

## Empirical Bayes Analysis on the Power Law Process with Natural Conjugate Priors

Zhao Chen

*Florida Gulf Coast University*

*Abstract:* The power law process has been used extensively in software reliability models, reliability growth models and more generally reliable systems. In this paper we work on the Power Law Process via empirical Bayes (EB) approach. Based on a two-hyperparameter natural conjugate prior and a more generalized three-hyperparameter natural conjugate prior, which was stated in Huang and Bier (1998), we work out an empirical Bayes (EB) procedure and provide statistical inferences based on the natural conjugate priors. Given past experience about the parameters of the model, the empirical Bayes (EB) approach uses the observed data to estimate the hyperparameters of priors and then proceeds as though the prior were known.

*Key words:* Empirical Bayes, natural conjugate prior, power law process.

### 1. Introduction

Bayesian inference on the power law process was also studied during the past two decades. Bayesian point and interval estimates were obtained by Guida, M. (1989); Kyparisis and Singpurwalla (1985) and Calabria and Pulcini (1997). Informative and noninformative priors were both employed on failure truncated data case. Bar-Lev *et al.* (1991) discussed both time and failure truncated data by using noninformative priors. They derived prediction distributions of future failure times and the number of failures in some future time interval. Calabria (1990) also derived prediction distribution by using noninformative and informative priors. These references are given on a single system and usually assume parameters are independent. Crow (1974) and Bain (1978) analyzed independent equivalent multi-system by employing power law process. Power bounds for a test of equality of trends in several independent power law processes were discussed by Calabria, R., Guida, M. and Pulcini, G.(1992). Huang and Bier (1998) developed a natural conjugate prior for the PLP. The prior they proposed

has relatively simple closed-form expression for its moments.

When we only consider a single system, Bayesian framework can be utilized appropriately. Now suppose we have  $k$  similar but non-identical systems and each is modeled by PLP. This assumption is reasonable if  $k$  systems are made from the same manufacturing process. Parametric empirical Bayes model is proposed for the PLP with estimation of prior hyperparameters. We assume that the parameters (in our case the  $(\eta_i, \beta_i)$ s) of multiple systems are drawn from certain prior distribution  $\pi(\eta, \beta)$ . In section 2, we work on the EB procedure with employing two-hyperparameter natural conjugate prior and in section 3 three-hyperparameter natural conjugate prior. In our paper, we applied Huang's (1998) prior to obtain Bayesian maximum likelihood estimate of a hyperparameter  $a$ , posterior means of the shape parameter  $\beta$  and the scale parameter  $\eta$  in closed forms. In section 4, several contour plots and three-dimension plots for the conjugate prior illustrate some properties of the hyperparameters in the natural conjugate prior by changing hyperparameters' value. Finally in section 5, we make some remarks and conclusions.

## 2. Parametric Empirical Bayes on the PLP–Two Hyperparameters $(a, m)$

Nonhomogeneous Poisson Process (NHPP) is a Poisson process in which intensity function is not a constant. Let  $N(t)$  denote the cumulative number of failures from time 0 to time  $t$  and suppose we have a sequence of successive failure times. The power law process (PLP) can be described as a nonhomogeneous Poisson process  $\{N(t), t \geq 0\}$  with intensity function

$$v(t) = \frac{\beta}{\alpha^\beta} t^{\beta-1}, \text{ for } \alpha > 0, \beta > 0,$$

where  $\alpha$  and  $\beta$  are the scale parameter and shape parameter respectively. Hence, the mean value function  $\lambda(t)$  of the process is

$$\lambda(t) = E(N(t)) = \int_0^t v(s) ds = \int_0^t \frac{\beta}{\alpha^\beta} s^{\beta-1} ds = (t/\alpha)^\beta.$$

It can be shown that the number of failures in any interval  $(t_1, t_2]$  has a Poisson distribution with mean  $\int_{t_1}^{t_2} v(t) dt = (t_2/\alpha)^\beta - (t_1/\alpha)^\beta$ . When  $\beta < 1$ , the intensity function  $\nu(t)$  is decreasing and the system is improving. Under this situation, the power law process can be applied as a reliability growth model.

