BAYESIAN PREDICTION OF FUTURE LIFETIMES FROM THE PARETO MODEL WITH INCOMPLETE DATA

Zeinab H. Amin

Faculty of Economics and Political Sciences, Department of Statistics, Cairo University, Giza, Egypt.

Key Words and Phrases: Pareto distribution; prediction interval; censored sampling; conjugate priors.

ABSTRACT

This paper discusses the problem of predicting, on the basis of an incomplete sample from a two parameter Pareto distribution, the life time of the j-th item to fail either in the same sample or in a future sample from the same distribution. Prediction intervals are based on doubly censored data when lifetimes are left and right censored, of which complete sampling and Type II censoring are special cases, as well as the case when an extra observation is missing elsewhere.

1. INTRODUCTION

The life distribution under consideration in this study is the two parameter Pareto distribution with probability density function

$$f(x;\alpha,\sigma)=\alpha\sigma^{\alpha}x^{-(\alpha+1)} \quad (x\geq\sigma) \quad (1.1)$$

where $\alpha>0$ and $\sigma>0$.

The Pareto distribution has found wide spread use as a model for various socio-economic phenomena; see, for example, Johnson et al. (1994). It has also been used in reliability and lifetime modeling; see, for example, Davis and Feldstein (1979).

This article is concerned with prediction of the behavior of further observations from the distribution (1.1). Censoring is common in life distribution work because of time limits and other restrictions on data collection. Lifetimes can be censored either on the left, on the right, or both; Lawless (1982). Prediction of unknown observables will be considered under four different sampling plans

One sample prediction

A) Suppose n items are simultaneously put on test and observed until there have been r failures. However $k-1$ lifetimes are censored on the left. The actual observed lifetimes are the middle $r-k+1$ observations. Denote by $X_{(j)}$ the lifetime of the j-th item to fail. Prediction intervals will be derived for $X_{(j)}$ $(r < j \leq n)$.

B) Another sampling plan is considered when data is missing from both extremes, left and right censored, and one extra observation is missing elsewhere. Prediction intervals will be derived for $X_{(j)}$ $(r < j \leq n)$ given the failures at $X_{(k)}$, $X_{(k+1)}$, ..., $X_{(k+l-1)}$, $X_{(k+l+1)}$, ..., $X_{(r)}$; that is, $X_{(k+l)}$ is missing for $l = 1, 2, ...,$ $r-k-1$.

Two sample prediction

A) Suppose n items are simultaneously put on test and only the middle $r-k+1$ observations were observed. Let Y_1, Y_2, ..., Y_N be a second independent random sample of size N of future observations from the same distribution (1.1). Prediction intervals will be derived for $Y_{(j)}$ $j = 1,2,....,N$, the lifetime of the j-th item to fail in the future experiment given the failures at $X_{(k)}$, $X_{(k+1)}$, ..., $X_{(r)}$.

B) Data is missing from both extremes, lifetimes are left and right censored, and one extra observation is missing elsewhere. Prediction intervals will be derived for $Y_{(j)}$ $j = 1,2,....,N$ given the failures at $X_{(k)}$, $X_{(k+1)}$, ..., $X_{(k+l-1)}$, $X_{(k+l+1)}$, ..., $X_{(r)}$ for $l = 1, 2, ..., r-k-1$.

Applications of lifetime distribution methodology range from investigations into the endurance of manufactured items to research involving human diseases. The following two applications can be thought of as applications for sampling plans A and B respectively.

Manufactured items such as mechanical or electronic components are often subjected to life tests in order to obtain information on their endurance. Some types of manufactured items can be repaired should they fail. In this case one might be interested in the length of time between successive failures of an item and refer to these times as lifetimes. Suppose an item (for example a mechanical or electronic component of an automobile or of a computer) is put into operation and observed until it fails (ceases operating satisfactorily). If this component is known to have failed twice for example before the beginning of our study (studies will not always be performed on new cars or new computers etc.), then for these first two life times (the length of time between successive failures referred to as failure times in this case) only an upper bound on lifetime is available. Hence, information available on their lifetimes is partial. The next failure time is treated as the third lifetime denoted by $X_{(3)}$ where $X_{(1)}$ and $X_{(2)}$ are left censored. By treating the third lifetime as the first observed lifetime $X_{(1)}$, some information is lost as a result of treating the component of the automobile or of the computer in this case as a new component first time to fail. Also, it may not be feasible to continue experimentation until this mechanical or electronic component stops work completely. It may take a very long time for this unit to completely fail (could not be repaired any more) and it is deemed necessary to terminate the experiment before this can happen. In this case some of the lifetimes are right censored and only a lower bound on lifetime is available. Hence, the information available on their lifetimes are partial.

