ADJUSTED PROFILE LIKELIHOODS FOR THE WEIBULL SHAPE PARAMETER

SILVIA L.P. FERRARI†
Departamento de Estatística, IME, Universidade de São Paulo
Caixa Postal 66281, São Paulo/SP, 05311–970, Brazil
email: sferrari@ime.usp.br

MICHEL FERREIRA DA SILVA
Departamento de Estatística, IME, Universidade de São Paulo
Caixa Postal 66281, São Paulo/SP, 05311–970, Brazil
email: michel@ime.usp.br

FRANCISCO CRIBARI–NETO
Departamento de Estatística, CCEN, Universidade Federal de Pernambuco
Cidade Universitária, Recife/PE, 50740–540, Brazil
email: cribari@ufpe.br

Abstract. This paper presents several different adjusted profile likelihoods for the Weibull shape parameter. These adjustments aim at reducing the impact of the nuisance parameter on the likelihood-based inference regarding the parameter of interest. Both point estimation and hypothesis testing are considered. We also show that the ratio between the estimators and the shape parameter are pivotal quantities and that the size properties of the usual and adjusted profile likelihood ratio tests depend neither on the scale parameter nor on the value of the shape parameter set at the null hypothesis. The numerical results suggest that the adjustment obtained by Yang and Xie (2003) outperforms not only the profile likelihood inference but also inference based on competing adjusted profile likelihoods.

Keywords: Censoring; failure time data; likelihood ratio test; maximum likelihood estimator; nuisance parameter; profile likelihood; Weibull distribution.

1. INTRODUCTION

The Weibull distribution is commonly used to model failure time data since it generalizes the exponential distribution allowing for a power dependence of the hazard function on time. This power dependence is controlled by the distribution shape parameter. It is thus important to reliably perform inference on such a parameter. It is also noteworthy that the Weibull distribution admits a closed-form expression for tail area probabilities and thereby simple formulas for survival and hazard functions. For applications to industrial life testing

† Corresponding author.

and biomedical data, see Mann et al. (1974) and Gross and Clark (1975). The reader is also referred to Lawless (1982) and Klein and Moeschberger (1997).

Let y be a random variable that follows a Weibull distribution, denoted $W(\alpha, \beta)$. Its density function is given by

$$p(y; \alpha, \beta) = \frac{\beta}{\alpha}\left(\frac{y}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{y}{\alpha}\right)^{\beta}\right], \quad \alpha, \beta, y > 0,$$

where β is the shape parameter and α is the scale parameter. The distribution function is $F(y; \alpha, \beta) = 1 - S(y; \alpha, \beta)$, where

$$S(y; \alpha, \beta) = \exp\left[-\left(\frac{y}{\alpha}\right)^{\beta}\right]$$

is the survival function. The hazard function, $\Lambda(y; \alpha, \beta) = p(y; \alpha, \beta)/S(y; \alpha, \beta)$, is thus given by $(\beta/\alpha)(y/\alpha)^{\beta-1}$. Note that $\beta < 1$, $\beta = 1$ and $\beta > 1$ imply decreasing, constant and increasing hazard functions, respectively. For example, after a major surgery the mortality hazard is typically decreasing over time, at least in the short term. There are situations where increasing and even constant hazard functions are plausible. They can all be obtained when the modeling is done via the Weibull distribution.

Parameter estimation can be performed by maximum likelihood. However, the resulting estimators may be considerably biased in small samples, including the shape parameter maximum likelihood estimator. Some recent papers discussing the estimation of Weibull shape parameters are Ageel (2002), Bar-Lev (2004), Lu et al. (2004) and Maswadah (2003). Yang and Xie (2003) obtained the Cox and Reid (1987) adjustment to the Weibull shape parameter profile likelihood function. Their Monte Carlo results suggest that the adjusted profile likelihood estimator is nearly unbiased and is also more efficient than the usual maximum likelihood estimator in small samples both with and without censoring.

The chief goal of this paper is to obtain several different adjustments to the profile likelihood for the Weibull shape parameter, and to evaluate their finite-sample behavior. We shall consider both estimation and hypothesis testing. We show that, for all estimators of β considered in this paper and for all sampling schemes (complete samples and types I and II censoring), the distribution of the ratio 'estimator/β' does not depend on (α, β), i.e., this is a pivotal quantity. We also show that the size properties of the usual and adjusted profile likelihood ratio tests depend neither on α nor on the value of β set at the null hypothesis. The numerical results suggest that the Cox and Reid (1987) adjustment derived by Yang and Xie (2003) outperforms competing inference strategies both with noncensored and censored data.

The remainder of the paper unfolds as follows. Section 2 presents the profile likelihood and its properties. Section 3 presents the Barndorff–Nielsen (1980, 1983) adjustment and some alternative adjusted profile likelihoods, for example, the Cox and Reid (1987) adjustment. Section 4 derives the adjustments for inference on the Weibull shape parameter.

Monte Carlo results are presented in Section 5 and numerical examples with real data sets are presented in Section 6. Finally, Section 7 concludes the paper. Some distributional properties of the different estimators and test statistics are derived in the Appendix.

2. PROFILE LIKELIHOOD

Let $\mathcal{Y} = (y_1, \ldots, y_n)^\top$ be an n-vector of independent random variables having a distribution that is indexed by two (possibly vector-valued) parameters, namely α and β.[1] Suppose that the interest lies in performing inference on β in the presence of the nuisance parameter α. In some situations, it is possible to perform inference on β using a marginal or a conditional likelihood function. However, there are a number of situations where these functions cannot be obtained. The standard approach is to use the profile likelihood function, which is defined as $L_p(\beta) = L(\widehat{\alpha}_\beta, \beta)$, where $L(\cdot)$ is the usual likelihood function and $\widehat{\alpha}_\beta$ is the maximum likelihood estimate of α for a given, fixed β. For instance, the usual likelihood ratio statistic,

$$LR(\beta) = 2[\ell(\widehat{\alpha}, \widehat{\beta}) - \ell(\widehat{\alpha}_\beta, \beta)] = 2[\ell_p(\widehat{\beta}) - \ell_p(\beta)],$$

is based on the profile likelihood function. Here, $\widehat{\beta}$ and $\widehat{\alpha}$ are the maximum likelihood estimates of β and α respectively, $\ell(\cdot)$ is the log-likelihood function and $\ell_p(\cdot)$ is the profile log-likelihood function. It is noteworthy, however, that $L_p(\cdot)$ is not a genuine likelihood. For example, for $\theta = (\alpha^\top, \beta^\top)^\top$, properties such as

$$E(u(\theta)) = 0 \quad \text{and} \quad E\{u(\theta)u(\theta)^\top\} + E\Big\{\frac{\partial u(\theta)}{\partial \theta^\top}\Big\} = 0$$

do not hold when $u_p(\beta)$ is used instead of $u(\theta)$. Here, $u(\theta) = \partial\ell(\theta)/\partial\theta$ is the score function and $u_p(\beta) = \partial\ell_p(\beta)/\partial\beta$ is the profile score function. The profile score and information biases are only guaranteed to be $O(1)$.

3. MODIFIED PROFILE LIKELIHOODS

3.1. BARNDORFF–NIELSEN'S MODIFIED PROFILE LIKELIHOOD

Several adjustments to the profile likelihood function have been proposed in the literature; see, e.g., Severini (2000, Chapter 9). Barndorff–Nielsen's (1983) modified profile likelihood is obtained as an approximation to a marginal or conditional likelihood for β, if either exists. In both cases, one uses the p^* formula (Barndorff–Nielsen, 1980, 1983), which approximates

[1] Throughout this section α and β are allowed to be vector-valued. However, in the model of interest in Section 4 these paramaters are scalars.

the probability density function of the maximum likelihood estimator conditional on an ancillary statistic. The modified profile likelihood proposed by the author is

$$L_{BN}(\beta) = \left|\frac{\partial\widehat{\alpha}_\beta}{\partial\widehat{\alpha}}\right|^{-1} |j_{\alpha\alpha}(\widehat{\alpha}_\beta, \beta)|^{-1/2} L_p(\beta),$$

where $j_{\alpha\alpha}(\alpha, \beta) = -\partial^2\ell/\partial\alpha\partial\alpha^\top$ is the observed information for α. The corresponding score and information biases are of order $O(n^{-1})$. Additionally, $L_{BN}(\beta)$ is invariant under reparameterization of the form $(\alpha, \beta) \rightarrow (\nu, \xi)$, where $\nu = \nu(\alpha, \beta)$ and $\xi = \xi(\beta)$.

