Interval Prediction For Pareto and
Exponential Observables

by

Seymour Geisser[1]
University of Minnesota
Technical Report No. 443

[1] This work was supported in part by NIH grant NIGMS25271.

1. Introduction.

The two-parameter Pareto distirubtion has served as a model for such economic variables as income, wealth and business size. An extensive historical survey regarding Pareto distribution application is found in Arnold (1982). The translated exponential distribution has been used to describe certain survival data in the Biomedical area c.f. Kalbfleisch and Prentice (1980) and compent failure or equipment reliability in industrial applications c.f. Gnedenko et al. (1969).

As is well known a very simple relationship exists between the two distributions. The Pareto variable, W, has distribution function

$$F(w|\alpha,\beta) = \begin{cases} 0 & w < \beta, \\ 1 - (\frac{\beta}{w})^{\alpha}, & w > \beta > 0, \end{cases} \qquad (1.1)$$

and density

$$f(w|\alpha,\beta) = \frac{\alpha\beta^{\alpha}}{w^{\alpha+1}}, \qquad \alpha > 0,\ w > \beta > 0, \qquad (1.2)$$

with survival function

$$\Pr[W > w|\alpha,\beta] = \min[(\frac{\beta}{w})^{\alpha},1]. \qquad (1.3)$$

The translated exponential variable X has distribution function

$$F(x|\alpha,\gamma) = \begin{cases} 1 - e^{-\alpha(x-\gamma)} & \text{for } x > \gamma, \\ 0 & \text{otherwise.} \end{cases} \quad (1.4)$$

and density

$$f(x|\alpha,\gamma) = \alpha e^{-a(x-y)} \quad \text{for } \alpha > 0,\ x > \gamma > -\infty. \quad (1.5)$$

with survival function

$$\Pr(X > x|\alpha,\gamma) = \min[e^{-\alpha(x-\gamma)},1] \quad (1.6)$$

By setting $\log W = X$ and $\log \beta = \gamma$, we note that a translated exponential variate is the logarithm of Pareto variate. Hence results for the translated exponential can be readily applied to the Pareto distribution since

$$\Pr[X \leq x] = \Pr[W \leq w]. \quad (1.7)$$

The two parameter Pareto distribution has been subjected to a Bayesian analysis regarding the Pareto parameters by Lwin (1972) and Arnold and Press (1983). In recent papers, Geisser (1984a,b) studied the problem of predicting future observables from the Pareto/Exponential distribution. The censored case of a type most encountered in practice was considered. Results were obtained for the probability that a fraction of M future observables survive beyond a certain threshold. In this paper we extend those results to the fraction that fall within a prescribed interval. Previous results which were restricted to the non-censored infinite interval case with a non-informative prior distribution for the parameters, Geisser (1982), and then extended to the censored case with a conjugate prior distribution for the parameters Geisser (1984a,b) are then special cases of this work.

2. The Posterior Density.

Let $X_1,\ldots,X_N$ be realizations from

$$f(x|\alpha,\gamma) = \begin{cases} \alpha e^{-\alpha(x-\gamma)} & \alpha > 0 \quad x > \gamma \\ 0 & \text{otherwise} \end{cases}$$

and

$$\Pr(x_1 \le X \le x_2|\alpha,\gamma) = \begin{cases} 0 & \text{if } \gamma \ge x_2 \\ 1 - e^{-\alpha(x_2-\gamma)} & x_2 > \gamma \ge x_1 \\ e^{-\alpha(x_1-\gamma)} - e^{-\alpha(x_2-\gamma)} & x_2 > x_1 \ge \gamma \end{cases} \tag{2.1}$$

We assume that $x_1,\ldots,x_d$ are the uncensored realizations of $X_1,\ldots X_d$ and that $X_{d+1},\ldots, X_N$ are censored at $x_{d+1},\ldots,x_N$ respectively. Let $\bar{x}_d = d^{-1}(x_1+\ldots+x_d)$ and $m_d = \min(x_1,\ldots,x_d)$, and further order the censored values $x_{d+1},\ldots,x_{N+d}$ as follows:

$x_{(1)} \ge x_{(2)} \ge \ldots \ge x_{(N-d)}$.

