A simple procedure for Bayesian estimation of the Weibull distribution.

2 authors:

Mark Kaminskiy
National Aeronautics and Space Administration
42 PUBLICATIONS **1,324** CITATIONS

Vasiliy V Krivtsov
The Ford Motor Company
51 PUBLICATIONS **1,388** CITATIONS

A Simple Procedure for
Bayesian Estimation of the Weibull Distribution

M.P. Kaminskiy, and V.V. Krivtsov, *Member, IEEE*

Abstract – Practical use of Bayesian estimation procedures is often associated with difficulties related to elicitation of prior information, and its formalization into the respective prior distribution. The two-parameter Weibull distribution is a particularly difficult case, because it requires a two-dimensional joint prior distribution of the Weibull parameters. The novelty of the procedure suggested here is that the prior information can be presented in the form of the interval assessment of the reliability function (as opposed to that on the Weibull parameters), which is generally easier to obtain. Based on this prior information, the procedure allows constructing the continuous joint prior distribution of Weibull parameters as well as the posterior estimates of the mean & standard deviation of the estimated reliability function (or the CDF) at any given value of the exposure variable. A numeric example is discussed as an illustration. We additionally elaborate on a new parametric form of the prior distribution for the scale parameter of the exponential distribution. This distribution is not a Gamma (as might intuitively be expected); its mode is available in a closed form, and the mean is obtainable through a series approximation.

Index Terms – Weibull distribution, Bayesian estimation, conjugate prior.

ACRONYMS[1]

CDF   cumulative distribution function

HPD   highest posterior density

PDF   probability density function

RRA   reliability & risk assessment

M. Kaminskiy is with University of Maryland at College Park, USA.
V. Krivtsov is with Ford Motor Company in Dearborn, USA.
[1] The singular and plural of an acronym are always spelled the same.

N.B. The procedure proposed in this paper was implemented by SAS in V.12 release of their SAS-JMP® software.

t	exposure variable (e.g., time)
$F(t)$	CDF
$\hat{F}(t_k)$	estimate of the CDF at fixed exposure t_k
$\hat{\sigma}\{\hat{F}(t_k)\}$	estimate of the standard deviation of $\hat{F}(t_k)$
p	Beta distributed random variable
$f(p)$	probability density function of p
$x_0(t_k)$, $n_0(t_k)$	parameters of the Beta distribution, representing uncertainty of the estimated CDF at fixed exposure t_k
$\Gamma(.)$	gamma function
λ	scale parameter of the exponential distribution
$f(\lambda)$	probability density function of λ

I. INTRODUCTION

Weibull is one of the most widely used distributions in reliability & risk assessment. Practical use of Bayes' estimation is often associated with difficulties related to elicitation of prior information, and its formalization into the respective prior distribution. Opposite to the binomial & exponential distributions, which are also popular in RRA, the two-parameter Weibull distribution requires a two-dimensional joint prior distribution. It should be noted that in the framework of the Bayesian approach, both the exponential & binomial distributions have their respective conjugate prior distributions, which makes their practical use more convenient.

The situation with the Weibull distribution is different. In the most realistic case, when both parameters of the distribution are considered as random variables, the fundamental result obtained by Soland [6] states that the Weibull distribution does not have a conjugate *continuous* joint prior distribution.

Later, Soland [7] proved that a conjugate *continuous-discrete* joint prior distribution exists for the Weibull distribution parameters. The continuous "component" of this distribution is related to the scale parameter of the Weibull distribution (similar to Bayesian estimation of the exponential distribution, for which gamma is used as the conjugate prior distribution), and the discrete one is related to the shape parameter. Although of a great academic interest, this result

may be challenging to apply to real life problems due to the difficulties of evaluating the prior information needed (mostly related to the shape parameter). A reliability application example of this approach can be found in [4].

In the sequel, we consider a new procedure for the Bayesian estimation of the Weibull distribution which allows constructing *continuous* joint prior distribution of Weibull parameters as well as the posterior estimates of the mean & standard deviation of the estimated CDF (or the reliability function) at any given value of the exposure variable.

