Stat Papers (2010) 51:325–336
DOI 10.1007/s00362-008-0172-x

NOTE

Testing for the Marshall–Olkin extended form of the Weibull distribution

Chrys Caroni

Received: 8 October 2007 / Revised: 8 September 2008 / Published online: 27 September 2008

Abstract Marshall–Olkin extended distributions offer a wider range of behaviour than the basic distributions from which they are derived and therefore may find applications in modeling lifetime data, especially within proportional odds models, and elsewhere. The present paper carries out a simulation study of likelihood ratio, Wald and score tests for the parameter that distinguishes the extended distribution from the basic one, for the Weibull and exponential cases, allowing for right censored data. The likelihood ratio test is found to perform better than the others. The test is shown to have sufficient power to detect alternatives that correspond to interesting departures from the basic model and can be useful in modeling.

Keywords Weibull distribution · Marshall–Olkin extension · Proportional odds · Likelihood

1 Introduction

The basic distributions used widely in reliability and survival analysis (Weibull, Gamma, log normal and others) have a limited range of behaviour and cannot represent all the situations found in applications. For example, although the Weibull is often described as flexible, its hazard function is in fact restricted to being monotonically increasing or monotonically decreasing, or constant. The limitations of standard distributions led naturally to interest in developing extended or generalized distributions by adding further parameters. One such system of extended distributions was introduced by Marshall and Olkin (1997). Let $\bar{F}(x)$ denote the reliability (or survival) function of

C. Caroni (✉)
National Technical University of Athens, 9 Iroon Polytechniou St, 157 80 Athens, Greece
e-mail: ccar@math.ntua.gr

the continuous random variable X. Then the corresponding Marshall–Olkin extended distribution has reliability function

$$\bar{G}(x;\alpha) = \frac{\alpha \bar{F}(x)}{1 - \bar{\alpha}\bar{F}(x)}, \quad -\infty < x < \infty, \ 0 < \alpha < \infty \tag{1}$$

where $\bar{\alpha} = 1 - \alpha$. Particular cases that have been considered in the literature include the Marshall–Olkin extensions of the Weibull distribution (Ghitany et al. 2005; Zhang and Xie 2007), Pareto distribution (Ghitany 2005), Lomax distribution (Ghitany et al. 2007) and linear failure-rate distribution (Ghitany and Kotz 2007).

Marshall and Olkin (1997) remarked that distributions of this extended form have a stability property, in the sense that applying the transformation (1) to $\bar{G}$ gives nothing new—the resulting reliability function is still defined by (1) in terms of $\bar{F}(x)$ but with a different value of the parameter α. From this, Economou and Caroni (2007) show that Marshall–Olkin extended distributions have a proportional odds property. The proportional odds model (Bennett 1983) for representing the effect of covariates on survival is

$$\frac{1 - S(t;\mathbf{x})}{S(t;\mathbf{x})} = g(\mathbf{x}) \frac{1 - S_b(t)}{S_b(t)}, \quad t > 0 \tag{2}$$

where $S(t;\mathbf{x})$ is the survival function of units with covariates $\mathbf{x}$, $S_b(t)$ is some baseline survival function and $g(\mathbf{x})$ is a non-negative function of the covariates, such as $g(\mathbf{x}) = \exp(\mathbf{x}'\boldsymbol{\beta})$. A similar proportional odds frailty model can represent heterogeneity by introducing unobserved frailty Z in place of (or possibly as well as) the observed covariates $\mathbf{x}$ (Economou and Caroni 2007):

$$\frac{1 - S(t|z)}{S(t|z)} = z \frac{1 - S_b(t)}{S_b(t)}. \tag{3}$$

This gives the odds on failure by time t of an individual with frailty $Z = z$, in terms of the baseline odds corresponding to individuals with frailty Z equal to the expected population frailty $E(Z)$, assumed to be 1. It follows that

$$S(t|z) = \frac{\frac{1}{z} S_b(t)}{1 - \left(1 - \frac{1}{z}\right) S_b(t)} \tag{4}$$

which has the same form as Marshall and Olkin's formula, with $\alpha = 1/z$, $\bar{F}(t) = S_b(t)$ and $\bar{G}(t;\alpha) = S(t|z)$ in (1). Thus all Marshall–Olkin extended distributions have the proportional odds property (Economou and Caroni 2007) and will be particularly useful in proportional odds models. It is easy to see that they do not have proportional hazards or accelerated life properties, so they would not be useful when applying the corresponding models.