Sometimes the events of interest are deaths of individuals and lifetime here is the actual length of life of an individual measured from some particular starting point. In medical studies dealing with fatal diseases one might be interested in the survival time of individuals with the disease, measured from the date of diagnosis or some other starting point. For example, a study might focus on comparing the effects of two chemotherapy treatments on advanced lung cancer patients in prolonging survival time. Patients are randomly assigned to one of the two treatments. Survival times from the start of treatment for each patient are recorded. Suppose a patient in one of the groups died in an accident during the period of the study (any cause other than lung cancer) his lifetime $X_{(k+l)}$ cannot be used in the study because it is not indicative of the treatment effect. However, neglecting this lifetime totally leads to a loss of information because it is known that the lifetime of this patient exceeded $X_{(k+l-1)}$, hence should be treated as a missing observation.

A number of authors have considered prediction intervals for the two parameter Pareto distribution within a Bayesian framework. Nigm and Hamdy (1987) considered the problem of predicting $X_{(j)}$ $(r < j \le n)$ given the failures at $X_{(1)}, X_{(2)}, \ldots, X_{(r)}$ from the distribution (1.1). Geisser (1984) and Arnold and Press (1989) considered the problem of predicting the fraction out of N future observations that survive beyond a certain threshold, when the present and future observations are Pareto distributed and where N is the size of the future sample. Geisser (1985) extended the work of Geisser (1984) to predicting the fraction that fall within a prescribed interval. Dunsmore and Amin (1998) presented a method which incorporates a missing data approach within a Gibbs sampling routine to establish predictive densities for sums of the total amount of testing time that remains until all the items have failed given the failures at $X_{(1)}, X_{(2)}, \ldots, X_{(r)}$ from the distribution (1.1) as well as the total amount of testing time up to the j^{th} failure in a future sample from the same distribution. Al-Hussaini (1999) proposed a general class of distributions that includes the Pareto, among others, as the population model and obtained the predictive density under a proper general prior density, imposing type II censoring on the informative sample. Soliman (2000) considered the problem of predicting $X_{(j)}$ $(r < j \le n)$ given the failures at $X_{(1)}, X_{(2)}, \ldots, X_{(r)}$ from the distribution (1.1) where the sample size n is a random variable having a Poisson or Binomial distribution.

2. BAYESIAN MODELS FOR THE PARETO DISTRIBUTION

When both α and σ are unknown, a natural conjugate prior for (α,σ) was first suggested by Lwin (1972) and later generalized by Arnold and Press (1983) to include broader classes of prior distributions. The generalized Lwin prior or the Power-Gamma prior, denoted by $PG(\nu,\lambda,\mu,\theta)$, is given by

$$g(\alpha,\sigma) \propto \sigma^{\lambda\alpha-1}\alpha^{\nu}\mu^{-\alpha} \qquad (\alpha > 0,\ 0 < \sigma < \theta) \qquad (2.1)$$

where μ, θ, ν and λ are positive constants, and $\theta^{\lambda} < \mu$. Such a prior specifies $g(\alpha)$ as $Ga(\nu, \ln\mu - \lambda\ln\theta)$ and $g(\sigma|\alpha)$ as a power function distribution $PF(\lambda\alpha,\theta)$ of form $\lambda\alpha\sigma^{\lambda\alpha-1}\theta^{-\lambda\alpha}$ $(0 < \sigma < \theta)$.

Vague prior information about α and σ is specified through $\nu = -1$, $\lambda = 0$, $\mu = 1$ and $\theta \to \infty$.

Under double censoring where $k-1$ lifetimes are left censored and $n-r$ lifetimes are right censored, the $r-k+1$ middle observations are the actual observed lifetimes. The likelihood for this data configuration assumes the form

$$L(\alpha,\sigma)=\frac{n!}{(k-1)!(n-r)!}\left\{1-\left(\frac{\sigma}{x_{(k)}}\right)^{\alpha}\right\}^{k-1}\left(\frac{\sigma}{x_{(r)}}\right)^{(n-r)\alpha}\alpha^{r-k+1}\sigma^{(r-k+1)\alpha}\prod_{l=k}^{r}x_{(l)}^{-(\alpha+1)}$$

$$\left(\sigma\le x_{(k)}<x_{(k+1)}<\dots<x_{(r)}\right).$$

Applying the power gamma prior given in (2.1), the corresponding posterior density under this sampling plan is given by

$$g\left(\alpha,\sigma\middle|\mathbf{x}^{(r-k+1)}\right)=\frac{\alpha^{r+v-k+1}\sum_{i=0}^{k-1}\binom{k-1}{i}(-1)^{i}\sigma^{(n+\lambda+i-k+1)\alpha-1}\left(\mu x_{(r)}^{(n-r)}x_{(k)}^{i}\prod_{l=k}^{r}x_{(l)}\right)^{-\alpha}}{\Gamma(r+v-k+1)\sum_{j=0}^{k-1}\binom{k-1}{j}\frac{(-1)^{j}}{(n+\lambda+j-k+1)}\{A(j)\}^{-(r+v-k+1)}}$$

$$(\alpha>0,\sigma<w) \quad (2.2)$$

where $\mathbf{X}^{(r-k+1)}=\left(X_{(k)},X_{(k+1)},\dots,X_{(r)}\right)$, $w=\min\left(\theta,x_{(k)}\right)$ and

$$A(j)=\sum_{l=k}^{r}\ln x_{(l)}+\ln\mu+(n-r)\ln x_{(r)}+j\ln x_{(k)}-(n+\lambda+j-k+1)\ln w \text{ for}$$

$j=0,1,\dots,k-1$.

