An ALT Proportional Hazard-Proportional Odds Model

T. Huang[1], E. A. Elsayed[2] and T. Jiang[1]

[1]*Department of System Engineering of Engineering Technology, Beihang University, Beijing 100083, China.*
E-mail: huangtt@dse.buaa.edu.cn
E-mail: jtm@buaa.edu.cn
[2]*Department of Industrial and Systems Engineering, Rutgers University, Piscataway, NJ, 08854-0818, USA.*
E-mail: elsayed@rci.rutgers.edu

Abstract

Accelerated life testing (ALT) is used to obtain failure data of products in a short time period in order to extrapolate the reliability and lifetime of the products under normal conditions. Estimation of the reliability and lifetime is based on models, which can be classified as parametric and non-parametric models. Non-parametric models are commonly used because of the distribution-free property. Proportional hazard model (PHM) and proportional odds model (POM) are two widely used non-parametric models for reliability prediction based on ALT data. These two models perform well if the underlying distributions are Weibull and log-logistic respectively. However, in some situations the failure time data may be obtained from a mixed population and using PHM or POM will result in inaccurate estimates of the reliability at normal operating conditions. In this paper, we develop a proportional hazard-proportional odds (PH-PO) model in order to obtain more accurate estimation for both mixed and non-mixed populations when the failure time distributions follow Weibull and log-logistic or a mix of the two. This PH-PO model is based on a parameter family which makes proportional hazard model and proportional odds model special cases of the proposed model. Baseline function and stress effect function of the PH-PO model are determined and maximum likelihood function is defined for parameter estimation.

The performance of the PH-PO model is verified numerically using simulated data. The results show that in general the PH-PO model gives more accurate estimations of the reliability.

Keywords: *accelerated life testing, proportional hazard model, proportional odds model, parameter family, transformation parameter*

Nomenclature

ALT: accelerated life testing

MLE: maximum likelihood estimation

PHM: proportional hazard model

POM: proportional odds model

PH-POM: proportional hazard-proportional odds model

$\lambda(t;z)$: hazard rate function under stress z

$\lambda_0(t)$: baseline hazard rate function

$\theta(t;z)$: odds function under stress z

$\theta_0(t)$: baseline odds function

$g(t;z)$: function of the PH-PO model under stress z

$g_0(t)$: baseline function of the PH-PO model

z: stress vector

c: transformation parameter

β: unknown regression model parameter vector

γ_1, γ_2: unknown parameters of baseline function of the PH-PO model

w: weight parameter

$e^{\alpha z}$: acceleration factor

1 Introduction

The purpose of accelerated life testing is to induce failures in a short time period under higher stresses and use the obtained failure time data to estimate the reliability and lifetime under normal operating conditions. Models of accelerated life testing are used for such extrapolations. Two statistics-based models are widely used (Elsayed, 1996): parametric models which require a priori specified lifetime distribution; and non-parametric models that relax the assumption of the lifetime distribution.

One of the widely used non-parametric models is the proportional hazard model presented by Cox (1972). It assumes constant ratio of the hazard rates under different stresses. This model performs well especially if the underlying lifetime distribution of the test units is described by a Weibull distribution. However, the assumption of proportional hazard model does not always hold. Brass (1971) shows that the death rates (or hazard rates) of two groups of patients under different stresses (e.g. smoke or non-smoke) converge nearly to unity with time in some conditions (such as one group of young people and one group of old people). Hence, he proposes the odds function model. Zhang and Elsayed (2007) present the proportional odds model in accelerated life testing based on the odds function and show that it gives a more accurate reliability estimates than proportional hazard model when the underlying lifetime distribution is log-logistic. Other models that generalize the PH model are developed such as the generalized additive models (Hastie and Tibshirani, 1990), and the extended linear hazard regression model (Wang and Elsayed, 2001).

