UNIVERSIDADE DE SÃO PAULO
CAMPUS DE SÃO CARLOS
INSTITUTO DE CIÊNCIAS MATEMÁTICAS DE SÃO CARLOS
I.C.M.S.C.

An useful reparametrization for the extreme value distribution

ACHCAR, JORGE ALBERTO

nº 58

Notas do ICMSC - USP

ISSN 0103-2577

An useful reparametrization for the extreme value distribution

ACHCAR, JORGE ALBERTO

nº 58

São Carlos (SP)
1990

AN USEFUL REPARAMETRIZATION FOR THE EXTREME VALUE DISTRIBUTION

JORGE ALBERTO ACHCAR
ICMSC-Universidade de São Paulo
Caixa Postal, 668
13560, São Carlos, SP, Brazil

SUMMARY

"The extreme value density is one of the most useful models in reliability. An usual problem is related to the accuracy of the normal approximations for the maximum likelihood estimators of the parameters of the extreme value density, specially for small or moderate sample sizes. In this paper, we present an useful reparametrization that improves the accuracy of the asymptotical results even for small or moderate sample sizes".

Key words: extreme value density, reparametrization, normality of the likelihood function.

1. INTRODUCTION

One of the most useful survival models for reliability data is given by the extreme value distribution for the logarithms of the life times of unities submitted to a reliability experiment. This model corresponds to a Weibull distribution for the survival data.

Considering T a random variable representing the survival time of an unity, we assume the model:

$$y = \ln(T) = \mu + \sigma W \tag{1}$$

where the error term W has an extreme value density $\exp\{w - e^{w}\}$, $-\infty < w < \infty$. In the Weibull form of the model, the random variable T has a Weibull density with parameters $\alpha = e^{\mu}$ and $\beta = 1/\sigma$.

For inferences about the parameters μ and σ, or even functions of these parameters, we usually consider asymptotical results. One of these results is given by the asymptotical normality of the maximum likelihood estimators for μ and σ. A practical problem of great interest for statisticians is related to the accuracy of these asymptotical results considering small or moderate sample sizes, specially when σ is small (that is, $\sigma < 1$).

In this paper, we present a reparametrization of the parameters μ and σ that can improve the normal approximation of the likelihood function for small or moderate sample sizes. Thus, we find an one-to-one transformation $\theta_1(\mu,\sigma)$ and $\theta_2(\mu,\sigma)$ of the parameters μ and σ such that the likelihood function for θ_1 and θ_2 is close to a bivariate normal density. We explore the method proposed by Anscombe (1964) and by Sprott (1973, 1980) considering a reparametrization θ_1 and θ_2 such that the third derivates of the logarithm of the likelihood function for θ_1 and θ_2 locally in the maximum likelihood estimators for θ_1 and θ_2 are close to zero to get approximate normal likelihoods.

2. NORMAL APPROXIMATION FOR THE MAXIMUM LIKELIHOOD ESTIMATORS FOR μ AND σ

Let $T_1, T_2, \ldots, T_n$ be the survival times of a random sample of size n of unities submitted to a reliability experiment and assume that the model (1) is appropriate. Thus, the logarithm of the likelihood function for μ and σ is given by:

$$\ell(\mu,\sigma) = -n\ell n\,\sigma + \frac{1}{\sigma}\sum_{i=1}^{n} y_i - \frac{n\mu}{\sigma} - e^{-\mu/\sigma}\sum_{i=1}^{n} e^{y_i/\sigma} \qquad (2)$$

Considering $W_i = (y_i - \mu)/\sigma$, we have:

$$\frac{\partial^2 \ell}{\partial \mu^2} = -\frac{1}{\sigma^2}\sum_{i=1}^{n} e^{w_i} \qquad (3)$$

$$\frac{\partial^2 \ell}{\partial \mu \partial \sigma} = \frac{n}{\sigma^2} - \frac{1}{\sigma^2}\sum_{i=1}^{n} e^{w_i} - \frac{1}{\sigma^2}\sum_{i=1}^{n} w_i e^{w_i}$$

and

$$\frac{\partial^2 \ell}{\partial \sigma^2} = \frac{n}{\sigma^2} - \frac{2}{\sigma^2}\sum_{i=1}^{n} w_i e^{w_i} + \frac{2}{\sigma^2}\sum_{i=1}^{n} w_i -$$

$$- \frac{1}{\sigma^2}\sum_{i=1}^{n} w_i^2 e^{w_i}$$

The random variable W with the extreme value density $\exp\{w - e^w\}$, $-\infty < w < \infty$, has $E(W) = \psi(1)$ where $\psi(k)$ is the digamma function $d\log\Gamma(k)/dk$, $E(We^W) = \psi(2)$, $E(W^2e^W) =$

