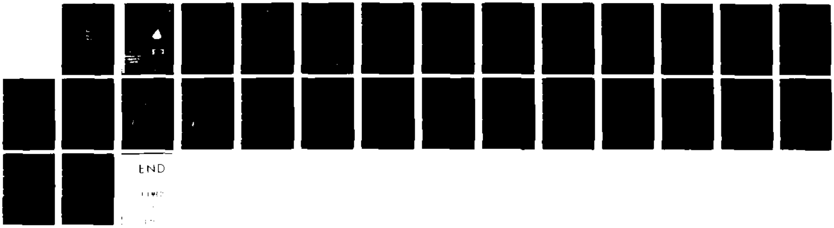
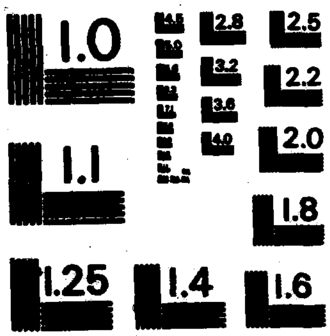
AD-A120 878 EXPECTED INFORMATION FROM A LIFE TEST EXPERIMENT(U) CALIFORNIA UNIV BERKELEY OPERATIONS RESEARCH CENTER R E BARLOW ET AL. MAY 82 ORC-82-5 AFOSR-81-0122 1/1

UNCLASSIFIED F/G 9/4 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

ORC 82-5
MAY 1982

12

EXPECTED INFORMATION FROM A LIFE TEST EXPERIMENT

by
RICHARD E. BARLOW
and
JAW HUAN HSIUNG

OPERATIONS
RESEARCH
CENTER

UNIVERSITY OF CALIFORNIA · BERKELEY

EXPECTED INFORMATION FROM A LIFE TEST EXPERIMENT

by

Richard E. Barlow
Department of Industrial Engineering
and Operations Research
University of California, Berkeley

and

Jaw Huan Hsiung
Department of Statistics
University of California, Berkeley

MAY 1982 ORC 82-5

This research was supported by the Air Force Office of Scientific Research (AFSC), USAF, under Grant AFOSR-81-0122 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ORC 82-5	2. GOVT ACCESSION NO. AD-A120 898	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) EXPECTED INFORMATION FROM A LIFE TEST EXPERIMENT		5. TYPE OF REPORT & PERIOD COVERED Research Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Richard E. Barlow and Jaw Huan Hsiung		8. CONTRACT OR GRANT NUMBER(s) AFOSR-81-0122
9. PERFORMING ORGANIZATION NAME AND ADDRESS Operations Research Center University of California Berkeley, California 94720		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS United States Air Force Air Force Office of Scientific Research Bolling Air Force Base, D.C. 20332		12. REPORT DATE May 1982
		13. NUMBER OF PAGES 28
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Sample Size Determination
Life Testing
Information
Bayesian Inference

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(SEE ABSTRACT)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ACKNOWLEDGEMENT

The computer program for computing information was originally developed for Dr. Ed Wu at Lawrence Livermore National Laboratory. The excellent typing was done by Mariko Kubik.

ABSTRACT

Expected information gain as a result of life testing n units for time t is calculated for the time transformed exponential model and a utility function based on entropy. We show that the expected information gain is concave increasing in n and a transform of the test time t. A computer program for calculating expected entropy for the Weibull distribution model is given. This may provide practical guidance in designing life test experiments.

EXPECTED INFORMATION FROM A LIFE TEST EXPERIMENT

by

Richard E. Barlow and Jaw Huan Hsiung

1. INTRODUCTION

In considering a life test experiment, two questions to be answered are:

How many items should be tested?

and

How long should we be prepared to wait before analyzing the data?

As Lindley (1956) pointed out, "the object of experimentation is (often) not to reach decisions, but rather to gain knowledge about the world." Hence, we do not consider the cost of experimentation directly in a conventional decision analysis approach to the solution of our problems. Instead, we consider the influence of sample size and test time on various measures of expected information to be gained.

Since the objective of testing is to gain *information* about life times of similar items, we need to determine how our expected measure of information to be gained depends on sample size as well as test time. By *information*, we mean anything which *changes* our probability distribution about unknown quantities. To measure this change we use a utility function, $u(\lambda, d(D))$, where λ is the unknown life distribution parameter of interest and the decision taken, $d(D)$, based on observed data D, will (in this paper) usually be identified with the posterior mean or the posterior density. The expected gain in information based on n observations can then be measured by

$$g(n) = E\left\{\max_d \int_\Lambda u(\lambda,d)\pi(\lambda \mid D,n)d\lambda\right\} - \max_d \int_\Lambda u(\lambda,d)\pi(\lambda)d\lambda \qquad (1.0)$$

where Λ is the parameter space, π is a prior density for λ and $\pi(\lambda \mid D)$ is the posterior density for λ given data D and d belongs to some appropriate decision space. Raiffa and Schlaifer (1961) call (1.0) the expected value of sample information. The expression is easily seen to be nonnegative. This idea of measuring expected information as expected utility has been discussed by DeGroot [(1970), pp. 429-433] and more recently by Bernardo (1979).

