

SOME THEOREMS RELEVANT TO LIFE TESTING FROM AN
EXPONENTIAL DISTRIBUTION¹

BY B. EPSTEIN AND M. SOBEL²

Wayne University

1. Introduction and Summary. A life test on N items is considered in which the common underlying distribution of the length of life of a single item is given by the density

$$(1) \quad p(x; \theta, A) = \begin{cases} \frac{1}{\theta} e^{-(x-A)/\theta}, & \text{for } x \geq A \\ 0, & \text{otherwise} \end{cases}$$

where $\theta > 0$ is unknown but is the same for all items and $A \geq 0$. Several lemmas are given concerning the first r out of n observations when the underlying p.d.f. is given by (1). These results are then used to estimate θ when the N items are divided into k sets S_j (each containing $n_j > 0$ items, $\sum_{j=1}^k n_j = N$) and each set S_j is observed only until the first r_j failures occur ($0 < r_j \leq n_j$). The constants r_j and n_j are fixed and preassigned. Three different cases are considered:

1. The n_j items in each set S_j have a common known A_j ($j = 1, 2, \dots, k$).
2. All N items have a common unknown A .
3. The n_j items in each set S_j have a common unknown A_j ($j = 1, 2, \dots, k$).

The results for these three cases are such that the results for any intermediate situation (i.e. some A_j values known, the others unknown) can be written down at will. The particular case $k = 1$ and $A = 0$ is treated in [2].

The constant A in (1) can be interpreted in two different ways:

(i) A is the minimum life, that is life is measured from the beginning of time, which is taken as zero.

(ii) A is the "time of birth", that is life is measured from time A . Under interpretation (ii) the parameter θ , which we are trying to estimate, represents the expected length of life.

2. Statement of results. Three lemmas are given concerning the smallest r -ordered observations out of n independent observations on the common distribution (1). Although they are called lemmas because of their relation to the problem at hand, they are of interest in themselves.

A uniformly minimum variance unbiased estimate θ_i^* of θ together with its distribution is given for each case $i = 1, 2, 3$. This estimate is the unique unbiased estimate based on a sufficient statistic. In each case i ($i = 1, 2, 3$) it is given by $\theta_i^* = C_i \hat{\theta}_i$, where C_i is a constant and $\hat{\theta}_i$ is the maximum likelihood

Received 5/6/53, revised 1/28/54.

¹ Work supported by the Office of Naval Research.

² Now at Cornell University.

(m.l.) estimate, given in (13), (14), and (17) below. If $R = \sum_{j=1}^k r_j$ is the total number of failures observed, then it is shown that $2R\hat{\theta}_i/\theta$ is distributed as χ^2 with $2R$, $2(R - 1)$, and $2(R - k)$ degrees of freedom in cases 1, 2, and 3, respectively.

In each case the estimate θ_i^* (or $\hat{\theta}_i$) depends on the k -tuples (r_1, r_2, \dots, r_k) and (n_1, n_2, \dots, n_k) and in case 1 on the known A values. But it is shown that the distribution of the estimate depends only on R , θ (and in case 3 also on k) and is otherwise independent of the k -tuple (r_1, r_2, \dots, r_k) . The distribution is independent also of the k -tuple (n_1, n_2, \dots, n_k) , of N , and in case 1 of the known A values. Clearly this means that there are many ways of dividing the N items into k sets and of taking a total of R observations, all of which give equivalent estimates of θ . This equivalence is not with respect to the time required to obtain the estimate, but with respect to any properties depending on the distribution of the estimate.