$= \psi^{(1)}(2) + |\psi(2)|^2$ where $\psi^{(1)}$ is the trigamma function $d^2 \log \Gamma(k)/dk^2$ and $E(e^w) = 1$ (see for example Lawless, 1982).

Thus, $E(-\partial^2 \ell/\partial\mu^2) = n/\sigma^2$, $E(-\partial^2 \ell/\partial\sigma^2) =$ $= \frac{n}{\sigma^2}(1 + \psi(2) + (\psi(2))^2)$, and $E(-\partial^2 \ell/\partial\mu\partial\sigma) = \frac{n}{\sigma^2}(1+\psi(1))$.

Therefore, the Fisher information matrix for μ and σ is given by,

$$I(\mu,\sigma) = \begin{pmatrix} \frac{n}{\sigma^2} & \frac{n}{\sigma^2}(1+\psi(1)) \\ \frac{n}{\sigma^2}(1+\psi(1)) & \frac{n}{\sigma^2}(1+\psi(2)+(\psi(2))^2) \end{pmatrix} \quad (4)$$

For large values of n, the maximum likelihood estimators $\hat{\mu}$ and $\hat{\sigma}$ have an asymptotical bivariate normal distribution $N\{(\mu,\sigma);\ I^{-1}(\hat{\mu},\hat{\sigma})\}$. In practical work, it is usual to consider the observed information matrix I_0 in place of the Fisher information matrix I, given by,

$$I_0 = \begin{pmatrix} -\left.\frac{\partial^2 \ell}{\partial\mu^2}\right|_{(\hat{\mu},\hat{\sigma})} & -\left.\frac{\partial^2 \ell}{\partial\mu\partial\sigma}\right|_{(\hat{\mu},\hat{\sigma})} \\ -\left.\frac{\partial^2 \ell}{\partial\sigma\partial\mu}\right|_{(\hat{\mu},\hat{\sigma})} & -\left.\frac{\partial^2 \ell}{\partial\sigma^2}\right|_{(\hat{\mu},\hat{\sigma})} \end{pmatrix} \quad (5)$$

From $\partial\ell/\partial\mu = 0$ and $\partial\ell/\partial\sigma = 0$, we have $\sum_{i=1}^{n} e^{\hat{w}_i} = n$ and $\sum_{i=1}^{n} \hat{w}_i e^{\hat{w}_i} = n + \sum_{i=1}^{n} \hat{w}_i$, where $\hat{w}_i = (y_i - \hat{\mu})/\sigma$. Therefore,

locally in $\hat{\mu}$ and $\hat{\sigma}$, we have:

$$\frac{\partial^3 \ell(\hat\mu,\hat\sigma)}{\partial\mu^3} = \frac{n}{\hat\sigma^3}$$

$$\frac{\partial^3 \ell(\hat\mu,\hat\sigma)}{\partial\mu^2\partial\sigma} = \frac{1}{\hat\sigma^3}\{3n + \sum_{i=1}^{n} \hat w_i\} \tag{6}$$

$$\frac{\partial^3 \ell(\hat\mu,\hat\sigma)}{\partial\sigma^3} = \frac{1}{\hat\sigma^3}\{4n + 6\sum_{i=1}^{n} \hat w_i^2 e^{\hat w_i} + \sum_{i=1}^{n} \hat w_i^3 e^{\hat w_i}\}$$

and,

$$\frac{\partial^3 \ell(\hat\mu,\hat\sigma)}{\partial\sigma^2\partial\mu} = \frac{1}{\hat\sigma^3}\{4n + 4\sum_{i=1}^{n} \hat w_i + \sum_{i=1}^{n} \hat w_i^2 e^{\hat w_i}\}.$$

From (6), we observe that the values of the third derivatives of the logarithm of the likelihood function for μ and σ locally in the maximum likelihood estimators $\hat\mu$ and $\hat\sigma$ can be very large specially for small values of σ. This implies that the normal approximation for the likelihood function for μ and σ can be very poor for small or moderate sample sizes.

