Pertanika J. Sci. & Technol. **16 (2)**: 83 - 95 (2008)
ISSN: 0128-7680

Parallel Weibull Regression Model

Jayanthi Arasan
Department of Mathematics, Faculty of Science,
Universiti Putra Malaysia, 43400 UPM,
Serdang, Selangor, Malaysia
E-mail: jayanthi@fsas.upm.edu.my

ABSTRACT

This paper focuses on the lifetime analysis of parallel systems consisting of Weibull components with independent failures and covariates. The performance of the parameter estimates of two and three-component parallel systems at different values of the shape parameter, σ, are compared and some confidence interval procedures are analyzed via a coverage probability study for m = 2, using simulated data.

Keywords: Wald, censoring, covariates

THE WEIBULL AND THE EXTREME VALUE DISTRIBUTION

There are numerous studies involving parallel systems in reliability. Many multivariate models have been developed, in particular, for the life testing of multi-component systems. Unfortunately, they may not be compatible with the lifetime analysis of medical data, which often involves other factors that affect survival times, more popularly known as covariates or concomitant variables. The estimation of the parameters of multivariate models is also usually very difficult and complicated, especially if it involves the estimation of a high number of parameters. More details on parallel systems can be found in Høyland and Rausand (1994), Kececioglu (1991) and in most books on reliability analysis.

Early work on parallel system models in the biomedical field was done by Gross *et al.* (1972), but this model does not include covariates and censored lifetimes. The same model was also discussed by Høyland and Rausand (1994) to study the reliability and mean time to failure of standby units in a parallel system. Elandt-Johnson and Johnson (1980) described another parallel system using the multi-hit models of Carcinogenesis via the Weibull distribution. Baklizi (1997) analyzed the likelihood inference in a parallel system regression model involving both censored and uncensored data. Arasan and Daud (2004) extended his work to analyze the efficiency of the parameter estimates of the same model with multiple covariates.

The Weibull distribution accommodates increasing, decreasing and also constant hazard rates as it reduces to the exponential as a special case. It is well known for modeling lifetimes and for its equivalence to extreme value distribution. The Weibull distribution can also be extended to include covariates by allowing its scale parameter, λ or shape parameter δ to depend on these variables, where $\lambda > 0$ and $\delta > 0$. The density and survivor functions of the Weibull are,

$$f(t) = \lambda\delta(\lambda t)^{\delta-1} exp\left[-\lambda t^{\delta}\right], t > 0 \quad (1)$$

Received : 8 January 2007
Accepted :14 April 2008

$$S(t) = exp[-(\lambda t)^{\delta}], \quad t > 0. \tag{2}$$

We know that if the lifetime, T has a Weibull distribution, then $Y = \log T$ has the extreme minimum value distribution, whose parameters could easily be modified to obtain the parameters of its Weibull counterpart. If Θ has extreme minimum value distribution, denoted as G(0,1), its density and survivor functions are,

$$f(\theta) = e^{(\theta - \exp(\theta))}, -\infty < \theta < \infty, \tag{3}$$

$$S(\theta) = e^{-\exp(\theta)}, -\infty < \theta < \infty \tag{4}$$

Suppose $x' = (x_0, x_1, ..., x_p)$ is the vector of covariate values, where $x_0 = 1$ and $\beta' = (\beta_0, \beta_1, ..., \beta_p$, are unknown parameters. If $\lambda = e^{-\beta' x}$, then the log lifetime, can be written as $Y = \beta' x + \sigma\Theta$, where $\sigma = \frac{1}{\delta}$. Since $\dfrac{Y - \beta' x}{\sigma}$ is equal to Θ, its distribution is extreme minimum value with the following distribution function,

$$F(y, \beta, x) = 1 - e^{-exp(\frac{y - \beta' x}{\sigma})}, -\infty < y < \infty \tag{5}$$

If the shape parameter of the Weibull distribution σ equals 1, then the Weibull is reduced to an exponential distribution. Having exponential lifetime means that the components of the parallel system would have constant failure rates and not strictly monotonic ones as in the Weibull case.

PARALLEL WEIBULL SYSTEM WITH COVARIATES

A parallel system can function as long as at least one of its components is still functioning. If the unit failures in a parallel system are assumed to be independent, then, this simply means that failure in one component will not affect the hazard rate of the remaining components. Although, this assumption may seem unrealistic, especially in the biomedical area, it can be very useful in making an initial interpretation of the data because the statistical analysis is much simpler and faster.

