

# Estimating Parameters of the Force of Mortality in Actuarial Studies

Malwane M. Ananda, Rohan J. Dalpatadu, Ashok K. Singh

Department of Mathematical Sciences

University of Nevada, Las Vegas

NV 89154.

## Abstract

The two-parameter Weibull model and two-parameter Gompertz model are commonly used as survival time distributions in actuarial studies. The estimation of the parameters of these models is numerically involved. We consider the estimation problem in a Bayesian framework and give the Bayesian estimators of parameters in terms of single numerical integrations. We propose an adaptive Bayesian estimation procedure by putting a prior only on one parameter and finding the other parameter by minimizing the distance between empirical and parametric cumulative distribution functions. This easily computable adaptive Bayesian procedure is compatible with the exact Bayesian procedure. In particular for large samples numerical integration for computing the exact Bayesian procedure is difficult. In both cases, noninformative generalized Bayes estimators and relevant adaptive noninformative Bayes estimators are also given. Some examples will be provided.

*Key words: Force of mortality, two parameter Weibull distribution, Two parameter Gompertz distribution; Bayesian estimation; Generalized Bayes estimators; Anderson and Darling  $A^2$  statistic*

## 1. INTRODUCTION AND SUMMARY

The two-parameter Weibull distribution and the two-parameter Gompertz distribution are widely used distributions in actuarial science and life-testing. We consider the problem of estimating parameters of the two parameter Weibull distribution and two parameter Gompertz distribution. Suppose  $X$  is the life-time random variable, then the force of mortality  $\mu_x$  for the two-parameter Weibull distribution is

$$\mu_x = kx^m, x > 0 \tag{1}$$

with the probability density  $f_X(x)$ ,

$$f_X(x) = kx^m \exp\left\{-\frac{k}{m+1}x^{m+1}\right\}, x > 0 \tag{2}$$

Here both  $k$  and  $m$  are unknown parameters and  $m > 0$  and  $k > 0$ . For the Gompertz distribution, force of mortality  $\mu_x$ ,

$$\mu_x = bc^x, x > 0 \tag{3}$$

with the density  $f_X(x)$ ,

$$f_X(x) = bc^x \exp\left\{-\frac{b}{\ln c}(1 - c^x)\right\}, x > 0 \tag{4}$$

where both  $b$  and  $c$  are unknown parameters and  $b > 0$  and  $c > 1$ . Based on a random complete sample  $x_1, x_2, \dots, x_n$  we would like to estimate these parameters. There are lot of estimation methods available in the literature for these parameters. For instance, maximum likelihood estimators, percentile estimators (estimators based on two selected percentiles; for details

see London (1988)), minimum chi-squared estimators, minimum modified chi-squared estimators, least square estimators, (for details see Johnson and Johnson (1980) or London (1988)) etc. But none of these estimators has closed forms and estimation of these parameters involve numerical computations. For example, maximum likelihood estimators of weibull parameters can be found by direct maximization of the likelihood function

$$L = k^n \left( \prod_{i=1}^n x_i^m \right) \exp \left\{ -\frac{k}{m+1} \left( \sum_{i=1}^m x_i^{m+1} \right) \right\}$$

or alternatively finding the solution for the following two nonlinear equations

$$n(m+1) - k \left( \sum x_i^{m+1} \right) = 0$$

$$\sum_{i=1}^n \ln x_i + \frac{k}{(m+1)^2} \left( \sum x_i^{m+1} \right) - \frac{k}{m+1} \left( \sum_{i=1}^n (m+1) x_i^m \right) = 0$$

and percentile estimators for gompertz model can be evaluated by solving nonlinear equations

$$b(1 - c^{x_{p_1}}) = \ln c \ln(1 - p_1)$$

$$b(1 - c^{x_{p_2}}) = \ln c \ln(1 - p_2)$$

where  $x_{p_1}$  is the  $p_1^{\text{th}}$  percentile and  $x_{p_2}$  is the  $p_2^{\text{th}}$  percentile.

We consider the problem from the Bayesian point of view using two independent priors; a gamma prior on one parameter and any other prior on the other parameter. The Bayesian estimators for these parameters are given as a ratio of one-dimensional integrals. For small samples, one can evaluate these integrals numerically. But for large samples, evaluating these integrals or getting an approximation to these integrals are difficult.