The purpose of the present article is to consider how to test for the extra parameter α in (1) in order to confirm that the Marshall–Olkin extended distribution is needed, instead of the basic distribution from which it is derived. The Weibull distribution and its special case, the exponential distribution, will be examined. Only tests based on likelihood functions will be considered (likelihood ratio, Wald, score tests). All the applications of Marshall–Olkin extended distributions that were cited above employed maximum likelihood estimation and carried out likelihood ratio tests. Zhang and Xie (2007) also suggested a graphical method based on the probability plot.

2 Maximum likelihood estimation

Assume a general uninformative right-censoring scheme in a simple random sample of n independent observations, giving data $\{(x_i, \delta_i), i = 1, 2, \ldots, n\}$ where δ is the usual indicator taking the value one for an uncensored observation and zero for a right-censored one. From (1) and the corresponding expression for the pdf $g(x)$ of the Marshall–Olkin extended distribution, the log likelihood of the sample is

$$n \ln \alpha + \sum \left[\delta_i \ln f(x_i) + (1 - \delta_i) \ln \bar{F}(x_i) - (1 + \delta_i) \ln \left(1 - \bar{\alpha} \bar{F}(x_i)\right)\right] \tag{5}$$

where $f(x)$ is the pdf of the original, unmodified distribution. In the case of the exponential distribution and its Marshall-Olkin extension, the log likelihood is

$$n \ln \alpha + n_U \ln \lambda - \lambda \sum x_i - \sum (1 + \delta_i) \ln \left(1 - \bar{\alpha} e^{-\lambda x_i}\right) \tag{6}$$

where n_U is the number of uncensored observations. Maximum likelihood estimates of the parameters are obtained by solving numerically the usual system of equations involving the first deriviatives.

In the case of the Weibull and its extension, if the pdf of the basic 2-parameter Weibull distribution with positive scale and shape parameters λ and β, respectively, is expressed as

$$\beta \lambda^\beta x^{\beta - 1} \exp\left[-(\lambda x)^\beta\right], \quad x > 0 \tag{7}$$

then the 3-parameter extended Weibull distribution has pdf

$$g(x) = \frac{\alpha \beta \lambda (\lambda x)^{\beta - 1} \exp[-(\lambda x)^\beta]}{\left\{1 - \bar{\alpha} \exp[-(\lambda x)^\beta]\right\}^2}, \quad x > 0, \quad \alpha, \beta, \lambda > 0, \quad \bar{\alpha} = 1 - \alpha \tag{8}$$

and survival function

$$\bar{G}(x) = \frac{\alpha \exp[-(\lambda x)^\beta]}{1 - \bar{\alpha} \exp[-(\lambda x)^\beta]}. \tag{9}$$

The log-likelihood from Eqs. (8) and (9) is

$$\ell(\alpha, \beta, \lambda) = n \ln \alpha + n_U \{\ln \beta + \beta \ln \lambda\} + (\beta - 1) \sum_U \ln x_i \\ - \sum \left\{ (\lambda x_i)^\beta + (\delta_i + 1) \ln\{1 - \bar{\alpha} \exp[-(\lambda x_i)^\beta]\} \right\} \tag{10}$$

where U on the summation denotes that it is over the uncensored values. The first derivatives are written out in Ghitany et al. (2005) and will not be repeated here. Zhang and Xie (2007) present the likelihood equations for Type II censoring.