3. ONE SAMPLE PRÉDICTION

This section considers the problem of predicting $X_{(j)}$, $r<j\le n$, based on the failures at $X_{(k)}$, $X_{(k+1)}$, ..., $X_{(r)}$ from the distribution (1.1). The predictive density function of $X_{(j)}$ given $X^{(r-k+1)}=(X_{(k)},X_{(k+1)},\dots,X_{(r)})$ is given as

$$f\left(x_{(j)}\middle|x^{(r-k+1)}\right)$$

$$=\frac{(r+v-k+1)x_{(j)}^{-1}\sum_{i=0}^{k-1}\sum_{m=0}^{j-r-1}\binom{k-1}{i}\binom{j-r-1}{m}\frac{(-1)^{i+m}}{(n+\lambda+i-k+1)}\left\{A(i)+(n-j+m+1)\ln\frac{x_{(j)}}{x_{(r)}}\right\}^{-(r+v-k+2)}}{\beta(j-r,n-j+1)\sum_{i=0}^{k-1}\binom{k-1}{i}\frac{(-1)^{i}}{(n+\lambda+i-k+1)}\{A(i)\}^{-(r+v-k+1)}}$$

$$\left(x_{(j)}>x_{(r)}\right) \quad (3.1)$$

A Bayesian interval (a_1, a_2) $x_{(r)} \le a_1 < a_2$ of cover $1-\gamma$ could then be derived where a_1 and a_2 are the solutions of

$$\int_{a_1}^{a_2} f\left(x_{(j)} \middle| x^{(r-k+1)}\right) dx_{(j)} = 1-\gamma$$

That is the solutions of

$$\frac{\sum_{i=0}^{k-1}\sum_{m=0}^{j-r-1}\binom{k-1}{i}\binom{j-r-1}{m}\frac{(-1)^{i+m}}{(n+\lambda+i-k+1)(n-j+m+1)}\left[\left\{A(i)+(n-j+m+1)\ln\frac{a_1}{x_{(r)}}\right\}^{-(r+v-k+1)} - \left\{A(i)+(n-j+m+1)\ln\frac{a_2}{x_{(r)}}\right\}^{-(r+v-k+1)}\right]}{\beta(j-r,n-j+1)\sum_{i=0}^{k-1}\binom{k-1}{i}\frac{(-1)^{i}}{(n+\lambda+i-k+1)}\{A(i)\}^{-(r+v-k+1)}}$$

$$= 1-\gamma$$

$$\left(x_{(r)} \le a_1 < a_2\right) \qquad (3.2)$$

In practice a_1 and a_2 should be chosen such that a_2 - a_1 is shortest.

When $k = 1$, (3.1) and (3.2) reduce to the case of Type II censored data, the results derived by Nigm and Hamdy (1987).

Now prediction intervals for $X_{(j)}$, $r < j \le n$, are considered in the case of incomplete data; that is, where data is missing from both extremes, left and right censored, and one extra observation is missing elsewhere. Hence, the problem considered is the problem of predicting $X_{(j)}$, $r < j \le n$ based on the ordered statistics $X^{(r-k)} = \left(X_{(k)}, X_{(k+1)}, \ldots, X_{(k+l-1)}, X_{(k+l+1)}, \ldots, X_{(r)}\right)$; that is, $X_{(k+l)}$ is missing for $l = 1, 2, \ldots, r-k-1$.

$$f\left(x_{(k)},x_{(k+1)},\ldots,x_{(k+l-1)},x_{(k+l+1)},\ldots,x_{(r)}\middle|\alpha,\sigma\right)=\frac{n!}{(k-1)!(n-r)!}\left\{1-\left(\frac{\sigma}{x_{(k)}}\right)^{\alpha}\right\}^{k-1}\left(\frac{\sigma}{x_{(r)}}\right)^{(n-r)\alpha}$$

$$\alpha^{r-k}\sigma^{(r-k+1)\alpha}\left(x_{(k+l-1)}^{-\alpha}-x_{(k+l+1)}^{-\alpha}\right)\prod_{\substack{i=k\\i\neq k+l}}^{r}x_{(i)}^{-(\alpha+1)}$$