The main difficulty in computing the modified profile likelihood function $L_{BN}(\beta)$ lies in obtaining $|\partial\widehat{\alpha}_\beta/\partial\widehat{\alpha}|$. There is an alternative expression for $L_{BN}(\beta)$ that does not involve this term, but it involves a sample space derivative of the log-likelihood function and the specification of an ancillary a such that $(\widehat{\alpha}, \widehat{\beta}, a)$ is a minimal sufficient statistic.

3.2. AN APPROXIMATION BASED ON POPULATION COVARIANCES

Several approximations to Barndorff–Nielsen's adjustment were proposed in order to simplify its evaluation. Severini (1998) proposed the following approximation to the modified profile log-likelihood function:

$$\bar{\ell}_{BN}(\beta) = \ell_p(\beta) + \frac{1}{2}\log\left|j_{\alpha\alpha}(\widehat{\alpha}_\beta, \beta)\right| - \log\left|I_\alpha(\widehat{\alpha}_\beta, \beta; \widehat{\alpha}, \widehat{\beta})\right|, \tag{3.1}$$

where

$$I_\alpha(\alpha, \beta; \alpha_0, \beta_0) = E_{(\alpha_0,\beta_0)}\{\ell_\alpha(\alpha, \beta)\ell_\alpha(\alpha_0, \beta_0)^\top\}, \tag{3.2}$$

with $\ell_\alpha(\alpha, \beta) = \partial\ell/\partial\alpha$. Note that $I_\alpha(\widehat{\alpha}_\beta, \beta; \widehat{\alpha}, \widehat{\beta})$ does not depend on the ancillary statistic a and that $I_\alpha(\alpha, \beta; \alpha_0, \beta_0)$ is the covariance between $\ell_\alpha(\alpha, \beta)$ and $\ell_\alpha(\alpha_0, \beta_0)$. The corresponding maximum likelihood estimator shall be denoted as $\widehat{\overline{\beta}}_{BN}$.

3.3. AN APPROXIMATION BASED ON EMPIRICAL COVARIANCES

An alternative approximation to Barndorff–Nielsen's (1983) modified profile likelihood function, say $\breve{\ell}_{BN}$, was proposed by Severini (1999); it was obtained replacing $I(\alpha, \beta; \alpha_0, \beta_0)$ by

$$\breve{I}(\alpha, \beta; \alpha_0, \beta_0) = \textstyle\sum_{j=1}^n \ell_\alpha^{(j)}(\alpha, \beta)\ell_\alpha^{(j)}(\alpha_0, \beta_0)^\top, \tag{3.3}$$

where $\ell_\theta^{(j)}(\theta) = (\ell_\alpha^{(j)}(\theta), \ell_\beta^{(j)}(\theta))$ is the score function for the jth observation. This approximation is particularly useful when the computation of expected values of products of log-likelihood derivatives is cumbersome. The corresponding estimator shall be denoted as $\widehat{\breve{\beta}}_{BN}$.

3.4. AN APPROXIMATION BASED ON AN ANCILLARY STATISTIC

A third approximation to Barndorff–Nielsen's (1983) modified profile log-likelihood function is (Fraser and Reid (1995), and Fraser et al. (1999))

$$\tilde{\ell}_{BN}(\beta) = \ell_p(\beta) + \frac{1}{2}\log|j_{\alpha\alpha}(\widehat{\alpha}_\beta, \beta)| - \log|\ell_{\alpha;\mathcal{Y}}(\widehat{\alpha}_\beta, \beta)\widehat{V}_\alpha|, \tag{3.4}$$

where $\ell_{\alpha;\mathcal{Y}}(\alpha, \beta) = \partial\ell_\alpha(\alpha, \beta)/\partial\mathcal{Y}^\top$ and

$$\widehat{V}_\alpha = \begin{pmatrix} -\frac{\partial F_1(y_1;\widehat{\alpha},\widehat{\beta})/\partial\widehat{\alpha}}{p_1(y_1;\widehat{\alpha},\widehat{\beta})} & \cdots & -\frac{\partial F_n(y_n;\widehat{\alpha},\widehat{\beta})/\partial\widehat{\alpha}}{p_n(y_n;\widehat{\alpha},\widehat{\beta})} \end{pmatrix}^\top,$$

$p_j(y;\alpha, \beta)$ being the probability density function of y_j and $F_j(y;\alpha, \beta)$ being the cumulative distribution function of y_j. The corresponding estimator shall be denoted as $\widehat{\widehat{\beta}}_{BN}$. The construction of the matrix $\widehat{V}_\alpha$ is based on an approximately ancillary statistic; see Severini (2000, p. 216).

3.4. AN APPROXIMATION BASED ON ORTHOGONAL PARAMETERS

We shall now consider a different adjustment to the profile likelihood function. Suppose that the parameters that index the model are orthogonal, that is, that the elements of the score vector, $\partial\ell/\partial\beta$ and $\partial\ell/\partial\alpha$, are uncorrelated. Cox and Reid (1987) proposed an adjustment that can be applied to the profile likelihood function in this setting. It is an approximation to a conditional probability density function of the observations given the maximum likelihood estimator of α and can be written as

$$L_{CR}(\beta) = |j_{\alpha\alpha}(\widehat{\alpha}_\beta, \beta)|^{-1/2}L_p(\beta).$$

Their modified profile log-likelihood function is

$$\ell_{CR}(\beta) = \ell_p(\beta) - \frac{1}{2}\log|j_{\alpha\alpha}(\widehat{\alpha}_\beta, \beta)|; \tag{3.5}$$

the maximizer of $\ell_{CR}(\beta)$ shall be denoted as $\widehat{\beta}_{CR}$. The corresponding score bias is $O(n^{-1})$ but, in general, the information bias remains $O(1)$.

The Cox and Reid (1987) adjustment has been proposed under the assumption of orthogonality of β and α. It is not always possible, however, to find a parameterization that delivers orthogonality.[2] Additionally, the Cox and Reid adjustment is not invariant

[2] We were not able to obtain an orthogonal parameterization under types I and II censoring (Section 3) when performing inference on the Weibull shape parameter. Following Yang and Xie (2003), we have used orthogonal parameterizations obtained under noncensoring and they delivered reliable inference in small samples; see the numerical results in Section 5.

under reparameterizations of the form $(\alpha, \beta) \to (\nu, \xi)$, where $\nu = \nu(\alpha, \beta)$ and $\xi = \xi(\beta)$ unlike $L_{BN}(\beta)$, for which the invariance property is guaranteed by the term $|\partial\widehat{\alpha}_\beta/\partial\widehat{\alpha}|$. Note that if $\widehat{\alpha}_\beta = \widehat{\alpha}$ for all β, then $L_{BN}(\beta) = L_{CR}(\beta)$. In this case, β and α are orthogonal parameters (Cox and Reid, 1987). Also, it is possible to show that $\ell_{BN}(\beta) - \ell_{BN}(\widehat{\beta}) = \ell_{CR}(\beta) - \ell_{CR}(\widehat{\beta}) + O_p(n^{-1})$. As a consequence, the likelihood ratio statistics obtained from $\ell_{BN}(\beta)$ and $\ell_{CR}(\beta)$ differ by a term of order $O_p(n^{-1})$.

4. PROFILE LIKELIHOODS FOR THE WEIBULL SHAPE PARAMETER

4.1. UNCENSORED DATA

At the outset, consider noncensored data. Let $y_1, \ldots, y_n$ be independent and identically distributed (i.i.d.) Weibull random variables. The log-likelihood function for the (α, β) parameters is given by

$$\ell(\alpha, \beta) = n \log\left(\frac{\beta}{\alpha}\right) + (\beta - 1)\sum_{j=1}^n \log\left(\frac{y_j}{\alpha}\right) - \sum_{j=1}^n \left(\frac{y_j}{\alpha}\right)^\beta. \tag{4.1}$$

The restricted maximum likelihood estimator of α, for fixed β (parameter of interest), is

$$\widehat{\alpha}_\beta = \left(\frac{1}{n}\sum_{j=1}^n y_j^\beta\right)^{1/\beta}.$$

Plugging this expression into the expression for $\ell(\alpha, \beta)$ we obtain

$$\ell_p(\beta) = n\log\beta - n\log\sum_{j=1}^n y_j^\beta + \beta\sum_{j=1}^n \log y_j.$$

The observed information $j_{\alpha\alpha}(\alpha, \beta)$ evaluated at $(\widehat{\alpha}_\beta, \beta)$ can be written as

$$j_{\alpha\alpha}(\widehat{\alpha}_\beta, \beta) = \beta^2\Big(\sum_{j=1}^n y_j^\beta\Big)^{-2/\beta} n^{1+2/\beta}.$$

In what follows we shall obtain the adjusted profile likelihoods described in Section 3. We shall omit the derivation details in the interest of space. From (3.2), we obtain

$$I_\alpha(\widehat{\alpha}_\beta, \beta; \widehat{\alpha}, \widehat{\beta}) = \frac{n\beta^2\widehat{\alpha}^{\beta-1}}{\widehat{\alpha}_\beta^{\beta+1}}\,\Gamma\Big(\frac{\beta}{\widehat{\beta}} + 1\Big).$$

Thus, from (3.1),

$$\overline{\ell}_{BN}(\beta) = (n-1)\log\big(\beta/\sum_{j=1}^n y_j^\beta\big) - \beta\Big(\log\widehat{\alpha} - \sum_{j=1}^n \log y_j\Big) - \log\Gamma(1 + \beta/\widehat{\beta}), \tag{4.2}$$

where $\widehat{\alpha}$ and $\widehat{\beta}$ are the maximum likelihood estimators; these estimators do not have closed-form.