To illustrate ideas, first consider a non-life test situation where n normally distributed measurements are to be made. Suppose our uncertainty about measurement X given θ and σ^2 is measured by a $N(\theta,\sigma^2)$ distribution. For convenience, suppose σ^2 is known but θ is unknown so that we wish to learn about θ . Let our prior uncertainty for θ be measured by a $N\left(\theta_o,\gamma^2\right)$ distribution. Let $x_1,x_2, \ldots, x_n$ be n independent (given θ) observations so that

$$\bar{x} = (x_1 + \ldots + x_n)/n$$

given θ has a $N\left(\theta,\frac{\sigma^2}{n}\right)$ distribution while θ given $\bar{x}$ has a $N\left(\mu(\bar{x}),\tau_n^2\right)$ distribution with mean

$$\mu(\bar{x}) = (1 - w)\theta_o + w\bar{x}$$

and variance

$$\tau_n^2 = \left(\frac{1}{\gamma^2} + \frac{n}{\sigma^2}\right)^{-1}$$

where $w = \dfrac{\gamma^2}{\gamma^2 + \frac{\sigma^2}{n}}$. To measure information gained as a result of these n measurements, let our utility function be

$$u(\theta,d) = \begin{cases} 1 & \text{if } |\theta - d| < \varepsilon \\ 0 & \text{otherwise}. \end{cases} \tag{1.1}$$

Then

$$\underset{d}{\text{Maximum}} \int_{-\infty}^{\infty} u(\theta,d)\pi(\theta \mid \bar{x},n)d\theta$$

$$= \underset{d}{\text{Maximum}}\ P[|\theta - d| < \varepsilon \mid \bar{x},n]$$

$$= P[|\theta - \mu(\bar{x})| < \varepsilon \mid \bar{x},n]$$

so that $d = \mu(\bar{x})$ is our optimum "decision" in this case. Therefore, if we take n measurements, our expected utility will be

$$\begin{aligned} E\left\{\int_{\Theta} u(\theta,\mu(\bar{x}))\pi(\theta \mid \bar{x},n)d\theta\right\} &= E\{P[|\theta - \mu(\bar{x})| < \varepsilon \mid \bar{x},n]\} \\ &= \Phi\left(\frac{\varepsilon}{\tau_n}\right) - \Phi\left(-\frac{\varepsilon}{\tau_n}\right). \end{aligned} \tag{1.2}$$

This is our expected posterior probability that, after n measurements are made, θ will be within ε of the posterior mean $\mu(\bar{x})$. Φ is the cumulative $N(0,1)$ distribution. One way to determine n irrespective of cost considerations is to specify a probability p and require that

$$\Phi\left(\frac{\varepsilon}{\tau_n}\right) - \Phi\left(-\frac{\varepsilon}{\tau_n}\right) = p .$$

If

$$\int_{-z(p)}^{z(p)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du = p ,$$

then

$$\frac{\varepsilon}{\tau_n} = z(p)$$

and

$$n = \left[\left(\left(\frac{z(p)}{\varepsilon}\right)^2 - \frac{1}{\gamma^2}\right)\sigma^2\right]^+$$

where $[\cdot]^+$ denotes greater nonnegative integer in the quantity within brackets. If we let $\gamma = \infty$, then we have the *non-Bayesian* solution $n = \left(\frac{z(p)}{\varepsilon}\right)^2 \sigma^2$.

From (1.1) we see that our expected measured gain in information based on n observations will be

$$g(n) = E\left\{\int_{\textcircled{H}} u(\theta,\mu(\bar{x}))\pi(\theta \mid \bar{x},n)d\theta\right\} - \int_{\textcircled{H}} u(\theta,\theta_o)\pi(\theta)d\theta \qquad (1.3)$$

where θ_o is the prior mean and $\mu(\bar{x})$ is the posterior mean. It is easy to verify from (1.2) that $g(n)$ is *concave* increasing in n so that marginal gain is *decreasing* in sample size n .

There are several reasons why a utility function such as (1.1) and the expected measured gain (1.3) based on (1.1) might *not* serve as an adequate measure of information gained.

1. The measure (1.3) is *not* invariant under a 1-1 transformation of the parameter space. Hence an experiment based on n observations will produce a different measure of information gained for θ than for, say θ^3. (The same comment would apply if we were to use $u(\theta,d) = -(\theta - d)^2$.)

2. The "decision" produced from (1.1) is the posterior mean whereas we know that the posterior *density* carries all relevant information about θ based on our experiment. A utility function which produces the posterior density would be intuitively superior.

There is an essentially unique utility function which does produce the posterior density from the space of "decisions" corresponding to densities on the parameter space. Bernardo (1979) showed that this utility function is

$$u(\theta,d) = \log p(\theta)$$

where $p(\theta)$ is a density on Θ. The corresponding expected value of sample information is

$$g(n) = E\int_{\Theta}[\log \pi(\theta \mid D,n)]\pi(\theta \mid D,n)d\theta - \int_{\Theta}[\log \pi(\theta)]\pi(\theta)d\theta \qquad (1.4)$$

where $\pi(\theta \mid D)$ is the posterior density for θ. This measure was first introduced and studied by Lindley (1956). It is the negative change in entropy of our probability density for θ. Rewriting (1.4), we have

$$g(n) = \iint p(D,\theta) \log\left[\frac{p(D,\theta)}{p(D)\pi(\theta)}\right]dDd\theta \qquad (1.5)$$

where $p(D,\theta)$ is the joint density of data D and parameter θ while $p(D) = \int_{\Theta} p(D,\theta)d\theta$. If $\theta = g(w)$ is a 1-1 map of Θ onto Θ and J is the Jacobian of the transformation, then

$$g(n) = \iint p(D,g(w))|J| \log \left[\frac{p(D,g(w))}{p(D)\pi(g(w))}\right] dDdw$$

$$= \iint p(D,g(w))|J| \log \left[\frac{p(D,g(w))|J|}{p(D)\pi(g(w))|J|}\right] dDdw$$

$$= \iint p^*(D,w) \log \left[\frac{p^*(D,w)}{P(D)\pi^*(w)}\right] dDdw$$

where p^* is the joint density of D and w . Hence (1.4) is invariant under 1-1 transformations of the parameter space.

We will show for the time-transformed exponential life distribution model and the utility function

$$u(\theta,p(\cdot)) = \log p(\theta)$$

where

$$\max_{p(\cdot)} \int_{\Theta} [\log p(\theta)]\pi(\theta \mid D)d\theta$$

$$= \int_{\Theta} [\log \pi(\theta \mid D)]\pi(\theta \mid D)d\theta$$

that our expected measured gain in information is concave increasing in both sample size n and a transform of the test time t . Methods for calculating expected information with respect to a Weibull life distribution model are discussed.

