J. R. Statist. Soc. B (1994)
56, *No.* 3, *pp.* 491–500

Estimation and Tests of Fit for the Three-parameter Weibull Distribution

By RICHARD A. LOCKHART and MICHAEL A. STEPHENS†

Simon Fraser University, Burnaby, Canada

[Received April 1989. Final revision May 1993]

SUMMARY

Estimation techniques are given for the three-parameter Weibull distribution, with all parameters unknown. Tables are given for the empirical distribution function statistics W^2, U^2 and A^2, for testing for the distribution.

Keywords: EMPIRICAL DISTRIBUTION FUNCTION; EMPIRICAL DISTRIBUTION FUNCTION TESTS; GOODNESS OF FIT; RELIABILITY; SURVIVAL ANALYSIS

1. INTRODUCTION

In this paper estimation procedures and tests of fit (based on the empirical distribution function (EDF)) are given for the three-parameter Weibull distribution:

$$F(x;\ \alpha,\ \beta,\ m) = 1 - \exp[-\{(x-\alpha)/\beta\}^m], \qquad x > \alpha, \tag{1}$$

where β and m are positive constants. When α is known, the distribution is called the two-parameter Weibull distribution, and estimation procedures and goodness-of-fit tests are then very straightforward; see, for example, sections 4.4 and 4.11 of Stephens (1986). Here we concentrate on tests for use when all three parameters are unknown and must be estimated from the sample.

It is worthwhile to observe that the three-parameter Weibull distribution is a member of a wider class, the generalized extreme value (or Jenkinson) distribution. This distribution is

$$F^*(x;\ a,\ b,\ c) = 1 - \exp\left[-\left\{1+c\left(\frac{x-a}{b}\right)\right\}^{1/c}\right], \qquad x > a. \tag{2}$$

The parameter b must be positive, whereas a and c may be any real numbers. The three-parameter Weibull distribution is the subfamily of F^* with $c > 0$. The special case $c=0$ is the usual extreme value distribution

$$F^*(x;\ a,\ b,\ 0) = 1 - \exp[-\exp\{(x-a)/b\}], \qquad -\infty < x < \infty; \tag{3}$$

it arises as the limit of the three-parameter Weibull family (1) as $m \to \infty$, $\alpha \to -\infty$ and $\beta \to \infty$.

The goodness-of-fit procedures depend on first estimating the parameters in distribution (1) by an efficient method, such as maximum likelihood (ML). However, it is well known that, when α is unknown, there are problems with ML estimation—for example, for $m < 1$ or for m unknown (as here), the likelihood can be made infinite. There are many papers on this problem (for example,

†*Address for correspondence*: Department of Mathematics and Statistics, Simon Fraser University, Burnaby, British Columbia, V5A 1S6, Canada.

 0035–9246/94/56491

Smith (1985), Smith and Weissman (1985) and Cheng and Iles (1987)). Smith and Naylor (1987) compare Bayesian and ML estimators in a case study. It is also possible, for some data sets, that there is no local maximum for the likelihood. If one were willing to fit the larger family (2), ML estimates could be found, but c will be negative, and the resulting fitted distribution will not be Weibull. If a Weibull fit is nevertheless required, albeit a limiting Weibull, then c should be taken as 0 (the nearest non-negative c), and the extreme value distribution (3) should be fitted. We discuss how to recognize this case, called case C below, in Section 2.

2. ESTIMATION PROCEDURES

The estimation procedures will depend on the *profile likelihood*. For a given sample $x_1, x_2, \ldots, x_n$, the likelihood is

$$L(\alpha, \beta, m) = \prod_{i=1}^{n} \frac{m}{\beta}\left(\frac{x_i-\alpha}{\beta}\right)^{m-1} \exp\left\{-\left(\frac{x_i-\alpha}{\beta}\right)^{m}\right\}. \tag{4}$$

