Inferences Based on Censored Sampling
From the Weibull or Extreme-Value Distribution

LEE J. BAIN*
Department of Mathematics
University of Missouri-Rolla
Rolla, Missouri

A simple, unbiased estimator, based on a censored sample, is proposed for the scale parameter of the extreme-value distribution. The exact distribution of the estimator is determined for the cases in which only the first two or only the first three ordered observations are available. The asymptotic distribution is derived, and an approximate distribution for small sample size is also provided. Interval estimation for the scale parameter is developed and a conservative interval estimate for reliability is also obtained.

KEY WORDS
Weibull Distribution
Extreme-value Distribution
Censored Sampling
Estimation
Reliability

1. INTRODUCTION

The Weibull distribution denoted by

$$F_x(x) = 1 - \exp \left[ -\left(\frac{x}{\alpha}\right)^\beta \right], \quad 0 < x < \infty; \quad \alpha, \beta > 0,$$

is considered in this paper. The variable $$Y = \ln X$$ follows the extreme-value distribution

$$F_Y(y) = 1 - \exp \left[ -\exp \left[ \left(\frac{y - u}{b}\right) \right] \right], \quad -\infty < y < \infty,$$

where $$b = 1/\beta$$ and $$u = \ln \alpha$$. The variables $$(X/\alpha)^\beta$$ and $$(Y - u)/b$$ are the corresponding reduced variates whose distributions are given by letting $$\alpha = \beta = 1$$ and $$b = 1, u = 0$$, respectively. Equivalent procedures can be developed under either model, but the extreme-value distribution has the advantage that its parameters appear as location and scale parameters.

Point and interval estimation procedures are in general quite complicated for these models, especially under censored sampling. Maximum likelihood esti-

* Research was sponsored by the Aerospace Research Laboratories, Air Force Systems Command, United States Air Force, Contract Number F 33165-71-C-1040.
Received Nov. 1970; revised May 1971.
mators of the parameters can be determined with the aid of a computer [4, 5]; and, inferential procedures based on the maximum likelihood estimators have been rather extensively developed for complete samples [16, 17]. Theoretically these methods can be extended to censored samples; however, the amount of computer time needed in order to determine the appropriate distributions becomes excessive. Thus other estimation and hypothesis testing techniques need to be developed for the censored sampling case.

For point estimation one common approach has been to apply the generalized least squares method or some related method to obtain linear estimators of the location and scale parameters $\mu$ and $\sigma$, see for example [9, 10, 11]. This approach for the most part requires knowledge of the variances and covariances of the ordered observations of the extreme-value distribution, which are available up to $n = 25$ [13]. It does not appear convenient to determine tests or confidence intervals based on these point estimators for sample sizes larger than 25. Some simple alternate point and interval estimation procedures have been presented in [11, 14] for censored sampling. Also, Johns and Lieberman [8] have a notable paper concerned with determining lower bounds for reliability based on censored samples. An attempt is made in this paper to develop procedures which are simple, reasonably good and widely applicable without the necessity of generating an undue number of tables.

2. Inferences Concerning $\sigma$

2.1 Unbiased estimator of $\sigma$

Suppose $x_1, \cdots, x_n$ denote the $r$ smallest ordered observations in a sample of size $n$ from the Weibull distribution. Also $y_1, \cdots, y_r$, where $y_i = \ln x_i$, will represent the $r$ smallest observations in a sample of size $n$ from the extreme-value distribution. The corresponding reduced observations will be denoted by $z_i = (x_i/\sigma)^{\kappa}$ and $w_i = (y_i - \mu)/\sigma$.

