DOI: 10.2195/lj_Proc_fu_en_201210_01
URN: urn:nbn:de:0009-14-34525

Integration of Truncated Distributions in Exponential Family with Simulation Models of Logistics and Supply Chain Management

Integration von gekürzten Verteilungen der exponentialen Familie mit Simulationsmodellen in Logistik und Supply Chain Management

Haichen Fu
Bernd Noche

Chair for Transport Systems and Logistics
Faculty for Engineering
University Duisburg-Essen

Truncated distributions of the exponential family have great influence in the simulation models. This paper discusses the truncated Weibull distribution specifically. The truncation of the distribution is achieved by the Maximum Likelihood Estimation method or combined with the expectation and variance expressions. After the fitting of distribution, the goodness-of-fit tests (the Chi-Square test and the Kolmogorov-Smirnov test) are executed to rule out the rejected hypotheses. Finally the distributions are integrated in various simulation models, e. g. shipment consolidation model, to compare the influence of truncated and original versions of Weibull distribution on the model.

[Keywords: Truncated Weibull distribution, Supply Chain Management, shipment consolidation policy]

Gekürzte Verteilungen der exponentialen Familie haben großen Einfluss auf die Simulationsmodelle. Dieser Beitrag konzentriert sich auf die gekürzte Weibull-Verteilung. Die Verkürzung der Verteilung wird durch die Maximum-Likelihood-Estimation-Methode oder eine Kombination mit Erwartungswert und Varianz erreicht. Danach werden die Anpassungstests (z. B. Chi-Quadrat-Test und Kolmogorov-Smirnov-Test) durchgeführt, um die falschen Hypothesen zu beseitigen. Weiterhin wird die gekürzte Weibull-Verteilung in einem Lieferungs-Konsolidation-Modell integriert, um den Einfluss von gekürzten und originalen Verteilungen zu vergleichen.

[Schlüsselwörter: Gekürzte Weibull-Verteilung, Supply Chain Management, Lieferungs-Konsolidierungs-Strategie]

1 Introduction and Overview

Simulation and modeling is a popular topic in many industrial fields. The source component of the simulation model comes from the distribution model which is induced from the empirical data. The majority of the important distributions used in the simulation come from the exponential family. Three members of the exponential family are the normal distribution, gamma distribution and Weibull distribution. These distributions are used in many simulation models to serve as the reflection of the real world. However, the truncated versions of these distributions are utilized less in practice. This paper discusses the truncation of Weibull distribution in simulation models.

First of all, the importance of the truncation should be discussed for the necessity of this research. There are multiple reasons for the truncation of distributions, especially in the simulation of the supply chains or the production systems. One most commonly seen reason is to discard the unreliable data from the sample pool. Douglas J. Depriest discussed the singly truncated normal distribution in the analysis of satellite data [DD83]. The infrared sensor from the satellite could have extremely distorted data reading because of the cloud in the view. So the sample data that are extracted from the data pool are contaminated by these falsely read data. In order to get a more accurate simulation input, a truncation point was set to rule out all the unreliable data. The truncation served this purpose and also maintained the properties of a distribution. Another reason to apply the truncation of distributions in the simulation model is that the truncation would reflect the real world in a better way than the original distributions. An example for this scenario would be a simulation of the breakdown times in a production system. A simple two-server system, which is composed by a source, two servers, and a sink, is simulated using a system with all the empirical data for each component

DOI: 10.2195/lj_Proc_fu_en_201210_01
URN: urn:nbn:de:0009-14-34525

provided. The servers are working under about 90% workload utility and they suffer from random breakdowns. For a simulation of the breakdowns, two sets of data are required, namely, the duration of the breakdowns and the interval between the breakdowns. The duration of the breakdowns is one important aspect of the model and could have great influence on the outcome of the simulation. For example, the empirical data collected show that the duration of breakdowns obeys a Weibull distribution, which would then be implemented to the simulation model as the duration of the breakdowns. Like any distributions, the data that are generated from this distribution would cover the whole possible range that distribution is defined on. This would cause some extreme values as the breakdown duration to be generated, which can have a great influence on the simulation result.

In the real industrial scenario, the breakdown of the machines is a devastating factor of the production process. Therefore, any extreme values that are generated for the duration of the breakdown should be considered as unreliable data because such long breakdown times would not happen in real industrial scenes. Having these ideas in mind, multiple approaches are made to avoid these extreme values in the simulation models. One of the most commonly used methods is to simply discard all the data that are generated beyond a certain value. This method could effectively rule out all the extreme values in a quite simple manner. However, it could also result in some problems that might affect the simulation itself. First of all, this method changes the property and integrity of a probability distribution. Another problem is that when the value generated is removed, it would influence the sequence of the seeding process at the random number generation.