The profile likelihood $L^*(\alpha_t)$, abbreviated L^*, is $L(\alpha, \beta, m)$ maximized, for given α_t, with respect to β and m. Suppose that $Z(\alpha_t)$, abbreviated Z, is $\log L^*(\alpha_t)$. A plot of Z against α_t can take any of three possible forms, similar to plots given in Smith and Weissman (1985), related to estimation of α using only k lower order statistics. Here we always assume a complete sample of size n. In one of these plots, case A say, there is a local minimum for α_t close to $x_{(1)}$, which gives a saddlepoint for the likelihood, and a local maximum for α_t further from $x_{(1)}$, giving the desired ML solution. In case B, a local minimum occurs, but no local maximum, and in case C there are no turning points—the likelihood steadily decreases as α_t moves away from $x_{(1)}$ towards $-\infty$. These three cases were noted by Rockette *et al.* (1974), who conjectured that they exhaust the possibilities; our own extensive Monte Carlo studies confirm this conjecture.

The likelihood equations, obtained by setting to 0 the partial derivatives of the log-likelihood with respect to α, β and m, will give solutions corresponding to the maxima and minima in the three cases. By eliminating β, we obtain two equations in α and m, which can be written

$$\frac{1}{m} - \frac{\sum (x_i-\alpha)^m \log(x_i-\alpha)}{\sum (x_i-\alpha)^m} + \frac{\sum \log(x_i-\alpha)}{n} = 0, \tag{5}$$

$$\frac{m-1}{m}\sum (x_i-\alpha)^{-1} - n\frac{\sum (x_i-\alpha)^{m-1}}{\sum (x_i-\alpha)^m} = 0. \tag{6}$$

When these are solved for $\hat{\alpha}$ and $\hat{m}$, the estimate of β is given by

$$\hat{\beta} = \left\{\sum (x_i-\hat{\alpha})^{\hat{m}} \Big/ n\right\}^{1/\hat{m}}. \tag{7}$$

Again it is useful to fix α_t and to plot the solutions m of equation (5) and m^* of equation (6) against α_t. As $\alpha_t \to x_{(1)}$ from below, it is easily shown that $m^* \to 1$ and $m \to 0$. The graph of m then rises steeply as α_t becomes more negative. If the graphs of m and m^* cross, we have either the minimum or the maximum in case

A, or the minimum in case B; if they do not cross, we have case C. Fig. 1 illustrates these three situations; the data for Figs 1(a) and 1(b) are from the examples in Section 4, and the data for the third, less likely, case have been artificially constructed.

2.1. *Case C*

Given a data set, it will be advantageous to decide at once whether the situation is case C without trying to solve the likelihood equations. This decision may be made straightforwardly. We have shown (Lockhart and Stephens, 1992) that, as $\alpha \rightarrow -\infty$, the plots of m and m^* have parallel asymptotes. Then suppose that the limiting difference is $\Delta = \lim_{\alpha_t \rightarrow -\infty}(m^* - m)$. The value of Δ is found as follows. Let $\bar{x} = \Sigma x_i/n$, and let $s = \Sigma x_i^2/n$; also define $T_r = \Sigma(x_i)^r \exp(-\gamma x_i)$; in these expressions the sums are over $i = 1, 2, \ldots, n$. The quantity γ is the solution of

$$\frac{1}{\gamma} = \bar{x} - \frac{T_1}{T_0}. \tag{8}$$

The value of γ can easily be found by iteration, starting, for example, with $\gamma = 1$ in the right-hand side quantities T_0 and T_1. Define

$$D = \bar{x}T_0 + \gamma(T_2 - \bar{x}T_1); \tag{9}$$

γ is the limiting slope of the lines, and Δ is given by

$$\Delta = \frac{\bar{x}T_0 - \gamma(sT_0 - T_2)/2}{D}. \tag{10}$$

A negative value of Δ means that we have case C, in which case the Weibull fit should be abandoned in favour of the extreme value fit (3).

Another method of discriminating between the various situations is to consider a plot of $L_1(c)$, the profile likelihood of the generalized extreme value distribution, against the parameter c. If there is a local maximum corresponding to a negative c, the derivative of $L_1(c)$ at $c = 0$ must be negative. Cheng and Iles (1990) give a discriminant based on this derivative, and our Δ is equivalent to their L.