As stated earlier, the failure time data at accelerated conditions might be the result of mixtures of different distributions and it becomes difficult to use standard models for reliability prediction. Under this condition, a generalized model is likely to provide more accurate estimates if properly developed. Aranda-Ordaz (1981) proposes a parameter family which makes the proportional hazard model and the proportional odds model special cases when some conditions are met. In medical field, regression models based on this parameter family are presented to analyze the lifetime of patients who suffer from different kinds of diseases. Examples are the link-based models presented by Younes and Lachin (1997) and spline-based parametric survival models proposed by Royston and Parmar (2002). However, the generalized models based on parameter family have not been developed for ALT area.

In this paper we develop and present a PH-PO model for accelerated life testing, which generalizes both the proportional hazard model and the proportional odds model. We then validate the model using a simulation study and assess the reliability estimation accuracy at normal operating conditions.

2 PH-PO model

2.1 Formulation of the PH-PO model

To specify a model for accelerated life testing, we need to identify its three functions: reliability, stress effect and baseline functions. Table 1 shows these functions for both the proportional hazards model and proportional odds model.

Table 1 Model functions for PHM and POM

Function	PHM	POM
Definition in reliability	$\lambda(t;z)=\dfrac{f(t;z)}{R(t;z)}$	$\theta(t;z)=\dfrac{F(t;z)}{R(t;z)}$
Stress effect function	$\lambda(t;z)=\lambda_0(t)e^{\beta z}$	$\theta(t;z)=\theta_0(t)e^{\beta z}$
Baseline function with one degree of freedom (1 d.f.)	$\lambda_0(t)=\dfrac{m}{\eta}\left(\dfrac{t}{\eta}\right)^{m-1}$ (Weibull distribution)	$\theta_0(t)=(\lambda t)^p$ (Log-logistic distribution)
Logarithm of baseline function	$\ln\lambda_0(t)=(m-1)\ln t+\ln(\dfrac{m}{\eta^m})$	$\ln\theta_0(t)=p\ln t+p\ln\lambda$

* m and η are the shape and scale parameters of Weibull distribution respectively. $\lambda=e^{-\mu}$ and $p=\dfrac{1}{\sigma}$, while μ and σ are the scale and shape parameters of log-logistic distribution respectively.

We develop a PH-PO model based on a parametric family defined by Aranda-Ordaz (1981),

$$g(t;z)=\frac{\dfrac{1}{R^c(t;z)}-1}{c} \tag{1}$$

Where c is the transformation parameter. In this PH-PO model, when $c\to 0$, it becomes the proportional hazard model as $g_{c\to 0}(t;z)=-\ln R(t;z)$; and when $c=1$, it becomes the proportional odds model as $g_{c=1}(t;z)=\dfrac{F(t;z)}{R(t;z)}$.

Stress effect function describes how stresses affect the baseline function of a product. In the proportional hazard model, it is expressed as $\lambda(t;z)=\lambda_0(t)e^{\beta z}$ while it is expressed as $\theta(t;z)=\theta_0(t)e^{\beta z}$ for the proportional odds model. In general, they have the similar form,

$$g(t;z)=g_0(t)e^{\beta z} \tag{2}$$

Which implies that the stresses have multiplicative effects on the baseline function $g_0(t)$.

From Table 1 we observe that the 1 d.f. baseline functions of PHM and POM are in a similar form. Therefore, we set the baseline function of the PH-PO model as follows,

$$\ln g_0(t) = a \ln t + b \tag{3}$$

Where a and b are unknown parameters.

Rewriting Eq.(3),

$$g_0(t) = t^a e^b \tag{4}$$

Accordingly, we describe the baseline function of the PH-PO model as,

$$g_0(t) = \gamma_1 t^{\gamma_2} \tag{5}$$

Where $\gamma_1 > 0$ and $\gamma_2 > 0$ are unknown parameters.

From Eq.(1), we obtain the reliability function and probability density function using $g(t;z)$ as,

$$R_c(t;z) = [cg(t;z)+1]^{-\frac{1}{c}} \tag{6}$$

$$f_c(t;z) = [cg(t;z)+1]^{-\frac{1}{c}-1} g'(t;z) \tag{7}$$

Where $g(t;z) = g_0(t)e^{\beta z}$, $g_0(t) = \gamma_1 t^{\gamma_2}$.