3. AN USEFUL TRANSFORMATION OF THE PARAMETERS

To search for an one-to-one transformation $\theta_1(\mu,\sigma)$ and $\theta_2(\mu,\sigma)$ of the parameters μ and σ such that the likelihood function for θ_1 and θ_2 is approximately normal, we should check if the third derivates of the logarithm of the likelihood function $\ell(\theta_1,\theta_2)$ are close to zero locally in the maximum likelihood estimators for θ_1 and θ_2. That is, we should verify if $\partial^3\ell/\partial\theta_1^3\big|_{(\hat\theta_1,\hat\theta_2)}\approx 0$, $\partial^3\ell/\partial\theta_1^2\partial\theta_2\big|_{(\hat\theta_1,\hat\theta_2)}\approx 0$, $\partial^3\ell/\partial\theta_2^2\partial\theta_1\big|_{(\hat\theta_1,\hat\theta_2)}\approx 0$,

and $\partial^3\ell/\partial\theta^3\big|_{(\hat\theta_1,\hat\theta_2)}\approx 0$, where $\hat\theta_1$ and $\hat\theta_2$ are the maximum likelihood estimators for θ_1 and θ_2 (see Anscombe, 1964).

Alternatively, we could search for an appropriate reparametrization θ_1 and θ_2 to get approximate normality of the likelihood function if the expected values of the second derivatives of the logarithm of the likelihood function are constants (see Sprott, 1973, 1980). In the Bayesian context, this parametrization θ_1 and θ_2 gives a Jeffreys prior density for θ_1 and θ_2 locally uniform, that is,

$$\pi(\theta_1,\theta_2) \propto \{\det I(\theta_1,\theta_2)\}^{1/2} \tag{7}$$

$$\propto \text{Constant}$$

Where $I(\theta_1,\theta_2)$ is the Fisher information matrix for θ_1 and θ_2 (see for example, Box and Tiao, 1973).

If θ_1 and θ_2 is an one-to-one transformation of the parameters μ and σ, such that the Jeffreys prior density for θ_1 and θ_2 is locally uniform, we verify that, the prior density for μ and σ satisfies,

$$\pi(\mu,\sigma) \propto \pi(\theta_1,\theta_2)\,\{\det \begin{vmatrix} \frac{\partial\theta_1}{\partial\mu} & \frac{\partial\theta_1}{\partial\sigma} \\ \frac{\partial\theta_2}{\partial\mu} & \frac{\partial\theta_2}{\partial\sigma} \end{vmatrix}\}$$

$$\propto \text{constant}\ \{\frac{\partial\theta_1}{\partial\mu}\frac{\partial\theta_2}{\partial\sigma} - \frac{\partial\theta_1}{\partial\sigma}\frac{\partial\theta_2}{\partial\mu}\} \tag{8}$$

Since the Jeffreys prior for μ and σ is proportional to σ^{-2}, $-\infty < \mu < \infty$ and $\sigma > 0$, we verify that the parametrization θ_1 and θ_2 such that $E\{-\partial^2\ell/\partial\theta_1^2\} = c_1$, $E\{-\partial^2\ell/\partial\theta_2^2\} = c_2$ and $E\{-\partial^2\ell/\partial\theta_1\partial\theta_2\} = c_3$, where c_1, c_2 and c_3 are constants, satisfies the differential equation,

$$\left\{ \frac{\partial\theta_1}{\partial\mu}\frac{\partial\theta_2}{\partial\sigma} - \frac{\partial\theta_1}{\partial\sigma}\frac{\partial\theta_2}{\partial\mu} \right\} = \frac{\text{Constant}}{\sigma^2} \tag{9}$$

Therefore, we should search for an appropriate one-to-one transformation θ_1 and θ_2 of μ and σ among many possible solutions of the differential equation (9) such that the expected values of the second derivatives of $\ell(\theta_1,\theta_2)$ are constants, or such that the third derivatives of $\ell(\theta_1,\theta_2)$ locally in $\hat{\theta}_1$ and $\hat{\theta}_2$ are close to zero.