2. INFORMATION FROM A LIFE TEST EXPERIMENT

Let $\bar{F}_o(x) = e^{-R_o(x)}$ be a specified absolutely continuous life distribution with hazard function R_o and failure rate $r_o(x) = \frac{d}{dx} R_o(x)$. Consider the life distribution model

$$\bar{F}(x \mid \lambda) = e^{-\lambda R_o(x)} \tag{2.1}$$

where λ is the unknown "proportional hazard" but R_o is specified. (2.1) is called the time-transformed exponential life distribution model. Suppose n similar units are put on life test for the time interval $[0,t]$ and we judge the model (2.1) to be an appropriate description of our uncertainty concerning the life length. If we observe k failures with lifetimes $x_1, x_2, \ldots, x_k$ and $n - k$ survivors in $[0,t]$, then the likelihood is

$$\begin{aligned} L(\lambda \mid x_1, x_2, \ldots, x_k, t) = \binom{n}{k} \lambda^k \left[\prod_{i=1}^{k} r_o(x_i) \right] \\ \cdot \exp\left[-\lambda \left[\sum_{i=1}^{k} R_o(x_i) + (n-k) R_o(t) \right] \right] . \end{aligned} \tag{2.2}$$

Clearly k and $s = \sum_{i=1}^{k} R_o(x_i) + (n-k) R_o(t)$ together constitute a sufficient statistic for λ. For some results, we will use the prior density

$$\pi(\lambda) = \frac{b^a \lambda^{a-1} e^{-b\lambda}}{\Gamma(a)} \qquad \lambda, a, b > 0 \tag{2.3}$$

and posterior density

$$\pi(\lambda \mid k,s) = \frac{(b+s)^{a+k}\lambda^{a+k-1}}{\Gamma(a+k)} e^{-(b+s)\lambda} .$$

2.1 A Measure of Information Based on Entropy

Lindley (1956) introduced the following measure of expected information gain as a result of performing an experiment $\mathcal{E}$ resulting in data D :

$$I(\mathcal{E},\pi(\lambda)) = E\int_\Lambda [\log \pi(\lambda \mid D)]\pi(\lambda \mid D)d\lambda - \int_\Lambda [\log \pi(\lambda)]\pi(\lambda)d\lambda \qquad (2.4)$$

where the expectation operator, E , is with respect to the unconditional distribution of the data D . Bernardo (1979) pointed out the connection with expected utility where the utility function

$$u(\lambda,\pi(\lambda \mid D)) = \log \pi(\lambda \mid D) \qquad (2.5)$$

depends on λ and the decision variable is $\pi(\lambda \mid D)$, the posterior density at λ . The entropy is $-\int_\Lambda [\log \pi(\lambda \mid D)]\pi(\lambda \mid D)d\lambda$ and (2.4) is the negative expected change in entropy as a result of performing $\mathcal{E}$.

Information measures based on (2.4) are dimensionless and as such may be difficult to interpret. However, (2.4) does provide a way of ordering proposed experiments by assigning information values which are invariant under 1-1 transformations of the parameter space. For example, suppose we life test n units for time t and use the Weibull life distribution model

$$P(X > x \mid \alpha,\lambda) = e^{-\lambda x^{\alpha}} \qquad (2.6)$$

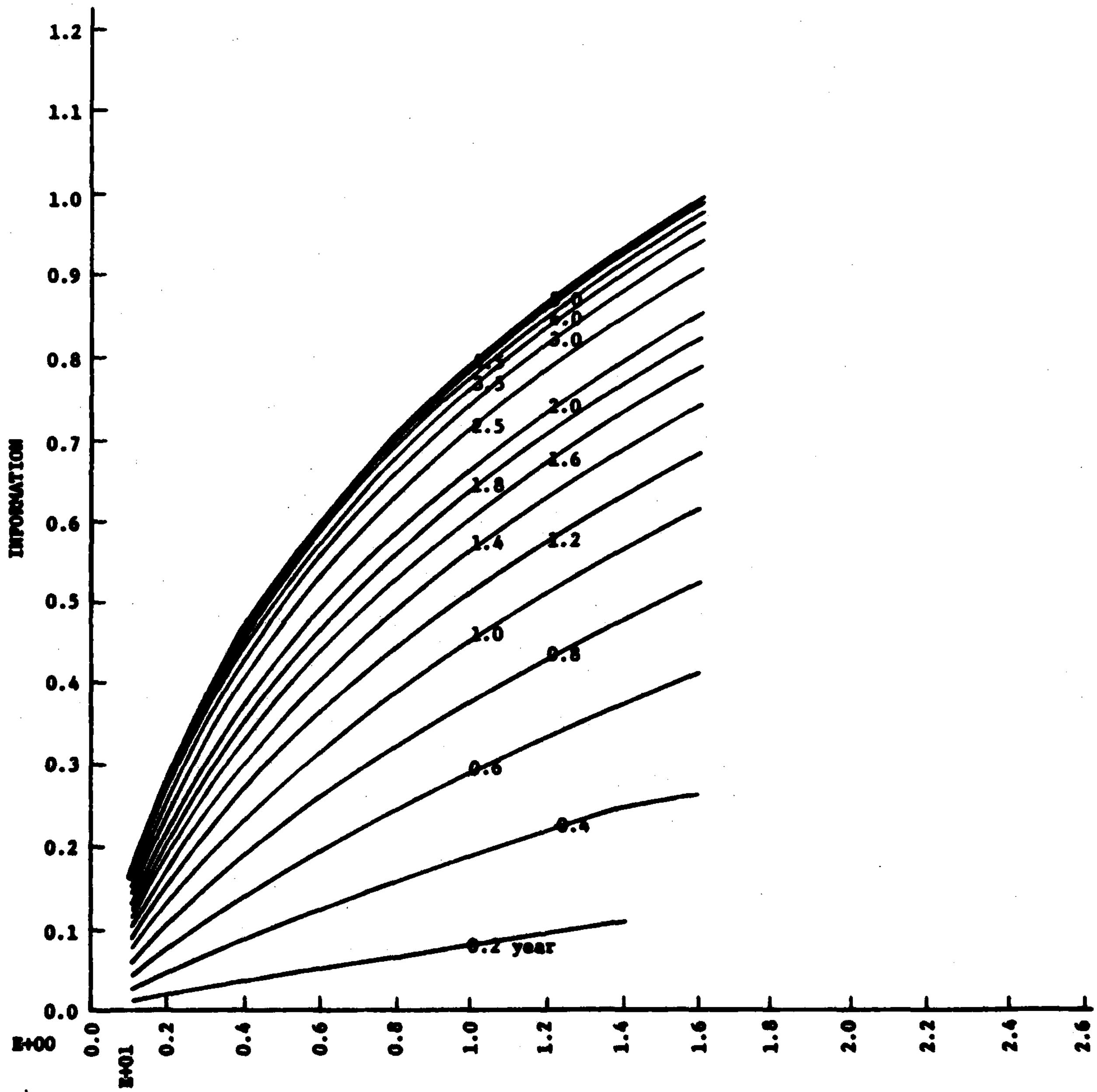
where α is known but λ unknown. Also let the prior for λ be