Using (3.3) it follows that

$$\breve{I}_\alpha(\widehat{\alpha}_\beta, \beta; \widehat{\alpha}, \widehat{\beta}) = \frac{\beta}{\widehat{\alpha}_\beta}\left(\frac{\widehat{\beta}}{\widehat{\alpha}^{\widehat{\beta}+1}} \frac{\sum_{j=1}^n y_j^{\beta+\widehat{\beta}}}{\widehat{\alpha}_\beta^\beta} - \frac{n\widehat{\beta}}{\widehat{\alpha}}\right),$$

and, hence,

$$\breve{\ell}_{BN}(\beta) = n\log\beta - (n-1)\log\sum_{j=1}^n y_j^\beta + \sum_{j=1}^n \log y_j^\beta - \log\left[\sum_{j=1}^n y_j^\beta\left(y_j^{\widehat{\beta}} - \widehat{\alpha}^{\widehat{\beta}}\right)\right]. \quad (4.3)$$

Additionally, the jth components ($j = 1, \ldots, n$) of the vectors $\ell_{\alpha;y}(\widehat{\alpha}_\beta, \beta)$ (row vector) and $\widehat{V}_\alpha$ (column vector) are, respectively, $\beta^2 y_j^{\beta-1}/\widehat{\alpha}_\beta^{\beta+1}$ and $y_j/\widehat{\alpha}$, and, thus, the adjusted profile likelihood given in (3.4) can be written as

$$\tilde{\ell}_{BN}(\beta) = (n-1)\log\beta - n\log\sum_{j=1}^n y_j^\beta + \beta\sum_{j=1}^n \log y_j. \quad (4.4)$$

From (3.5), under an orthogonal parameterization (λ, β), Yang and Xie (2003) showed that

$$\ell_{CR}(\beta) = \ell_p(\beta) - 2\log\beta = (n-2)\log\beta - n\log\sum_{j=1}^n y_j^\beta + \beta\sum_{j=1}^n \log y_j, \quad (4.5)$$

where $\lambda = \log(\alpha)/\beta + (1-\gamma)/\beta^2$ e $\gamma = 0.577215\ldots$ is Euler's constant.

4.2. TYPE II CENSORED DATA

We shall now consider type II censored data, where observation ceases after the rth ($r < n$) failure. Let $y_{(1)}, \ldots, y_{(r)}$ be the smallest r order statistics from a sample of size n of a Weibull distribution $W(\alpha, \beta)$. The likelihood function is

$$L(\alpha, \beta) = \prod_{j=1}^r p(y_{(j)}; \alpha, \beta)\prod_{j=r+1}^n S(y_{(r)}; \alpha, \beta) = \left[S(y_{(r)}; \alpha, \beta)\right]^{n-r}\prod_{j=1}^r p(y_{(j)}; \alpha, \beta).$$

Hence,

$$\ell(\alpha, \beta) = r\log\beta - \beta r\log\alpha + \beta\sum_{j=1}^r \log y_{(j)} - \frac{1}{\alpha^\beta}\left[\sum_{j=1}^r y_{(j)}^\beta - (n-r)y_{(r)}^\beta\right]. \quad (4.6)$$

The restricted maximum likelihood estimator of α can be written as

$$\widehat{\alpha}_\beta = \left[\frac{1}{r}\left(\sum_{j=1}^r y_{(j)}^\beta + (n-r)y_{(r)}^\beta\right)\right]^{1/\beta}.$$

Note that under no censoring ($r = n$), this estimator reduces to the one given previously.

Therefore, we obtain

$$\ell_p(\beta) = \ell(\widehat{\alpha}_\beta, \beta) = r\log\beta - r\log\left(\sum_{j=1}^r y_{(j)}^\beta + (n-r)y_{(r)}^\beta\right) + \beta\sum_{j=1}^r \log y_{(j)}.$$

Here,

$$j_{\alpha\alpha}(\widehat{\alpha}_\beta, \beta) = \beta^2\left(\sum_{j=1}^r y_{(j)}^\beta + (n-r)y_{(r)}^\beta\right)^{-2/\beta} r^{1+2/\beta}.$$

It has not been possible to derive $\bar{\ell}_{BN}(\beta)$ under type II censoring. It can be shown that

$$\breve{\ell}_{BN}(\beta) = r\log\beta - (r-1)\log\left[\sum_{j=1}^r y_{(j)}^\beta + (n-r)y_{(r)}^\beta\right] + \beta\sum_{j=1}^r \log y_{(j)}$$

$$-\log\left[r\left(\sum_{j=1}^r y_{(j)}^{\beta+\widehat{\beta}} + (n-r)y_{(r)}^{\beta+\widehat{\beta}}\right) + (n-r)^2 y_{(r)}^{\beta+\widehat{\beta}} - \left(\sum_{j=1}^r y_{(j)}^{\widehat{\beta}}\right)\left(\sum_{j=1}^r y_{(j)}^\beta\right)\right], \quad (4.7)$$

where $\widehat{\beta}$ is the maximum likelihood estimator of β under type II censoring. We also obtain

$$\tilde{\ell}_{BN}(\beta) = (r-1)\log\left[\beta/(\sum_{j=1}^{r} y_{(j)}^{\beta} + (n-r)y_{(r)}^{\beta})\right] + \sum_{j=1}^{r}\log y_{(j)}^{\beta} - \log\sum_{j=1}^{r} y_{(j)}^{\beta}. \quad (4.8)$$

Under the same orthogonal parameterization used under noncensoring, Yang and Xie (2003) obtained

$$\ell_{CR}(\beta) = (r-2)\log\beta - r\log\left[\sum_{j=1}^{r} y_{(j)}^{\beta} + (n-r)y_{(r)}^{\beta}\right] + \beta\sum_{j=1}^{r}\log y_{(j)}. \quad (4.9)$$

4.3. TYPE I CENSORED DATA

We shall now move to the situation where there is type I censoring, that is, one observes $y_j = \min(t_j, c)$, where $t_j \sim W(\alpha, \beta)$ $(j = 1, \ldots, n)$ and c is a preassigned (fixed) censoring time. Here the data can be represented by n independent pairs of random variables (y_j, δ_j), where

$$y_j = \min(t_j, c) \qquad \text{and} \qquad \delta_j = \begin{cases} 1, & t_j \leq c, \\ 0, & t_j > c. \end{cases}$$

The likelihood function is

$$L(\alpha, \beta) = \prod_{j=1}^{n} p(y_j; \alpha, \beta)^{\delta_j}\ S(c; \alpha, \beta)^{1-\delta_j}.$$

Thus,

$$L(\alpha, \beta) = \left(\frac{\beta}{\alpha}\right)^{r} \exp\left[-\sum_{j=1}^{n}\left(\frac{y_j}{\alpha}\right)^{\beta}\right] \prod_{j\in\overline{C}}\left(\frac{y_j}{\alpha}\right)^{\beta-1},$$

where $\overline{C}$ denotes the set of noncensored observation indices and r is the number of censored cases. (Note that here, unlike type II censoring, r is random.) The log-likelihood function is

$$\ell(\alpha, \beta) = r\log\beta - r\log\alpha - \sum_{j=1}^{n}\left(\frac{y_j}{\alpha}\right)^{\beta} + (\beta-1)\sum_{j\in\overline{C}}\log\left(\frac{y_j}{\alpha}\right), \quad (4.10)$$

where $r = \sum\delta_j$ is the observed number of failures.