For example, considering $\theta_1 = \ell n\ \sigma$, we verify from (9), that θ_2 is given by the differential equation,

$$\frac{1}{\sigma}\frac{\partial\theta_2}{\partial\mu} = \frac{\text{Constant}}{\sigma^2} \tag{10}$$

A solution for this differential equation (10) is given by,

$$\theta_2 = \int \frac{d\mu}{\sigma} + c(\sigma)$$

$$= \frac{\mu}{\sigma} + c(\sigma) \tag{11}$$

where $c(\sigma)$ is an arbitrary function of σ.

Considering $c(\sigma) = k/\sigma$, where k is a constant, we have the reparametrization $\theta_1 = \ell n\ \sigma$ and $\theta_2 = (\mu+k)/\sigma$. That is, $\sigma = e^{\theta_1}$ and $\mu = \theta_2 e^{\theta_1} - k$. The logarithm of the likelihood function for θ_1 and θ_2 is given by:

$$\ell(\theta_1,\theta_2) = -n\theta_1 + \sum_{i=1}^{n} w_i - \sum_{i=1}^{n} e^{w_i} \tag{12}$$

where $w_i = y_i e^{-\theta_1} - \theta_2 + k e^{-\theta_1}$.

The third derivatives of the logarithm of the likelihood function for θ_1 and θ_2 (12) are given by:

$$\frac{\partial^3 \ell}{\partial \theta_1^3} = -n\theta_2 - \sum_{i=1}^{n} w_i + (3\theta_2^2 + 6\theta_2 + 1) \sum_{i=1}^{n} w_i e^{w_i} + \theta_2 (\theta_2^2 + 3\theta_2 + 1) \sum_{i=1}^{n} e^{w_i} + 3(\theta_2 + 1) \sum_{i=1}^{n} w_i^2 e^{w_i} + \sum_{i=1}^{n} w_i^3 e^{w_i}$$

$$\frac{\partial^3 \ell}{\partial \theta_2^3} = \sum_{i=1}^{n} e^{w_i} \tag{13}$$

$$\frac{\partial^3 \ell}{\partial \theta_1^2 \, \partial \theta_2} = \sum_{i=1}^{n} w_i^2 e^{w_i} + (2\theta_2 + 1) \sum_{i=1}^{n} w_i e^{w_i} + \theta_2 (\theta_2 + 1) \sum_{i=1} e^{w_i}$$

and $$\frac{\partial^3 \ell}{\partial \theta_2^2 \, \partial \theta_1} = \sum_{i=1}^{n} w_i e^{w_i} + \theta_2 \sum_{i=1}^{n} e^{w_i}$$

Considering $k = \hat{\mu}$ in (13) we have $\hat{\theta}_2 = 0$. In the data parametrization $\theta_1 = \ell n \sigma$ and $\theta_2 = (\mu - \hat{\mu})/\sigma$, we observe from $\partial \ell / \partial \theta_1 = 0$ and $\partial \ell / \partial \theta_2 = 0$ that $\sum_{i=1}^{n} \hat{w}_i e^{\hat{w}_i} = n + \sum_{i=1}^{n} \hat{w}_i$ and $\sum_{i=1}^{n} e^{\hat{w}_i} = n$, where $\hat{w}_i = y_i e^{\hat{\theta}_i} + \hat{\mu} e^{-\hat{\theta}_1}$. Therefore, locally in $\hat{\theta}_1$ and $\hat{\theta}_2$, we have:

$$\frac{\partial^3 \ell(\hat{\theta}_1,\hat{\theta}_2)}{\partial \theta_1^3} = n + 3\sum_{i=1}^{n} \hat{w}_i^2 e^{\hat{w}_i} + \sum_{i=1}^{n} \hat{w}_i^3 e^{\hat{w}_i}$$

$$\frac{\partial^3 \ell(\hat{\theta}_1,\hat{\theta}_2)}{\partial \theta_2^3} = n \tag{14}$$

$$\frac{\partial^3 \ell(\hat{\theta}_1,\hat{\theta}_2)}{\partial \theta_1^2\,\partial \theta_2} = n + \sum_{i=1}^{n} \hat{w}_i + \sum_{i=1}^{n} \hat{w}_i^2 e^{\hat{w}_i}$$

and
$$\frac{\partial^3 \ell(\hat{\theta}_1,\hat{\theta}_2)}{\partial \theta_2^2\,\partial \theta_1} = n + \sum_{i=1}^{n} \hat{w}_i$$