We propose adaptive bayesian estimators for the above problem by putting a prior only on one parameter and finding the other unknown parameter by

minimizing the distance between empirical and parametric cumulative distribution functions. This procedure for the Gompertz distribution is originally proposed by Ananda, Dalpatadu and Singh (1992). The proposed procedure works well for large  $n$  as well as for small  $n$ . Furthermore, we derive the estimators corresponding to noninformative priors. Couple of examples are given.

## 2. ESTIMATION OF WEIBULL PARAMETERS

Consider the gamma prior distribution (with parameters  $\alpha > 0$  and  $\beta > 0$ ) on the parameter  $k$  and any other prior  $g(m)$  on the parameter  $m$ . Assuming that the parameters  $k$  and  $m$  are independent, the joint prior distribution of  $k$  and  $m$  is

$$\pi(k, m) = \frac{e^{-k/\beta} k^{\alpha-1}}{\Gamma(\alpha)\beta^\alpha} g(m) \quad (5)$$

and one can show that the posterior distribution of  $k$  and  $m$  given the data  $\pi(k, m/\text{data})$  is proportional to

$$\pi(k, m/\text{data}) \propto k^{n+\alpha-1} \left( \prod_{i=1}^n x_i^m \right) \exp \left\{ -k \left[ \frac{1}{\beta} + \frac{1}{m+1} (\sum x_i^{m+1}) \right] \right\} g(m). \quad (6)$$

Easy calculation shows that the Bayesian estimators  $\hat{k}$  and  $\hat{m}$  (assuming the quadratic loss) of the parameters  $k$  and  $m$  are

$$\hat{k} = \frac{(n + \alpha) \int_0^\infty (\prod x_i^m) g(m) \left[ \frac{1}{\beta} + \frac{1}{m+1} (\sum x_i^{m+1}) \right]^{-(n+\alpha+1)} dm}{\int_0^\infty (\prod x_i^m) g(m) \left[ \frac{1}{\beta} + \frac{1}{m+1} \sum x_i^{m+1} \right]^{-(n+\alpha)} dm} \quad (7)$$

and

$$\hat{m} = \frac{\int_0^\infty m (\prod x_i^m) g(m) \left[ \frac{1}{\beta} + \frac{1}{m+1} (\sum x_i^{m+1}) \right]^{-(n+\alpha)} dm}{\int_0^\infty (\prod x_i^m) g(m) \left[ \frac{1}{\beta} + \frac{1}{m+1} (\sum x_i^{m+1}) \right]^{-(n+\alpha)} dm}. \quad (8)$$

If one use the noninformative prior

$$\pi(k, m) = k^{\alpha-1} dk dm; k > 0, m > 0, \quad (9)$$

the generalized Bayes estimators of these parameters are

$$\hat{k} = \frac{(n + \alpha) \int_0^{\infty} (\prod x_i^m) g(m) \left[ (\sum x_i^{m+1}) / (m + 1) \right]^{-(n+\alpha+1)} dm}{\int_0^{\infty} (\prod x_i^m) g(m) \left[ (\sum x_i^{m+1}) / (m + 1) \right]^{-(n+\alpha)} dm} \quad (10)$$

and

$$\hat{m} = \frac{\int_0^{\infty} m (\prod x_i^m) g(m) \left[ (\sum x_i^{m+1}) / (m + 1) \right]^{-(n+\alpha)} dm}{\int_0^{\infty} (\prod x_i^m) g(m) \left[ (\sum x_i^{m+1}) / (m + 1) \right]^{-(n+\alpha)} dm}. \quad (11)$$

These integrals must be evaluated numerically. For small  $n$ , this can be accomplished by using any integration program such as QDAGS in IMSL. For large  $n$ , calculating and getting accurate answers for these integrals are difficult. We look at the problem in a slightly different approach, which is proposed by Ananda, Dalpatadu and Singh (1992) for the Gompertz distribution.