$$\sigma\leq x_{(k)}<\ldots<x_{(r)}.$$

The posterior density of (α,σ) is given by

$$g\left(\alpha,\sigma\middle|x^{(r-k)}\right)=\frac{\dfrac{1}{\Gamma(r+v-k)}\alpha^{r+v-k}\left(x_{(k+l-1)}^{-\alpha}-x_{(k+l+1)}^{-\alpha}\right)\displaystyle\sum_{q=0}^{k-1}\binom{k-1}{q}(-1)^{q}\sigma^{(n+\lambda+q-k+1)\alpha-1}\left(\mu x_{(r)}^{n-r}x_{(k)}^{q}\prod_{\substack{i=k\\i\neq k+l}}^{r}x_{(i)}\right)^{-\alpha}}{\displaystyle\sum_{m=0}^{k-1}\binom{k-1}{m}\frac{(-1)^{m}}{(n+\lambda+m-k+1)}\left\{\left(A^{*}(m)+\ln x_{(k+l-1)}\right)^{-(r+v-k)}-\left(A^{*}(m)+\ln x_{(k+l+1)}\right)^{-(r+v-k)}\right\}}$$

$$(\alpha>0,\sigma<w)$$

where $w=\min\left(\theta,x_{(k)}\right)$ and

$$A^{*}(m)=\sum_{\substack{i=k\\i\neq k+l}}^{r}\ln x_{(i)}+\ln\mu+(n-r)\ln x_{(r)}+m\ln x_{(k)}-(n+\lambda+m-k+1)\ln w$$

for $m=0,1,\ldots,k\text{-}1$.

The predictive density of $X_{(j)}$ given $X^{(r-k)}$ is given by

$$f\left(x_{(j)}\middle|x^{(r-k)}\right)=\frac{(r+v-k)}{\beta(j-r,n-j+1)}x_{(j)}^{-1}\cdot$$

$$\frac{\displaystyle\sum_{p=0}^{k-1}\sum_{q=0}^{j-r-1}\binom{k-1}{p}\binom{j-r-1}{q}\frac{(-1)^{p+q}}{(n+\lambda+p-k+1)}\left\{\begin{array}{l}\left(A^{*}(p)+(n-j+q+1)\ln\dfrac{x_{(j)}}{x_{(r)}}+\ln x_{(k+l-1)}\right)^{-(r+v-k+1)}\\-\left(A^{*}(p)+(n-j+q+1)\ln\dfrac{x_{(j)}}{x_{(r)}}+\ln x_{(k+l+1)}\right)^{-(r+v-k+1)}\end{array}\right\}}{\displaystyle\sum_{m=0}^{k-1}\binom{k-1}{m}\frac{(-1)^{m}}{(n+\lambda+m-k+1)}\left\{\left(A^{*}(m)+\ln x_{(k+l-1)}\right)^{-(r+v-k)}-\left(A^{*}(m)+\ln x_{(k+l+1)}\right)^{-(r+v-k)}\right\}}$$

$$\left(x_{(j)}>x_{(r)}\right).\quad(3.3)$$

A Bayesian interval (a_1, a_2) of cover $1-\gamma$ could then be derived where a_1 and a_2 $(x_{(r)} \le a_1 < a_2)$ are the solutions of

$$\frac{1}{\beta(j-r, n-j+1)\sum_{m=0}^{k-1}\frac{\binom{k-1}{m}(-1)^m}{(n+\lambda+m-k+1)}\left\{\left(A^*(m)+\ln x_{(k+l-1)}\right)^{-(r+v-k)}-\left(A^*(m)+\ln x_{(k+l+1)}\right)^{-(r+v-k)}\right\}}$$

$$\cdot\sum_{p=0}^{k-1}\sum_{q=0}^{j-r-1}\frac{\binom{k-1}{p}\binom{j-r-1}{q}(-1)^{p+q}}{(n+\lambda+p-k+1)(n-j+q+1)}\left\{\begin{array}{l}\left(A^*(p)+(n-j+q+1)\ln\frac{a_1}{x_{(r)}}+\ln x_{(k+l-1)}\right)^{-(r+v-k)}\\ -\left(A^*(p)+(n-j+q+1)\ln\frac{a_1}{x_{(r)}}+\ln x_{(k+l+1)}\right)^{-(r+v-k)}\\ -\left(A^*(p)+(n-j+q+1)\ln\frac{a_2}{x_{(r)}}+\ln x_{(k+l-1)}\right)^{-(r+v-k)}\\ +\left(A^*(p)+(n-j+q+1)\ln\frac{a_2}{x_{(r)}}+\ln x_{(k+l+1)}\right)^{-(r+v-k)}\end{array}\right\}$$

$$=1-\gamma. \qquad (3.4)$$

4. TWO SAMPLE PREDICTION

Let X_1, X_2, ..., X_n be a random sample from the distribution (1.1). The items have been tested simultaneously and only the $r-k+1$ life times $X^{(r-k+1)} = (X_{(k)}, X_{(k+1)}, \ldots, X_{(r)})$ for $(k < r \le n)$ are observed. The interest now lies in predicting $Y_{(j)}$ $(j = 1,2,\ldots,N)$ where Y_1, Y_2, ..., Y_N is a set of N independent observations from (1.1).