Comparing (6) with (14), we observe that the data parametrization $\theta_1 = \ell n\ \sigma$ and $\theta_2 = (\mu - \hat{\mu})/\sigma$ usually gives better normal approximation for the likelihood function, specially for small values of σ, since the third derivatives of $\ell(\theta_1,\theta_2)$ locally in the maximum likelihood estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ will be very small. We also observe that locally in the maximum likelihood estimators $\hat{\theta}_1 = \ell n\hat{\sigma}$ and $\hat{\theta}_2 = 0$, the expected values of the second derivatives of $\ell(\theta_1,\theta_2)$ are constants. This parametrization can be very usefull if the parameter of interest is the scale parameter σ.

4. AN EXAMPLE

In table 1, we have the voltage levels at which failures occurred when specimens were subjected to an increasing voltage stress in a laboratory experiment. The test involved 20 specimens and the failure voltages T_i were given in Kilovolts per millimiter (data in Lawless, 1982, page 189).

32.0	35.4	36.2	39.8	41.2
43.3	45.5	46.0	46.2	46.4
46.5	46.8	47.3	47.3	47.6
49.2	50.4	50.9	52.4	56.3

Table 1. Voltage Levels in Kilovolts/Millimiter

Assuming that the model (1) is appropriate, the maximim likelihood estimators for μ and σ are given by $\hat{\mu} = 3.86663$ and $\hat{\sigma} = 0.10658$. We observe that $n = 20$,

$$\sum_{i=1}^{20} \hat{w}_i = -11.517 , \qquad \sum_{i=1}^{20} e^{\hat{w}_i} = 20.00,$$

$$\sum_{i=1}^{20} \hat{w}_i e^{\hat{w}_i} = 8.4807, \qquad \sum_{i=1}^{20} \hat{w}_i^2 e^{\hat{w}_i} = 17.084 \quad \text{and}$$

$$\sum_{i=1}^{20} \hat{w}_i^3 e^{\hat{w}_i} = 13.2369 \quad \text{where} \quad \hat{w}_i = (y_i - \hat{\mu})/\hat{\sigma}.$$

Considering the observed information matrix (5), the maximum likelihood estimators for μ and σ have an asymptotic bivariate normal distribution given by,

$$(\hat{\mu}, \hat{\sigma}) \overset{a}{\sim} N\left((\mu,\sigma) ; \begin{pmatrix} 0.000629 & -0.000144 \\ -0.000144 & 0.000339 \end{pmatrix}\right) \qquad (15)$$

The third derivatives of the logarithm of the likelihood function $\ell(\mu,\sigma)$ given in (6), locally in the maximum likelihood estimators $\hat{\mu}$ and $\hat{\sigma}$ are given by $\partial^3\ell(\hat{\mu},\hat{\sigma})/\partial\mu^3 = 16519.73$,

$\partial^3\ell(\hat{\mu},\hat{\sigma})/\partial\mu^2\partial\sigma = 40044.39$,

$\partial^3\ell(\hat{\mu},\hat{\sigma})/\partial\sigma^3 = 161667.91$ and

$\partial^3\ell(\hat{\mu},\hat{\sigma})/\partial\sigma^2\partial\mu = 42130.92$. We observe large values of the third derivatives of $\ell(\mu,\sigma)$ locally in $\hat{\mu}$ and $\hat{\sigma}$, indicating that the normality of the likelihood function for μ and σ is not adequate. In figure 1, we have contour plots for the exact likelihood function considering $\ell(\mu,\sigma) = -1$ and for the normal approximation (15). We observe that in this parametrization μ and σ , the normal approximation for the likelihood function is not good, and we can get bad inferences, specially for the scale parameter σ.

