x\ln(1-p_2) - G\}\right]} \\ \quad \times \left[1 - e^{-\frac{1}{\theta}[g - x\ln(1-p_1) - G]}\right]^{i-1} \\ \quad \times \left[e^{-\frac{1}{\theta}[g - x\ln(1-p_1) - G]} - e^{-\frac{1}{\theta}[g - x\ln(1-p_2) - G]}\right]^{j-i-1}, \\ \qquad g > G + x\ln(1-p_1)\ \&\ x > 0. \\ 0, \quad \text{otherwise.} \end{cases}$$

From the joint distribution of x and g it does not seem convenient to obtain expression for covariance between x and g. However, since both x and g are linear functions of t_i and t_j we can compute covariance of x and g by knowing covariance of t_i and t_j. Thus,

$$\mathrm{Cov}(x,g) = \mathrm{Cov}\left[a(t_j - t_i),\, bt_i + (1-b)t_j\right]$$

$$= a(1-b)\,\mathrm{Var}\,t_j - ab\,\mathrm{Var}\,t_i + a(2b-1)\,\mathrm{Cov}(t_i, t_j).$$

Noting that in the present case expression for $\mathrm{Cov}(t_i, t_j) = \mathrm{Var}\,t_i$, the above expression reduces to

$$\mathrm{Cov}(x,g) = a(1-b)\left[\mathrm{Var}\,t_j - \mathrm{Var}\,t_i\right]$$

$$= a(1-b)\theta^2\left[\sum_{k=0}^{j}\frac{1}{(n-j+k)^2} - \sum_{k=0}^{i}\frac{1}{(n-i+k)^2}\right].$$

which is asymptotically equal to

$$\ln(1-p_1)\,\ln^{-2}\left(\frac{1-p_1}{1-p_2}\right)\frac{\theta^2}{n}\left[\frac{p_2-p_1}{(1-p_1)(1-p_2)}\right].$$

V. WEIBULL FAILURE LAWS

The probability density function (p.d.f.) of the 3-parameter Weibull failure law is given by

$$f(t) = \begin{cases} \dfrac{m(t-G)^{m-1}}{\theta} \; e^{-\frac{1}{\theta}(t-G)^m} , & t \geq G \in (-\infty, \infty) \;\; \& \; \theta, m \in (0, \infty) . \\ 0 \;, & \text{otherwise} \end{cases}$$

where G is the location parameter, known as guarantee time, θ, the scale parameter and m the shape parameter. When $m = 1$, the 3-parameter Weibull law reduces to the 2-parameter exponential law. When G is known we have the 2-parameter Weibull law and when both G and θ are known we have the 1-parameter Weibull law. Even here the shape parameter, m, being unknown, presents rather a difficult problem of estimation.

In the present investigation we shall work with the 3-parameter Weibull law. The results will, of course, remain valid for special cases of this law. It will also be possible to obtain some more interesting results in special cases.

The rth moment of the 3-parameter Weibull law is given by

$$Et^r = \sum_{k=0}^{r} \binom{r}{k} G^{r-k} \theta^{\frac{k}{m}} \Gamma\left(\frac{k}{m}+1\right) .$$

This gives

$$Et = G + \theta^{\frac{1}{m}} \Gamma\left(\frac{1}{m}+1\right) ;$$

$$\mathrm{Var}\, t = \theta^{\frac{2}{m}} \left[\Gamma\left(\frac{2}{m}+1\right) - \Gamma^2\left(\frac{1}{m}+1\right)\right]$$

β_1 (measure of skewness) = $\dfrac{\mu_3^2}{\mu_2^3}$

$$= \frac{\left\{\Gamma\left(\frac{3}{m}+1\right) - 3\Gamma\left(\frac{1}{m}+1\right)\Gamma\left(\frac{2}{m}+1\right) + 2\Gamma^3\left(\frac{1}{m}+1\right)\right\}^2}{\left\{\Gamma\left(\frac{2}{m}+1\right) - \Gamma^2\left(\frac{1}{m}+1\right)\right\}^3}$$

and

β_2 (measure of Kurtosis) = $\dfrac{\mu_4}{\mu_2^2}$

$$= \frac{\left\{\Gamma\left(\frac{4}{m}+1\right) - 4\Gamma\left(\frac{3}{m}+1\right)\Gamma\left(\frac{1}{m}+1\right) + 6\Gamma\left(\frac{2}{m}+1\right)\Gamma^2\left(\frac{1}{m}+1\right) - 3\Gamma^4\left(\frac{1}{m}+1\right)\right\}}{\left\{\Gamma\left(\frac{2}{m}+1\right) - \Gamma^2\left(\frac{1}{m}+1\right)\right\}^2}$$

where μ_2, μ_3 and μ_4 are the second, third and the fourth central moments.

We may note that the variance of the 3-parameter Weibull random

variable (r.v.) is a function of θ and m and its β_1 and β_2 functions of m. Recalling that β_1 and β_2 are measures relating to the shape of a frequency curve it seems appropriate to call m the shape parameter of Weibull law. The following relationship between the rth moment and the rth power of the first moment can be of use in investigating properties of the 3-parameter Weibull law.

Lemma 1: If

$$P_B(J = j) = \begin{cases} \binom{r}{j}\left[\frac{\theta^{\frac{1}{m}}\Gamma(\frac{1}{m}+1)}{Et}\right]^j\left(\frac{G}{Et}\right)^{r-j}, & j = 1, 2, \cdots r. \\ 0, & \text{otherwise} \end{cases}$$

then

$$Et^r = E_B \frac{\Gamma(\frac{J}{m}+1)}{\Gamma^J(\frac{1}{m}+1)} \cdot (Et)^r.$$

Proof:

$$Et^r = \sum_{j=0}^{r}\binom{r}{j} G^{r-j}\,\theta^{\frac{j}{m}}\,\Gamma\left(\frac{j}{m}+1\right)$$

$$= \sum_{j=0}^{r}\binom{r}{j}\frac{\Gamma(\frac{j}{m}+1)}{\Gamma^j(\frac{1}{m}+1)}\,p^j(1-p)^{r-j}\cdot(Et)^r,$$

with $p = \dfrac{\theta^{\frac{1}{m}}\Gamma(\frac{1}{m}+1)}{G+\theta^{\frac{1}{m}}\Gamma(\frac{1}{m}+1)}$

$$= E_B \frac{\Gamma(\frac{J}{m}+1)}{\Gamma^J(\frac{1}{m}+1)}(Et)^r.$$

Q.E.D.