On examining a table of coefficients for determining best linear unbiased estimators (BLUE's) of $\sigma$, one sees that the statistic

$$\hat{\sigma} = -\sum_{i=1}^{r-1} (Y_i - Y_{i+1})/nk_{r,n} = T/k_{r,n},$$

is an appropriate unbiased estimator of $\sigma$, where

$$k_{r,n} = -(1/n)E\sum_{i=1}^{r-1} (W_i - W_{i+1}).$$

The statistic $\hat{\sigma}$ is the BLUE for $r = 2$, and it is somewhat similar to the BLUE for larger $r$. The exact moments of the $W_i$ are given in [18] up to $n = 100$, and $k_{r,n}$ can be calculated easily from these values for any prescribed combination of $r$ and $n$. For illustration purposes, values of $k_{r,n}$ are presented in Table I for $n = 5, 10, 15, 20, 30, 60, 100$ and integer values of $r = np$ for $p = .1(.1)1.0$. Asymptotic results can be utilized for larger values of $n$. If $r/n \rightarrow p$ as $n \rightarrow \infty$, expressions for the asymptotic values, say $k_p$, of the constants have been derived, and numerical values of $k_p$ are also presented in Table I. Values of the asymptotic
CENSORED SAMPLING FROM THE WEIBULL OF EXTREME-VALUE DISTRIBUTION

Table I

<table>
<thead>
<tr>
<th>n</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>5.2231</td>
<td>4.813</td>
<td>4.018</td>
<td>1.267</td>
<td>1.054</td>
<td>.2172</td>
<td>.2669</td>
<td>.4567</td>
<td>.6099</td>
<td>.7735</td>
</tr>
<tr>
<td>n</td>
<td>1.406</td>
<td>.3752</td>
<td>.6536</td>
<td>1.019</td>
<td>1.710</td>
<td>.1054</td>
<td>.2172</td>
<td>.3369</td>
<td>.4667</td>
<td>.6099</td>
</tr>
<tr>
<td>n</td>
<td>.0513</td>
<td>.1583</td>
<td>.2721</td>
<td>.3944</td>
<td>.5277</td>
<td>.6576</td>
<td>.8448</td>
<td>1.048</td>
<td>1.316</td>
<td>1.800</td>
</tr>
<tr>
<td>n</td>
<td>.0684</td>
<td>.1759</td>
<td>.2904</td>
<td>.4137</td>
<td>.5482</td>
<td>.6979</td>
<td>.8697</td>
<td>1.077</td>
<td>1.357</td>
<td>1.916</td>
</tr>
<tr>
<td>n</td>
<td>.0924</td>
<td>.2007</td>
<td>.3162</td>
<td>.4407</td>
<td>.5770</td>
<td>.7292</td>
<td>.9050</td>
<td>1.120</td>
<td>1.417</td>
<td>2.196</td>
</tr>
<tr>
<td>n</td>
<td>.1027</td>
<td>.2113</td>
<td>.3272</td>
<td>.4523</td>
<td>.5894</td>
<td>.7427</td>
<td>.9203</td>
<td>1.138</td>
<td>1.444</td>
<td>2.224</td>
</tr>
<tr>
<td>n</td>
<td>.1027</td>
<td>.2117</td>
<td>.3288</td>
<td>.4560</td>
<td>.6001</td>
<td>.7667</td>
<td>.9725</td>
<td>1.258</td>
<td>1.782</td>
<td>2.256</td>
</tr>
<tr>
<td>n</td>
<td>9.746</td>
<td>4.742</td>
<td>3.070</td>
<td>2.232</td>
<td>1.728</td>
<td>1.390</td>
<td>1.148</td>
<td>.971</td>
<td>.655</td>
<td>2.224</td>
</tr>
<tr>
<td>n</td>
<td>.9999</td>
<td>.9994</td>
<td>.9984</td>
<td>.9966</td>
<td>.9934</td>
<td>.9878</td>
<td>.9774</td>
<td>.9559</td>
<td>.8972</td>
<td>.8141</td>
</tr>
</tbody>
</table>

efficiency of \( \hat{b} \) are also given in Table I, and these indicate that \( \hat{b} \) is a simple, relatively efficient point estimator for \( b \) under censored sampling. A comparison of the variance of \( \hat{b} \) with the variance of the best linear unbiased estimator of \( b \) is also given in Table II for some small sample sizes and various censoring fractions. It may be worth noting that if mean squared error rather than variance is used as a goodness criterion, then a value of \( c \) can be found such that the MSE (c\( \hat{b} \)) is minimized. This value of \( c \) is given by

\[
c = \left[ 1 + \text{var} \left( \frac{\hat{b}}{b} \right) \right]^{-1} \approx \frac{n_k_{r,n}}{(1 + nk_{r,n})}.
\]

A comparison of the mean squared error of c\( \hat{b} \) with the mean squared error of the best linear invariant estimator (BLIE) [9] of \( b \) is also given in Table II.

