

Digital Experiment for Estimating Three Parameters and Their Confidence Intervals of Weibull Distribution

Jiajin Xu

L&Z International Leasing Co. Ltd, Ganton, China

Email address:

xujiajin666@163.com

To cite this article:

Jiajin Xu. Digital Experiment for Estimating Three Parameters and Their Confidence Intervals of Weibull Distribution. *International Journal of Science, Technology and Society*. Vol. 10, No. 2, 2022, pp. 72-81. doi: 10.11648/j.ijsts.20221002.16

Received: April 6, 2022; **Accepted:** April 22, 2022; **Published:** April 28, 2022

Abstract: Weibull distribution plays a very important role in fatigue statistics, because structural fatigue life mostly conforms to Weibull distribution rather than Gaussian distribution. However, the biggest obstacle to the application of Weibull distribution is the complexity of Weibull distribution, especially the estimation of its three parameters is difficult, and it is even more difficult to determine the confidence interval of each parameter. In order to solve the first problem, many methods have been proposed, such as the graph method, the analytical method, and the Gao Zhentong method proposed by me recently using the characteristics of Python. So, the question arises which method is better? To this end, referring to the idea and methods of machine learning, the concept of digital experiment is introduced. The digital experiment is carried out to determine the three parameters and their confidence interval of Weibull distribution. Experiments show that the Gao Zhentong method is indeed a better method for estimating the three parameters of Weibull distribution. Further, the bootstrap method that can be used for a digital experiment is introduced, and it is used to determine the three parameters and their confidence intervals of Weibull distribution. The results show that the three parameters and their confidence intervals of Weibull distribution can basically be determined at the same time.

Keywords: Three-Parameter Weibull Distribution, Confidence Interval, Gao Zhentong Method, Digital Experiment, Bootstrap, Python

1. Introduction

The PDF of the three-parameter Weibull distribution can be written as,

$$f(x)=(b/\lambda)[(x-x_0)/\lambda]^{b-1}\exp\{-(x-x_0)/\lambda\}^b\}$$

Among them, b is the shape parameter, λ is the scale parameter, and x_0 is the position parameter (called safe life in fatigue statistics). And the Weibull distribution not only plays a big role in fatigue statistics, but also, as "Wikipedia" points out, "is the theoretical basis for reliability analysis and life testing" [1]. The problem as for the complexity of the Weibull distribution, people often change the three-parameter Weibull distribution into a two-parameter Weibull distribution for convenience, that is, the position parameter x_0 is zeroed [1]. However, this parameter has a clear physical meaning in fatigue life, that is, the safety life under 100% reliability, referred to as "safe life", which

cannot be reset to zero [2]. Furthermore, from a mathematical point of view, the Weibull distribution is a "full state distribution" [2]. That is, it can show different shapes according to its shape parameter b , "When $0 < b < 1$, it is similar to the $1/x$ function, while $1 < b < 3$ is a left-biased distribution, and $3 < b < 4$ is approximately Gaussian distribution, and $b > 4$ is a right-skewed distribution" [2]. This is why some fatigue life data are sometimes fitted with a Gaussian distribution to get a better fitting effect. In this sense, the Gaussian distribution can also be considered to be a first-order approximation of the Weibull distribution.

Therefore, it becomes an unavoidable problem to determine the three parameters of Weibull distribution through sample data. Zhentong Gao pointed out that the graph method using Weibull graph paper can be used to determine these three parameters; and the analytical method of solving three simultaneous transcendental equations can also be used [3]. Obviously, the graph method is inconvenient to use, and the accuracy is difficult to control,

and the analytical method also has the problem of self-conflict, that is, the obtained safe life will be greater than the minimum life in the sample [2]. In order to solve this problem, the author proposes the Gao Zhentong method, which uses the characteristics of Python to easily determine the estimation of the three parameters of the Weibull distribution. [2].

Although Gao Zhentong method solves the problem of self-conflict, it does not prove that the result obtained from this method can better fit the population itself. Therefore, I want to use the digital experiment to compare the advantages and disadvantages of each method through the effect of fitting the population by the analytical method, the Gao Zhentong method and the Gaussian distribution.

The criterion for judging what kind of statistical distribution the population of the sample data conforms to should be the coefficient of determination of the fit, the bigger the better [4]. Looks very reasonable, there are still some flaws. In fact, there will still be the so-called "overfitting" problem [5, 6]. The main reason is that the number of samples is relatively small, so that there are more fitting curves with a relatively large coefficient of determination, that is, there is a relatively large chance of mixing fish and dragons. To avoid this phenomenon, it is necessary to make the sample size relatively large. The question is how to quantify this "big"? This requires digital experimental techniques. After a brief introduction to digital experiments, this paper mainly discusses various possibilities in the application of the three-parameter Weibull distribution, and draws some valuable conclusions, which are undoubtedly beneficial to the practical application of the three-parameter Weibull distribution. In addition, through the bootstrap method [7, 10], we experience the effect of digital experiments in the process of determining the three parameters and their confidence intervals of Weibull distribution, and obtain some constructive results.

2. Brief Description About Digital Experiment

As we all know, the essence of modern electronic computers is to use physical components, such as electronic tubes at first, then semiconductor transistors, integrated circuits and now chips, to simulate people's computing process. Although it is a digital computer, it is essentially a kind of human computing and the simulation of various human thinking activities. In fact, the earliest computers were also simulators, but now the so-called simulators are almost extinct. The reason is that with the development of computer software and hardware, the current simulation is already a digital simulation. However, the digital experiment proposed in this paper is not digital simulation in the general sense, but refers to the use of digital to do "experiments". There are two main purposes of the so-called experiments. One is to test whether the theoretical hypothesis is correct. For example,

the famous Galileo free-fall experiment; the second is to provide reliable material for new discoveries; such as the discovery of Roentgen rays. Inspired by "machine learning" [11-14], this paper will do a digital experiment for each parameter of the three-parameter Weibull distribution: use a computer to make a random generator of Weibull distribution, which resulting data are used to test the fitting advantages and disadvantages of Gao Zhentong method, the analytical method and Gaussian distribution. On this basis, the bootstrap method, which is essentially a digital experimental method, is further introduced to simultaneously determine the three parameters and their confidence interval of Weibull distribution.