For a parallel system consisting of m identical and independent components, the probability of survival is equivalent to the probability of at least one component still operating. If t_k is the survival time of component k, where $k = 1, 2, ... m$, the time to failure of the system t is then,

$$t = \max\{t_1, t_2, ..., t_m\}.$$

It follows that the survival function of the entire system is,

$$S_T(t) = P(T > t) = 1 - \prod_{k=1}^{m} P(t_k \le t) = 1 - [F(t)]^m \tag{6}$$

For the i^{th} observation, if $y_i = \log t_i$ and $x'_i = (x_{i0}, x_{i1}, ..., x_{ip})$, the density and survival functions of the system are,

$$f(y_i) = \tfrac{m}{\sigma}\left\{e^{\frac{y_i - \beta' x_i}{\sigma} - \exp\left(\frac{y_i - \beta' x_i}{\sigma}\right)}\right\}\left\{1 - e^{-\exp\left(\frac{y_i - \beta' x_i}{\sigma}\right)}\right\}^{m-1}, \quad -\infty < y_i < \infty \tag{7}$$

$$S(y_i) = 1 - \left\{1 - e^{-exp\left(\frac{y_i - \beta' x_i}{\sigma}\right)}\right\}^m, \quad -\infty < y_i < \infty \tag{8}$$

LIKELIHOOD EQUATIONS AND ESTIMATION

Suppose we have both censored and uncensored lifetimes for $i + 1, 2, ..., n$ observations, accompanied by the information on p covariates. Let us say the following indicator variables were used to identify whether the data was censored or otherwise,

$$s_i = \begin{cases} 1 \ for\ t_i \text{ uncensored,} \\ 0 \ for\ t_i \text{ censored.} \end{cases} \tag{9}$$

The log-likelihood function of the full sample of a system consisting of m identical and independent Weibull components and p covariates is,

$$L(\beta,\sigma) = \sum_{i=1}^{n} \left\{ s_i \log f(y_i) + (1 - s_i) \log S(y_i) \right\}.$$

It follows that if $Z_i = \frac{y_i - \beta' x_i}{\sigma}$

$$L(\beta,\sigma) = \sum_{i=1}^{n} s_i \left\{ \log \tfrac{m}{\sigma} + z_i - e^{z_i} + (m-1)\log(1 - e^{-exp(z_i)}) \right\} +$$
$$(1 - s_i)\log\left\{1 - (1 - e^{-\exp(z_i)})^m\right\}. \tag{10}$$

The first and second derivatives of the log-likelihood function would be as follows,

$$\frac{\partial L(\beta,\sigma)}{\partial \beta_j} = \sum_{i=1}^{n} x_{ij}(-\tfrac{s_i}{\sigma} A_i + (1 - s_i)h_i),$$

$$A_i = 1 - e^{z_i} + \frac{(m-1)e^{z_i - \exp(z_i)}}{1 - e^{-\exp(z_i)}},$$

$$h_i = \frac{f(z_i)}{S(z_i)} = \frac{\frac{m}{\sigma}(e^{z_i - \exp(z_i)})\,(1 - e^{-\exp(z_i)})^{m-1}}{1 - (-1 - e^{-\exp(z_i)})^m},$$

$j = 0, 1 \, . \, . \, . \, , p.$

$$\frac{\partial L(\beta,\sigma)}{\partial \sigma} = \sum_{i=1}^{n} \tfrac{-s_i}{\sigma} B_i + (1 - s_i) z_i h_i,$$

$$B_i = 1 - z_i + z_i e^{z_i} + \frac{(m-1)z_i e^{-\exp(z_i)}}{1 - e^{-\exp(z_i)}}.$$

$$\frac{\partial^2 L(\beta,\sigma)}{\partial \beta_j \partial \beta_k} = \sum_{i=1}^{n} x_{ij} x_{ik} \left\{ \tfrac{-s_i}{\sigma^2} D_i - (1 - s_i)\,(\tfrac{A_i h_i}{\sigma} + h_i^2) \right\},$$

$$D_i = e^{z_i} + \frac{(m-1)(e^{z_i - \exp(z_i)})\,(-1 + e^{z_i} + e^{-\exp(z_i)})}{(1 - e^{-\exp(z_i)})^2},$$

$j, \ k = 0, \ 1 \ ...,p.$

$$\frac{\partial^2 L(\beta,\sigma)}{\partial \beta_j \partial \sigma} = \sum_{i=1}^{n} x_{ij} \left\{ -\tfrac{s_i}{\sigma^2}(-A_i + z_i D_i) - (1 - s_i)\,(\tfrac{1}{\sigma} B_i h_i + z_i h_i^2) \right\}$$

$j = 0, \ 1 \ ...,p.$

$$\frac{\partial^2 L(\beta,\sigma)}{\partial\sigma^2} = \sum_{i=1}^{n} - \frac{-s_i}{\sigma^2}\left\{-B_i + z_i\left((-1 + e^{zi} + z_i e^{zi}) + (m-1)E_i\right)\right\}$$
$$-(1-s_i)z_i\left(\frac{-h_i}{\sigma} - \frac{B_i h_i}{\sigma} - z_i h_i^2\right),$$
$$E_i = \frac{e^{z_i - \exp(z_i)}(-1 + e^{-\exp(z_i)} - z_i + z_i e^{-\exp(z_i)} + z_i e^{z_i})}{(1 - e^{-\exp(z_i)})^2}.$$