$$\pi(\lambda \mid A,B) = B^A \lambda^{A-1} e^{-B\lambda}/\Gamma(A) \ .$$

Figure 2.1 and 2.2 are example graphs of expected information versus sample size and test time respectively. For example, from the graphs, we can see that testing 3 units for 3 years results in the same information as testing 10 units for about 0.75 years. Thus, we have a means of comparing experiments. Information values can be related to specified experiments.

The parameters of the gamma prior used (A and B) were originally specified based on the pressure vessel data analyzed in Barlow, Toland and Freeman (1979). The shape parameter $\alpha = 1.5$ was used. The graphs show that for these parameter values (A and B) there is little to be gained by testing more than 3 years.

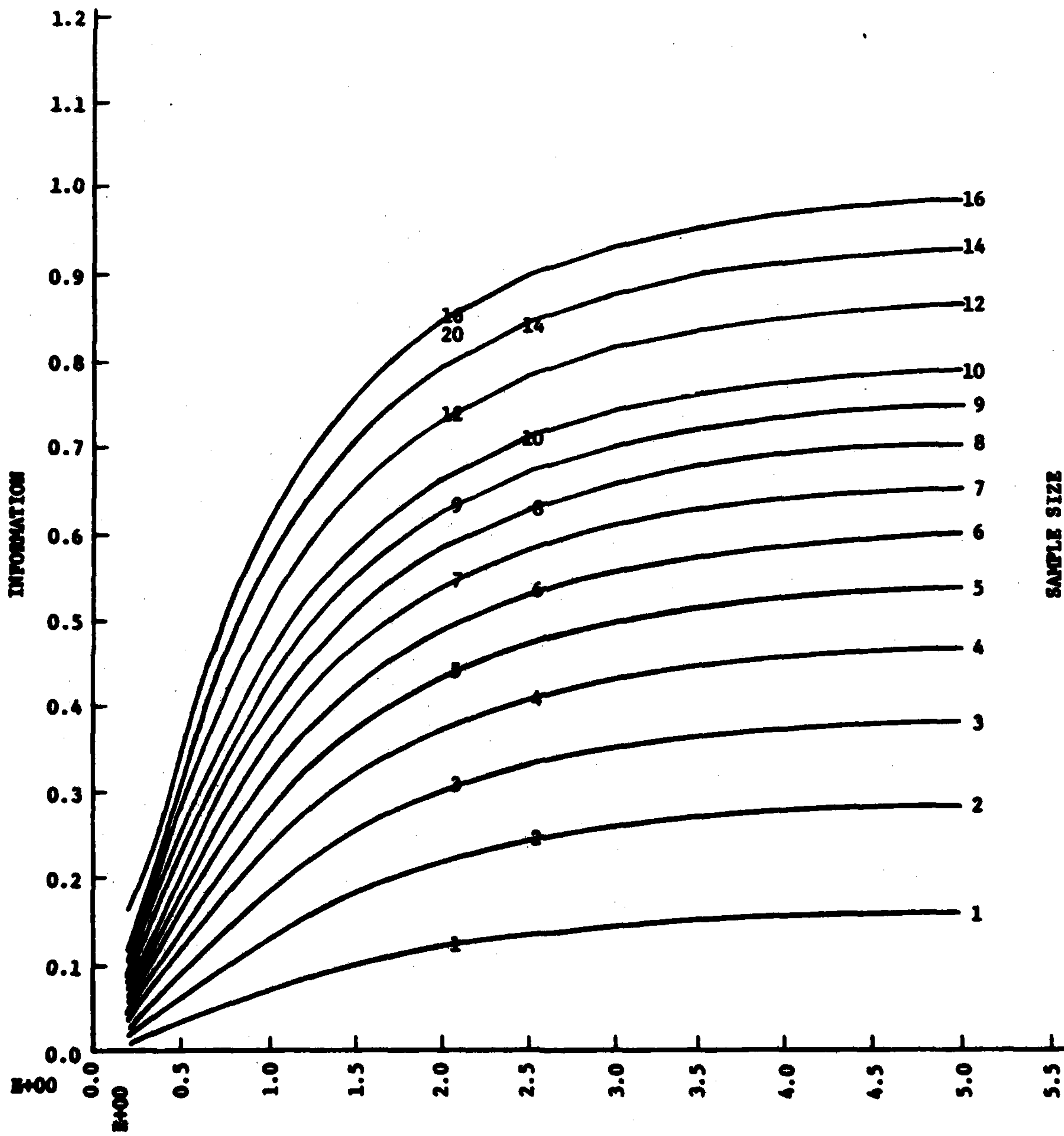
In order to obtain our main results we define the experiment $\mathcal{E}$ as a quadruple $\{\mathcal{D},\mathcal{B},\Lambda,\mathcal{P}\}$, where $\mathcal{D}$ is the space of observations x of the random vector X, $\mathcal{B}$ is the σ-field of the subsets of $\mathcal{D}$, the probability measure (or density of X belongs to a family $\mathcal{P}$ indexed by a parameter $\lambda \in \Lambda$. Suppose that the observation x in our experiment $\mathcal{E}$ consists of a pair of observations x_1,x_2, that is, $\mathcal{D} = \mathcal{D}_1 \times \mathcal{D}_2$. Let $\mathcal{B}_i$ be the σ-field over $\mathcal{D}_i$ induced from $\mathcal{B}$ by the transformation $x_i = x_i(X)$ and let $\mathcal{P}_i$ be the set of probability measures on $\mathcal{B}_i$ $(i = 1,2)$. Then $\mathcal{E}_i = \{\mathcal{D}_i,\mathcal{B}_i,\Lambda,\mathcal{P}_i\}$ $(i = 1,2)$ are two experiments. Denote the sum of the experiments $\mathcal{E}_1$ and $\mathcal{E}_2$, by $\mathcal{E} = (\mathcal{E}_1,\mathcal{E}_2)$. Now we consider a related experiment $\mathcal{E}_2(x_1) = \{D_2,B_2,\Lambda,P_2(x_1)\}$, where $\mathcal{P}_2(x_1)$ is the set of probability measures