3. A Random Generator of Weibull Distribution

In order to perform a digital experiment on the three-parameter Weibull distribution, a basic condition is to generate a variety of sample data at will, that is, to construct a random "generator" of the three-parameter Weibull distribution. There are no three-parameter Weibull distribution random generators in general software libraries, for the simple reason that it is "natural" for the average user to treat the position parameter x_0 as zero, because from a mathematical point of view It seems that it only reflects the translation of one coordinate axis and does not seem to affect other properties. However, as pointed out in the introduction, the position parameter in the Weibull distribution is called "safe life" in fatigue statistics, which is generally not zero, so it is unrealistic to assume zero, and it is also not advisable. It cannot and should not be set to zero for serious study of the Weibull distribution. According to this principle, the "randomly generated three-parameter Weibull distribution generator" is composed of the following code in Python:

```
def randomWN(NN, b, λ, x0, G):
    np.random.seed(G)
    NR=np.random.rand(NN)
    N=x0+λ*np.exp(np.log(np.log(1/NR))/b)
    N.sort()
    return N
```

Among them, NN is the capacity (size) of the sample to be generated, b , λ , x_0 are the shape parameters, scale parameters and position parameters (in the study of fatigue life, it is generally expressed by N_0 , which is called the safety life parameter) of the Weibull distribution respectively. And G represents the "seed" for generating random numbers, so that the generated samples can be repeated, which is very important for digital experiments. Section 4 is to change the samples of different Weibull distributions by changing different parameters to test the pros and cons of Gao Zhentong method, the analytical method and Gaussian distribution fitting, and the confidence interval of each parameter of the Weibull distribution is determined by Gao Zhentong method [4, 16, 17].

4. Digital Experiments on Variation of Parameters of Weibull Distribution

Like physical experiments, only one parameter is changed and other parameters remain unchanged, so that the function of this parameter can be found. The following experiments

have a confidence level of 95%.

4.1. Experiment 1: Change Only the Position Parameter x_0

For the convenience of fixing first, $NN=50$, $b=2.5$, $\lambda=5$, $G=1$, and let x_0 take 0, 1, and 5 respectively, and the following results can be obtained by using Python code:

Table 1. Comparison of three different fits with a sample of Weibull distributed random numbers changing only the location parameter x_0 .

	b	x_0	λ	r	R^2
$x_0=0$					
GZT Method	2.19	0.00	5.37	0.99482	0.99259
Confidence Interval	(1.94, 2.22)	(0.00, 0.59)	(4.51, 5.55)	---	---
Analysis Methods	3.53	-2.30	7.81	0.98528	0.99247
Gaussian Distribution	---	---	---	---	0.99194
$x_0=1$					
GZT Method	2.25	0.91	5.46	0.99489	0.99293
Confidence Interval	(2.22, 2.56)	(0.00, 1.59)	(5.61, 5.55)	---	---
Analysis Methods	3.53	-1.30	7.81	0.98528	0.99247
Gaussian Distribution	---	---	---	---	0.99194
$x_0=5$					
GZT Method	2.25	4.91	5.46	0.99489	0.99293
Confidence Interval	(2.22, 2.64)	(3.86, 5.59)	(5.55, 5.76)	---	---
Analysis Methods	3.53	2.70	7.81	0.98528	0.99247
Gaussian Distribution	---	---	---	---	0.99194

Note: GZT is the abbreviation of Gao Zhentong; r is the correlation coefficient of fitting the sample data in logarithmic coordinates; R^2 represents the coefficient of determination for the ideal reliability fit [4]. Subsequent tables are the same, so the same notes are not given.

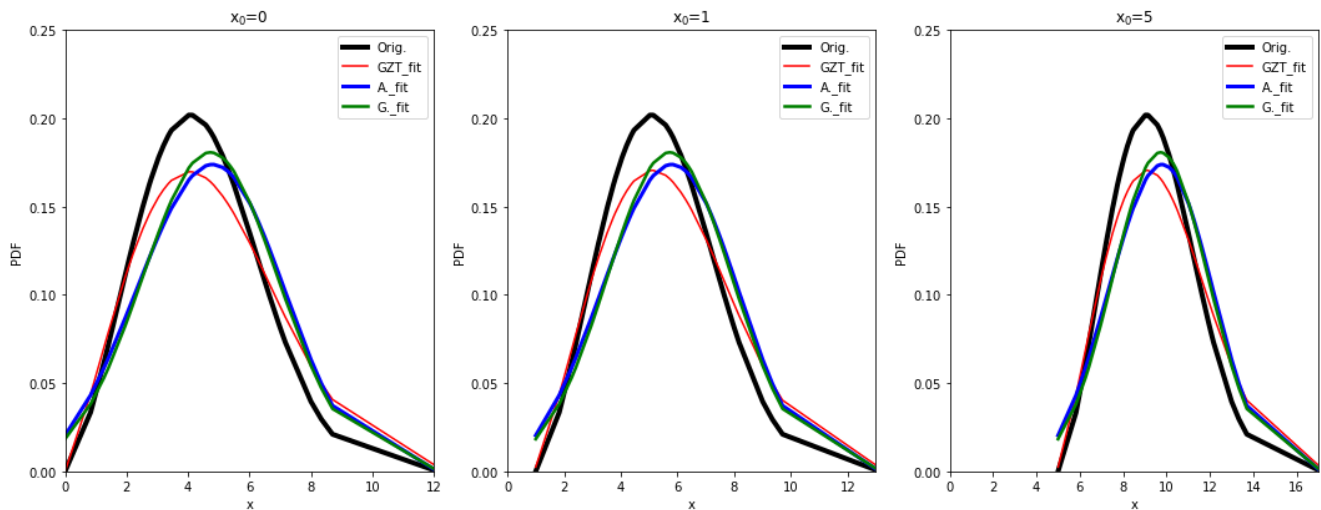


Figure 1. Comparison plot of three different fits when only changing the location parameter x_0 when taking a random sample from Weibull distribution.