Let  $\eta = \alpha^{-\beta}$ . Parametric empirical Bayes procedures on the Power Law Process are easier to work with if the intensity function is parametrized as

$$\nu(t) = \eta\beta t^{\beta-1}, \quad t > 0.$$

The power law process has been widely used in reliability growth (Crow (1982)), and software reliability models (Kyparisis and Singpurwalla (1985)), and in repairable systems (Ascher and Feingold (1984); Engelhardt and Bain (1986); Rigdon and Basu (1989)).

Assume there are  $k$  similar systems from the same manufacturing process. Let  $\vec{t}_i$  denote the vector of failure times for system  $i$ , and let

$$T = [\vec{t}_1, \vec{t}_2, \dots, \vec{t}_k]$$

denote the two-dimensional array of failure times. Suppose that the  $i$ th system has been operated from time 0 until the  $n_i$ th failure and the  $n_i$ th failure occurs at time  $t_{in_i}$ . Clearly  $t_{i1} < t_{i2} < \dots < t_{in_i}$ . This is the failure truncated case, where  $n_i$  is a constant. The likelihood function of  $\eta, \beta$  associated with the first  $n_i$  failure times for system  $i$  is

$$L(\vec{t}_i|\eta, \beta) = L(t_{i1}, t_{i2}, \dots, t_{in_i}|\eta, \beta) = \eta^{n_i} \beta^{n_i} \left( \prod_{j=1}^{n_i} t_{ij} \right)^{\beta-1} \exp(-\eta t_{in_i}^{\beta}). \quad (2.1)$$

A natural conjugate prior distribution for the parameter  $\eta, \beta$  is given by

$$\pi_0(\eta, \beta|m, a) = c^{-1} \eta^{m-1} \beta^{m-1} (\exp(-a) t_{in_i}^m)^{\beta-1} \exp(-\eta t_{in_i}^{\beta}), \quad (2.2)$$

here  $(m, a)$  are positive hyperparameters.  $c$  is a constant.

$$\begin{aligned} c &= \int \int \eta^{m-1} \beta^{m-1} (\exp(-a) t_{in_i}^m)^{\beta-1} \exp(-\eta t_{in_i}^{\beta}) d\eta d\beta \\ &= \Gamma^2(m) a^{-m} [\exp(-a) t_{in_i}^m]^{-1}. \end{aligned}$$

Then the marginal distribution of failure time  $\vec{t}_i$  from system  $i$  given  $(m, a)$  is

$$\begin{aligned} m(\vec{t}_i|m, a) &= \int \int L(t_{i1}, t_{i2}, \dots, t_{in_i}|\eta, \beta) \pi_0(\eta, \beta|m, a) d\eta d\beta \\ &= c^{-1} \int \int \eta^{n_i+m-1} \beta^{n_i+m-1} [\exp(-a) t_{in_i}^m \prod_{j=1}^{n_i} t_{ij}]^{\beta-1} \exp(-2\eta t_{in_i}^{\beta}) d\eta d\beta \\ &= \frac{2^{-(m+n_i)} \Gamma^2(n_i + m) [a + n_i \ln(t_{in_i}) - \ln \prod_{j=1}^{n_i} (t_{ij})]^{-(n_i+m)}}{\prod_{j=1}^{n_i} t_{ij} \Gamma^2(m) a^{-m}}. \end{aligned} \quad (2.3)$$

Suppose  $k$  systems are independent, we have the marginal distribution of  $T$  given  $(m, a)$  is

$$\begin{aligned} m(T|m, a) &= m(\vec{t}_1|m, a)m(\vec{t}_2|m, a)\cdots m(\vec{t}_k|m, a) \\ &= \prod_{i=1}^k \left\{ \frac{\Gamma^2(n_i + m)[a + n_i \ln(t_{in_i}) - \ln \prod_{j=1}^{n_i} (t_{ij})]^{-(n_i+m)}}{2^{m+n_i} \prod_{j=1}^{n_i} t_{ij} \Gamma^2(m) a^{-m}} \right\}. \end{aligned} \quad (2.4)$$