A = 2.79 B = 4.78

INFORMATION VS SAMPLE SIZE

FIGURE 2.1



TEST TIME

A = 2.79 B = 4.78

INFORMATION VS TEST TIME

FIGURE 2.2

of x_2 conditional on x_1. Now consider the expected information for $\mathcal{E}_2$ were we to know the observation x_1 from performing $\mathcal{E}_1$:

$$I(\mathcal{E}_2(x_1),\pi(\lambda \mid x_1)) = E_{x_2}\int_\Lambda [\log \pi(\lambda \mid x_2,x_1)]\pi(\lambda \mid x_2,x_1)d\lambda$$
$$- \int_\Lambda [\log \pi(\lambda \mid x_1)]\pi(\lambda \mid x_1)d\lambda \; . \tag{2.7}$$

Since $\pi(\lambda \mid x_1)$ is the posterior density of λ after x_1 has been observed, $I(\mathcal{E}_2(x_1),\pi(\lambda \mid x_1))$ is the measure of expected information gain to be provided by our observation x_2 after $\mathcal{E}_1$ has been performed and x_1 observed. $I(\mathcal{E}_2 \mid \mathcal{E}_1) = E_{x_1}[I(\mathcal{E}_2(x_1),\pi(\lambda \mid x_1))]$, the average of $I(\mathcal{E}_2(x_1),\pi(\lambda \mid x_1))$ over x_1, is defined to be the average information to be provided by $\mathcal{E}_2$ after $\mathcal{E}_1$ has been performed. From now on we shall often denote the expected information by $I(\mathcal{E})$ when the particular prior distribution does not have to be stressed. This measure of information has the following properties:

1. $I(\mathcal{E}) \geq 0$.

2. $I(\mathcal{E}_2 \mid \mathcal{E}_1) \geq 0$.

3. $I(\mathcal{E}_1) + I(\mathcal{E}_2 \mid \mathcal{E}_1) = I(\mathcal{E})$ where $\mathcal{E} = (\mathcal{E}_1,\mathcal{E}_2)$.

4. If x_1 is sufficient for λ, then $I(\mathcal{E}_1) = I(\mathcal{E})$.

5. If $p(x_1,x_2 \mid \lambda) = p(x_1 \mid \lambda)p(x_2 \mid \lambda)$, i.e., x_1 and x_2 are independent when λ is given, then $I(\mathcal{E}_2 \mid \mathcal{E}_1) \le I(\mathcal{E}_2)$.

6. Let $\mathcal{E}_{(1)} = \mathcal{E}_1$ be any experiment and let $\mathcal{E}_2, \mathcal{E}_3 \ldots$ be independent identical experiments. Let $\mathcal{E}_{(2)} = (\mathcal{E}_1, \mathcal{E}_2)$ and generally $\mathcal{E}_{(n)} = (\mathcal{E}_n, \mathcal{E}_{(n-1)})$. Then $I(\mathcal{E}_{(n)})$ is a concave increasing function of n .

See Lindley (1956) for proofs of the above properties.

Let $\mathcal{E}_{n,t_1,t_2}$ be the experiment wherein n units aged t_1 and with identical life distributions are put on life test to age t_2 $(t_2 > t_1)$. Assume statistical independence among the n units conditional on λ .

<u>Theorem 2.2</u>

For the time-transformed exponential model, $\bar{F}(x \mid \lambda) = e^{-\lambda R_o(x)}$, $I(\mathcal{E}_{n,0,t})$ is concave increasing in n and also concave increasing in $R_o(t)$. The prior density $\pi(\lambda)$ is arbitrary.

<u>Proof</u>:

Since $R_o(\cdot)$ is known and continuous, $Y = R_o(X)$ is exponentially distributed with parameter λ . Therefore, performing an experiment for a period $[0,t]$ under the time-transformed exponential model is the same as performing an experiment for a time period $[0,R_o(t)]$ under the exponential model. Let $\pi(\lambda)$ be the prior in both cases. Hence, the measures of expected information gain provided by these two experiments are the same. It is therefore sufficient to prove this theorem for the exponential model.

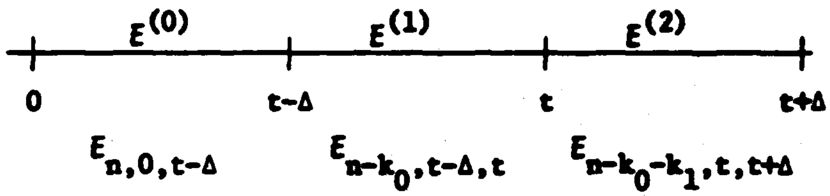
Now define

$\mathcal{E}^{(0)} = \mathcal{E}_{n,0,t-\Delta}$,

$\mathcal{E}^{(1)} = \mathcal{E}_{n-k_0,t-\Delta,t}$ given k_0 failures in $[0,t-\Delta]$,

$\mathcal{E}^{(2)} = \mathcal{E}_{n-k_0-k_1,t,t+\Delta}$ given k_0 failures in $[0,t-\Delta]$

and k_1 failures in $[t-\Delta,t]$.