Note: orig. in the legend is the abbreviation of original, which means sample data; GZT_fit, which means the fitting curve of the PDF of Weibull distribution obtained by the three parameters of Weibull distribution obtained by Gao Zhentong method; A_fit, which means the fitting curve of the PDF of the Weibull distribution obtained by the Weibull distribution parameters obtained by the analytical method; G_fit, the curve of fitting the sample data with the PDF obtained by the Gaussian distribution. Subsequent figures are the same, so the same notes are not given.

It can be seen from the sample data generated by the Weibull distribution random number generator that their standard deviation has nothing to do with the position parameters, and the mean and median only reflect the "translation" of the coordinates. Therefore, the sample data will not affect the Gaussian distribution and the shape of the Weibull distribution

calculated by the Gao Zhentong method and the analytical method. For example, for different x_0 , the shape parameters and scale parameters obtained by the analytical method are exactly the same, while the position parameters are only "translated"; and when Gao Zhentong method is $x_0=0$, due to its own characteristics, the shape parameters and scale parameters are

somewhat different from $x_0=1$ and 5; but when $x_0=1$ and 5, b and λ remains the same. In this sense, the correlation coefficient and the coefficient of determination obtained by these three methods are independent of changes in x_0 .

The coefficient of determination obtained by the Gao Zhentong method is the largest among the three methods, and the estimated values of the three parameters obtained are also the closest to the true values.

4.2. Experiment 2: Change Only the Random Seed G

G represents the "seed" of random numbers. The seed changes, which means that the sample has changed, and other conditions are unchanged, that is, fixed at this time, $NN=50$, $b=2.5$, $\lambda=5$, $x_0=1$, And changing the case of $G=0$, 1 and 5, the result is:

Table 2. Comparison of three different fits when only random seed (G) is changed when taking random samples from Weibull distribution.

	b	x_0	λ	r	R^2
$G=0$					
GZT Method	2.27	1.27	4.49	0.99288	0.98318
Confidence Interval	(2.13, 2.66)	(0.41, 1.90)	(4.50, 4.73)	---	---
Analysis Methods	1.86	1.98	3.65	0.96465	0.98657
Gaussian Distribution	---	---	---	---	0.97720
$G=1$					
GZT Method	2.25	0.91	5.46	0.99489	0.99293
Confidence Interval	(2.22, 2.56)	(0.00, 1.59)	(5.61, 5.55)	---	---
Analysis Methods	3.53	-1.30	7.81	0.98528	0.99247
Gaussian Distribution	---	---	---	---	0.99194
$G=5$					
GZT Method	2.18	0.90	5.33	0.99379	0.98831
Confidence Interval	(1.92, 2.50)	(0.00, 1.73)	(5.34, 5.47)	---	---
Analysis Methods	2.35	0.77	5.44	0.99370	0.99063
Gaussian Distribution	---	---	---	---	0.99171

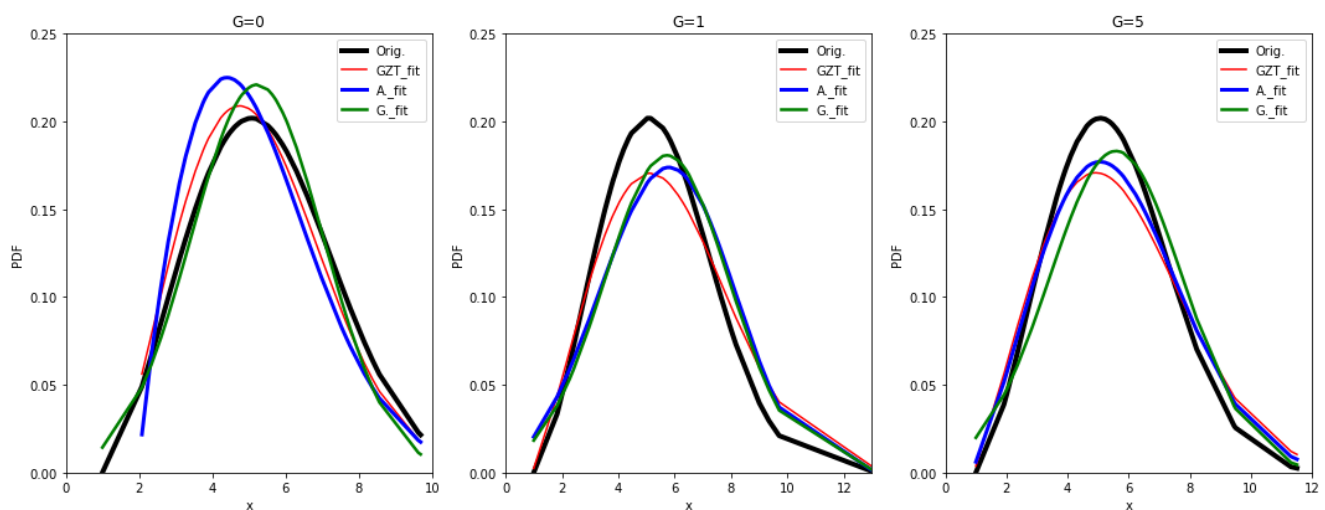


Figure 2. Comparison plot of three different fits that only change the random seed (G) when taking a random sample from Weibull distribution.