The probability density function of the life time of the j-th item to fail is given by

$$f(y_{(j)}|\alpha,\sigma)=\frac{1}{\beta(j,N-j+1)}\alpha\sigma^{(N-j+1)\alpha}y_{(j)}^{-(N-j+1)\alpha-1}\left\{1-\left(\frac{\sigma}{y_{(j)}}\right)^{\alpha}\right\}^{j-1} \qquad (y_{(j)}>\sigma).$$

The predictive density function for $Y_{(j)}$ $(j = 1,2,\ldots,N)$ given $x^{(r-k+1)}$ is then expressed as

$$f(y_{(j)}|x^{(r-k+1)})=\iint f(y_{(j)}|\alpha,\sigma)g(\alpha,\sigma|x^{(r-k+1)})d\sigma d\alpha \qquad (y_{(j)}>\sigma, \sigma<w).$$

Two cases are considered separately $Y_{(j)} < w$ and $Y_{(j)} \ge w$:

$Y_{(j)} < w$

$$f\left(y_{(j)}\middle|x^{(r-k+1)}\right)=\frac{(r+v-k+1)}{\beta(j,N-j+1)}y_{(j)}^{-1}\cdot\frac{\sum_{p=0}^{k-1}\sum_{q=0}^{j-1}\binom{k-1}{p}\binom{j-1}{q}\frac{(-1)^{p+q}}{(N-j+q+n+\lambda+p-k+2)}\left\{A(p)-(n+\lambda+p-k+1)\ln\frac{y_{(j)}}{w}\right\}^{-(r+v-k+2)}}{\sum_{i=0}^{k-1}\binom{k-1}{i}\frac{(-1)^{i}}{(n+\lambda+i-k+1)}(A(i))^{-(r+v-k+1)}}\tag{4.1}$$

$Y_{(j)} \geq w$

$$f\left(y_{(j)}\middle|x^{(r-k+1)}\right)=\frac{(r+v-k+1)}{\beta(j,N-j+1)}y_{(j)}^{-1}\cdot\frac{\sum_{p=0}^{k-1}\sum_{q=0}^{j-1}\binom{k-1}{p}\binom{j-1}{q}\frac{(-1)^{p+q}}{(N-j+q+n+\lambda+p-k+2)}\left\{A(p)+(N-j+q+1)\ln\frac{y_{(j)}}{w}\right\}^{-(r+v-k+2)}}{\sum_{i=0}^{k-1}\binom{k-1}{i}\frac{(-1)^{i}}{(n+\lambda+i-k+1)}(A(i))^{-(r+v-k+1)}}\tag{4.2}$$

A Bayesian interval (a_1,a_2) of cover $1-\gamma$ could then be derived where $a_1 < w$ and $a_2 > w$ are the solutions of

$$\frac{1}{\beta(j,N-j+1)\sum_{i=0}^{k-1}\frac{\binom{k-1}{i}(-1)^{i}}{(n+\lambda+i-k+1)}(A(i))^{-(r+v-k+1)}}\sum_{p=0}^{k-1}\sum_{q=0}^{j-1}\frac{\binom{k-1}{p}\binom{j-1}{q}(-1)^{p+q}}{(N-j+q+n+\lambda+p-k+2)}\left[\frac{1}{(n+\lambda+p-k+1)}\left\{\begin{array}{l}(A(p))^{-(r+v-k+1)}\\-\left\{A(p)-(n+\lambda+p-k+1)\ln\frac{a_1}{w}\right\}^{-(r+v-k+1)}\end{array}\right\}-\frac{1}{(N-j+q+1)}\left\{\begin{array}{l}\left\{A(p)+(N-j+q+1)\ln\frac{a_2}{w}\right\}^{-(r+v-k+1)}\\-(A(p))^{-(r+v-k+1)}\end{array}\right\}\right]$$

$$=1-\gamma.\tag{4.3}$$

For $k = 1$ the results of (4.1) to (4.3) reduce to the case of Type II censored data.

The interest now lies in predicting $Y_{(j)}$ $\left(j=1,2,\ldots,N\right)$ in the case of incomplete data. Data are left and right censored and one extra observation is missing elsewhere. The predictive density function for $Y_{(j)}$ $\left(j=1,2,\ldots,N\right)$ given $x^{(r-k)}$ is expressed as

$$f\left(y_{(j)}\middle|x^{(r-k)}\right)=\iint f\left(y_{(j)}\middle|\alpha,\sigma\right)g\left(\alpha,\sigma\middle|x^{(r-k)}\right)d\sigma d\alpha \qquad \left(y_{(j)}>\sigma,\sigma<w\right).$$

Two cases are considered separately $Y_{(j)}<w$ and $Y_{(j)}\geq w$:

$\underline{Y_{(j)}<w}$

$$f\left(y_{(j)}\middle|x^{(r-k)}\right)=\frac{(r+v-k)}{\beta(j,N-j+1)}y_{(j)}^{-1}\cdot\frac{\displaystyle\sum_{p=0}^{k-1}\sum_{q=0}^{j-1}\frac{\binom{k-1}{p}\binom{j-1}{q}(-1)^{p+q}}{(N-j+q+n+\lambda+p-k+2)}\left[\left\{A^{*}(p)-(n+\lambda+p-k+1)\ln\frac{y_{(j)}}{w}+\ln x_{(k+l-1)}\right\}^{-(r+v-k+1)}-\left\{A^{*}(p)-(n+\lambda+p-k+1)\ln\frac{y_{(j)}}{w}+\ln x_{(k+l+1)}\right\}^{-(r+v-k+1)}\right]}{\displaystyle\sum_{i=0}^{k-1}\binom{k-1}{i}\frac{(-1)^{i}}{(n+\lambda+i-k+1)}\left\{\left(A^{*}(i)+\ln x_{(k+l-1)}\right)^{-(r+v-k)}-\left(A^{*}(i)+\ln x_{(k+l+1)}\right)^{-(r+v-k)}\right\}}$$

(4.4)

$\underline{Y_{(j)}\geq w}$

$$f\left(y_{(j)}\middle|x^{(r-k)}\right)=\frac{(r+v-k)}{\beta(j,N-j+1)}y_{(j)}^{-1}\cdot\frac{\displaystyle\sum_{p=0}^{k-1}\sum_{q=0}^{j-1}\frac{\binom{k-1}{p}\binom{j-1}{q}(-1)^{p+q}}{(N-j+q+n+\lambda+p-k+2)}\left[\left\{A^{*}(p)+(N-j+q+1)\ln\frac{y_{(j)}}{w}+\ln x_{(k+l-1)}\right\}^{-(r+v-k+1)}-\left\{A^{*}(p)+(N-j+q+1)\ln\frac{y_{(j)}}{w}+\ln x_{(k+l+1)}\right\}^{-(r+v-k+1)}\right]}{\displaystyle\sum_{i=0}^{k-1}\binom{k-1}{i}\frac{(-1)^{i}}{(n+\lambda+i-k+1)}\left\{\left(A^{*}(i)+\ln x_{(k+l-1)}\right)^{-(r+v-k)}-\left(A^{*}(i)+\ln x_{(k+l+1)}\right)^{-(r+v-k)}\right\}}$$

(4.5)

A Bayesian interval $\left(a_1,a_2\right)$ of cover $1-\gamma$ could then be derived where $a_1<w$ and $a_2>w$ are the solutions of

$$\frac{1}{\beta(j,N-j+1)\sum_{i=0}^{k-1}\binom{k-1}{i}\frac{(-1)^{i}}{(n+\lambda+i-k+1)}\left\{\left(A^{*}(i)+\ln x_{(k+l-1)}\right)^{-(r+v-k)}-\left(A^{*}(i)+\ln x_{(k+l+1)}\right)^{-(r+v-k)}\right\}}$$

$$\sum_{p=0}^{k-1}\sum_{q=0}^{j-1}\frac{\binom{k-1}{p}\binom{j-1}{q}(-1)^{p+q}}{(N-j+q+n+\lambda+p-k+2)}\left(\begin{array}{l}\dfrac{(N-j+q+n+\lambda+p-k+2)}{(n+\lambda+p-k+1)(N-j+q+1)}\left[\begin{array}{l}\left\{A^{*}(p)+\ln x_{(k+l-1)}\right\}^{-(r+v-k)}\\-\left\{A^{*}(p)+\ln x_{(k+l+1)}\right\}^{-(r+v-k)}\end{array}\right]\\-\dfrac{1}{(n+\lambda+p-k+1)}\left[\begin{array}{l}\left\{A^{*}(p)-(n+\lambda+p-k+1)\ln\dfrac{a_1}{w}+\ln x_{(k+l-1)}\right\}^{-(r+v-k)}\\-\left\{A^{*}(p)-(n+\lambda+p-k+1)\ln\dfrac{a_1}{w}+\ln x_{(k+l+1)}\right\}^{-(r+v-k)}\end{array}\right]\\-\dfrac{1}{(N-j+q+1)}\left[\begin{array}{l}\left\{A^{*}(p)+(N-j+q+1)\ln\dfrac{a_2}{w}+\ln x_{(k+l-1)}\right\}^{-(r+v-k)}\\-\left\{A^{*}(p)+(N-j+q+1)\ln\dfrac{a_2}{w}+\ln x_{(k+l+1)}\right\}^{-(r+v-k)}\end{array}\right]\end{array}\right)$$

$$=1-\gamma. \tag{4.6}$$

Another interval which is more relevant for higher order statistics is of the form $(a_1^{\bullet},a_2^{\bullet})$ where $w<a_1^{\bullet}<a_2^{\bullet}$ then $a_1^{\bullet}$ and $a_2^{\bullet}$ are the solutions of

$$\frac{1}{\beta(j,N-j+1)\sum_{i=0}^{k-1}\binom{k-1}{i}\frac{(-1)^{i}}{(n+\lambda+i-k+1)}\left\{\left(A^{*}(i)+\ln x_{(k+l-1)}\right)^{-(r+v-k)}-\left(A^{*}(i)+\ln x_{(k+l+1)}\right)^{-(r+v-k)}\right\}}$$