In order to obtain Maximum likelihood estimates (MLEs) of  $m$  and  $a$ , we need to take a logarithm of (2.4)

$$\begin{aligned} \ln[m(T|m, a)] &= \sum_{i=1}^k \left\{ - (m + n_i) \ln 2 + 2 \ln \Gamma(n_i + m) \right. \\ &\quad \left. - (n_i + m) \ln [a + n_i \ln(t_{in_i}) - \ln \prod_{j=1}^{n_i} (t_{ij})] \right. \\ &\quad \left. - \ln \prod_{j=1}^{n_i} t_{ij} - 2 \ln \Gamma(m) + m \ln a \right\}. \end{aligned} \quad (2.5)$$

Then after taking the derivative with respect to  $a$  in (2.5) and setting it equal to zero, we have

$$\sum_{i=1}^k \left[ \frac{n_i + m}{a + n_i \ln(t_{in_i}) - \ln \prod_{j=1}^{n_i} t_{ij}} - \frac{m}{a} \right] = 0. \quad (2.6)$$

Similarly, after taking the derivative with respect to  $m$  in (5) and setting it equal to zero, we have

$$\sum_{i=1}^k \left[ -\ln 2 + \frac{2\Gamma'(n_i + m)}{\Gamma(n_i + m)} - [a + n_i \ln(t_{in_i}) - \ln \prod_{j=1}^{n_i} (t_{ij})] - 2 \frac{\Gamma'(m)}{\Gamma(m)} + \ln a \right] = 0. \quad (2.7)$$

According to the following well known property of Gamma function:

$$-\frac{\Gamma'(z)}{\Gamma(z)} = \frac{1}{z} + \gamma + \sum_{i=1}^{\infty} \left( \frac{1}{n+z} - \frac{1}{n} \right),$$

equation (2.7) can be simplified to

$$2 \sum_{i=1}^k \sum_{j=1}^{n_i} \frac{1}{m+j-1} + k \ln \frac{a}{2} - ka - \sum_{i=1}^k \left[ n_i \ln(t_{in_i}) - \ln \prod_{j=1}^{n_i} (t_{ij}) \right] = 0. \quad (2.8)$$

MLEs of  $a$  and  $m$  can be obtained by solving equations (2.6) and (2.8) numerically. In general, the likelihood equations do not admit a closed form solution and a numerical method must be employed to approximate the MLEs of the hyperparameters  $(a, m)$ .

In the special case that we consider one system, which is  $k = 1$ , with observations  $(t_1, t_2, \dots, t_n)$ , we are able to obtain the estimate of  $a$  in a closed form from Bayesian maximum likelihood approach. The following results shall be considered as Bayesian inference.

$$\hat{a} = \hat{m} \ln t_n - \frac{\hat{m}}{n} \ln \prod_{i=1}^n t_i.$$

Then we can use Newton-Raphson method to obtain MLE for  $m$ . The posterior distribution is

$$\pi(\eta, \beta | \vec{t}, n, \hat{m}, \hat{a}) = \frac{c^{-1} \eta^{n+\hat{m}-1} \beta^{n+\hat{m}-1} [\exp(-\hat{a}) t_n^{\hat{m}} \prod_{i=1}^n t_i]^{\beta-1} \exp(-2\eta t_n^\beta)}{m(\vec{t} | \hat{m}, \hat{a})}. \tag{2.9}$$

From the form of equation (2.9), the prior distribution and posterior distribution are from the same family, therefore, priors are natural conjugate priors. The posterior mean for  $\eta$  has a closed form. The posterior mean is considered as a Bayesian estimation of  $\eta$  which is given by

$$\begin{aligned} \tilde{\eta}_B &= E(\eta) = \int \int \eta \pi(\eta, \beta) d\beta d\eta \\ &= \frac{\Gamma(n + \hat{m} + 1)}{2\Gamma(\hat{m} + n)} \left\{ \frac{\hat{a} + (n + 1) \ln t_n - \ln \prod_{i=1}^n t_i}{\hat{a} + n \ln t_n - \ln \prod_{i=1}^n t_i} \right\}^{-(n+\hat{m})}. \end{aligned}$$

The posterior mean for  $\beta$  also has a closed form and is given by

$$\begin{aligned} \tilde{\beta}_B &= E(\beta) = \int \int \beta \pi(\eta, \beta) d\eta d\beta \\ &= \frac{\Gamma(n + \hat{m} + 1)}{\Gamma(n + \hat{m})(\hat{a} + n \ln t_n - \sum_{i=1}^n \ln t_i)}. \end{aligned}$$