Suppose that for some $k \in [0,N - d]$ the sample is such that

$$x_{(k+1)} < m_d < x_{(k)} \tag{2.2}$$

where $x_{(0)} = \infty$ and $x_{(N-d+1)} = -\infty$. Then the likelihood is, for $\alpha > 0$,

$$L(\alpha,\gamma) = \begin{cases} \alpha^d e^{-\alpha d(\bar{x}_d-\gamma)-\alpha k(\bar{x}_{(k)}-\gamma)} & \text{for } x_{(k+1)} < \gamma \le m_d, \\ \alpha^d e^{-\alpha d(\bar{x}_d-\gamma)-\alpha j(\bar{x}_{(j)}-\gamma)} & \text{for } x_{(j+1)} < \gamma \le x_j, \\ 0 & \text{elsewhere} \end{cases} \qquad (2.3)$$

Further, we assume a conjugate prior density for α and γ to be

$$p(\gamma|\alpha) = N_0 \alpha e^{\alpha N_0(\gamma-m_0)} \qquad \text{for } \gamma < m_0 \qquad (2.4)$$

$$p(\alpha) = [N_0(\bar{x}_0-m_0)]^{d_0-1} \alpha^{d_0-2} e^{-\alpha N_0(\bar{x}_0-m_0)}/\Gamma(d_0-1) \qquad (2.5)$$

where $\alpha > 0$, $\bar{x}_0 > m_0$, and $1 < d_0 \le N_0$ to insure that the distributions are proper. This prior is translatable into the conjugate prior for the parameters α and $\beta = e^{\gamma}$ of the Pareto distribution which was used by Lwin (1972) in that context. This is a consequence of the fact that the translated exponential variate is the logarithm of the two parameter Pareto variate, as noted in the previous section. For another formulation of the prior distribution for these Pareto parameters which can also be transformed for use with the translated exponential parameters, see Arnold and Press (1983).

From the above we can calculate

$$p(\gamma) = \frac{(d_0-1)(\bar{x}_0-m_0)^{d_0-1}}{(\bar{x}_0-\gamma)^{d_0}} \qquad (2.6)$$

and

$$p(\alpha|\gamma) = [N_0(\bar{x}_0-\gamma)]^{d_0}\alpha^{d_0-1}e^{-\alpha N_0(\bar{x}_0-\gamma)}/\Gamma(d_0). \qquad (2.7)$$

Let $d^* = d + d_0$, $m^* = \min(m_d, m_0)$ then for $x_{(t+1)} \leq m^* < x_{(t)}$ the posterior density of γ and α is

$$p(\alpha,\gamma) \propto \begin{cases} \alpha^{d^*-1}e^{-\alpha[N_0(\bar{x}_0-\gamma) + d(\bar{x}_d-\gamma) + t(\bar{x}_{(t)}-\gamma)]} & x_{(t+1)}<\gamma\leq m^* \\ \alpha^{d^t-1}e^{-\alpha[N_0(\bar{x}_0-\gamma) + d(\bar{x}_d-\gamma) + j(\bar{x}_{(j)}-\gamma)]} & x_{(j+1)}<\gamma\leq x_{(j)} \\ 0 & \text{elsewhere} \end{cases} \qquad (2.8)$$

where $j = t+1,\ldots,N-d$.

Although there is no intrinsic difficulty in carrying on with this piecewise posterior density, it simplifies what is already a somewhat complex calculation to make the reasonable assumption that $m_d = m = \min(x_1,\ldots,x_N)$ i.e. the minimum of all the observations including the censored ones. Such a situation is most often met in practice, especially in well-controlled experiments where censoring is a device to terminate a time consuming experiment. In certain less well controlled studies, those subjects, who drop out or are lost prior to any fully observed value for reasons unrelated to the experiment itself, will negligibly influence inferences involving future observables. The reason being that there is relatively little information in these censored values because their effect is confined to regions of relatively low density of the parameters.

In the light of this we shall now continue with the assumption $m = m_d$ which implies that

$$m^* \le \min(x_1, \ldots, x_N). \tag{2.9}$$

For $\bar{x}^* = (N_0\bar{x}_0 + N\bar{x})/N^*$, and $N^* = N_0 + N$, the posterior density of α and γ can now be expressed simply as the product of the densities

$$p(\gamma|\underset{\sim}{x}) = \frac{(d^*-1)(\bar{x}^*-m^*)^{d^*-1}}{(\bar{x}^*-\gamma)^{d^*}}, \qquad \gamma < m^*,\ \underset{\sim}{x} = (x_1, \ldots, x_N) \tag{2.8}$$

and

$$p(\alpha|\gamma,\underset{\sim}{x}) = [N^*(\bar{x}^*-\gamma)]^{d^*}\alpha^{d^*-1}e^{-\alpha N^*(\bar{x}^*-\gamma)}/\Gamma(d^*), \qquad \alpha > 0 \tag{2.9}$$

or as the product of

$$p(\gamma|\alpha,\underset{\sim}{x}) = \alpha N^* e^{\alpha N^*(\gamma-m^*)}, \qquad \gamma < m^* \tag{2.10}$$

and

$$p(\alpha|\underset{\sim}{x}) = [N^*(\bar{x}^*-m^*)]^{d^*-1}\alpha^{d^*-2}e^{-\alpha N^*(\bar{x}^*-m^*)}/\Gamma(d^*-1), \quad \alpha > 0. \tag{2.11}$$