3 Test statistics

The null value of α corresponding to the basic distribution F is $\alpha = 1$. Since $\alpha > 0$, this means that the null value is an interior point of the parameter space so likelihood-based tests should have regular behaviour.

In general, suppose that the parameter vector of any Marshall–Olkin extended distribution is partitioned as (α, ϕ), where ϕ is the vector of the parameters of the basic distribution. Standard likelihood theory states that the null hypothesis of $\alpha = \alpha_0$ can be tested using a likelihood ratio test by comparing the statistic

$$-2 \left\{ \ell(\alpha_0, \hat{\phi}_0) - \ell(\hat{\alpha}, \hat{\phi}) \right\} \tag{11}$$

against the χ_1^2 distribution, where ℓ is the log-likelihood, $\hat{\phi}_0$ is the maximum likelihood estimate of ϕ under the null hypothesis $\alpha = \alpha_0$ and $\hat{\alpha}$, $\hat{\phi}$ are maximum likelihood estimates under the alternative hypothesis. Other tests based on likelihood theory include the Wald and score (or Lagrange multiplier) tests. Let

$$I = \begin{bmatrix} I_{\alpha\alpha} & I_{\alpha\phi} \\ I_{\phi\alpha} & I_{\phi\phi} \end{bmatrix} \tag{12}$$

be the information matrix partitioned in the corresponding way to the parameter vector, and let

$$I^{-1} = \begin{bmatrix} I^{\alpha\alpha} & I^{\alpha\phi} \\ I^{\phi\alpha} & I^{\phi\phi} \end{bmatrix} \tag{13}$$

be its partitioned inverse.

Then the Wald test statistic is

$$\left(\hat{\alpha} - \alpha_0\right)^2 / I^{\alpha\alpha}(\hat{\alpha}, \hat{\phi}) \tag{14}$$

and the score test statistic is

$$U_\alpha(\alpha_0, \hat{\phi}_0)^2 I^{\alpha\alpha}(\alpha_0, \hat{\phi}_0) \tag{15}$$

where $U_\alpha = \partial\ell/\partial\alpha$. Like the likelihood ratio statistic, both can be compared to the χ_1^2 distribution which holds asymptotically.

All three types of test are widely used. The Wald and score tests are often said to have the advantage, compared to the likelihood ratio test, that they only require the fitting of one model (under the alternative hypothesis for Wald and under the null for score) but with modern computing facilities this is not such an important factor as it used to be. Both have the potential disadvantage, compared to the likelihood ratio test, of being based on approximations to the likelihood surface (Buse 1982) and this may affect the applicability of the asymptotic results in small samples. The Wald test has the additional disadvantage of being dependent on the parametrisation of the model, unlike the other two tests. Thus the likelihood ratio may be preferable in general (Lindsey 1996), but this needs to be examined in detail in each application. Consequently, a simulation study was carried out firstly to examine how closely the null distributions of the three test statistics follow the χ_1^2 distribution in the case of the Weibull and exponential distributions, and secondly to estimate their powers in detecting departures from the null hypothesis.

4 Simulation results: exponential distribution

For each combination of sample size, parameter values and censoring scheme, 10 000 simulated samples were generated. This gives a binomial standard error of about 0.22 for an estimated percentage equal to about 5%. The scale parameter of the exponential distribution was taken equal to one without loss of generality. Right censoring was introduced when required by censoring all generated data values at a fixed value (Type I censoring). For studying the power, data were generated from the Marshall–Olkin extended exponential distribution using the inversion method, which gives

$$x = -\frac{1}{\lambda} \ln \left(u/(\alpha + u\bar{\alpha})\right) \tag{16}$$

where u is a pseudorandom random deviate generated from $U[0, 1]$. The simulations were carried out in Microsoft Fortran using IMSL Library routines. Pseudorandom numbers were generated by the routines RNUN and RNEXP for the uniform and exponential distributions, respectively. Function minimisation was performed by the routine BCOAH which employs a modified Newton method with user-supplied Hessian.