2.2 Summary of distribution results

In considering similar tests concerning \( b \) relative to the nuisance parameter \( u \), one observes that attention can be restricted to functions of the \( r - 1 \) statistics \( Y_i - Y_r = \ln (X_i/X_r); i = 1, \ldots, r - 1 \). The statistic \( \hat{b} \) is suggested for use since it is a natural type of function of these statistics to consider, and it is also a good point estimator of \( b \) under censored sampling. Another possible indication of its desirability is that, although it is not a sufficient statistic, it does appear as a quantity in the joint density of the \( Y_i - Y_r \). Also, since \( \hat{b}/b \) is distributed independently of all parameters, tests based on \( \hat{b} \) are convenient to express, if the appropriate percentage points are available. The derivation of the distribution of \( S = \exp \left( -nk_{r,n} \hat{b}/b \right) = \prod_i \left( x_i/x_r \right)^{\hat{b}} \) is considered in section 5.1. For \( r = 2 \), the distribution of \( S \) is shown to be

\[
F_S(s) = \frac{ns}{(s + n - 1)}, \quad 0 < s < 1.
\]
### Table II

**Comparison of \( b \) with the BLUE and BLIE**

<table>
<thead>
<tr>
<th>( n )</th>
<th>( r )</th>
<th>( V(b/b) )</th>
<th>( V(\text{BLUE}/b) )</th>
<th>( C )</th>
<th>( nk_{r,n} / (1 + nk_{r,n}) )</th>
<th>MSE(cb/b)</th>
<th>MSE(BLIE/b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>5</td>
<td>.4175</td>
<td>.4168</td>
<td>.7055</td>
<td>.7064</td>
<td>.2945</td>
<td>.2942</td>
</tr>
<tr>
<td>4</td>
<td>.2553</td>
<td>.2538</td>
<td>.7966</td>
<td>.8004</td>
<td>.2033</td>
<td>.2024</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.1725</td>
<td>.1666</td>
<td>.8529</td>
<td>.8637</td>
<td>.1471</td>
<td>.1428</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>.4609</td>
<td>.4607</td>
<td>.6845</td>
<td>.6847</td>
<td>.3155</td>
<td>.3154</td>
</tr>
<tr>
<td>4</td>
<td>.2979</td>
<td>.2975</td>
<td>.7705</td>
<td>.7711</td>
<td>.2295</td>
<td>.2293</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.2161</td>
<td>.2155</td>
<td>.8223</td>
<td>.8235</td>
<td>.1777</td>
<td>.1773</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>.0795</td>
<td>.0716</td>
<td>.9264</td>
<td>.9400</td>
<td>.0736</td>
<td>.0668</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>.4009</td>
<td>.4009</td>
<td>.6753</td>
<td>.6753</td>
<td>.3247</td>
<td>.3247</td>
</tr>
<tr>
<td>4</td>
<td>.3162</td>
<td>.3161</td>
<td>.7598</td>
<td>.7599</td>
<td>.2402</td>
<td>.2402</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.2338</td>
<td>.2337</td>
<td>.8105</td>
<td>.8107</td>
<td>.1895</td>
<td>.1894</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>.0960</td>
<td>.0956</td>
<td>.9124</td>
<td>.9134</td>
<td>.0876</td>
<td>.0872</td>
<td></td>
</tr>
</tbody>
</table>