3. Derivation of results in Section 2. Let $X_1 \leq X_2 \leq \dots \leq X_r$ denote the r smallest ordered observations from a set of n independent observations on the common distribution (1). In life testing, X_i , the i th smallest failure, is also the i th observation taken so that a sample like the above is obtained by merely stopping the experiment immediately after the r th observation. The set of n random variables under discussion represents a typical set S_j described above with the subscript j dropped. The joint p.d.f. of X_1, X_2, \dots, X_r is

$$(2) \quad p(x_1, x_2, \dots, x_r; \theta, A) = \begin{cases} \frac{n!}{(n-r)! \theta^r} e^{\frac{1}{\theta} [\sum_{i=1}^r (x_i - A) + (n-r)(x_r - A)]} & \text{for } A \leq x_1 \leq x_2 \leq \dots \leq x_r < \infty \\ 0, & \text{otherwise.} \end{cases}$$

Unless explicitly stated otherwise, any set X_1, X_2, \dots, X_r of the first r of n observations considered below will have density (2).

We now state a series of preliminary lemmas and corollaries.

Most proofs are direct and hence omitted.

LEMMA 1. For $1 \leq s < r \leq n$, the conditional joint density of

$$(3) \quad Y_i = X_{i+1} - X_s, \quad i = s, s + 1, \dots, r - 1$$

given $X_s = x_s$ (as well as the unconditional joint density) is (2) with (n, r, A) replaced by $(n - s, r - s, 0)$ respectively.

LEMMA 2. For $1 \leq r \leq n$ and for any preassigned constant $c \geq A$ the conditional joint density of the set

$$(4) \quad X_i^* = X_i - c \quad (i = 1, 2, \dots, r)$$

given that $X_1 \geq c$, is (2) with A replaced by zero.

COROLLARY 1. If c is replaced by a random variable C , independent of the X_i , whose range is the interval $[A, \infty]$, then the conditional joint density of X_i^* given that $X_1 \geq C$ is the same as in Lemma 2.

LEMMA 3. For $1 \leq r \leq n$ the set of random variables

$$(5) \quad W_i = (n - i + 1)(X_i - X_{i-1}) \quad i = 1, 2, \dots, r$$

(where X_0 is defined as the constant A) are mutually independent with common p.d.f. (1) except that $A = 0$.

PROOF. Utilizing the fact that for $r = 1, 2, \dots, n$

$$(6) \quad \sum_{i=1}^r (X_i - A) + (n - r)(X_r - A) = \sum_{i=1}^r W_i$$

the result is immediate if the transformation (5) is carried out in (2).

COROLLARY 2. For $1 \leq r \leq n$ if

$$(7) \quad V = \sum_{i=1}^r (X_i - A) + (n - r)(X_r - A)$$

then $2V/\theta$ is distributed as $\chi^2(2r)$.

PROOF. By Lemma 3 and (6), V is a sum of r independent, identically distributed exponential variables W_i . Since $2W_i/\theta$ is a $\chi^2(2)$ for each i , the corollary follows.

COROLLARY 3. For $1 < r \leq n$, if

$$(8) \quad V' = \sum_{i=1}^r (X_i - X_1) + (n - r)(X_r - X_1),$$

then the conditional distribution of $2V'/\theta$ given $X_1 = x_1$ (as well as the unconditional distribution) is $\chi^2(2r - 2)$. The random variables V' and X_1 are independent.

PROOF. The "unconditional" result follows from the fact that

$$(9) \quad V' = \sum_{i=2}^r W_i.$$

By Lemma 3 each of W_2, W_3, \dots, W_r is independent of W_1 and hence of X_1 and the corollary follows.

COROLLARY 4. For $1 \leq r \leq n$ and any preassigned constant $c \geq A$, if

$$(10) \quad V^* = \sum_{i=1}^r (X_i - c) + (n - r)(X_r - c),$$

then the conditional distribution of $2V^*/\theta$ given $X_1 \geq c$ (as well as the unconditional distribution) is $\chi^2(2r)$.

PROOF. By Lemma 2 the conditional joint density of $X_i^* = X_i - c$ given $X_1 \geq c$ is the same as the joint density of $X_i - A$ ($i = 1, 2, \dots, r$). Hence the conditional distribution of V^* must be the same as the distribution of V , namely $\chi^2(2r)$. Since the result is independent of c it is also the unconditional distribution.