2. PROCEDURE

Consider the prior information, which is available at two values (t_1 & t_2, $t_2 > t_1$) of the exposure variable as the estimates of CDF, $\hat{F(t_k)}$, and its standard deviation, $\hat{\sigma}\{\hat{F(t_k)}\}$; see Figure 1. This prior information can be divided in the following two cases. In the first case, it comes in the traditional form as life data sufficient for the Kaplan-Meier & Greenwood estimators to be used. It might be attributed to the empirical Bayesian inference. In the second case, the prior information is obtained using the expert elicitation methods [1].
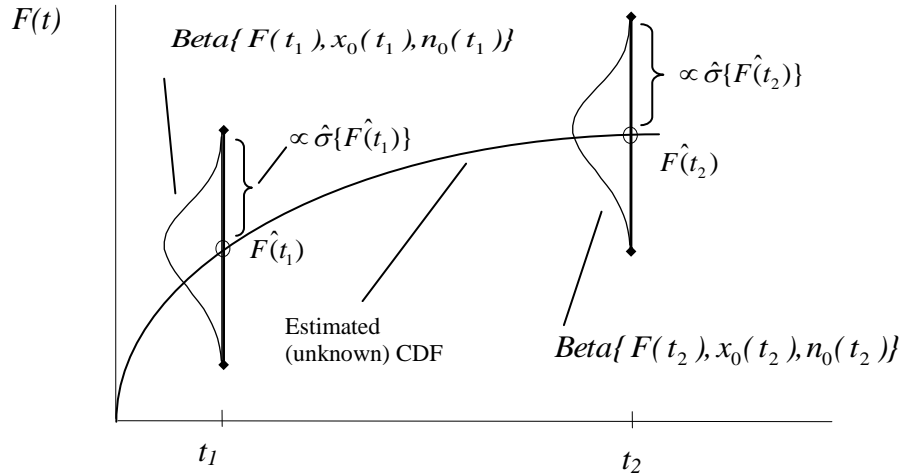

Figure 1: Prior information, and Beta distributions
approximating uncertainty of the estimated CDF at a fixed exposure, t_k.

Based on these prior data, we are interested in obtaining the following:

- joint *prior* distribution of the Weibull distribution parameters,
- prior point estimates of the Weibull parameters,
- *prior* distribution of the estimated CDF at any fixed exposure t_k,

- joint *posterior* distribution of the Weibull distribution parameters (based on data observed at t_3),

- posterior point estimates of the Weibull parameters as the location of the *highest posterior probability density* (mode) of their joint distribution, similar to the HPD estimates considered in [2], and

- *posterior* distribution of the estimated CDF at exposure t_3 (based on data observed at t_3).

A. Prior Distribution

The Beta prior distribution with the following probability density function is assumed to characterize uncertainty of the estimated CDF at a fixed exposure:

$$f(p) = \frac{\Gamma(n_0)}{\Gamma(x_0)\Gamma(n_0 - x_0)} p^{x_0 - 1}(1 - p)^{n_0 - x_0 - 1}, \quad n_0 > 0, n_0 - x_0 > 0 \tag{1}$$

where p is the random variable representing the estimated CDF at a fixed exposure, $F(t_k)$; n_0, x_0 are the parameters of the Beta distribution, both positive quantities; and $\Gamma(.)$ is the gamma function.

