HUNGARIAN JOURNAL
OF INDUSTRIAL CHEMISTRY
VESPRÉM
Vol. 33(1-2)., pp. 69-73. (2005)

ESTIMATION OF PARAMETERS IN 3-PARAMETER WEIBULL PROBABILITY DISTRIBUTION FUNCTIONS

R. LUUS and M. JAMMER

Department of Chemical Engineering, University of Toronto, Toronto, ON M5S 3E5, CANADA

This paper was presented at the 10th International Workshop on Chemical Engineering Mathematics, Budapest, Hungary, August 18-20, 2005

Three-parameter Weibull probability distribution function is used to represent the time-to-failure data. For parameter estimation, the errors-in-variables, maximum likelihood, and least squares methods are compared. The results obtained from five data sets show that maximum likelihood method gives the most reliable parameter estimates which are close to those obtained by errors-in-variables approach. The least squares method gave the poorest results in most cases. For parameter estimation, the Luus-Jaakola direct search optimization procedure yielded the optimal parameter values in negligible computation time, taking less than a second of computation time on a Pentium4/2.66MHz personal computer for 50 data points.

Keywords: parameter estimation, Weibull, maximum likelihood, least squares, errors-in-variables

Introduction

Similar items, such as 100-watt incandescent light bulbs produced by the same company, subjected to similar environmental conditions tend to fail at different times. To predict the life of such items, the three-parameter Weibull probability distribution function has been found to represent the time-to-failure data reasonably well.

The Weibull cumulative distribution function is

$$F(t) = 1 - \exp[-(\frac{t-\mu}{\alpha})^{\beta}], \quad \mu < t \tag{1}$$

and its derivative is the Weibull probability density function

$$f(t) = (\frac{\beta}{\alpha^{\beta}}(t-\mu)^{\beta-1}) \exp[-(\frac{t-\mu}{\alpha})^{\beta}]. \tag{2}$$

The data are in the form of time to failure for n similar items arranged in order so that t_1 is the time to failure of the first item and t_n is the time to failure for the n^{th} item.

Statistically, a good estimate for the Weibull cumulative distribution function for n data points, as suggested by O'CONNER [1] and used recently by HUNG [2], is

$$F_s(i) = \frac{i-0.3}{n+0.4}. \tag{3}$$

The problem is to estimate the parameters μ, α, and β from a given reliability data set.

There are three approaches that may be used to obtain the parameters, depending on the criterion used for optimization.

Errors-in-Variables Estimation (EIV)

Recently, errors-in-variables approach has been found useful for parameter estimation where there is a significant error in the independent variables [3]. Instead of estimating only the parameters, the method also estimates the values of the independent variables.

Here we choose the performance index to be minimized as

$$I = \sum_{i=1}^{n} (t_i - t_{si})^2 \tag{4}$$

where t_i are the actual measured times to failure and

$$t_{si} = \mu + \alpha[-\ln(1 - F_s(i))]^{1/\beta} \tag{5}$$

which, with the use of *Eq.(3)*, becomes

$$t_{si} = \mu + \alpha[-\ln(1-\frac{i-0.3}{n+0.4})]^{1/\beta}. \qquad (6)$$

Maximum likelihood estimation (ML)

The likelihood function is the product of the individual probability density functions. Thus the aim is to maximize the likelihood function

$$L = \prod_{i=1}^{n} f(t_i) \qquad (7)$$

which is written as

$$L = \prod_{i=1}^{n} (\frac{\beta}{\alpha^{\beta}}(t_i-\mu)^{\beta-1}) \exp[-(\frac{t_i-\mu}{\alpha})^{\beta}]. \qquad (8)$$

To prevent computational difficulties, it is further specified that $\beta \geq 1$. Otherwise, if $\beta < 1$ and if μ is very close to t_1, the term $(t_1 - \mu)^{\beta-1}$ becomes very large.