### Table III

**Exact \( Pr(2nT/b < \chi_{1-a}^2(2nk_{r,n})) \)**

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( r/n )</th>
<th>( 2 )</th>
<th>( 10 )</th>
<th>( 20 )</th>
<th>( 3 )</th>
<th>( 5 )</th>
<th>( 10 )</th>
<th>( 30 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>.01</td>
<td>.0098</td>
<td>.0099</td>
<td>.0100</td>
<td>.0097</td>
<td>.0099</td>
<td>.0102</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.05</td>
<td>.0501</td>
<td>.0501</td>
<td>.0501</td>
<td>.0502</td>
<td>.0500</td>
<td>.0501</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.10</td>
<td>.1012</td>
<td>.1007</td>
<td>.1003</td>
<td>.1017</td>
<td>.1003</td>
<td>.1003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.25</td>
<td>.2541</td>
<td>.2520</td>
<td>.2510</td>
<td>.2538</td>
<td>.2518</td>
<td>.2508</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.50</td>
<td>.5020</td>
<td>.5007</td>
<td>.5004</td>
<td>.5039</td>
<td>.5019</td>
<td>.5007</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.75</td>
<td>.7417</td>
<td>.7454</td>
<td>.7475</td>
<td>.7469</td>
<td>.7483</td>
<td>.7495</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.90</td>
<td>.8864</td>
<td>.8927</td>
<td>.8962</td>
<td>.8926</td>
<td>.8956</td>
<td>.8986</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.95</td>
<td>.9374</td>
<td>.9433</td>
<td>.9465</td>
<td>.9426</td>
<td>.9460</td>
<td>.9486</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.99</td>
<td>.9833</td>
<td>.9865</td>
<td>.9882</td>
<td>.9859</td>
<td>.9878</td>
<td>.9893</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
For $r = 3$,

$$F_3(s) = 1 - \frac{n!}{(n-3)!} \left\{ \frac{1}{2n} - \frac{1}{2(s+n-1)} - \frac{(n-2)(s-1)}{2q(s+n-1)} - \frac{s}{q!} \left[ \ln \left( \frac{n - \sqrt{q}(2s+n-2+\sqrt{q})}{(n-\sqrt{q}(2s+n-2-\sqrt{q})} \right) \right] \right\},$$

where $q = (n-2)^2 - 4s$.

The exact distribution becomes intractable for larger $r$; however, the chi-square distribution can be used to provide an extremely good approximation. It is shown in section 5.2 that $-2 \ln S = 2nk_{r,\alpha} \delta/b$ is distributed approximately as a chi-square variable with $2nk_{r,\alpha}$ degrees of freedom, for $n$ smaller than about 12 or $r/n$ about .5 or less.

The asymptotic distribution of $T/b = -\ln S$ is also derived in section 4. It is shown that the distribution of $\sqrt{n}(T/b - \mu_\beta)/\sigma_\beta$ approaches a standard normal distribution as $n \to \infty$ and $r/n \to p$, where

$$\mu_\beta = \sum_{i=1}^{\infty} (-\lambda_i) / (i+1)(i+1)!,$$

$$\sigma_\beta^2 = p^2/(1-p)^2 \lambda_1^2 - \mu_\beta^2 + 2\mu_\beta p/\lambda_1 + 2 \sum \lambda_i (-1)^{i+1}/i^2 i!,$$

and $\lambda_\beta = -\ln(1 - p)$. Numerical values of $\mu_\beta$ and $\sigma_\beta^2$ are tabulated in Table I for $p = .1(.1).9$.

### 2.3 Test of hypothesis concerning $b$ and power of the test

Consider, for example, the test of $H_0 : b \leq b_0$ against the alternative $H_A : b > b_0$ at the $\alpha$ significance level. Using the chi-square approximation stated in section 2.2, one rejects the hypothesis $H_0$ if

$$-2 \sum_{i=1}^{\infty} (Y_i - \bar{Y})/b_0 > \chi^2_{\alpha}(2nk_{r,\alpha}), \text{ where } \Pr[\chi^2(\alpha) > \chi^2_{\alpha}(\alpha)] = \alpha.$$

Linear interpolation can be used for non-integer degrees of freedom.