Figure 1 shows the simulated size of the three tests for the null hypothesis $\alpha = 1$ in the extended exponential distribution, at the nominal (χ_1^2) 5 and 1 % levels, for uncensored samples of up to n $=$ 150 observations. At the nominal 5% level, the LR and Wald tests have quite accurate actual sizes for $n \geq 30$, but the score test is liberal and remains so even for $n = 150$. At the 1% level, the LR test is again accurate but now the Wald test is liberal, and the score test even more so. Thus, only the LR test seems to be satisfactory overall. Similar conclusions hold for right-censored data. With about 25% Type I censoring, the simulated sizes of the tests for $n = 50$ were LR 4.79%, score 7.92% and Wald 6.54%, and for $n = 100$, LR 5.24%, score 7.88% and Wald 6.36%. Thus the score test seemed to be the most strongly affected by censoring. Further results for the exponential distribution and its Marshall–Olkin extension will be given

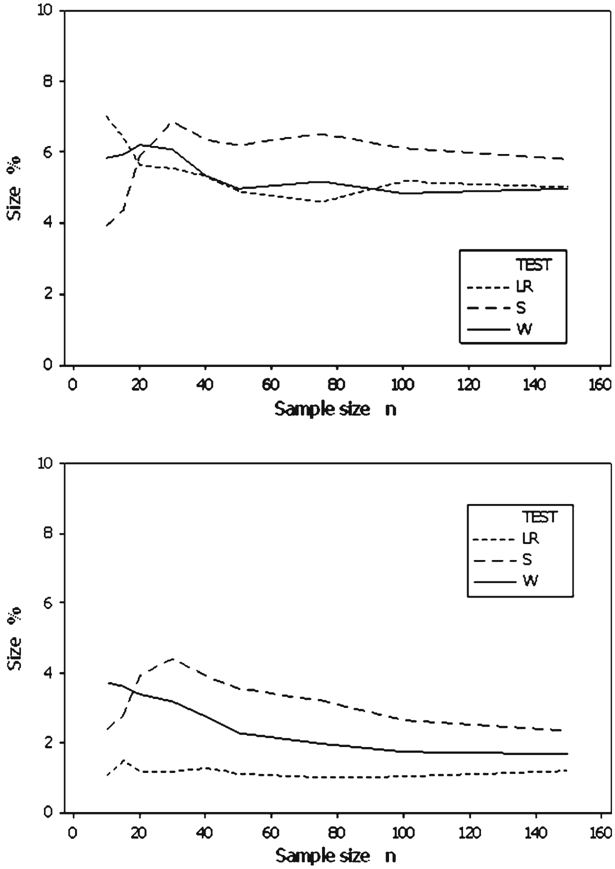

Fig. 1 Simulated sizes at nominal 5% (*upper plot*) and 1% (*lower plot*) levels of likelihood ratio (LR), Wald (W) and score (S) tests for the null hypothesis $\alpha = 1$ in the Marshall–Olkin extended exponential distribution with uncensored data. Exponential parameter $\lambda = 1$

for the likelihood ratio test only, since it is clearly superior to the other tests. The size of the LR test was not noticeably affected by up to about 35% of Type I censoring, and even with 50% censoring the simulated size at the nominal 5% level was still about 4%.

Figure 2 shows the simulated power of the LR test at the nominal 5% level as a function of α, for sample sizes of 20, 50 and 100. Power increases steeply away from $\alpha = 1$ for the larger sample sizes but, as could be expected, power is low for a small sample size ($n = 20$). Figure 3 shows the effect of right censoring on power. There appears to be almost a linear decline in power in proportion to the percentage of right censoring.