The inverse of the observed information matrix, which can be obtained from the second partial derivatives of the log-likelihood function evaluated at $\hat{\beta}$ and $\hat{\sigma}$ provides us with the estimators for the variance and covariance,

$$\widehat{var}(\widehat{\beta}, \widehat{\sigma}) = [i(\widehat{\beta}, \widehat{\sigma})]^{-1} \quad (11)$$

Simulation Study

Study Design

A simulation study was conducted to see how well the estimation procedure works with the parallel Weibull regression model with two components and two covariates at different values of sigma. In addition, we were also interested in investigating how the proportion of censoring in the data affected the parameter estimates. The study was conducted by using 1000 simulations, each with sample size of 100. The survival times were obtained by drawing 100 random numbers from the uniform distribution between 0 and 1, $U(0,1)$. These numbers were later used to produce the log survival time, y_i. For the i^{th} observation,

$$y_i = \beta' x_i + (\log(-\log(1 - u^{\frac{1}{m}})))\sigma. \quad (12)$$

The value of 1 was used as the parameters of β_0, β_1 and β_2. As the parameter of σ, three different values of 0.5, 0.8 and 1.15 were used to enable a comparative study. We know that δ is the shape parameter of the Weibull distribution and its value determines whether the distribution has an increasing, constant or decreasing failure rate. Since, $\sigma = \frac{1}{\delta}$ a value of σ less than 1 would indicate an increasing hazard rate while value of more than 1 would indicate otherwise.

In addition, a simulation study for a three-component system with two covariates was carried out for σ = 0.5 to see how the bias, standard error and root mean square error (rmse) changes with increase in the number of components. The two covariates used in the model were simulated from the Bernoulli and standard normal distribution. Six different values of approximate censoring proportions (cp), between 0 and 0.5 were used to investigate the relationship between censoring proportion and efficiency of the parameter estimates. A value of cp=0.3 means that approximately 30% of the survival times were censored before the actual failure times.

The censoring time for the i^{th} observation, $c_i \sim \exp(\mu)$, is where the value of μ would be adjusted to obtain the desired approximate censoring proportion in our data. If $t_i = \exp(y_i)$ then t_i will be censored at c_i according to the following,

$$s_i = \begin{cases} 1 & if\ t_i \le c_i, \\ 0 & if\ t_i \le c_i. \end{cases} \quad (13)$$

Parameter Estimation and Calculations

The estimates of β and σ can be obtained by solving the likelihood equations using any iterative procedure for solving non linear equations. In this research, the maximum likelihood estimators of all the parameters were computed using the Newton Raphson iterative method, which was implemented using the FORTRAN programming language. In this section the true parameter values were used as the starting values for all the estimates.

Simulation Results and Conclusion

The results of the simulation study are given in the following section. Both bias and standard error contribute to the average error size of an estimator, thus the rmse, $\sqrt{s.e^2 + bias^2}$ is used to measure the average overall error of the parameter estimates. Tables 1-3 display the bias and standard error of the estimates at different values of σ.

The values of the rmse are illustrated in *Figs. 1-2*. It is clear that both standard error and rmse of all parameter estimates increase with the increase in censoring proportion. This was expected because increase in censoring proportion means less data with complete failure times and thus the likelihood contribution would depend more on the survival function and censored times instead of the density function and the exact failure times.

As for the bias, although it seems to increase with the increase in censoring proportion, the trend is not very clear, probably because of the increasing standard error values. It should be borne in mind that a low value of bias at higher levels of censoring proportion does not imply that the resulting estimates are better than the ones at lower censoring proportion. This is because the increasing standard error values at higher censoring proportion suggest that the estimates are still typically far from the real value even though the average is close to the parameter value.

However, none of the bias values were significant at the 5% level. In addition, the rmse and standard error also appear to be higher at higher values of σ. The reason for this can be explained as follows. In the original Weibull distribution the events will be sparse for smaller δ so the parameters will be estimated less efficiently when δ is smaller. Similarly, with the parallel Weibull model the parameters will be estimated less efficiently when σ is larger, because $\sigma = \frac{1}{\delta}$.

TABLE 1
Bias and standard error when σ = 0.5

	Bias				Std. Error			
cp	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\sigma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\sigma}$
0.0	0.000581	0.000075	0.001155	-0.007104	0.057150	0.074871	0.038976	0.038511
0.1	0.001473	-0.002033	0.000003	-0.007816	0.058883	0.079985	0.041491	0.039747
0.2	0.000442	-0.000598	0.001274	-0.007355	0.060513	0.083162	0.043113	0.040508
0.3	-0.000347	-0.001084	0.000610	-0.008624	0.062857	0.087451	0.047459	0.041316
04	0.003094	-0.001318	0.001884	-0.009578	0.065999	0.100208	0.054448	0.042908
0.5	0.002043	-0.002524	0.002759	-0.011823	0.070825	0.107113	0.063674	0.044408

