Faculty Research

University of Michigan Business School

WORKING PAPER SERIES

A Mixture Model for Automobile Warranty Data

Karl D. Majeske
University of Michigan Business School

Working Paper 03-008

Research Support • Ann Arbor, MI 48109-1234 • 734/763-9266

A mixture model for automobile warranty data

Karl D. Majeske
The University of Michigan Business School
701 Tappan Street
Ann Arbor, MI 48109-1234
Phone: 734-647-6978
E-mail: kdm@umich.edu

Abstract

This paper proposes a general mixture model framework for automobile warranty data that includes parameters for product field performance, the manufacturing and assembly process, and dealer preparation process. The model fits warranty claims as a mixture of manufacturing or assembly defects (quality problems) and usage related failures (reliability problems). The model also estimates the fraction of vehicles containing a manufacturing or assembly defect when leaving the assembly plant. This parameter measures the quality of the entire vehicle production process, i.e., component manufacturing and final assembly. The model also measures the proportion of manufacturing or assembly defects repaired by the automobile dealer prior to customer delivery. This conditional probability quantifies the ability of the vehicle preparation process to identify and repair defects prior to customer delivery. To apply the model to field failure or warranty data, the practitioner must identify parametric distributions for each of the two failure processes. To demonstrate the model, this paper develops a Weibull-Uniform mixture for manufacturer supplied warranty claim data.

Key Words: Automobile Warranty Data, Mixture Distributions, Field Failure Data, Maximum Likelihood Estimation.

A Mixture Model for Automobile Warranty Data

1. INTRODUCTION

Many consumer durable goods - such as automobiles, appliances and personal computers - include a manufacturer's warranty to insure product quality and reliability. The resulting warranty claims contain field performance data, obtained under actual operating conditions, which manufacturers use to track product lifetimes. Majeske, Lynch-Caris and Herrin [1] show how automobile manufacturers apply statistical models to warranty data to make inference regarding product design changes. Manufacturers also use warranty data to compare actual field performance with pseudo-lifetime data generated in laboratory or bench test settings [2].

Automobile manufacturers provide customers with a basic two-dimensional warranty that quantifies vehicle lifetime with two metrics: time and mileage. However, many manufacturers model field performance in the time domain due to the uncertainty associated with mileage accumulation. Wasserman [3] and Robinson and McDonald [4] define the statistic R(t) - claims per thousand vehicles reported cumulatively by month in service – a non-parametric model that automobile manufacturers use to track warranty performance. Wasserman [3] develops a dynamic linear predictive model for R(t) using data from multiple model years of a given vehicle. Robinson and McDonald [4] suggest plotting R(t) on log-log paper and fitting a line to the observed data. Singpurwalla and Wilson [5] develop a bi-variate failure model for automobile warranty data indexed by time and mileage. They derive the two marginal failure distributions and present a method for predicting R(t) using a log-log model.

Manufacturers also use parametric models for automobile warranty data. Kalbfleisch, Lawless and Robinson [6] develop a Poisson model for predicting automobile warranty claims in the time domain. Moskowitz and Chun [7] suggest using a bi-variate Poisson model to predict claims for a two dimensional warranty by fitting the cumulative Poisson parameter λ with various functions of time and mileage. Hu and Lawless [8] suggest a technique for modeling warranty claims as truncated data that assumes warranty claims follow a Poisson process. Oh and Bai [9] present a method for augmenting parametric warranty data models with selected observations from products whose lifetime exceeded the warranty period.

A Mixture Model for Automobile Warranty Data

The warranty modeling techniques cited above fit warranty data using parametric and non-parametric techniques. To make inference on product features or design changes, the authors suggest stratifying data or making model parameters a function of co-variates. Unfortunately, none of these models differentiate between failure types or provide feedback or assessment of supporting business processes. For example, a manufacturer may want to know if field failures (warranty claims) are a result of inadequate design, defects generated during component manufacturing, or errors in the assembly process. The manufacturers response or corrective actions for these various failure types are quite different.

Automobile manufacturers sell their products to automobile dealers, who in turn sell them to the end customer, which defines time zero for warranty duration [10]. Some vehicles are repaired under warranty in the interval between the two sales, generating a pre-delivery claim [4] or a claim with a negative lifetime. These negative values invalidate many parametric models – Weibull, exponential, log-Normal, etc. [11] – that reliability engineers fit to lifetime data. Many manufacturers remove pre-delivery claims from the population to allow fitting data with traditional methods, rather than developing techniques to accommodate this unconventional data that regularly occurs in automobile warranty claims.

This paper develops a warranty data model that provides manufacturers with a variety of information. The model captures information on product reliability by including a parametric model of vehicle lifetime and assesses product quality using a parameter to indicate if a vehicle functioned properly at time of manufacture. The model also has a parameter to measure the dealer preparation process – a conditional probability assessing the ability to detect defects. The mixture model presented in this paper allows including pre-delivery claims by assuming they represent manufacturing or assembly defects. The remainder of this paper has the following organization. Section 2 develops the generic mixture model framework for automobile warranty data. Section 3 outlines how to estimate the model parameters using maximum likelihood estimation and provides the log-likelihood function for the model. Section 4 demonstrates the methodology by deriving a specific Uniform-Weibull mixture for the manufacturer provided warranty data. Section 5 presents the conclusions and recommendations from this research. Appendix A & B contain the first order (gradient) and the second order (hessian matrix) partial derivatives of the Uniform-Weibull mixture distribution likelihood function. The gradient and Hessian matrix can be used for parameter estimation and standard error calculations respectively.