To illustrate what these detectable departures from $\alpha = 1$ mean in terms of the hazard function, Fig. 4 plots the hazard functions for the values of α that are detectable with 80% power when the uncensored sample size is 50. These are quite different from the constant hazard $h(t) = 1$ of the unmodified exponential distribution. One has a very high hazard initially, decreasing very rapidly so that by $t = 1$ (the mean lifetime)

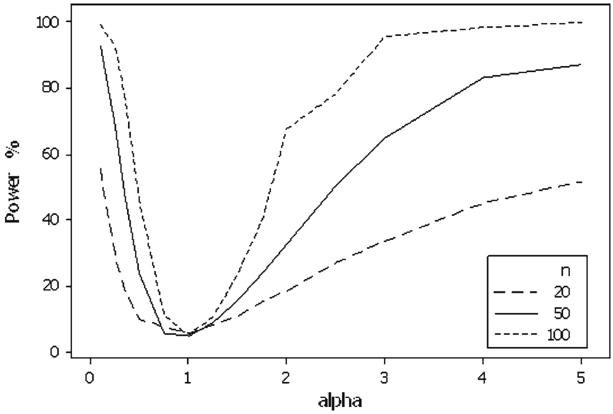

Fig. 2 Simulated power of likelihood ratio test at the nominal 5% level as a function of α in the Marshall–Olkin extended exponential distribution with uncensored data. Exponential parameter $\lambda = 1$

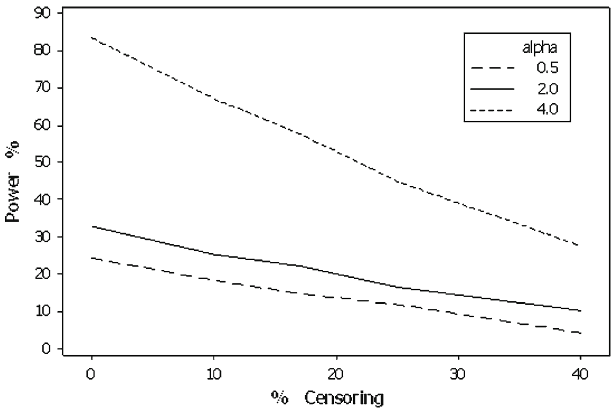

Fig. 3 Simulated power of likelihood ratio test at the nominal 5% level for selected values of alpha in the Marshall–Olkin extended exponential distribution, as a function of average percentage of right censoring in samples of size 50. Exponential parameter $\lambda = 1$

it is not much higher than the constant hazard. The other has a lower hazard initially, increasing steadily with time.

5 Simulation results: Weibull distribution

The simulation study of the Marshall–Olkin extension of the Weibull distribution was carried out along the same lines as for the exponential distribution. Pseudorandom deviates from the Weibull distribution were obtained from the IMSL routine RNWIB. Without loss of generality, the scale parameter was taken as one always. Applying the inversion method of generating a value x from the Marshall–Olkin extended Weibull

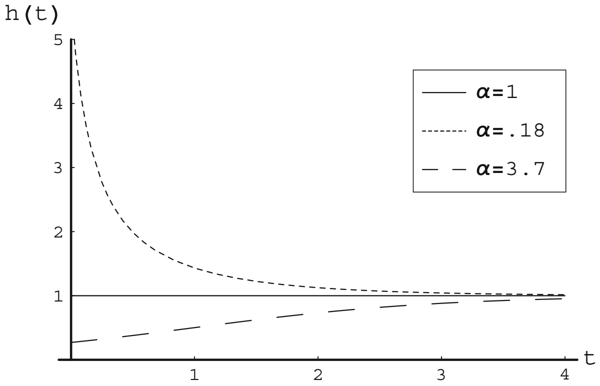

Fig. 4 Hazard functions of Marshall–Olkin extended exponential distributions for alpha values for which the likelihood ratio test has 80% power in uncensored samples of size 50. Exponential parameter $\lambda = 1$

distribution, gives

$$x = \frac{1}{\lambda}\{-\ln(u/(\alpha + u\bar{\alpha}))\}^{1/\beta} \tag{17}$$

where u is from $U[0, 1]$.