Having these two disadvantages at mind, another method of dealing with this problem is used to truncate the unwanted values. Instead of removing all the values beyond a certain limit, this method changes the values that are beyond the limit to that limit value, so that the probability distribution would still keep the integrity and the random number generation process would not be messed up as well. This method seems to have solved the above mentioned problems quite well and also in a relatively simple manner. However, when it comes to the simulation process, this method would bring other problems to the modeling and the result analysis. One of the most obvious problems is that the probability at the truncation point would be abnormally high due to the truncation method. And the simulation behavior would be compromised due to the unexpected high probability at the truncation points. The drawbacks of these truncation methods call for an improved method of truncating probability functions which would restore the integrity of the probability functions and keep the shape of the probability function according to the histogram provided by the empirical data. This paper focuses on the truncation versions of the exponential family, especially the Weibull distribution. A literature review of the truncated distribution of the exponential family is discussed in the following paragraphs.

A. Clifford Cohen Jr. [AC50] worked on the estimation of the mean and variance of the normal distribution with both the singly and doubly truncated samples. Cohen used the maximum likelihood estimation and the standard table to estimate the parameters of the truncated distribution. He also discussed the situations where the truncation point or the number of unmeasured observations in each “tail”. Following his work, Douglas J. Depriest discussed the truncated normal distribution in the analysis of the satellite data in his paper [DD83]. The truncated distribution is calculated from a set of raw data with the maximum likelihood estimation. After the calculation, the author examined the goodness of fit using the Kolmogorov-Smirnov test. He also gave the estimation from both parameters of a singly truncated normal distribution, which could be numerically solved when the truncation point is given. The reason that truncated normal distribution is used to estimate the radiance measurements from satellite-borne infrared sensors is to discard the unreliable samples which could lead to the inaccurate estimation. This is one common reason to use the truncated distributions in parameter estimation.

Gamma distribution is another important member of exponential family. A. Clifford Cohen Jr. discussed the method of moments for estimating the parameters of the Pearson Type III samples [CO50]. J. Arthur Greenwood and David Durand also discussed parameter estimation using the maximum likelihood estimation for gamma distribution. He also provided a tabulated solution for the general type as well as the Erlang distribution. For the computational convenience, polynomial and rational approximations are also given in the paper [GD60]. With the aid of the works above, S. C. Choi and R. Wette [CW69] discussed two numerical methods for the parameters estimation of the gamma distribution, namely, the Newton-Raphson Method and the M.L. scoring method. Based on these works, D.V. Kliche, P.L. Smith, and R.W. Johnson [KSR08] used the maximum likelihood estimation and the L-moment estimators, which are widely used in the field of hydrology, to reduce the bias from the method of moment. They also provide the method to estimate the parameters of left truncated gamma distribution [RKS09] in the scenario where some samples are missing.

Weibull distribution, another distribution that takes on the exponential form, could be used to describe the survival and failure analysis especially in the extreme situations. Lee J. Bain and Max Engelhardt [BE80] worked on the time truncated Weibull process by estimating the parameters of the distribution and the system reliability, which is a support for the tabulated value for confidence intervals in the failure truncated process [FM76]. D. R. Wingo [DRW89] used the maximum likelihood method to estimate the parameters of left truncated Weibull distribution with the known truncation point. It should be pointed out that the inference of derivatives of incomplete gamma integrals is made possible by the work of R. J. Moore [MO82]. Robert P. McEwen and Bernard R. Parresol [MRB91] discussed the method of moments in detail to induce the moment expression of both standard Weibull distribution and the three-parameter Weibull distribution. More importantly, they gave the moment expression of the left truncated Weibull distribution, the right truncated Weibull distribution, and the doubly truncated Weibull distribution. In the following chapters, both the maximum likelihood estimation method and the method of moments are both used for the parameter inference of the truncated Weibull distribution. A simple production system is integrated with the truncated Weibull distributions to compare the effect of the truncated and the original distributions on the system. A breakdown analysis of the inner modeling mechanism is presented as well. The truncation of the distributions could also influence the shipment consolidation models. The consolidation policies differ in the total cost and each cost component. To choose the time policy or the quantity policy could be decided by the different truncation alternatives. The following chapter will discuss the truncated Weibull distribution in the exponential family.

2 THE FITTING OF WEIBULL DISTRIBUTION AND TRUNCATED WEIBULL DISTRIBUTIONS

The three-parameter Weibull distribution is:

$$f(x,a,b,c) = \frac{c}{b}\left(\frac{x-a}{b}\right)^{c-1} e^{-((x-a)/b)^c} \tag{2.1}$$

with $x \geq a, a > 0, b > 0, c > 0$

where a is the location parameter, b is the scale parameter and c is the shape parameter. The standard form of Weibull distribution is f(x,0,1,c), where it could be simply transformed to the three-parameter form by replacing x with x=a+bx' [MRB91].