TABLE 2
Bias and standard error when σ = 0.80

	Bias				Std. Error			
cp	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\sigma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\sigma}$
0.0	0.000772	0.001110	0.001402	-0.011430	0.091543	0.119754	0.062798	0.061781
0.1	0.001261	-0.003418	0.002256	-0.012283	0.093123	0.126553	0.065503	0.063672
0.2	-0.000451	-0.003492	0.000325	-0.015526	0.098181	0.130567	0.070673	0.063754
0.3	0.000202	-0.000566	0.002043	-0.002524	0.100325	0.137630	0.077110	0.065856
04	0.005657	-0.002862	0.003901	-0.028687	0.106841	0.149999	0.084543	0.068156
0.5	0.002043	-0.005297	0.007270	-0.030128	0.108170	0.169781	0.091523	0.071096

TABLE 3
Bias and standard error when σ = 1.15

	Bias				Std. Error			
cp	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\sigma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\sigma}$
0.0	0.000662	0.000578	0.003038	-0.015808	0.131393	0.173225	0.089642	0.087425
10.0	0.001260	-0.002929	0.001837	-0.019161	0.134265	0.185366	0.096405	0.087970
20.0	0.008061	-0.005796	-0.001241	-0.023398	0.139283	0.196294	0.098891	0.0888667
30.0	0.003024	0.002277	0.013622	-0.022709	0.139601	0.206300	0.110511	0.090894
40.0	0.001900	0.008285	0.016151	-0.024623	0.151129	0.229006	0.126382	0.099512
50.0	0.000750	0.017021	0.017841	-0.021945	0.156191	0.245132	0.131873	0.132947

There were no serious problems encountered in the estimation process although there were a few samples with convergence problems when the censoring proportion in the data was high, usually when cp=0.5. Since efficiency of an estimator depends on both precision and accuracy, based on the rmse measures obtained, we conclude that the estimation procedure is most efficient when the censoring proportion in the data and the value of σ are both low.

Table 4 compares the bias, standard error and rmse rmse of the parameter estimates when m = 3, where m is the number of components in the system. It appears that the bias values increase whereas the standard error values decrease with an increase in the number of components in the system. The rmse, which is the total error, seems to decrease with the increase in number of components in the system. However, at high censoring proportions some of the rmse values for m = 3 are higher than m = 2.

CONFIDENCE INTERVAL ESTIMATES

It is rather common to resort to confidence interval estimates based on the asymptotic normality of maximum likelihood estimates when it is impossible to compute the exact confidence intervals of the parameters of a model. This is also known as basing the intervals on the Wald statistic or simply Wald intervals. Other popular interval estimates are those based on the likelihood ratio test and the score test. In this section, we will be focusing on intervals based on the Wald statistics and recommend suggestions on how to improve these estimates using parameterizations. The method will be assessed by conducting a coverage probability study.