3. Interval Prediction.

Suppose we focus our interest on calculating the probability that a single future observation Z, from this process lies in the interval $I = [z_1, z_2]$ then

$$Pr[z_1 \le Z \le z_2 | \underline{x}] = F(z_2) - F(z_1) \tag{3.1}$$

where F(z) is the predictive distribution of Z. The calculation yields

$$Pr[z_1 \le Z \le z_2|\underline{x}] = \begin{cases} \dfrac{(N^*)^{d^*}(\bar{x}^*-m^*)^{d^*-1}}{N^*+1} \displaystyle\sum_{i=1}^{2} \dfrac{(-1)^{i+1}}{[z_i-m^*+N^*(\bar{x}^*-m^*)]^{d^*-1}} & \text{for } z_1>m^* \\ 1 - \dfrac{1}{N^*+1}\left[\left(\dfrac{\bar{x}^*-m^*}{\bar{x}^*-z_1}\right)^{d^*-1} + N^*\left(\dfrac{N^*(\bar{x}^*-m^*)}{N^*(\bar{x}^*-m^*)+z_2-m^*}\right)^{d^*-1}\right] & \text{for } z_2>m^*>z_1 \\ \dfrac{(\bar{x}^*-m^*)^{d^*-1}}{N^*+1}\left[(\bar{x}^*-z_2)^{-(d^*-1)}-(\bar{x}^*-z_1)^{-(d^*-1)}\right] & \text{for } m^*>z_2 \end{cases} \tag{3.2}$$

Consider a set of future values $Z_1, \ldots, Z_M$ from the given translated exponential and the proportion of them that lie in the interval I. Let

$$Y_i = \begin{cases} 1 & \text{if } Z_i \varepsilon I \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

for $i = 1, \ldots, M$, and set

$$\bar{Y} = M^{-1} \sum_{i=1}^{M} Y_i \ .$$

Now let

$$\theta(\alpha,\gamma) = Pr[z_1 \le Z \le z_2|\alpha,\gamma] = \begin{cases} 0 & \text{if } \gamma \ge z_2 \\ 1 - e^{-\alpha(z_2-\gamma)} & \text{if } z_2>\gamma\ge z_1 \\ e^{-\alpha(z_1-\gamma)} - e^{-\alpha(z_2-\gamma)} & \text{if } z_2>z_1\ge\gamma. \end{cases} \tag{3.4}$$

Therefore, for $M\bar{Y} = R$ the number of Z_i's that lie in I,

$$\Pr[\bar{Y} = \frac{r}{M}|I] = \int \binom{M}{r}\theta^r(1-\theta)^{M-r}p(\alpha,\gamma|\underset{\sim}{x})d\alpha d\gamma. \tag{3.5}$$

For $v = \min(m^*,z_2)$, evaluation of the above yields

$$\Pr[\bar{Y}=\frac{r}{M}|I] = \begin{cases} N^*\binom{M}{r}\sum_{j=0}^{r}\binom{r}{j}\frac{(-1)^j[N^*(\bar{x}^*-m^*)]^{d^*-1}}{N^*+M-r+j}\{[N^*(\bar{x}^*-v)+(M-r+j)(z_2-v)]^{-(d^*-1)} \\ \quad -[N^*(\bar{x}^*-z_1)+(M-r+j)(z_2-z_1)]^{-(d^*-1)}\} + N^*\binom{M}{r}\sum_{j=0}^{M-r}\binom{M-r}{j}\frac{(-1)^j}{(N^*+r+j)} \\ \quad \cdot\sum_{k=0}^{r+j}\binom{r+j}{k}(-1)^k\left(\frac{N^*(\bar{x}^*-m^*)}{N^*(\bar{x}^*-z_1)+k(z_2-z_1)}\right)^{d^*-1} \qquad \text{for } v > z_1 \\ \\ N^*\binom{M}{r}\sum_{j=0}^{M-r}\binom{M-r}{j}\frac{(-1)^j}{N^*+r+j}\sum_{k=0}^{r+j}(-1)^k\binom{r+j}{k}\left(1+\frac{k(z_2-z_1)+(r+j)(z_1-m^*)}{N^*(\bar{x}^*-m^*)}\right)^{-(d^*-1)} \\ \qquad \text{for } z_1 > v \end{cases} \tag{3.6}$$