2. MIXTURE MODEL FOR TIME TO FIRST WARRANTY CLAIM

This paper develops a mixture model that assumes (automobile) warranty claims represent a combination of product reliability issues (usage related failures) and quality problems (manufacturing defects). A usage related failure assumes that the product functioned properly when purchased. Then, after some amount of customer usage, the product experienced a failure. It is common to model these usage related failures with lifetime distributions such as Weibull, Exponential, log-Normal, etc. The other type of warranty claim results from items that did not function properly when the customer purchased the product. Technically, there is no associated time to failure for these warranty claims; rather, the time associated with the customer identifying the defect and having it repaired.

Chung, Wu, and Herrin [12] and Meeker [13] proposed mixed Weibull models equivalent to

$$F(t)=\begin{cases} p(1-e^{-(\alpha t)^{\beta}}) & 0\le t<\infty \\ 1 & t=\infty \end{cases}$$

to fit occupational injury data and integrated circuit data respectively. These authors motivate the need for the mixtures by providing intuitive arguments regarding properties of lifetime and I take a similar approach for the automobile warranty data. Mori, Arai, Kaneko and Yoshikawa [14] model lifetime data when a product is exposed to two failure modes. In their example, the multiple failure modes result in higher early failure rates than the single failure mode case. Martin, O'Sullivan and Mathewson [15] model two types of failures (extrinsic and intrinsic) for MOS gate oxide capacitors.

A Mixture Model for Automobile Warranty Data

To develop the model, let T represent the lifetime (in the time domain) for the item under study, e.g., a complete vehicle (Ford Taurus) or vehicle component (V-8 engine). Defining the warranty coverage as t_c, the first failure results in a warranty claim when the failure occurs during the warranty coverage or $T \le t_c$. Let Z, with cumulative distribution function

$$F_Z(z) \quad z > 0,$$

represent the time to first usage related failure. This represents the failure distribution generally modeled by warranty and reliability engineers. Let Y, with cumulative distribution function,

$$F_Y(y) \quad 0 \le y \le t^*,$$

represent the time from delivery until repair for an assembly or manufacturing defect. This second distribution does not represent a time to failure in the conventional sense; rather, the time for the customer to diagnose an existing defect. Notice that this approach assumes customers have manufacturing or assembly related defects repaired early in warranty lifetime, designated as the interval (0 to t^*]. This model formulation does not include t* as a parameter estimated from the sample data. To use this approach, one must select t* prior to fitting the observational data. When selecting t*, one must address the question "How long will it take the customer to determine there is a defect and return the product for repair?" To a certain extent, this depends on the severity – how significantly it impacts product performance – of the defect. Notice that the time it takes to identify these defects and have them repaired may not follow standard lifetime distributions.

The mixture model contains two parameters that provide addition information regarding the complete vehicle production and delivery process. Let p represent the probability a given vehicle contained a manufacturing or assembly defect when leaving the assembly plant, a measure of vehicle quality. Let θ represent the fraction of defective vehicles repaired by the dealership. This conditional probability captures the dealership's ability to identify and correct vehicle defects prior to customer delivery. Assuming independence, the proportion of vehicles that experience a pre-delivery claim is $p\theta$. The probability of a manufacturing or assembly defect related claim occurring in the interval (0, t^*] is $p(1-\theta)$. While not direct parameters of this model, these two probabilities may be useful to reliability engineers, sales and marketing staff and financial planners.

The mixture model divides the time axis into three regions. In the pre-delivery region, $T \le 0$, the model assumes that all warranty claims represent manufacturing or assembly defects. Warranty claims in the interval (0, t^*] represent a mixture of the two failure types. Lastly, all claims with a lifetime $T > t^*$ represent usage related failures. From the above definitions and assumptions, I directly derive the cumulative distribution function for T as

$$F_T(t) = \begin{cases} p\theta & t \le 0 \\ p\theta + p(1-\theta)F_y(t) + (1-p)F_z(t) & 0 < t \le t^* \\ p + (1-p)F_z(t) & t > t^* \end{cases} \qquad (1)$$

One can easily extend the model by allowing p, θ and the parameters that characterize $F_Z(z)$ and $F_Y(y)$ to be functions of covariates.

3. ESTIMATING MODEL PARAMETERS

Jiang and Kececioglu [16] describe the two types of observations that exist with mixture processes. Postmortem describes the situation when an observation indicates the process that contributed the data while a non-postmortem observation provides no information on sub-population. For estimation purposes, claims prior to customer delivery and after t^* are post-mortem observations since pre-delivery claims represent assembly or manufacturing defects and claims after t^* represent usage related failures. All failures in the interval (0, t^*) represent non-postmortem observations unless the warranty claim records contain sufficient information to determine failure type. Vehicles in the population that do not have a warranty claim can be modeled as right censored observations. Estimating parameters via maximum likelihood allows incorporating the right censored observations.