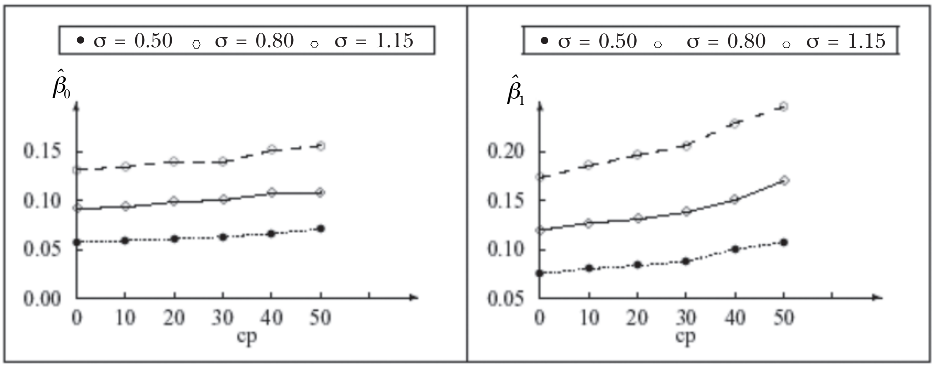

Fig. 1: Values of rmse for $\hat{\beta}_0$ and $\hat{\beta}_1$ vs. cp

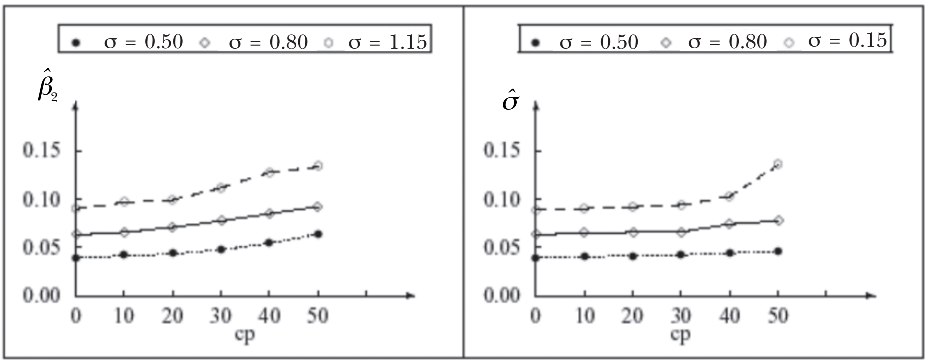

Fig. 2: Values of rmse for $\hat{\beta}_2$ and $\hat{\sigma}$ vs. cp

TABLE 4
Comparison between different number of component models

		m =2			*m* = 3		
Est.	cp	Bias	Std.err	rmse	Bias	Std.err	rmse
$\hat{\beta}_0$	0.0	0.000581	0.057150	-0.057153	0.002730	0.050849	0.050922
	0.2	-0.000442	0.060513	-0.0060515	0.016328	0.055214	0.057578
	0.4	-0.003094	-0.065999	-0.066071	0.033482	0.057741	0.066746
$\hat{\beta}_1$	0.0	0.000075	0.074871	0.074871	0.000326	0.063108	0.063109
	0.2	-0.000598	0.083162	0.083164	0.010462	0.071568	0.072329
	0.4	-0.001318	0.100208	0.100217	0.019873	0.085562	0.087839
$\hat{\beta}_2$	0.0	0.001155	0.038976	0.038993	0.001009	0.032872	0.032887
	0.2	0.001274	0.043113	0.043132	0.009371	0.038748	0.039865
	0.4	0.001884	0.054448	0.054480	0.018654	0.049361	0.052768
$\hat{\sigma}$	0.0	-0.007104	0.038511	0.039161	-0.007110	0.037757	0.038421
	0.2	-0.007355	0.040508	0.041170	-0.019685	0.038561	0.043295
	0.4	-0.009578	0.042908	0.043964	-0.027013	0.039846	0.048139

Wald Confidence Intervals

Let $\hat{\theta}$ be the maximum likelihood estimator for parameter θ and $l(\theta)$ the log-likelihood function of θ. Under mild regularity conditions, $\hat{\theta}$ is asymptotically normally distributed with mean θ and covariance matrix, $I^{-1}(\theta)$ where $I(\theta)$ is the Fisher information matrix evaluated at the true value of the parameter, θ (Cox and Hinkley, 1979). The matrix, $I(\theta)$ which is not available can be replaced by the observed information matrix, $i(\hat{\theta})$ whose $(j, k)^{th}$ element can be obtained from the second partial derivatives of the log-likelihood function evaluated at $\hat{\theta}$ as given below,

$$i(\hat{\theta})_{jk} = \left(-\frac{\partial^2 l(\hat{\theta})}{\partial\hat{\theta}_j \partial\hat{\theta}_k}\right), \quad j,\ k = 0,\ 1\ \ldots,\ p. \tag{14}$$

The estimate of var$(\hat{\theta}_j)$ is then given by the $(j, j)^{th}$ element of $i^{-1}(\hat{\theta})$ If $z_{1-\frac{\alpha}{2}}$ is the $(1-\frac{\alpha}{2})$ quartile of the standard normal distribution, then the $100(1-\alpha)\%$ confidence interval for θ_j is given by the following,

$$\hat{\theta}_j - z_{1-\frac{\alpha}{2}}\sqrt{i^{-1}(\hat{\theta})_{jj}} < \theta_j < \hat{\theta}_j + z_{1-\frac{\alpha}{2}}\sqrt{i^{-1}(\hat{\theta})_{jj}}\,. \tag{15}$$

Application of the Wald Intervals

The parallel system model with two covariates discussed in the previous section has a total of 4 parameters to be estimated, β_0, β_1, β_2 and σ. Estimates of var$(\hat{\beta}_j)$ and var$(\hat{\sigma})$ can be obtained from $[i(\hat{\beta},\ \hat{\sigma})]^{-1}$. The $100(1-\alpha)$ % confidence interval for β_j and σ where $j = 0,\ 1,\ 2$ are,

$$\widehat{\beta}_j - z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}(\widehat{\beta}_j)} < \beta_j < \widehat{\beta}_j + z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}(\widehat{\beta}_j)}, \tag{16}$$

$$\widehat{\sigma} - z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}(\widehat{\sigma})} < \sigma < \widehat{\sigma} + z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}(\widehat{\sigma})}. \tag{17}$$

It is expected that Wald intervals work rather well for the covariate parameters but not so well for σ, the shape parameter. The Wald interval for σ will most probably be highly asymmetrical and will not have desirable statistical properties due to a sharp boundary in the parameter space. This is usually the case with other similarly bounded parameters, such as the odds ratio, as reported by Hosmer and Lemeshow (1999).

By applying a suitable parameterization such as $1/\sigma$ or log σ, it is expected that the confidence interval estimates will have better symmetry in its left and right estimated error probabilities because the parameterization will help in achieving more symmetry in the log-likelihood function. The following part deals with a suitable parameterization for σ and intervals based on these transformations.