From the data of $G=0$, it can be seen that the Gao Zhentong method is still better than the Gaussian distribution, although the coefficient of determination of the analytical method is still a little larger than that of the Gao Zhentong method. However, the correlation coefficient obtained by the Gao Zhentong method is much larger than that obtained by the analytical method; as for the case of $G=5$, it seems that the results of the analytical method are better than those obtained by the Gao Zhentong method, but the coefficient of determination of the Gaussian distribution is higher than that of the Gao Zhentong method. This indicates that the sample data is relatively symmetrical, so the shape parameters obtained by the analytical method are better, but in order to obtain a more reliable safety life (position parameter), it is

better to use the Gao Zhentong method, and the other two parameters are also not too bad.

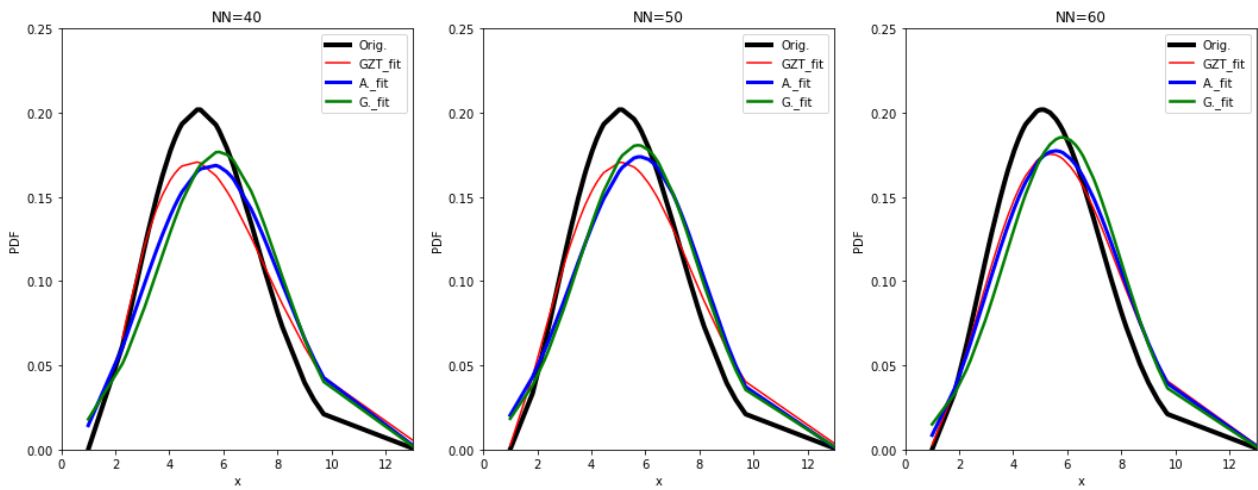
Also note that the confidence interval for λ is unsatisfactory, not including the estimated value but also the true value. This shows that the λ confidence interval obtained by Gao Zhentong method is still problematic [4], which is also the problem to be solved by the bootstrap method to be introduced in section 5.

4.3. Experiment 3: Change Only the Sample Size NN

According to general practice, the capacity of 50 can be regarded as a "large sample". Now fix $b=2.5$, $\lambda=5$, $x_0=1$, $G=1$ and change $NN=40$, 50 and 60 the result is,

Table 3. Comparison of three different fits that vary only in sample size (NN) when taking random samples from Weibull distribution.

	b	x_0	λ	r	R²
NN=40					
GZT Method	1.94	1.50	4.92	0.99365	0.99308
Confidence Interval	(1.91, 2.25)	(0.64, 2.09)	(4.93, 5.27)	---	---
Analysis Methods	2.77	0.04	6.49	0.98536	0.99159
Gaussian Distribution	---	---	---	---	0.98832
NN=50					
GZT Method	2.25	0.91	5.46	0.99489	0.99293
Confidence Interval	(2.22, 2.56)	(0.00, 1.59)	(5.61, 5.55)	---	---
Analysis Methods	3.53	-1.30	7.81	0.98528	0.99247
Gaussian Distribution	---	---	---	---	0.99194
NN=60					
GZT Method	2.44	0.83	5.65	0.99575	0.99604
Confidence Interval	(2.37, 2.7)	(0.00, 1.50)	(5.58, 5.80)	---	---
Analysis Methods	2.74	0.37	6.13	0.99478	0.99595
Gaussian Distribution	---	---	---	---	0.99375

**Figure 3.** Comparison plot of three different fits that change only the sample size (NN) when taking a random sample from the Weibull distribution.

It can be seen from Table 3 that both the correlation coefficient and the coefficient of determination by the Gao Zhentong method are better than the analytical method at this time. As the sample size increases, the results of the Gao Zhentong method are closer to the parameters of the sample, that is, the effect is better. And the estimated λ also falls

within the confidence interval.

4.4. Experiment 4: Change Only the Shape Parameters b

Also fix $NN=50$, $\lambda=5$, $x_0=1$, $G=1$ and change b to 1.5, 2.5, 3.5 and 4.5. The result is:

Table 4. Comparison of three different fits that vary only in sample shape paramant b when taking random samples from Weibull distribution.

	b	x_0	λ	r	R²
b=1.5					
GZT Method	1.34	0.96	5.68	0.99495	0.99295
Confidence Interval	(1.39, 1.48)	(0.61, 1.19)	(4.75, 6.85)	---	---
Analysis Methods	1.87	-0.95	7.92	0.96986	0.98539
Gaussian Distribution	---	---	---	---	0.97101
b=3.5					
GZT Method	3.16	0.88	5.39	0.99487	0.99288
Confidence Interval	(2.79, 3.47)	(0.00, 1.99)	(4.80, 5.69)	---	---
Analysis Methods	5.95	-2.52	8.85	0.98770	0.99216
Gaussian Distribution	---	---	---	---	0.99190
b=4.5					
GZT Method	4.09	0.85	5.35	0.99485	0.99285
Confidence Interval	(3.25, 4.40)	(0.00, 2.34)	(4.26, 5.75)	---	---
Analysis Methods	9.80	-4.68	10.93	0.98829	0.99127
Gaussian Distribution	---	---	---	---	0.99025

Note: The sub-table of $b=2.5$ is the same as the second sub-table ($x_0=1$) of Table 1, and will not be given here.