From the above recurrence relationship, we have the following interesting results.

For $r = 1$, $E_B \left(\dfrac{\Gamma\left(\frac{J}{m}+1\right)}{\Gamma^J\left(\frac{1}{m}+1\right)} \right) = 1$.

Writing $\Gamma\left(\frac{j}{m}+1\right) = \frac{j}{m}\Gamma\left(\frac{j}{m}\right)$ and $\Gamma\left(\frac{1}{m}+1\right) = \frac{1}{m}\Gamma\left(\frac{1}{m}\right)$ we get,

for $r = 1$, $E_B \dfrac{J m^J \Gamma\left(\frac{J}{m}\right)}{\Gamma^J\left(\frac{1}{m}\right)} = m$. In general, for any $r > 1$,

$Et^r = \frac{1}{m} E_B \left(\dfrac{J m^J \Gamma\left(\frac{J}{m}\right)}{\Gamma^J\left(\frac{1}{m}\right)} \right) (Et)^r$. The following lemma is true.

<u>Lemma 2</u>: $\dfrac{\Gamma\left(\frac{j}{m}+1\right)}{\Gamma^j\left(\frac{1}{m}+1\right)} = \dfrac{j m^{j-1}}{\Pi B\left(\frac{j-i}{m}, \frac{1}{m}\right)}$, where

$$B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1}\, dx \quad \text{for} \quad a > 0, \quad b > 0,$$

is a Beta function.

<u>Proof</u>:

$$\Gamma\left(\frac{j}{m}+1\right) = \frac{j}{m}\Gamma\left(\frac{j}{m}\right) = \frac{j}{m}\Gamma\left(\frac{j-1}{m}+\frac{1}{m}\right) = \frac{j}{m}\,\frac{\Gamma\left(\frac{j-1}{m}\right)\Gamma\left(\frac{1}{m}\right)}{B\left(\frac{j-1}{m}, \frac{1}{m}\right)},$$

since $B(a, b) = \dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$. Writing $\Gamma\left(\frac{j-1}{m}\right) = \Gamma\left(\frac{j-2}{m}+\frac{1}{m}\right)$

and repeating this process several times we finally get

$$\Gamma\left(\frac{j}{m}+1\right) = \frac{j}{m}\,\frac{\Gamma^{j}\left(\frac{1}{m}\right)}{\prod_{i=1}^{j-1} B\left(\frac{j-i}{m}, \frac{1}{m}\right)}$$

and we have $\Gamma^{j}\left(\frac{1}{m}+1\right) = \left(\frac{1}{m}\Gamma\left(\frac{1}{m}\right)\right)^{j}$, hence

$$\frac{\Gamma\left(\frac{j}{m}+1\right)}{\Gamma^{j}\left(\frac{1}{m}+1\right)} = \frac{j\,m^{j-1}}{\prod_{i=1}^{j-1} B\left(\frac{j-i}{m}, \frac{1}{m}\right)}.$$

Q.E.D.

Now we can write the above recurrence relationship as

$$Et^{r} = E_{B}\left\{\frac{J\,m^{J-1}}{\prod_{i=1}^{J-1} B\left(\frac{J-i}{m}, \frac{1}{m}\right)}\right\}\cdot(Et)^{r}.$$

From the above results it also follows that

$$\frac{\Gamma\left(\frac{j}{m}\right)}{\Gamma^{j}\left(\frac{1}{m}\right)} = \frac{1}{\prod_{i=1}^{j-1} B\left(\frac{j-i}{m}, \frac{1}{m}\right)}.$$

The quantities β_1 and β_2 can be expressed in terms of Beta function as

$$\beta_1 = \frac{B\left(\frac{1}{m},\frac{1}{m}\right)\left[3m^2 - 2mB\left(\frac{2}{m},\frac{1}{m}\right) + 2B\left(\frac{1}{m},\frac{1}{m}\right)B\left(\frac{2}{m},\frac{1}{m}\right)\right]^2}{B^2\left(\frac{2}{m},\frac{1}{m}\right)\left[2m - B\left(\frac{1}{m},\frac{1}{m}\right)\right]^3}$$

and,

$$\beta_2 = \frac{B\left(\frac{1}{m},\frac{1}{m}\right)\left[4m^3 - 12m^2 B\left(\frac{3}{m},\frac{1}{m}\right) + 12mB\left(\frac{3}{m},\frac{1}{m}\right)B\left(\frac{2}{m},\frac{1}{m}\right) - 3\prod_{i=1}^{3} B\left(\frac{4-i}{m},\frac{1}{m}\right)\right]}{\prod_{i=1}^{3} B\left(\frac{4-i}{m},\frac{1}{m}\right)\left[2m - B\left(\frac{1}{m},\frac{1}{m}\right)\right]^2}$$

When $\frac{1}{m} = c$ is a positive integer, it is ture [21] that

$$B\left(\frac{1}{m},\frac{1}{m}\right) = B(c, c) = \left[c\binom{2c-1}{c-1}\right]^{-1} .$$

For the 2-parameter Weibull law whose p.d.f. is given by

$$f(t) = \begin{cases} \frac{mt^{m-1}}{\theta} e^{-\frac{1}{\theta}t^m}, & t > 0 \ \& \ \theta, m \in (0,\infty). \\ 0, & \text{otherwise} \end{cases}$$

we have

$$Et^r = \theta^{\frac{r}{m}} \Gamma\left(\frac{r}{m} + 1\right) .$$

This gives

$$Et = \theta^{\frac{1}{m}} \Gamma\left(\frac{1}{m} + 1\right) , \qquad \text{and}$$

$$\mathrm{Var}\ t = \theta^{\frac{2}{m}} \left[\Gamma\left(\frac{2}{m} + 1\right) - \Gamma^2 \left(\frac{1}{m} + 1\right)\right] .$$