### 3. Parametric Empirical Bayes on the PLP–Three Hyperparameters $(m, a, y_m)$

Now we consider a more general conjugate prior with three hyperparameters  $(m, a, y_m)$ , which adds a hyperparameter  $y_m$ . Recall the likelihood function:

$$L(\vec{t}_i | \eta, \beta) = L(t_{i1}, t_{i2}, \dots, t_{in} | \eta, \beta) = \eta^{n_i} \beta^{n_i} \left( \prod_{j=1}^{n_i} t_{ij} \right)^{\beta-1} \exp(-\eta t_{in_i}^\beta).$$

The natural conjugate prior distribution for the power law failure model is given by

$$\pi_0(\eta, \beta | m, a, y_m) \propto \eta^{m-1} \beta^{m-1} \left( \prod_{i=1}^m y_i \right)^{\beta-1} \exp(-\eta y_m^\beta). \quad (3.1)$$

The parameters  $y_1 \dots y_m$  can be interpreted as a pseudo-data set, where  $m$  is the number of failures and  $y_i$  is the time of the  $i$ th failure. For simplicity and without loss of generality, we can choose  $\prod_{i=1}^m y_i = \exp(-a) y_m^m$ . Thus the natural conjugate prior becomes

$$\pi_0(\eta, \beta | m, a, y_m) = c^{-1} \eta^{m-1} \beta^{m-1} (\exp(-a) y_m^m)^{\beta-1} \exp(-\eta y_m^\beta), \quad (3.2)$$

here  $(m, a, y_m)$  are positive hyperparameters.  $c$  is a constant and

$$\begin{aligned} c &= \int \int \eta^{m-1} \beta^{m-1} (\exp(-a) y_m^m)^{\beta-1} \exp(-\eta y_m^\beta) d\eta d\beta \\ &= \Gamma^2(m) a^{-m} [\exp(-a) y_m^m]^{-1}. \end{aligned}$$

$$\begin{aligned} & m(\vec{t}_i | m, a, y_m) \\ &= \int \int L(t_{i1}, t_{i2}, \dots, t_{in_i} | \eta, \beta) \pi_0(\eta, \beta | m, a, y_m) d\eta d\beta \\ &= \frac{1}{c} \int \Gamma(n_i + m) [t_{in_i}^\beta + y_m^\beta]^{-(m+n_i)} [\exp(-a) y_m^m \prod_{j=1}^{n_i} t_{ij}]^{\beta-1} \beta^{n_i+m-1} d\beta. \end{aligned}$$

Since  $k$  systems are independent, the likelihood distribution of  $T$  given  $(m, a, y_m)$  is

$$m(T | m, a, y_m) = m(\vec{t}_1 | m, a, y_m) m(\vec{t}_2 | m, a, y_m) \cdots m(\vec{t}_k | m, a, y_m).$$

Again the likelihood equations do not admit a closed form solution and a numerical method need to be employed to approximate the maximum likelihood estimates of  $(a, m, y_m)$ . However, suppose we only have observations  $(t_1, t_2, \dots, t_n)$  from one system, which means we only have a random sample of size one  $(\eta, \beta)$  from the prior  $\pi(\eta, \beta | a, m, y_m)$ , our inference shall be regarded as Bayesian maximum likelihood approach. For this special case, we have the posterior distribution of  $(\eta, \beta)$  is

$$\pi(\eta, \beta | \vec{t}, n, \hat{m}, \hat{a}, \hat{y}_m) = \frac{c^{-1} \eta^{n+\hat{m}-1} \beta^{n+\hat{m}-1} [\exp(-\hat{a}) \hat{y}_m^{\hat{m}} \prod_{i=1}^n t_i]^{\beta-1} \exp(-\eta(\hat{y}_m^\beta + t_n^\beta))}{m(\vec{t} | \hat{m}, \hat{a}, \hat{y}_m)}. \quad (3.3)$$