2.2. *Cases A and B: Solutions for Estimates*

We now turn to cases A and B, Figs 1(a) and 1(b). To distinguish these cases, it is recommended to establish whether or not a saddlepoint exists; if so, we have case A, and an ML estimate. The saddlepoint is usually very close to $x_{(1)}$. Thus it can be detected by starting with $\alpha_t = x_{(1)} - \epsilon$, where ϵ must be sufficiently small that $m^* - m$ is positive, and then decreasing α_t in very small steps until the saddlepoint is passed ($m^* - m$ becomes negative). The steps in α_t can then be increased until once again $m^* - m = 0$, when the values of α_t and m are the ML solutions. If no saddlepoint exists, $m^* - m$ will pass through a minimum positive value and then start to increase again. We then have case B, and, formally, ML gives $x_{(1)}$ as the estimate for α. This estimate is clearly biased, and there are many publications on improvements to the estimate and also on methods for obtaining estimates and confidence intervals for a wider class of distributions with a threshold parameter. See, for example, Smith and Weissman (1985) and Weissman (1982), and their references.

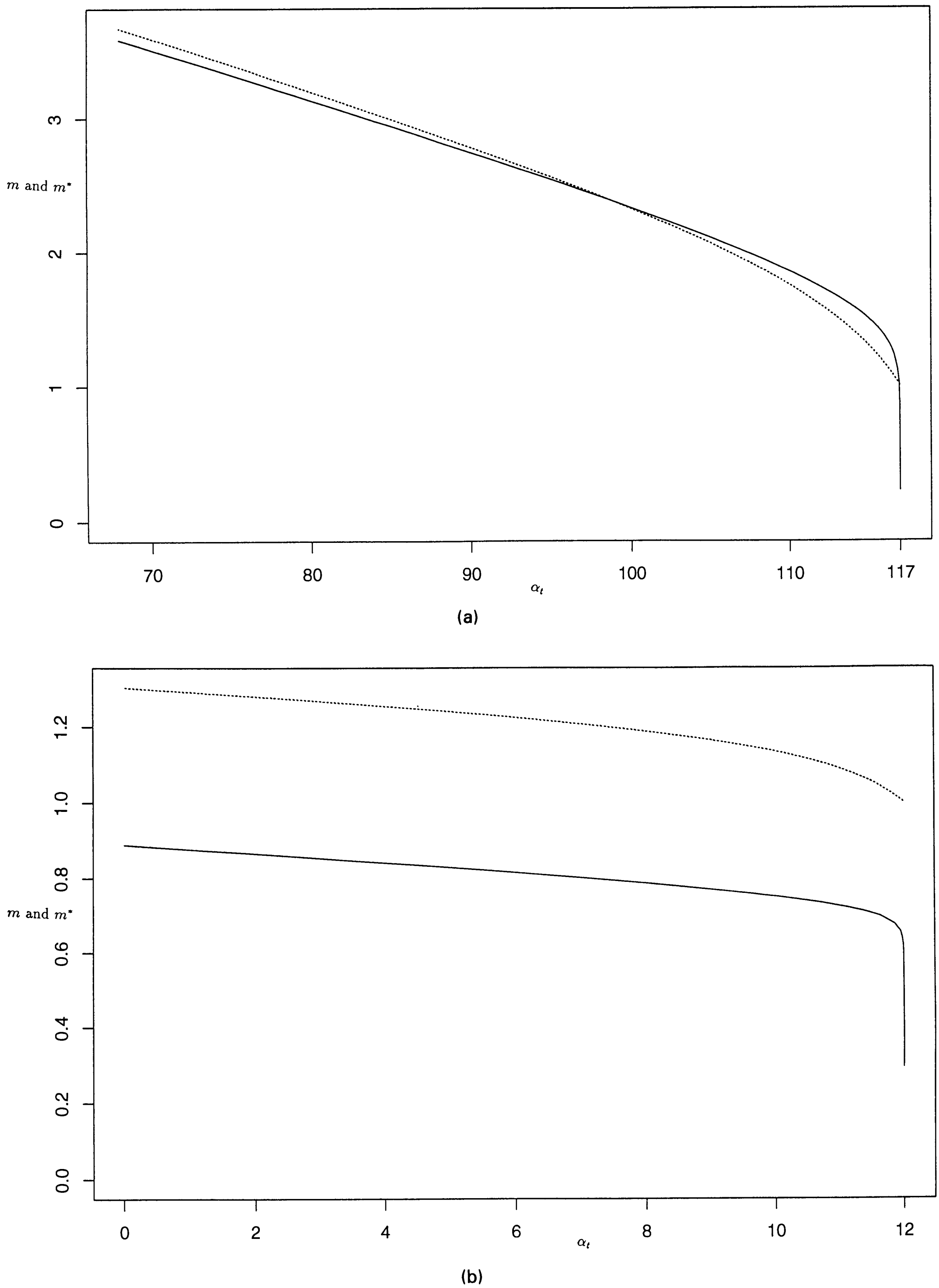
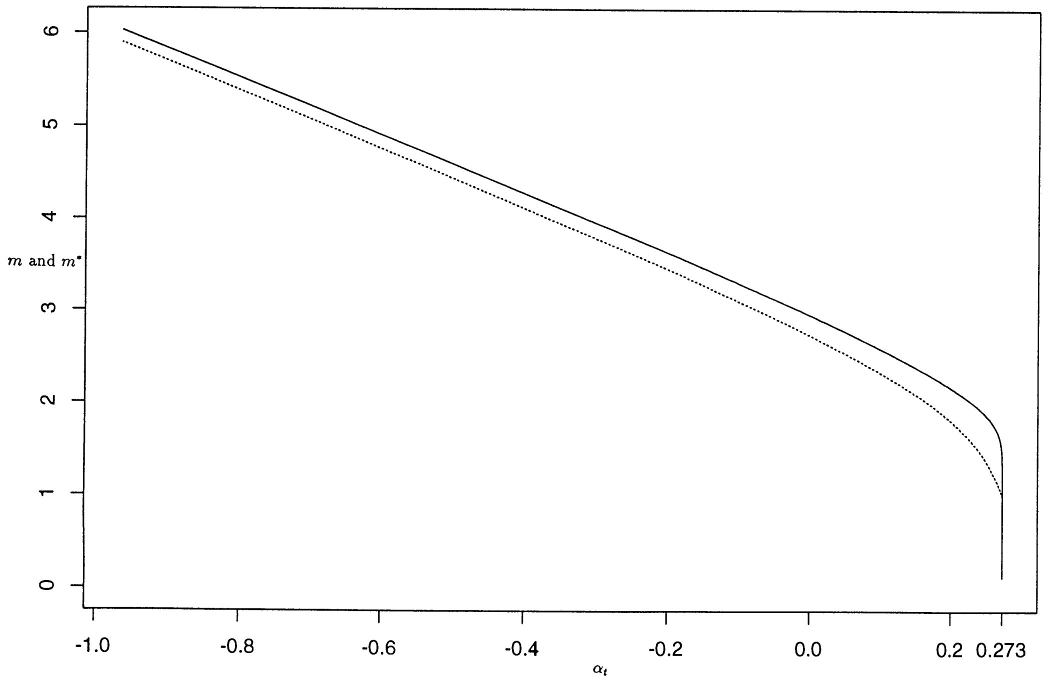


Fig. 1. m and m^* for (a) data set 1 (from Cox and Oakes (1984)), (b) data set 2 (from Proschan (1963)) and (c) the generated data set: ———, m from equation (5); ········, m^* from equation (6)



(c)

Fig. 1 (*continued*)

We give an iterative bias reduction procedure, which, although based on previous ideas, appears itself to be new. The bias in the estimate $\hat{\alpha}=x_{(1)}$ is approximately β/n^c, where $c=1/m$ (Smith and Weissman, 1985); since β and c are not known, they must be estimated and then used to reduce the bias in the estimate $x_{(1)}$ of α. For this we need the likelihood equation $\partial(\log L)/\partial m=0$, namely

$$\frac{n}{m}+\sum \log(x_i-\alpha)-\frac{\sum (x_i-\alpha)^m \log(x_i-\alpha)}{\beta^m}-\left\{n-\frac{\sum (x_i-\alpha)^m}{\beta^m}\right\}\log\beta=0. \quad (11)$$

Suppose that α_r, β_r and m_r are estimates (we omit the circumflex symbol) at iteration r; find estimates α_{r+1}, β_{r+1} and m_{r+1} as follows.