Let T_i, $i = 1,2,\ldots,n$, represent the time to first claim for a population of n vehicles. Let the censor code $\delta_i = 1$ if the ith vehicle has an observed failure and $\delta_i = 0$ if the observation is right censored. The observed data will then consist of the pairs (δ_i, t_i) for each of the n vehicles. The likelihood function is the probability of observing the sample data or

$$L = \prod_{i=1}^{n} [f(t_i)]^{\delta_i} [S(t_i)]^{1-\delta_i} \ , \tag{2}$$

with an associated log-likelihood function of

$$\log(L) = \sum_{i=1}^{n} \delta_i \log(f(t_i)) + (1-\delta_i)\log(S(t_i)) \ . \tag{3}$$

To construct the log-likelihood function for the mixture model, let L_i represent the contribution of vehicle i to the likelihood function and condition on t_i, resulting in

$$\begin{aligned}
\log(L_i | t_i = 0) &= \log(p) + \log(\theta) \\
\log(L_i \mid 0 < t_i \le T^*) &= \delta_i \log\{p(1-\theta) f_Y(t_i) + (1-p) f_Z(t_i)\} \\
&\quad + (1-\delta_i)\log\{1 - p\theta - p(1-\theta)F_Y(t_i) - (1-p)F_Z(t_i)\} \\
\log(L_i \mid t_i > T^*) &= \delta_i \log\{(1-p) f_Z(t_i)\} + (1-\delta_i)\log(1-p)F_Z(t_i)\} \ .
\end{aligned} \tag{4}$$

Maximizing the log-likelihood with respect to the parameters results in the maximum likelihood estimates.

4. APPLICATION TO AUTOMOBILE WARRANTY DATA

This paper uses a population of 9532 luxury cars, representing one month of production, to demonstrate the methodology. One subsystem of these vehicles was experiencing higher than expected levels of warranty claims, and the manufacturer provided me with the data available about two years after final vehicle assembly. I constructed a database containing one row for each vehicle, with columns defined by Table 1, to analyze the luxury car warranty data. Figure 1 shows the empirical hazard function $\hat{h}(t)$ for the luxury cars calculated with life tables using an interval width of 7 days [17]. Unfortunately, the method does not allow for negative values so Figure 1 does not capture the 107 vehicles – representing 1.12% of the population – that experienced pre-delivery claims.

To apply the mixture, one must identify parametric models for the two distributions comprising the mixture. Product design and reliability engineers model the field performance in the time domain with the Weibull distribution. Therefore, I model the usage related failure component of the mixture with a Weibull distribution characterized by cumulative distribution function

$$F_Z(t) = 1 - \exp[-(\alpha t)^{\beta}] \qquad t > 0, \alpha > 0, \beta > 0.$$

Manufacturing or assembly defects existed at customer delivery, therefore, the lifetime associated with these warranty claims represents the time to identify the problem and have the vehicle repaired. The manufacturer suggested that the customer is equally likely to have these defects repaired over the first few months of ownership. Notice from Figure 1 that two distinct claim levels exist: the initial high failure rate prior to 119 days and the ongoing relatively lower failure rate above 119 days. After showing this figure to the manufacturer, we agreed to use t^{*} = 119 days and model time to repair for inherent defects with a uniform distribution characterized by

$$F_Y(t) = \frac{t}{t^*} \qquad 0 \le t \le t^*.$$

Using the Weibull and uniform as inputs to the generic model of equation (1) leads to the specific Weibull-Uniform mixture

$$F_T(t) = \begin{cases} p\theta & t = 0 \\ p\theta + p(1\text{-}\theta)\dfrac{t}{t^*} + (1-p)(1-e^{-(\alpha t)^{\beta}}) & 0 < t \le t^* \\ p + (1-p)(1-e^{-(\alpha t)^{\beta}}) & t > t^* \end{cases} \tag{5}$$

with an associated log-likelihood of

$$\begin{aligned}
\log(L_i | t_i = 0) &= \log(p) + \log(\theta) \\
\log(L_i | 0 < t_i \le T^*) &= \delta_i \log\left(\frac{p(1-\theta)}{T^*} + (1-p)[\alpha\beta(\alpha t_i)^{\beta-1} e^{-(\alpha t_i)^{\beta}}]\right) \\
&\quad + (1-\delta_i)\log\left(1 - p\left(\theta + \frac{(1-\theta)t_i}{T^*}\right) - (1-p)(1-e^{-(\alpha t_i)^{\beta}})\right) \\
\log(L_i | t_i > T^*) &= \delta_i[\log((1-p)\alpha\beta(\alpha t_i)^{\beta-1} e^{-(\alpha t_i)^{\beta}})] \\
&\quad + (1-\delta_i)[\log((1-p)e^{-(\alpha t_i)^{\beta}})] \\
&= \log(1-p) - (\alpha t_i)^{\beta} + \delta_i(\beta\log(\alpha t_i) + \log(t_i) + \log(\beta)).
\end{aligned} \tag{6}$$

Appendix A contains the gradient of equation (6) required by iterative maximization search routines that could be used for parameter estimation. To make inference on model parameters one can use large sample theory, that states under certain regularity conditions, the estimated parameter vector $(\hat{\alpha},\hat{\beta},\hat{p},\hat{\theta})$ follows a multivariate normal with mean (α,β,p,θ) and variance-covariance matrix equal to the inverse of the Fisher information matrix. Estimating the Fisher information matrix with the observed data allows making statistical inference (hypothesis tests and confidence intervals) on model parameters. Appendix B contains the second order partial derivatives for equation (6) needed to calculate the observed Fisher information matrix.