It is also easy to show that irrespective of M

$$E(\bar{Y}) = \Pr[z_1 \le Z \le z_2|\underset{\sim}{x}]$$

where the right hand side was given in (3.2). For the usual non-informative prior density,

$$p(\alpha,\gamma) \propto \alpha^{-1},$$

the results are obtained by letting $d^* \to d$, $N^* \to N$, $\bar{x}^* \to \bar{x}$ and $m^* \to m$.

For the case where γ is known, one can proceed directly or use (2.9) and obtain

$$\Pr[\bar{Y}=\tfrac{r}{M}|I] = \begin{cases} \binom{M}{r}\sum_{j=0}^{M-r}\binom{M-r}{j}(-1)^j\sum_{k=0}^{r+j}(-1)^k\binom{r+j}{k}\left(1+\frac{k(z_2-z_1)+(r+j)(z_1-\gamma)}{N^*\bar{x}^*}\right)^{-d^*} & \text{for } z_1>\gamma \\ 1-\binom{M}{r}\sum_{j=0}^{M-r}\binom{M-r}{j}(-1)^j\left(1+\frac{(z_2-\gamma)(r+j)}{N^*\bar{x}^*}\right)^{-d} & \text{for } z_2>\gamma>z_1 \\ 1 & \text{for } r=0,\ \gamma>z_2 \\ 0 & \text{for } r>0,\ \gamma>z_2 \end{cases} \quad (3.7)$$

A special case of the above for a non-informative prior and $z_2 \to \infty$ was given by Geisser (1982).

From (3.6) the distribution function of $\bar{Y}$ can be calculated i.e.

$$P[\bar{Y} \le \tfrac{r}{M}|I] = \sum_{j=0}^{r}\Pr[\bar{Y} = \tfrac{j}{M}|I]. \quad (3.8)$$

However the asymptotic distribution function of $\theta = \lim_{M\to\infty}\bar{Y}$ appears to be rather difficult to obtain explictedly except in infinite interval cases i.e. where either $z_1 = -\infty$ or $z_2 = \infty$. Results for these special cases are given in the next section.

4. Special Cases - Infinite Intervals.

In (3.6) if we let $z_2 \to \infty$, then we are computing the probability that a given fraction exceeds z_1. The result, which may be obtained from (3.6) is

$$
\Pr(\bar{Y}=\tfrac{r}{M}|z_1) = \begin{cases} \left(\dfrac{\bar{x}^*-m^*}{\bar{x}^*-z_1}\right)^{d^*-1} \binom{N^*+r-1}{r} \Big/ \binom{N^*+M}{M} & r<M,\ z_1<m^* \\ 1 - \dfrac{M}{N^*+M}\left(\dfrac{\bar{x}^*-m^*}{\bar{x}^*-z_1}\right)^{d^*-1} & r=M,\ z_1<m^* \\ N^*\binom{M}{r} \displaystyle\sum_{j=0}^{M-r} \binom{M-r}{j} \dfrac{(-1)^j}{(N^*+r+j)} \left(1 + \dfrac{(r+j)(z_1-m^*)}{N^*(\bar{x}^*-m^*)}\right)^{-(d^*-1)} & m^*<z_1 \end{cases} \quad (4.1)
$$

The asymptotic distribution function of $\bar{Y} \longrightarrow \theta$ is

$$
\Pr[\theta \le \theta | z_1] = \begin{cases} \theta^{N^*\left(\frac{\bar{x}^*-m^*}{\bar{x}^*-z_1}\right)^{d^*-1}} & \text{for } z_1 \le m^* \\ \theta^{N^*\left(\frac{\bar{x}^*-m^*}{|\bar{x}^*-z_1|}\right)^{d^*-1}} G\left(2N^*\left(\frac{|\bar{x}^*-z_1|}{m^*-z_1}\right)\log\theta\right)+1-G\left(2N^*\left(\frac{\bar{x}^*-m^*}{m^*-z_1}\right)\log\theta\right) & \text{for } m^*<z_1 \ne x^* \\ 1 - G^*(-2N^*\log\theta) & \text{for } z_1 = \bar{x}^* \end{cases} \quad (4.2)
$$

for $0 \le \theta < 1$ and

$$
\Pr[\theta = 1|\underset{\sim}{x}] = \begin{cases} 1 - \left(\dfrac{\bar{x}^*-m^*}{\bar{x}^*-z_1}\right)^{d^*-1} & z_1 < m^* \\ 0 & z_1 \le m^*, \end{cases} \quad (4.3)
$$

where $G(u)$ represents the distribution function of a χ^2 variate with $2d^*-2$ degrees of freedom and $G^*(u)$ one with $2d^*$ degrees of freedom. These results appear in Geisser (1984a,b).