In this case the recurrence relation between moments is relatively simple. In fact,

$$Et^r = \theta^{\frac{r}{m}} \Gamma\left(\frac{r}{m} + 1\right) = \theta^{\frac{r}{m}} \Gamma^r\left(\frac{1}{m} + 1\right) \cdot \frac{\Gamma\left(\frac{r}{m} + 1\right)}{\Gamma^r\left(\frac{1}{m} + 1\right)} = \frac{\Gamma\left(\frac{r}{m} + 1\right)}{\Gamma^r\left(\frac{1}{m} + 1\right)} \cdot (Et)^r .$$

This is a special case of Lemma 1 . Here we define

$$P(J = j) = \begin{cases} 1 & \text{if } j = r \\ 0 , & \text{otherwise.} \end{cases}$$

The above recurrence formula yields the following identities.

$$Et^r = \frac{\Gamma\left(\frac{r}{m}+1\right)}{\Gamma^r\left(\frac{1}{m}+1\right)} \cdot (Et)^r = \frac{r m^{r-1}}{\prod_{i=1}^{r-1} B\left(\frac{r-i}{m}, \frac{1}{m}\right)} (Et)^r = \frac{r m^{r-1} \Gamma\left(\frac{r}{m}\right)}{\Gamma^r\left(\frac{1}{m}\right)} (Et)^r.$$

Now without loss of generality, assuming $G = 0$ and $\theta = 1$, the p.d.f. of 1-parameter Weibull law is given by

$$f(t) = \begin{cases} mt^{m-1} e^{-t^m}, & t > 0, \quad m > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Here $Et^r = \Gamma\left(\frac{r}{m} + 1\right) = \frac{r}{m} \Gamma\left(\frac{r}{m}\right)$.

This gives

$$Et = \frac{1}{m} \Gamma\left(\frac{1}{m}\right) \quad \text{and}$$

$$\text{Var } t = \Gamma\left(\frac{2}{m} + 1\right) - \Gamma^2 \left(\frac{1}{m} + 1\right) \quad .$$

The recurrence formula for moments is

$$Et^r = \Gamma\left(\frac{r}{m} + 1\right) = \frac{\Gamma\left(\frac{r}{m}+1\right)}{\Gamma^r\left(\frac{1}{m}+1\right)} \cdot \Gamma^r\left(\frac{1}{m}+1\right) = \frac{\Gamma\left(\frac{r}{m}+1\right)}{\Gamma^r\left(\frac{1}{m}+1\right)} \cdot (Et)^r$$

which is identical with the recurrence formula established earlier in case of the 2-parameter Weibull law. This recurrence formula is also a special case of Lemma 1 with degenerate probability law at $j = r$.

Now we proceed to discuss the problem of the estimation of the parameters of Weibull laws. It is clear from the functional representation of Weibull laws that if m, the shape parameter of Weibull laws is known, the transformation, $u = (t-G)^m$ reduces the 3-parameter Weibull law to 1-parameter exponential law. With G also known, the above transformation being a parameter free transformation causes no difficulty in getting the maximum likelihood estimator (m.l.e.) of θ, the scale parameter of the Weibull law, based on the first $r(\leq n)$ ordered observations out of a random sample of size n. In fact, such m.l.e. of θ is found to be

$$\hat{\theta}_{r,n} = \frac{\sum_{i=1}^{r} (t_i - G)^m + (n-r)\ (t_r - G)^m}{r} .$$

It possesses all desirable properties of a good estimator, namely, consistency, unbiasedness, sufficiency, completeness and asymptotic normality. The proofs are exactly the same as given by Epstein and Sobel [1] and [19]. In case G is unknown but m known we suggest that G be estimated by the smallest sample observation (m.l.e.) which in life testing case is the first sample observation. The m.l.e. of θ is now found to be

$$\hat{\theta}_{r,n} = \frac{\sum_{i=2}^{r} (t_i - t_1)^m + (n-r)\ (t_r - t_1)^m}{r} .$$

It may be added that the m.l.e. of θ in case of the scale parameter Weibull law is a unique minimum variance unbiased estimator which follows from a theorem of Lehmann and Scheffe ([3], p. 61). In this case a single-observation minimum variance unbiased percentile estimator of θ can be obtained in exactly the same manner as has been explained in Chapter IV. of this work. When $m = 1$ and $G = G_0$, the 1-parameter Weibull law can be immediately reduced to the 1-parameter exponential law.

In this case the parameter of the Weibull law can be estimated most efficiently by the maximum likelihood method. And a single-observation minimum variance unbiased percentile estimator of the parameter has been the subject matter of discussion in Chapter IV. When $m = 1$ and G unknown, the 3-parameter Weibull law becomes the 2-parameter exponential law. The most efficient estimators for the parameters of the 2-parameter exponential law based on the first $r(\leq n)$ observations have been found by Epstein and Sobel [19]. But again if we wish to derive estimators of G and θ based on only two observations, percentile unbiased estimators for them have been obtained in Chapter IV. and we can insure minimum variance for this type of estimation by proper choice of cumulative probabilities. However when m, the shape parameter, is unknown and we are interested in getting good estimators for all the 3-parameters of Weibull law we face several difficulties. The likelihood equations to obtain m.l.e. for G, θ and m fail to provide explicit solutions for them. Kao [12], assuming $G = 0$, proceeds to derive m.l.e. for θ and m on the basis of the first $r(\leq n)$ ordered observations from a random sample of size n. His likelihood equations, namely,

$$\hat{\theta} = \frac{1}{r}\left\{ \sum_{i=1}^{r} t_i^{\hat{m}} + (n-r)\, t_r^{\hat{m}} \right\}$$

and

$$\theta = \frac{\sum_{i=1}^{r} t_i^{\hat{m}} \ln t_i + (n-r) t_r^{\hat{m}} \ln t_r}{\frac{r}{\hat{m}} + \sum_{i=1}^{r} \ln t_i}$$

clearly reveal the need for use of the successive approximation method. Of course, the similar situation will arise in case of the 3-parameter Weibull law. Here we shall first estimate G by the smallest sample observation which is the m.l.e. for G and then the m.l.e. for θ and m can be obtained in the above manner.