Least squares estimation (LS)

An alternative to the above two methods is to consider as a performance index the sum of squares of deviations of the cumulative distribution functions

$$S = \sum_{i=1}^{n} [F(t_i) - F_s(i)]^2 \qquad (9)$$

Substituting *Eq.(1)* and *Eq.(3)* into *Eq.(9)* gives

$$S = \sum_{i=1}^{n} [1 - \exp[-(\frac{t_i-\mu}{\alpha})^{\beta}] - \frac{i-0.3}{n+0.4}]^2 . \qquad (10)$$

The values of the three parameters α, β, and μ obtained from a set of data will depend on the choice of the performance index. The goal here is to examine the parameters obtained by these three performance indices I, L, and S.

Optimization procedure

To obtain the three parameters, we decided to use direct search optimization, so that no transformations or auxiliary variables would have to be calculated. There are numerous direct search optimization procedures that may be used. Recently, genetic algorithm has been favored by some, but recent comparison of the genetic algorithm to the direct search procedure introduced by LUUS and JAAKOLA [4], and refined recently by LUUS [5-7], showed that the LJ optimization procedure tends to be somewhat faster and more reliable than the genetic algorithm [8]. Therefore, we used the LJ optimization procedure for the optimization.

The LJ optimization procedure involves taking a number of random sampling points over a region, finding the best point, and then using the best point as the centre of the region for the next iteration. However, at the beginning of this iteration the region size is reduced by a factor $\gamma < 1$ to make the search more intensive around the best point. This procedure is continued for a number of iterations to finish a pass. Since it is desirable to obtain the optimum value of the performance index very accurately, a number of passes is usually necessary. The algorithm for the LJ optimization procedure is given by LUUS [9].

We used a multi-pass procedure, involving 25 passes, each consisting of 21 iterations. The initial values for the parameters were: $\mu = 0.5\ t_1$, $\alpha = t_1$, $\beta = 2.0$, and the initial region size for the first pass was taken as 0.5 times the initial value. After every iteration the region sizes were reduced by $\gamma = 0.95$. At the beginning of the second pass, the region sizes were put to 0.01 times the parameter values. For the remaining passes, the region size for each parameter at the beginning of the pass was put equal to the amount by which the parameter changed during the previous pass. If this change, divided by the parameter value, was less than 10^{-6}, then the region size was put to 10^{-6} times the parameter value. The parameter estimates were run with $R = 25$ and with $R = 100$ random points per iteration.

Numerical results

All computations were done in double precision using WATCOM Fortran compiler version 9.5 on a Pentium4/2.66GHz personal computer.

Example 1: Lifetime of light bulbs

As the first example we chose the data of WALPOLE and MYERS [10] relating to the lifetimes of 50 internally frosted incandescent 40-watt 110-volt light bulbs. The shortest lifetime was 702 h, so we chose the initial values: $\mu = 351$, $\alpha = 702$, and $\beta = 2$. Using the EIV approach, after 25 passes, we obtained the minimum value $I = 1.754306594335 \times 10^4$ with $R = 100$ in a total computation time of 0.83 s, and $I = 1.754306594336 \times 10^4$ with $R = 25$ in a computation time of 0.22 s. The convergence profile given in *Fig.1* shows that 25 randomly chosen test points per iteration are quite sufficient to get accurate values for the parameters within 25 passes. However, when 100 points are used, convergence to 12 figures is obtained already in 11 passes.

Maximization of the likelihood function given in *Eq.(8)* yielded with $R = 100$ in a computation time of 1.49 s the maximum value $L = 1.0285396 \times 10^{-139}$ with parameter values which are very close to those obtained by EIV, as is shown in *Table 1*. To calculate such a small number, each probability density function was multiplied by 10^3 and then the resulting likelihood function at the end was multiplied by 10^{-150}. This avoided any possibility of underflow problems.

Minimization of the sum of squares of deviations given in *Eq.(9)* yielded with $R = 100$ in a computation time of 0.50 s the minimum $S = 5.6267167 \times 10^{-2}$ with parameter values given in *Table 1*. These values are quite

different from those obtained from the other two methods. Also, the deviations of the estimated times to failure and the measured times to failure are substantially higher, as is shown in the last column of *Table 1*.