2.2 Special cases

The parameter c is called the transformation parameter as it determines the form of the PH-PO model. As shown in Table 1, the PHM with 1 d.f. is a Weibull model and the POM with 1 d.f. is a log-logistic model, so the estimation of proportional hazard model ($c \to 0$) is more accurate when the underlying lifetime distribution is Weibull, and it is also more accurate for the proportional odds model ($c=1$) when the underlying lifetime distribution is log-logistic. Therefore, the transformation parameter c (unknown) is an indicator of the nature of the failure times. For example, if the estimated value of c approaches 0, then the failure time distribution is Weibull. On the other hand, if c approaches 1, then the distribution is log-logistic. Other values approaching 0.5 from above or under indicate that the test units are a mixture of different distributions.

3 Estimation of the parameters

The unknown parameters of the PH-PO model are $\gamma_1, \gamma_2, \beta$, and c. We use the maximum likelihood estimation procedure to estimate these parameters. The likelihood function is,

$$L(\gamma_1, \gamma_2, \beta, c) = \prod_{i=1}^{r} f_c(t_i; z_i) \prod_{i=n-r+1}^{n} R_c(t_i; z_i) \tag{8}$$

Log-likelihood function is,

$$l(\gamma_1, \gamma_2, \beta, c) = \sum_{i=1}^{r} \ln f_c(t_i; z_i) + \sum_{i=n-r+1}^{n} \ln R_c(t_i; z_i) \tag{9}$$

Where n is the total number of the units under test, r is the number of failures before censoring and z is the related stress vector. The maximization of Eq. (9) results in the estimates of the unknown parameters. The variance of the reliability estimate is obtained by construct the Fisher information matrix below.

$$F_c = \sum_{i=1}^{n} \begin{bmatrix} -E[\frac{\partial^2 l}{\partial \gamma_1^{\ 2}}] & -E[\frac{\partial^2 l}{\partial \gamma_1 \gamma_2}] & 0 & 0 \\ -E[\frac{\partial^2 l}{\partial \gamma_1 \gamma_2}] & -E[\frac{\partial^2 l}{\partial \gamma_2^{\ 2}}] & 0 & 0 \\ 0 & 0 & -E[\frac{\partial^2 l}{\partial \beta^2}] & 0 \\ 0 & 0 & 0 & -E[\frac{\partial^2 l}{\partial c^2}] \end{bmatrix} \tag{10}$$

Following the delta optimality criterion, we obtain the asymptotic variance of the reliability estimation $\operatorname{var} \hat{R}_c(t; z_D)$ as,

$$\begin{aligned} \operatorname{var} \hat{R}_c(t; z_D) = & \left[\frac{\partial R_c(t; z_D)}{\partial \gamma_1}, \frac{\partial R_c(t; z_D)}{\partial \gamma_2}, \frac{\partial R_c(t; z_D)}{\partial \beta}, \frac{\partial R_c(t; z_D)}{\partial c} \right] \\ & \times F_c^{-1} \left[\frac{\partial R_c(t; z_D)}{\partial \gamma_1}, \frac{\partial R_c(t; z_D)}{\partial \gamma_2}, \frac{\partial R_c(t; z_D)}{\partial \beta}, \frac{\partial R_c(t; z_D)}{\partial c} \right]^T \Bigg|_{(\gamma_1, \gamma_2, \beta, c) = (\hat{\gamma}_1, \hat{\gamma}_2, \hat{\beta}, \hat{c})} \end{aligned} \tag{11}$$

Where z_D is the stress vector under normal operating condition.

4 Simulation and model validation

4.1 Data generation

Consider n units which are divided into k groups to be tested under k different stress levels. There are m different types of stresses expressed as vector z. The units are from a mixed or non-mixed distribution of Weibull and log-logistic. Based on accelerated failure time model, the reliability of the units can be expressed as the weighted sum of two distributions: Weibull and log-logistic,

$$R(t; z) = w \frac{1}{1 + (\lambda t e^{\alpha z})^p} + (1 - w) e^{-(\frac{t}{\eta} e^{\alpha z})^m} \tag{12}$$

Where $e^{\alpha z}$ is the acceleration factor, and w is the weight parameter. When w=0, the underlying distribution is Weibull; when w=1, the underlying distribution is log-logistic. When w is between 0 and 1, the underlying distribution is a mixture of

Weibull and log-logistic.

