A Simple Analysis Method of Products Lifetime Distribution

Yueming Wang

Fiber Optic Sensing Technology Research Center

Wuhan University of Technology

Wuhan 430070, China

Abstract

The basis of reliability analysis is determining products lifetime distribution type, which is also a tedious job. In accordance with the failure of parametric estimation using probability papers and hypothetical examination, products lifetime distribution type is studied with linear regression analysis and Office Excel.

Keywords: Cumulative distribution function, Reliability, Linear regression analysis, Correlation coefficient, Office Excel

The lifetime distribution type of a product is determined for analyzing reliability of it.Parametric estimation using probability papers and linear hypothesis examination are two commonly methods (*ZHU Min-bo,WANG Shi-ping.Study and Application on Analysis Method Using Probability Papers under Reliability Test[J].COMPUTER AND PERIPHERALS,1997,21(6),66.*), but both of them have deficiencies. The former is intuitive, but difficult for manual drawing, more artificially imposed factor analysis of low level precision.The test data of some products is approximately close to a linear line on the various kinds of probability papers,and it is difficult to determine its most suitable distribution type.The latter is usually supposing some distribution type suitable for the product lifetime and then making correlation examination.If correlation coefficient is greater than the minimum value,the hypothesis about lifetime distribution type is tenable .But this method may miss more reasonable distribution type.

The linear regression analysis method is accurate calculation that is widely used in various fields,and also can be used to determine the distribution types of products lifetime.

1. linear regression correlation coefficient calculation with Excel

Basic thoughts of linear regression analysis is to analyze a non-linear problem by turning it into a linear problem and to make linear correlation coefficient **R** as examination,**R** is defined as (*CHEN Min-xiang.Lifetime Distribution Analysis and Reliability Estimation of Electric Motors[J].ELECTRIC MACHINES AND CONTROL,1998,2(2),89.*):

$$R = \frac{\sum_{i=1}^{n}(t_i-\bar{t})(Z_i-\bar{Z})}{\sqrt{\sum_{i=1}^{n}(t_i-\bar{t})^2 \sum_{i=1}^{n}(Z_i-\bar{Z})^2}} \tag{1}$$

In above formula, (t_i, Z_i) is test data,$\bar{t}$, $\bar{Z}$ are the the average values,namely:

$$\bar{t} = \tfrac{1}{n}\sum_{i=1}^{n} t_i$$

$$\bar{Z} = \tfrac{1}{n}\sum_{i=1}^{n} Z_i \qquad (i=1,2,3,...n)$$

From formula (1), it is tedious to calculate **R** value, but the Office Excel tool will make it simple.If scattered points sketch is drawn with Excel,it is much more simple, it is simple to get R^2 and **R** value directly.

2. Product lifetime distribution type analysis method

In reliability analysis,product lifetime distribution type usually include an exponential distribution,normal distribution,lognormal distribution and the Weibull distribution.These cumulative distribution functions are all nonlinear and very complex,they can be transformed into linear equations through variable transformation, Scattered points sketch about products test data for $(t_i, F(t_i))$ can be drawn directly with Excel and **R** value can be given directly with Excel.By comparing these **R** values,the distribution type correspondence with the maximum value of **R** is the most

reasonable lifetime distribution function for the product.

3. Linearization of cumulative distribution function and determination of products lifetime distribution type

To estimate the distribution type and characteristic parameters of all products through the experimental results for large samples,histogram with frequency usually used.But in engineering practice,There are usually small samples.It is inaccurate to estimate the result of small samples with estimation method for a large samples.To estimate the probability of small samples,methods of calculating the cumulative distribution probability are usually used,namely,mid-rank being calculated as follows:

$$F(t_i) = \frac{i}{n+1} \tag{2}$$

In above formula, *n* is the total number of samples under test, t_i is the sequence number of failure samples, which arranges from short lifetime to long lifetime.

Now motor accelerated lifetime data shown in table 1(*CHEN Min-xiang.Lifetime Distribution Analysis and Reliability Estimation of Electric Motors[J].ELECTRIC MACHINES AND CONTROL,1998,2(2),89*) is studied as an example, $F(t_i)$ in the table 1 can be given immediately with Excel according to formula (2).

3.1 Linearization of an exponential distribution function

Cumulative distribution function of two-parameter exponential distribution is as follows:

$$F(t) = 1 - e^{-\frac{t-\gamma}{\theta}} \qquad (t \geq \gamma) \tag{3}$$

The above formula is arranged in order and changed in natural logarithm,it can become a standard linear equation:

$$Z_e(t) = at + b \tag{4}$$

In above equation,

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

$$a = \frac{1}{\theta}$$

$$b = -\frac{\gamma}{\theta}$$

$Z_e(t)$ can be given immediately with Excel according to formula (5), shown in table 2.