The left truncated three-parameter Weibull distribution is

$$f(x,a,b,c) = \frac{c}{b}\left(\frac{x-a}{b}\right)^{c-1} e^{[(t-a)/b^c - ((x-a)/b)^c]} \tag{2.2}$$

with $x \geq t, 0 < a < t, b > 0, c > 0$.

The right truncated three-parameter Weibull distribution is

$$f(x,a,b,c) = \frac{\frac{c}{b}\left(\frac{x-a}{b}\right)^{c-1} e^{-((x-a)/b)^c}}{1 - e^{\left[-((T-a)/b)^c\right]}} \tag{2.3}$$

with $a \leq x \leq T, a > 0, b > 0, c > 0$.

The doubly truncated three-parameter Weibull distribution is

$$f_{t,T}(x,a,b,c) = \frac{\frac{c}{b}\left(\frac{x-a}{b}\right)^{c-1} e^{[(t-a)/b^c - ((x-a)/b)^c]}}{1 - e^{\left[-((T-a)/b)^c\right]}} \tag{2.4}$$

with $t \leq x \leq T, a > t, b > 0, c > 0$.

To illustrate the effect of the truncation on the distribution, a sample data pool with a size of 113 is drawn. After the fitting of distribution, the Weibull distribution is chosen to be the one which can describe the sample data properly. The Weibull probability density function is

$$f(x) = 0.1573x^{0.57} e^{[-0.1002x^{1.57}]} \tag{2.5}$$

And the Weibull cumulative density function is

$$F(x) = 1 - e^{[-0.1002x^{1.57}]} \tag{2.6}$$

Now the focus is the fitting of truncated Weibull distributions. The probability density function of left truncated Weibull distribution:

$$f(x,a,b,c) = \frac{c}{b}\left(\frac{x-a}{b}\right)^{c-1} e^{[(t-a)/b^c - ((x-a)/b)^c]} \tag{2.7}$$

For a more convenient calculation and the differentiation, the above expression is transformed into another expression and a generalized form of probability density function of left truncated Weibull distribution is solved:

$$f(x,a,b,t) = abx^{b-1} e^{[at^b - ax^b]} \tag{2.8}$$

The cumulative probability function is

$$F(x,a,b,t) = 1 - e^{[at^b - ax^b]} \tag{2.9}$$

With a sample of x, the log likelihood function of the sample is

$$L(a,b) = n\log a + n\log b + (b-1)\sum \log x_i - \sum a(x_i^{\,b} - t^b) \tag{2.10}$$

To find the maximum likelihood estimates of the parameters, the global maximum of the above LLF is differentiated into these two functions:

$$\begin{cases} \frac{\partial L}{\partial a} = n/a - \sum (x_i^{\,b} - t^b) \\ \frac{\partial L}{\partial b} = n/b + \sum \log x_i - a\sum (x_i^{\,b}\log x_i - t^b \log t) \end{cases} \tag{2.11}$$

By solving the above non-linear equation system, estimated parameters a and b could be induced. After the calculation, the sample data could be fitted in a LTWD with the truncation point chosen as t = 0.5. The LTWD probability density function with t = 0.5 is

$$f(x) = 0.1672x^{0.5338}e^{[0.0376-0.109x^{1.5338}]} . \quad (2.12)$$

And the LTWD cumulative density function with t = 0.5 is

$$F(x) = 1 - e^{[0.0376-0.109x^{1.5338}]} . \quad (2.13)$$

The right truncated Weibull distribution (RTWD) has the following probability density function:

$$f(x,a,b,T) = \frac{abx^{b-1}e^{[-ax^b]}}{1-e^{[-aT^b]}} . \quad (2.14)$$

The cumulative RTWD probability function is

$$F(x,a,b,T) = \frac{1-e^{[-ax^b]}}{1-e^{[-aT^b]}} \quad (2.15)$$

where T is the right truncation point. The log-likelihood function of the RTWD has the following form:

$$L(a,b) = n\log a + n\log b + (b-1)\sum \log x_i - \sum a x_i^b - n\log[1-\exp(-aT^b)] . \quad (2.16)$$

To find the maximum likelihood estimates of the parameters, the global maximum of the above LLF is differentiated into these two functions:

$$\begin{cases} \frac{\partial L}{\partial a} = \frac{n}{a} - \sum x_i^b - \frac{nT^b\exp(-aT^b)}{1-\exp(-aT^b)} \\ \frac{\partial L}{\partial b} = \frac{n}{b} + \sum \log x_i - a\sum x_i^b \log x_i - \\ \frac{na\log(T)T^b\exp(-aT^b)}{1-\exp(-aT^b)} \end{cases} . \quad (2.17)$$