First let us assume that the value of the parameter  $m$  is known. Also we assume that the prior information about the parameter  $k$  can be expressed using the gamma prior with parameters  $\alpha$  and  $\beta$ ,

$$\pi(k) = e^{-k/\beta} k^{\alpha-1} / \{\Gamma(\alpha)\beta^\alpha\}. \quad (12)$$

Easy calculation shows that the posterior distribution of  $k$  given the data,  $\pi(k/\text{data})$  is proportional to

$$\pi(k/\text{data}) \propto k^{n+\alpha-1} \exp \left\{ -k \left[ 1/\beta + (\sum x_i^{m+1}) / (m + 1) \right] \right\} \quad (13)$$

and the Bayesian estimator of  $k$ ,  $\hat{k}$  is

$$\hat{k} = \frac{(n + \alpha)}{[1/\beta + (\sum x_i^{m+1}) / (m + 1)]}. \quad (14)$$

Now we find the optimum  $m$  value by minimizing the “distance” between empirical cdf ( $F_{\text{emp}}(x)$ ) and the parametric cdf  $F(x)$ . Here

$$F_{\text{emp}}(x) = (\#\text{of obs.} \leq x)/n$$

and

$$F(x) = 1 - \exp\left\{-\frac{k}{m+1}x^{m+1}\right\}.$$

We use following distance criteria (used in Ananda, Dalpatadu and Singh (1992)) in our analysis.

(a). Area ( $A_1$ ) between  $F_{\text{emp}}(x)$  and  $\hat{F}(x)$ :

$$A_1 = \int_0^{\infty} |F_{\text{emp}}(x) - \hat{F}(x)| dx \quad (15)$$

(b). Anderson & Darling (1954)  $A^2$  Statistic:

$$A^2 = \int_0^{\infty} \frac{n}{(1 - F(x))F(x)} [F_{\text{emp}}(x) - F(x)]^2 f(x) dx$$

This is the expected squared deviation of  $F_{\text{emp}}(x)$  and  $F(x)$ , with the deviations are weighted by the inverse variance of  $F_{\text{emp}}(x)$ . By replacing  $F(x)$  and  $f(x)$  by their estimated values  $\hat{F}(x)$  and  $\hat{f}(x)$ , one gets the  $A^2$  statistic as

$$A^2 = \int_0^{\infty} \frac{n}{(1 - \hat{F}(x))\hat{F}(x)} [F_{\text{emp}}(x) - \hat{F}(x)]^2 \hat{f}(x) dx$$

which can be written as (see, London (1988))

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i - 1) \{ \ln[\hat{F}(x_i) \cdot (1 - \hat{F}(x_{n-i+1}))] \}. \quad (16)$$

Studies by Stephens (1974) show that, in many important situations  $A^2$  statistic provide a better measure for the departure between  $\hat{F}(x)$  and  $F_{\text{emp}}(x)$  than the Kolmogorove-Smirnov  $D$  Statistic which is defined as

$$D = \sup_x |F_{\text{emp}}(x) - \hat{F}(x)|.$$

If no prior information is available about the parameter  $k$  one can use the noninformative generalized prior which has the density

$$\pi(k) = k^{\alpha-1} \quad 0 < b < \infty. \quad (17)$$

The posterior density of  $k$  with respect to the prior given in (17) is

$$\pi(k/\text{data}) \propto k^{n+\alpha-1} \left( \prod_{i=1}^n x_i^m \right) \exp \left\{ -\frac{k}{m+1} (\sum x_i^{m+1}) \right\}$$

and the Generalized Bayes estimator of  $k$  is

$$\hat{k} = \frac{(n+\alpha)(m+1)}{(\sum x_i^{m+1})}$$

Again one can find  $m$  by minimizing the area  $A_1$  given in (15) or Anderson & Darling  $A^2$  statistic given in (16).