According to table 2,take *t* as x-axis and $Z_e(t)$ as y-axis,scattered points sketch about products test data is drawn with Excel,and then,selecting Linear Type in Adding the Trend Line, clicking on Showing Formula button and Showing Square Of R button in the function of Select,pressing Enter,linear tread chart can be drawn automatically,its linear formula and the value of R^2 can be given immediately,as shown in Figure 1, the linear equations of exponential distribution function is as bellow:

$$y = 8 \times 10^{-7} x - 1.4674 \tag{6}$$

$$R^2 = 0.9506 \tag{7}$$

From equation (6), $a = 8 \times 10^{-7}, b = -1.4674$,From equation (4)and(5), the parameters of θ and γ can be given.

From equation (7), the correlation coefficient **R** is given,namely: $R = 0.9750$.

3.2 Linearization of normal distribution functions

The cumulative distribution function of normal distribution is as follows:

$$F(t) = \int_0^t \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{t-\mu}{\sigma})^2} dx = \Phi(\frac{t-\mu}{\sigma}) = \Phi(Z_N) \tag{8}$$

In it,

$$Z_N = \frac{1}{\sigma}t - \frac{\mu}{\sigma} \tag{9}$$

Taking the inverse of the standard normal cumulative distribution,selecting NORMSINV(n/(N+1)) in some spreadsheets,transforming $F(t_i)$ (or $\Phi(Z_N)$) into $Z_N(i)$,listed in table 3.

According to the same method as subclauses 3.1, Figure of Linearization and correlation coefficient of normal distribution functions can be drawn with Excel.

3.3 Linearization of lognormal distribution function

The accumulative distribution function of lognormal distribution(*ZHANG Zhi-hua.Statistical Analysis of Accelerated Life Test on Lognormal Distribution Case Under Type I Censoring[J].ACTA MATHEMATICAE APPLICATAE SINICA,2000,23(4),489.*) is as follows:

$$F(t) = \int_0^t \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{1}{2}(\frac{\ln x - \mu}{\sigma})^2} dx = \Phi(\frac{\ln t - \mu}{\sigma}) = \Phi(z_{LN}) \tag{10}$$

In it,

$$Z_{LN}(t') = \frac{1}{\sigma} t' - \frac{\mu}{\sigma} \tag{11}$$

$$t' = \ln t \tag{12}$$

According to the same method as subclauses 3.2,select NORMSINV(n/(N+1)),transform $\Phi(z_{LN})$ (or $F(t_i)$) into $Z_{LN}(i)$, $Z_{LN}(i)$ can be given as the method in table 3,and according to the equation(12),change t into t' ,,listed in table 4 with Excel.

According to the same method as subclauses 3.1, Figure of Linearization and correlation coefficient of lognormal distribution function can be drawn with Excel.

3.4 Linearization of Weibull distribution function

The cumulative distribution function of three-parameter Weibull distribution(*TONG Xiao-jun,MA QUN.General Regression Analysis Method for Estmating The Weibull Three Parameters[J].JOURNAL OF TANGSHAN INSTITUTE OF TECHNOLOGY,1995,17(2),78.*) is as follows:

$$F(t) = 1 - e^{-(\frac{t-\gamma}{\eta})^m} \tag{13}$$

In it, γ , η and m are the location parameter, the scale parameter and the shape parameter (or Weibull slope) of Weibull distribution, respectively.

The equation (13)is arranged in order and taken natural logarithm, it is:

$$Z_w(t) = at' + b \tag{14}$$

In it,

$$Z_w(t) = \ln \ln \frac{1}{1-F(t)}$$

$$a = m$$

$$t' = \ln(t - \gamma)$$

$$b = -m \ln \eta$$

From equation (14), if (t_i , $F(t_i)$) obey Weibull distribution, relationship between $Z_w(t_i)$ and t_i' should be linear,but its linear correlation relates to the location parameter γ . According to the drawing principle of Weibull probability paper, γ value correspondence with maximum R value is the most reasonable location parameter estimation of Weibull distribution.

Due to the complex R value calculation, trial calculation by equation (1) is very simple and accurate with Excel.The method of trial calculation(*FANG Zhi-qiang,GAO Lian-hua.Estimation of Parameters of Three-parameter Weibull Distribution in Life Analysis[J].Journal of Armored Force Engineering Institute,1999,13(1),72.*) in Excel spreadsheets:Suppose a minimum lifetime t_1 was location parameter γ ,Then,a start from $\gamma_2 = 0.95\gamma_1$, step down from $\Delta = 0.05\gamma_1$,The respective corresponding correlation coefficient R is calculated,the location parameter correspondence with the maximum R value is γ . In this example, $\gamma = 1.73 \times 10^6$.

According to equation (14), $Z_w(t_i)$ can be calculated directly with Excel, shown in table 5.

According to the above method in subclauses 3.1,Figure 2 is drawn with Excel.

From figure 1 to figure 4, each of the four distribution functions can be fit a linear relationship sketch,and among them the linear correlation coefficient R value correspondence with Weibull distribution function is the maximum.It indicates that Weibull distribution is the most appropriate function to describe products lifetime in this example.

4. Conclusion

Based on the linear regression analysis,products lifetime distribution function is linearized,fitting-straight line,its equation and its correlation coefficient R value are given quickly and easily by using Excel auxiliary drawing method.Through comparing the R values,the distribution function correspondence with the maximum of R values describes the most reasonably lifetime distribution,which is a simple method to determine the distribution type of product lifetime. According to the fit linear equations, correlative parameters can also accurately calculated.