By solving the above non-linear equation system, estimated parameters a and b could be induced. After the calculation, the sample data could be fitted in a RTWD with the truncation point chosen as T = 12. The RTWD probability density function converts to

$$f(x) = 0.1517x^{0.6517}e^{[-0.0915x^{1.6517}]} . \quad (2.18)$$

And the RTWD cumulative density function with T = 12 is

$$F(x) = \frac{1-e^{[-0.0915x^{1.6517}]}}{0.9961} . \quad (2.19)$$

The doubly truncated Weibull distribution (DTWD) has the following probability density function:

$$f(x,a,b,t,T) = \frac{abx^{b-1}e^{[at^b-ax^b]}}{1-e^{[-aT^b]}} . \quad (2.20)$$

The cumulative DTWD probability function is

$$F(x,a,b,t,T) = \frac{1-e^{[at^b-ax^b]}}{1-e^{[-aT^b]}} \quad (2.21)$$

where t and T are left and right truncation points respectively.

The log-likelihood function of the DTWD has the following form:

$$L(a,b) = n\log a + n\log b + (b-1)\sum \log x_i - \sum a(x_i^b - t^b) - n\log[1-\exp(-aT^b)] . \quad (2.22)$$

After differentiation of the above LLF with respect to a and b, the functions are:

$$\begin{cases} \frac{\partial L}{\partial a} = \frac{n}{a} - \sum (x_i^b - t^b) - \frac{nT^b\exp(-aT^b)}{1-\exp(-aT^b)} \\ \frac{\partial L}{\partial b} = \frac{n}{b} + \sum \log x_i - a\sum (x_i^b \log x_i - t^b \log t) \\ - \frac{na\log(T)T^b\exp(-aT^b)}{1-\exp(-aT^b)} \end{cases} . \quad (2.23)$$

With the truncation points $t = 0.5$ and $T = 12$, the DTWD which fits this sample is:

$$f(x) = 0.1559x^{0.6384}e^{[0.0305-0.0915x^{1.6384}]} . \quad (2.24)$$

And the DTWD cumulative density function with t = 0.5 and T = 12 is

$$F(x) = \frac{1-e^{[0.0305-0.0948x^{1.6384}]}}{0.9961} . \quad (2.25)$$

The histogram and the fitted distributions are put in the following graphs to see the difference between these alternatives. It could be observed that the shape of all the probability functions and the cumulative probability functions show some difference between each other. In the next chapters, the impact they have on the system performance when integrated in the production systems are discussed and analyzed.

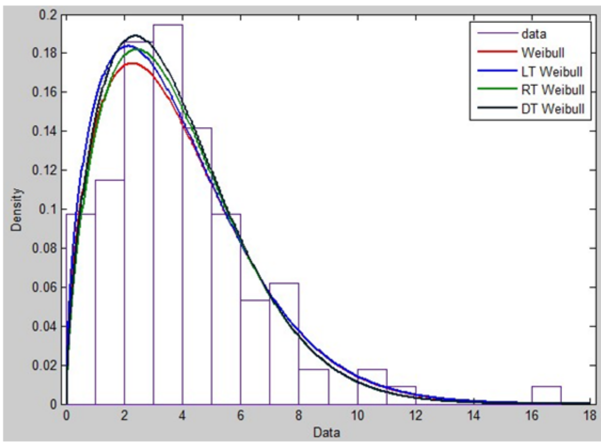

Figure 1. Histogram and probability density distributions

DOI: 10.2195/lj_Proc_fu_en_201210_01
URN: urn:nbn:de:0009-14-34525

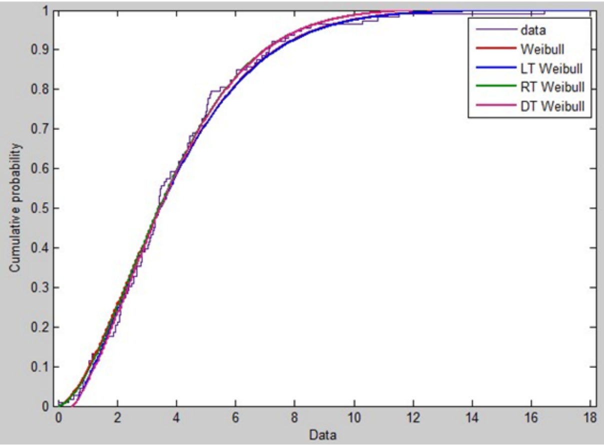

Figure 2. Histogram and cumulative probability distributions

3 TRUNCATED DISTRIBUTIONS IN PRODUCTION SYSTEMS

Although the difference in parameters is not obvious, the effect of the truncation would be shown in the simulation. Here a simple model with two servers is introduced as an example. To illustrate the effect of the different distributions on the model, two sets of comparison simulations are made with all the distributions at source and distributions at the servers. Before moving on to the numerical results section, an effective variation reduction technique adopted in the simulation should be explained briefly. Common Random Number (CRN) [AL07, p. 578] is a technique which uses exactly the same stream of random numbers when comparing alternate model configurations. Put simply, the same stream of random numbers in the system gives all the alternatives the same condition. Moreover, the same random seed is used in different random number generations of all the distributions. This guarantees that the only reason that would lead to the difference in the final result is the distribution itself.