By processing the inverse transform of Eq. (12) under the condition that w is equal to 0 and 1, we obtain

$$t_{w=0} = \eta e^{-\alpha z} (\ln \frac{1}{1-F(t;z)})^{\frac{1}{m}} \tag{13}$$

$$t_{w=1} = \frac{1}{\lambda} e^{-\alpha z} (\frac{1}{1-F(t;z)} - 1)^{\frac{1}{p}} \tag{14}$$

$F(t;z)$ is generated by random numbers (*rand*) of a uniform distribution.

The simulated failure time data can then be obtained as,

$$t = \begin{cases} t_{w=0}, F(t) = rand1(wn) \\ t_{w=1}, F(t) = rand2[(1-w)n] \end{cases} \tag{15}$$

Here we set $n = 150$, $k = 3$, z (the applied stress value) are selected to be three temperature levels: 80℃, 150℃ and 200℃ which are transformed to $100K^{-1}$ with corresponding values of 0.283, 0.236, 0.211 respectively. We also generate 50 failure time data under normal operating temperature of 25℃. We set w equals to 0, 1, and 0.5, respectively. The values of the other parameters are set as follows: $\lambda = 12.18$, $p = 10$, $m = 3$, $\eta = 0.1$, $\alpha = -30$.

4.2 Weibull distribution

Set w=0 and obtain failure time data according to the Weibull distribution. The parameters of the model are obtained as $\gamma_1 = 6884.64$, $\gamma_2 = 3.67$, $\beta = -111.93$, $c = 0.07$. We use Kaplan-Meier estimation method to analyze the failure time data at the normal operating condition of 25℃ and estimate the reliability as $c \to 0$ (PHM).

Figure 1 shows the estimations of reliability for the three cases.

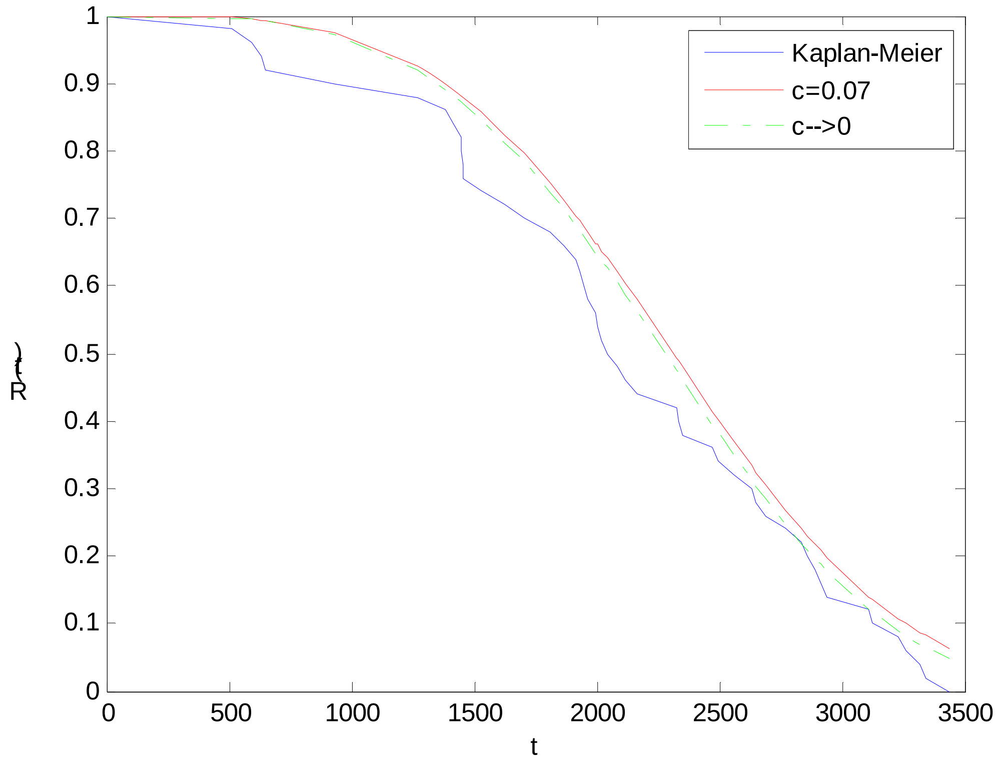

Figure 1 Comparison of reliability estimations under different *c* values of Weibull distribution

The variances of the estimations are shown in Table 2.