It is natural for EIV to outperform the other two methods if the deviations of the same data points are used for estimating the parameters and then used for the evaluation, since in EIV the evaluation function is used for estimating the parameters. We therefore used only half of the data to obtain the parameters. Starting with the first data point, every second data point was used, plus the last data point, giving 26 data points.

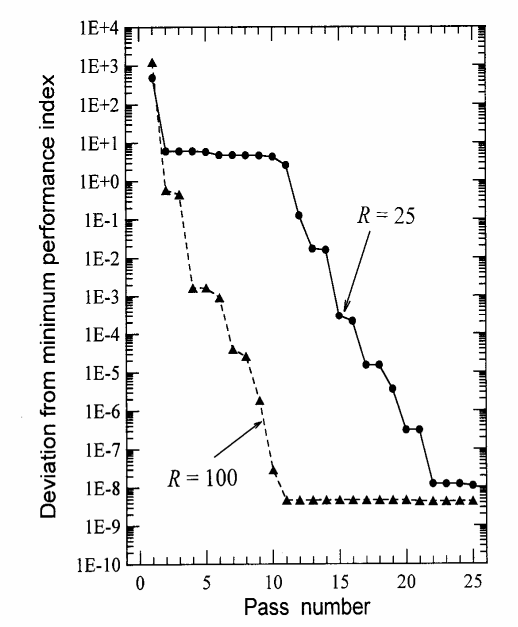

Fig.1 Convergence profile for Example 1, showing the effect of the number of random points used per iteration

For evaluation of the results, the entire data length was used. As is shown in *Table 2*, the maximum likelihood method yielded the best results. Such evaluation was carried out with several additional data sets available in the literature.

Example 2: Lifetime of batteries

We now consider the lives of 40 similar batteries recorded to the nearest tenth of a year, as reported by WALPOLE AND MYERS [9]. Convergence was easily obtained, yielding $I = 0.36569261$, $L = 7.5001463 \times 10^{-19}$, and $S = 3.724944 \times 10^{-2}$ with the parameter values given in *Table 3* for all the data.

Table 1 Parameters obtained for Example 1 with the use of 50 time-to-failure data points: 702, 765, 785, 811, 832, 855, 896, 902, 905, 918, 919, 920, 923, 929, 936, 938, 948, 950, 956, 958, 958, 970, 972, 978, 1009, 1009, 1022, 1035, 1037, 1045, 1067, 1085, 1092, 1102, 1122, 1126, 1151, 1156, 1157, 1157, 1162, 1170, 1195, 1195, 1196, 1217, 1237, 1311, 1333, 1340

	μ	α	β	$\sum_{i=1}^{50}(t_i - t_{si})^2$
EIV	626.155	450.129	2.90623	1.7543×10^4
ML	623.527	452.020	3.00294	1.8188×10^4
LS	702.000	371.347	2.25438	2.3295×10^4

Table 2 Parameters obtained for Example 1 with the use of every second data point and the last one (26 data points) but evaluated with all data points

	μ	α	β	$\sum_{i=1}^{50}(t_i - t_{si})^2$
EIV	577.18	510.72	3.0347	2.8713×10^4
ML	590.06	495.22	3.0953	2.1177×10^4
LS	702.00	380.64	2.1505	4.0526×10^4

Table 3 Parameters obtained for Example 2, given the time-to-failure data: 1.6, 1.9, 2.2, 2.5, 2.6, 2.6, 2.9, 3.0, 3.0, 3.1, 3.1, 3.1, 3.1, 3.2, 3.2, 3.2, 3.3, 3.3, 3.3, 3.4, 3.4, 3.4, 3.5, 3.5, 3.6, 3.7, 3.7, 3.7, 3.8, 3.8, 3.9, 3.9, 4.1, 4.1, 4.2, 4.3, 4.4, 4.5, 4.7, 4.7

	μ	α	β	$\sum_{i=1}^{40}(t_i - t_{si})^2$
EIV	0.00000	3.69330	5.51662	0.36569
ML	0.10346	3.58331	5.49813	0.38317
LS	1.60000	2.05230	3.18526	0.82921

In *Table 4* are the parameter values for half of the data. By comparing the estimates in these two tables, again the maximum likelihood method is seen to give the best results if half of the data are used to obtain the parameters, and the parameter values are evaluated with the entire data length. Again, the least squares method yielded the worst results.