The restricted maximum likelihood estimator of α, for fixed β, is

$$\widehat{\alpha}_{\beta} = \left(\frac{1}{r}\sum_{j=1}^{n} y_j^{\beta}\right)^{1/\beta},$$

and it thus follows that the profile log-likelihood function can be written as

$$\ell_p(\beta) = r\log\beta - r\log\sum_{j=1}^{n} y_j^{\beta} + \beta\sum_{j\in\overline{C}}\log y_j.$$

Here,

$$j_{\alpha\alpha}(\widehat{\alpha}_{\beta}, \beta) = \beta^2\left(\sum_{j=1}^{n} y_j^{\beta}\right)^{-2/\beta} r^{1+2/\beta}.$$

As with type II censoring, it has not been possible to derive $\overline{\ell}_{BN}(\beta)$. However, we have obtained the following two adjusted log-likelihood functions:

$$\breve{\ell}_{BN}(\beta) = r\log\beta - (r-1)\log\sum_{j=1}^{n} y_j^{\beta} + \beta\sum_{j\in\overline{C}}\log y_j$$
$$-\log\Big[r\Big(\sum_{j\in\overline{C}} y_j^{\beta+\widehat{\beta}} + (n-r)c^{\beta+\widehat{\beta}}\Big) + (n-r)^2 c^{\beta+\widehat{\beta}} - \Big(\sum_{j\in\overline{C}} y_j^{\widehat{\beta}}\Big)\Big(\sum_{j\in\overline{C}} y_j^{\beta}\Big)\Big] \quad (4.11)$$

and

$$\tilde{\ell}_{BN}(\beta) = (r-1)\Big[\log\beta - \log\sum_{j=1}^{n} y_j^{\beta}\Big] + \beta\sum_{j\in\overline{C}}\log y_j - \log\sum_{j\in\overline{C}} y_j^{\beta}, \quad (4.12)$$

where $\widehat{\beta}$ is the maximum likelihood estimator of β under type I censoring.

Following Yang and Xie (2003), we now use the parameterization (λ, β) with $\lambda = \log(\alpha) + (1-\gamma)/\beta$, which is orthogonal under noncensoring. They have shown that

$$\ell_{CR}(\beta) = \ell_p(\beta) - \log\beta = (r-1)\log\beta - r\log\sum_{j=1}^{n} y_j^{\beta} + \beta\sum_{j\in\overline{C}}\log y_j. \quad (4.13)$$

It is important to note that the two orthogonal parameterizations obtained under noncensoring were also used, both by Yang and Xie (2003) and by us, under censoring, although it is not possible to guarantee parameter orthogonality when observations were recorded under either censoring scheme.

5. MONTE CARLO RESULTS

This section presents Monte Carlo results on the finite-sample inference regarding the Weibull shape parameter under both noncensoring and (types I and II) censoring. The sample size is $n = 20$. Under type II censoring, we considered samples with 10 and 15 failures, and under type I censoring we considered samples with 25% and 50% expected censored data. The true value of α was set at 100, and the following values for the parameter of interest were considered: $\beta = 0.5, 1.0, 5.0$; these correspond, respectively, to decreasing, constant and increasing hazard functions. All simulations were performed using the `Ox` matrix programming language (Doornik, 2001) and are based on 100,000 replications.

The following measures are presented for all point estimators: mean, variance, bias, mean squared error (MSE), relative bias (RB), skewness and kurtosis. Relative bias is defined as 100 × (bias / true parameter value)%. Since the results (relative bias, skewness and kurtosis) are the same for the three values of β, we shall only present them for $\beta = 5.0$ (increasing hazard rate). Also, the null rejection rates of the likelihood ratio tests based on the profile and the modified profile likelihoods are presented. The null hypotheses under test are $\mathcal{H}_0 : \beta = 0.5, 1.0, 5.0$. The test of $\mathcal{H}_0 : \beta = 1.0$ is of particular interest since under the null hypothesis the Weibull distribution reduces to the exponential distribution and the hazard function becomes constant over time. The results for the three null hypotheses were identical both without and with censoring.

5.1. UNCENSORED DATA

Table 1 contains results for estimation of the Weibull shape parameter in the noncensoring case. We consider the estimators obtained by maximizing the profile likelihoods given in Section 4.1: $\widehat{\beta}$ (4.1), $\widehat{\overline{\beta}}_{BN}$ (4.2), $\widehat{\breve{\beta}}_{BN}$ (4.3), $\widehat{\tilde{\beta}}_{BN}$ (4.4) e $\widehat{\beta}_{CR}$ (4.5). The figures in the table show that the different estimators have approximately the same skewness and kurtosis (0.9 and 4.7, respectively; the corresponding asymptotic values are 0 and 3). The estimator with smallest relative bias was $\widehat{\beta}_{CR}$ (0.726%), followed by $\widehat{\overline{\beta}}_{BN}$ (4.136%), $\widehat{\tilde{\beta}}_{BN}$ (4.152%), $\widehat{\breve{\beta}}_{BN}$ (4.587%) and $\widehat{\beta}$ (7.529%). Note that the relative bias of $\widehat{\beta}$ is nearly ten times larger than that of $\widehat{\beta}_{CR}$. It is also noteworthy that the Cox and Reid estimator has the smallest mean squared error (0.920).

Table 1. Point estimation, noncensoring, $n = 20$.

estimator	mean	variance	bias	MSE	RB(%)	skewness	kurtosis
$\widehat{\beta}$	5.376	1.049	0.376	1.191	7.529	0.928	4.708
$\widehat{\overline{\beta}}_{BN}$	5.207	0.983	0.207	1.025	4.136	0.922	4.694
$\widehat{\breve{\beta}}_{BN}$	5.229	0.993	0.229	1.046	4.587	0.930	4.710
$\widehat{\tilde{\beta}}_{BN}$	5.208	0.983	0.208	1.026	4.152	0.924	4.699
$\widehat{\beta}_{CR}$	5.036	0.919	0.036	0.920	0.726	0.920	4.689

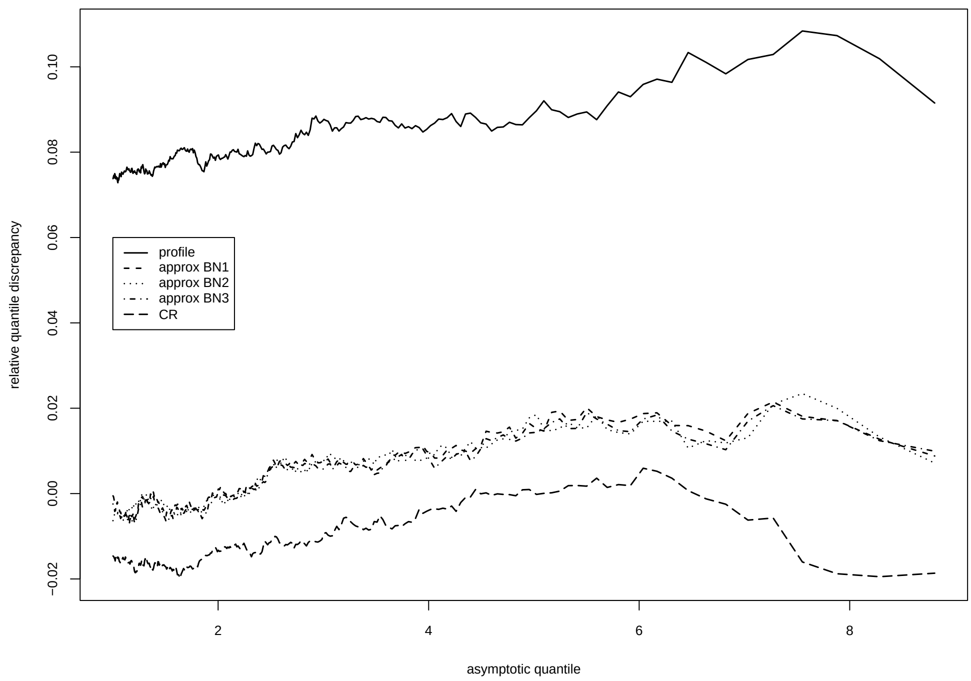
Table 2 gives the null rejection rates of the likelihood ratio test and of the corresponding tests from the adjusted likelihoods. More precisely, such tests are based on ℓ (4.1), $\overline{\ell}_{BN}$ (4.2), $\breve{\ell}_{BN}$ (4.3), $\tilde{\ell}_{BN}$ (4.4) and ℓ_{CR} (4.5). The figures in this table show that the tests derived from adjusted profile likelihoods ($\overline{\ell}_{BN}$, $\breve{\ell}_{BN}$, $\tilde{\ell}_{BN}$ e ℓ_{CR}) displayed smaller size distortions than the profile likelihood ratio test. No adjusted test clearly outperformed the others.