Parameterization for σ

The usual procedure for applying any parameterization is rather straightforward. If a transformation $\varsigma = \varsigma(\sigma)$ is chosen, then $\hat{\sigma}$, the maximum likelihood estimator for σ must be obtained and transformed to the ς scale, $\hat{\varsigma} = \varsigma(\hat{\sigma})$. Then calculate $i(\hat{\varsigma})$ which is the observed information matrix evaluated at $\hat{\varsigma}$,

$$i(\hat{\varsigma}) = \frac{i(\hat{\sigma})}{[\varsigma'(\hat{\sigma})]^2}. \tag{18}$$

where $\varsigma'(\hat{\sigma})$ is the first derivative of ς with respect to σ. The $100(1 - \alpha)\%$ confidence interval for ς is,

$$\hat{\varsigma} - z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}(\hat{\varsigma})} < \varsigma < \hat{\varsigma} + z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}(\hat{\varsigma})}. \tag{19}$$

where $\widehat{var}(\hat{\varsigma})$ can be obtained from the inverse of the observed information matrix, $[i(\hat{\varsigma})]^{-1}$ The equivalent confidence interval for σ can be obtained by transforming the confidence interval for ς back to the original scale. For example, if $\varsigma = \log\sigma$ then the transformation back is $\sigma = \exp(\varsigma)$.

Simulation Study

Study Design

A simulation study was conducted using 2000 samples of size n=100, 150, 200 and 250 to evaluate the performance of the confidence intervals based on the Wald statistics. Since there are 4 parameters to be estimated in this model and data is censored, a larger sample size is required in order to have any meaningful results. Samples were generated from the parallel Weibull regression model with two components and two covariates when $\sigma = 0.5$. The samples were produced and censored using methods similar to the ones described in section (2.2.2). Two levels of approximate censoring proportions, cp=0.10 and cp=0.30 were used to see how the level of censoring affected the interval estimates. The values of cp=0.10 and cp=0.30 were chosen to represent both low and high levels of censoring proportions respectively.

These samples were then used to obtain the maximum likelihood estimators of the parameter and the estimators of their variances to carry out the coverage probability study. The coverage probability is the probability that an interval contains the true parameter value. For parameters, β_0, β_1 and β_2 only the coverage probability was analyzed using the untransformed Wald interval estimates. For the parameter σ the coverage probability study was conducted using both the untransformed and two transformed Wald interval estimates. The first parameterization is using $\varsigma = \log\sigma$ and the other is using $\varsigma = \frac{1}{\sigma}$.

Parameter Estimation and Calculations

The parameter estimation process is similar to that described in section (2.2.2). The coverage probability study was conducted by calculating the left and right estimated error probabilities for each of the parameter estimates. The estimated left(right) error probability is calculated by adding the number of times the left(right) endpoint was more(less) than the true parameter value divided by the total number of samples, *N*.

For the covariate parameters β_j, where j = 0, 1, 2, the left and right error probabilities are,

$$\text{Left} = \frac{\#\left\{\hat{\beta}_j - z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}(\hat{\beta}_j)} > \beta_j\right\}}{2000},$$

$$\text{Right} = \frac{\#\left\{\hat{\beta}_j + z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}(\hat{\beta}_j)} < \beta_j\right\}}{2000}.$$

For the untransformed σ

$$\text{Left} = \frac{\#\left\{\hat{\sigma} - z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}(\hat{\sigma})} > \sigma\right\}}{2000},$$

$$\text{Right} = \frac{\#\left\{\hat{\sigma} + z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}(\hat{\sigma})} < \sigma\right\}}{2000}.$$

For the transformed $\varsigma = \varsigma(\sigma)$,

$$\text{Left} = \frac{\#\left\{\hat{\varsigma} - z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}(\hat{\varsigma})} > \varsigma\right\}}{2000},$$

$$\text{Right} = \frac{\#\left\{\hat{\varsigma} + z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}(\hat{\varsigma})} < \varsigma\right\}}{2000}.$$

When the nominal error probability α is 0.05(0.1), the ideal left and right error probabilities should be equal to $\frac{\alpha}{2} = 0.025(0.05)$. Similarly, the ideal total error probabilities should be equal to $\alpha = 0.05(0.1)$. Following Doganaksoy and Schmee (1993), if α was the nominal error probability, then the standard error of the estimated error probability $\hat{\alpha}$ assuming that the observed and nominal error are close, is approximately, $\sqrt{\frac{\alpha(1-\alpha)}{N}}$.

Using a normal approximation, a 99.75% confidence interval for α would be $\hat{\alpha} \pm 2.58\ s.e(\hat{\alpha})$. So, if this interval contains α, then the total error probability, $\hat{\alpha}$ is considered to have actually converged to the nominal error probability, $\overline{\alpha}$. If the total error probability is greater than $\alpha + 2.58\ s.e(\hat{\alpha})$ then the method is termed anticonservative and if it is lower than $\alpha - 2.58\ s.e(\hat{\alpha})$, the method is termed conservative. The estimated error probabilities are called symmetric when the larger error probability is less than 1.5 times the smaller one.

The overall performance of the different methods will be judged based on their total number of anticonservative, conservative and asymmetrical intervals. When an interval is conservative (anticonservative), it means that it generates coverage probability that is greater (smaller) than $(1 - \alpha)$. So, a conservative (anticonservative) confidence interval procedure leads to confidence intervals, which are generally wider (shorter) than they need to be. Thus, a conservative interval is still considered valid and a higher penalty is attached to an anticonservative interval. In addition, methods that are robust in handling censored data and behave well at different nominal levels are also of interest.