Based on the available estimates of the *prior* CDF at a fixed exposure ($t=const$), $\hat{F}(t)$, and its standard deviation, $\hat{\sigma}\{\hat{F}(t)\}$, the parameters of the respective *prior* beta-distribution can be obtained through the method of moments [5] as

$$x_0(t) = \frac{\hat{F}(t)^2[1 - \hat{F}(t)]}{\hat{\sigma}^2\{\hat{F}(t)\}} - \hat{F}(t), \quad n_0(t) = \frac{x_0(t)}{\hat{F}(t)} \tag{2}$$

By random sampling from the estimated Beta distributions, one can obtain pairs of random realizations of the estimated CDF at exposures t_1 & t_2: $\{F_i(t_1), F_i(t_2)\}$, $i = 1,2, ..., n$. For any such pair, as long as $F(t_1) < F(t_2)$, the *shape* & the *scale* parameters of the respective Weibull CDF can be obtained as

$$\beta = \frac{F^*(t_2) - F^*(t_1)}{Ln(t_2/t_1)}$$
$$\alpha = Exp\left(Ln(t_1) - \frac{F^*(t_1)}{\beta}\right) \tag{3}$$

Here, $F^*(t_k) = Ln(Ln(1-F(t_k))^{-1})$, $k = 1, 2$.

It is obvious that n simulated pairs of $\{F_i(t_1), F_i(t_2)\}$ would define n pairs of Weibull distribution parameters $\{\alpha_i, \beta_i\}$, which are used to construct their joint prior distribution.

Further, the obtained realizations of the Weibull CDF, $F_i(t; \alpha_i, \beta_i)$, can be extrapolated to a given exposure, say, t_3, to obtain the distribution of the CDF at t_3. Using Equation (2), this distribution can again be approximated by the Beta-*prior* distribution with parameters $x_0(t_3)$ & $n_0(t_3)$.

B. Posterior Distributions

Posterior distribution of Weibull CDF at a fixed exposure

The estimated Beta-prior at t_3 is now available for Bayesian estimation. Assume that observation data are available for exposure t_3 in the binary form of r failures out of N trials. Then, the *posterior* distribution is also the Beta distribution with parameters

$$x_0(t_3)^* = x_0(t_3) + r,$$
$$n_0(t_3)^* = n_0(t_3) + N \qquad\qquad (4)$$

Posterior distribution of Weibull distribution parameters

The joint posterior distribution of the Weibull parameters can be numerically obtained by sampling random realizations from the estimated Beta distributions at exposures t_1, t_2, & (the newly obtained) t_3; and then using the probability paper method to obtain the sample of estimates of the Weibull distribution parameters.

3. NUMERIC EXAMPLE

Consider expert estimates of the failure probability of a mechanical component at $t_1=1$, and $t_2=3$ years of operation as $\hat{F}(t_1) = 0.01$, and $\hat{F}(t_2) = 0.05$, with 10% error. Treating the error as the coefficient of variation, the estimates of the standard deviation of the failure probability at the two exposures are $\hat{\sigma}\{\hat{F}(t_1)\} = 0.001$, and $\hat{\sigma}\{\hat{F}(t_2)\} = 0.005$, respectively. Using Equation (2), the estimates of the Beta distribution parameters at the two exposures are $\{x_0(t_1) = 99, n_0(t_1) = 9899\}$, and $\{x_0(t_2) = 95, n_0(t_2) = 1899\}$, respectively. By sampling random realizations ($n = 10000$) of the CDF from the two estimated Beta distributions, and the subsequent use of Equation (3), one