The power of the test for an alternative $b$ is

$$\Pr[\text{reject } H_0] = P[-2 \sum (Y_i - \bar{Y})/b_0 > \chi^2_{\alpha}(2nk_{r,\alpha})] = P[\chi^2(2nk_{r,\alpha}) > (b_0/b)\chi^2_{\alpha}(2nk_{r,\alpha})].$$

Note that a test on $b$ is analogous to a test on the variance of a normal distribution, since in that case $(n-1)s^2/\sigma^2$ is distributed as a chi-square variable with $n-1$ degrees of freedom. Thus material developed for the normal case can be applied to this case. For example, the sample size table and o.c. curves on pages 299–303 of [2] are applicable by simply replacing $\sigma^2$ by $b$ and $n$ by $2nk_{r,\alpha} + 1$.

### 3. Inferences on the Reliability

The problem of determining a test for $\xi = \alpha^{1/n}$, or equivalently the reliability, $R = \exp \left[ -\left( t/\alpha \right)^n \right]$, will now be considered. It is well known that


\[ 2r^2 = 2 \left[ \sum_{i=1}^{r-1} X_i^{1/n} + (n - r + 1)X_r^{1/n} \right] \]

is a complete, sufficient statistic for \( \alpha \) if \( b \) is known, and that \( 2r^2/\xi \) is distributed as a chi-square variable with \( 2r \) degrees of freedom. Since the distribution of \( \hat{b} \) is independent of \( \alpha \), it follows from Basu's Theorem \([1]\) that \( \hat{\xi} \) and \( \hat{b} \) are stochastically independent. Thus a joint confidence region can be determined for \( b \) and \( \hat{R} \), and a conservative limit for \( R \) can be obtained. A similar approach has been followed by Mann \([9]\) to obtain a test for \( R \) based on \( X_r/X_1 \) and \( 2r^2/\xi \).

In terms of \( R \),

\[ 2r^2/\xi = 2(-\ln R)[\sum (X_i/t)^{1/b} + (n - r + 1)(X_r/t)^{1/b}], \]

and

\[ P \left[ 2r^2/\xi < \chi^2_s(2r), \chi^2_s(2nk_{r,a}) < 2nk_{r,a}b/b < \chi^2_s(2nk_{r,a}) \right] = (1 - \alpha_1)(1 - \alpha_2 - \alpha_3). \]

This gives the joint confidence region \( \{R > R(b), \hat{b} < b < \hat{b} \} \), where

\[ R(b) = \exp \left\{ -\chi^2_s(2r)/2 \left[ \sum (X_i/t)^{1/b} + (n - r + 1)(X_r/t)^{1/b} \right] \right\}, \]

\[ \hat{b} = 2nk_{r,a}b/b, \]

\[ \hat{b} = 2nk_{r,a}b/b \]

A conservative \((1 - \alpha_1)(1 - \alpha_2 - \alpha_3)\) confidence limit for \( R \) is then given by

\[ R = \min_{R(b)} \hat{R}(b). \]

It is shown by Mann \([9]\) that \( \hat{R}(b) \) is a monotonically decreasing function of \( b \), if the time \( t \) is sufficiently small, and at least if \( t < (\prod_{i=1}^r x_i)^{1/r} \), in which case \( R = R(\hat{b}) \). This would ordinarily be the situation if \( p \) is not too small, since the expected content between any two order statistics is \( 1/(n + 1) \). It is also clear that if \( t > x_r \), then \( \hat{R}(b) \) is an increasing function of \( b \), and \( R = \hat{R}(b) \). If \( t < x_r \) but near \( x_r \), then \( \hat{R}(b) \) may not be monotonic; but it would have a single minimum, and \( \hat{R}(b) \) would approach 1 as \( b \to 0 \), and \( \hat{R}(b) \) would approach \( \exp \left\{ -\chi^2_s(2r)/2n \right\} \) as \( b \to \infty \). Thus, if \( t < x_r \) and \( \hat{R}(b) \) is decreasing at \( \hat{b} \), then \( \hat{R} = \hat{R}(\hat{b}) \). If \( \hat{R}(b) \) is increasing at \( \hat{b} \), then a search between \( \hat{b} \) and \( \hat{b} \) would be needed to determine the minimum.