COROLLARY 5. If c is replaced by a random variable C , independent of the X_i , whose range is the interval $[A, \infty]$ then again the conditional distribution of $2V^*/\theta$ given $X_1 \geq C$ is $\chi^2(2r)$. The random variables V^* and C are independent.

THEOREM 1. *The distribution of $\hat{\theta}$, the m.l. estimate, depends only on R, θ (and in case 3 also on k). The random variable $2R\hat{\theta}/\theta$ is distributed as $\chi^2(2R)$, $\chi^2(2R - 2)$ and $\chi^2(2R - 2k)$ in cases 1, 2, and 3 respectively.*

PROOF. In case 1 the joint p.d.f. of the R observed x 's is

$$(11) \quad \begin{matrix} B\theta^{-R}e^{-\sum V_i/\theta} & \text{if } A_i \leq X_{i1} \leq \dots \leq X_{ir_i} < \infty, i = 1, \dots, k \\ 0 & \text{otherwise} \end{matrix}$$

where B is independent of θ and

$$(12) \quad V_j = \sum_{i=1}^{r_j} (X_{ji} - A_j) + (n_j - r_j)(X_{jr_j} - A_j), \quad j = 1, 2, \dots, k.$$

The m.l. estimate $\hat{\theta}_1$ of θ is easily shown to be

$$(13) \quad \hat{\theta}_1 = \sum_{j=1}^k V_j/R.$$

From Corollary 2 and the independence of the V_j , it follows that $2R\hat{\theta}_1/\theta = \sum_{j=1}^k 2V_j/\theta$ is distributed as $\chi^2(2R)$.

In case 2 it can be readily verified that

$$(14) \quad \hat{\theta}_2 = \sum_{j=1}^k V_j^*/R$$

where

$$(15) \quad V_j^* = \sum_{i=1}^{r_j} (X_{ji} - \hat{A}) + (n_j - r_j)(X_{jr_j} - \hat{A})$$

and \hat{A} is the smallest of the R observed X 's. Let S_{j_0} denote the set containing \hat{A} . By Corollary 3 the distribution of $2V_{j_0}^*/\theta$ is $\chi^2(2r_{j_0} - 2)$ where $\chi^2(0)$ is to be interpreted as the sure constant zero. For any other set S_j ($j \neq j_0$) it follows from Corollary 5 that the distribution of $2V_j^*/\theta$ is $\chi^2(2r_j)$ and is independent of \hat{A} . All the random variables V_j^* are independent and hence

$$(16) \quad 2R\hat{\theta}_2/\theta = \sum_{j=1}^k 2V_j^*/\theta$$

is distributed as $\chi^2(2R - 2)$. Since $V_{j_0}^*$ is also independent of \hat{A} by Corollary 3 it follows that $\hat{\theta}_2$ and \hat{A} are independent.

In case 3 one easily computes

$$(17) \quad \hat{\theta}_3 = \sum_{j=1}^k V'_j/R$$

where

$$(18) \quad V'_j = \sum_{i=1}^{r_j} (X_{ji} - X_{j1}) + (n_j - r_j)(X_{jr_j} - X_{j1}) \quad (j = 1, 2, \dots, k).$$

By Corollary 3 the distribution of $2V'_j/\theta$ is $\chi^2(2r_j - 2)$ for each j (where $\chi^2(0)$ is to be interpreted as the sure constant zero). Hence

$$(19) \quad 2R\hat{\theta}_3/\theta = \sum_{j=1}^k 2V'_j/\theta$$

is distributed as $\chi^2(2R - 2k)$. In this case one needs $R > k$ observations to obtain an estimate of θ or, since $r_j \geq 1$ for all j , one needs $r_j > 1$ for at least one j . This completes the proof of Theorem 1.