obtains the joint prior distribution of Weibull parameters shown in Figure 2. The prior estimates of the Weibull parameters corresponding to the mode of their joint distribution are found as $\{\beta_{pr} = 1.51, \alpha_{pr} = 22.95\}$. The failure probability, and its standard deviation at a future exposure of t_3=5 years of operation, are estimated as 0.105, and 0.016, respectively; which using Equation (2), can be translated into respective Beta distribution parameters at t_3: $\{x_0(t_3) = 40, n_0(t_3) = 386\}$.
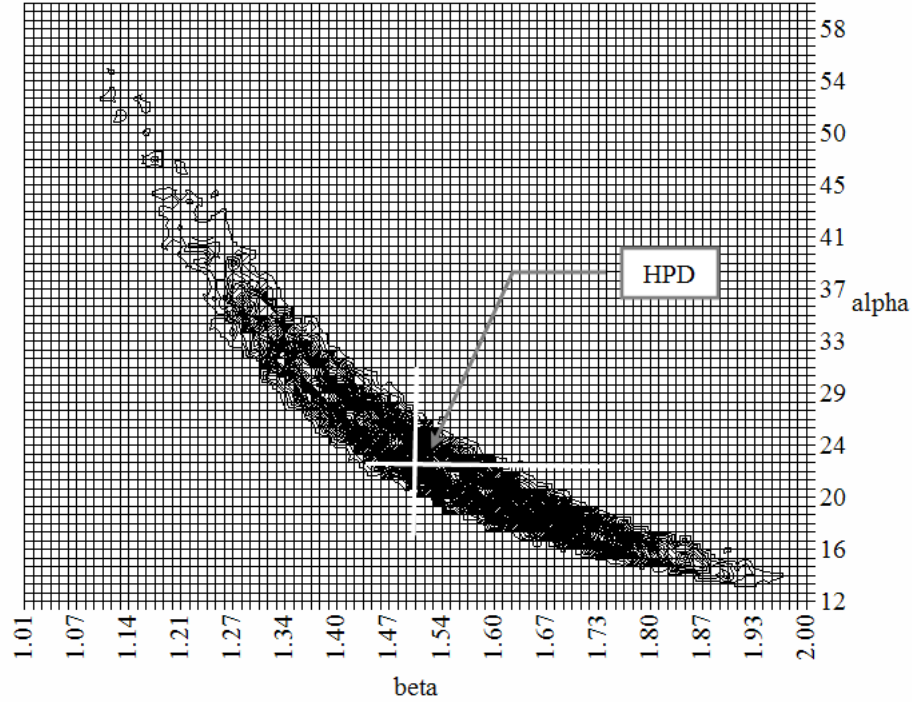

Figure 2: Contours of Joint Prior Distribution of the Weibull Parameters.

Further, consider the data observed at t_3, which are 1 failure out of 12 components observed. The parameters of the posterior Beta distribution parameters at t_3, according to Equation (4), are found as $\{x_0(t_3)^* = 41, n_0(t_3)^* = 398\}$. Figure 3 shows the joint posterior distribution of the Weibull distribution parameters obtained by sampling random realizations ($n = 10000$) from the estimated Beta distributions at exposures t_2 & t_3, and the subsequent use of Equation (3). The posterior HPD estimates of Weibull parameters are found as $\{\beta_{post} = 1.70, \alpha_{post} = 16.55\}$. The posterior estimates of the failure probability (CDF), and its standard deviation at t_3 are obtained as 0.104, and 0.069 respectively. The estimation results are summarized in Table I.
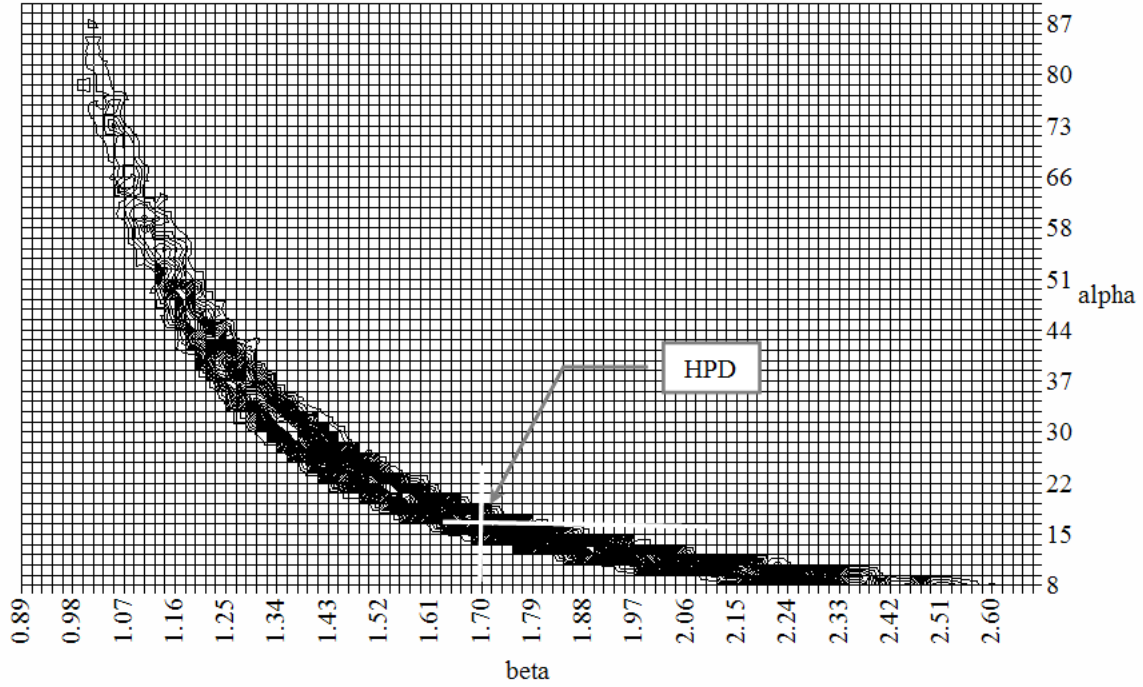
Figure 3: Contours of Joint Posterior Distribution of the Weibull Parameters.