Table 2. Null rejection rates, noncensoring, $n = 20$.

nominal level	ℓ	$\overline{\ell}_{BN}$	$\breve{\ell}_{BN}$	$\tilde{\ell}_{BN}$	ℓ_{CR}
15%	16.575	15.003	14.949	14.987	14.747
10%	11.353	10.122	10.115	10.105	9.800
5%	6.020	5.113	5.105	5.090	4.937
1%	1.390	1.058	1.040	1.047	0.997
0.1%	0.164	0.099	0.097	0.097	0.087
0.05%	0.078	0.049	0.048	0.048	0.048

Figure 1 plots the relative quantile discrepancies versus the corresponding asymptotic quantiles. Relative quantile discrepancy is defined as the difference between exact (esti-

Figure 1. Relative quantile discrepancies plot, $\mathcal{H}_0 : \beta = 1$, noncensoring, $n = 20$.



mated by simulation) and asymptotic quantiles divided by the latter. The closer to zero the relative quantile discrepancy, the better the approximation of the exact null distribution of the test statistic by the limiting χ^2_1 distribution. The tests based on $\bar{\ell}_{BN}$ (approx BN1), $\breve{\ell}_{BN}$ (approx BN2) and $\tilde{\ell}_{BN}$ (approx BN3) display similar finite-sample behavior; their relative discrepancy curves are all close to a horizontal line at zero. It is also clear that the null distribution of the usual likelihood ratio statistic is poorly approximated by the limiting distribution.

5.2. TYPE II CENSORED DATA

Consider now type II censoring with $n = 20$ and $r = 15$ (number of failures). The figures in Table 3 correspond to the estimators obtained by maximizing the profile likelihoods derived in Section 4.2: $\widehat{\beta}$ (4.6), $\widehat{\breve{\beta}}_{BN}$ (4.7), $\widehat{\tilde{\beta}}_{BN}$ (4.8) e $\widehat{\beta}_{CR}$ (4.9). The figures show that the merits of $\widehat{\beta}_{CR}$ relative to the remaining estimators are even more pronounced here than in the noncensoring case both in terms of relative bias and mean squared error. For instance, the relative bias and mean squared errors of $\widehat{\beta}$ are 12.204% and 2.358, respectively, whereas for $\widehat{\breve{\beta}}_{BN}$, 10.201% and 2.176, and for $\widehat{\tilde{\beta}}_{BN}$, 7.951% and 1.993. The relative bias of $\widehat{\beta}$ is nearly 20 times larger than that of $\widehat{\beta}_{CR}$ (0.600%).

Results on hypothesis testing derived from ℓ (4.6), $\breve{\ell}_{BN}$ (4.7), $\tilde{\ell}_{BN}$ (4.8) and ℓ_{CR} (4.9)

Table 3. Point estimation, type II censoring, $\beta = 5$, $(n, r) = (20, 15)$.

estimator	mean	variance	bias	MSE	RB(%)	skewness	kurtosis
$\widehat{\beta}$	5.610	1.985	0.610	2.358	12.204	1.146	5.658
$\breve{\widehat{\beta}}_{BN}$	5.510	1.916	0.510	2.176	10.201	1.147	5.660
$\tilde{\widehat{\beta}}_{BN}$	5.398	1.835	0.398	1.993	7.951	1.143	5.646
$\widehat{\beta}_{CR}$	5.030	1.593	0.030	1.594	0.600	1.143	5.644

are given in Table 4. All tests based on adjusted profile likelihoods ($\breve{\ell}_{BN}$, $\tilde{\ell}_{BN}$ and ℓ_{CR}) outperform the usual likelihood ratio test, the one based on ℓ_{CR} displaying superior small sample behavior. Indeed, at the 5% nominal level the size distortion of the best performing test is nearly 40 times smaller than that of the worst performing one.

Table 4. Null rejection rates, type II censoring, $(n, r) = (20, 15)$.

nominal level	ℓ	$\breve{\ell}_{BN}$	$\tilde{\ell}_{BN}$	ℓ_{CR}
15%	17.649	16.463	15.385	14.894
10%	12.259	11.279	10.348	9.971
5%	6.500	5.845	5.248	5.040
1%	1.591	1.352	1.097	1.019
0.1%	0.193	0.155	0.116	0.110
0.05%	0.097	0.083	0.062	0.060

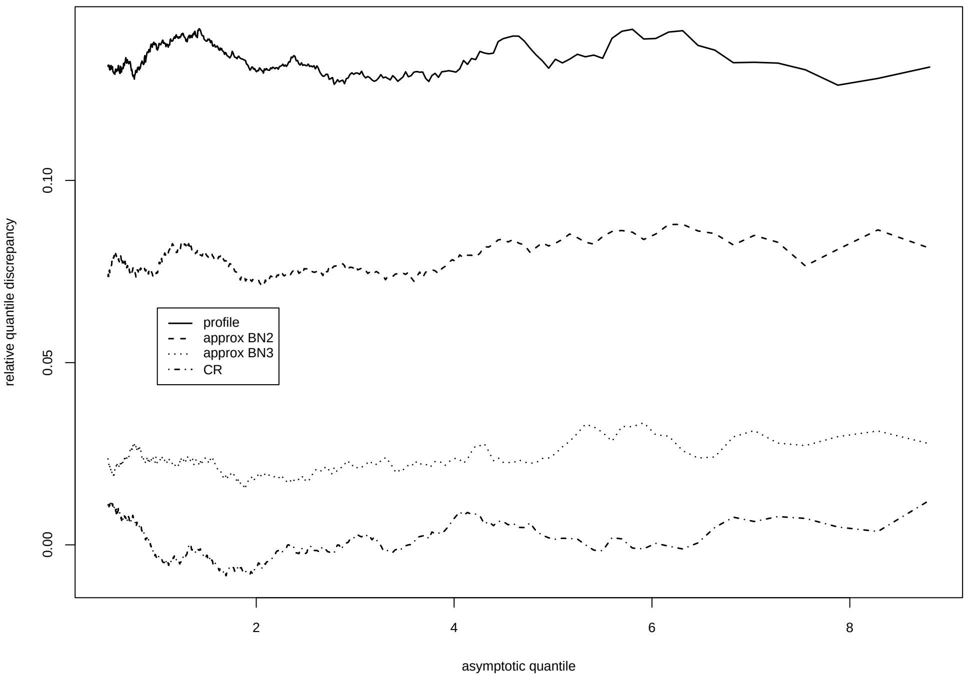
The relative quantile discrepancy plot under type II censoring is presented in Figure 2. Visual inspection of this figure shows that the best performing tests are the ones based on ℓ_{CR} and $\tilde{\ell}_{BN}$ (approx BN3), with slight advantage for the former.

We have also considered type II censoring with increased level of censoring: $(n, r) = (20, 10)$. For brevity, the results are not shown. The merits of the estimator obtained from $\widehat{\beta}_{CR}$ and of the corresponding test relative to the competitors are even more pronounced under heavier censoring. For example, the relative biases of $\breve{\widehat{\beta}}_{BN}$ and $\widehat{\beta}$ were 19.248% and 21.000%, respectively; they are more than 500 times larger than the relative bias of $\widehat{\beta}_{CR}$ (0.037%). The corresponding measure for $\tilde{\widehat{\beta}}_{BN}$ was 14.133%; this was the second best performing estimator. As for hypothesis testing, the best test was again the one based on ℓ_{CR}, followed by the tests obtained from $\tilde{\ell}_{BN}$, $\breve{\ell}_{BN}$ and ℓ_p. For instance, at the 5% nominal level the respective null rejection rates were 4.932%, 5.260%, 6.662% and 7.199%.