Figure 5 shows the simulated sizes of the three likelihood-based tests for the extended Weibull distribution. It was found empirically that these results did not depend on the value of the Weibull shape parameter. The LR and Wald tests are slightly conservative even for sample sizes as big as 100. The Wald test is very conservative in small samples and is still clearly inferior to the LR test in moderately large samples (about $n = 75$). For larger samples its performance is very close to the LR test at both the 1 and 5% levels. The score test is very liberal. Thus only the LR test is everywhere satisfactory and further results will be given for this test only.

Figure 6 shows that the size of the LR test is more sensitive to censoring than in the case of the exponential distribution. It is not affected by right censoring up to about 20% in samples of size $n = 100$, but starts to deteriorate rapidly (becoming very conservative) beyond that point. Figure 7 shows the power of the LR test at the nominal 5% level in uncensored samples of size $n = 100$ as a function of α. Note that in contrast to Fig. 2 for the exponential case, a logarithmic scale for α is used here. The power seems to be rather low for a moderately large sample size ($n = 100$). Figure 8 shows the steady decline in power as censoring increases. Similar results were obtained when the Weibull shape parameter was equal to 4, which we use in the following because of its relevance to the first of the applications that will be mentioned. To see what kind of changes in the distribution are implied by the values of α that have a reasonable chance of being detected, Fig. 9 plots the hazard functions of the extended Weibull distributions for the values of α for which the LR test has 80% power at the nominal 5% level in samples of size 250. Both functions are monotonically increasing, like the hazard function of the unmodified Weibull distribution, which may explain the moderate power of the test. But more extreme values of α can give rise to

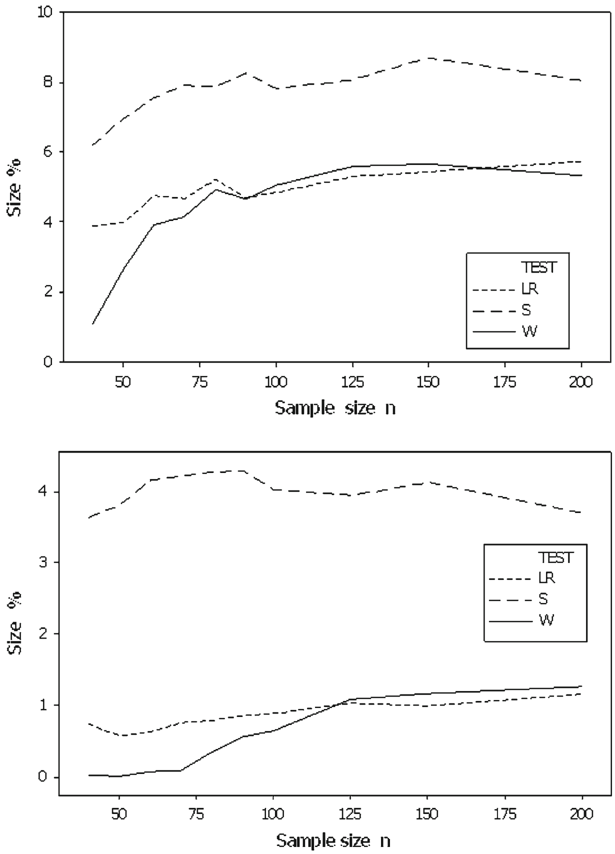

Fig. 5 Simulated sizes at nominal 5% (*upper plot*) and 1% (*lower plot*) levels of likelihood ratio (LR), Wald (W) and score (S) tests for the null hypothesis $\alpha = 1$ in the Marshall–Olkin extended Weibull distribution with uncensored data. Weibull shape parameter = 0.5

hazard functions that are entirely different from the Weibull hazard, as in the following example, and the test's power will be very high to detect such departures from the null.