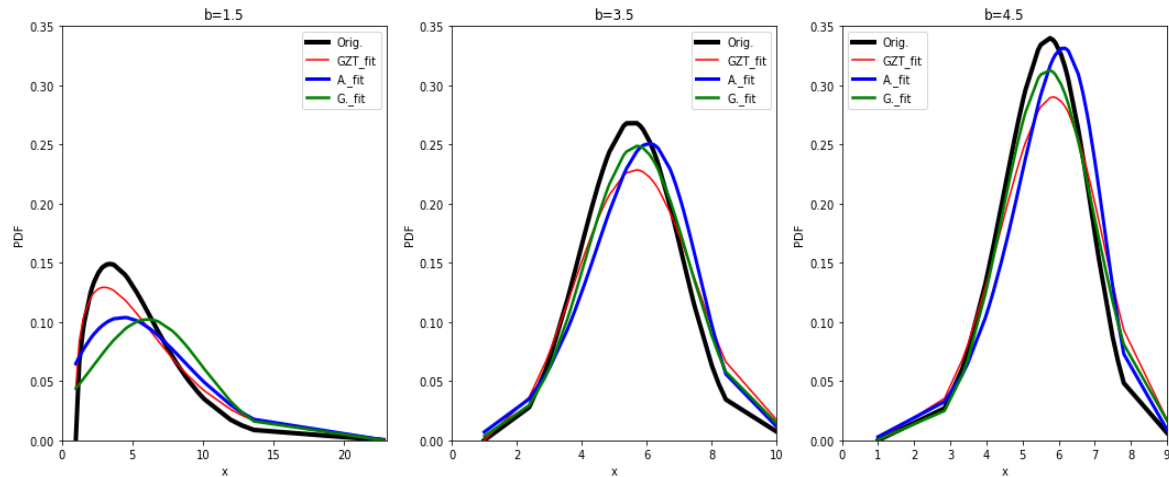


Figure 4. Comparison plot of three different fits that change only the sample shape parameter b when taking a random sample from the Weibull distribution.

It is not difficult to see from Table 4 that both the correlation coefficient and the coefficient of determination of the Gao Zhentong method are larger than those obtained by the analytical method, and the estimated parameters obtained are closer to the parameters of the sample data than those obtained by the analytical method. So Gao Zhentong method also is better.

As in the case of Experiment 1, that is, the coefficient of determination obtained by the Gao Zhentong method is the

largest among the three methods, and the estimated values of the three parameters obtained are also the closest to the true value.

4.5. Experiment 5: Change Only the Scale Parameter λ

Also fixed $NN=50$, $b=2.5$, $x_0=1$, $G=1$. $\lambda=5$ and change: 2, 5 and 10, the result is,

Table 5. Comparison of three different fits that vary only in sample scale parameter λ when taking random samples from Weibull distribution.

	b	x_0	λ	r	R^2
$\lambda=2$					
GZT Method	2.26	0.96	2.19	0.99489	0.99296
Confidence Interval	(2.20, 2.63)	(0.55, 1.24)	(2.30, 2.22)	---	---
Analysis Methods	3.53	0.08	3.12	0.98528	0.99247
Gaussian Distribution	---	---	---	---	0.99194
$\lambda=5$					
GZT Method	2.25	0.91	5.46	0.99489	0.99293
Confidence Interval	(2.22, 2.56)	(0.00, 1.59)	(5.61, 5.55)	---	---
Analysis Methods	3.53	-1.30	7.81	0.98528	0.99247
Gaussian Distribution	---	---	---	---	0.99194
$\lambda=10$					
GZT Method	2.25	0.82	10.93	0.99489	0.99293
Confidence Interval	(2.22, 2.26)	(0.00, 2.18)	(10.13, 11.09)	---	---
Analysis Methods	3.53	-3.61	15.61	0.98528	0.99247
Gaussian Distribution	---	---	---	---	0.99194

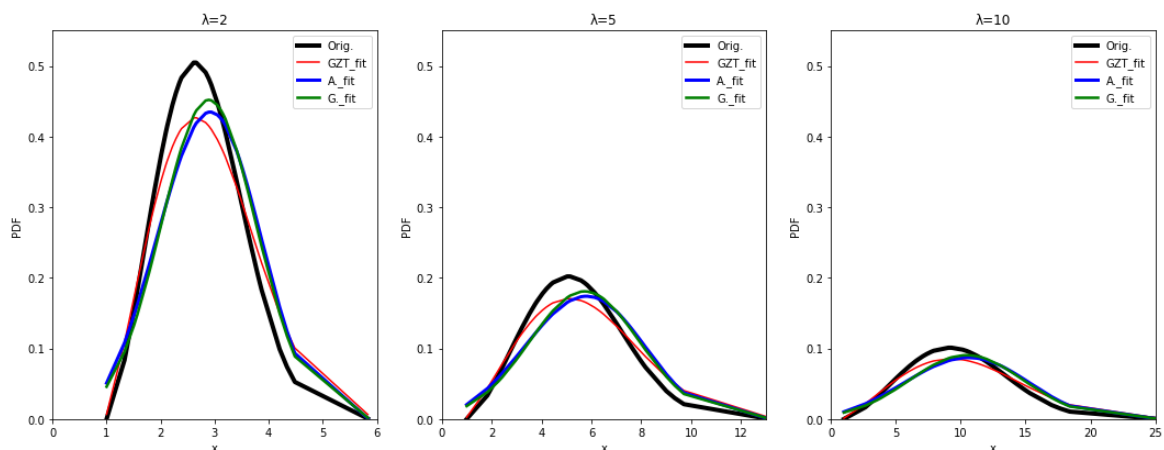


Figure 5. Comparison plot of three different fits that change only the sample scale parameter λ when taking a random sample from the Weibull distribution.