(a) Let $\alpha_{r+1}=x_{(1)}-\beta_r/n^k$, where $k=1/m_r$.
(b) Then solve equation (11) for m_{r+1}, using $\alpha=\alpha_{r+1}$ and $\beta=\beta_r$.
(c) Use equation (7) to give β_{r+1}, using $\alpha=\alpha_{r+1}$ and $m=m_{r+1}$.

Iteration of these three steps continues until the required accuracy for $\hat{m}$ is obtained. Initial estimates m_0 and β_0 may be found by setting $\alpha_0=x_{(1)}$ and continuing with steps (b) and (c) above, but using only the $n-1$ sample values $x_{(2)}$, $x_{(3)}$, . . ., $x_{(n)}$. The final estimates will be the estimates $\hat{\alpha}$, $\hat{\beta}$ and $\hat{m}$ for case B.

This procedure has been examined by using extensive Monte Carlo studies. We have found that it *always* converges when case B occurs (i.e. when $\hat{m}$ is small) and gives a better fit, as measured by the goodness-of-fit statistics, to the Weibull distribution (1). It will also sometimes converge when case A occurs (usually when m is small), but since, in this case, there is also an ML estimate, the ML estimate should be taken.

3. GOODNESS-OF-FIT TESTS

In this section, the EDF tests are described. The null hypothesis is H_0: the random sample $x_1, x_2, \ldots, x_n$ comes from distribution (1). Let $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ be the order statistics of the sample.

(a) First, estimates of the unknown parameters must be found, as described above. Then, for $i=1, 2, \ldots, n$, make the transformation $z_{(i)}=F(x_{(i)}; \hat{\alpha}, \hat{\beta}, \hat{m})$.

(b) The EDF statistics are calculated as follows:

$$W^2 = \sum\left(z_{(i)} - \frac{2i-1}{2n}\right)^2 + \frac{1}{12n},$$

$$U^2 = W^2 - n(\bar{z}-0.5)^2,$$

$$A^2 = -n - \frac{1}{n}\sum (2i-1)\{\log z_{(i)} + \log(1-z_{(n-i+1)})\}$$

where the sums are over $i=1, 2, \ldots, n$, log means natural logarithm and $\bar{z}=\Sigma z_{(i)}/n$.

(c) Let $c=1/\hat{m}$. Enter Table 1, using the subtable for the appropriate statistic. When $\hat{m} > 2$, we have $0 < c < 0.5$ and Table 1 is entered at the line corresponding to c; when $\hat{m} \leqslant 2$, so that $c \geqslant 0.5$, the *last line*, labelled $c=0.5$, should be used. The null hypothesis is rejected at significance level p if the statistic used is greater than the value given for level p. Table 1 has been given using c rather than $\hat{m}$ because linear interpolation for c will give good accuracy. The points given are for the asymptotic distributions of the statistics; however, Monte Carlo studies show that they can be used with good accuracy for smaller values of n, say $n \geqslant 10$; for $n < 10$ a goodness-of-fit test would in any case have very little power.

4. EXAMPLES

We illustrate the tests with two examples.

4.1. *Example 1*

Data set 1 is taken from Table 3 of Cox and Oakes (1984) and consists of 10 values of the number of cycles to failure when springs are subjected to various stress levels. For these data, the stress level is 950 N mm^{-2}, and the values, given in units of 1000 cycles, are

225, 171, 198, 189, 189, 135, 162, 135, 117, 162.