I used Hide and Seek, a simulated annealing algorithm developed by Romeijn and Smith [18], to estimate parameters by directly maximizing the log-likelihood function. Table 2 contains the parameter estimates and associated standard errors for the Weibull-Uniform mixture fit to the luxury car warranty data. Figure 2 adds the mixture hazard function characterized with the parameter estimates from Table 2 to the empirical hazard plot of Figure 1. Figure 2 provides a visual assessment of the model aptness or goodness of fit. To further assess the model appropriateness, I used simulate a set of 9532 observations from a mixture distribution with cumulative distribution function of equation 5 using the parameter values from Table 2. Figure 3 shows the empirical hazard function for this data that appears quite similar to the observed automobile warranty data.

The fit model suggests that 5.6% of the vehicles left the assembly plant with the component under study not functioning properly. This rate of more than one in 20 defects lead the manufacturer to take two courses of action. First, the manufacturer investigated ways to reduce vehicle defects and implemented changes in the final assembly process. Secondly, until the vehicle defect rate decreases, the manufacturer added a test procedure to identify defective vehicles for repair at the assembly plant. This analysis also suggests that the dealer preparation process only identified 26.1% of the defective vehicles. This led the manufacturer to review the ability of the dealership to identify defective vehicles.

5. CONCLUSIONS AND RECOMMENDATIONS

A Mixture Model for Automobile Warranty Data

This research suggests that manufacturers can develop richer statistical models for warranty data that provide better fits and additional feedback. For example, the mixture model presented in this paper provides the manufacturer the ability to separate two failure modes and include pre-delivery claims. In addition, this mixture provides the manufacturer quantitative measures of vehicle quality when leaving the assembly plant and the dealer preparation process. Using a similar approach, one could develop higher order mixtures or models with parameters to represent other business processes of interest. This paper develops a specific Uniform / Weibull mixture to demonstrate how to apply this modeling framework to a specific problem. The distribution easily extends to include co-variates by making p, θ, t^* and the Weibull parameters α, β functions of engineering design and assembly process parameters.

APPENDIX A

GRADIENT OF THE LOG LIKELIHOOD FUNCTION FOR THE THREE REGIONS

Assume a set of N vehicles: n_1 vehicles with pre-delivery claims, n_2 vehicles with unknown failure types, and n_3 vehicles with usage related failure claims. Then $N = n_1 + n_2 + n_3$.

The n_1 vehicles with pre-delivery claims ($t_i \le 0$):

$$\frac{\partial \log(L)}{\partial \alpha} = 0$$

$$\frac{\partial \log(L)}{\partial \beta} = 0$$

$$\frac{\partial \log(L)}{\partial p} = \frac{n_1}{p}$$

$$\frac{\partial \log(L)}{\partial \theta} = \frac{n_1}{\theta}$$

The n_2 vehicles with unknown failure types ($0 < t_i \le t^*$):

$$\frac{\partial \log(L)}{\partial \alpha} = \sum_{i=1}^{n_2} \delta_i \left\{ \frac{(1-p)\beta^2 (\alpha t_i)^{\beta-1} e^{-(\alpha t_i)^\beta} [1-(\alpha t_i)^\beta]}{\frac{p(1-\theta)}{t^*} + (1-p)[\alpha\beta(\alpha t_i)^{\beta-1} e^{-(\alpha t_i)^\beta}]} \right\}$$

$$+ (1-\delta_i)\left\{ \frac{(p-1)\beta t_i (\alpha t_i)^{\beta-1} e^{-(\alpha t_i)^\beta}}{1 - p(\theta + \frac{(1-\theta)t_i}{t^*}) - (1-p)(1-e^{-(\alpha t_i)^\beta})} \right\}$$

$$\frac{\partial \log(L)}{\partial \beta} = \sum_{i=1}^{n_2} \delta_i \left\{ \frac{(1-p)\alpha(\alpha t_i)^{\beta-1} e^{-(\alpha t_i)^\beta} [1-\beta(\alpha t_i)^\beta \log(\alpha t_i) + \beta \log(\alpha t_i)]}{\frac{p(1-\theta)}{t^*} + (1-p)[\alpha\beta(\alpha t_i)^{\beta-1} e^{-(\alpha t_i)^\beta}]} \right\}$$

$$+ (1-\delta_i)\left\{ \frac{(p-1)(\alpha t_i)^\beta e^{-(\alpha t_i)^\beta} \log(\alpha t_i)}{1 - p(\theta + \frac{(1-\theta)t_i}{t^*}) - (1-p)(1-e^{-(\alpha t_i)^\beta})} \right\}$$