3.1 VALIDATION WITH M/TR/1 QUEUEING SYSTEMS

The purpose of this chapter is to validate the truncated distributions with the theoretical knowledge of queueing theory. If a queueing system consists of a source with an exponentially distributed inter-arrival time, one server with a generally distributed service time, this system is denoted as an M/G/1 system. Denote the average rate of customers as λ, the average rate of service station as μ, the service rate as $\rho = \lambda / \mu$, the mean waiting time as W, and the mean number of customers in the system as L, then the following equation holds [RC81, p. 178]:

$$L = \lambda W \quad (3.1)$$

The above equation is also known as the Little's Theorem or the Little's formula.

For an M/G/1 system, the length of the system is as follows [GH98, p. 212]:

$$L = \rho + \frac{\rho^2 + \lambda^2 \sigma_s^2}{2(1-\rho)} \quad (3.2)$$

where σ_s^2 is the variance of the service time. This equation is also referred to as the Pollaczek - Khintchine (PK) formula. With the above formula the expected waiting time in the queue could also be calculated [HT91, p. 8]:

$$E[T] = E[L] / \lambda \, . \quad (3.3)$$

For an M/Tr/1 system, the specifics are listed as follows [RC81, p. 209]:

- The expected waiting time

$$W = \frac{\rho}{1-\rho}\left(\frac{1}{2\mu} + \frac{\mu\sigma_s^2}{2}\right), \quad (3.4)$$

- The expected system length:

$$L = \rho + \frac{\rho^2 + \lambda^2 \sigma_s^2}{2(1-\rho)} \quad (3.5)$$

, and

- The expected queue length [PHB93, p. 369]:

$$L_q = \frac{\rho^2 + \lambda^2 \sigma_s^2}{2(1-\rho)} \, . \quad (3.6)$$

For the service station, it follows the Weibull distribution with a mean of 3.89 and a standard deviation of 2.57.

The Weibull probability density function is

$$f(x) = 0.1573 x^{0.57} e^{[-0.1002 x^{1.57}]} \, . \quad (3.7)$$

And the Weibull cumulative density function is

$$F(x) = 1 - e^{[-0.1002 x^{1.57}]} \, . \quad (3.8)$$

When the above model is run for 1,000,000 time units, the average queue length from the simulation result is 4.1074. The theoretical value of the average queue length is 4.1093. The expected waiting time is 18.4088 when we read directly from the simulation results, while the theoretical value of average waiting time is 18.4238.

Other alternative distributions are chosen to test the model consistency. If the left truncated Weibull distribution using the mean and variance method with the truncation point at $t = 0.5$ is chosen, the average queue length is 4.0994. The theoretical value of the average queue length is 4.0839. The expected waiting time is 18.3732 when read directly from the simulation results,

DOI: 10.2195/lj_Proc_fu_en_201210_01
URN: urn:nbn:de:0009-14-34525

while the theoretical value of average waiting time is 18.3100.

The next truncated distribution is the doubly truncated Weibull distribution using the mean variance method and the maximum likelihood estimation method. The simulation results show the average queue length is 4.1291 and the theoretical value of the expected queue length is 4.0694. The average waiting time read from the model results is 18.5065. The waiting time calculated from the formula is 18.2447. If the model is run for 10,000,000 time units, the results of the scenario with the doubly truncated Weibull distribution using the mean variance method and the maximum likelihood estimation method show that the average queue length is 3.9587 and the average waiting time is 17.7755. The theoretical values are 3.99 and 17.9169 respectively.

3.2 Truncated Distributions in Production Systems

After validating the truncated distributions, one production system with one source and two servers is simulated to test the effect of truncations. Firstly, the effect of the different distributions as source on the model is compared. The only modification in each round is the random number generation at the source component (but still with the same seed). A simulation of 100,000 time units is made for each alternative. The results of interest are the average waiting time (AWT) of both queues, the average queue length (AQL) of both queues, average dwelling time (ADT) in both servers, the Utilization (Ut) of both servers, the intergeneration time (IT), and the throughput (TP). The first round of simulation is made under the condition that the queue capacity is infinite.