Table 4 Parameters obtained for Example 2 with the use of every second data point and the last one (21 data points) but evaluated with all data

	μ	α	β	$\sum_{i=1}^{40}(t_i - t_{si})^2$
EIV	0.00000	3.7487	4.9521	0.65398
ML	0.00000	3.7382	5.2934	0.45611
LS	1.60000	2.1089	2.8953	0.85610

Example 3: Data used by Lockhart and Stephens

The data given by COX and OAKES [11], and used by LOCKHART and STEPHENS [12] gave $I = 261.79902$, $L = 9.0746400 \times 10^{-22}$, and $S = 2.0017367 \times 10^{-2}$ with a wide range of values for the parameters, as is seen in *Table 5*. The ML values obtained here correspond very closely to the ML values obtained by Lockhart and Stephens, namely, 99.02, 78.23 and 2.38.

Table 5 Parameters obtained for Example 3, given the time-to-failure data: 117, 135, 135, 162, 162, 171, 189, 189, 198, 225

	μ	α	β	$\sum_{i=1}^{10}(t_i - t_{si})^2$
EIV	71.5445	108.641	3.0214	261.799
ML	99.0109	78.240	2.3755	412.692
LS	13.9179	167.491	4.7922	321.376

When only half of the data were used, we see in *Table 6* that the maximum likelihood method yielded the most consistent values for the parameters.

Table 6 Parameters obtained for Example 3 with the use of every second data point and the last one (6 data points) but evaluated with entire data length

	μ	α	β	$\sum_{i=1}^{10}(t_i - t_{si})^2$
EIV	17.323	169.824	3.8339	1068.61
ML	102.506	77.016	1.8877	423.78
LS	0.000	188.474	4.0675	1499.66

Example 4: Lifetime of fruit flies

For the data on the lives of 50 fruit flies in seconds when exposed to a spray in a controlled laboratory experiment as given by WALPOLE AND MYERS [10], we obtained the parameter values shown in *Table 7* with $I = 18.879469$, $L = 4.8751765 \times 10^{-18}$, and $S = 3.8023533 \times 10^{-2}$.

Table 7 Parameters obtained for Example 4, given the time-to-failure data: 3, 4, 5, 6, 6, 6, 7, 7, 7, 7, 7, 7, 7, 8, 8, 9, 9, 9, 9, 10, 10, 10, 10, 10, 10, 11, 12, 12, 13, 13, 13, 13, 13, 14, 14, 15, 15, 16, 16, 17, 18, 18, 18, 19, 19, 20, 23, 24, 27, 32

	μ	α	β	$\sum_{i=1}^{50}(t_i - t_{si})^2$
EIV	3.0000	10.3405	1.4947	18.879
ML	2.6548	10.8189	1.6606	28.646
LS	3.0000	10.3032	1.5949	27.829

When only half of the data were used, as is seen in *Table 8*, the best results were obtained with the least squares method.

Table 8 Parameters obtained for Example 4 with the use of every second data point and the last one (26 data points) but evaluated with entire data length

	μ	α	β	$\sum_{i=1}^{50}(t_i - t_{si})^2$
EIV	3.0000	10.6904	1.3285	84.548
ML	2.5974	11.2482	1.5300	36.322
LS	3.0000	10.6047	1.4892	23.355

Example 5: Lifetime of fuel pumps

For the data on the lives in years of 30 similar fuel pumps as presented by WALPOLE AND MYERS [10], convergence to $I = 14.135236$, $L = 3.456841 \times 10^{-26}$, and $S = 6.2525950 \times 10^{-2}$ was easily obtained, yielding the parameter values in *Table 9*.

When only half of the data were used, *Table 10* shows that EIV yielded the best results. Here it is noted that for the maximum likelihood method μ is very close to t_1 and β is very close to 1, so the maximum likelihood estimation procedure is not very reliable. For this example, therefore, the errors-in-variables approach gives the most reliable parameter estimates.