$$\frac{\partial \log(L)}{\partial p} = \sum_{i=1}^{n_2} \delta_i \left\{ \frac{\left(\frac{1-\theta}{t^*}\right) - \alpha\beta(t_i\alpha)^{\beta-1} e^{-(\alpha t_i)^\beta}}{\frac{p(1-\theta)}{t^*} + (1-p)[\alpha\beta(\alpha t_i)^{\beta-1} e^{-(\alpha t_i)^\beta}]} \right\} + (1-\delta_i) \left\{ \frac{1-\theta - \left(\frac{(1-\theta)t_i}{t^*}\right) - e^{-(\alpha t_i)^\beta}}{1 - p(\theta + \frac{(1-\theta)t_i}{t^*}) - (1-p)(1-e^{-(\alpha t_i)^\beta})} \right\}$$

$$\frac{\partial \log(L)}{\partial \theta} = \sum_{i=1}^{n_2} \delta_i \left\{ \frac{\frac{-p}{t^*}}{\frac{p(1-\theta)}{t^*} + (1-p)[\alpha\beta(\alpha t_i)^{\beta-1} e^{-(\alpha t_i)^\beta}]} \right\} + (1-\delta_i) \left\{ \frac{p(\frac{t_i}{t^*} - 1)}{1 - p(\theta + \frac{(1-\theta)t_i}{t^*}) - (1-p)(1-e^{-(\alpha t_i)^\beta})} \right\}$$

The n_3 vehicles with usage related failure claims ($t_i > t^*$):

$$\frac{\partial \log(L)}{\partial \alpha} = \sum_{i=1}^{n_3} -\beta\alpha^{\beta-1} t_i^{\beta} + \delta_i \frac{\beta}{\alpha}$$

$$\frac{\partial \log(L)}{\partial \beta} = \sum_{i=1}^{n_3} -(\alpha t_i)^\beta \ln(\alpha t_i) + \delta_i (\ln(\alpha) + \frac{1}{\beta} + \ln(t_i))$$

$$\frac{\partial \log(L)}{\partial p} = \frac{-n_3}{1-p}$$

$$\frac{\partial \log(L)}{\partial \theta} = 0$$

A Mixture Model for Automobile Warranty Data

APPENDIX B

SECOND ORDER PARTIAL DERIVATIVES OF THE LOG LIKELIHOOD FUNCTION FOR THE THREE REGIONS

Second order partial derivatives of the log likelihood function for each of the three regions. Assume a set of N vehicles, n_1 vehicles with pre-delivery claims, n_2 vehicles with unknown failure types, n_3 vehicles with usage related failures type claims. Then $N = n_1 + n_2 + n_3$. Below appear the upper right entries of the symmetric matrix by region.

The n_1 vehicles with pre-delivery claims ($t_i \leq 0$):

$$\frac{\partial^2 \log(L)}{\partial \alpha^2} = \frac{\partial^2 \log(L)}{\partial \alpha \partial \beta} = \frac{\partial^2 \log(L)}{\partial \alpha \partial p} = \frac{\partial^2 \log(L)}{\partial \alpha \partial \theta} = 0$$
$$\frac{\partial^2 \log(L)}{\partial \beta^2} = \frac{\partial^2 \log(L)}{\partial \beta \partial p} = \frac{\partial^2 \log(L)}{\partial \beta \partial \theta} = 0$$
$$\frac{\partial^2 \log(L)}{\partial p^2} = \frac{-n_1}{p^2} \qquad \frac{\partial^2 \log(L)}{\partial p \partial \theta} = 0$$

$$\frac{\partial^2 \log(L)}{\partial \theta^2} = \frac{-n_1}{\theta^2}$$

The n_2 vehicles with unknown failure types ($0 < t_i \leq t^*$):

Define the following identities :

$$C_{1i} = (1-p)\beta(\alpha t_i)^{\beta-1} e^{-(\alpha t_i)^\beta}$$

$$C_{2i} = -\theta - \frac{(1-\theta)t_i}{T^*} + 1 - e^{-(\alpha x_i)^\beta}$$

$$C_{3i} = (1-p)(\alpha t_i)^\beta \log(\alpha t_i)^2 e^{-(\alpha t_i)^\beta}$$

$$C_{4i} = \frac{1}{\beta} + \log(\alpha t_i) - (\alpha t_i)^{\beta} \log(\alpha t_i)$$

$$C_{5i} = \frac{1-\theta}{t^*} - \alpha\beta(\alpha t_i)^{\beta-1} e^{-(\alpha t_i)^{\beta}}$$

$$C_{6i} = (\alpha t_i)^{\beta} \log(\alpha t_i) - 3\log(\alpha t_i) - \frac{2}{\beta}$$

$$D_{1i} = \frac{p(1-\theta)}{t*} + (1-p)[\alpha\beta(\alpha t_i)^{\beta-1} e^{-(\alpha t_i)^{\beta}}]$$

$$D_{2i} = 1 - p(\theta + \frac{(1-\theta)t_i}{t*}) - (1-p)(1-e^{-(\alpha t_i)^{\beta}})$$