Then

$$I(\mathcal{E}_{n,0,t},\pi(\lambda)) - I(\mathcal{E}_{n,0,t-\Delta},\pi(\lambda))$$

$$= I(\mathcal{E}^{(1)} \mid \mathcal{E}^{(0)}) = E_{D_0} I(\mathcal{E}_{n-k_0,t-\Delta,t},\pi(\lambda \mid D_0))$$

using the memoryless property of the exponential and $D_0 = (k_0$, total time on test in $(0,t-\Delta))$, the sufficient statistic for λ . Similarly

$$I(\mathcal{E}^{(2)} \mid \mathcal{E}^{(1)},\mathcal{E}^{(0)}) = E_{D_1} E_{D_0} I(\mathcal{E}_{n-k_0-k_1,t,t+\Delta},\pi(\lambda \mid D_0,D_1)) ,$$

where $D_1 = (k_1$, total time on test in $(t-\Delta,t))$. To show concavity and the increasing property we need only show

$$0 < I(\mathcal{E}^{(2)} \mid \mathcal{E}^{(1)},\mathcal{E}^{(0)}) \le I(\mathcal{E}^{(1)} \mid \mathcal{E}^{(0)}) .$$

By definition and property (5), we have

$$E_{D_1}(I(E_{n-k_0,t,t+\Delta},\pi(\lambda \mid D_0,D_1)))$$

$$= I(E_{n-k_0,t,t+\Delta},\pi(\lambda \mid D_0) \mid E_{n-k_0,t-\Delta,t},\pi(\lambda \mid D_0))$$

$$\leq I(E_{n-k_0,t,t+\Delta},\pi(\lambda \mid D_0)) .$$

By using the memoryless property of the exponential and using the prior $\pi(\lambda \mid D_0)$ on the parameter space, we have $I(E_{n-k_0,t,t+\Delta},\pi(\lambda \mid D_0)) = I(E_{n-k_0,t-\Delta,t},\pi(\lambda \mid D_0))$. Therefore,

$$E_{D_1}(I(E_{n-k_0,t,t+\Delta},\pi(\lambda \mid D_0,D_1))) \leq I(E_{n-k_0,t-\Delta,t},\pi(\lambda \mid D_0)) .$$

But information is increasing in sample size, so that if k_1 is the (random) number of failures in $E^{(1)}$ then

$$E_{D_1}(I(E_{n-k_0-k_1,t,t+\Delta},\pi(\lambda \mid D_0,D_1))) \leq I(E_{n-k_0,t-\Delta,t},\pi(\lambda \mid D_0)) .$$

Now take the expectation with respect to D_0 . Then

$$0 \leq I(E^{(2)} \mid E^{(1)},E^{(0)}) \leq I(E^{(1)} \mid E^{(0)}) .$$

Hence, $I(E_{n,0,t})$ is concave increasing in t which completes the proof.||

2.2 A Computerized Method for Calculating Entropy in the Case of a Weibull Distribution

Let $R_o(x) = x^\alpha$, $\alpha > 0$ in the time-transformed exponential model. That is, the life distribution of the test unit is

$$P(\text{lifetime} > x) = \exp(-\lambda x^{\alpha}) \quad , \text{ where } x > 0, \alpha > 0, \lambda > 0 .$$

This is the Weibull distribution survival probability. Assume α is known. From (2.2), the likelihood function of λ is

$$L(\lambda \mid x_1, \ldots, x_k, t) = \binom{n}{k} \lambda^k \alpha^k \left[\prod_{i=1}^{k} x_i\right]^{\alpha-1} \exp\left\{-\lambda\left[\sum_{i=1}^{k} x_i^{\alpha} + (n-k)t^{\alpha}\right]\right\} .$$

The pair k and $s = \sum_{i=1}^{k} x_i^{\alpha} + (n-k)t^{\alpha}$ constitute a sufficient statistic for λ. Let K and S be the random quantities corresponding to the number of failures and total time on test, respectively. Bartholomew (1963) has obtained the joint density of K and S given λ as follows:

$$P(K=k, S=s \mid \lambda) = \binom{n}{k} \frac{\lambda^k}{(k-1)!} e^{-\lambda s} \sum_{i=0}^{k} \binom{k}{i} (-1)^i \left\{\max\,[0, s - t^{\alpha}(n-k+i)]\right\}^{k-1}$$

$$\equiv D_k(s) \lambda^k e^{-\lambda s}$$

where

$$D_k(s) = \binom{n}{k} \frac{1}{(k-1)!} \sum_{i=0}^{k} \binom{k}{i} (-1)^i \left\{\max\,[0, s - t^{\alpha}(n-k+i)]\right\}^{k-1} .$$

The probability of observing no failure in $[0,t]$ is

$$P[K = 0 \ , \ S = nt^{\alpha} \mid \lambda] = e^{-\lambda n t^{\alpha}} \equiv p(0, nt^{\alpha} \mid \lambda)$$

where '$\equiv$' means definition.

