Tech. Bull. Fac. Agr. Kagawa Univ., Vol. 31, No. 2, 127～136, 1980

# BEST LINEAR INVARIANT ESTIMATORS FOR PARAMETERS OF THE EXTREME-VALUE DISTRIBUTION UNDER PROGRESSIVE CENSORING

Yoshikazu Kusanagi and Kiyoshi Fukuda

# マルチ・センサリングにおける極値分布母数の最良線形不変推定量について

草 薙 義 一, 福 田 清

This paper is concerned with the progressive censoring model which arises in hydrological situations. The problem of constructing the best linear invariant estimators for parameters of the largest extreme-value distribution under Type II progressive censoring is considered. Weights for obtaining best linear invariant estimates under this model are tabulated for all possible censorings for sample sizes 2 through 6. Use of weights is illustrated with an example.

マルチ・センサード・サンプルにおける極値分布母数の最良線形不定推定量を求める．最良線形不定推定値を求めるための重みを，サンプル数 2—6 のセンサリングについて表にする．また重みの使い方を例示する．

## 1. Introduction

In this paper we shall consider the linear estimation set up; that is, for a given sample, unknown parameters are estimated. Suppose that we are given a sample of size $n$ from a population with the largest extreme-value distribution

$$F(x)=\exp\left[-\exp\left(-\frac{x-u}{b}\right)\right],\ -\infty<x<\infty. \tag{1}$$

Suppose furthermore that observations are the quantity and they are rearranged in order from least to greatest. Incomplete data occur in hydrological situations where one or more observations are spurious and coming from a different source. In statistical applications it may be necessary and can be desirable to remove these outliers.

The linear estimation problem which we shall consider is the following: How should the incomplete data be handled in order to give rise to the best estimation. A number of models have been proposed in the literature (for example, see Montfort[25],) but none of these are entirely satisfactory. In dealing with the incomplete data we shall introduce the Type II progressive censoring model, which seems to overcome some of drawbacks of the models.

Many investigators, in particular Herd[10, 11] and Cohen[2-4], have considered the progressive censoring model. Mann[19] [22] [23] have investigated the linear invariant estimation of extreme-value parameters. A similar statement can be made for the case of the "largest" extreme-value distribution. The main point of this paper will be to construct the linear invariant estimators for parameters of the largest extreme-value distribution under Type II progressive censoring.

Before discussing the linear invariant estimation problem for Type II progressive censoring, we shall state some properties of best linear unbiased and best linear invariant estimators for extreme-value parameters based on complete and censored samples. Lieblein[13] and Lieblein and Zelen[14] have applied the Gauss-Markov theorem to the linear estimation of extreme-value parameters. Harter and Moore[8] [9] and Mann[20], using Monte Carlo methods, have compared the mean squared errors of the maximum likelihood and best linear invariant estimators, and they have shown that both estimators are good for parameter estimation from the point of view of the mean square error. The principle difficulty with adopting the method of maximum likelihood is that iterative methods must be employed to find the estimates. To avoid this difficulty, Mann[17] [18] [20] [21] suggested choosing the best linear invariant estimators for extreme-value parameters. Mann[20] has proved that the best linear invariant estimators are asymptotically efficient and asymptotically normal and hence are asymptotically equal to their respective Cramér-Rao lower bounds. Mann[15] [16], Mann, Schafer and Singpurwalla[24] have tabulated the weights for obtaining best linear invariant and best linear unbiased estimates based on complete and censored samples.

We use the following notation throughout. Let $X_1, \ldots, X_n$ be a random sample of size $n$ from a population with *cumulative distribution function* (*cdf*) $F(x)$. The $i$th smallest ordered progressively censored sample is denoted by $X_{i,n}$. Let us denote by $O_i(X^{(\cdot)})$ the value of the $i$th smallest coordinate in $X^{(\cdot)}=(X_1^{(\cdot)}, \ldots, X_m^{(\cdot)})$.

## 2. The Progressively Censored Sample

Suppose that $X_{1,n} \leq X_{2,n} \leq \cdots \leq X_{m,n}$ is the $m$ stage Type II progressively censored sample from a population with *c.d.f.* $F(x)$ and *p.d.f.* $f(x)$, and $R_i$ is the number of samples removed at the $i$th stage of censoring. The joint probability density function of $(x_{1,n}, \ldots, x_{m,n})$ is, as in Cohen[3],

$$\prod_{i=1}^{m} \{n^* f(x_{i,n})(1-F(x_{i,n})^{R_i}\} \tag{2}$$

for $-\infty < x_{1,n} \leq \cdots \leq x_{m,n} < \infty$, where $n^* = n - \sum_{k=1}^{i-1} R_k - i + 1$. As a special case, when $R_1 = \cdots = R_{m-1} = 0$, $R_m = n-m$, (2) reduces to the *p.d.f.* for the singly censored sample. By carrying out the necessary integrations, we can evaluate the *p.d.f.* $f(x_{i,n})$ and joint *p.d.f.* $f(x_{i,n} x_{j,n})$ as follows:

$$f(x_{i,n}) = \kappa_i \sum_{r=1}^{i} (-1)^{r+1} \zeta_1(r, 0, i) \zeta_2(r, i) f(x_{i,n})(1-F(x_{i,n}))^{\zeta_3(r)-1}, \tag{3}$$

$$f(x_{i,n}, x_{j,n}) = \kappa_j \sum_{r=1}^{i} \sum_{s}^{j-i} (-1)^{r+s} K(r, s) f(x_{i,n})(1-F(x_{i,n}))^{\zeta_4(r,s)-1} f(x_{j,n})(1-F(x_{j,n}))^{\zeta_5(s)-1}, \tag{4}$$

where

$$\kappa_i = n(n-R_1-1)\cdots(n-\sum_{k=1}^{i}(R_k+1)+R_i+1),$$

$$\zeta_1(r, l, i) = [\prod_{u=2}^{i-r+1-l} \sum_{k=i-r-u+2}^{i-r} (R_k+1)]^{-1}, \qquad r=1, 2, \ldots, i-l-1,\ i-l \geq 2,$$