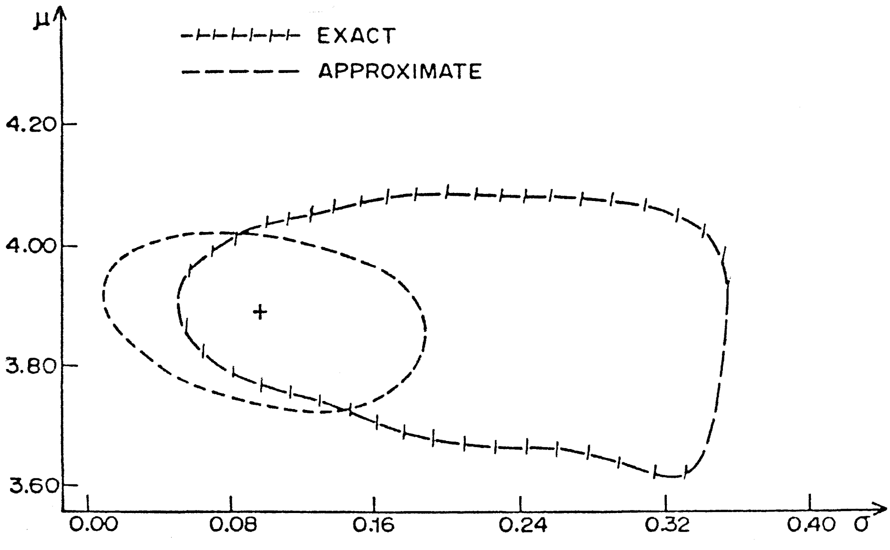

FIGURE 1 - CONTOURS OF THE LIKELIHOOD FUNCTION FOR μ AND σ (ℓ(μ,σ)=-1)

Considering the data dependent reparametrization $\theta_1 = \ell n \sigma$ and $\theta_2 = (\mu - 3.90)/\sigma$, the maximum likelihood estimators for θ_1 and θ_2 are given by $\hat{\theta}_1 = -2.23886$ and $\hat{\theta}_2 = 0$. From (14), we have $\partial^3 \ell(\hat{\theta}_1, \hat{\theta}_2)/\partial\theta_2^3 = 20.00$, $\partial^3 \ell(\hat{\theta}_1, \hat{\theta}_2)/\partial\theta_2^2 \partial\theta_1 = 8.4807$, $\partial^3 \ell(\hat{\theta}_1, \hat{\theta}_2)/\partial\theta_1^3 = 84.4866$ and $\partial^3 \ell(\hat{\theta}_1, \hat{\theta}_2)/\partial\theta_1^2 \partial\theta_2 = 25.5647$.

In this parametrization, we observe small values for the third derivatives of the logarithm of the likelihood function for θ_1 and θ_2 locally in $\hat{\theta}_1$ and $\hat{\theta}_2$, indicating better normality for $\ell(\theta_1, \theta_2)$.

With the observed information matrix for θ_1 and θ_2, the maximum likelihood estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ have an approximate bivariate normal distribution given by,

$$(\hat{\theta}_1, \hat{\theta}_2) \overset{a}{\sim} N\left((\theta_1, \theta_2) ; \begin{pmatrix} 0.0300 & -0.0127 \\ -0.0127 & 0.0554 \end{pmatrix}\right) \qquad (16)$$