Table I: Prior and Posterior Estimates.

	Prior			Posterior
	$t_1 = 1$ (given)	$t_2 = 3$ (given)	$t_3 = 5$ (inferred)	$t_3 = 5$ (inferred)
$F(t)$	0.010	0.050	0.105	0.104
$\sigma\{F(t)\}$	0.001	0.005	0.016	0.069
α	22.95			16.55
β	1.51			1.70

We additionally investigated the sensitivity of the posterior distribution to the initial data on the estimated CDF. Using the data in the above example, we increased the CDF error at t_1, and t_2 to 50%, and 20%, respectively. As could be expected, the contours of the prior & the posterior
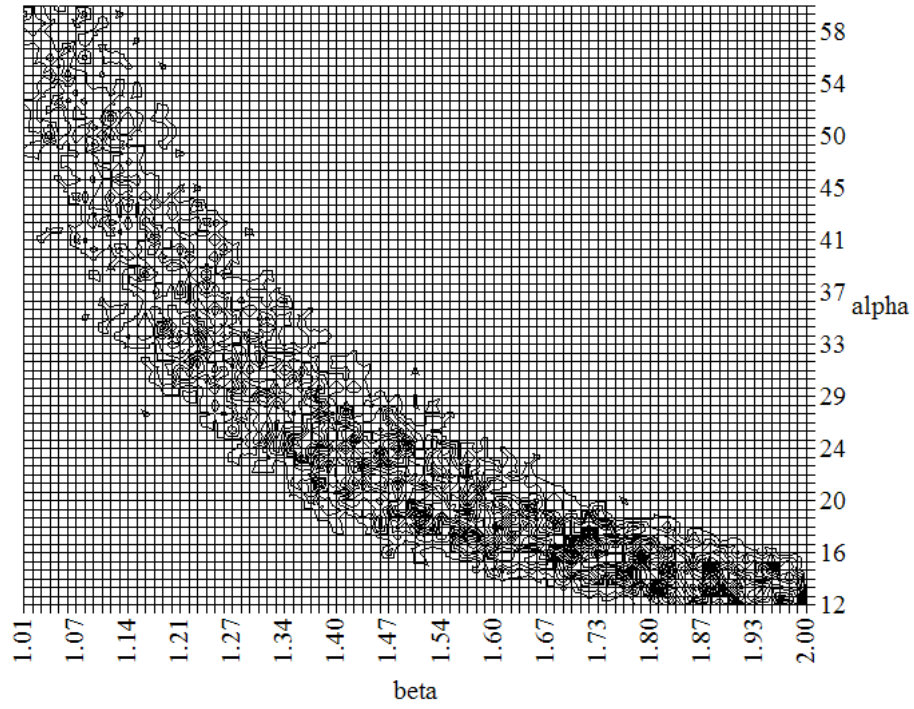
7