Table 9 Parameters obtained for Example 5, given the time-to-failure data: 0.2, 0.2, 0.2, 0.3, 0.3, 0.4, 0.5, 0.7, 1.0, 1.2, 1.3, 1.5, 1.5, 1.8, 2.0, 2.3, 2.5, 3.0, 3.3, 4.0, 4.5, 4.7, 5.0, 5.5, 5.6, 5.9, 6.0, 6.0, 6.0, 6.5

	μ	α	β	$\sum_{i=1}^{30}(t_i - t_{si})^2$
EIV	0.00000	3.2353	1.4346	14.135
ML	0.20000	2.5965	1.0000	26.517
LS	0.00000	3.1446	0.9248	68.778

Table 10 Parameters obtained for Example 5 with the use of every second data point and the last one (16 data points) but evaluated with entire data length

	μ	α	β	$\sum_{i=1}^{50}(t_i - t_{si})^2$
EIV	0.00000	3.3732	1.3757	15.642
ML	0.20000	2.6999	1.0000	28.337
LS	0.00000	3.2830	0.8826	105.122

Conclusions

The Luus-Jaakola optimization procedure is easy to use for parameter estimation for the Weibull distribution, since no transformations are required. The optimal parameter values are readily obtained in negligible computation time on a personal computer.

Errors-in-variables approach is easy to use with direct search optimization, and no computational difficulties were encountered with the use of the LJ optimization procedure. However, with the maximum likelihood function, great care is required if $\beta < 1$, and μ is very close to t_1. The parameter estimates obtained by EIV are closer to those obtained by maximum likelihood than by least squares. The least squares approach tends to give the least reliable parameter estimates. For most of the data sets, the maximum likelihood method gave the most consistent estimates for the parameters when a shorter data length was used. It is recommended to use both the errors-in-variables and the maximum likelihood methods to obtain the parameters in the Weibull distribution and then to examine and evaluate the results.

Acknowledgement

Financial support from the Natural Sciences and Engineering Research Council of Canada is gratefully appreciated.

SYMBOLS

f - Weibull probability density function
F - Weibull cumulative distribution function
F_s - statistically determined Weibull cumulative distribution function
I - performance index for errors-in-variables
L - likelihood function
n number of items tested
R - number of random points used in each iteration
S - sum of squares of deviations of cumulative distribution functions
t_i - time for the i^{th} item to fail
α - parameter to be determined
β - parameter to be determined
γ - region size reduction after every iteration in the LJ optimization procedure
μ - parameter to be determined

Acronyms

EIV - errors-in-variables estimation
LJ - Luus-Jaakola optimization procedure
LS - least squares estimation
ML - maximum likelihood estimation

REFERENCES

1. O'CONNER P.D.T.: Practical Reliability Engineering, Wiley, New York, 1985
2. HUNG W.L.: Quality and Reliability Engineering International, *2001*,17(6),467-469
3. LUUS R. and HERNAEZ H.: Hungarian J. Industrial Chemistry, *2000*, 28(3), 201-206
4. LUUS R. and JAAKOLA T.H.I.: AIChE J., *1973*, 19(4), 760-766.
5. LUUS R.: Recent Research Developments in Chemical Engineering, *2000*, 4, 45-64.
6. LUUS R.: Iterative Dynamic Programming, Chapman & Hall/CRC, London, UK, 2000
7. LUUS R.: Direct search Luus-Jaakola optimization procedure, LJ optimization procedure, in FLOUDAS C.A. and PARDALOS P.M. (Eds.), Encyclopedia of Optimization, Kluwer, Dordrecht, The Netherlands, 2001, Vol. 1, 440-444
8. LIAO B. and LUUS R.: Engineering Optimization, *2005*, 37(4), 381-398
9. LUUS R.: Hungarian J. Industrial Chemistry, *2000*, 28(3), 217-223
10. WALPOLE R.E. and MYERS R.H.: Probability and Statistics for Engineers and Scientists, Macmillan, New York, 1993
11. COX D.R. and OAKES D.: Analysis of Survival Data, Chapman and Hall, London, UK, 1984
12. LOCKHART R.A. and STEPHENS M.A.: J. Royal Statistical Society, Series B, *1994*, 56(3), 491-500.