### 3. ESTIMATION OF GOMPERTZ PARAMETERS

Consider the gamma prior distribution (with parameters  $\alpha > 0$  and  $\beta > 0$ ) on the parameter  $b$  and any other prior  $g(c)$  on the parameter  $c$ . Assuming that the parameters  $b$  and  $c$  are independent, the joint prior distribution of  $b$  and  $c$  is

$$\pi(b, c) = \frac{e^{-b/\beta} b^{\alpha-1}}{\Gamma(\alpha) \beta^\alpha} g(c) \quad (18)$$

and one can show that the posterior distribution of  $b$  and  $c$  given the data  $\pi(b, c/\text{data})$  is proportional to

$$\pi(b, c/\text{data}) \propto b^{n+\alpha-1} c^{\sum x_i} g(c) \exp \{ b(n - \sum c^{x_i}) / \ln c - b/\beta \}.$$

Easy calculation shows that the Bayesian estimators  $\hat{b}$  and  $\hat{c}$  (assuming the quadratic loss) of the parameters  $b$  and  $c$  are

$$\hat{b} = \frac{(n+\alpha) \int_1^\infty c^{\sum x_i} g(c) \{1/\beta + (\sum c^{x_i} - n)/(\ln c)\}^{-(n+\alpha+1)} dc}{\int_1^\infty c^{\sum x_i} g(c) \{1/\beta + (\sum c^{x_i} - n)(\ln c)\}^{-(n+\alpha)} dc}$$

and

$$\hat{c} = \frac{\int_1^\infty c^{\sum x_i+1} g(c) \{1/\beta + (\sum c^{x_i} - n)/(\ln c)\}^{-(n+\alpha)} dc}{\int_1^\infty c^{\sum x_i} g(c) \{1/\beta + (\sum c^{x_i} - n)(\ln c)\}^{-(n+\alpha+1)} dc}.$$

If one use the noninformative prior

$$\pi(b, c) = b^{\alpha-1} db dc; b > 0, c > 1, \quad (19)$$

the generalized Bayes estimator for these parameters are

$$\hat{b} = \frac{(n + \alpha) \int_1^{\infty} c^{\sum x_i} g(c) \{(\sum c^{x_i} - n)/(\ln c)\}^{-(n+\alpha+1)} dc}{\int_1^{\infty} c^{\sum x_i} g(c) \{(\sum c^{x_i} - n)/(\ln c)\}^{-(n+\alpha)} dc}$$

and

$$\hat{c} = \frac{\int_1^{\infty} c^{\sum x_i + 1} g(c) \{(\sum c^{x_i} - n)/(\ln c)\}^{-(n+\alpha)} dc}{\int_1^{\infty} c^{\sum x_i} g(c) \{(\sum c^{x_i} - n)/(\ln c)\}^{-(n+\alpha+1)} dc}.$$

These integrals must be evaluated numerically. For small  $n$ , this can be done quite easily using any integration program. For large  $n$ , calculating these integrals are difficult.

As in section 2, one can find the adaptive Bayes estimators as follows. First assume that the value of the parameter  $c$  is known and the prior information about the parameter  $b$  can be expressed using the gamma prior with parameters  $\alpha$  and  $\beta$ ,

$$\pi(b) = e^{-b/\beta} b^{\alpha-1} / \{\Gamma(\alpha) \beta^\alpha\}. \quad (20)$$

Easy calculation shows that the posterior distribution of  $b$  given the data,  $\pi(b/\text{data})$  is proportional to

$$\pi(b/\text{data}) \propto b^{n+\alpha-1} \exp\{-b/\beta + b(n - \sum c^{x_i})/\ln c\}$$

and the Bayesian estimator of  $b$ ,  $\hat{b}$  is

$$\hat{b} = \frac{(n + \alpha)}{\{1/\beta + (\sum c^{x_i} - n)/\ln c\}}.$$

Now one can find the optimum  $c$  value by minimizing the "distance" between empirical cdf ( $F_{\text{emp}}(x)$ ) and the parametric cdf  $F(x)$  given in (15) or in (16).

Here



$$F(x) = 1 - \exp\{b(1 - c^x)/\ln c\}.$$

If no prior information is available about the parameter  $b$  one can use the noninformative generalized prior which has the density

$$\pi(b) = b^{\alpha-1} \quad 0 < b < \infty. \quad (21)$$

When  $c$  is known, the Jeffreys noninformative prior corresponds to the case  $\alpha = 0$ . The posterior density of  $b$  with respect to the prior given in (21) is

$$\pi(b/\text{data}) \propto b^{n+\alpha-1} \exp\{-b(\sum c^{x_i} - n)/\ln c\}$$

and the Generalized Bayes estimator of  $b$  is

$$\hat{b} = \frac{(n + \alpha) \ln c}{(\sum c^{x_i} - n)}.$$

Again one can find  $c$  by minimizing the area  $A_1$  given in (15) or Anderson & Darling  $A^2$  statistic given in (16). Notice that, when  $c$  is known, the generalized Bayes estimator of  $b$  with respect to the Jeffreys noninformative prior is the ML estimator of  $b$ .