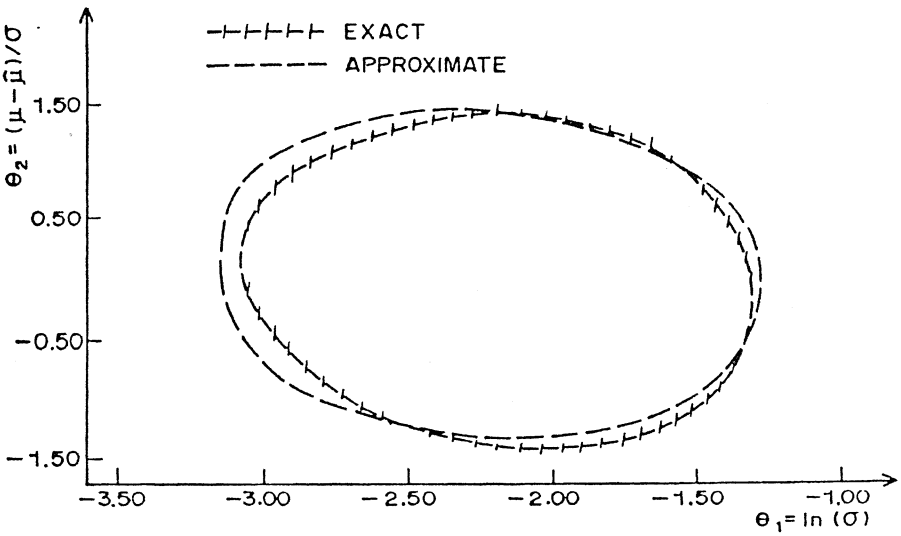

FIGURE 2 - CONTOURS OF THE LIKELIHOOD FUNCTION FOR θ_1 AND θ_2 ($\ell(\theta_1, \theta_2) = -1$)

Considering $\ell(\theta_1,\theta_2) = -1$, we have in figure 2 the contour plots for the exact likelihood function for θ_1 and θ_2, and for the normal approximation (16). We observe very good normal approximation for the likelihood function for θ_1 and θ_2.

5. AN EXAMPLE WITH CENSORED DATA

In table 2, we have a type II censored data set where the observations y_i are given in extreme value form, with $r = 28$ uncensored observations and $n = 40$ (data in Lawless, 1982, page 152).

-2.982	-2.849	-2.546	-2.350	-1.983
-1.492	-1.443	-1.394	-1.386	-1.269
-1.195	-1.174	-0.845	-0.620	-0.576
-0.548	-0.247	-0.195	-0.056	-0.013
0.006	0.033	0.037	0.046	0.084
0.221	0.245	0.296	0.296+	0.296+
0.296+	0.296+	0.296+	0.296+	0.296+
0.296+	0.296+	0.296+	0.296+	0.296+

Table 2. Type II Censored Data (Observations with '+' are Censored Observations)

The maximum likelihood estimators for μ and σ are given by $\hat{\mu} = 0.1563$ and $\hat{\sigma} = 0.9104$. The logarithm of the likelihood function for μ and σ is given by:

$$\ell(\mu,\sigma) = -r\,\ell n\sigma + \sum_{i\varepsilon D} w_i - \sum_{i=1}^{n} e^{w_i} \qquad (17)$$

where $w_i = (y_i - \mu)/\sigma$ and D is the set of unities whose lifetime are uncensored.

The maximum likelihood estimators $\hat{\mu}$ and $\hat{\sigma}$ have an asymptotical normal distribution $N\{(\mu,\sigma);\ I_o^{-1}\}$ where the observed information matrix I_o is given by:

$$I_o = \frac{1}{\hat{\sigma}^2} \begin{pmatrix} r & \sum_{i=1}^{n} \hat{w}_i e^{\hat{w}_i} \\ \sum_{i=1}^{n} \hat{w}_i e^{\hat{w}_i} & r + \sum_{i=1}^{n} \hat{w}_i e^{\hat{w}_i} \end{pmatrix} \qquad (18)$$

where r is the number of uncensored observations and $\hat{w}_i = (y_i - \hat{\mu})\hat{\sigma}$ (see Lawless, 1982).

With the data of table 2, we have:

$$(\hat{\mu},\hat{\sigma}) \overset{a}{\sim} N\{(\mu,\sigma);\ \begin{pmatrix} 0.032889 & 0.003101 \\ 0.003101 & 0.025654 \end{pmatrix}\} \qquad (19)$$

Considering the reparametrization $\theta_1 = \ln\sigma$ and $\theta_2 = (\mu - 0.16)/\sigma$, the logarithm of the likelihood function for θ_1 and θ_2 is given by:

$$\ell(\theta_1,\theta_2) = -r\,\theta_1 + \sum_{i\in D} w_i - \sum_{i=1}^{n} e^{w_i} \qquad (20)$$