$$\zeta_2(r, i) = [\prod_{u=2}^{r} \sum_{k=i-r+1}^{i-r+u-1} (R_k+1)]^{-1}, \qquad r=2, 3, \ldots, i,\ i \geq 2,$$

$$\zeta_1(i-l, l, i) = \zeta_2(1, i) = 1,$$

$$\zeta_3(r) = \sum_{k=i-r+1}^{m} (R_k+1), \qquad r=1, 2, \ldots, i,$$

$$\zeta_4(r, s) = \sum_{k=i-r+1}^{j-s} (R_k+1), \qquad r=1, 2, \ldots, i, \quad s=1, 2, \ldots, j-i,$$

$$\zeta_5(s) = \sum_{k=j-s+1}^{m} (R_k+1), \qquad s=1, 2, \ldots, j-i,$$

and

$$K(r, s) = \zeta_1(r, 0, i) \zeta_2(r, i) \zeta_1(s, i, j) \zeta_2(s, j).$$

## 3. Moment of Order Statistics

Without loss of generality, we shall consider the reduced extreme-value distribution, with $Y_{i,n} = (X_{i,n} - u)/b$. From (3) we obtain the $k$th moment

$$E(y_{i,n}^k)=\kappa_j \sum_{r=1}^{i} \sum_{s=0}^{\zeta_3(r)-1} (-1)^{r+s+1}\zeta_1(r, 0, i)\zeta_2(r, i)\binom{\zeta_3(r)-1}{s} g_k(s+1) \tag{5}$$

where

$$g_k(c)=\int_{-\infty}^{\infty} x^k e^{-x-ce^{-x}} dx, \quad c>0, \tag{6}$$

is the polygamma function. In particular, we have

$$g_1(c)=\frac{1}{c}(\gamma+\log c) \tag{7}$$

and

$$g_2(c)=\frac{1}{c}\left(\frac{\pi^2}{6}+(\gamma+\log c)\right)^2 \tag{8}$$

where $\gamma$ is Euler's constant .5772156649...... From (4) we similarly obtain the moments $E(y_{i,n}y_{j,n})$

$$E(y_{i,n}y_{j,n})=\kappa_j \sum_{r=1}^{i} \sum_{s=1}^{j-i} \sum_{t=0}^{\zeta_4(r,s)-1} \sum_{u=0}^{\zeta_5(s)-1} (-1)^{r+s+t+u}\binom{\zeta_4(r, s)-1}{t}\binom{\zeta_5(s)-1}{u}\phi(t+1, s+1) \tag{9}$$

where

$$\phi(t, u)=\int_{-\infty}^{\infty}\int_{\infty}^{y} xy e^{-x-te^{-x}} e^{-y-ue^{-y}} dx dy, \ t, u>0$$

$$=\frac{1}{2tu}\left\{(u-t)g_2(t+u)+t^2(g_1(t))^2+2L\left(1+\frac{t}{u}\right)-\left(\log\frac{t}{u}\right)^2-\frac{\pi^2}{6}\right\}, \ t<u \tag{10}$$

$$=\frac{1}{2tu}\left\{(u-t)g_2(t+u)+t^2(g_1(t))^2-2L\left(1+\frac{t}{u}\right)+\frac{\pi^2}{6}\right\}, \ t>u,$$

where

$$L(1+x)=\int_1^{1+x}\frac{\log w}{w-1}dx=\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n^2}x^n \tag{11}$$

is spence's integral. By denoting the variance of $y_{i,n}$ by $\sigma_{i,n}^2$ and the covariance between $y_{i,n}$ and $y_{j,n}$ by $\sigma_{i,j,n}$, we have

$$\sigma_{i,n}^2=E(y_{i,n}^2)-[E(y_{i,n})]^2 \tag{12}$$

and

$$\sigma_{i,j,n}=E(y_{i,n}y_{j,n})-E(y_{i,n})E(y_{j,n}) \tag{13}$$

## 4. Computation of Weights

Following Lieblein and Zelen[14], we now consider the problem of constructing a general linear estimation for the 100 $p$th percentile $x_p=u-b\, ln\, ln(1/p)$. Denoting the weights by $a_{i,m,n}$ and $c_{i,m,n}$, we have

$$\hat{X}_{p,m,n}=\hat{u}_{m,n}-\hat{b}_{m,n}\, ln\, ln(1/p)$$

$$=\sum_{i=1}^{m}[a_{i,m,n}-ln\, ln(1/p)c_{i,m,n}]X_{i,n}. \tag{14}$$

For $\hat{u}_{m,n}$ and $\hat{b}_{m,n}$ to be unbiased we require that $a_{i,m,n}$ and $c_{i,m,n}$ be chosen so as to satisfy

$$E(\hat{u}_{m,n})=E(\sum_{i=1}^{m} a'_{i,m,n}X_{i,n})=\sum_{i=1}^{m} a'_{i,m,n}E(u+bY_{i,n})=u, \tag{15}$$

$$E(\hat{b}_{m,n})=E(\sum_{i=1}^{m} c'_{i,m,n}X_{i,n})=\sum_{i=1}^{m} c'_{i,m,n}E(u+bY_{i,n})=b, \tag{16}$$

that is,

$$\sum_{i=1}^{m} a'_{i,m,n}=1, \quad \sum_{i=1}^{m} a'_{i,m,n}E(Y_{i,n})=0, \tag{17}$$

Table 1 Weights for obtaining best linear invariant estimates of parameters of the largest extreme-value distribution