References

Chen, Minxiang. Lifetime Distribution Analysis and Reliability Estimation of Electric Motors[J]. *ELECTRIC MACHINES AND CONTROL*, 1998, 2(2), 88-95. (in Chinese)

Fang, Zhiqiang, Gao, Lianhua. Estimation of Parameters of Three-parameter Weibull Distribution in Life Analysis[J]. *Journal of Armored Force Engineering Institute*, 1999, 13(1), 72-74. (in Chinese)

Tong, Xiaojun, Ma, Qun. General Regression Analysis Method for Estmating The Weibull Three Parameters[J]. *JOURNAL OF TANGSHAN INSTITUTE OF TECHNOLOGY*, 1995, 17(2),78. (in Chinese)

Zhang, Zhihua.Statistical Analysis of Accelerated Life Test on Lognormal Distribution Case Under Type I Censoring[J]. *ACTA MATHEMATICAE APPLICATAE SINICA*, 2000, 23(4), 489. (in Chinese)

Zhu, Minbo,Wang, Shiping. Study and Application on Analysis Method Using Probability Papers under Reliability Test[J]. *COMPUTER AND PERIPHERALS*, 1997, 21(6), 66-67. (in Chinese)

Table 1. motor accelerated lifetime test data, (t_i is reversal times)

i	t_i/h	$F(t_i)$	i	t_i/h	$F(t_i)$
1	1.82E+06	0.0714	8	2.76E+06	0.5714
2	1.95E+06	0.1429	9	2.92E+06	0.6429
3	2.20E+06	0.2143	10	3.55E+06	0.7143
4	2.18E+06	0.2857	11	3.90E+06	0.7857
5	2.25E+06	0.3571	12	4.20E+06	0.8571
6	2.27E+06	0.4286	13	4.44E+06	0.9286
7	2.43E+06	0.5000			

Table 2. corresponding relation between $Z_e(t)$ and t

i	t_i/h	$z_e(t)$	i	t_i/h	$z_e(t)$
1	1.82E+06	0.0741	8	2.76E+06	0.8473
2	1.95E+06	0.1542	9	2.92E+06	1.0296
3	2.20E+06	0.2412	10	3.55E+06	1.2528
4	2.18E+06	0.3365	11	3.90E+06	1.5404
5	2.25E+06	0.4418	12	4.20E+06	1.9459
6	2.27E+06	0.5596	13	4.44E+06	2.6391
7	2.43E+06	0.6931			

Table 3. corresponding relationship between $Z_N(t)$ and t

i	t_i/h	$z_N(t)$	i	t_i/h	$z_N(t)$
1	1.82E+06	-1.4657	8	2.76E+06	0.1800
2	1.95E+06	-1.0676	9	2.92E+06	0.3600
3	2.20E+06	-0.7917	10	3.55E+06	0.5659
4	2.18E+06	-0.5659	11	3.90E+06	0.7917
5	2.25E+06	-0.3660	12	4.20E+06	1.0676
6	2.27E+06	-0.1800	13	4.44E+06	1.4657
7	2.43E+06	0.0000			

Table 4. corresponding relationship between $Z_{LN}(t')$ and t'

i	t'_i/h	$z_{LN}(t'_i)$	i	t'_i/h	$z_{LN}(t'_i)$
1	14.415	-1.4657	8	14.831	0.1800
2	14.481	-1.0676	9	14.888	0.3600
3	14.606	-0.7917	10	15.082	0.5659
4	14.596	-0.5659	11	15.177	0.7917
5	14.626	-0.3660	12	15.249	1.0676
6	14.634	-0.1800	13	15.306	1.4657
7	14.702	0.0000			

Table 5. corresponding relation between $Z_w(t_i)$ and t'_i

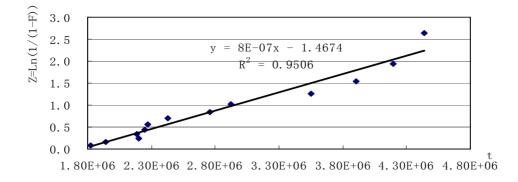| i	t'_i/h	$Z_w(t_i)$	i	t'_i/h	$Z_w(t_i)$
1	11.3144	-2.6022	8	13.8373	-0.1657
2	12.2307	-1.8698	9	13.9836	0.0292
3	13.0476	-1.4223	10	14.4083	0.2254
4	13.0013	-1.0892	11	14.5861	0.4321
5	13.1402	-0.8168	12	14.7136	0.6657
6	13.1768	-0.5805	13	14.8084	0.9704
7	13.4386	-0.3665			

Figure 1. Linear and correlation coefficient of exponential distribution function
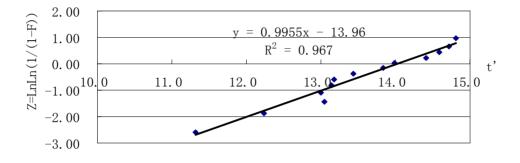

Figure 2. Linearization and correlation coefficient of Weibull distribution function