Assume (2.3) as the prior density of λ. Using Equation (10) in Lindley (1956), we have

$$I(E) = \sum_{k=0}^{n} \int_{s=0}^{nt^\alpha} \int_{\lambda=0}^{\infty} p(k,s,\lambda) \log \frac{p(k,s,\lambda)}{p(k,s)\pi(\lambda)} \, ds d\lambda$$

$$= \sum_{k=0}^{n} \int_{s=0}^{nt^\alpha} \int_{\lambda=0}^{\infty} p(k,s \mid \lambda)\pi(\lambda) \log p(k,s \mid \lambda) ds d\lambda$$

$$- \sum_{k=0}^{n} \int_{s=0}^{nt^\alpha} p(k,s) \log p(k,s) ds$$

$$= \int_{\lambda=0}^{\infty} p(0,nt^\alpha \mid \lambda)\pi(\lambda) \log p(0,nt^\alpha \mid \lambda) d\lambda$$

$$+ \sum_{k=1}^{n} \int_{s=0}^{nt^\alpha} \int_{\lambda=0}^{\infty} p(k,s \mid \lambda)\pi(\lambda) \log p(k,s \mid \lambda) ds d\lambda$$

$$- p(0,nt^\alpha) \log p(0,nt^\alpha) - \sum_{k=1}^{n} \int_{s=0}^{nt^\alpha} p(k,s) \log p(k,s) ds$$

$$\equiv A_0 - B_0 + \int_{s=0}^{nt^\alpha} [A_k(s) - B_k(s)] ds \ ,$$

where

$$A_0 = \int_{\lambda=0}^{\infty} p(0,nt^\alpha \mid \lambda)\pi(\lambda) \log p(0,nt^\alpha \mid \lambda) d\lambda$$

$$= \int_{\lambda=0}^{\infty} e^{-\lambda nt^\alpha} \frac{b^a \lambda^{a-1} e^{-b\lambda}}{\Gamma(a)} (-\lambda nt^\alpha) d\lambda$$

$$= - \frac{ab^a nt^\alpha}{(b + nt^\alpha)^{a+1}} \ ,$$

$$B_0 = p(0,nt^\alpha) \log p(0,nt^\alpha)$$

$$= \left[\int_{\lambda=0}^{\infty} e^{-\lambda nt^\alpha} \frac{b^a \lambda^{a-1} e^{-b\lambda}}{\Gamma(a)} d\lambda\right] \log \left[\int_{\lambda=0}^{\infty} e^{-\lambda nt^\alpha} \frac{b^a \lambda^{a-1} e^{-b\lambda}}{\Gamma(a)} d\lambda\right]$$

$$= \frac{b^a}{(b + nt^\alpha)^a} \log \left[\frac{b^a}{(b + nt^\alpha)^a}\right] .$$

Also

$$A_k(s) = \int_{\lambda=0}^{\infty} p(k,s \mid \lambda)\pi(\lambda) \log p(k,s \mid \lambda)d\lambda ,$$

and

$$B_k(s) = p(k,s) \log p(k,s) .$$

Now

$$A_k(s) = \int_{\lambda=0}^{\infty} D_k(s)\lambda^k e^{-\lambda s} \frac{b^a \lambda^{a-1} e^{-b\lambda}}{\Gamma(a)} [\log D_k(s) + k \log \lambda - \lambda s]d\lambda$$

$$= D_k(s) \log D_k(s) \int_0^{\infty} \frac{b^a \lambda^{a+k-1} e^{-(b+s)\lambda}}{\Gamma(a)} d\lambda$$

$$+ kD_k(s) \int_0^{\infty} \frac{b^a \lambda^{a+k-1} e^{-(b+s)\lambda}}{\Gamma(a)} \log \lambda d\lambda$$

$$- sD_k(s) \int_0^{\infty} \frac{b^a \lambda^{a+k} e^{-(b+s)\lambda}}{\Gamma(a)} d\lambda$$

$$= \frac{\Gamma(a + k)b^a D_k(s) \log D_k(s)}{\Gamma(a)(b + s)^{a+k}} + A_k'(s) - \frac{sD_k(s)b^a\Gamma(a + k + 1)}{\Gamma(a)(b + s)^{a+k+1}} ,$$

where

$$A_k'(s) = kD_k(s) \int_0^\infty \frac{b^a \lambda^{a+k-1} e^{-(b+s)\lambda}}{\Gamma(a)} \log \lambda d\lambda$$

$$= \frac{kD_k(s) b^a \Gamma(a+k)}{\Gamma(a)(b+s)^{a+k}} \left[\psi(a+k) - \log(b+s)\right] ,$$

where $\psi(x)$ is a digamma function, defined as the derivative of $\log \Gamma(x)$.

$$E_k(s) = p(k,s) \log p(k,s)$$

$$= \left[\frac{D_k(s) b^a \Gamma(a+k)}{\Gamma(a)(b+s)^{a+k}}\right] \log \left[\frac{D_k(s) b^a \Gamma(a+k)}{\Gamma(a)(b+s)^{a+k}}\right] .$$

Therefore,

$I(s) =$ [illegible]

[illegible]

[illegible] .

[illegible]