It is not difficult to see that after the scale parameter changes, the shape parameter, the correlation coefficient and the coefficient of determination have not changed. In this sense, the scale coefficient does only affect the proportion of the Weibull distribution. At the same time, it can be seen that the parameters obtained by Gao Zhentong method are closer to the original parameters of the sample than those obtained by the analytical method, and the correlation coefficient and the coefficient of determination are larger than those obtained by the analytical method. As in the case of Experiment 1, that is, the coefficient of determination obtained by the Gao Zhentong method is the largest among the three methods, and the estimated values of the three parameters obtained are also the closest to the true value.

5. Estimating Three Parameters and Their Confidence Intervals of Weibull Distribution by Bootstrap

The Bootstrap was researched and developed by Efron in 1997 [7]. The basic idea is as follows. Assuming that there are n observations to obtain the regression parameter b^{\wedge} , how to estimate the accuracy of this regression parameter? The bootstrap method randomly selects n samples (called bootstrap samples) in these n samples with replacement. Obviously, there are repeated selections in the selected samples, and there are also unselected observations, but still regression parameter estimates can be obtained using the

original procedure. Such bootstrap calculation can be performed m times, and m bootstrap-estimated regression parameters b^{\wedge} can be obtained. Their mean value $b^{*\wedge}$ can be used as the estimation of parameter b , and its standard deviation is recorded as $s(b^{*\wedge})$, which is called is the bootstrap standard deviation and is an estimate of the standard deviation of the sampling distribution of b^{\wedge} , which gives the precision of the regression parameter b . This so-called bootstrap method is essentially a digital experiment that can be extended to determine the three parameters and their intervals of Weibull distribution, although these parameters are not really "regression parameters".

Now we mainly do digital experiments on small samples and large samples respectively for the sampling times m ; because the digital experiments in the previous section have proved that the Gao Zhentong method is a better method for determining the three parameters of Weibull distribution from the current situation, so future digital experiments will use the Gao Zhentong method. m is taken as 500, 1000, and 1500 respectively. In order to ensure the repeatability of the experiment, the random seed is taken as 0 here. The confidence level is still taken as 95%.

5.1. Experiment 6: Using the Fatigue Life in Appendix I as the Small Sample Data

After modifying the Python code, Table 6 can be obtained. Replace x_0 with N_0 in Table 6, this is because N_0 is safe life in fatigue statistics, highlighting its physical significance.

Table 6. Comparing the results of the three-parameter Weibull distribution with the small sample by the bootstrap method and the non-bootstrap method.

	b	N_0	λ	r	R^2
GZT Method	2.04	277.00	320.55	0.99922	0.99823
Confidence Interval	(1.99, 2.80)	(159.0, 333.0)	(320.6, 343.3)	---	---
Bootstrap (500 times)	2.37	255.68	340.58	0.98898	0.99702
Confidence Interval	(0.28, 4.47)	(79.0, 432.4)	(128.4, 552.8)	---	---
Bootstrap (1000 times)	2.37	254.02	342.67	0.98927	0.99728
Confidence Interval	(0.30, 4.44)	(72.5, 435.6)	(124.2, 561.2)	---	---
Bootstrap (1500 times)	2.36	255.27	341.41	0.96666	0.99726
Confidence Interval	(0.28, 4.44)	(72.7, 437.9)	(122.9, 559.9)	---	---

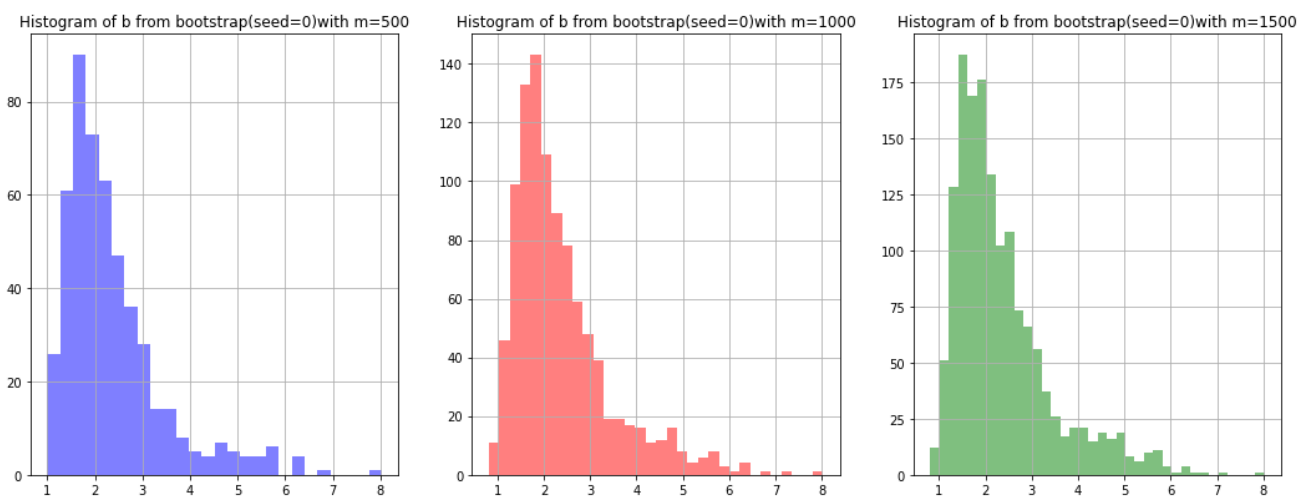


Figure 6. Histogram of shape parameter b of small sample by the bootstrap method.