Duggan [13] has worked out the moment estimators for G, θ and m of the Weibull law. Again we do not have explicit solutions for G, θ and m . However his table seems to be convenient for computing such moment estimators. His numerical example based on the data pertaining to life of 34 automobile batteries provides negative estimate for G .

The recurrence formulas for moments of Weibull laws established earlier appear to throw more light on obtaining moment estimators for the parameters of Weibull laws. In case of the 1-parameter Weibull law our recurrence formula is:

$$E t^r = \Gamma\left(\frac{r}{m}+1\right) = \frac{r m^{r-1} \Gamma\left(\frac{r}{m}\right)}{\Gamma^r\left(\frac{1}{m}\right)} (Et)^r = \frac{r m^{r-1} (Et)^r}{\prod_{i=1}^{r-1} B\left(\frac{r-i}{m}, \frac{1}{m}\right)}$$

Equating population moments to sample moments we get, for $r = 1$,

i) $\bar{t}$ (sample mean) $= \frac{1}{m} \Gamma\left(\frac{1}{m}\right)$.

For $r = 2$,

ii) $\frac{\bar{t}^2}{(\bar{t})^2}$ (sample estimate) $= \frac{2m}{B\left(\frac{1}{m}, \frac{1}{m}\right)} = \frac{2m\Gamma\left(\frac{2}{m}\right)}{\Gamma^2\left(\frac{1}{m}\right)}$, where

$$\bar{t} = \frac{1}{n} \sum_{i=1}^{n} t_i \quad \text{and} \quad \bar{t}^2 = \frac{1}{n} \sum_{i=1}^{n} t_i^2 \text{, and so on.}$$

Thus for every r we have an equation in m which provides moment estimator for m, the shape parameter of the 1-parameter Weibull law. This raises a problem of investigating the effect of properties of moment estimator with respect to (w.r.t.) r, the order of moment an investigation which we do not intend to take up at the present time. While investigating this problem it seems fruitful to consider the consequences of directly computing moment estimator for the shape parameter from the expression of the rth moment since $Et^r = \Gamma\left(\frac{r}{m} + 1\right) = \frac{r}{m} \Gamma\left(\frac{r}{m}\right)$. For instance, if the 2nd moment is found to provide a better estimator for the shape parameter than the 1st moment, in that case the moment estimator should be computed from the equation, $\bar{t}^2 = \frac{2}{m} \Gamma\left(\frac{2}{m}\right)$, a relatively simpler expression to handle than ii) mentioned above.

From the p.d.f. of the 3-parameter Weibull law we obtain the following expressions for median and mode.

$$\lambda \text{ (median)} = G + \theta^{\frac{1}{m}} \, (\ln 2)^{\frac{1}{m}}, \qquad \text{and}$$

$$\eta \text{ (mode)} = \begin{cases} G + \theta^{\frac{1}{m}} \left(1 - \frac{1}{m}\right)^{\frac{1}{m}} & \text{when} \quad m > 1 \\ G, \quad \text{otherwise.} & \end{cases}$$

This gives the following expressions for median and mode for the 1-parameter Weibull law.

$$\lambda = (\ln 2)^{\frac{1}{m}} \qquad \text{and} \qquad \eta = \begin{cases} \left(1 - \frac{1}{m}\right)^{\frac{1}{m}} & \text{when} \quad m > 1 . \\ 0 & \text{otherwise} \end{cases}$$

Equating population median to sample median we get

$$t_{med} \text{ (sample median)} = (\ln 2)^{\frac{1}{m}} \qquad \text{which gives} \quad \hat{m} = \frac{\ln \ln 2}{\ln t_{med}} .$$

This is indeed a simple estimator for m whose exhaustive investigation should be taken up on a subsequent occasion. Equating population mode to sample mode does not provide such an explicit estimator for m as we have with the median. Here we have

$$t_{mode} \text{ (sample mode)} = \left(1 - \frac{1}{m}\right)^{\frac{1}{m}} .$$

Since the recurrence formula for the moments of the 2-parameter Weibull law is identical with that of the 1-parameter Weibull law we shall obtain the moment estimator of its shape parameter in a similar fashion and then proceed to derive moment estimator of its scale parameter. As for instance, if m is the moment estimator of its shape parameter then one moment estimator of the scale parameter, derived from the expression of its first moment, is found to be

$$\hat{\theta} \text{ (moment estimator)} = \left[\frac{\bar{t}}{\Gamma\left(\frac{1}{\hat{m}}+1\right)}\right]^{\hat{m}} = \frac{\hat{m}^{\hat{m}}\,\bar{t}^{\hat{m}}}{\Gamma^{\hat{m}}\left(\frac{1}{\hat{m}}\right)}$$

In case of the 2-parameter Weibull law,

$$\lambda \text{ (median)} = \theta^{\frac{1}{m}} (\ln 2)^{\frac{1}{m}} \qquad \text{and}$$

$$\eta \text{ (mode)} = \begin{cases} \theta^{\frac{1}{m}} \left(1 - \frac{1}{m}\right)^{\frac{1}{m}} & \text{if} \quad m > 1 . \\ 0 , & \text{otherwise.} \end{cases}$$

Equating population median and mode to sample median and mode we get

$$\frac{t_{med}}{t_{mode}} = \left(\frac{m}{m-1} \ln 2\right)^{\frac{1}{m}} \tag{1}$$

This gives, $m\,(\ln t_{med} - \ln t_{mode}) - \ln\left(\frac{m}{m-1}\right) - \ln\ln 2 = 0$.

Solving the above equation for m we have $\hat{\theta} = \frac{t_{med}^{\hat{m}}}{\ln 2}$.

From the relation (1) it is clear that the median and the mode of the Weibull law are quite apart provided m is less than two and away from one. In this situation it seems reasonable to use sample median and sample mode to estimate the parameters, θ and m of the Weibull law. When m is large it is not desirable to obtain estimators from sample median and mode for the parameters of Weibull law. Because in that case median and mode are very close to each other.