The second derivatives of $\ell(\theta_1,\theta_2)$ locally in $\hat{\theta}_1 = \ln\hat{\sigma}$ and $\hat{\theta}_2 = 0$ are given by:

$$\left.\frac{\partial^2 \ell}{\partial\theta_1^2}\right|_{(\hat{\theta}_1,\hat{\theta}_2)} = \sum_{i\in D} \hat{w}_i - \sum_{i=1}^{n} \hat{w}_i^2 e^{\hat{w}_i} - \sum_{i=1}^{n} \hat{w}_i e^{\hat{w}_i}, \quad \left.\frac{\partial^2 \ell}{\partial\theta_2^2}\right|_{(\hat{\theta}_1,\hat{\theta}_2)} =$$

$$= - \sum_{i=1}^{n} e^{\hat{w}_i} \tag{21}$$

and $\left.\dfrac{\partial^2 \ell}{\partial\theta_1 \partial\theta_2}\right|_{(\hat{\theta}_1,\hat{\theta}_2)} = - \sum_{i=1}^{n} \hat{w}_i e^{\hat{w}_i}$

Since $\sum_{i\varepsilon D} \hat{w}_i = -31.383$, $\sum_{i=1}^{n} e^{\hat{w}_i} = 28.001$, $\sum_{i=1}^{n} \hat{w}_i e^{\hat{w}_i} = -3.3837$

and $\sum_{i=1} \hat{w}_i^2 e^{\hat{w}_i} = 7.8966$, we have

$\partial^2 \ell(\hat{\theta}_1,\hat{\theta}_2)/\partial\theta_1^2 = -35.8959$,

$\partial^2 \ell(\hat{\theta}_1,\hat{\theta}_2)/\partial\theta_2^2 = -28.001$ and

$\partial^2 \ell(\hat{\theta}_1,\hat{\theta}_2)/\partial\theta_1\partial\theta_2 = 3.3837$.

In the parametrization $\theta_1 = \ell n\sigma$ and $\theta_2 = (\mu - 0.16)/\sigma$, the maximum likelihood estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ have an asymptotical normal distribution $N\{(\hat{\theta}_1,\hat{\theta}_2);\ I_o^{-1}(\hat{\theta}_1,\hat{\theta}_2)\}$, where,

$$I_o^{-1}(\hat{\theta}_1, \hat{\theta}_2) = \begin{vmatrix} 0.02818 & 0.003405 \\ 0.003405 & 0.036612 \end{vmatrix} \tag{22}$$

Considering σ the parameter of interest, an approximate 90% confidence interval for θ_1 is given by (-0.37085; 0.18311). Therefore, an approximate 90% confidence interval for $\sigma = e^{\theta_1}$ is given by (0.69014; 1.20095). We observe that this confidence interval is very close to the 90% confidence interval for σ

obtained with the likelihood ratio method that is invariant to parametrization (see table 3).

CONFIDENCE INTERVALS FOR σ	
ASYMPTOTICAL NORMALITY OF $\hat{\mu}$ AND $\hat{\sigma}$ (OBSERVED INFORMATION)	(0.65; 1.17)
ASYMPTOTICAL NORMALITY OF $\hat{\theta}_1$ AND $\hat{\theta}_2$ (OBSERVED INFORMATION)	(0.69; 1.20)
LIKELIHOOD RATIO METHOD	(0.70; 1.22)
EXACT METHOD	(0.72; 1.28)

Table 3. 90% Confidence Intervals for σ Obtained by Four Methods

6. CONCLUSIONS

Since the extreme value distribution is a very usefull model in reliability with censored or uncensored data, the proposed reparametrization $\theta_1 = \ell n\ \sigma$ and $\theta_2 = (\mu-\hat{\mu})/\sigma$ can be of great practical interest, specially if the parameter of interest is the scale parameter σ, since the use of the asymptotical normality of the maximum likelihood estimators can simplify the determination of confidence intervals for σ or observed significance levels.

REFERENCES

ANSCOMBE,F.J. (1964). Normal Likelihood Functions, Ann. Inst. Stat. Math. 16, 1-19.

BOX,G.E.P.; TIAO,G.C. (1973). Bayesian Inference in Statistical Analysis, Addison-Wesley.

LAWLESS,J.F. (1982). Statistical Models and Methods for Lifetime Data, John Wiley & Sons.

SPROTT,D.A. (1973). Normal Likelihoods and Relation to a Large Sample Theory of Estimation, Biometrika, 60, 457-465.

SPROTT,D.A. (1980). Maximum Likelihood in Small Samples: Estimation in The Presence of Nuisance Parameters, Biometrika, 67, 515-523.