$$\begin{aligned}\frac{\partial^2 \log(L)}{\partial\alpha^2} = \sum_{i=1}^{n_2} & \delta_i \frac{C_{1i}}{\alpha D_{1i}}(\beta^2 - \beta + (\alpha t_i)^{\beta}[\beta^2((\alpha t_i)^{\beta} - 3) + \beta]) \\ & - \delta_i \left(\frac{C_{1i}}{D_{1i}}\right)^2 (\beta - \beta(\alpha t_i)^{\beta})^2 \\ & + (1-\delta_i)\frac{t_i C_{1i}}{\alpha D_{2i}}(1 - \beta + (\alpha t_i)^{\beta}\beta - \frac{\alpha t_i C_{1i}}{D_{2i}})\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 \log(L)}{\partial\alpha\partial\beta} = \sum_{i=1}^{n_2} & \delta_i \frac{C_{1i}}{D_{1i}}(C_{4i}\beta(1 - 2(\alpha t_i)^{\beta}) - \beta\left((\alpha t_i)^{\beta}\right)^2 \log(\alpha t_i)) \\ & - \delta_i \left(\frac{C_{1i}}{D_{1i}}\right)^2 (1 - (\alpha t_i)^{\beta})\alpha C_{4i} \\ & - (1-\delta_i)\frac{t_i C_{1i}}{D_{2i}}(C_{4i} + \frac{\alpha t_i C_{1i} \log(\alpha t_i)}{\beta D_{2i}})\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 \log(L)}{\partial\alpha\partial p} = \sum_{i=1}^{n_2} & \delta_i \left(\frac{C_{1i}\beta[(\alpha t_i)^{\beta} - 1]}{(1-p)D_{1i}} - \frac{C_{1i}C_{5i}\beta[1 - (\alpha t_i)^{\beta}]}{D_{1i}^{\,2}}\right) \\ & + (1-\delta_i)\left(\frac{C_{1i}t_i}{(1-p)D_{2i}} + \frac{C_{1i}t_i C_{2i}}{D_{2i}^{\,2}}\right)\end{aligned}$$

$$\frac{\partial^2 \log(L)}{\partial\alpha\partial\theta} = \sum_{i=1}^{n_2} \delta_i \frac{C_{1i}\, p\beta[1-(\alpha t_i)^\beta]}{D_{1i}^2 t^*} - (1-\delta_i)\frac{C_{1i}\alpha t_i\, p(1-\frac{t_i}{t^*})}{\alpha D_{2i}^2}$$

$$\frac{\partial^2 \log(L)}{\partial\beta^2} = \sum_{i=1}^{n_2} \delta_i \frac{C_{1i}\alpha\log(\alpha t_i)}{D_{1i}}\left(\frac{2}{\beta} + \log(\alpha t_i) + (\alpha t_i)^\beta [C_{6i}]\right)$$
$$- \delta_i\left(\frac{\alpha C_{1i}C_{4i}}{D_{1i}}\right)^2 + (1-\delta_i)\frac{C_{3i}}{D_{2i}}\left((\alpha t_i)^\beta - 1 - \frac{C_{1i}\alpha t_i}{\beta D_{2i}}\right)$$

$$\frac{\partial^2 \log(L)}{\partial\beta\partial p} = \sum_{i=1}^{n_2} -\delta_i\left(\frac{\alpha C_{1i}C_{4i}}{(1-p)D_{1i}} + \frac{\alpha C_{1i}C_{4i}C_{5i}}{D_{1i}^2}\right)$$
$$+ (1-\delta_i)\left(\frac{C_{3i}}{(1-p)\log(\alpha t_i)D_{2i}} + \frac{C_{2i}C_{3i}}{\log(\alpha t_i)D_{2i}^2}\right)$$

$$\frac{\partial^2 \log(L)}{\partial\beta\partial\theta} = \sum_{i=1}^{n_2} \delta_i \frac{\alpha C_{1i}\, pC_{4i}}{D_{1i}^2 t^*} - (1-\delta_i)\frac{C_{3i}\, p(1-\frac{t_i}{t^*})}{\log(\alpha t_i)D_{2i}^2}$$

$$\frac{\partial^2 \log(L)}{\partial p^2} = \sum_{i=1}^{n_2} -\delta_i \frac{C_{5i}^2}{D_{1i}^2} - (1-\delta_i)\frac{C_{2i}^2}{D_{2i}^2}$$

$$\frac{\partial^2 \log(L)}{\partial p\partial\theta} = \sum_{i=1}^{n_2} \delta_i\left(\frac{pC_{5i}}{t^* D_{1i}^2} - \frac{1}{t^* D_{1i}}\right) + (1-\delta_i)\left(\frac{\frac{t_i}{t^*} - 1}{D_{2i}} + \frac{C_{2i}\, p\left(1-\frac{t_i}{t^*}\right)}{D_{2i}^2}\right)$$

$$\frac{\partial^2 \log(L)}{\partial\theta^2} = \sum_{i=1}^{n_2} -\delta_i\left(\frac{p}{t^* D_{1i}}\right)^2 - (1-\delta_i)\left(\frac{p\left(1-\frac{t_i}{t^*}\right)}{D_{2i}}\right)^2$$