6 Comments

The simulation results presented above show clearly that only the LR test is sufficiently accurate for use in testing for the extra parameter α in the extensions of the exponential and Weibull distributions and it is therefore recommended. Testing for these distributions is a matter of practical importance. Depending on the value of α, the extended distribution can have an interesting hazard function that is quite different from the hazard functions of the basic exponential and Weibull distributions and therefore can model real situations that the simple distributions cannot. For example, smaller values of α in Fig. 9 would lead to an increase in the height of the bump in the left-hand curve and eventually give an increasing–decreasing–increasing behaviour in the hazard function, which may be particularly useful in applications. It appears in

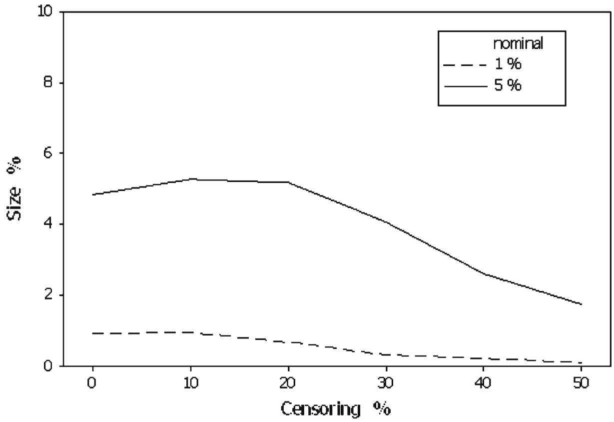

Fig. 6 Simulated size of likelihood ratio test for α in the Marshall–Olkin extended Weibull distribution, at the nominal 5 and 1% levels, as a function of average percentage of right censoring in samples of size 100. Weibull shape parameter = 0.5

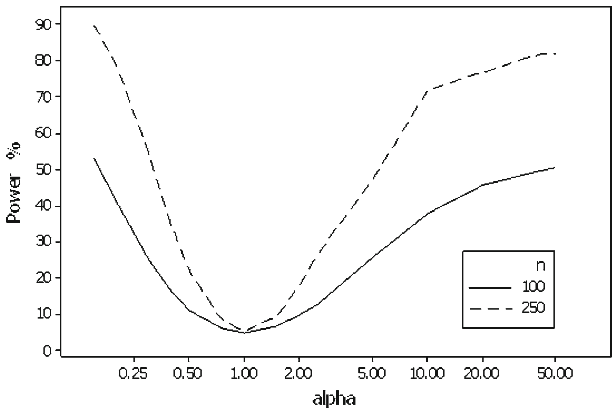

Fig. 7 Simulated power of likelihood ratio test at the nominal 5% level as a function of α in the Marshall–Olkin extended Weibull distribution with uncensored data, Weibull shape = 0.5

Ghitany et al. (2005), who found that an extended Weibull distribution of this form fitted very well to a set of data on bladder cancer ($n = 137$, Lee and Wang 2003). The fitted value of α was 0.074 (estimated SE 0.071). The value of the LR test statistic was 7.71 ($p = 0.006$), the Wald test gives 7.22 ($p = 0.007$) and the score test 3.29 ($p = 0.070$). Zhang and Xie (2007) also reported that an extended Weibull distribution improved the fit to a real data set compared to the unmodified Weibull, although our test statistics are not statistically significant in this case with only 30 observations (LR 0.46, $p = 0.50$; Wald 0.45, $p = 0.50$; score 0.62, $p = 0.43$). They went on to consider how to determine optimum burn-in times and replacement times when the hazard function takes this complex shape.