4. Asymptotic Distribution of \( \hat{R} \)

Results given in \([3]\) will be applied to determine the asymptotic distribution of

\[ -T = \sum_{i=1}^r (Y_i - Y_r)/n, \]

where \( r/n \to p \) as \( n \to \infty \). Following the notation of \([3]\),

\[ -T/b = \sum_{i=1}^r \ln X_i - r \ln X_r = \sum_{i=1}^r c_i h(X_i), \]

where the \( X_i \) are ordered exponential variables. Also, \( c_i = 1, i = 1, \cdots, r - 1; c_r = -r, c_{r+1} = 0, i = r + 1, \cdots, n; F(x) = 1 - \exp (-x), h(x) = \ln x = \hat{H}(x), \]

Also

\[ E(X_i) = \sum_{j=0}^{i-1} 1/(n - j + 1). \]

\[ \alpha_{ij} = [1/(n - j + 1)] \sum_{i=1}^r c_i h'(E(X_i)) \]
\[
(n - j + 1)^{-1} \left[ \sum_{i=1}^{n} \left[ (E(X_i))^{-1} - r(E(X_i))^{-1} \right] \right].
\]

Then
\[
\mu_n = (1/n) \sum_{i=1}^{n} c_i \ln E(X_i) \]
\[
= (1/n) \sum_{i=1}^{n} \ln E(X_i) - (r/n) \ln E(X_i),
\]
\[
\sigma_n^2 = (1/n) \sum_{i=1}^{n} \alpha_i^2.
\]

and the distribution of \(\sqrt{n}((-T/b) - \mu_n)/\sigma_n\) approaches a standard normal distribution. Furthermore, \(\mu_n \rightarrow \mu_b\) and \(\sigma_n^2 \rightarrow \sigma_b^2\), where \(r/n \rightarrow p\) as \(n \rightarrow \infty\). Let
\[
J(x) = 1, \quad x \leq r/(n + 1)
\]
\[
= 0, \quad \text{otherwise},
\]
and let \(a_1 = p\), \(\lambda_p = F^{-1}(p) = -\ln(1 - p)\). Then
\[
\mu_p = \int_0^1 J(u)H(u) \, du + a_1\lambda_p
\]
\[
= \int_0^p \ln(-\ln(1 - u)) \, du - p \ln \lambda_p
\]
\[
= \sum_{i=1}^{\infty} (-\lambda_p)^i/(i + 1)(i + 1)!
\]

Also,
\[
\alpha(u) = (1 - u)^{-1} \left\{ \int_0^1 J(u)H'(w)(1 - w) \, dw + a_1(1 - p)H'(p) \right\}
\]

and
\[
\sigma_p^2 = \int_0^1 [\alpha(u)]^2 \, du.
\]

It can be shown that
\[
\sigma_p^2 = p^3/(1 - p)\lambda_p^2 - \mu_p^2 + 2\mu_p \lambda_p / \lambda_p + 2 \sum_{i=1}^{\infty} \lambda_p^2(1)^{i+1}/i!
\]

Then the distribution of \(\sqrt{n}((-T/b) - \mu_b)/\sigma_b\) approaches a standard normal distribution.

Also, \(k_{p,n} \rightarrow -\mu_b\) as \(n \rightarrow \infty\); and the asymptotic variance of \(\hat{b}\) is \(b^2 \sigma_p^2/\lambda_b^2\).

The asymptotic variance of the maximum likelihood estimator of \(b\) is provided in \([6]\) for \(p = .1(.1).9\); and, since this corresponds to the Cramér–Rao lower bound for an unbiased estimator of \(b\), the asymptotic relative efficiency of \(\hat{b}\) can be calculated. Some numerical values of \(k_{p,n} = -\mu_b\), \(\sigma_p^2\), \(n\) var \(\hat{b}/b^2\), and the relative efficiency of \(\hat{b}\) are presented in Table I. Note that, for complete samples,
where \( y \) denotes Euler's constant and numerical values of \( E(\ln X_c) \) are given in [15]. Also, from [15, page 711] \( k_{c..} \) is approximated by \( y + \ln \ln n + \gamma/\ln n \) for large \( n \).