$$\sum_{p=0}^{k-1}\sum_{q=0}^{j-1}\frac{\binom{k-1}{p}\binom{j-1}{q}(-1)^{p+q}}{(N-j+q+n+\lambda+p-k+2)(N-j+q+1)}\left[\begin{array}{l}\left\{A^{*}(p)+(N-j+q+1)\ln\dfrac{a_1^{\bullet}}{w}+\ln x_{(k+l-1)}\right\}^{-(r+v-k)}\\-\left\{A^{*}(p)+(N-j+q+1)\ln\dfrac{a_1^{\bullet}}{w}+\ln x_{(k+l+1)}\right\}^{-(r+v-k)}\\-\left\{A^{*}(p)+(N-j+q+1)\ln\dfrac{a_2^{\bullet}}{w}+\ln x_{(k+l-1)}\right\}^{-(r+v-k)}\\-\left\{A^{*}(p)+(N-j+q+1)\ln\dfrac{a_2^{\bullet}}{w}+\ln x_{(k+l+1)}\right\}^{-(r+v-k)}\end{array}\right].$$

$$=1-\gamma.$$

Again $a_1^{\bullet}$ and $a_2^{\bullet}$ should be chosen such that $a_2^{\bullet}$ - $a_1^{\bullet}$ is shortest.

5. NUMERICAL EXAMPLE

Consider a life test where 20 units whose lifetimes follow the same Pareto distribution (1.1) are put on test simultaneously. The times of failure of the third to the elevnth items measured in an informative experiment are shown in Table 1.

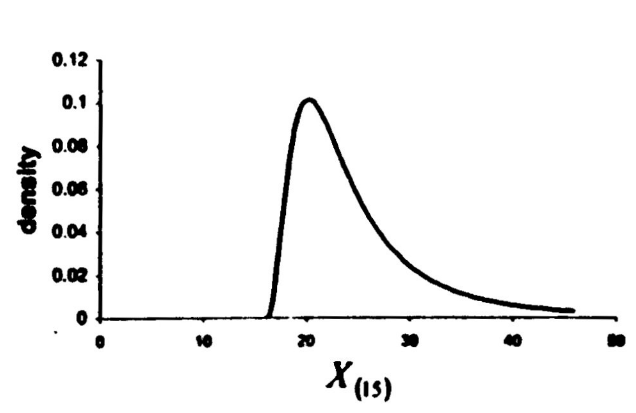
Table 1 Times of failure of the third to the elevnth items

10.425	10.757	10.946	11.433	11.663
11.945	14.712	15.279	16.121	

The results of section (3) are used with $n = 20$, $k = 3$ and $r = 11$ and assume little is known a priori about (α, σ); that is, results of these sections are used under the settings $\nu = -1$, $\lambda = 0$, $\mu = 1$ and $\theta \to \infty$ $(w = x_{(k)})$.

The predictive density function of $X_{(15)}$ given $X^{(9)}$ derived from (3.1) under these settings is shown in figure (1) below.

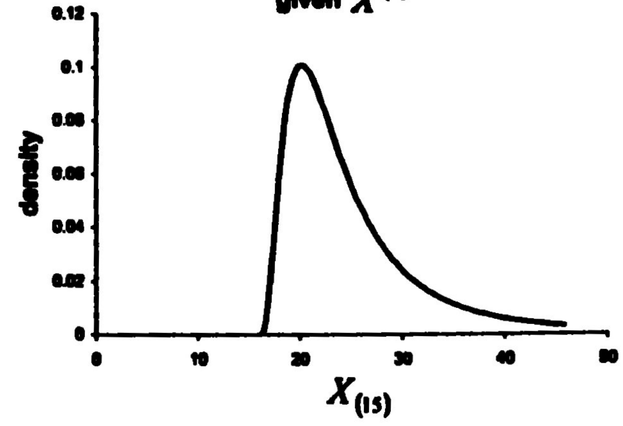
FIG. (1) Predictive density function of $X_{(15)}$ given $X^{(9)}$



From (3.2), a 95% Bayesian interval for $X_{(15)}$ based on $X^{(9)}$ is given by (16.121, 42.8307).

Suppose now the fifth observation, $X_{(5)}$, is missing; that is, $X^{(8)} = (X_{(3)}, X_{(4)}, X_{(6)}, \ldots, X_{(11)})$. The predictive density function of $X_{(15)}$ given $X^{(8)}$ derived from (3.3) is shown in figure (2) below.