For more details and performances of these procedures for the Gompertz distribution, see Ananda, Dalpatadu and Singh (1992). In the next section, couple of examples taken from the above paper are given to compare the performances of this proposed adaptive procedures with the other procedures.

#### 4. CASE EXAMPLES

**Example 1.** The following data from Hoel (1972) represents time (in days) at death of 39 irradiated mice. These life times in days are as follows:

40, 42, 51, 62, 163, 179, 206, 222, 228, 249, 252, 282, 324, 333, 341, 366, 385, 407, 420, 431, 441, 461, 462, 482, 517, 517, 524, 564, 567, 586, 619, 620, 621, 622, 647, 651, 686, 761, 763.

Johnson and Johnson (1980) showed that this data follows a two parameter Gompertz distribution. Since we do not have any prior information, we can't calculate the exact Bayesian procedure. We calculate the adaptive generalized Bayes estimators (with respect to the noninformative prior in (5) with  $\alpha = 0$  and  $\beta = \infty$ ) by minimizing 1. area and 2. Anderson and Darling Statistic. Chi-squared goodness of fit test was used to compare these methods with the other methods. Results of the analysis are given in the Table I. The p-value with the maximum likelihood procedure is 0.307. The p-values with the adaptive generalized Bayes procedures are 0.791 (by minimizing the area) and 0.729 (by minimizing the Anderson and Darling Statistic). Here, performances of the adaptive noninformative procedures are much better than all the other procedures.

**Example 2.** For the second example, we use data from page 136, Johnson and Johnson (1980) (original data given by Kimball (1960)). These are mortality data for 208 mice, which were exposed to gamma radiation. Maximum likelihood, Minimum chi-square, Minimum modified chi-square estimators of parameters  $b$  and  $c$  are given in Johnson and Johnson (1980). Since we do not have any prior information, we analyze this data set using the noninformative prior. Results of these analysis (estimated parameter values and chi-square comparison) are given in the Table II. As minimum chi-square estimators and minimum modified chi-square estimators are readily available from Johnson and Johnson, those methods also used in the chi-square comparison. While the p-value for the ML procedure is 0.178, p-values for the adaptive procedures are 0.188 (by minimizing area) and 0.200 (by minimizing A & D statistics). Here too, adaptive noninformative procedures are doing slightly better than all the other procedures.

Table I: Results from example 1

Time interval	Observed number	Expected number of deaths			
		$M_1$	$M_2$	$M_3$	$M_4$
0-100	4	6.901	4.444	2.592	2.723
100-200	2	6.763	5.306	3.729	3.845
200-300	6	6.330	6.002	5.104	5.165
300-400	5	5.603	6.306	6.458	6.422
400-500	7	4.635	5.986	7.232	7.090
500-600	6	3.532	4.946	6.700	6.527
600-700	7	2.436	3.381	4.632	4.557
700-800	2	1.490	1.785	2.048	2.096
df=		3	3	3	3
$\chi^2 =$		10.29	3.601	1.044	1.299
p-value=		0.016	0.307	0.791	0.729