$A(N, M, I)$: Weight for esrimating $\tilde{u}$

$C(N, M, I)$: Weight for estimating $\tilde{b}$

$b^2E(LU)$: Mean squared error

$b^2E(CP)$: $E(\tilde{u}-u)\ (\tilde{b}-b)$

$b^2E(LB)$: Mean squared error of $\tilde{b}$

| | N | M | R | I | A(N,M,I) | C(N,M,I) |
|---|---|---|---|---|---|---|
| | 2 | 2 | 0 | 1 | 0.889269 | -0.421383 |
| | | | 0 | 2 | 0.110731 | 0.421383 |
| E(LU) | | | | | 0.6571300 | |
| E(CP) | | | | | -0.0375742 | |
| E(LB) | | | | | | 0.4158392 |
| | 3 | 3 | 0 | 1 | 0.667936 | -0.468904 |
| | | | 0 | 2 | 0.251001 | 0.190239 |
| | | | 0 | 3 | 0.081063 | 0.278666 |
| E(LU) | | | | | 0.4024074 | |
| E(CP) | | | | | 0.0184217 | |
| E(LB) | | | | | | 0.2563462 |
| | 3 | 2 | 0 | 1 | 0.603352 | -0.690920 |
| | | | 1 | 2 | 0.396648 | 0.690920 |
| E(LU) | | | | | 0.4148771 | |
| E(CP) | | | | | 0.0612878 | |
| E(LB) | | | | | | 0.4037041 |
| | 3 | 2 | 1 | 1 | 0.785559 | -0.409868 |
| | | | 0 | 2 | 0.214441 | 0.409868 |
| E(LU) | | | | | 0.4463369 | |
| E(CP) | | | | | 0.0881192 | |
| E(LB) | | | | | | 0.3969849 |
| | 4 | 4 | 0 | 1 | 0.526813 | -0.455910 |
| | | | 0 | 2 | 0.261510 | 0.070109 |
| | | | 0 | 3 | 0.147340 | 0.182749 |
| | | | 0 | 4 | 0.064336 | 0.203052 |
| E(LU) | | | | | 0.2924765 | |
| E(CP) | | | | | 0.0283121 | |
| E(LB) | | | | | | 0.1838619 |
| | 4 | 3 | 0 | 1 | 0.483201 | -0.593557 |
| | | | 0 | 2 | 0.270739 | 0.099235 |
| | | | 1 | 3 | 0.246060 | 0.494322 |
| E(LU) | | | | | 0.2988436 | |
| E(CP) | | | | | 0.0484073 | |
| E(LB) | | | | | | 0.2472848 |
| | 4 | 2 | 0 | 1 | 0.340634 | -0.879965 |
| | | | 2 | 2 | 0.659366 | 0.879965 |
| E(LU) | | | | | 0.3371736 | |
| E(CP) | | | | | 0.1254102 | |
| E(LB) | | | | | | 0.4019794 |
| | 4 | 3 | 1 | 1 | 0.616043 | -0.457117 |
| | | | 0 | 2 | 0.258140 | 0.185473 |
| | | | 0 | 3 | 0.125817 | 0.271644 |
| E(LU) | | | | | 0.3244425 | |
| E(CP) | | | | | 0.0663331 | |
| E(LB) | | | | | | 0.2452919 |
| | 4 | 2 | 1 | 1 | 0.517929 | -0.668951 |
| | | | 1 | 2 | 0.482071 | 0.668951 |
| E(LU) | | | | | 0.3550144 | |
| E(CP) | | | | | 0.1323390 | |
| E(LB) | | | | | | 0.3878015 |
| | 4 | 3 | 0 | 1 | 0.471845 | -0.586201 |
| | | | 1 | 2 | 0.405890 | 0.312466 |
| | | | 0 | 3 | 0.122265 | 0.273735 |
| E(LU) | | | | | 0.3069883 | |
| E(CP) | | | | | 0.0578293 | |
| E(LB) | | | | | | 0.2506743 |
| | 4 | 2 | 2 | 1 | 0.727987 | -0.402018 |
| | | | 0 | 2 | 0.272013 | 0.402018 |
| E(LU) | | | | | 0.3905527 | |
| E(CP) | | | | | 0.1561616 | |
| E(LB) | | | | | | 0.3831827 |
| | 5 | 5 | 0 | 1 | 0.433593 | -0.431259 |
| | | | 0 | 2 | 0.246092 | 0.005600 |
| | | | 0 | 3 | 0.163808 | 0.111820 |
| | | | 0 | 4 | 0.103531 | 0.155707 |
| | | | 0 | 5 | 0.052975 | 0.158131 |
| E(LU) | | | | | 0.2304050 | |
| E(CP) | | | | | 0.0291352 | |
| E(LB) | | | | | | 0.1428429 |
| | 5 | 4 | 0 | 1 | 0.402049 | -0.525419 |
| | | | 0 | 2 | 0.247811 | 0.010733 |
| | | | 0 | 3 | 0.173070 | 0.139467 |
| | | | 1 | 4 | 0.177069 | 0.375219 |
| E(LU) | | | | | 0.2342760 | |
| E(CP) | | | | | 0.0406902 | |
| E(LB) | | | | | | 0.1773343 |
| | 5 | 3 | 0 | 1 | 0.324417 | -0.689924 |
| | | | 0 | 2 | 0.255634 | 0.027310 |
| | | | 2 | 3 | 0.419949 | 0.662615 |
| E(LU) | | | | | 0.2492961 | |
| E(CP) | | | | | 0.0725185 | |
| E(LB) | | | | | | 0.2447802 |
| | 5 | 2 | 0 | 1 | 0.112593 | -1.024151 |
| | | | 3 | 2 | 0.887407 | 1.024151 |
| E(LU) | | | | | 0.3127041 | |
| E(CP) | | | | | 0.1725667 | |
| E(LB) | | | | | | 0.4026408 |
| | 5 | 4 | 1 | 1 | 0.503355 | -0.447571 |
| | | | 0 | 2 | 0.248199 | 0.069238 |
| | | | 0 | 3 | 0.160251 | 0.180500 |
| | | | 0 | 4 | 0.088196 | 0.197833 |
| E(LU) | | | | | 0.2538155 | |
| E(CP) | | | | | 0.0527322 | |
| E(LB) | | | | | | 0.1768164 |
| | 5 | 3 | 1 | 1 | 0.444725 | -0.579085 |
| | | | 0 | 2 | 0.260686 | 0.097249 |
| | | | 1 | 3 | 0.294589 | 0.481836 |
| E(LU) | | | | | 0.2658274 | |
| E(CP) | | | | | 0.0796762 | |
| E(LB) | | | | | | 0.2372548 |