Fig. 1(a) comes from this data set. The ML estimates are $\hat{\alpha}=99.02$, $\hat{\beta}=78.23$ and $\hat{m}=2.38$, so that $c=0.420$. The value of A^2 is 0.260 and this is not significant,

TABLE 1
Critical points for W^2, U^2 and A^2

c	*Critical points for the following values of α:* 0.500	0.750	0.850	0.900	0.950	0.975	0.990	0.995
Critical points for W^2								
0.0	0.044	0.062	0.075	0.085	0.103	0.120	0.144	0.162
0.05	0.044	0.063	0.076	0.086	0.104	0.122	0.145	0.163
0.10	0.044	0.063	0.077	0.087	0.105	0.123	0.147	0.165
0.15	0.045	0.064	0.077	0.088	0.106	0.125	0.149	0.168
0.20	0.045	0.065	0.079	0.089	0.108	0.127	0.152	0.170
0.25	0.046	0.066	0.080	0.091	0.110	0.129	0.154	0.174
0.30	0.047	0.067	0.081	0.093	0.112	0.132	0.157	0.177
0.35	0.047	0.068	0.083	0.094	0.114	0.134	0.161	0.181
0.40	0.048	0.069	0.085	0.097	0.117	0.138	0.165	0.186
0.45	0.049	0.071	0.087	0.099	0.120	0.141	0.170	0.191
0.50	0.050	0.073	0.089	0.102	0.124	0.146	0.175	0.197
Critical points for U^2								
0.0	0.043	0.061	0.074	0.084	0.102	0.119	0.143	0.160
0.05	0.043	0.062	0.075	0.085	0.103	0.121	0.144	0.162
0.10	0.044	0.062	0.076	0.086	0.104	0.122	0.146	0.164
0.15	0.044	0.063	0.077	0.087	0.105	0.123	0.148	0.166
0.20	0.045	0.064	0.077	0.088	0.107	0.125	0.150	0.168
0.25	0.045	0.065	0.078	0.089	0.108	0.127	0.152	0.171
0.30	0.046	0.065	0.080	0.091	0.110	0.129	0.154	0.173
0.35	0.046	0.066	0.081	0.092	0.111	0.131	0.157	0.176
0.40	0.047	0.067	0.082	0.094	0.113	0.133	0.159	0.180
0.45	0.048	0.068	0.083	0.095	0.115	0.136	0.162	0.183
0.50	0.048	0.070	0.085	0.097	0.118	0.138	0.166	0.187
Critical points for A^2								
0.0	0.292	0.395	0.467	0.522	0.617	0.711	0.836	0.931
0.05	0.295	0.399	0.471	0.527	0.623	0.719	0.845	0.941
0.10	0.298	0.403	0.476	0.534	0.631	0.728	0.856	0.954
0.15	0.301	0.408	0.483	0.541	0.640	0.738	0.869	0.969
0.20	0.305	0.414	0.490	0.549	0.650	0.751	0.885	0.986
0.25	0.309	0.421	0.498	0.559	0.662	0.765	0.902	1.007
0.30	0.314	0.429	0.508	0.570	0.676	0.782	0.923	1.030
0.35	0.320	0.438	0.519	0.583	0.692	0.802	0.947	1.057
0.40	0.327	0.448	0.532	0.598	0.711	0.824	0.974	1.089
0.45	0.334	0.469	0.547	0.615	0.732	0.850	1.006	1.125
0.50	0.342	0.472	0.563	0.636	0.757	0.879	1.043	1.167

using Table 1, at the 50% level. The other two statistics give values $W^2=0.041$ and $U^2=0.040$; all three statistics indicate a very good Weibull fit.

4.2. *Example 2*

Data set 2 consists of 15 times to failure of air-conditioning equipment in aircraft, measured in hours; the data are taken from Table 1 of Proschan (1963) and are the data for aircraft 7910. The values are as follows:

74, 57, 48, 29, 502, 12, 70, 21, 29, 386, 59, 27, 153, 26, 326.