Table 2 Variances of estimations for Weibull distribution

c value	Variance
0.07	0.08
$c \to 0$	0.08

From Figure 1 and variances in Table 2, we conclude that the PH-PO model has the same reliability and variance estimates as those the PHM developed based on the Weibull distribution failure time data.

4.3 Log-logistic distribution

Set w=1 and obtain failure time data according to the log-logistic distribution. The parameters of the model are obtained as: $\gamma_1 = 1.72 \times 10^{12}$, $\gamma_2 = 11.20$, $\beta = -337.57$, $c = 0.92$. We use Kaplan-Meier estimation method to analyze the failure time data at the normal operating condition of 25℃ and estimate the reliability as $c \to 1$ (POM). Figure 2 shows the estimations of reliability for the three cases.

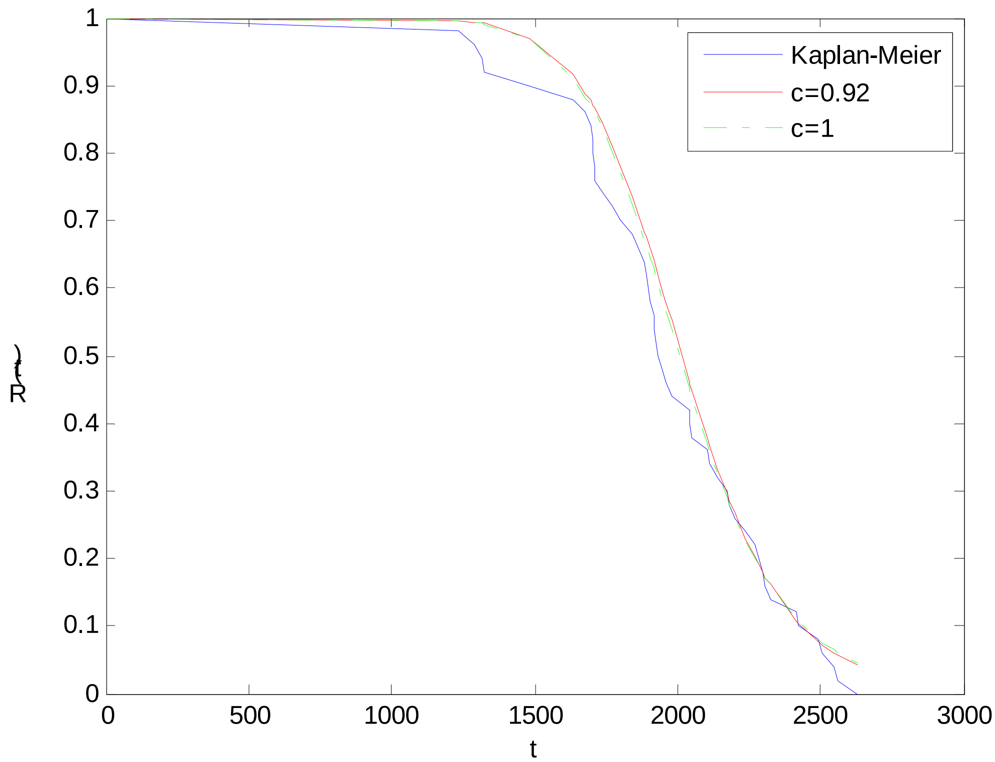

Figure 2 Comparison of reliability estimations under different *c* values of log-logistic distribution

The variances of the estimations are shown in Table 3.