For $z_1 \longrightarrow -\infty$ and z_2 finite, we obtain

$$\Pr[\bar{Y}=\tfrac{r}{M}|z_2] = \begin{cases} \left(\dfrac{\bar{x}^*-m^*}{\bar{x}^*-z_2}\right)^{d^*-1} \binom{N^*+M-r-1}{M-r} \Big/ \binom{N^*+M}{M} & r>0,\ z_2<m^* \\ 1 - \dfrac{M}{N^*+M}\left(\dfrac{\bar{x}^*-m^*}{\bar{x}^*-z_2}\right)^{d^*-1} & r=0,\ z_2<m^* \\ N^*\binom{M}{r} \displaystyle\sum_{j=0}^{r} \binom{r}{j} \frac{(-1)^j}{(N^*+M-r+j)} \left(1 + \frac{(M-r+j)(z_2-m^*)}{N^*(\bar{x}^*-m^*)}\right)^{-(d^*-1)} & \end{cases} \tag{4.4}$$

Here the limiting distribution of θ is, for $0 < \theta \le 1$,

$$\Pr[\Theta\le\theta|\underset{\sim}{x}] = \begin{cases} 1 - (1-\theta)^{N^*}\left(\dfrac{\bar{x}^*-m^*}{\bar{x}^*-z_2}\right)^{d^*-1} & z_2 \le m^* \\ G\left(2N^*\left(\dfrac{\bar{x}^*-m^*}{m^*-z_2}\right)\log(1-\theta)\right) - (1-\theta)^{N^*}\left(\dfrac{\bar{x}^*-m^*}{|\bar{x}^*-z_2|}\right)^{d^*-1} G\left(2N^*\left(\dfrac{|\bar{x}^*-z_2|}{m^*-z_2}\right)\log(1-\theta)\right) & m^*<z_2\ne\bar{x}^* \\ G^*\left(-2N^*\log(1-\theta)\right) & z = \bar{x}^* \end{cases} \tag{4.5}$$

and

$$\Pr[\Theta=0|\underset{\sim}{x}] = \begin{cases} 1 - \left(\dfrac{\bar{x}^*-m^*}{\bar{x}^*-z_2}\right)^{d^*-1} & z_2 < m^* \\ 0 & z_2 \ge m^*. \end{cases} \tag{4.6}$$

5. Remarks.

As was noted an explicit expression for the asymptotic distribution of $\bar{Y}$ in the finite interval case seems difficult to achieve although the exact expression (3.7) for finite M is available. Further, the asymptotic distribution as given for the infinite interval in either 4.2-4.3 or 4.5-4.6 often requires M to exceed 100 for it to be an adequate approximation for the exact probability for a finite M, Geisser (1984a). The usefulness of asymptotic expressions then is for those values of M which are so large as to render high speed computers inadequate to the task of precise calculation or when θ itself is actually of interest. For intermediate values of M in the infinite interval case, approximations to the exact distribution of $\bar{Y}$ involving the F distribution have been proposed in Geisser (1984a).

This research was supported in part by NIH grant NIGMS25271.

REFERENCES

Arnold, B.C., 1982, Pareto distributions. Preliminary version of a monograph to be published by the international Co-operative Publishing House.

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

Geisser, S. (1982). Aspects of the predicitive and estimative approaches in the determination of probabilities. Biometrics (supplement), 38, no. 1, 75-93.

__________ (1984a). Predicting Pareto and exponential observables. The Canadian Journal of Statistics, 12, 2 (in press).

__________ (1984b). On the predicting of observables: A selective update. Proceedings of Second International Conference on Bayesian Statistics (with discussion) University Press, Valencia Spain (in press).

Gnedenko, B.B., Belyayev, Y.K., and Solovyev, A.D. (1969). Mathematical Methods of Reliability Theory, New York and London: Academic Press.

Kalbfleisch, J.D., and Prentice, R.L. (1980). The Statistical Analysis of failure time data. John Wiley and Sons.

Lwin, T. (1972). Estimation of the tail of Paretian law, Scand. Actuar. J., 55, 170-178.