5. Derivation of Exact and Approximate Distributions

5.1. Exact distribution for \( r = 2 \) and \( r = 3 \)

The joint density of \( (x_1, \ldots, x_r) \) is given by

\[
 f(x_1, \ldots, x_r) = \frac{1}{n!/(n - r)!} \prod_{i=1}^{r} (\beta/\alpha)(x_i/\alpha)^{\beta-1} \exp \left( x_i/\alpha \right) \exp \left\{ -\left( n - r \right)(x_i/\alpha)^\beta \right\},
\]

where \( 0 < x_1 < \cdots < x_r < \infty \).

On letting \( U_i = x_i/X_r, i = 1, \ldots, r - 1 \), the joint density of the \( U_i \) is found to be

\[
 f(U_1, \ldots, U_{r-1}) = \frac{n!/(n - r)!}{\Gamma(\beta)(\alpha)^{\beta-1}} \prod_{i=1}^{r-1} U_i^{\beta-1} \left( \sum_{i=1}^{r-1} U_i^{\beta} + n - r + 1 \right),
\]

which involves the two quantities \( \sum_{i=1}^{r-1} U_i^{\beta} \) and \( \prod_{i=1}^{r-1} U_i = \exp \left\{ -n(k_{c..}, \beta) \right\} = S^{1/\beta} \).

Since \( S^{1/\beta} \) is distributed independently of \( \beta, \beta \) may be set equal to one without loss of generality. For \( r = 2, S = U_1 \) and the distribution of \( S \) is given directly (see also [7], [11], [12]). For \( r = 3 \), the distribution of \( S \) may be determined by making a change of variable. The work is simplified by noting first that the joint density of \( U_i \)’s and also the function \( S \) are symmetric in the variables. Thus, for \( r = 3, \)

\[
 P[S \geq s] = P[U_1 U_2 > s ], \quad \text{where} \quad 0 < U_1 < U_2 < 1,
\]

\[
 = 1/2 \int_{U_1}^{s} \int_{U_1}^{U_2} f(u_1, u_2) \, du_1 \, du_2.
\]

Direct integration yields the result given earlier in the paper. The integration becomes quite tedious for larger values of \( r \), so that an approximation is needed.

5.2. Approximate distribution

The variable \( T = -\sum_{i=1}^{r-1} (Y_i - Y_r)/n \) takes on positive values, and the mean of \( nT/b \) is approximately equal to the variance of \( nT/b \), especially for
small \( p \). This is verified by Table I for large \( n \), since \( \text{var} (nT/b) \approx n \alpha^2 \), \( E(nT/b) \approx nk \), and \( k \approx \alpha^2 \) in Table I. Thus an approximate distribution with nearly the correct first two moments is obtained by assuming that \( 2nT/b \) is distributed approximately as a chi-square variable with \( 2nk \) degrees of freedom. The approximate percentage points were determined for \( r = 2, n = 5, 10, 20; r = 3, n = 5, 10, 30, \) and \( \gamma = 0.01, 0.05, 0.25, 0.5, 0.75, 0.90, 0.95 \) and \( 0.99; \) and the exact distributions were then used to determine the true probabilities for these approximate percentage points. As seen in Table III, the exact and approximate probabilities are in very close agreement.

This chi-square approximation is also consistent with the asymptotic results, at least to the extent that \( k \approx \alpha^2 \). This follows since \( ((2nT/b) - 2nk) / \sqrt{4nk} = (\chi^2(r) - v) / \sqrt{2r} \), which becomes normally distributed as \( r \) increases; but \( ((2nT/b) - 2nk) / \sqrt{4nk} \approx \sqrt{n}((T/b) - k)/\alpha \) which corresponds to the asymptotic result.

Thus the chi-square approximation seems appropriate if substantial censoring is involved. Further work is needed to determine the amount of error if \( r/n \) is near 1.

References