This result has a significant sense in the fact that the intergeneration time in this table reflects the means of each distribution. From this table, we could see that the means of these alternatives are different from each other. The means of the LTWD is the highest of all, while the RTWD is the lowest. The DTWD is the closest to the original Weibull distribution. The average queue length and the average waiting time is another important aspect of the model. The reason why the queue length of RTWD is higher than the other alternatives lies not only in the fact that the means of RTWD intergeneration time is the lowest. We take the first queue as an example. The queue length before the first server is dependent on two factors: the state of the server and the state of the arrival station. The server time obeys the exponential distribution with the means of 3.46, as shown in the ADT S1. So the decisive aspect of the queue length is the inter-arrival time of the source. There are two factors in the inter-arrival time: the relieving factor and the aggravating factor. If the intergeneration time is extremely small, this would put an aggravation to the waiting line. On the other hand, the large inter-arrival time is a relief to the waiting queue since it gives the system more time to digest the block in the queue. These two factors are the main reason for the difference in Table 1.

Table 1. Weibull distribution as sources with infinite QC

Weibull distributions as sources				
QC = infinite no breakdown				
	Weibull	**LT 0.5**	**RT 12**	**DT 0.5 12**
IT	3.8805	3.962	3.8113	3.8495912
AWT Q1	20.336	16.115	23.557	19.222669
AQL Q1	5.2404	4.0669	6.1951	4.9927802
ADT S1	3.4646	3.4625	3.4699	3.4642568
Ut S1	0.8928	0.8738	0.909	0.8997541
AWT Q2	26.587	23.525	31.918	28.620898
AQL Q2	6.8505	5.9365	8.3599	7.4329963
ADT S2	3.5011	3.5043	3.4993	3.4991884
Ut S2	0.9021	0.8841	0.9164	0.908554
TP	25765	25229	26188	25963

To illustrate the difference between the original Weibull distribution and the truncated version, another more extreme case is taken where the left truncation point is chosen to be 1. The choice of the truncation point can be significant in fitting the sample to a distribution. The left truncation point should be set to less than 0.5 in this case. An extreme truncation point would only lead to an extreme outcome. The consequences caused by this choice are listed in Table 2:

Table 2. Weibull distribution as sources with QC = infinite without breakdowns

Weibull distributions as Sources			
QC=inf no breakdown			
	Weibull	**LT 0.5**	**LT 1**
IT	3.88046	3.96201	4.50271
AWT Q1	20.3359	16.1147	7.21497
AQL Q1	5.24036	4.06691	1.6023
ADT S1	3.46461	3.46247	3.4731
Ut S1	0.89276	0.87382	0.77127
AWT Q2	26.5871	23.5249	10.0238
AQL Q2	6.85046	5.93655	2.22601
ADT S2	3.50111	3.50429	3.50311
Ut S2	0.90208	0.88411	0.77778
TP	25765	25229	22202

DOI: 10.2195/lj_Proc_fu_en_201210_01
URN: urn:nbn:de:0009-14-34525

The above discussion also shows another point of view regarding the source of these different systems. That is, the mean inter-generation time of the source also plays an important role in the system performance. The following chapter deals with the truncation of Weibull distribution focused on the mean time.

3.3 TRUNCATION METHOD COMBINED WITH EXPECTATION EXPRESSION

First of all, the expressions of mean and variance of the different truncated distributions should be discussed. For the sake of brevity, the mean and variance of 3-parameter Weibull distribution and the 3-parameter left truncated Weibull distribution are listed below [MRB91]:

- Original Weibull distribution:

$$\begin{cases} E[X] = b\Gamma\left(\frac{1}{c}+1\right)+a \\ Var[X] = b^2\Gamma\left(\frac{2}{c}+1\right)-\left[b\Gamma\left(\frac{1}{c}+1\right)^2\right] \end{cases} \tag{3.9}$$

- Left truncated Weibull distribution:

$$\begin{cases} E[X] = e^{((t-a)/b)^c}\left[b\Gamma\left(\frac{1}{c}+1,\left((t-a)/b\right)^c\right)\right] \\ Var[X] = e^{((t-a)/b)^c}\left[b^2\Gamma\left(\frac{2}{c}+1,\left((t-a)/b\right)^c\right)\right]. \\ -\left[e^{((t-a)/b)^c}b\Gamma\left(\frac{1}{c}+1,\left((t-a)/b\right)^c\right)\right]^2 \end{cases} \tag{3.10}$$

When we combine the criteria of these two methods together, we have the two maximum likelihood estimation (MLE) functions, and two functions about the mean and the variance (M-V). So to utilize them to the fullest, not only the scale and shape parameters but also the truncation point is considered to be the unknown element here.