Table of Weights (continued)

| | N | M | R | I | A(N,M,I) | C(N,M,I) |
|---|---|---|---|---|---|---|
| | 5 | 2 | 1 | 1 | 0.276898 | -0.853587 |
| | | | 2 | 2 | 0.723102 | 0.853587 |
| E(LU) | | | | | 0.3223637 | |
| E(CP) | | | | | 0.1721483 | |
| E(LB) | | | | | | 0.3885043 |
| | 5 | 4 | 0 | 1 | 0.390321 | -0.517347 |
| | | | 1 | 2 | 0.363219 | 0.140168 |
| | | | 0 | 3 | 0.160530 | 0.176145 |
| | | | 0 | 4 | 0.085929 | 0.201034 |
| E(LU) | | | | | 0.2430970 | |
| E(CP) | | | | | 0.0494248 | |
| E(LB) | | | | | | 0.1802482 |
| | 5 | 3 | 0 | 1 | 0.321333 | -0.678748 |
| | | | 1 | 2 | 0.387470 | 0.196903 |
| | | | 1 | 3 | 0.291197 | 0.481845 |
| E(LU) | | | | | 0.2547245 | |
| E(CP) | | | | | 0.0766277 | |
| E(LB) | | | | | | 0.2438902 |
| | 5 | 4 | 0 | 1 | 0.394388 | -0.525590 |
| | | | 0 | 2 | 0.249757 | 0.013506 |
| | | | 1 | 3 | 0.270654 | 0.311980 |
| | | | 0 | 4 | 0.085201 | 0.200104 |
| E(LU) | | | | | 0.2375532 | |
| E(CP) | | | | | 0.0449391 | |
| E(LB) | | | | | | 0.1800071 |
| | 5 | 3 | 2 | 1 | 0.582581 | -0.448689 |
| | | | 0 | 2 | 0.262538 | 0.182535 |
| | | | 0 | 3 | 0.154881 | 0.266153 |
| E(LU) | | | | | 0.2927889 | |
| E(CP) | | | | | 0.0963995 | |
| E(LB) | | | | | | 0.2367323 |
| | 5 | 2 | 2 | 1 | 0.463774 | -0.652851 |
| | | | 1 | 2 | 0.536226 | 0.652851 |
| E(LU) | | | | | 0.3396732 | |
| E(CP) | | | | | 0.1769673 | |
| E(LB) | | | | | | 0.3751830 |
| | 5 | 3 | 1 | 1 | 0.436106 | -0.571951 |
| | | | 1 | 2 | 0.412428 | 0.304011 |
| | | | 0 | 3 | 0.151466 | 0.267940 |
| E(LU) | | | | | 0.2753763 | |
| E(CP) | | | | | 0.0890285 | |
| E(LB) | | | | | | 0.2414673 |
| | 5 | 3 | 0 | 1 | 0.308995 | -0.670600 |
| | | | 2 | 2 | 0.540490 | 0.399652 |
| | | | 0 | 3 | 0.150515 | 0.270948 |
| E(LU) | | | | | 0.2647504 | |
| E(CP) | | | | | 0.0862437 | |
| E(LB) | | | | | | 0.2472479 |

| | N | M | R | I | A(N,M,I) | C(N,M,I) |
|---|---|---|---|---|---|---|
| | 5 | 2 | 3 | 1 | 0.690189 | -0.396152 |
| | | | 0 | 2 | 0.309811 | 0.396152 |
| E(LU) | | | | | 0.3773561 | |
| E(CP) | | | | | 0.1993557 | |
| E(LB) | | | | | | 0.3724040 |
| | 6 | 6 | 0 | 1 | 0.368179 | -0.405733 |
| | | | 0 | 2 | 0.226486 | -0.031796 |
| | | | 0 | 3 | 0.163591 | 0.064666 |
| | | | 0 | 4 | 0.117541 | 0.111951 |
| | | | 0 | 5 | 0.079377 | 0.132102 |
| | | | 0 | 6 | 0.044826 | 0.128810 |
| E(LU) | | | | | 0.1903043 | |
| E(CP) | | | | | 0.0277157 | |
| E(LB) | | | | | | 0.1165767 |
| | 6 | 5 | 0 | 1 | 0.344243 | -0.474513 |
| | | | 0 | 2 | 0.225354 | -0.035049 |
| | | | 0 | 3 | 0.168128 | 0.077704 |
| | | | 0 | 4 | 0.124635 | 0.132333 |
| | | | 1 | 5 | 0.137640 | 0.299524 |
| E(LU) | | | | | 0.1928985 | |
| E(CP) | | | | | 0.0351703 | |
| E(LB) | | | | | | 0.1379977 |
| | 6 | 4 | 0 | 1 | 0.294288 | -0.583222 |
| | | | 0 | 2 | 0.224688 | -0.036498 |
| | | | 0 | 3 | 0.179161 | 0.101713 |
| | | | 2 | 4 | 0.301863 | 0.518007 |
| E(LU) | | | | | 0.2006959 | |
| E(CP) | | | | | 0.0521384 | |
| E(LB) | | | | | | 0.1749228 |
| | 6 | 3 | 0 | 1 | 0.186635 | -0.767957 |
| | | | 0 | 2 | 0.227525 | -0.031630 |
| | | | 3 | 3 | 0.585840 | 0.799587 |
| E(LU) | | | | | 0.2242859 | |
| E(CP) | | | | | 0.0926197 | |
| E(LB) | | | | | | 0.2443900 |
| | 6 | 2 | 0 | 1 | -0.086117 | -1.140225 |
| | | | 4 | 2 | 1.086117 | 1.140225 |
| E(LU) | | | | | 0.3099333 | |
| E(CP) | | | | | 0.2095161 | |
| E(LB) | | | | | | 0.4039368 |
| | 6 | 5 | 1 | 1 | 0.424637 | -0.426052 |
| | | | 0 | 2 | 0.229636 | 0.006710 |
| | | | 0 | 3 | 0.164432 | 0.111500 |
| | | | 0 | 4 | 0.113812 | 0.153673 |
| | | | 0 | 5 | 0.067483 | 0.154169 |
| E(LU) | | | | | 0.2080218 | |
| E(CP) | | | | | 0.0436594 | |
| E(LB) | | | | | | 0.1379928 |