5.3. TYPE I CENSORED DATA

We shall now move to type I censoring. The censoring time was set at $c = \alpha(-\log p)^{1/\beta}$

Figure 2. Relative quantile discrepancies plot, $\mathcal{H}_0 : \beta = 1$, type II censoring, $(n, r) = (20, 15)$.



which yielded the desired probability of censoring p. Here, $p = 0.25, 0.50$. Tables 5 and 6 present results on estimation and testing for $p = 0.25$. The likelihood ratio tests are derived from ℓ (4.10), $\breve{\ell}_{BN}$ (4.11), $\tilde{\ell}_{BN}$ (4.12) and ℓ_{CR} (4.13). As before, the estimators obtained by maximizing these functions are denoted as $\widehat{\beta}$, $\widehat{\breve{\beta}}_{BN}$, $\widehat{\tilde{\beta}}_{BN}$ and $\widehat{\beta}_{CR}$, respectively. Again, the estimator $\widehat{\beta}_{CR}$ outperformed all others in terms of relative bias: 6.792% for $\widehat{\beta}$, 4.722% for $\widehat{\breve{\beta}}_{BN}$, 3.097% for $\widehat{\tilde{\beta}}_{BN}$ and 1.301% for $\widehat{\beta}_{CR}$. That is, the relative bias of the profile maximum likelihood estimator was over five times larger than that of the Cox and Reid estimator, and over twice that of $\widehat{\tilde{\beta}}_{BN}$. Additionally, the figures in Table 6 show that the usual likelihood ratio test has reliable small sample behavior, with little room left for improvement. The adjusted tests display slightly superior behavior. Overall, the results in Table 6 and in Figure 3 slightly favor the test obtained from $\breve{\ell}_{BN}$.

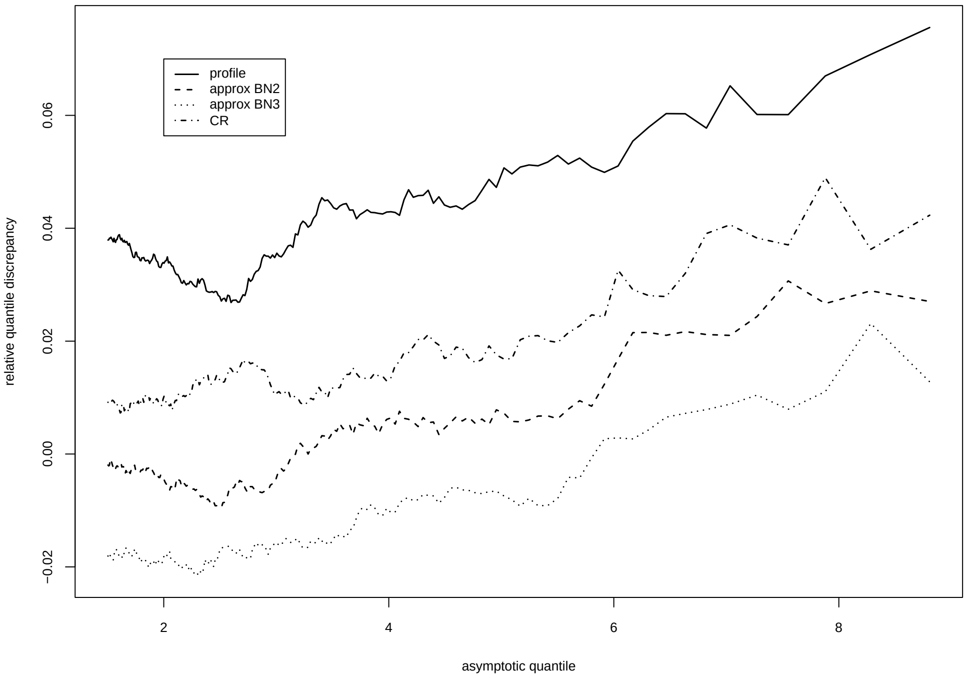
Table 5. Point estimation, type I censoring with $p = 0.25$, $\beta = 5.0$, $n = 20$.

estimator	mean	variance	bias	MSE	RB(%)	skewness	kurtosis
$\widehat{\beta}$	5.340	1.625	0.340	1.740	6.792	0.914	4.707
$\widehat{\breve{\beta}}_{BN}$	5.236	1.556	0.236	1.612	4.722	0.916	4.715
$\widehat{\tilde{\beta}}_{BN}$	5.155	1.526	0.155	1.550	3.097	0.907	4.674
$\widehat{\beta}_{CR}$	5.065	1.498	0.065	1.502	1.301	0.899	4.643

Table 6. Null rejection rates, type I censoring with $p = 0.25$, $n = 20$.

nominal level	ℓ	$\breve{\ell}_{BN}$	$\tilde{\ell}_{BN}$	ℓ_{CR}
15%	15.687	14.867	14.603	15.170
10%	10.465	9.899	9.684	10.254
5%	5.499	5.069	4.892	5.146
1%	1.226	1.073	1.033	1.104
0.1%	0.145	0.110	0.095	0.128
0.05%	0.072	0.055	0.053	0.057

Figure 3. Relative quantile plot, $\mathcal{H}_0 : \beta = 1$, type I censoring with $p = 0.25$, $n = 20$.



Under type I censoring with $p = 0.50$ (results not shown), the relative advantage of the Cox and Reid estimator was even more noticeable. The usual likelihood ratio test proved reliable, again leaving little room for improvement. For instance, at the 5% nominal level the null rejection rates of the tests based on ℓ_p, $\breve{\ell}_{BN}$, $\tilde{\ell}_{BN}$ and ℓ_{CR} were 5.066%, 4.830%, 4.422% and 4.883%, respectively. It is noteworthy that the test with poorest small sample behavior was the one obtained from $\tilde{\ell}_{BN}$, which was conservative.

5.4. DISCUSSION

We shall now compare some of our results to those presented by Yang and Xie (2003). We note that the empirical relative biases of $\widehat{\beta}$ and $\widehat{\beta}_{CR}$ obtained by these authors seem to depend on the true β value in contrast with our results. For example, for type II censored samples with $n = 20$ and $r = 10$ (see Table I in their paper), Yang and Xie (2003) obtained empirical relative biases of $\widehat{\beta}$ equal to 22.21% for $\beta = 5.0$ and 21.27% for $\beta = 0.5$; the corresponding figures for $\widehat{\beta}_{CR}$ were 1.03% and 0.52%, respectively. We obtained 21.000% and 0.037% for $\widehat{\beta}$ and $\widehat{\beta}_{CR}$, respectively, regardless of the value of β. A more pronounced difference in the empirical relative biases of $\widehat{\beta}_{CR}$ depending on the value of β is observed for $n = 30$ and $r = 15$ in Yang and Xie's (2003) simulations (-0.17% for $\beta = 3.0$ and 0.66% for $\beta = 4.0$). As another example, consider the type I censoring scheme with $n = 20$ and $p = 0.5$. Yang and Xie (2003; Table II) obtained empirical relative biases of $\widehat{\beta}$ equal to 10.87% for $\beta = 0.5$ and 10.37% for $\beta = 1.0$, the correspondig figures for $\widehat{\beta}_{CR}$ being 0.81% and 0.35%, respectively. We obtained 10.536% and 0.481% for $\widehat{\beta}$ and $\widehat{\beta}_{CR}$, respectively, regardless of the value of β. In fact, the results we obtained are expected since it can be shown that the distributions of $\widehat{\beta}/\beta$ and $\widehat{\beta}_{CR}/\beta$ do not depend on (α, β). The same holds for the other estimators (see Appendix).[3] We noticed that Yang and Xie (2003) do not mention the number of replications they used in their simulations. Perhaps the unexpected differences in the relative biases they report are due to a not large enough number of replications; as mentioned previously, our simulations are based on 100,000 replications. We note, however, that their simulation study is in line with ours in the sense that both lead to the conclusion that estimation of β based on the Cox and Reid adjusted profile likelihood is more accurate than the usual maximum likelihood estimation. Finally, it is noteworthy that our simulations include results on hypothesis testing while their numerical results only cover parameter estimation.

6. NUMERICAL EXAMPLES

We shall now consider two well know data sets obtained from Lawless (1982). In both cases, we assume that the data are independent and follow a Weibull distribution. We perform inference on the shape parameter β using the profile and the modified profile likelihoods derived in Section 4.

The following data are times between successive failures of air conditioning equipment in a Boeing 720 airplane (Proschan, 1963, Lawless, p. 134): 74, 57, 48, 29, 502, 12, 70, 21, 29, 386, 59, 27, 153, 26 and 326. Here, $n = 15$ and there is no censoring. The null hypothesis of interest is $\mathcal{H}_0 : \beta = 1$. The likelihood ratio statistics based on ℓ (4.1), $\overline{\ell}_{BN}$ (4.2), $\breve{\ell}_{BN}$ (4.3), $\tilde{\ell}_{BN}$ (4.4) and ℓ_{CR} (4.5) are, respectively, 0.409, 0.658, 0.655, 0.682 and

[3] In the Appendix, we also show that the size properties of the usual and the adjusted profile likelihood ratio tests depend neither on α nor on the value of β set at the null hypothesis.