Recognizing the fact that the moment estimators are usually not as good as maximum likelihood estimators and furthermore realizing that both moment and maximum likelihood estimators have failed to provide equations explicitly solvable for the estimators of the parameters of Weibull laws, we proceed to present some other estimators for the parameters of Weibull laws in formula form so that it may be possible to improve these estimators by following the technique of generating BAN estimators from them. In Chapter IV., we have taken up the problem of deriving percentile estimators for the parameters of the exponential law and have investigated their properties. There we have mentioned that the subject matter of Chapter IV. has been the consequence of getting percentile estimator of the shape parameter of Weibull laws. Here we give such percentile estimators of the parameters of Weibull laws.

Corresponding to the given cumulative probability p, the population percentile γ_p for the 3-parameter Weibull law is found to be

$$\gamma_p = G + \theta^{\frac{1}{m}} [\ln(1-p)^{-1}]^{\frac{1}{m}} .$$

In case of the 1-parameter Weibull law where $G = 0$ and $\theta = 1$, we have

$$\tau_p = \ln^{\frac{1}{m}} (1-p)^{-1} = \ln^{\frac{1}{m}} \left(\frac{1}{1-p}\right) .$$

Equating population percentile to sample percentile we have

$$t_p \text{ (sample percentile)} = \ln^{\frac{1}{m}} \left(\frac{1}{1-p}\right) \qquad \text{which gives}$$

$$\hat{m} \text{ (percentile estimator)} = \frac{\ln\ln\left(\frac{1}{1-p}\right)}{\ln t_p} .$$

We may note that $p = \frac{1}{2}$ corresponds to sample median in which case we have shown earlier that $\frac{\ln\ln 2}{\ln t_{med}}$, which follows from the above percentile estimator when we put $p = \frac{1}{2}$.

The 1-parameter Weibull law admits any positive known value of scale parameter. If scale parameter is known θ_0, the percentile estimator of the shape parameter is given by

$$m = \frac{\ln\theta_0 + \ln\ln\left(\frac{1}{1-p}\right)}{\ln t_p} .$$

Here $\theta_0 = 1$ reduces this $\hat{m}$ to the former $\hat{m}$ which provides a check on the accuracy of the expression. In case of the 2-parameter Weibull law we have population percentile τ_p given by

$$\tau_p = \theta^{\frac{1}{m}} \ln^{\frac{1}{m}} \left(\frac{1}{1-p}\right) .$$

To obtain percentile estimators of θ and m we choose two cumulative probabilities p_1 and p_2 such that $[np_1] < [np_2]$ where n is the number of sample observations. When population percentiles γ_{p_1} and γ_{p_2} are equated to sample percentiles t_{p_1} and t_{p_2} we get

$$\hat{m} = \frac{\ln\ln\left(\frac{1}{1-p_1}\right) - \ln\ln\left(\frac{1}{1-p_2}\right)}{\ln t_{p_1} - \ln t_{p_2}} \quad \text{and} \quad \hat{\theta} = t_{p_1}^{\hat{m}} \cdot \ln^{-1}\left(\frac{1}{1-p}\right).$$

In case of the 3-parameter Weibull law let us pick up three cumulative probabilities, namely, p_1, p_2 and p_3 such that $[np_1] < [np_2] < [np_3]$. Equating population percentiles to sample percentiles we get the following equations,

$$t_{p_1} = G + \theta^{\frac{1}{m}} [\ln(1-p_1)^{-1}]^{\frac{1}{m}}$$

$$t_{p_2} = G + \theta^{\frac{1}{m}} [\ln(1-p_2)^{-1}]^{\frac{1}{m}}$$

$$t_{p_3} = G + \theta^{\frac{1}{m}} [\ln(1-p_3)^{-1}]^{\frac{1}{m}} \quad .$$

These equations give

$$\frac{t_{p_3} - t_{p_2}}{t_{p_2} - t_{p_1}} = \frac{[\ln(1-p_3)^{-1}]^{\frac{1}{m}} - [\ln(1-p_2)^{-1}]^{\frac{1}{m}}}{[\ln(1-p_2)^{-1}]^{\frac{1}{m}} - [\ln(1-p_1)^{-1}]^{\frac{1}{m}}} \quad ,$$

which provides estimator for m by the successive approximation procedure. Then,

$$\hat{\theta} = (t_{p_3} - t_{p_2})^{\hat{m}} \left(\left\{ \ln(1-p_3)^{-1} \right\}^{\frac{1}{\hat{m}}} - \left\{ \ln(1-p_2)^{-1} \right\}^{\frac{1}{\hat{m}}} \right)^{-\hat{m}} .$$

and

$$\hat{G} = t_{p_1} - \hat{\theta}^{\frac{1}{\hat{m}}} [\ln(1-p_1)^{-1}]^{\frac{1}{\hat{m}}} .$$

It is clear that the above percentile estimator of m can be obtained by successive approximations. This may not be convenient in many instances. But, we can derive a modification of the percentile estimator for m in formula form if we use an indirect satisfactory estimator for G. The smallest sample observation is the sufficient statistic for G and can be used as its estimator. Denoting a satisfactory estimator for G by G^*, modified percentile estimators for m and θ are

$$\hat{m} = \frac{\ln\ln(1-p_1)^{-1} - \ln\ln(1-p_2)^{-1}}{\ln(t_{p_1} - G^*) - \ln(t_{p_2} - G^*)}$$

and

$$\hat{\theta} = \frac{(t_{p_1} - G^*)^{\hat{m}}}{\ln(1-p_1)^{-1}}$$

respectively with t_{p_1} and t_{p_2} as sample percentiles corresponding to predetermined cumulative probabilities p_1 and p_2 satisfying $[np_1] < [np_2]$.