The truncation point is no longer a constant before the simulation and the modeling, which means, besides the maximum likelihood estimators, another one or two functions are needed to determine the additional variable. For the left truncated and right truncated Weibull distribution, we take the maximum likelihood estimators and the mean or variance of the distribution. For the doubly truncated distribution, we need the mean and the variance as well as the maximum likelihood estimators, because the doubly truncated distribution has two truncation points to estimate.

For the left truncated Weibull distribution, the maximum likelihood estimators and the mean expression (LTMM) are listed below:

$$\begin{cases} \frac{\partial L}{\partial a} = n/a - \sum(x_i^{\ b} - t^b) \\ \frac{\partial L}{\partial b} = n/b + \sum \log x_i - a\sum(x_i^{\ b}\log x_i - t^b \log t)\,. \\ E[x] = e^{((t-a)/b)^c}\left[b\Gamma\left(\frac{1}{c}+1,\left((t-a)/b\right)^c\right)\right] \end{cases} \tag{3.11}$$

After the substitution of the parameters and the solution of the parameters of interest, the left truncated Weibull probability density function with t = 0.7261 is

$$f(x) = 0.2380x^{0.2721}e^{[0.1245-0.109x^{1.2721}]}\,. \tag{3.12}$$

And the LTWD cumulative density function with t=0.7261 is

$$F(x) = 1 - e^{[0.1245-0.1871x^{1.2721}]}\,. \tag{3.13}$$

Similarly, all the alternatives of the truncated Weibull distribution using both maximum likelihood estimation and the mean-variance method could be induced. After running the same model where the queue has infinite capacity and the servers are without breakdowns, the result of the simulation is listed in Table 3.

Table 3. Truncated Weibull distributions using combined method as sources

Weibull distributions as sources				
QC=inf with no breakdown				
	Weibull	**LTMM**	**RTMM**	**DTMM**
IT	3.88046	3.88709	3.88521	3.90049
AWT Q1	20.3359	24.0523	21.2174	19.3045
AQL Q1	5.24036	6.18744	5.46078	4.94865
ADT S1	3.46461	3.46538	3.46525	3.46368
Ut S1	0.89276	0.89145	0.89184	0.88787
AWT Q2	26.5871	27.4793	26.0267	25.0933
AQL Q2	6.85046	7.06791	6.69738	6.43199
ADT S2	3.50111	3.50111	3.50167	3.50239
Ut S2	0.90208	0.90053	0.90102	0.89736
TP	25765	25718	25730	25620

This table reveals the effect of the truncation. Under the condition where the inter-generation times and the service times are almost the same, the queue length and the waiting time of with each truncated distributions still yield quite different results.

DOI: 10.2195/lj_Proc_fu_en_201210_01
URN: urn:nbn:de:0009-14-34525

3.4 GOODNESS-OF-FIT TESTS OF VARIOUS DISTRIBUTIONS

There are seven alternative distribution fittings to the empirical data, including the original Weibull distribution. Some of these distributions seem to be inappropriate to be mentioned as a "distribution fitting" as the parameters or the probability density function graph is far away from the histogram of the data. However, extra efforts need to be taken to test the fitness of distributions especially when some extreme cases are dealt. For example, when the n is very large, the test almost always rejects the hypothesis that the given data obey the target distribution [GD85]. There are some methods to test the goodness of fit of distributions. Two of the most commonly used ones are the chi-square test and the Kolmogorov–Smirnov test, or the K-S test.

Firstly, the level of significance is set to be 0.05 and the number of levels is set to be 13 in all the following chi-square test.

Table 4. Comparison of the total cost using two policies

	chi-square statistic	**critical value**	**Accept Hypothesis**
Weibull	11.3628	18.307	Yes
LT05	12.6018	18.307	Yes
RT12	10.9469	18.307	Yes
DT	13.1062	18.307	Yes
LTMM	15.4956	16.919	Yes
RTMM	13.0177	16.919	Yes
DTMM	11.4513	15.5073	Yes

The results of the K-S test with the same level of significance and calculated K-S statistics is shown in Table 5.

Table 5. K-S test results

	Reject Hypothesis?		
	unequal	larger	smaller
Weibull	no	no	no
LT05	no	no	no
RT12	no	no	no
DT	no	no	no
LTMM	yes	yes	no
RTMM	no	no	no
DTMM	no	no	no

From the results of these two goodness-of-fit tests, the left-truncated Weibull distribution with the combined method should be rejected. The left six distributions are integrated in a shipment consolidation model in the following chapter.