The posterior mean for  $\eta$  is

$$\begin{aligned}
 E(\eta) &= \int \int \eta \pi(\eta, \beta) d\beta d\eta \\
 &= \frac{\Gamma(n + \hat{m} + 1) \int \beta^{n+\hat{m}-1} [\exp(-\hat{a}) \hat{y}_m^{\hat{m}} \prod_{i=1}^n t_n]^{\beta-1} (\hat{y}_m^\beta + t_n^\beta)^{-(n+\hat{m}+1)} d\beta}{\Gamma(n + \hat{m}) \int \beta^{n+\hat{m}-1} [\exp(-\hat{a}) \hat{y}_m^{\hat{m}} \prod_{i=1}^n t_n]^{\beta-1} (\hat{y}_m^\beta + t_n^\beta)^{-(n+\hat{m})} d\beta}.
 \end{aligned}$$

The posterior mean for  $\beta$  is

$$\begin{aligned}
 E(\beta) &= \int \int \beta \pi(\eta, \beta) d\eta d\beta \\
 &= \frac{\int \beta^{n+\hat{m}} [\exp(-\hat{a}) \hat{y}_m^{\hat{m}} \prod_{i=1}^n t_n]^{\beta-1} (\hat{y}_m^\beta + t_n^\beta)^{-(n+\hat{m})} d\beta}{\int \beta^{n+\hat{m}-1} [\exp(-\hat{a}) \hat{y}_m^{\hat{m}} \prod_{i=1}^n t_n]^{\beta-1} (\hat{y}_m^\beta + t_n^\beta)^{-(n+\hat{m})} d\beta}.
 \end{aligned}$$

It should be addressed that the problem with parametric Empirical Bayes (PEB) is that we assume that the estimates of the prior parameters are the prior parameters themselves. The PEB approach does not account for uncertainty in the estimates of these hyperparameters. Variation in these estimates would lead to more variation in the estimates of function of parameters, such as intensity and reliability etc.

#### 4. Prior Plots

The joint prior density is given by (3.2). By taking integral with respect to  $\eta$ ,

$$\pi_0(\beta|m, a, y_m) = \frac{a^m \beta^{m-1} \exp(-a\beta)}{\Gamma(m)}. \tag{4.1}$$

The marginal prior distribution of  $\beta$  has a Gamma distribution with mean  $m/a$  and variance  $m/a^2$ . The conditional prior distribution of  $\eta$  given  $\beta$  is

$$\begin{aligned}
 \pi_0(\eta|\beta) &= \frac{\pi_0(\eta, \beta)}{\pi_0(\beta)} \\
 &= \frac{y_m^{\beta m} \eta^{m-1} \exp(-\eta y_m^\beta)}{\Gamma(m)},
 \end{aligned}$$

which is Gamma distribution with mean  $m/y_m^\beta$  and variance  $m/y_m^{2\beta}$ . Figure 1 and 2 are the prior density plots given different value of hyperparameters.