Define

$$(20) \quad T_1 = \sum_{j=1}^k \left[\sum_{i=1}^{r_j} X_{ji} + (n_j + r_j)X_{jr_j} \right] = R\hat{\theta}_1 + \sum_{j=1}^k n_j A_j.$$

$$(21) \quad T_2 = (T_{20}, T_{21}) \quad \text{where} \quad T_{20} = T_1 \quad \text{and} \quad T_{21} = \min_j X_{j1}.$$

$$(22) \quad T_3 = (T_{30}, T_{31}, \dots, T_{3k}) \quad \text{where} \quad T_{30} = T_1 \quad \text{and} \quad T_{3j} = X_{j1}$$

for $j = 1, 2, \dots, k$.

The unbiased estimates $\hat{\theta}_i^*$ for cases 1, 2, and 3 respectively are given by

$$(23) \quad \theta_1^* = \hat{\theta}_1, \theta_2^* = R\hat{\theta}_2/(R - 1) \quad \text{and} \quad \theta_3^* = R\hat{\theta}_3/(R - k).$$

It can be quickly verified that θ_i^* depends on the observations only through T_i ($i = 1, 2, 3$). Hence, to show that θ_i^* are uniformly minimum variance unbiased estimates it suffices [3] to show that T_i is complete and sufficient for estimating θ in each case i ($i = 1, 2, 3$). The proof for case 3 is similar to that for case 2 and is omitted. To prove completeness we will need the following uniqueness theorem for one-sided Laplace transforms (see [1] and [3]): "If

$$(24) \quad \int_0^\infty f(t)e^{-t\theta} dt = 0 \quad \text{for all } \theta > 0$$

then $f(t) = 0$ for almost all $t > 0$."

THEOREM 2. T_1 is sufficient and complete for estimating θ .

PROOF. The sufficiency follows from the fact that the joint density in case 1 can be written as

$$(25) \quad C\theta^{-R} \exp \left[-\left(T_1 - \sum_{j=1}^k n_j A_j \right) / \theta \right] \prod_{j=1}^k f_j(X_{j1}, X_{j2}, \dots, X_{jr_j}; A_j)$$

where C is constant, and for each j ($j = 1, 2, \dots, k$)

$$(26) \quad f_j(X_{j1}, \dots, X_{jr_j}; A_j) = \begin{cases} 1 & \text{if } A_j \leq X_{j1} \leq \dots \leq X_{jr_j} < \infty \\ 0 & \text{otherwise.} \end{cases}$$

If we let $A^* = \sum_{j=1}^k n_j A_j$ then (since $X_{ji} \geq A_j$ for each i and j) $T_1 \geq A^*$. Let $p_\theta(t)$ denote the density of T_1 . To prove completeness it has to be shown that if

$$(27) \quad \int_{A^*}^{\infty} f(t)p_{\theta}(t) dt = 0 \quad \text{for all } \theta > \theta$$

then $f(t) = 0$ for almost all $t > A^*$. Letting $t_1 = t - A^*$ and

$$f^*(t_1) = {}^*t_1^{R-1}f(t_1 + A^*)$$

and using the result of Theorem 1 that $2t_1/\theta$ is distributed as $\chi^2(2R)$, then (27) takes the form

$$(28) \quad \int_0^{\infty} f^*(t_1)e^{-t_1/\theta} dt_1 = 0 \quad \text{for all } \theta > \theta.$$

It follows from the uniqueness theorem for one-sided Laplace transforms that $f^*(t_1) = 0$ for almost all $t_1 > 0$. Hence $f(t) = f(t_1 + A^*) = 0$ for almost all $t > A^*$. This proves that T_1 is complete.

COROLLARY 6. $\theta_1^* = \hat{\theta}_1 = (T_1 - A^*)/R$ is the unique uniformity minimum variance unbiased estimate of θ .

PROOF. This is a direct consequence of Theorem 2 and the theorem on page 321 of [3].

THEOREM 3. $T_2 = (T_{20}, T_{21})$ is sufficient and complete for estimating the pair (θ, A) .