Fig. 1(b) comes from this data set. The ML estimates are $\hat{\alpha}=9.313$ (using the bias reduction procedure in Section 2), $\hat{\beta}=93.50$ and $\hat{m}=0.763$. The value of A^2

is 0.54; since $c=1/\hat{m}=1.31$ is greater than 0.5, Table 1 is entered at the last line ($c=0.5$), giving a significance level of approximately 0.17. Statistic $W^2=0.099$ and $U^2=0.093$; both give approximate significance level 0.12.

5. ASYMPTOTIC THEORY OF EMPIRICAL DISTRIBUTION FUNCTION TESTS

In this section the asymptotic theory of EDF tests is summarized. The calculation of asymptotic distributions of EDF statistics follows a well-known procedure (see, for example, Durbin (1973) or Stephens (1976)). The procedure rests on the fact that $y_n(x) = \{F_n(z) - z\}\sqrt{n}$, where $F_n(z)$ is the EDF of the set of $z_{(i)}$, tends to a Gaussian process $y(z)$ as $n \to \infty$, and the statistics are functionals of this process. The mean of $y(z)$ is 0: we need the covariance function $\rho(s,t) = E\{y(s)\,y(t)\}$. When all parameters are known (case 0), this covariance is $\rho_0(s,t) = \min(s,t) - st$. When the parameters are estimated, the covariance will not depend on the true values of location or scale parameters α or β, provided that these parameters are estimated efficiently, but it will depend on the true shape parameter m.

Suppose that the parameters are components of a vector θ, with $\theta_1=\alpha$, $\theta_2=\beta$ and $\theta_3=m$. Let $F(x;\theta)$ now denote the distribution $F(x;\ \alpha,\ \beta,\ m)$ and let $f(x;\theta)$ be the corresponding density. Suppose that a vector $g(s)$, with components $g_i(s)$, is constructed as follows:

$$g_i(s) = \frac{\partial F(x;\theta)}{\partial \theta_i}, \qquad i=1,\ 2,\ 3, \tag{12}$$

where the right-hand side is written as a function of s using the transformation $s=F(x;\theta)$. Let $\{g(s)\}'$ denote the transpose of $g(s)$. Let D be the symmetric matrix with entries

$$\delta_{ij} = E\left[-\frac{\partial^2\{\log f(x;\theta)\}}{\partial\theta_i\,\partial\theta_j}\right], \qquad i,j=1,\ 2,\ 3,$$

where E denotes expectation, and let Σ be the inverse of D. Then

$$\rho(s,t) = \rho_0(s,t) - \{g(s)\}'\Sigma\{g(s)\}. \tag{13}$$

From equation (12), the components of $g(s)$ become, after some algebra, and using F for $F(x;\ \theta)$

$$\left.\begin{aligned} g_1(s) &= \partial F/\partial\alpha = -m(1-s)/\beta - \{-\log(1-s)\}^{(m-1)/m}, \\ g_2(s) &= \partial F/\partial\beta = \{m(1-s)/\beta\}\log(1-s), \\ g_3(s) &= \partial F/\partial m = \{-(1-s)/m\}\log(1-s)\log\{-\log(1-s)\}. \end{aligned}\right\} \tag{14}$$

Also, for $m > 2$, D has the upper right-hand terms

$$D = \begin{pmatrix} \dfrac{(m-1)^2}{\beta^2}\Gamma\left(1-\dfrac{2}{m}\right) & \dfrac{m(m-1)}{\beta^2}\Gamma\left(1-\dfrac{1}{m}\right) & \dfrac{\Gamma(m-1)-\Gamma(2-1/m)-\Gamma'(2-1/m)}{\beta} \\ & \dfrac{m^2}{\beta^2} & -\dfrac{\Gamma'(1)+1}{\beta} \\ & & \dfrac{\Gamma''(1)+2\Gamma'(1)}{m^2} \end{pmatrix} \tag{15}$$

When Σ is calculated, and $g(s)$ and Σ are inserted into equation (13), $\rho(s, t)$ will be independent of α and β. The Cramér–von Mises statistic W^2 is based directly on the process $y(z)$, and U^2 is based on

$$u(z) = y(z) - \int_0^1 y(z)\, \mathrm{d}z;$$

A^2 is based on the process $a(z) = y(z)/\{z(1-z)\}^{1/2}$; asymptotically, the values of the statistics are given by

$$W^2 = \int_0^1 y^2(z)\, \mathrm{d}z, \qquad U^2 = \int_0^1 u^2(z)\, \mathrm{d}z, \qquad A^2 = \int_0^1 a^2(z)\, \mathrm{d}z.$$