It can be seen from Table 6 that in the case of small samples, the results obtained by the non-bootstrap method (i.e. Gao Zhentong method) are better than the bootstrap method except for the scale confidence interval. The advantage of the bootstrap method is that it simultaneously gives confidence intervals for all three parameters of the Weibull distribution, albeit quite large. It can also be seen from the histogram that it is not a strict Gaussian distribution, so the standard deviation is still relatively large. Of course, this is related to the small

sample size. At the same time, it would not certainly be seen that the more m the better, the fact is that the effect of $m=1000$ is better than that of $m=1500$.

5.2. Experiment 7: Using the Fatigue Life in Appendix II as a Large Sample

Modify the Python code to obtain Table 7. N_0 in Table 7 has the same meaning as N_0 in Table 6.

Table 7. Comparing the results of the three-parameter Weibull distribution with the large sample by the bootstrap method and the non-bootstrap method.

	b	N_0	λ	r	R^2
GZT Method	2.147	2.780	2.867	0.99376	0.98903
Confidence Interval	(2.2, 2.27)	(2.5, 3.0)	(2.8, 2.9)	---	---
Bootstrap (500 times)	2.160	2.801	2.844	0.97162	0.98996
Confidence Interval	(1.59, 2.73)	(2.4, 3.2)	(2.4, 3.3)	---	---
Bootstrap (1000 times)	2.159	2.795	2.849	0.97939	0.98982
Confidence Interval	(1.59, 2.73)	(2.4, 3.2)	(2.4, 3.3)	---	---
Bootstrap (1500 times)	2.158	2.795	2.852	0.98735	0.98970
Confidence Interval	(1.60, 2.72)	(2.4, 3.2)	(2.4, 3.3)	---	---

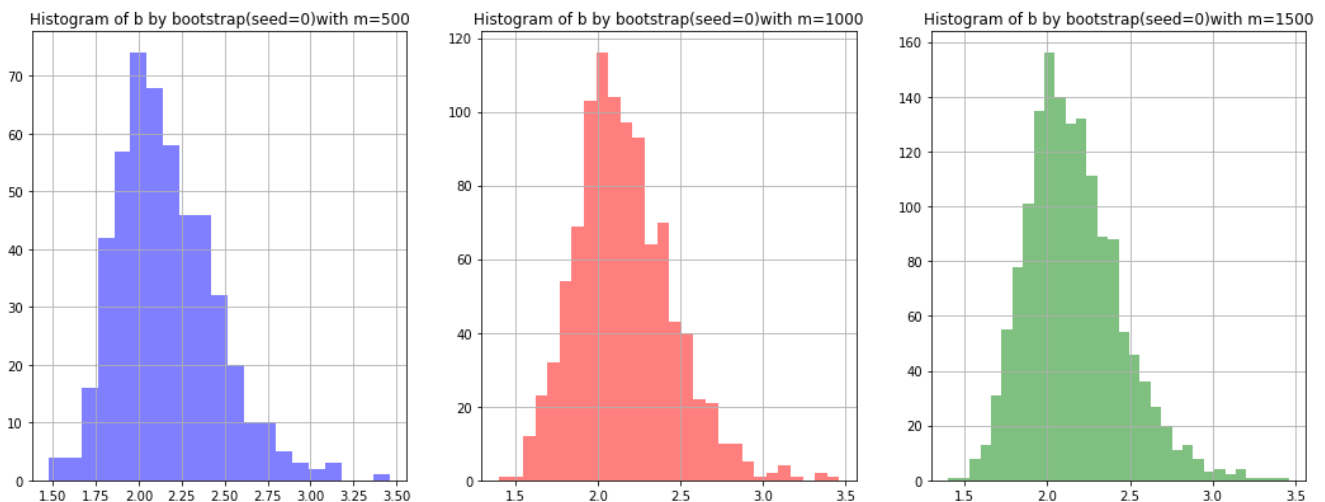


Figure 7. Histogram of shape parameter b of large sample by the bootstrap method.

It can be seen from Table 7 that for the so-called large sample, the bootstrap method has little to do with the number of extractions. The values of the three parameters of the Weibull distribution and the coefficient of determination obtained by the three different times of the bootstrap method are almost indistinguishable! In other words, 500 draws are good enough. Moreover, the estimated values of the three parameters obtained are quite close to those obtained by Gao Zhentong method, and the coefficient of determination is a little better than that of the Gao Zhentong method, but the confidence interval obtained is a little worse than that obtained by Gao Zhentong method.

This in turn proves that the Gao Zhentong method is indeed a better method for estimating the three parameters of the Weibull distribution and the corresponding confidence intervals. This also shows that the problem of Gao Zhentong method in the confidence interval of the previous scale parameters is related to the fact that the sample is not large

enough. This is consistent with the conclusion of experiment 3.

It is natural to continue with digital experiments, such as changing the random seed, or following the normal practice of machine learning [6], a sample set is often divided into 7:3 parts, the former is used as the training set, and the latter is used as the test set. Then, the training set is extracted by the bootstrap, and the parameters obtained from this are then tested on the test set. But I feel that basically the application of digital experiments to the Weibull distribution has been made clear, and that's it.