1.033. At the 10% significance level, none of the tests reject the null hypothesis. That is, all five tests point to the same conclusion, namely that the exponential model cannot be rejected as the data generating mechanism. The maximum likelihood estimates obtained by maximizing the profile and modified profile likelihods are $\widehat{\beta} = 0.888$, $\widehat{\widehat{\beta}}_{BN} = 0.858$, $\breve{\widehat{\beta}}_{BN} = 0.857$, $\tilde{\widehat{\beta}}_{BN} = 0.856$ and $\widehat{\beta}_{CR} = 0.823$.

We shall now turn to our second numerical example. Mann and Fertig (1973) — see also Lawless (1982, p. 86) — present failure times of airplane components subjected to a life test. The data are type II censored: 13 components were placed under test until the tenth failure took place. The observed failure times (in hours) are 0.22, 0.50, 0.88, 1.00, 1.32, 1.33, 1.54, 1.76, 2.50, 3.00. Here, the likelihood ratio statistics for the test of $\mathcal{H}_0 : \beta = 1$ obtained from ℓ (4.6), $\breve{\ell}_{BN}$ (4.7), $\tilde{\ell}_{BN}$ (4.8) and ℓ_{CR} (4.9) are 1.435, 1.160, 0.960 and 0.354, respectively. Regardless of the statistic chosen for the test, the null hypothesis is not rejected at the usual significance levels, that is, one cannot reject the exponential model. Consider, however, the null hypothesis $\mathcal{H}_0 : \beta = 2$. The tests based on ℓ, $\breve{\ell}_{BN}$ and $\tilde{\ell}_{BN}$ do not reject $\mathcal{H}_0$ at the 10% significance level, unlike the test based on the Cox and Reid (1987) modified profile likelihood ℓ_{CR}. The Cox and Reid (1987) test is the only that distinguishes between $\beta = 1$ and $\beta = 2$. The parameter estimates obtained by maximizing the profile and modified profile likelihoods are $\widehat{\beta} = 1.417$, $\breve{\widehat{\beta}}_{BN} = 1.374$, $\tilde{\widehat{\beta}}_{BN} = 1.343$ and $\widehat{\beta}_{CR} = 1.204$.

7. CONCLUDING REMARKS

Inference on the Weibull shape parameter is particularly important when modeling failure data, since the dependence of the hazard function on time is controled by such a parameter. Yang and Xie (2003) obtained an adjustment to the profile likelihood function using the approach developed by Cox and Reid (1987, 1989). Their numerical results showed that estimation was considerably more precise when performed via the Cox and Reid likelihood function both without and with (types I and II) censoring. We have derived alternative adjustments to the Weibull shape parameter profile likelihood function. Our numerical results, which cover both estimation and testing under noncensoring and (types I and II) censoring, show that inference based on the adjusted profile likelihood of Cox and Reid outperforms not only the usual profile likelihood inference, as suggested by Yang and Xie (2003), but also inference yielded by alternative adjusted profile likelihoods. The relative advantage of the Cox and Reid approach is more pronounced when used for point estimation than for hypothesis testing; it is also more pronounced under noncensoring and type II censoring than under type I censoring. Overall, the results in this paper strengthen those in Yang and Xie (2003). Indeed, the results in our paper and in theirs taken together strongly suggest that practitioners should base inference on the Weibull shape parameter on the Cox and Reid adjusted profile likelihood functions obtained through the orthogonal parameterizations presented in Section 4. In future research, we shall derive some of the adjusted profile likelihoods described in this paper in Weibull regression models. Also, it will be interesting to extend our results to extensions of the traditional two-parameter Weibull distribution. Lai et al. (2004) describe in a unified manner some of these extensions; see also Murthy et al. (2003), Xie et al. (2003) and Xie et al. (2002).

APPENDIX

In this appendix we show that, for all estimators of β considered in this paper and for all sampling schemes (complete samples and types I and II censoring), the distribution of the ratio 'estimator/β' does not depend on (α, β), i.e., it is a pivotal quantity. We also show that the size properties of the usual and the adjusted profile likelihood ratio tests depend neither on α nor on the value of β set at the null hypothesis. We thus generalize the results obtained by Thoman et al. (1969) which are confined to maximum likelihood estimation in complete samples. In what follows we shall use the fact that if $y_1, \ldots, y_n$ are i.i.d. random variables having a $W(\alpha, \beta)$ distribution, then $x_j = x_j(y_j, \alpha, \beta) = (y_j/\alpha)^\beta$, for $j = 1, \ldots, n$, are independently distributed with a standard exponential distribution. Note that $(x_{(1)}, \ldots, x_{(n)})$, the vector of order statistics relative to $(x_1, \ldots, x_n)$, is a pivotal quantity.

Consider the type II censoring scheme with n observations and r failures (for noncensored samples, let $n = r$). The profile score function for β is

$$u_p(\beta) = \frac{r}{\beta} - r\frac{\sum_{j=1}^r y_{(j)}^\beta \log y_{(j)} + (n-r)y_{(r)}^\beta \log y_{(r)}}{\sum_{j=1}^r y_{(j)}^\beta + (n-r)y_{(r)}^\beta} + \sum_{j=1}^r \log y_{(j)}.$$

The maximum likelihood estimator of β satisfies $u_p(\beta) = 0$, i.e.

$$\frac{r}{\widehat{\beta}} - r\frac{\sum_{j=1}^r y_{(j)}^{\widehat{\beta}} \log y_{(j)} + (n-r)y_{(r)}^{\widehat{\beta}} \log y_{(r)}}{\sum_{j=1}^r y_{(j)}^{\widehat{\beta}} + (n-r)y_{(r)}^{\widehat{\beta}}} + \sum_{j=1}^r \log y_{(j)} = 0.$$

Since $y_j = \alpha x_j^{1/\beta}$, for $j = 1, \ldots, n$, then $y_{(j)} = \alpha x_{(j)}^{1/\beta}$ and $\log y_{(j)} = \log\alpha + (1/\beta)\log x_{(j)}$. Replacing these relations in the above equation and multiplying both sides by β, we get

$$\frac{r}{\widehat{\beta}/\beta} - r\frac{\sum_{j=1}^r x_{(j)}^{\widehat{\beta}/\beta} \log x_{(j)} + (n-r)x_{(r)}^{\widehat{\beta}/\beta} \log x_{(r)}}{\sum_{j=1}^r x_{(j)}^{\widehat{\beta}/\beta} + (n-r)x_{(r)}^{\widehat{\beta}/\beta}} + \sum_{j=1}^r \log x_{(j)} = 0. \qquad \text{(A.1)}$$

The solution of (A.1) for $\widehat{\beta}/\beta$ depends on the data and the parameters only through the vector $(x_{(1)},\ldots,x_{(n)})$, a pivotal quantity, and hence the distribution of $\widehat{\beta}/\beta$ does not depend on (α,β). Analogously, the likelihood ratio statistic may be written as

$$LR(\beta) = 2\Bigg\{ r\log(\widehat{\beta}/\beta) - r\log\left[\sum_{j=1}^{r} x_{(j)}^{\widehat{\beta}/\beta} + (n-r)x_{(r)}^{\widehat{\beta}/\beta}\right] + \left(\frac{\widehat{\beta}}{\beta}-1\right)\sum_{j=1}^{r}\log x_{(j)} \qquad \text{(A.2)}$$

$$+r\log\left[\sum_{j=1}^{r} x_{(j)} + (n-r)x_{(r)}\right]\Bigg\},$$

and it is then clear that $\Pr_{(\alpha,\beta)}(LR(\beta)\leq q)$ is the same for all (α,β). Notice that this result implies that the null distribution of the likelihood ratio statistic does not depend on α and is the same regardless the value of β set at the null hypothesis.

For the Cox and Reid adjusted profile likelihood inference the estimating equation is given by (A.1) with r replaced by $r-2$ in the first term and $\widehat{\beta}$ replaced by $\widehat{\beta}_{CR}$. Also, the corresponding likelihood ratio statistic is given by (A.2) with r replaced by $r-2$ in the first term between braces and $\widehat{\beta}$ replaced by $\widehat{\beta}_{CR}$. Hence, $\widehat{\beta}_{CR}/\beta$ is distributed independently of (α,β). Also, the null distribution of the Cox and Reid adjusted profile likelihood ratio statistic does not depend on α and does not change with changes in the value of the shape parameter set at the null hypothesis.