Table 3 Variances of estimations for log-logistic distribution

c value	Variance
0.92	0.06
1	0.06

From Figure 2 and variances in Table 3, we conclude that the PH-PO model has the same reliability and variance estimates as those the POM developed based on the log-logistic distribution failure time data.

4.4 Mixture of distributions

Set w=0.5 then we obtain failure time data from a mixed distributions with 50% of the data is generated from a Weibull distribution and 50% of the data is generated from a log-logistic distribution. The parameters of the model are obtained as: $\gamma_1 = 99243.23$, $\gamma_2 = 5.41$, $\beta = -155.92$, $c = 0.47$.

We use Kaplan-Meier estimation method to analyze the failure time data at the normal operating condition of 25℃ and estimate the reliability. Figure 3 shows the estimations of reliability for four cases: Kaplan-Meier, PH-PO model, PHM for $c \to 0$ and POM for $c \to 1$. It is clear that the PH-PO model results in reliability estimates close to the Kaplan-Meier model.

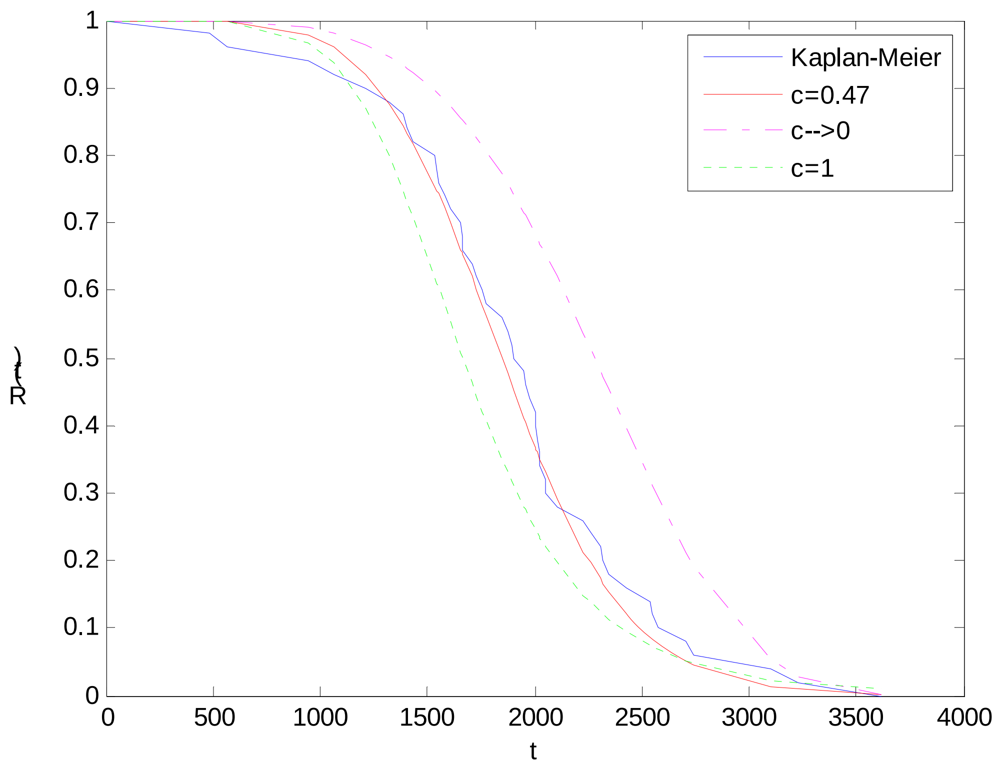

Figure 3 Comparison of reliability estimations under different *c* values of mixed distribution

The variances of the estimations are shown in Table 4.

Table 4 Variances of estimations for mixed distribution

c value	Variance
0.47	0.03
$c \to 0$	0.21
1	0.13

5 Conclusions and future work

PHM and POM are two widely used non-parametric models and they obtain accurate estimates of reliability at normal operating conditions when the underlying distributions are Weibull and log-logistic respectively. However, when the test data come from a mixture of distributions, the reliability estimate may be highly inaccurate. In this case, the PH-PO model (based on a parameter family proposed by Aranda-Ordaz) provides more accurate estimates. Simulations study shows that this PH-PO model gives a more accurate estimation than PHM and POM in general.