The n_3 vehicles with usage related failure claims ($t_i > t^*$):

$$\frac{\partial^2 \log(L)}{\partial \alpha^2} = \sum_{i=1}^{n_3} \frac{\beta}{\alpha^2}\left((\alpha t_i)^\beta - \beta(\alpha t_i)^\beta - \delta_i\right)$$

$$\frac{\partial^2 \log(L)}{\partial \alpha \partial \beta} = \sum_{i=1}^{n_3} \frac{1}{\alpha}\left(\delta_i - \beta(\alpha t_i)^\beta \log(\alpha t_i) - (\alpha t_i)^\beta\right)$$

$$\frac{\partial^2 \log(L)}{\partial \alpha \partial p} = \frac{\partial^2 \log(L)}{\partial \alpha \partial \theta} = 0$$

$$\frac{\partial^2 \log(L)}{\partial \beta^2} = \sum_{i=1}^{n_3} -\frac{\delta_i}{\beta^2} - (\alpha t_i)^\beta \log(\alpha t_i)^2$$

$$\frac{\partial^2 \log(L)}{\partial \beta \partial p} = \frac{\partial^2 \log(L)}{\partial \beta \partial \theta} = 0$$

$$\frac{\partial^2 \log(L)}{\partial p^2} = \frac{-n_3}{(1-p)^2} \qquad \frac{\partial^2 \log(L)}{\partial p \partial \theta} = 0$$

$$\frac{\partial^2 \log(L)}{\partial \theta^2} = 0$$

REFERENCES

[1] K. D. Majeske, T. Lynch-Caris, and G. D. Herrin, "Evaluating Product and Process design changes with warranty data", *International Journal of Production Economics*, Vol. 50, 1997, pp. 79-89.

[2] K. D. Majeske, and G. D. Herrin, "Determining Warranty Benefits for Automobile Design Changes", *Proceedings of the Annual Reliability and Maintainability Symposium*, 1998.

[3] G. S. Wasserman, "An Applicaiton of Dynamic Linear Models for Predicting Warranty Claims", *Computers & Industrial Engineering*, Vol. 22, No. 1, 1992, pp. 37-47.

[4] J. A. Robinson and G. C. McDonald, "Issues Related to Field Reliability and Warranty Data", *Data Quality Control: Theory and Pramatics, 1991*, Marcel Dekker, New York.

[5] N.D. Singpurwalla and S. Wilson, "The Warranty Problem: Its Statistical and Game Theoretic Aspects", *SIAM Review*, Vol. 35, 1993, pp. 17-42.

[6] J.D. Kalbfleisch, J.F. Lawless, and J.A. Robinson, "Methods for the Analysis and Prediction of Warranty Claims", *Technometrics*, Vol. 33, 1991, pp. 273-285.

[7] H. Moskowitz and Y.H. Chun, "A Poisson Regression Model for Two-Attribute Warranty Policies", *Naval Research Logistics*, Vol. 41, 1994, pp. 355-376.

[8] X.J. Hu and J.F. Lawless, "Estimations of Rate and Mean Functions From Truncated Recurrent Event Data", *Journal of the American Statistical Association,* Vol. 91, No. 433, 1996, pp. 300-310.

[9] Y.S. Oh and D.S. Bai, "Field data analysis with additional after-warranty failure data", *Reliability Engineering and System Safety,* Vol. 72, No. 1, 2001, pp. 1-8.

[10] J. P. Womack, D. T. Jones, and D. L. Roos, *The Machine That Changed the World: The Story of Lean Production*, 1990; HarperPerennial.

[11] W. Q. Meeker and L. E. Escobar, *Statistical Methods for Reliability Data*, 1998; John Wiley and Sons.

[12] M.K. Chung, S. H. Wu, and G. D. Herrin, "The use of a Mixed Weibull Model in Occupational Injury Analysis", *Journal of Occupational Accidents*, Vol. 7, 1986, pp. 239-250.

[13] W. Q. Meeker, "Limited Failure Population Life Tests: Application to Integrated Circuit Reliability", *Technometrics*, Vol. 29, No. 1, 1987, pp. 51-65.

[14] Mori, S., Arai, N., Kaneko, Y., and Yoshikawa, K., "Polyoxide Thinning Limitation and Superior ONO Interpoly Dielectric of Nonvolatile Memory Device," *IEEE Transactions on Electron. Devices*, Vol. 38, 1991, pp. 270-276.

[15] Martin, A., O'Sullivan, P., and Mathewson, A., "Study of Unipolar Pulsed Ramp and Combined Ramped/ Constant Voltage Stress on Mos Gate Oxides," *Microelectron Reliability*, Vol. 37, 1997, pp. 1045-1051.

[16] S. Jiang, and D. Kececioglu, "Maximum Likelihood Estimates, from Censored Data, for Mixed-Weibull Distributions", *IEEE Transactions on Reliability*, Vol. 41, No. 2, 1992, pp. 248-255.

[17] J. F. Lawless, *Statistical Models and Methods for Lifetime Data*, 1982; John Wiley and Sons.

[18] H. E. Romeijn, and R. L. Smith, "Simulated Annealing for Constrained Global Optimization", *Journal of Global Optimization*, Vol. 5, 1994, pp. 101-126.