Simulation Results and Conclusion

Tables 5-8 give the results of the simulation study at different levels of censoring proportion and two levels of α when $\sigma = 0.5$. The Wald procedure seems to work rather well for parameters β_0, β_1 and β_2. When α is 0.05, all total error probabilities of the parameters seem to be close to the nominal level except in a few cases when the sample size is low, n = 100, where there were two anticonservative intervals even at low censoring proportion, as shown in Table 5.

Unexpectedly, there are fewer anticonservative and asymmetrical intervals when the censoring proportion is high and n is low compared to when censoring proportion is low. This is probably caused by the combined effect of both censoring and small sample size on the estimated parameters and its standard errors. The small sample size and high censoring proportion cause the intervals to be much wider (due to larger standard errors of the parameter estimates).

In addition, the high level of censoring in the data generated estimates that were very biased resulting in intervals that sometimes accidentally included the true parameter value. The coverage probabilities were also better at higher nominal levels, α. When censoring proportion and were high, there were no anticonservative, conservative or asymmetrical intervals.

For the parameter, σ , as expected the untransformed Wald procedure generated intervals that were highly asymmetrical. This method also generated more anticonservative intervals compared to the transformed methods. From the two different parameterizations of σ the transformation $\varsigma = \frac{1}{\sigma}$ produced more symmetrical intervals than the transformation $\varsigma = \log\ \sigma$.

The transformation $\varsigma = \frac{1}{\sigma}$ worked very well especially when α = 0.05 because in addition to being more symmetrical, it also did not generate any conservative or anticonservative intervals. However, it generated some conservative and asymmetrical intervals when both α and censoring proportion were high as shown in Table 8. The untransformed Wald method generated many anticonservative intervals whereas the transformation $\varsigma = \log\ \sigma$ produced many conservative intervals. Both the untransformed Wald and the transformation method $\varsigma = \log\ \sigma$ failed to generate a single symmetrical interval.

TABLE 5
Estimated error probabilities of β at α = 0.05, A = "Anticonservative"

θ	n	cp = 0.10			cp = 0.03		
		Left Error	Right Error	Total Error	Left Error	Right Error	Total Error
β_0	100	0.0330	0.0240	0.0570	0.0285	0.0300	0.0585
	150	0.0265	0.0225	0.0490	0.0265	0.0195	0.0460
	200	0.0285	0.0265	0.0550	0.0225	0.0260	0.0485
	250	0.0275	0.0280	0.0555	0.0245	0.0240	0.0485
β_1	100	0.0380	0.0285	0.0665[A]	0.0325	0.0330	0.0655[A]
	150	0.0325	0.0280	0.0605	0.0295	0.0295	0.0590
	200	0.0345	0.0250	0.0595	0.0250	0.0255	0.0505
	250	0.0315	0.0250	0.0565	0.0275	0.0235	0.0510
β_2	100	0.0300	0.0330	0.0630[A]	0.0250	0.0305	0.0555
	150	0.0290	0.0315	0.0605	0.0255	0.0290	0.0545
	200	0.0220	0.0355	0.0575	0.0235	0.0350	0.0585
	250	0.0255	0.0320	0.0575	0.0275	0.0290	0.0565

TABLE 6
Estimated error probabilities of β at α = 0.10, A = "Anticonservative"

θ	n	cp = 0.10			cp = 0.03		
		Left Error	Right Error	Total Error	Left Error	Right Error	Total Error
β_0	100	0.0620	0.0435	0.1055	0.0555	0.0505	0.1060
	150	0.0570	0.0420	0.0990	0.0520	0.01445	0.0965
	200	0.0570	0.0475	0.1045	0.0455	0.0500	0.0955
	250	0.0515	0.0515	0.1030	0.0465	0.0480	0.0945
β_1	100	0.0635	0.0580	0.0665^A	0.0575	0.0595	0.1170
	150	0.0555	0.0515	0.1070	0.0510	0.0505	0.1015
	200	0.0625	0.0495	0.1120	0.0555	0.0465	0.1020
	250	0.0530	0.0420	0.0950	0.0570	0.0425	0.0995
β_2	100	0.0555	0.0605	0.1160	0.0480	0.0570	0.1050
	150	0.0560	0.0620	0.1180^A	0.0515	0.0585	0.1100
	200	0.0575	0.0585	0.1160	0.0455	0.0620	0.1075
	250	0.0485	0.0610	0.1095	0.0485	0.0605	0.1090