APPENDIX

```
      PROGRAM INFO(INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT)
C
C     THIS PROGRAM CALCULATES THE MEASURE OF INFORMATION (EXPECTED
C     ENTROPY) OF A LIFE TEST EXPERIMENT WHEREIN THE LIFE DISTRIBUTION
C     OF THE TESTING UNIT IS WEIBULL DISTRIBUTION AND PRIOR IS A GAMMA
C     DISTRIBUTION.  WE USED SUBROUTINE 'GAUSSQ' TO EVALUATE THE INTEGRAL.
C     'GAUSSQ' APPLIES GAUSSIAN QUADRATUE TECHNIQUES TO DO THE EVALUA-
C     TION OF THE INTEGRAL.
C     THIS PROGRAM IS GOOD FOR SAMPLE SIZE UP TO 50.
C
      DIMENSION BB(500),X(500),C(500),ENDPTS(2)
      COMMON FACTA(55),TALPHA,A,B,N,H
1     READ 100, N, A, ALPHA, T, B
      IF ( N.EQ.0) STOP
      IF (N. GT. 1) GO TO 3
C
C       INFORMATION CALCULATION FOR N=1
C
      TA=T**ALPHA
      V=B+TA
      VB=B/V
      VA=VB**A
      UA=-(A*TA*VA)/V
      UB=VA*ALOG(VA)
      A1=A+1
      UC=(1-VA)*(PSI(A1)-ALOG(A))
      UD=A*VA*(1-VB+ALOG(VB))
      XINFO=UA-UB+UC+UD
      PRINT 150,N,T,ALPHA,A,B
      PRINT 300,XINFO
      GO TO 1
C
C       INFORMATION CALCULATION FOR N GREATER THAN 1
C       GENERATE FACTORIAL FROM 0 TO N
C
3     FACTA(1)=1.
      FACTA(2)=1.
      NN=N+1
      DO 5 I=3,NN
      K=I-1
5     FACTA(I)=K*FACTA(K)
C
C       CALCULATE A0 AND B0
C
      TALPHA=T**ALPHA
      H=N*TALPHA
      Q=B+H
      QQ1=(B/Q)**A
      QQ2=(A*H)/Q
```

```
      AO=-(QQ1)*QQ2
      BO=(QQ1)*ALOG(QQ1)
      PRINT 200,N,T,ALPHA,A,B
      CO=AO-BO
      PRINT 250, CO
      KIND=1
      KPTS=0
      DO 25 I=2,3
      MM=90*I
C
C     WHEN N*TALPHA IS LARGE, IT IS BETTER TO CHARGE THE VALUE OF MM,
C       FOR EXAMPLE MM=90*I, MM=120*I,........., BUT THE VALUE OF MM
C       CANNOT GREATER THAN 500.
C
      CALL GAUSSQ(KIND,MM,BALPHA,BETA,KPTS,ENDPTS,BB,X,C)
C
C     'GAUSSQ' RETURNS THE NODES X(I) AND WEIGHTS C(I), THEN APPROXIMATES
C     THE INTEGRAL BY SUM OF C(I)*F(X(I)) (I FROM 1 TO N).
C
      D=0.0
      DO 20 J=1,MM
20    D=D+C(J)*F(X(J))
      PRINT 150, D
      XINFO=CO+D
      PRINT 300,XINFO
25    CONTINUE
100   FORMAT(I3,F6.3,F5.2,F4.1,F5.2)
150   FORMAT(//10X,9HINTEGRAL=,F21.14)
200   FORMAT(///10X,2HN=,I3,10X,2HT=,F4.1,10X,6HALPHA=,F5.2,10X,2HA=,F6
      .3,10X,2HB=,F5.2)
250   FORMAT(//10X,3HCO=,F21.14)
300   FORMAT(//10X,12HINFORMATION=,F21.14)
      GO TO 1
      END


      FUNCTION F(S)
C
C     TO USE 'GAUSSQ' AN INTEGRAL (FROM A TO B) OF F(X) MUST BE BROUGHT
C     TO THE STANDARD INTEGRAL FORM.  THIS IS DONE BY A SUITABLE CHANGE
C     OF VARIABLES, FOR EXAMPLE, INTEGRAL (FROM A TO B) OF F(X) EQUALS
C     TO (B-A)/2 TIMES THE INTEGRAL (FROM -1 TO 1) OF F(Z), WHERE
C     Z=(X+1)(B-A)/2+A.
C
      DIMENSION SUM(55),D(55),SUMM(55)
      COMMON FACTA(55),TALPHA,A,B,N,H
      S=((S+1)*H/2.
      TIND=TALPHA*(N-1)
      IF (S-TIND) 30,30,40
30    D(1)=0.
      GO TO 45
40    D(1)=N
```

```
45    DO 50 K=2,N
50    SUM(K)=0
      DO 70 K=2,N
      DO 60 J=1,K
      TEST=TALPHA*(N-K+J-1)
      IF (S.LE.TEST) GO TO 65
      TT1=(S-TEST)**(K-1)
      TT2=FACTA(K+1)/(FACTA(J)*FACTA(K-J+2))
      TT3=(-1)**(J-1)
60    SUM(K)=SUM(K)+TT1*TT2*TT3
65    TT4=FACTA(N+1)/(FACTA(K+1)*FACTA(N-K+1)*FACTA(K))
70    D(K)=TT4*SUM(K)
      DO 80 K=1,N
      TT5=(B/(S+B))**A
      TT6=GAMMA(A+K)/GAMMA(A)
      TT7=D(K)/((S+B)**K)
      TT8=TT7*TT5*TT6
      X=A+K
      Y=PSI(X)
      TT9=K*Y-(S*X)/(S+B)
      TT10=TT9-ALOG(TT5)-ALOG(TT6)
80    SUMM(K)=TT8*TT10
      XSUMM=0.
      DO 90 K=1,N
90    XSUMM=XSUMM+SUMM(K)
      F=(H*XSUMM)/2.
      RETURN
      END
```

REFERENCES

[1] Barlow, R. E., "Accelerated Life Tests and Informations," ORC 80-14, Operations Research Center, University of California, Berkeley, (1980).

[2] Barlow, R. E. and F. Proschan, STATISTICAL THEORY OF RELIABILITY AND LIFE TESTING, Holt, Rinehart and Winston, (1975).

[3] Barlow, R. E., R. H. Toland and T. Freeman, "Stress-Rupture Life of Kevlar/Epoxy Spherical Pressure Vessels," UCID-17755, Part 3, Lawrence Livermore Laboratory, Livermore, California, (1979).

[4] Bartholomew, D. J., "The Sampling Distribution of an Estimate Arising in Life Testing," Technometrics, Vol. 5, No. 3, pp. 361-374, (1963).

[5] Bernardo, J. M., "Expected Information as Expected Utility," Annals of Statistics, Vol. 7, No. 3, pp. 686-690, (1979).

[6] DeGroot, M. H., OPTIMAL STATISTICAL DECISIONS, McGraw-Hill, (1970).

[7] Lindley, D. V., "On a Measure of the Information Provided by an Experiment," Annals of Statistics, Vol. 27, pp. 986-1005, (1956).

[8] Raiffa, H. and R. Schlaifer, APPLIED STATISTICAL DECISION THEORY, The M.I.T. Press, Cambridge, Massachusetts, (1961).