The same results can be proved for the other estimators and statistics considered in this paper under type II censoring in a similar way. In the case of the adjusted profile likelihood $\overline{\ell}_{BN}$ (noncensored data) the corresponding likelihood ratio statistic involves $\widehat{\beta}\log(\widehat{\alpha}/\alpha)$. It can be shown, however, that it is distributed independently of (α,β) (see Thoman et al. 1969, Theorem B). To save space we omit the proofs, but they may be obtained from the authors upon request.

We now move to type I censoring. The profile score function for β is

$$u_p(\beta) = \frac{r}{\beta} - \frac{r\sum_{j=1}^{n} y_j^{\beta}\log y_j}{\sum_{j=1}^{n} y_j^{\beta}} + \sum_{j=1}^{n}\delta_j\log y_j.$$

The maximum likelihood estimator of β is the solution of $u_p(\beta)=0$, and hence satisfies

$$\frac{r}{\widehat{\beta}} - \frac{r\sum_{j=1}^{n} y_j^{\widehat{\beta}}\log y_j}{\sum_{j=1}^{n} y_j^{\widehat{\beta}}} + \sum_{j=1}^{n}\delta_j\log y_j = 0,$$

which can be rewritten as

$$\frac{r}{\widehat{\beta}} - r\frac{\sum_{j=1}^{n}\delta_j t_j^{\widehat{\beta}}\log t_j + (n-r)c^{\widehat{\beta}}\log c}{\sum_{j=1}^{n}\delta_j t_j^{\widehat{\beta}} + (n-r)c^{\widehat{\beta}}} + \sum_{j=1}^{n}\delta_j\log t_j = 0.$$

Note that if $\delta_j = 1$, then $t_j = \alpha x_j^{1/\beta}$, and recall that we set $c = \alpha(-\log p)^{1/\beta}$ in our simulations. Replacing these relations in the above equation and multiplying both sides by β, we have

$$\frac{r}{\widehat{\beta}/\beta} - r\frac{\sum_{j=1}^{n}\delta_j x_j^{\widehat{\beta}/\beta}\log x_j + (n-r)(-\log p)^{\widehat{\beta}/\beta}\log(-\log p)}{\sum_{j=1}^{n}\delta_j x_j^{\widehat{\beta}/\beta} + (n-r)(-\log p)^{\widehat{\beta}/\beta}} + \sum_{j=1}^{n}\delta_j \log x_j = 0. \quad \text{(A.3)}$$

In (A.3), n and p are fixed constants, $\delta_1, \ldots, \delta_n$ are i.i.d. random variables having a Bernoulli distribution with parameter $1-p$ and $r = \sum_{j=1}^{n}\delta_j$. Also, as mentioned before, the distribution of $(x_1, \ldots, x_n)$ does not depend on (α, β). Hence, the solution of (A.3) for $\widehat{\beta}/\beta$ is distributed independently of such parameters.

Analogously, the likelihood ratio statistic can be written as

$$LR(\beta) = 2\Bigg\{r\log(\widehat{\beta}/\beta) - r\log\left[\sum_{j=1}^{n}\delta_j x_j^{\widehat{\beta}/\beta} + (n-r)(-\log p)^{\widehat{\beta}/\beta}\right] \quad \text{(A.4)}$$

$$+r\log\left[\sum_{j=1}^{n}\delta_j x_j + (n-r)(-\log p)\right] + \left(\frac{\widehat{\beta}}{\beta} - 1\right)\sum_{j=1}^{n}\delta_j \log x_j\Bigg\}.$$

Using the same arguments as before, it can be seen that $\Pr_{(\alpha,\beta)}(LR(\beta) \leq q)$ does not change with changes in (α, β)

The results for the other adjusted profile likelihoods can be obtained in a similar fashion. For instance, for the Cox and Reid adjusted profile likelihood, the results follow by showing that the corresponding estimating equation is given by (A.3) with r replaced by $r-1$ in the first term and $\widehat{\beta}$ by $\widehat{\beta}_{CR}$. Moreover, the resulting likelihood ratio statistic is given by (A.4) with r replaced by $r-1$ in the first term between braces and $\widehat{\beta}$ replaced by $\widehat{\beta}_{CR}$. The omitted proofs may be obtained from the authors upon request.

ACKNOWLEDGEMENTS

We gratefully acknowledge partial financial support from CNPq and FAPESP. We also thank an anonymous referee for comments on an earlier draft.

REFERENCES

[1] Ageel, M. I. (2002). A novel means of estimating quantiles for 2-parameter Weibull distribution under the right random censoring model. *Journal of Computational and Applied Mathematics*, **149**, 373–380.

[2] Bar-Lev, S. K. (2004). Likelihood–based inference for the shape parameter of a two parameter Weibull distribution. *Lifetime Data Analysis*, **10**, 293–308.

[3] Barndorff-Nielsen, O.E. (1980). Conditionality resolutions. *Biometrika*, **67**, 293–310.

[4] Barndorff-Nielsen, O.E. (1983). On a formula for the distribution of the maximum likelihood estimator. *Biometrika*, **70**, 343–365.

[5] Cox, D.R. and Reid, N. (1987). Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society B*, **49**, 1–39.

[6] Cox, D.R. and Reid, N. (1989). On the stability of maximum-likelihood estimators of orthogonal parameters. *Canadian Journal of Statistics*, **17**, 229–233.

[7] Doornik, J.A. (2001). *Ox: an Object-oriented Matrix Programming Language*, 4th edition. London: Timberlake Consultants and Oxford: `http://www.doornik.com`.

[8] Fraser, D. A. S. and Reid, N. (1995). Ancillaries and third-order significance. *Utilitas Mathematica*, **47**, 33–53.

[9] Fraser, D. A. S., Reid, N. and Wu, J. (1999). A simple formula for tail probabilities for frequentist and Bayesian inference. *Biometrika*, **86**, 655–661.

[10] Gross, A.J. and Clark, V.A. (1975). *Survival Distribution: Reliability Applications in the Biomedical Sciences*. New York: Wiley.

[11] Klein, J.P. and Moeschberger, M. (1997). *Survival Analysis*. New York: Springer–Verlag.

[12] Lai, C. D., Zhang, L. and Xie, M. (2004). Mean residual life and other properties of Weibull related bathtub shape failure rate distributions. *International Journal of Reliability, Quality and Safety Engineering*, **11**, 113-132.

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

[14] Lu, H. L., Chen, C. H. and Wu, J. W. (2004). A note on weighted least-squares estimation of the shape parameter of the Weibull distribution. *Quality and Reliability Engineering International*, **20**, 579–586.

[15] Mann, N. R. and Fertig, K. W. (1973). Tables for obtaining confidence bounds and tolerance bounds based on best linear invariant estimates of parameters of the extreme value distribution. *Technometrics*, **15**, 87–101.

[16] Mann, N.R., Schafer, R.E. and Singpurwalla, N.D. (1974). *Methods for Statistical Analysis and Reliability and Life Data*. New York: Wiley.

[17] Maswadah, M. (2003). Conditional confidence interval estimation for the inverse Weibull distribution based on censored generalized order statistics. *Journal of Statistical Computation and Simulation*, **73**, 887–898.

[18] Murthy, D. N. P., Xie, M. and Jiang, R. (2003). *Weibull Models*. New York: Wiley.

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

[20] Severini, T.A. (1998). An approximation to the modified profile likelihood function. *Biometrika*, **85**, 403–411.

[21] Severini, T.A. (1999). An empirical adjustment to the likelihood ratio statistic. *Biometrika*, **86**, 235–247.

[22] Severini, T.A. (2000). *Likelihood Methods in Statistics*. Oxford: Oxford University Press.

[23] Thoman, D.R., Bain, L.J. and Antle, C.E. (1969). Inferences on the parameters of the Weibull distribution. *Technometrics*, **11**, 445–460.

[24] Xie, M., Goh, T. N. and Tang, Y. (2002). A modified Weibull extension with bathtub-shapedfailure rate function. *Reliability Engineering and Systems Safety*, **76**, 279–285.

[25] Xie, M., Lai, C. D. and Murthy, D. N. P. (2003). Weibull-related Distributions with Bathtub Shaped Failure Rate Functions. In: Mathematical and Statistical Methods in Reliability, vol. 7, eds. K. Doksum and B. Lindqvist, Singapore: World Scientific Publishing Co., 283–297.

[26] Yang, Z. and Xie, M. (2003). Efficient estimation of the Weibull shape parameter based on a modified profile likelihood. *Journal of Statistical Computation and Simulation*, **73**, 115–123.