Table of Weights (continued)

| | N | M | R | I | A(N,M,I) | C(N,M,I) |
|---|---|---|---|---|---|---|
| | 6 | 4 | 1 | 1 | 0.385032 | -0.516531 |
| | | | 0 | 2 | 0.231864 | 0.011800 |
| | | | 0 | 3 | 0.176219 | 0.138428 |
| | | | 1 | 4 | 0.206885 | 0.366303 |
| E(LU) | | | | | 0.2143050 | |
| E(CP) | | | | | 0.0580136 | |
| E(LB) | | | | | | 0.1707856 |
| | 6 | 3 | 1 | 1 | 0.295775 | -0.674565 |
| | | | 0 | 2 | 0.240956 | 0.027898 |
| | | | 2 | 3 | 0.463269 | 0.646667 |
| E(LU) | | | | | 0.2350371 | |
| E(CP) | | | | | 0.0947210 | |
| E(LB) | | | | | | 0.2357782 |
| | 6 | 2 | 1 | 1 | 0.065313 | -0.996263 |
| | | | 3 | 2 | 0.934687 | 0.996263 |
| E(LU) | | | | | 0.3146701 | |
| E(CP) | | | | | 0.2058790 | |
| E(LB) | | | | | | 0.3909415 |
| | 6 | 5 | 0 | 1 | 0.334156 | -0.467085 |
| | | | 1 | 2 | 0.321629 | 0.050716 |
| | | | 0 | 3 | 0.164551 | 0.106033 |
| | | | 0 | 4 | 0.113312 | 0.153765 |
| | | | 0 | 5 | 0.066351 | 0.156571 |
| E(LU) | | | | | 0.2009168 | |
| E(CP) | | | | | 0.0422665 | |
| E(LB) | | | | | | 0.1402241 |
| | 6 | 4 | 0 | 1 | 0.290232 | -0.570735 |
| | | | 1 | 2 | 0.329005 | 0.068121 |
| | | | 0 | 3 | 0.175736 | 0.132427 |
| | | | 1 | 4 | 0.205027 | 0.370187 |
| E(LU) | | | | | 0.2070530 | |
| E(CP) | | | | | 0.0567461 | |
| E(LB) | | | | | | 0.1743920 |
| | 6 | 3 | 0 | 1 | 0.187417 | -0.756373 |
| | | | 1 | 2 | 0.352977 | 0.111403 |
| | | | 2 | 3 | 0.459606 | 0.644970 |
| E(LU) | | | | | 0.2279285 | |
| E(CP) | | | | | 0.0944381 | |
| E(LB) | | | | | | 0.2424467 |
| | 6 | 5 | 0 | 1 | 0.334968 | -0.474071 |
| | | | 0 | 2 | 0.226994 | -0.031821 |
| | | | 1 | 3 | 0.258560 | 0.198010 |
| | | | 0 | 4 | 0.114431 | 0.150884 |
| | | | 0 | 5 | 0.065047 | 0.156998 |
| E(LU) | | | | | 0.1970046 | |
| E(CP) | | | | | 0.0398707 | |
| E(LB) | | | | | | 0.1406386 |

| | N | M | R | I | A(N,M,I) | C(N,M,I) |
|---|---|---|---|---|---|---|
| | 6 | 4 | 0 | 1 | 0.291007 | -0.580177 |
| | | | 0 | 2 | 0.225753 | -0.034817 |
| | | | 1 | 3 | 0.279310 | 0.248091 |
| | | | 1 | 4 | 0.203930 | 0.366903 |
| E(LU) | | | | | 0.2029703 | |
| E(CP) | | | | | 0.0542696 | |
| E(LB) | | | | | | 0.1753921 |
| | 6 | 5 | 0 | 1 | 0.339015 | -0.476295 |
| | | | 0 | 2 | 0.225877 | -0.033656 |
| | | | 0 | 3 | 0.169829 | 0.079403 |
| | | | 1 | 4 | 0.199919 | 0.274294 |
| | | | 0 | 5 | 0.065360 | 0.156253 |
| E(LU) | | | | | 0.1945469 | |
| E(CP) | | | | | 0.0374385 | |
| E(LB) | | | | | | 0.1397804 |
| | 6 | 4 | 2 | 1 | 0.486573 | -0.441353 |
| | | | 0 | 2 | 0.238698 | 0.068993 |
| | | | 0 | 3 | 0.169739 | 0.178817 |
| | | | 0 | 4 | 0.104990 | 0.193543 |
| E(LU) | | | | | 0.2342989 | |
| E(CP) | | | | | 0.0694903 | |
| E(LB) | | | | | | 0.1711390 |
| | 6 | 3 | 2 | 1 | 0.417844 | -0.568051 |
| | | | 0 | 2 | 0.253469 | 0.096222 |
| | | | 1 | 3 | 0.328687 | 0.471828 |
| E(LU) | | | | | 0.2513597 | |
| E(CP) | | | | | 0.1009410 | |
| E(LB) | | | | | | 0.2291164 |
| | 6 | 2 | 2 | 1 | 0.233147 | -0.833181 |
| | | | 2 | 2 | 0.766853 | 0.833181 |
| E(LU) | | | | | 0.3232776 | |
| E(CP) | | | | | 0.2041786 | |
| E(LB) | | | | | | 0.3773135 |
| | 6 | 4 | 1 | 1 | 0.375829 | -0.508619 |
| | | | 1 | 2 | 0.351839 | 0.137547 |
| | | | 0 | 3 | 0.169295 | 0.174495 |
| | | | 0 | 4 | 0.103037 | 0.196577 |
| E(LU) | | | | | 0.2239751 | |
| E(CP) | | | | | 0.0666306 | |
| E(LB) | | | | | | 0.1742935 |
| | 6 | 3 | 1 | 1 | 0.294593 | -0.663601 |
| | | | 1 | 2 | 0.380432 | 0.192098 |
| | | | 1 | 3 | 0.324975 | 0.471503 |
| E(LU) | | | | | 0.2407387 | |
| E(CP) | | | | | 0.0986124 | |
| E(LB) | | | | | | 0.2353090 |