Figure 4: Contours of Joint Prior Distribution of the Weibull Parameters with Increased CoV of Prior Data.
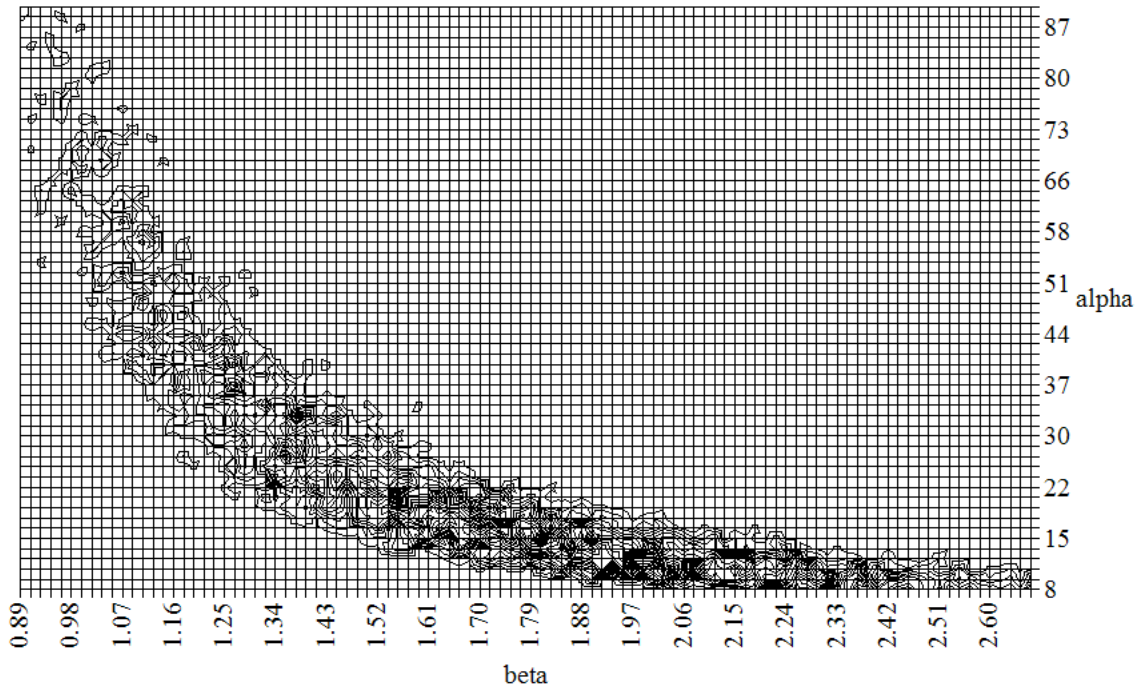

Figure 5: Contours of Joint Posterior Distribution of the Weibull Parameters with Increased CoV of Prior Data.

joint distribution of Weibull parameters became much wider; see Figures 3 & 4. For the convenience of comparative analysis, we maintained the same scales in Figures 4 & 5 as in Figures 2 & 3.

4. PARTICULAR CASE

Generally speaking, the suggested procedure for evaluating prior information does not result in a prior distribution available in a closed form. Nevertheless, for the particular case of the Weibull distribution with the shape parameter equal to one (the exponential distribution), the closed form *can* be obtained. This closed form of the prior distribution can be useful from a methodological standpoint showing that it is *not* the Gamma distribution, which might intuitively be anticipated in this case.

In this case, the prior information can be reduced to the estimates of the CDF, and its standard deviation at just *one* value of the nonrandom exposure, t, so that the random variable of interest, λ, can be represented as

$$\lambda = \frac{-\ln(1-F(t))}{t}, \tag{5}$$

where $F(t=const)$ is distributed according to the Beta distribution with PDF given in Equation (1).