Besides the Marshall–Olkin method of extending the Weibull distribution, there are other ways of adding an extra parameter to the usual 2-parameter Weibull distribution in order to obtain extra flexibility. Several methods that have been proposed in

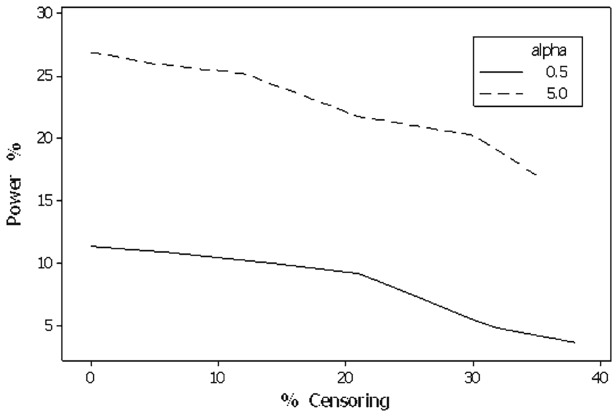

Fig. 8 Simulated power of likelihood ratio test at the nominal 5% level for selected values of alpha in the Marshall–Olkin extended Weibull distribution, as a function of average percentage of right censoring in samples of size 100. Weibull shape 0.5

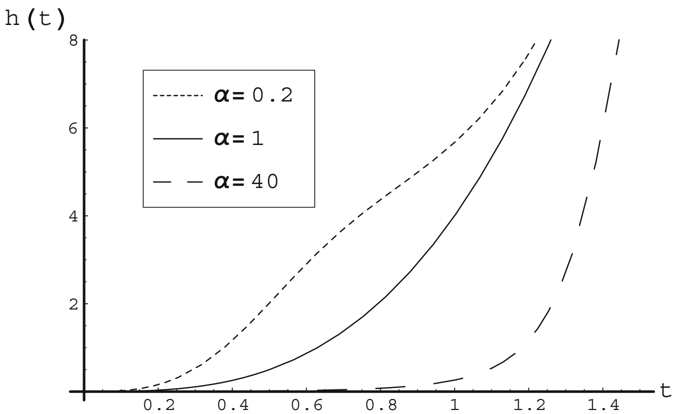

Fig. 9 Hazard functions of Marshall–Olkin extended Weibull distributions for α values for which the likelihood ratio test has 80% power in uncensored samples of size 250, Weibull shape = 4

the literature are discussed by Pham and Lai (2007). Since the standard 2-parameter distribution is nested inside the extended distribution, the likelihood-based tests can be applied to test whether the extended distribution is necessary. A study should be carried out in each case to establish which test is the most useful for that method of extension. Usually, this should be straightforward. Complications should arise only if the null value of the extra parameter lies on the boundary of the parameter space in which case the standard likelihood theory will not apply.

References

Bennett S (1983) Analysis of survival data by the proportional odds model. Statist Med 2:273–277

Buse A (1982) The likelihood ratio, Wald, and Lagrange multiplier tests: an expository note. Am Statist 36:153–157

Economou P, Caroni C (2007) Parametric proportional odds frailty models. Commun Statist Simul Comp 36:1295–1307

Ghitany ME (2005) Marshall–Olkin extended Pareto distribution and its application. Int J Appl Math 18: 17–32

Ghitany ME, Al-Awadhi FA, Alkhalfan LA (2007) Marshall–Olkin extended Lomax distribution and its application to censored data. Commun Statist Theory Meth 36:1855–1866

Ghitany ME, Al-Hussaini EK, Al-Jarallah RA (2005) Marshall–Olkin extended Weibull distribution and its application to censored data. J Appl Stat 32:1025–1034

Ghitany ME, Kotz S (2007) Reliability properties of extended linear failure-rate distributions. Prob Eng Inform Sci 21:441–450

Lee ET, Wang JW (2003) Statistical methods for survival data analysis, 3rd edn. Wiley, Hoboken, NJ

Lindsey JK (1996) Parametric statistical inference. Clarendon Press, Oxford

Marshall AW, Olkin I (1997) A new method of adding a parameter to a family of distributions with application to the exponential and Weibull families. Biometrika 84:641–652

Pham H, Lai C-D (2007) On recent generalizations of the Weibull distribution. IEEE Trans Reliab 56:454–458

Zhang T, Xie M (2007) Failure data analysis with extended Weibull distribution. Commun Statist Simul Comp 36:579–592