Table of Weights (continued)

| | N | M | R | I | A(N,M,I) | C(N,M,I) |
|---|---|---|---|---|---|---|
| | 6 | 4 | 1 | 1 | 0.378133 | -0.516858 |
| | | | 0 | 2 | 0.233933 | 0.014450 |
| | | | 1 | 3 | 0.285838 | 0.306904 |
| | | | 0 | 4 | 0.102096 | 0.195504 |
| E(LU) | | | | | 0.2181133 | |
| E(CP) | | | | | 0.0623135 | |
| E(LB) | | | | | | 0.1737208 |
| | 6 | 4 | 0 | 1 | 0.277936 | -0.562959 |
| | | | 2 | 2 | 0.450640 | 0.191850 |
| | | | 0 | 3 | 0.169062 | 0.171501 |
| | | | 0 | 4 | 0.102362 | 0.199608 |
| E(LU) | | | | | 0.2169906 | |
| E(CP) | | | | | 0.0654842 | |
| E(LB) | | | | | | 0.1774456 |
| | 6 | 3 | 0 | 1 | 0.185319 | -0.743564 |
| | | | 2 | 2 | 0.491001 | 0.270556 |
| | | | 1 | 3 | 0.323680 | 0.473008 |
| E(LU) | | | | | 0.2337926 | |
| E(CP) | | | | | 0.0982484 | |
| E(LB) | | | | | | 0.2413363 |
| | 6 | 4 | 0 | 1 | 0.281582 | -0.571948 |
| | | | 1 | 2 | 0.333212 | 0.072693 |
| | | | 1 | 3 | 0.283769 | 0.300589 |
| | | | 0 | 4 | 0.101437 | 0.198666 |
| E(LU) | | | | | 0.2109359 | |
| E(CP) | | | | | 0.0611577 | |
| E(LB) | | | | | | 0.1772666 |
| | 6 | 4 | 0 | 1 | 0.282202 | -0.580180 |
| | | | 0 | 2 | 0.227446 | -0.031784 |
| | | | 2 | 3 | 0.389321 | 0.413451 |
| | | | 0 | 4 | 0.101031 | 0.198513 |
| E(LU) | | | | | 0.2071079 | |
| E(CP) | | | | | 0.0588670 | |
| E(LB) | | | | | | 0.1780699 |
| | 6 | 3 | 3 | 1 | 0.558677 | -0.442243 |
| | | | 0 | 2 | 0.265764 | 0.180578 |
| | | | 0 | 3 | 0.175559 | 0.261664 |
| E(LU) | | | | | 0.2791696 | |
| E(CP) | | | | | 0.1170952 | |
| E(LB) | | | | | | 0.2298169 |
| | 6 | 2 | 3 | 1 | 0.425791 | -0.640304 |
| | | | 1 | 2 | 0.574209 | 0.640304 |
| E(LU) | | | | | 0.3399281 | |
| E(CP) | | | | | 0.2076534 | |
| E(LB) | | | | | | 0.3647904 |
| | 6 | 3 | 2 | 1 | 0.411120 | -0.561145 |
| | | | 1 | 2 | 0.416698 | 0.297961 |
| | | | 0 | 3 | 0.172182 | 0.263185 |
| E(LU) | | | | | 0.2619327 | |
| E(CP) | | | | | 0.1104114 | |
| E(LB) | | | | | | 0.2339878 |
| | 6 | 3 | 1 | 1 | 0.284690 | -0.655749 |
| | | | 2 | 2 | 0.543917 | 0.389715 |
| | | | 0 | 3 | 0.171393 | 0.266033 |
| E(LU) | | | | | 0.2517755 | |
| E(CP) | | | | | 0.1082550 | |
| E(LB) | | | | | | 0.2394112 |
| | 6 | 3 | 0 | 1 | 0.172283 | -0.735150 |
| | | | 3 | 2 | 0.656027 | 0.466005 |
| | | | 0 | 3 | 0.171690 | 0.269146 |
| E(LU) | | | | | 0.2451464 | |
| E(CP) | | | | | 0.1079544 | |
| E(LB) | | | | | | 0.2447258 |
| | 6 | 2 | 4 | 1 | 0.662928 | -0.391517 |
| | | | 0 | 2 | 0.337072 | 0.391517 |
| E(LU) | | | | | 0.3790750 | |
| E(CP) | | | | | 0.2294126 | |
| E(LB) | | | | | | 0.3636243 |

$$\sum_{i=1}^{m} c'_{i,m,n}=0, \quad \sum_{i=1}^{m} c'_{i,m,n}E(Y_{i,n})=1, \tag{18}$$

and also so as to minimize the variance of $\hat{a}_{m,n}$ and $\hat{b}_{m,n}$

$$\sum_{i=1}^{m}\sum_{j=1}^{m} a'_{i,m,n}a'_{j,m,n}Cov(X_{i,n}, X_{j,n}), \tag{19}$$

$$\sum_{i=1}^{m}\sum_{j=1}^{m} c'_{i,m,n}c'_{j,m,n}Cov(X_{i,n}, X_{j,n}). \tag{20}$$