6. Conclusion

From the above, the following conclusions can be drawn:

The digital experiments on Weibull distribution show that the results of the Gao Zhentong method are better than the analytical method in most cases, and the reason is very simple. Because Gao Zhentong method starts with the largest

correlation coefficient, and the analytical method is only guaranteed to be consistent with the mean, standard deviation and median of the sample data, but this is precisely the conflict situation where the estimated safe life may be greater than the minimum value in the sample data [2].

Once the shape coefficient is changed, the influence on the safe life and the scale coefficient is relatively large. In this sense, the shape coefficient has the greatest influence on the Weibull distribution. This conclusion is also natural, since the shape coefficient plays a decisive role in the shape of the Weibull distribution. It can be seen from the previous digital experiments that although the three parameters of the Weibull distribution are different, their correlation coefficients in the logarithmic case of fitting the ideal reliability can be very close, so at this time, the coefficient of determination in the non-logarithmic case is particularly important.

The influence of the sample size on the correct estimation

of the population parameters of the sample is relatively large. The digital experiments show that a sample with a capacity of 50 can basically reflect the population parameters better. When 60 is taken and the confidence level is 0.95, the confidence interval of the estimated value of the parameter can basically contain the estimated value.

The number of times m drawn by the bootstrap method is not as large as possible, but how much is appropriate depends on the digital experiment.

The digital experiments in this paper show that the Gao Zhentong method is a better method for estimating the three parameters and their confidence intervals of the Weibull distribution.

Acknowledgements

I would like to thank Mr. Weihao Wan for his unconditional support for the completion of this thesis.

Appendix

I. The following fatigue life data are from Table 8-4 of Reference [3].

Fatigue life data (10^3 cycles)									
350	380	400	430	450	470	480	500	520	540
550	570	600	610	630	650	670	730	770	840

II. The following fatigue life data are from Table 12-3 of Reference [3].

Fatigue life data (10^5 cycles)																			
3.08	3.26	3.32	3.48	3.49	3.56	3.69	3.7	3.78	3.79	3.8	3.87	3.95	4.07	4.08	4.1	4.12	4.2	4.24	4.25
4.28	4.31	4.31	4.36	4.54	4.58	4.6	4.62	4.63	4.65	4.67	4.67	4.72	4.73	4.75	4.77	4.8	4.82	4.84	4.9
4.92	4.93	4.95	4.96	4.98	4.99	5.02	5.03	5.06	5.08	5.06	5.1	5.12	5.15	5.18	5.2	5.22	5.38	5.41	5.46
5.47	5.53	5.56	5.6	5.61	5.63	5.64	5.65	5.68	5.69	5.73	5.82	5.86	5.91	5.94	5.95	5.99	6.04	6.08	6.13
6.16	6.19	6.21	6.26	6.32	6.33	6.36	6.41	6.46	6.81	7	7.35	7.82	7.88	7.96	8.31	8.45	8.47	8.79	9.87

References

[1] Wikipedia: Weibull Distribution, URL: https://en.wikipedia.org/wiki/Weibull_distribution

[2] Jiajin Xu. Gao Zhentong Method in the Fatigue Statistics Intelligence, Journal of Beijing University of Aeronautics and Astronautics. 47 (10): 2024-2033. DOI: 10.13700/j.bh.1001-5965. 2020. 0373 (in Chinese).

[3] Zhentong Gao. Fatigue applied statistics [M]. Beijing: National Defense Industry Press, 1986: 87, 136, 253 (in Chinese).

[4] Jiajin Xu. (2021) Further research on fatigue statistics intelligence, Acta Aeronautica et Astronautica Sinica has been accepted and first published online on 2020-12-04, which URL: <http://hkxb.buaa.edu.cn/CN/html/202203xx.html>. doi: 10.7527/S1000-6893.2021.25138 (in Chinese).

[5] Montgomery D. C., Peck E. A., Vining G. G. Introduction to Linear Regression Analysis [M]. Beijing: Machinery Industry Press, 2019: 61, 386 (in Chinese).

[6] Bowles M. (2017). Machine Learning in Python: Essential Techniques for Predictive Analysis [M], People Post Press: 101-103 (in Chinese).

[7] Efron B., Hastie T. (2019) Computer Age Statistical Inference [M]. Beijing: Machinery Industry Press: 110-112 (in Chinese).

[8] Hogg R. V., McKean J. W., Craig A. T. (2015) Introduction to Mathematical Statistics [M]. 7th, Beijing: Machinery Industry Press: 206-209 (in Chinese).

[9] XinTao Xian, YouZhi Xu., ect. (2015) Assessment of optimum confidence interval of reliability with three—parameter Weibull distribution using bootstrap weighted—norm method. Journal of Aerospace Power. 28 (3): 481-488 (in Chinese).

[10] Mooney C. Z., Duval R. D. (2017). Bootstrap. [M]. Shanghai People Press (in Chinese).

[11] Mitchell T. M. (2014) Machine Learning [M]. Beijing: Machinery Industry Press (in Chinese).

[12] Shalev S. S., David S. B. (2017) Understanding Machine Learning [M]. Beijing: Machinery Industry Press (in Chinese).

[13] Chollet F. (2018) Deep Learning with Python [M]. People Post Press (in Chinese).

[14] Haslwanter T. (2018) An Application to Statistics with Python [M]. People Post Press (in Chinese).

- [15] McClure N. (2018) Tensor Flow Machine Learning Cookbook [M]. Beijing: Machinery Industry Press (in Chinese).
- [16] DaoLi Chen. (1997) A new method to estimate the confidence limit of Weibull distribution parameters, reliable life and reliability [J]. Mechanical Design, No. 10: 18-20 (in Chinese).
- [17] HuiMing Fu, ZhenTong Gao, RiPing Xu. (1992) Confidence limits of the three-parameter Weibull distribution [J]. Acta Aeronautica et Astronautica Sinica. 13 (3): A153-A163 (in Chinese).