Now we present some other estimators for the parameters of Weibull laws. In case of the 3-parameter Weibull law, the expression for the cumulative density function (c.d.f.) is found to be

$$F(x) = 1 - e^{-\frac{1}{\theta}(x-G)^m}$$

which gives

$$m \ln(x-G) = \ln\theta + \ln\ln(1-F(x))^{-1} .$$

Noting that, $1 - F(x)$ is the probability that an item will survive beyond x , we call $1 - F(x) = R(x)$, the reliability of the item. The equation,

$$\ln\ln R^{-1}(x) = m\ln(x-G) - \ln\theta$$

is a linear function of $(x-G)$. It is known that any sample distribution function of a continuous random variable obeys the uniform law on the unit interval. For the sake of convenience, we denote $\ln\ln R^{-1}(x)$ by y . On the basis of sample observations: $t_1 < t_2 < \ldots < t_n$, we define

$$g(m, \theta, G) = \sum_{i=1}^{n} (y_i - m \ln(t_i - G) + \ln \theta)^2 \qquad (2)$$

Now, $\frac{\partial g}{\partial m} = 0$, $\frac{\partial g}{\partial \theta} = 0$, $\frac{\partial g}{\partial G} = 0$ (3)

yield 3 equations,

$$m \sum_{i=1}^{n} \ln^2 (t_i - G) - \ln \theta \sum_{i=1}^{n} \ln(t_i - G) = \sum_{i=1}^{n} y_i \ln(t_i - G)$$

$$m \sum_{i=1}^{n} \ln(t_i - G) - n \ln \theta = \sum_{i=1}^{n} y_i$$

and,

$$m \sum_{i=1}^{n} \frac{\ln(t_i - G)}{t_i - G} - \ln \theta \sum_{i=1}^{n} \frac{1}{t_i - G} = \sum_{i=1}^{n} \frac{y_i}{t_i - G} .$$

Here it is easy to get expressions for m and θ in terms of G . The real difficulty is in obtaining estimator for G . We can overcome this difficulty if we use an indirect satisfactory estimator for G . One such estimator for G has been pointed out earlier.

By means of the equations (2) and (3) we derive estimators for the parameters of Weibull laws under various situations.

i) 1-parameter Weibull law:

a) Special Case: $G = 0$ and $\theta = 1$.

$$\hat{m} = \frac{\sum_{i=1}^{n} y_i \ln t_i}{\sum_{i=1}^{n} \ln^2 t_i} .$$

b) General: $G = G_0$ and $\theta = \theta_0$.

$$\hat{m} = \frac{\sum_{i=1}^{n} y_i \ln(t_i - G_0) + \ln \theta_0 \sum_{i=1}^{n} \ln(t_i - G_0)}{\sum_{i=1}^{n} \ln^2 (t_i - G_0)}$$

ii) 2-parameter Weibull law:

a) Special Case: $G = 0$

$$\hat{m} = \frac{\sum_{i=1}^{n} (y_i - \bar{y}) \ln t_i}{\sum_{i=1}^{n} (\ln t_i - \overline{\ln t})^2}$$

and,

$$\hat{\theta} = e^{-\bar{y} + \hat{m} \, \overline{\ln t}} ,$$

where

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \ , \quad \bar{t} = \frac{1}{n} \sum_{i=1}^{n} t_i \quad \text{and} \quad \overline{\ln t} = \frac{1}{n} \sum_{i=1}^{n} \overline{\ln t_i}$$

b) General: $G = G_0$

$$\hat{m} = \frac{\sum_{i=1}^{n} (y_i - \bar{y}) \ln(t_i - G_0)}{\sum_{i=1}^{n} [\ln(t_i - G_0) - \overline{\ln(t - G_0)}]^2}$$

and,

$$\hat{\theta} = e^{-\bar{y} + \hat{m}\, \overline{\ln(t - G_0)}} \quad .$$

iii) 3-parameter Weibull law:

Here using G^* as a satisfactory indirect estimator for G ,

we have

$$m = \frac{\sum_{i=1}^{n} (y_i - \bar{y}) \ln(t_i - G^*)}{\sum_{i=1}^{n} [\ln(t_i - G^*) - \overline{\ln(t - G^*)}]^2}$$

and,

$$\hat{\theta} = e^{-\bar{y} + \hat{m}\, \overline{\ln(t - G^*)}} \quad .$$

In case of the 2-parameter Weibull law we can derive estimators for m and θ in formula form from another consideration as well. In the field of life testing, the concept of intensity function, $\lambda(t)$ (also called force of mortality or hazard rate) plays a very useful role. This is defined as

$$\lambda(t) = \frac{f(t)}{R(t)} = \frac{\text{probability density at } t \text{ of a failure time random variable}}{\text{reliability function at time } t \text{ of the item under consideration.}}$$

Now for the 2-parameter Weibull law with $G = 0$, $\lambda(t) = \frac{mt^{m-1}}{\theta}$. This gives $\ln\lambda(t) = \ln(m) + (m-1)\ln t - \ln\theta$ which is a linear function in t. We can convert sample observations to the data on intensity function by following Lomax [18]. Let us denote $\ln\lambda(t)$ by z. On the basis of sample observations: $t_1 < t_2 < \ldots < t_n$, we define

$$h(m, \theta) = \sum_{i=1}^{n} (z_i - \ln(m) - (m-1)\ln t_i + \ln\theta)^2 .$$

Here $\frac{\partial h}{\partial m} = 0$ and $\frac{\partial h}{\partial \theta} = 0$ yield

$$\hat{m} = 1 + \frac{\sum_{i=1}^{n} (z_i - \bar{z})\ln t_i}{\sum_{i=1}^{n} (\ln t_i - \overline{\ln t})^2}$$

and

$$\hat{\theta} = \hat{m}\, e^{(\hat{m}-1)\,\overline{\ln t} - \bar{z}}$$

where

$$\bar{z} = \frac{1}{n} \sum_{i=1}^{n} z_i \quad .$$

VI. INTENSITY FUNCTION: GENERATOR OF FAILURE LAWS

The statistical analysis of data pertaining to life, death or failure time of inanimate and animate objects (e.g. i. length of life of electric bulbs, electron tubes etc. which are specimens of industrial production and, ii. reaction time observed while determining the effect of drugs on mice, rats etc.) and also fatigue of men, machines etc. can be successfully conducted only when we correctly know the probability density function (p.d.f.) of the random variable (r.v.) concerned. The problem of actual determination of the p.d.f. of a r.v. arising in the field of life testing has not yet received due attention from the statisticians. On the basis of empirical evidence of Davis [15], the exponential law was taken as a good first approximation to the distribution of length of life. Epstein, Sobel and others have made useful statistical contributions which are valid under the assumption of exponentiality. Realizing the limitations of this assumption, some work has been done with the Weibull law [12]. The arguments put forward in favor of the use of the Weibull law appear in observing that the intensity function, defined in Chapter V. of the r.v. representing length of life can change with time in contrast to the exponential law whose intensity function is constant in time. Since several industrial products show aging effect, it becomes apparent that the intensity function of such r.v. must essentially be a function of time. The matter does not seem to end here. The intensity function, as a matter of fact, appears to be a very useful tool in generating a large number of p.d.f.'s appropriate to life testing data.