It can be shown that the probability density function of Equation (5) is given by

$$f(\lambda) = t\frac{\Gamma(n_0)}{\Gamma(x_0)\Gamma(n_0-x_0)}[z^{n_0-x_0}(1-z)^{x_0-1}], \tag{6}$$

where $z = e^{-t\lambda}$, and $\lambda > 0$. Figure 6 shows the PDF of the introduced distribution for various values of its parameters.

The mode of Equation (5), the HPD, happens to be available in a closed form as

$$\lambda = \frac{1}{t}Ln\left(\frac{n_0-1}{n_0-x_0}\right), \quad n_0 > 1, \ n_0 - x_0 > 0.$$
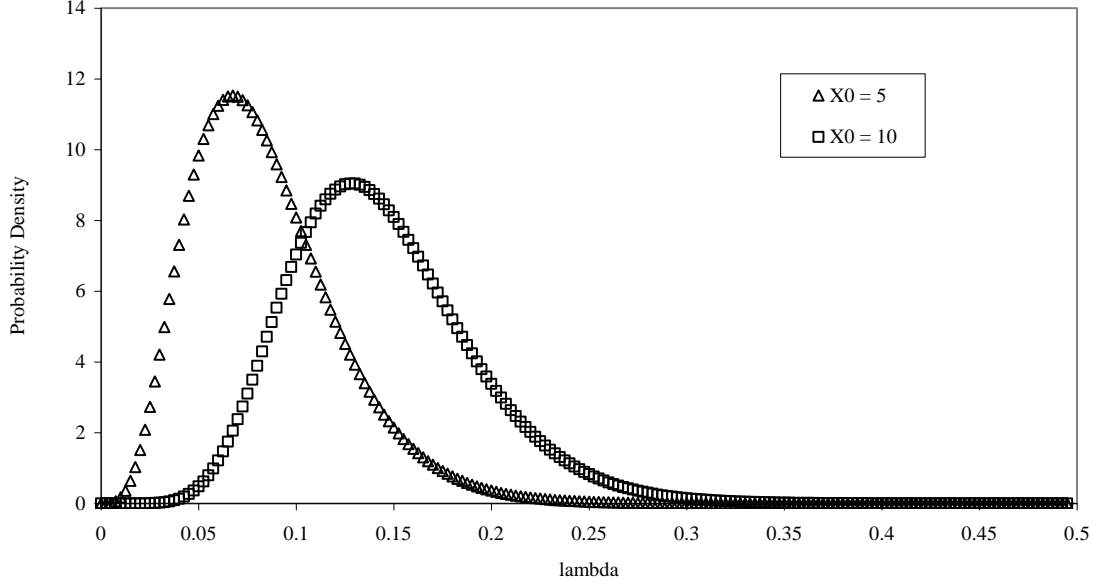
9

Figure 6: Probability Density Function of Equation (4)

with t=5, n_0 - x_0 = 10, and x_0 of 5, and 10.

The mean of Equation (5) can be obtained as

$$E(\lambda) = t \frac{\Gamma(n_0)}{\Gamma(x_0)\Gamma(n_0 - x_0)} \int_0^\infty \lambda [e^{-(n_0-x_0)\lambda t}(1 - e^{-\lambda t})^{x_0-1}] d\lambda,$$

Using the same variable $z = e^{-t\lambda}$, the above integral can be written in the form

$$E(\lambda) = \frac{1}{t}\left(\frac{\Gamma(\delta+\rho)}{\Gamma(\delta)\Gamma(\rho)}\right)\int_0^1 \ln(1/z)[z^{n_0-x_0-1}(1-z)^{x_0-1}]dz, \tag{7}$$

From Tables of Integrals [3], one finds

$$\int_0^1 \left(\ln\frac{1}{z}\right)^h z^{q-1}(1-z^b)^{\frac{c}{b}-1} dz = h!\left\{\frac{1}{q^{h+1}} + \sum_{k=1}^\infty \frac{(kb-c)[(k-1)b-c]...(b-c)}{k!b^k(q+bk)^{h+1}}\right\}, \tag{8}$$

where $q > 0$, $h + c/b > 0$.