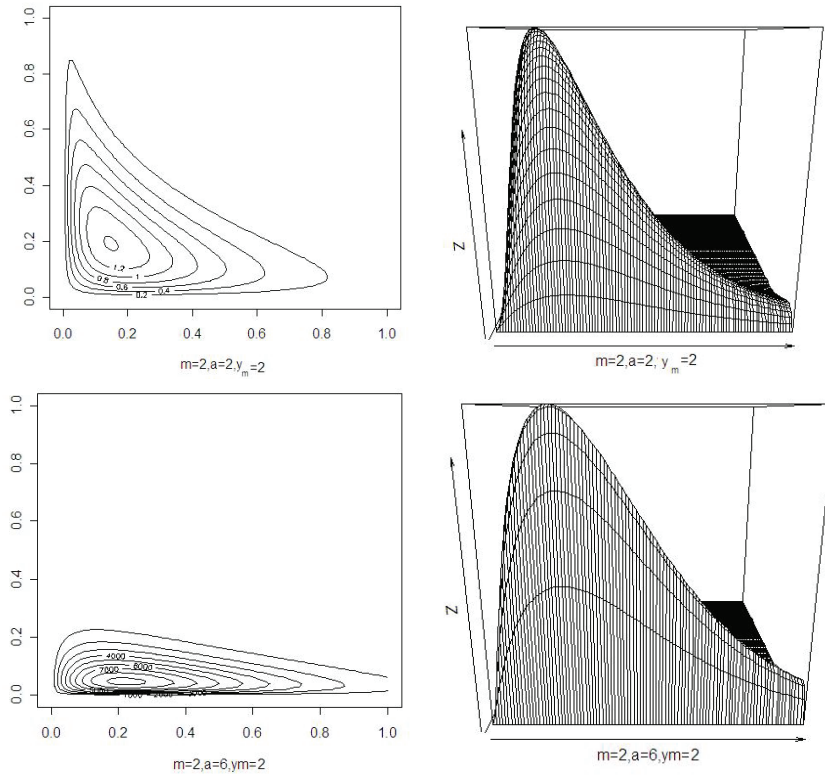


Figure 1: Prior Contour Plots and Three-Dimension Plots - Part I

- In Figure 1, the first two are prior's contour plot and three-dimension plot when  $m = 2$ ,  $a = 2$ ,  $y_m = 2$ . Only we only draw x-axis of  $\eta$  in  $[0,1]$  and y-axis of  $\beta$  in  $[0,1]$ .
- In Figure 1, the bottom two graphs show that if only  $a$  increases from 2 to 6 and the other parameters stay unchanged, whose threshold values are 2, then both the mean  $m/a$  and the variance  $m/a^2$  decrease. The graphs move closer to x-axis and become more concentrated.
- In Figure 2, the first two graphs show that if only  $y_m$  increases from 2 to 6 and the other parameters stay unchanged, whose threshold values are 2, then the mean  $m/y_m^\beta$  decreases and the variance  $m/y_m^{2\beta}$  decreases. The graph move closer to y -axis and become more concentrated.
- In Figure 2, the bottom two graphs show that if only  $m$  increases from 2 to 6 and the other parameters stay unchanged, whose threshold values are 2, then both the mean and the variance increase. The graphs move away from x-axis and y-axis and become more dispersed.



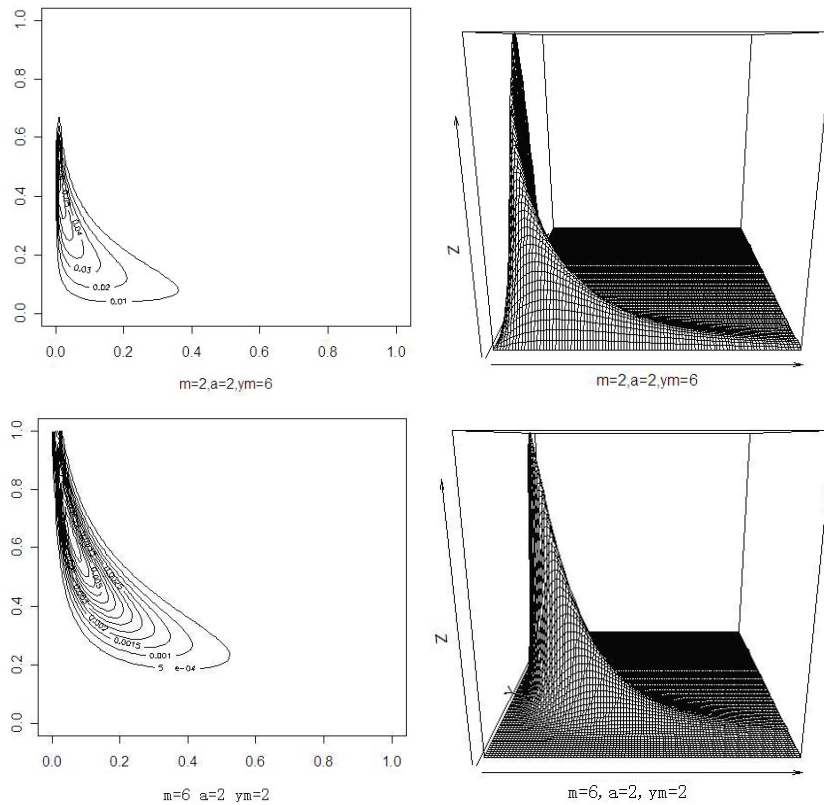


Figure 2: Prior Contour Plots and Three-Dimension Plots - Part II

### 5. Conclusions and Remarks

An empirical Bayes (EB) approach on the Power Law Process with natural conjugate priors has been developed in this paper. We obtained a closed form of marginal distribution of observed data in the two-hyperparameter case. For a special case of single system, we derived Bayesian maximum likelihood estimates of hyperparameters and also found a closed form of the posterior means of model parameters  $(\eta, \beta)$ . Similar empirical Bayes approach based on the three-hyperparameter prior has been shown and the marginal distribution of observed data and estimates of  $(\eta, \beta)$  have been derived in a simpler form.