The term, <u>intensity function</u>, is due to Gumbel [17]. It is synonymous to hazard rate or force of mortality in actuarial statistics. For the sake of convenience to the readers we restate the definition of the intensity function, $\lambda(t)$.

$$\lambda(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{R(t)} \quad , \qquad \text{provided} \quad t > G \quad ,$$

where $f(t)$ is the p.d.f. of a r.v. representing length of life of an item and $R(t) = 1-F(t)$ is the reliability function of the item which is the probability that an item will survive beyond a given time, t . Before we proceed further, it may be proper to list some of its simple properties.

i) $\lambda(t) \geq f(t)$ since $0 \leq F(t) \leq 1$.

This implies that $\lambda(t)$ is always non-negative.

ii) The reciprocal of the intensity function is called <u>Mills' ratio</u>. It has been studied by Mills, Gordon, Birnbaum, Des Raj etc. in different connections.

iii) $\lambda(t)$ may be independent of t ; it may increase with t without limit; it may converge toward a constant.

iv) $\lambda(t) = \frac{f(t)}{R(t)}$ with $t > G$, gives

$$\text{a)} \quad f(t) = \lambda(t)\, e^{-\int_G^t \lambda(x)\, dx} \qquad \text{and}$$

$$\text{b)} \quad f(t) = -R'(t) \qquad \text{where} \qquad t > G \; .$$

The proofs for a) and b) are immediate and hence we omit them. Unless otherwise specified G will, for convenience, generally be taken as zero in the following.

The intensity function of the 3-parameter Weibull law whose p.d.f. is

$$f(t) = \begin{cases} \dfrac{m(t-G)^{m-1}}{\theta}\, e^{-\frac{1}{\theta}(t-G)^m}, & t > G \in (-\infty, \infty) \;\&\; \theta, m \in (0, \infty) \; . \\ 0 \, , & \text{otherwise} \end{cases}$$

is found to be

$$\lambda(t) = \frac{m(t-G)^{m-1}}{\theta} \quad .$$

When $m = 1$, $\lambda(t) = \frac{1}{\theta}$ which is the intensity function for the 2-parameter exponential law. The simplicity of the intensity function for these failure laws, which have been found to agree well with empirical data in many cases, and the appeal of the idea that the intensity function, an instantaneous propensity to failure in an object with has survived to time t, should be a simple function of t, suggests that forms derived from other simple assumptions about the behavior of the intensity function may find application in a wider class of cases than those covered by the Weibull distributions. One naturally considers using a polynomial in t for the intensity function. If

$$\lambda(t) = \sum_{i=0}^{p} a_i t^i \qquad \text{then}$$

$$f(t) = \begin{cases} \sum_{i=0}^{p} a_i t^i e^{-\frac{\sum_{i=0}^{p} a_i t^{i+1}}{i+1}}, & t > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Unless the polynomial is restricted, we have the trouble of too many parameters to be able to tell without a large number of data whether the fit is good because of the appropriateness of the form or because of the number of parameters.

In some applications it is reasonable to assume that the intensity function is a decreasing function of time. Lomax [18] has pointed out that $\lambda(t) = \frac{a}{b+t}$ appears to be more appropriate for the data relating to retail, craft and service groups in business failure and $\lambda(t) = a\,e^{-bt}$ for manufacturing trades. Corresponding to $\lambda(t) = \frac{a}{b+t}$ we get

$$f(t) = \begin{cases} \frac{a}{b}\left(1 + \frac{t}{b}\right)^{-(a+1)}, & t > 0 \\ 0, & \text{otherwise,} \end{cases}$$

and corresponding to $\lambda(t) = a\,e^{-bt}$ we have

$$f(t) = \begin{cases} a\, e^{-\left[bt + \frac{a}{b}\left(1 - e^{-bt}\right)\right]}, & t > 0 \\ 0, & \text{otherwise.} \end{cases}$$

It is clear that $\frac{1}{\lambda(t)} = \frac{b}{a} + \frac{t}{a}$ is a linear function of t. Denoting $\frac{1}{\lambda(t)}$ by z we obtain some estimators for a and b on the basis of sample observations: $t_1 < t_2 < \dots < t_n$ in the following manner.

Let $$g(a, b) = \sum_{i=1}^{n} \left(z_i - \frac{b}{a} - \frac{t_i}{a} \right)^2 .$$

Now $\frac{\partial g(a, b)}{\partial a} = 0$ and $\frac{\partial g(a, b)}{\partial b} = 0$ yield two equations whose solutions are

$$\hat{a} = \frac{\bar{z} \sum_{i=1}^{n} t_i^2 - \bar{t} \sum_{i=1}^{n} t_i z_i}{\sum_{i=1}^{n} t_i z_i - \bar{z} \sum_{i=1}^{n} t_i}$$

and

$$\hat{b} = \frac{\hat{a} + \bar{t}}{\bar{z}}$$

where $\bar{z} = \frac{1}{n} \sum_{i=1}^{n} z_i$. Similarly from $\lambda(t) = a\, e^{-bt}$

we get $\ln \lambda(t) = \ln a - bt$ which is a linear function of t . Here

let $h(a, b) = \sum_{i=1}^{n} (y_i - \ln a + bt_i)^2$, where $y = \ln \lambda(t)$

and $t_1, t_2, \ldots, t_n$ are sample observations. Now $\frac{\partial h}{\partial a} = 0$ and $\frac{\partial h}{\partial b} = 0$ yield two equations whose solutions are

$$\hat{a} = e^{\bar{y} + \hat{b}\bar{t}} \quad , \qquad \text{and}$$

$$\hat{b} = \frac{\sum_{i=1}^{n} (t_i - \bar{t})(y_i - \bar{y})}{\sum_{i=1}^{n} (t_i - \bar{t})^2}$$

where $\bar{t}$ and $\bar{y}$ are arithmetic means.