PROOF. The sufficiency follows from the fact that the joint density in case 2 can be written as

$$(29) \quad C\theta^{-R} e^{-(T_{20}-NA)/\theta} f(T_{21}, A) \prod_{j=1}^k f_j(X_{j1}, X_{j2}, \dots, X_{jr_j}; T_{21}).$$

Here C is constant,

$$(30) \quad f(T_{21}, A) = \begin{cases} 1 & \text{if } T_{21} \geq A \\ 0 & \text{otherwise,} \end{cases}$$

and the f_j are defined in (26).

To show that T_2 is complete it has to be shown that if

$$(31) \quad \int_A^{\infty} \int_{Nt_{21}}^{\infty} f(t_{20}, t_{21})p_{\theta, A}(t_{20}, t_{21}) dt_{20} dt_{21} = 0 \quad \text{for all } \theta > 0 \text{ and all } A \geq 0,$$

then $f(t_{20}, t_{21}) = 0$ almost everywhere in the region $t_{21} > 0, t_{20} > Nt_{21}$. Let $u = t_{20} - Nt_{21}$ and $t = t_{21}$. By Theorem 1 we have that $2u/\theta$ is distributed as $\chi^2(2R - 2)$ and is independent of t . Moreover, since $t = \min_{j,i} X_{ji}, 2N(t - A)/\theta$ is distributed as $\chi^2(2)$ by Lemma 3. Then (31), after some cancellation, takes the form

$$(32) \quad \int_A^{\infty} \int_0^{\infty} e^{-(u+Nt)/\theta} u^{R-2} f(u + Nt, t) du dt = 0, \quad \text{for all } \theta > 0 \text{ and all } A \geq 0.$$

It thus follows directly from a two-dimensional uniqueness theorem for Laplace transforms that

$$(33) \quad f(t_{20}, t_{21}) = f(u + Nt, t) = 0, \quad \text{for all } \theta > 0 \text{ and } A \geq 0,$$

almost everywhere in the region $t_{21} > 0, t_{20} > Nt_{21}$. Thus completeness of T_2 is established.

COROLLARY 7. $\theta_2^* = R\hat{\theta}_2/(R - 1) = (T_{20} - NT_{21})/(R - 1)$ is the unique uniformly minimum variance unbiased estimate of θ .

PROOF. Unbiasedness of θ_2^* is easy to verify. The assertion is a consequence of Theorem 3 and the theorem in [3] cited in Corollary 6.

4. Confidence intervals on θ and A in case 2. Since $2(T_{20} - NT_{21})/\theta$ is distributed as $\chi^2(2R - 2)$, it is clear that confidence intervals on θ which do not involve A can be found. The following result concerning A is a corollary of Theorem 3.

COROLLARY 8. A unique uniformly minimum variance unbiased estimate of A in Case 2 based on (T_{20}, T_{21}) is given by

$$(34) \quad A^* = T_{21} - \frac{T_{20} - NT_{21}}{N(R - 1)}.$$

PROOF. It is readily verified that A^* has expectation A . Hence from the completeness of the sufficient pair (T_{20}, T_{21}) it follows as before that A^* is the unique uniformly minimum variance unbiased estimate of A . The minimum variance is $\sigma_{A^*}^2 = R\theta^2/N^2(R - 1)$.

To get confidence limits which do not involve θ , let us introduce the random variable U , where

$$(35) \quad U = N(T_{21} - A)/(T_{20} - NT_{21}).$$

Since the numerator and denominator are independent by Theorem 1, it is readily shown that the p.d.f. of U is given by

$$(36) \quad f(u) = (R - 1)/(1 + u)^R, \quad 0 < u < \infty.$$

Since $f(u)$ is independent of θ , for confidence coefficient α we solve the equation

$$(37) \quad 1 - \alpha = \int_0^c f(u) du = 1 - \frac{1}{(1 + c)^{R-1}}$$

or

$$(38) \quad c = \alpha^{-1/(R-1)} - 1.$$

Thus confidence limits on A are

$$(39) \quad t_{21} - c(t_{20} - NT_{21})/N < A < t_{21}.$$

These limits do not involve θ and are shortest in length for a given confidence in the class of confidence intervals based on U . The latter property is established by first noting from (35) that all possible confidence intervals on A are obtained by equating the probability in some interval of U values, say (c_1, c_2) , to $1 - \alpha$. The confidence interval then takes the form

$$(40) \quad t_{21} - c_2(t_{20} - NT_{21})/N < A < t_{21} - c_1(t_{20} - NT_{21})/N.$$