$M_1$ : Percentile Estimators ( based on 25th and 75th percentiles)

$$\hat{c} = 1.00195 \quad \hat{b} = 0.00176393$$

$M_2$ : Maximum likelihood Estimators

$$\hat{c} = 1.00321 \quad \hat{b} = 0.00102648$$

$M_3$ : Adaptive Generalized Bayes (By minimizing the area,  $\alpha = 0$  and  $\beta = \infty$ )

$$\hat{c} = 1.00453 \quad \hat{b} = 0.00054404$$

$M_4$ : Adaptive Generalized Bayes (By minimizing A & D Stat.,  $\alpha = 0$  and  $\beta = \infty$ )

$$\hat{c} = 1.00438 \quad \hat{b} = 0.00057717$$

Table II: Results from example 2

Time interval	Observed number	Expected number of deaths					
		$M_1$	$M_2$	$M_3$	$M_4$	$M_5$	$M_6$
0 - 50	3	27.341	6.446	6.664	7.031	6.595	5.717
50 - 60	3	5.024	3.497	3.580	3.721	3.542	3.188
60 - 70	6	4.888	5.054	5.152	5.316	5.095	4.667
70 - 80	6	4.752	7.229	7.335	7.515	7.256	6.766
80 - 90	16	4.621	10.183	10.286	10.461	10.176	9.667
90 - 100	14	4.494	14.02	14.10	14.23	13.95	13.51
100 - 110	25	4.371	18.68	18.70	18.74	18.53	18.29
110 - 120	20	4.249	23.71	23.63	23.51	23.45	23.60
120 - 130	32	4.132	28.01	27.81	27.50	27.68	28.33
130 - 140	25	4.019	29.79	29.50	29.05	29.48	30.59
140 - 150	27	3.908	27.15	26.89	26.46	27.04	28.22
150 - 160	13	3.799	19.76	19.66	19.47	19.96	20.65
160 - 170	11	3.694	10.390	10.459	10.536	10.77	10.76
170 - 180	7	3.593	3.417	3.522	3.676	3.705	3.427
df=		9	9	9	9	9	9
$\chi^2=$		668.7	12.66	12.46	12.25	12.60	13.82
p-value=		0.	0.178	0.188	0.200	0.181	0.129

$M_1$ : Percentile Estimators (based on 10<sup>th</sup> and 90<sup>th</sup> percentiles)

$$\hat{c} = 1.00003 \quad \hat{b} = 0.00281676$$

$M_2$ : Maximum likelihood Estimators

$$\hat{c} = 1.03975 \quad \hat{b} = 0.00020389$$

$M_3$ : Adaptive Generalized Bayes Est. (By minimizing the area,  $\alpha = 0$  and  $\beta = \infty$ )

$$\hat{c} = 1.03935 \quad \hat{b} = 0.00021348$$

$M_4$ : Adaptive Generalized Bayes Est. (By minimizing A & D Stat.,  $\alpha = 0$  and  $\beta = \infty$ )

$$\hat{c} = 1.03871 \quad \hat{b} = 0.000230039$$

$M_5$ : Minimum chi-square Estimators

$$\hat{c} = 1.03931 \quad \hat{b} = 0.0002115 \text{ (from Johnson and Johnson (1980))}$$

$M_6$ : Minimum Modified chi-square Estimators

$$\hat{c} = 1.04090 \quad \hat{b} = 0.000174 \text{ (from Johnson and Johnson (1980))}$$

## 5. BIBLIOGRAPHY

- Ananda, M. A.; Dalpatadu, R. J. and Singh, A. K. (1992); Adaptive Bayes Estimators for Parameters of the Gompertz Survival Model, Submitted to Communications in Statistics: Theory and Methods.
- Anderson, T. W. and Darling, D. A. (1954); A test of goodness of fit; J. Amer. Statist. Assoc. 49, 765-769.
- Bain, L.J. (1978); Statistical Analysis of Reliability and Life- Testing Models- Theory and Methods. Marcel Dekker, New York.
- Berger, James (1985); Statistical Decision Theory; Springer-Verlag, New York.
- Hoel, D. G. (1972); A representation of mortality data by competing risks; Biometrics 28, 475-488.
- Jeffreys, H. (1961); The Theory of Probability; Oxford University Press, Oxford.
- Elandt-Johnson, R. C. and Johnson, N. L. (1980); Survival Models and Data Analysis, New York, John Wiley and Sons.
- Kimball, A. W. (1960); Estimation of mortality intensities in animal experiments; Biometrics 16, 505-521.
- London, Dick (1988); Survival Models and Their estimation, ACTEX publications.
- Stephens, M. A. (1974); EDF statistics for goodness of fit and some comparisons, J. Amer. Statist. Assoc. 69, 730-737.