Finally we generate failure law from the consideration of growth cruves. Here $\lambda(t) = \frac{1}{1 + e^{-(\alpha + \beta t)}}$, which is known as a logistic function, gives

$$f(t) = \begin{cases} \dfrac{e^{\alpha + \beta t}(1 + e^{\alpha})^{\frac{1}{\beta}}}{(1 + e^{\alpha + \beta t})^{1 + \frac{1}{\beta}}}, & t > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Furthermore, $\ln\left(\dfrac{\lambda(t)}{1 - \lambda(t)}\right) = \alpha + \beta t$ is a linear function of t. Hence letting

$$\phi(\alpha, \beta) = \sum_{i=1}^{n} (u_i - \alpha - \beta\, t_i)^2$$

Where $u = \ln\left(\dfrac{\lambda(t)}{1 - \lambda(t)}\right)$ and $t_1, t_2, \ldots, t_n$ are sample observations, $\dfrac{\partial \phi}{\partial \alpha} = 0$ and $\dfrac{\partial \phi}{\partial \beta} = 0$ yield

$\hat{\alpha} = \bar{u} - \hat{\beta}\,\bar{t}$, and

$$\hat{\beta} = \frac{\sum_{i=1}^{n} (t_i - \bar{t})(u_i - \bar{u})}{\sum_{i=1}^{n} (t_i - \bar{t})^2}$$

where $\bar{t}$ and $\bar{u}$ are arithmetic means.

The transformation $u = \int_0^t \lambda(x)\, dx$ is helpful in reducing unwieldy expressions of failure laws to relatively simple forms provided the parameters involved in the intensity function are known. In particular, with extreme value distributions which have possibility of applications in life testing problems the above transformation may prove of immense value.

BIBLIOGRAPHY

[1] Benjamin Epstein and Milton Sobel, "Life Testing," Journal of the American Statistical Association, Vol. 48(1953), pp. 486-501.

[2] E. L. Lehmann, Testing Statistical Hypotheses, John Wiley and Sons, 1959.

[3] D. A. S. Fraser, Nonparametric Methods in Statistics, John Wiley and Sons, 1957.

[4] Edward Paulson, "On Certain Likelihood-Ratio Tests Associated with the Exponential distribution," Annals of Mathematical Statistics, Vol. 12, (1941), pp. 301-306.

[5] K. Pearson (Editor), Tables of the Incomplete Gamma Function, Biometric Laboratory, London, 1922.

[6] J. Neyman and E. S. Pearson, "Sufficient Statistics and Uniformly Most Powerful Tests of Statistical Hypotheses," Statistical Research Memoirs, Vol. 1(1936), pp. 113-137.

[7] Benjamin Epstein and Chia Kuei Tsao, "Some Tests Based on Ordered Observations from Two Exponential Populations," Annals of Mathematical Statistics, Vol. 24(1953), pp. 458-466.

[8] Phillip G. Carlson, "Tests of Hypothesis on the Exponential Lower Limit," Skandinavisk Aktuarietidskrift, 1958- Häft 1-2 (pp. 47-54).

[9] B. O. Peirce, A Short Table of Integrals, Ginn and Co., 1929 (page 65, formula 514), 3d rev. ed.

[10] Harald Cramér, Mathematical Methods of Statistics, Princeton University Press, 1957.

[11] A. E. Sarhan, "Estimation of the Mean and Standard Deviation by Order Statistics," Annals of Mathematical Statistics, Vol. 25(1954), pp. 317-328.

[12] John H. K. Kao, "The Weibull Distribution in Life Testing of Electron Tubes," Unpublished paper.

[13] John A. Duggan, "Utilizing the Weibull Distribution in Life Testing," Engineering Memorandum No. 20, Bendix Aviation Corporation.

[14] B. Epstein, "Estimates of Mean Life Based on the rth Smallest Value in a Sample of Size n Drawn from an Exponential Population," Wayne University Technical Report No. 2, prepared under ONR Contract Nonr-451 (00), NR-042-017, July 1952.

[15] D. J. Davis, "An Analysis of Some Failure Data," Journal of the American Statistical Association, Vol. 47(1952), pp. 113-150.

[16] Paul Gunther, "Techniques for Statistical Analysis of Life Test Data," General Electric Report No. 456GL278, Nov. 23, 1956.

[17] E. J. Gumbel, Statistics of Extremes, Columbia University Press, 1958.

[18] K. S. Lomax, "Business Failures: Another Example of the Analysis of Failure Data," Journal of the American Statistical Association, Vol. 49(1954), pp. 847-852.

[19] B. Epstein and M. Sobel, "Some Theorems Relevant to Life Testing from an Exponential Population," Annals of Mathematical Statistics, Vol. 25(1954), pp. 373-381.

[20] K. Pearson (Editor), Tables of the Incomplete Beta-Function, Biometric Laboratory, London, 1933.

[21] Bateman Manuscript Project, Higher Transcendental Functions, Vol. 1, McGraw-Hill Book Company, Inc., 1953.

[22] E. T. Copson, Theory of Functions of a Complex Variable, Oxford University Press, 1935.

[23] B. Epstein, "Simple Estimators of the Parameters of Exponential Distributions When Samples Are Censored," Annals of the Institute of Statistical Mathematics, Vol. III(No. 1, pp. 15-25).

[24] Charles E. Clark and G. Trevor Williams, "Distributions of the Members of an Ordered Sample," Annals of Mathematical Statistics, Vol. 29(1928), pp. 862-870.

[25] Gunnar Blom, Statistical Estimates and Transformed Beta-Variables, John Wiley and Sons, 1958.

[26] F. N. David and N. L. Johnson, "Statistical Treatment of Censored Data," Biometrika, Vol. 41(1954), pp. 228-240.

[27] Charles E. Clark and G. Trevor Williams, "Distributions of the Members of an Ordered Sample--An Addendum," Annals of Mathematical Statistics, Vol.30(1959), p. 610.

[28] N. G. De Bruijn, Asymptotic Methods in Analysis, North-Holland Publishing Co. - Amsterdam, 1958.