To minimize its length it suffices to minimize $c_2 - c_1$ (i.e. to find the shortest interval of U values containing probability $1 - \alpha$). Since the density (36) is strictly decreasing it is evident that the minimum is obtained by taking $c_1 = 0$.

5. Related results. We now indicate some connections between the results in Lemmas 1, 2, and 3 and some recent work [4] on ordered observations on a uniformly distributed random variable. It is easy to show that if Y is uniformly distributed on the interval $[0, B]$ then

$$(41) \quad X = A - \theta \log (Y/B)$$

has the exponential distribution (1). It follows from the monotonicity of the log that an initial ordered set of r out of n exponential random variables corresponds to a terminal ordered set of r out of n uniform random variables. S. Malmquist [4] has pointed out that by virtue of the transformation (41), independence in Lemma 3 implies and is implied by a corresponding result for rectangularly distributed variables. By using the transformation (41) one could prove (this is not done in [4]) analogues of Lemmas 1 and 2 for the rectangular case. Specifically let Y_ν be the ν th largest among n independent observations on the uniformly distributed random variable Y , then

(i) the random variables

$$(42) \quad Z_\nu = Y_{\nu+1}/Y_\nu \quad \nu = s, s + 1, \dots, r - 1; 1 \leq s < r \leq n$$

are jointly distributed like the $r - s$ largest (ordered) observations out of a set of $n - s$ independent uniform random variables on the unit interval $[0, 1]$.

(ii) for any preassigned constant $c \leq B$ the conditional random variables

$$(43) \quad Z_\nu^* = Y_\nu/c \quad \nu = 1, 2, \dots, r; 1 \leq r \leq n$$

given $Y_1 \leq c$ are jointly distributed like the r largest (ordered) observations out of a set of n independent uniform random variables on the unit interval $[0, 1]$.

Alternatively, if these results are shown independently they furnish another proof of the lemmas.

6. Conclusion and an application. In this paper we have given a number of results which are useful in making estimates of θ based on life test information from one or more sets of data, where the underlying probability law is the two-parameter exponential distribution (1). If (1) is the underlying p.d.f., then

$$(44) \quad \log \frac{1}{1 - P(x; \theta, A_j)} = \frac{x - A_j}{\theta}$$

where $P(x; \theta, A_j) = \Pr \{X \leq x; \theta, A_j\}$. Thus it is clear that cases 1, 2, and 3 are equivalent to assuming that the theoretical life distributions in the various sets S_j will plot either as parallel straight lines or as the same straight line on the semi-logarithmic scale suggested by (44). The results of this paper serve to give a procedure for estimating the slope (common slope) of the line (lines). A_j can be interpreted as the sensitivity limit at the appropriate stress level.

REFERENCES

- [1] G. DOETSCH, "Theorie und Anwendung der Laplace-Transformation," Dover, 1943.
- [2] B. EPSTEIN AND M. SOBEL, "Life testing I," *J. Amer. Stat. Assn.*, Vol. 48 (1953), pp. 486-502.
- [3] E. L. LEHMANN AND H. SCHEFFÉ, "Completeness, similar regions, and unbiased estimation I," *Sankhyā*, Vol. 10 (1950), pp. 305-340.
- [4] S. MALMQUIST, "On a property of order statistics from a rectangular distribution," *Skand. Aktuarietids.*, Vol. 33 (1950), pp. 214-222.