The asymptotic distributions of these statistics are sums of weighted independent χ_1^2-variables; the weights must be found from the eigenvalues of an integral equation with, for W^2, $\rho(s, t)$ as kernel. For U^2 and A^2, one must find the $\rho(s, t)$ of the $u(z)$ and $a(z)$ processes. Once the weights are known, the percentage points of the distributions can be calculated by Imhof's method. The techniques are straightforward once the $\rho(s, t)$ are known, and we omit the details; they are given in Lockhart and Stephens (1989). When $m \leqslant 2$, the ML estimate $\hat{\alpha}$ of α is superefficient in the sense of Darling (1955), and then the asymptotic percentage points are the same as for when α is known. These are the points in the last lines of the subtables in Table 1.

EDF statistics are known to provide powerful tests for many distributions; the powers naturally depend on the alternatives considered, and a study is being made on power properties for the various alternatives to the Weibull distribution usually encountered. On the whole, with the limited power results currently available, the statistic A^2 is suggested as the preferred statistic for overall Weibull testing. The other statistics have been included for completeness. Tables for tests where one or both of β or m are known are given by Lockhart and Stephens (1989).

ACKNOWLEDGEMENTS

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada and by the US Office of Naval Research; it was finished while the second author was partly supported by the Science and Engineering Research Council of Great Britain. All these agencies are thanked for their assistance.

REFERENCES

Cheng, R. C. H. and Iles, T. C. (1987) Corrected maximum likelihood in non-regular problems. *J. R. Statist. Soc.* B, **49**, 95–101.

——— (1990) Embedded models in three-parameter distributions and their estimation. *J. R. Statist. Soc.* B, **52**, 135–149.

Cox, D. R. and Oakes, D. (1984) *Analysis of Survival Data*. London: Chapman and Hall.

Darling, D. A. (1955) The Cramér–Smirnov test in the parametric case. *Ann. Math. Statist.*, **26**, 1–20.

Durbin, J. (1973) *Distribution Theory for Tests Based on the Sample Distribution Function*. Philadelphia: Society for Industrial and Applied Mathematics.

Lockhart, R. A. and Stephens, M. A. (1989) Tests of fit for the three-parameter Weibull distribution. *Research Report 89-08*. Department of Mathematics and Statistics, Simon Fraser University, Burnaby.

——— (1992) Estimation and tests of fit for the three-parameter Weibull distribution. *Research Report 92-10.* Department of Mathematics and Statistics, Simon Fraser University, Burnaby.

Proschan, F. (1963) Theoretical explanation of observed decreasing failure rate. *Technometrics*, **5**, 375–383.

Rockette, H., Antle, C. and Klimko, L. A. (1974) Maximum likelihood estimation with the Weibull model. *J. Am. Statist. Ass.*, **69**, 246–249.

Smith, R. L. (1985) Maximum Likelihood estimation in a class of non-regular cases. *Biometrika*, **72**, 67–90.

Smith, R. L. and Naylor, J. C. (1987) A comparison of Maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution. *Appl. Statist.*, **36**, 358–369.

Smith, R. L. and Weissman, I. (1985) Maximum likelihood estimation of the lower tail of a probability distribution. *J. R. Statist. Soc.* B, **47**, 285–298.

Stephens, M. A. (1976) Asymptotic results for goodness-of-fit statistics with unknown parameters. *Ann. Statist.*, **4**, 357–369.

——— (1986) Tests based on EDF statistics. In *Goodness-of-fit Techniques* (eds R. B. D'Agostino and M. A. Stephens), ch. 4. New York: Dekker.

Weissman, I. (1982) Confidence intervals for the threshold parameter: II, Unknown shape parameter. *Communs Statist. Theory Meth.*, **11**, 2451–2474.