Figures and Tables

Table 1: **Test Data for Luxury Cars**

Variable	*Description*
VIN	Vehicle Identification Number
t	Time from final sale to first claim in days
δ	Censor Code: 0 = Observed, 1 = Censored

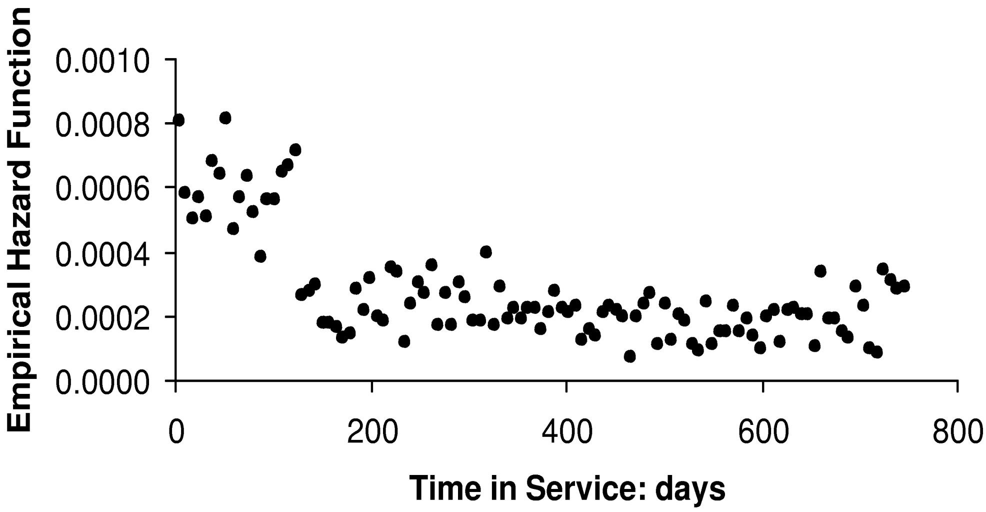

Figure 1: **Empirical Hazard Function of Time to First Warranty Claim**

Table 2: **Mixture Parameter Estimates for Time to First Claim**

Parameter	Estimate	Std error
α	0.00018	0.0000043
β	0.91626	0.0167756
p	0.05604	0.003939
θ	0.26081	0.0227802

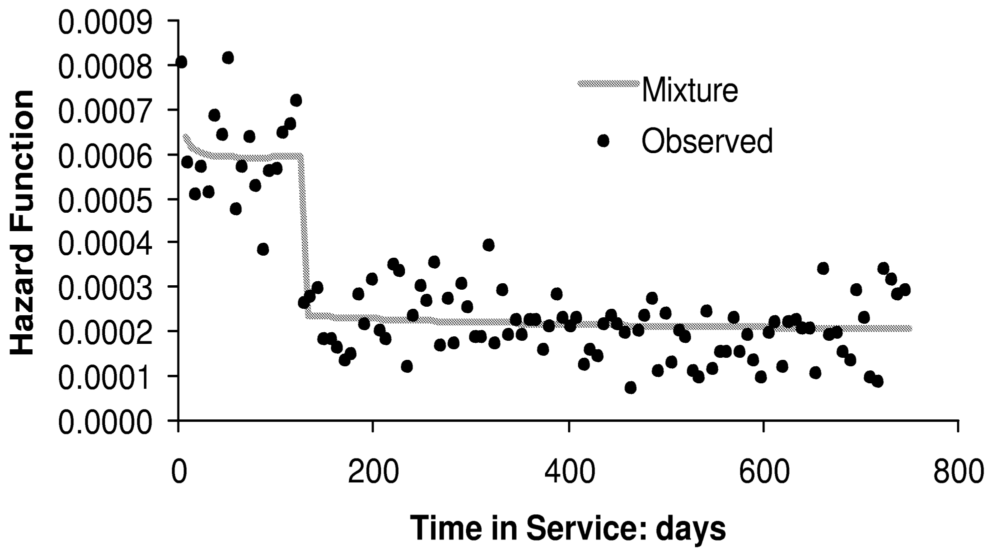

Figure 2: Mixture and Empirical Hazard Functions for Time to First Failure

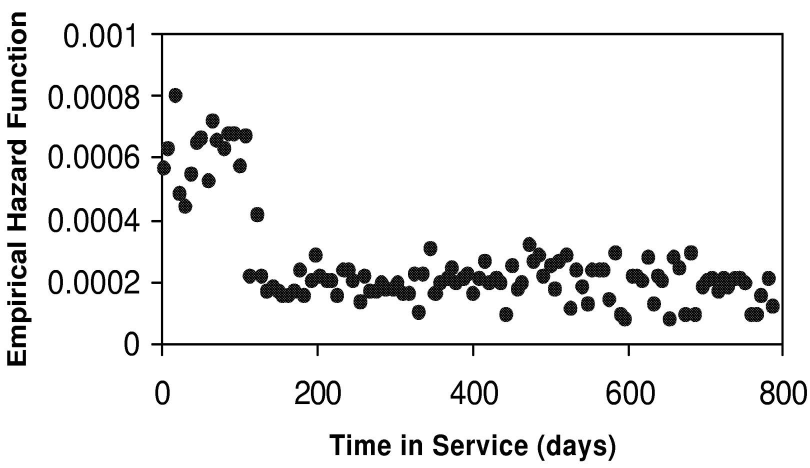

Figure 3: Empirical Hazard Function for Simulated Mixture Data