FIG. (2) Predictive density function of $X_{(15)}$ given $X^{(8)}$

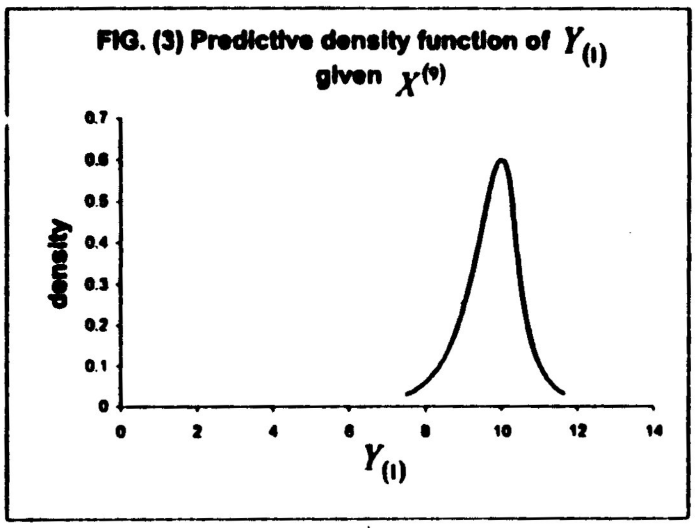


From (3.4), a 95% Bayesian interval for $X_{(15)}$ based on $X^{(8)}$ is given by (16.121, 42.7596).

Consider now the same life test where 20 units whose lifetimes follow the same Pareto distribution (1.1) are put on test simultaneously. The times of failure of the third to the elevnth items measured in the informative experiment are shown in Table 1. However, the interest now lies in predicting the life time of the first item to fail, $Y_{(1)}$, in a future independent sample of size 15 from the same distribution (1.1).

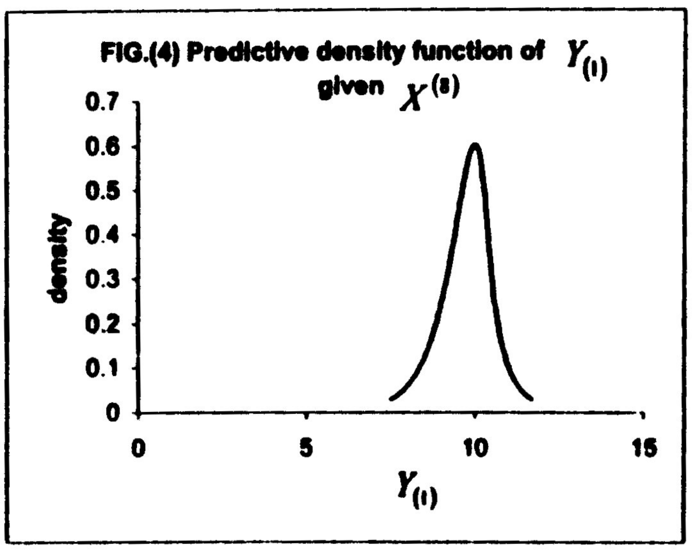
The results of section (4) are used with $n = 20$, $k = 3$, $r = 11$, $N = 15$ and $j = 1$ and assume little is known a priori about (α, σ); that is, results of that section are used under the settings $\nu = -1$, $\lambda = 0$, $\mu = 1$ and $\theta \to \infty$ $(w = x_{(k)})$.

The predictive density function of $Y_{(1)}$ given $X^{(9)}$ derived from (4.1) and (4.2) under these settings is shown in figure (3) below.



From (4.3), a 95% Bayesian interval for $Y_{(1)}$ based on $X^{(9)}$ is given by (7.8799, 11.8134).

When the fifth observation is missing, the predictive density function of $Y_{(1)}$ given $X^{(8)}$ derived from (4.4) and (4.5) is shown in figure (4) below.



From (4.6), a 95% Bayesian interval for $Y_{(1)}$ when the fifth observation is missing is given by (7.8759, 11.7894).

BIBLIOGRAPHY

Al-Hussaini, E. K. (1999). "Predicting observables from a general class of distributions," *J. Statist. Plann. Inf.*, 79, 79-91.

Arnold, B. C. and Press, S. J. (1983). "Bayesian inference for Pareto populations," *J. Econometrics*, 21, 287-306.

Arnold, B. C. and Press, S. J. (1989). "Bayesian estimation and prediction for Pareto data," *J. Amer. Statist. Assoc.*, 84, 1079-1084.

Davis, H. T. and Feldstein, M. L. (1979). "The generalized Pareto law as a model for progressively censored survival data," *Biometrika*, 66, 299-306.

Dunsmore, I. R. and Amin, Z. H. (1998). "Some prediction problems concerning samples from the Pareto distribution," *Comm. Statist.*, 27, 1221-1238.

Geisser, S. (1984). "Predicting Pareto and Exponential observables," *Canadian J. Statist.*, 12, 143-152.

Geisser, S. (1985). "Interval prediction for Pareto and Exponential observables," *J. Econometrics*, 29, 173-185.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions, Vol. 1*, 2nd edtn., New York: Wiley.

Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data.* New York: Wiley.

Lingappaiah, G. S. (1981). "Sequential life testing with spacings, Exponential model," *IEEE Trans. Reliability*, 30, 370-374.

Lwin, T. (1972). "Estimation of the tail of the Paretian law," *Skand. Aktuarietidskr.*, 55, 170-178.

Nigm, A. M. and Hamdy, H. I. (1987). "Bayesian prediction bounds for the Pareto life time model," *Comm. Statist.*, 16, 1761-1772.

Soliman, A. (2000). "Bayes prediction in a Pareto lifetime model with random sample size," *The Statistician*, 49, 51-62.