TABLE 7
Estimated error probabilities of σ at α = 0.05, A = "Anticonservative" C="Conservative"

cp	n	$\varsigma = \sigma$			$\varsigma = \log \sigma$			$\varsigma = \frac{1}{\sigma}$		
		Left Error	Right Error	Total Error	Left Error	Right Error	Total Error	Left Error	Right Error	Total Error
	100	0.0065	0.0570	0.0635^A	0.0120	0.0420	0.0540	0.0190	0.0265	0.0455
	150	0.0080	0.0430	0.0510	0.0105	0.0355	0.0460	0.0165	0.0290	0.0455
0.1	200	0.0110	0.0535	0.0645^A	0.0155	0.0430	0.0585	0.0220	0.0305	0.0525
	250	0.0120	0.0490	0.0490	0.0610	0.0170	0.0400	0.0570	0.0210	0.0555
	100	0.0025	0.0390	0.0415	0.0085	0.0280	0.0365^C	0.0190	0.0265	0.0455
	150	0.0035	0.0410	0.0445	0.0070	0.0290	0.0360^C	0.0165	0.0290	0.0455
0.3	200	0.0050	0.0380	0.0430	0.0105	0.0280	0.0385	0.0220	0.0305	0.0525
	250	0.0060	0.0415	0.0475	0.0070	0.0305	0.0375	0.0210	0.0345	0.0555

In fact, on average, the untransformed Wald method was almost five times more asymmetrical than the same method using $\varsigma = \frac{1}{\sigma}$, when σ was low. The confidence intervals for the $\varsigma = \log \sigma$ transformation were almost twice as asymmetrical as those for $\varsigma = \frac{1}{\sigma}$. The transformation, $\varsigma = \frac{1}{\sigma}$, should be preferred because the other two transformations gave very asymmetrical intervals. Thus, $\varsigma = \frac{1}{\sigma}$, in the original Weibull distribution works best.

The transformation, $\varsigma = \frac{1}{\sigma}$, worked very well especially when α =0.05 because in addition to being more symmetrical, it also did not generate any conservative or anticonservative intervals. However, it generated some conservative and asymmetrical intervals when both α and censoring proportion were high as shown in Table 8. The

TABLE 8
Estimated error probabilities of σ at α = 0.10, A = "Anticonservative" C="Conservative"

cp	n	$\varsigma = \sigma$ Left Error	Right Error	Total Error	$\varsigma = \log \sigma$ Left Error	Right Error	Total Error	$\varsigma = \frac{1}{\sigma}$ Left Error	Right Error	Total Error
0.1	100	0.0215	0.0980	0.1195[A]	0.0245	0.0820	0.1065	0.0305	0.0660	0.0965
	150	0.0215	0.0815	0.1030	0.0260	0.0665	0.0925	0.0335	0.0530	0.0865
	200	0.0295	0.0815	0.1110	0.0340	0.0735	0.1075	0.0380	0.0650	0.1030
	250	0.0310	0.0820	0.1130	0.0345	0.0745	0.1090	0.0400	0.0650	0.1050
0.3	100	0.0150	0.0695	0.0845	0.0220	0.0550	0.0770[C]	0.0270	0.0400	0.0670[C]
	150	0.0130	0.0715	0.0845	0.0190	0.0610	0.0800[C]	0.0255	0.0485	0.0740[C]
	200	0.0185	0.0670	0.0855	0.0230	0.0575	0.0805[C]	0.0290	0.0450	0.0740[C]
	250	0.0210	0.0695	0.0905	0.0240	0.0615	0.0855	0.0315	0.0535	0.0850

untransformed Wald method generated many anticonservative intervals whereas the transformation $\varsigma = \log \sigma$ produced many conservative intervals. Both the untransformed Wald and the transformation method $\varsigma = \log \sigma$ failed to generate a single symmetrical interval.

In fact, on average, the untransformed Wald method was almost five times more asymmetrical than the same method using $\varsigma = \frac{1}{\sigma}$ when α was low. The confidence intervals for the $\varsigma = \log \sigma$ transformation were almost twice as asymmetrical as those for $\varsigma = \frac{1}{\sigma}$. Thus, the transformation, $\varsigma = \frac{1}{\sigma}$, should be preferred because the other two transformations gave very asymmetrical intervals. So, basically $\varsigma = \frac{1}{\sigma}$ in the original Weibull distribution works best.

REFERENCES

Arasan, J. and Daud, I. (2004). Parallel system using the log-linear exponential model with censored data and two covariates. *Pertanika J. of Sci. and Technol., 12*(2), 265-275.

Baklizi, A.S.M. (1997). Likelihood inference in parallel systems regression models with censored data (PhD thesis, University Putra Malaysia, Department of Statistics, 1997).

Cox, D.R. and Hinkley, D.V. (1979). *Theoretical Statistics.* New York: Chapman and Hall.

Doganaksoy, N. and Schmee, J. (1993). Comparison of approximate confidence intervals for distributions used in life-data analysis. *Technometrics, 35*(2), 175-184.

Elandt-Johnson, R.C. and Johnson, N.L. (1980). *Survival Models and Data Analysis.* New

York: Wiley.

Høyland, A. and Rausand, M. (1994). *System Reliability Theory.* New York: Wiley.

Hosmer, D.W. and Lemeshow, S. (1999). *Applied Survival Analysis.* New York: Wiley.

Kececioglu, D. (1991). *Reliability Engineering Handbook* (Volume 1). New York. Prentice Hall.