Employing the method of Lagrange multipliers, we have

$$\sum_{i=1}^{m}\sum_{j=1}^{m} a'_{i,m,n}a'_{j,m,n}Cov(Y_{i,n}, Y_{j,n})+\lambda_1\sum_{i=1}^{m} a'_{i,m,n}+\lambda_2\sum_{i=1}^{m} a'_{i,m,n}E(Y_{i,n}), \tag{21}$$

$$\sum_{i=1}^{m}\sum_{j=1}^{m} c'_{i,m,n}c'_{j,m,n}Cov(Y_{i,n}, Y_{j,n})+\lambda_1\sum_{i=1}^{m} c'_{i,m,n}+\lambda_2\sum_{i=1}^{m} c'_{i,m,n}E(Y_{i,n}), \tag{22}$$

where the $\lambda$'s are Lagrange multipliers. Differentiating (21) and (22) with respect to $a'_{i,m,n}$ and $c'_{i,m,n}$, respectively, and letting $a_i=a_{i,m,n}$, $c_i=c_{i,m,n}$, $\sigma_{ii}=\sigma^2_{i,n}$, $\sigma_{ij}=\sigma_{i,j,n}$, $\mu_i=E(Y_{i,n})$, we have

$$\begin{bmatrix} \Sigma & H' \\ H & 0 \end{bmatrix}\begin{bmatrix} a \\ \lambda_{\mathrm{I}} \end{bmatrix}=\begin{bmatrix} 0 \\ b \end{bmatrix} \tag{23}$$

$$\begin{bmatrix} \Sigma & H' \\ H & 0 \end{bmatrix}\begin{bmatrix} c \\ \lambda_{\mathrm{II}} \end{bmatrix}=\begin{bmatrix} 0 \\ d \end{bmatrix} \tag{24}$$

where $\Sigma=\|\sigma_{ij}\|$, $H'=(1, \mu)$, $a'=(a_1,\dots, a_m)$, $c'=(c_1,\dots, c_m)$, $\mu'=(\mu_1,\dots, \mu_m)$, $\lambda_{\mathrm{I}}'=(\lambda_1, \lambda_2)$, $\lambda'_{\mathrm{II}}=(\lambda_3, \lambda_4)$, $b'=(1, 0)$, $d'=(0, 1)$.

It will be observed that $Var(\hat{u}_{m,n})=-\lambda_1 b^2=\alpha_{m,n}b^2$, $Var(\hat{b}_{m,n})=-\lambda_4 b^2=\gamma_{m,n}b^2$, and $\mathrm{Cov}(\hat{u}_{m,n}, \hat{b}_{m,n})=-\lambda_2 b^2=-\lambda_3 b^2=\beta_m b^2$. Exactly as in Mann[23], we consider weights $\{A_{i,m,n}\}$ and $\{C_{i,m,n}\}$ such that $\tilde{u}_{m,n}=\sum_{i=1}^{m}A_{i,m,n}X_{i,n}$ and $\tilde{b}_{m,n}=\sum_{i=1}^{m}C_{i,m,n}X_{i,n}$ have the smallest mean squared errors among linear estimators for $u$ and $b$, with mean squared independent of $u$. Then we have $A_{i,m,n}=a_{i,m,n}-\beta_{m,n}c_{i,m,n}/(1+\gamma_{m,n})$ and $C_{i,m,n}=c_{i,m,n}/(1+\gamma_{m,n})$. The mean squared errors of $\tilde{u}_{m,n}$ and $\tilde{b}_{m,n}$ are equal to $b^2E(LU)=[\alpha_{m,n}-\beta^2_{m,n}/(1+\gamma_{m,n})]b^2$ and $b^2E(LB)=[\gamma_{m,n}/(1+\gamma_{m,n})]b^2$, respectively. The expected value of $(\tilde{u}_{m,n}-u)(\tilde{b}_{m,n}-b)$ is equal to $b^2E(CP)=[\beta_{m,n}/(1+\gamma_{m,n})]b^2$. Hence the mean squared error of $\tilde{x}_{p,m,n}$ is equal to $MSE(\tilde{x}_{p,m,n})=\{\alpha_{m,n}-2\beta_{m,n}\ ln\ ln(1/p)+\gamma_{m,n}\cdot[ln\ ln(1/p)]^2-[\beta_{m,n}-\gamma_{m,n}\ ln\ ln(1/p)]^2/(1+\gamma_{m,n})\}b^2$. The values for $A_{i,m,n}$ and $C_{i,m,n}$ are given in Table 1 for the case of all possible censorings for sample sizes 2 through 6. For any sample size $n$, there are $2^{n-1}$ possible censorings to be considered. If unbiased estimates are desired, one can obtain by noting that $\hat{u}_{m,n}=\tilde{u}_{m,n}+E(CP)\tilde{b}_{m,n}/(1-E(LB))$ and $\hat{b}_{m,n}=\tilde{b}_{m,n}/(1-E(LB))$. These values were computed in ACOS900 (Computing Center, Osaka University, Osaka 56500) quadruple precision with about 36 significant digits of accuracy using explicit formulas in samples from the extreme-value distribution given by Lieblein[12].

## 5. Formulation of the Problem for a sample

Suppose we are given observations $X_1, X_2,\dots, X_n$. It is hoped that they are a random sample from a population with the largest extreme-value distribution. But possibly one or more observations are spurious, coming from a different source, and ought to be removed. The problem can be simplified by restricting our attention to excessively large observations. They may spoil the estimates. It is therefore desirable to find a rule and samples which meet the following consideration; that is, if the model is adequate, it should be possible to provide reasonably precise estimation. The rule which we shall now propose is the following:

Rule 0. If $X_i$ is removed, find $j$ such that $X_{j-1,n}<O_{m-1}(X)\leq X_{j,n}$. Consider the remaining observations as progressively censored samples, where censoring occurs progressively in the $j$th stage.

The justification for this rule is based on the fact that the asymptotic distribution of $O_m(X^{(i)})$ is the extreme-value distribution. (This distribution was originally derived by Fisher and Tippett[5] and has been studied by Gumbel[6] [7].) Thus this rule says "a compromise between the desire to include the data and the need for precise estimation."

The question remains, how should the rejection rules be formulated? In some situations it is suggested (see Anscombe[1], in particular) that the rejection rules are not significant tests but insurance policies. We will not further discuss the rejection rules here.