The range of the transformation parameter c used in this paper is form 0 to 1 which makes the PH-PO model suitable for failure time data obtained from Weibull, log-logistic distributions, and mixtures of the two distributions. Other models should be investigated for cases when c falls outside this range. The applicability of this model under the PH and PO assumptions for time varying stresses warrant further investigation.

References

Aranda-Ordaz, F. J., 1981. On Two Families of Transformations to Additivity for Binary Response Data, *Biometrics*, Vol.68, No.2.,pp.357-363.

Brass, W., 1971. On the Scale of Mortality, In: Brass, W., editor. Biological Aspects of Mortality, *Symposia of the Society for the Study of Human Biology*. Volume X. London: Taylor & Francis Ltd.: 69-110.

Cox, D. R., 1972. Regression Models and Life Tables (with Discussion), *Journal of the Royal Statistical Society*, Series B, 34, pp. 187-220

Elsayed, E. A., 1996. *Reliability Engineering*, Massachusetts: Addison-Wesley.

Hastie, T. and Tibshirani, R., 1990. *Generalized Additive Models*, London and New York: Chapman and Hall.

Royston, P. and Parmar, M. K. B., 2002. Flexible Parametric Proportional-Hazards and Proportional-Odds Models for Censored Survival Data, with Application to Prognostic Modeling and Estimation of Treatment Effects, *Statist. Med.*; **21**:2175-2197.

Wang, X., 2001. An Extended Hazard Regression Model for Accelerated Life Testing with Time Varying Coefficients. Thesis (PhD). Rutgers University.

Younes, N. and Lachin,J., 1997. Link-Based Models for Survival Data with Interval and Continuous Time Censoring, *Biometrics*, Vol.53, No.4.,pp.1199-1211.

Zhang, H. and Elsayed, E. A., 2005. Nonparametric Accelerated Life Testing Based on Proportional Odds Model, *Proceedings of the 11th ISSAT International Conference on Reliability and Quality in Design*, St. Louis, Missouri, USA.

Zhang, H., 2007. *Modeling and Planning Accelerated Life Testing with Proportional Odds*. Thesis (PhD). Rutgers University.

Appendix

Log-likelihood function is,

$$l(\gamma_1,\gamma_2,\beta,c)=\sum_{i=1}^{r}\ln\ f_c(t_i;z_i)+\sum_{i=n-r+1}^{n}\ln R_c(t_i;z_i)$$

$$=\sum_{i=1}^{r}\ln\{[cg(t_i;z_i)+1]^{-\frac{1}{c}-1}g'(t_i;z_i)\}+\sum_{i=n-r+1}^{n}\ln[cg(t_i;z_i)+1]^{\frac{1}{-c}}$$

$$=\sum_{i=1}^{r}(-\frac{1}{c}-1)\ln(c\gamma_1 t_i^{\gamma_2}e^{\beta z_i}+1)+r\ln\gamma_1+r\ln\gamma_2+\sum_{i=1}^{r}(\gamma_2-1)\ln t_i+\sum_{i=1}^{r}\beta z_i$$
$$+\sum_{i=n-r+1}^{n}(-\frac{1}{c})\ln(c\gamma_1 t_i^{\gamma_2}e^{\beta z_i}+1)$$

In data generation, we use the following strategy for mixtures of distributions,

$$t_{w=0}=\eta e^{-\alpha z}(\ln\frac{1}{1-F(t;z)})^{\frac{1}{m}}$$

$$t_{w=1}=\frac{1}{\lambda}e^{-\alpha z}(\frac{1}{1-F(t;z)}-1)^{\frac{1}{p}}$$

$$t=\begin{cases}t_{w=0}, F(t)=rand1(wn)\\ t_{w=1}, F(t)=rand2[(1-w)n]\end{cases}$$

This implies that wn observations are generated according to the Weibull distribution and $(1-w)n$ observations according to the log-logistic distribution. Moreover, these two sets of observations are generated in this way so that they do not have great differences in lifetime and reliability.