It is fundamental basis of Bayesian decision theory that statistical inference should start with the determination of three factors: the distribution family for the observations, the prior distribution for the parameters and the loss associated with decisions. Further work can be done to check the robustness of the

priors we used, such as noninformative priors, natural conjugate EB priors and nonparametric Empirical Bayes priors. We can slightly change the prior and see what happens to the decision. Two commonly used measures are the range of posterior decision and comparing Bayes risk criteria.

## References

- Ascher, H. and Feingold, H. (1984). *Repairable Systems Reliability*. Marcel Dekker.
- Bain, L. J. (1978). *Statistical Analysis of Reliability and Life Testing Models*. Marcel Dekker.
- Bain, L. J., and Engelhardt, M. (1991). *Statistical Analysis of Reliability and Life Testing Models* 2nd edition. Marcel Dekker.
- Bar-Lev, S. K., Lavi, I., and Reiser, B. (1992). Bayesian Inference for the Power Law Process. *Ann. Inst. Statist. Math.* **44**, 623-639.
- Calabria, R., Guida, M., and Pulcini, G. (1990). Bayes Estimation of Prediction Intervals for a Power Law Process, *Comm. Statist. Theory Methods* **19**, 3023-3035.
- Calabria, R., Guida, M. and Pulcini, G. (1992). Power Bounds for a Test of Equality of Trends in  $k$  Independent Power Law Processes. *Comm. Statist. Theory Methods* **21**, 3275-3290.
- Calabria, R. and Pulcini, G. (1997). Bayes Inference for the Modulated Power Law Process, *Comm. Statist. Theory Methods* **26**, 2421-2438.
- Crow, L. H. (1974). Reliability Analysis for Complex Repairable Systems. In *Reliability and Biometry* (Edited by F. Proschan and R.J. Serfling), 379-410. SIAM.
- Crow, L. H. (1982). Confidence Interval Procedures for the Weibull Process with Applications to Reliability Growth. *Technometrics* **24**, 67-72.
- Engelhardt, M. and Bain, L. J. (1986). On the Mean Time Between Failures for Repairable Systems. *IEEE Transactions on Reliability* **35**, 419-422.
- Guida, M., Calabria, R. and Pulcini, G. (1989). Bayes Inference for a Non-homogeneous Poisson Process with Power Law Intensity. *IEEE Transactions on Reliability* **38**, 603-609.
- Huang, Y. and Bier, V. (1998). A Natural Conjugate Prior for the Non-Homogeneous Poisson Process with a Power Law Intensity Function, *Commun. Statist.* **27**, 525-551.
- Kyparisis, J. and Singpurwalla, N. D., (1985). Bayesian inference for the Weibull process with applications to assessing software reliability growth and predicting software failures. In *Computer Science and Statistics* (Edited by Billard, L.), 57-64. Elsevier Applied Science.
- Lee, L. and Lee, K. (1978). Some Results on Inference for the Weibull Process, *Technometrics* **20**, 41-45.

- 
- Lindqvist, B. (1997). Statistical Modeling and Analysis of Repairable Systems, *Mathematical Methods in Reliability* **21**, 16-19.
- Singpurwalla, N. and Wilson, S. (1999). *Statistical Methods in Software Engineering*. Springer.
- Sinha, S. K. , John Wiley and Sons. (1986). *Reliability and Life Testing*. Wiley.
- Rigdon, S. E. and Basu, A. P. (1990). Estimating the Intensity Function of a Power Law Process at the Current Time: Time Truncated Case. *Communications in Statistics Simulation and Computation* **19**, 1079-1104.

Received January 9, 2008; accepted March 30, 2008.

Zhao Chen  
Department of Mathematics  
Florida Gulf Coast University,  
Fort Myers, Florida, 33965, USA  
kchen@fgcu.edu