Applying Equation (8) to the integral in Equation (7), one finally gets

$$\int_0^1 \ln(1/z)[z^{n_0-x_0-1}(1-z)^{x_0-1}]dz = \left\{\frac{1}{(n_0-x_0)^2} + \sum_{k=1}^\infty \frac{(k-x_0)[(k-1)-x_0]...(1-x_0)}{k!(n_0-x_0+k)^2}\right\}.$$

5. CONCLUSIONS

We considered a new procedure for the Bayesian estimation of the Weibull distribution. The novelty of the suggested procedure is that the prior information can be presented in the form of the interval assessment of *the reliability function* (as opposed to that on the Weibull *parameters*), which is generally easier to obtain. Based on this prior information, the procedure allows construction of the continuous joint prior distribution of Weibull parameters as well as the posterior estimates of the mean & standard deviation of the estimated reliability function (or the CDF) at any given value of the exposure variable. We also introduced a *new parametric form* of the prior distribution for the scale parameter of the exponential distribution. This distribution is not a Gamma (as might intuitively be expected); its mode is available in a closed form, and the mean is obtainable through a series approximation.

ACKNOWLEDGEMENT

REFERENCES

[1]   B. M. Ayyub, *Elicitation of Expert Opinions for Uncertainty and Risks*, CRC Press, 2002.

[2]   G. E. P. Box, and C. Tiao, *Bayesian Inference in Statistical Analysis,* Addison-Wesley Publishing Company, 1973.

[3]   I.S. Gradshteyn, and I.M. Ryzhik, *Table of integrals, series, and products* / Alan Jeffrey, editor. Translated from the Russian by Scripta Technica, Inc., Academic Press, 2000.

[4]   H. Martz, and R. Waller, *Bayesian Reliability Analysis*, John Wiley & Sons, 1982.

[5]   M. Modaress, M. Kaminksiy, and V. Krivtsov, *Reliability Engineering and Risk Analysis: A Practical Guide* - Marcel Dekker, Inc., 1999.

[6]   R. Soland, "Use of the Weibull Distribution in Bayesian Decision Theory", Report No. RAC-TP-225, Research Analysis Corporation, McLean, VA, 1966.

[7]   R. Soland, "Bayesian Analysis of the Weibull Process with Unknown Scale and Shape Parameters", *IEEE Transactions on Reliability*, vol. R-18, pp. 181-184, 1969.

Mark Kaminskiy is the chief statistician at the Center of Technology and Systems Management of the University of Maryland (College Park), USA. Dr. Kaminskiy is a researcher & consultant in reliability engineering, life data analysis, and risk analysis of engineering systems. He has conducted numerous research & consulting projects funded by the government & industrial companies, such as Department of Transportation, Coast Guards, Army Corps of Engineers, Navy, Nuclear Regulatory Commission, American Society of Mechanical Engineers, Ford Motor Company, Qualcomm Inc, and several other engineering companies. He has taught several graduate courses on Reliability Engineering at the University of Maryland. Dr. Kaminskiy is the author or co-author of over 50 publications in journals, conference proceedings, and reports.

Vasiliy Krivtsov is a Staff Technical Specialist in reliability & statistical analysis with Ford Motor Co. He holds M.S., and Ph.D. degrees in Electrical Engineering from Kharkov Polytechnic Institute, Ukraine; and a Ph.D. in Reliability Engineering from the University of Maryland, USA. Dr. Krivtsov is the author or co-author of over 40 professional publications, including a book on *Reliability Engineering and Risk Analysis*, 9 patented inventions, and 2 Ford trade secret inventions. He is an editor of the Elsevier's *Reliability Engineering and System Safety* journal, and is a member of the IEEE Reliability Society. Prior to Ford, Krivtsov held the position of Associate Professor of Electrical Engineering in Ukraine, and that of a Research Affiliate at the University of Maryland Center for Reliability Engineering. Further information on Dr. Krivtsov's professional activity is available at www.krivtsov.net