## 6. An Example: Estimation of the 100*p*th Percentile of Distribution

As an illustration of the preceding ideas for the extreme-value model consider the following problem. Suppose we are given the annual maximum twelve-hour rainfalls in millimeters, observed for 18 years in Uchinomi, Kagawa Prefecture, Japan; that is, 82, 94, 180, 59, 55, 74, 168, 66, 99, 123, 139, 194, 223, 146, 53, 279, 230, 472. Based on previous experience, we are willing to assume that the extreme-value distribution is indeed the appropriate one. We wish to find the best linear invariant estimates of $u$, $b$, and $x_p$. But the value $X_{18}=472$ may be judged excessively large and indicative of a spurious observation. Let us $O_{m-1}(X^{(18)})=190$. We find that $R_1=\cdots=R_{13}=0$, $R_{14}=1$, $R_{15}=R_{16}=R_{17}=0$. Then the 99th percentile of distribution $x_{.99}=u-b\ ln\ ln(1/.99)$ is estimated by $\tilde{x}_{.99}=\tilde{u}-\tilde{b}\ ln\ ln(1/.99)=104.2-55.50(-4.600)=359.5$. The mean squared error of $\tilde{x}_{.99}$ is $.9330\,b^2$.

The progressive censoring we have introduced here is an attempt to provide robust estimation in the presence of outliers. Of course, the Type II progressive censoring model considered is violated slightly in our problem. But the minor violation will make little difference in parameter estimation. In comparison, with $R_1=\cdots=R_{16}=0$, $R_{17}=1$ we have $\tilde{x}_{.99}=105.7-58.93(-4.600)=376.8$ and $MSE(\tilde{x}_{.99})=.9759b^2$. Thus the single stage censoring model gives a conservative estimates of $\tilde{x}_p$.

## References


(1) ANSCOMBE, F. J.: Rejection of outliers, *Technometrics*, **2**, 123–147 (1960).

(2) COHEN, A. C.: Progressively censored samples in life testing, *Technometrics*, **5**, 327–339 (1963).

(3) COHEN, A. C.: Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples, *Technometrics*, **7**, 579–588 (1965).

(4) COHEN, A. C.: Multi-censored sampling in the three parameter Weibull distribution, *Technometrics*, **17**, 347–351 (1975).

(5) FISHER, R. A. and TIPPETT, L. H. C.: Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Proc. Cambridge Philos. Soc.*, **24**, 180–190 (1928).

(6) GUMBEL, E. J.: The return period of flood flows, *AMS*, **12**, 163–190 (1941).

(7) GUMBEL, E. J.: *Statistics of extremes*, Columbia University Press (1958).

(8) HARTER, H. L. and MOORE, A. H.: Maximum-likelihood estimation of the parameters of the gamma and Weibull populations from complete and from censored samples, *Technometrics*, **7**, 639–643 (1965).

(9) HARTER, H. L. and MOORE, A. H.: Maximum-likelihood estimation, from doubly censored samples, of the parameters of the first asymptotic distribution of extreme values, *JASA*, **63**, 889–901 (1968).

(10) HERD, G. R.: Estimation of reliability functions, *Proceedings Third National Symposium on Reliability and Quality Control* (1957).

(11) HERD, G. R.: Estimation of reliability from incomplete data, *Proceedings Sixth National Symposium on Reliability and Quality Control* (1960).

(12) LIEBLEIN, J.: On the exact evaluation of the variances and covariances of order statistics in samples from the extreme-value distribution, *AMS*, **24**, 282–287 (1953).

(13) LIEBLEIN, J.: A new method of analyzing extreme-value data, *Technical Note* 3053, National Advisory Committee for Aeronautics.

(14) LIEBLEIN, J. and ZELEN, M.: Statistical investigation of the fatigue life of deep groove ball bearings, Research Paper 2719, *Journal of Research, National Bureau Standards*, Vol. 57, 273–316 (1956).

(15) MANN, N. R.: Results on location and scale

parameter estimation with application to the extreme-value distribution, *Aerospace Research Laboratories Report ARL* 67–0023, Office of Aerospace Research, U. S. Air Force, Wright-Patterson Air Force Base, Ohio (1967).

(16) MANN, N. R.: Tables for obtaining the best linear invariant estimates of parameters of the Weibull distribution, *Technometrics*, **9**, 629–645 (1967).

(17) MANN, N. R.: Point and interval estimation procedures for the two-parameter Weibull and extreme-value distributions, *Technometrics*, **10**, 231–256 (1968).

(18) MANN, N. R.: Results on statistical estimation and hypothesis testing with application to the Weibull and extreme-value distributions, *Aerospace Research Laboratories Report ARL* 68–0068, Office of Aerospace Research, U. S. Air Force, Wright-Patterson Air Force Base, Ohio (1968).

(19) MANN, N. R.: Exact three-order-statistic confidence bounds on reliable life for a Weibull model with progressive censoring, *JASA*, **64**, 306–315 (1969).

(20) MANN, N. R.: Cramér-Rao efficiencies of best linear invariant estimators of parameters of the extreme-value distribution under Type II censoring from above, *SIAM Journal of Applied Mathematics*, Vol. **17**, 1150–1162 (1969).

(21) MANN, N. R.: Optimum estimators for linear functions of location and scale parameters, *AMS*, **40**, 2149–2155 (1969).

(22) MANN, N. R.: Estimation of location and scale parameters under various models of censoring and truncation, *Aerospace Research Laboratories Report ARL* 70–0026, Office of Aerospace Research, U. S. Air Force, Wright-Patterson Air Force Base, Ohio (1970).

(23) MANN, N. R.: Best linear invariant estimation for Weibull parameters under progressive censoring, *Technometrics*, **13**, 521–533 (1971).

(24) MANN, N. R., SCHAFER, R. E. and SINGPURWALLA, N. D.: *Methods for statistical analysis of reliability and life data*, John Wiley and Sons (1974).

(25) VAN MONTFORT, M. A. J.: On testing that the distribution of extremes is of Type I when Type II is the alternative, *Journal of Hydrology*, **11**, 421–427 (1970).

(Received October 15, 1979)