4 TRUNCATED DISTRIBUTIONS IN SHIPMENT CONSOLIDATION MODELS

Shipment consolidation is a shipping policy using joint stock replenishment and dispatching outstanding orders based on different criteria, namely, time and quantity. The vendor adopts an (s; S) policy, where s means the reorder point and S stands for the order-up-to level. Instead of immediately sending out the deliveries after the orders come in, the vendor would wait until one of the criteria is met. If the vendor decides to go for the time-based policy, the goods would be delivered after a fixed amount of time units. If the quantity-based policy is chosen, the vendor would send out the goods when the outstanding order quantity reaches a certain amount. Both of these shipment consolidation policies are based on the (s; S) replenishment policy [CCY00].

Therefore, the total cost that the vendor needs to consider falls onto four main cost components: replenishment cost, dispatching cost, warehouse cost, and the waiting cost. Please note that the last cost component is an imaginary cost which does not physically exist. It is the potential lost that the vendor keeps the customers waiting until the criteria are met. In a shipment consolidation model where the vendor needs to make a decision of whether to take the time-based policy or the quantity-based policy, the truncation of the distribution could influence the decision making as the results of the simulation using different distributions are quite different. The basic parameters for the shipment consolidation scenario is pre-set and tested for both policies with the original Weibull distribution. For the sake of brevity, only the comparison of the total cost is listed in the following table.

Table 6. Comparison of the total cost using two policies

	Time	Quantity	Difference
OWD	442,320	442,960	-640
LTWD MLE	442,060	440,670	1,390
RTWD MLE	443,390	441,780	1,610
DTWD MLE	443,110	442,040	1,070
RTWD MM	442,490	442,980	-490
DTWD MM	442,440	442,760	-320

DOI: 10.2195/lj_Proc_fu_en_201210_01
URN: urn:nbn:de:0009-14-34525

It could be observed that the results of these runs with truncated Weibull distributions yield different decision making for the vendor. In the view of the total cost, the Time-based policy has less total cost in three scenarios and the Quantity-based policy outperforms its rival in the other three scenarios. The interpretation of these results of the simulation needs more detailed analysis than the count of the advantageous cases.

5 Conclusion and Future Works

The truncated distributions have shown great impact on the system performance. When some components of the system are integrated with the truncated distributions, the simulation model can be greatly influenced by the slightly changed parameters. Some simulation models are sensitive to the extreme situations where the small probability events have large influence on the system. For some simulation models, the rare events could cause severe blocking in the queues and the service stations. When the distributions are truncated, the extreme situations are changed to a certain amount. Therefore, the system couldn't handle the extra burden and the whole system is trapped with the blocking. In the shipment consolidation models, the truncated distributions could influence the final decision making of the customer. The truncation has effect on each cost component as well as the total cost, although the truncated Weibull distributions and the original Weibull distribution have similar properties in many ways.

However, an absolute comparison on the effects of truncated version and the original version could not be induced without the help of neural network and a huge sample database that could suffice the large time units which are needed for the simulation. Also, the method in this paper used to find the parameters of the truncated Weibull distribution could be improved to a method with higher efficiency and less complexity. In the future, an ideal simulation tool with truncated version would at least contain the above mentioned potential improvements and an accurate analysis of the results as well as the impact of the truncation should also be listed for reference. The truncation of the distributions is a powerful tool which could provide more significant insight of the simulation model with these improvements.

Literature

[AC50] Cohen, Jr. A. Clifford: *Estimating the Mean and Variance of Normal Populations from Singly Truncated and Doubly Truncated Samples*, In: The Annals of Mathematical Statistics, Vol. 21, No. 4, Edwards Brothers, Incorporated, 1950, p. 557-569

[AL07] Law, Avrill M.: *Simulation Modeling and Analysis*. Fourth edition, McGraw-Hill, 2007, ISBN 9780071255196

[BE80] Bain, Lee J.; Engelhardt, Max: *Inferences on the Parameters and Current System Reliability for a Time Truncated Weibull Process*. In: Technometrics, Vol. 22, No. 3, American Society for Quality Control, 1980, p. 421-426

[CCY00] Cetinkaya, S.; Lee, Chung-Yee: *Stock Replenishment and Shipment Scheduling for Vendor Managed Inventory Systems*. In: Management Science, 48, Institute of Management Sciences, 2000, p. 217-232

[CO50] Cohen, Jr. A. Clifford: *Estimating parameters of Pearson Type III populations from truncated samples*. In: Jour. Am. Statist. Assoc. 45, 1950, p. 411-423

[CW69] Choi, S. C.; Wette, R.: *Maximum likelihood estimation of the parameters of the gamma distribution and their bias*. In: Technometrics, 11, American Society for Quality Control, 1969, p. 683-690.

[DD83] Depriest, Douglas. J.: *Using the singly truncated normal distribution to analyze satellite data*. In: Communications in Statistics - Theory and Methods, Volume 12, Issue 3, Marcel Dekker, 1983, p. 263 - 272

[DRW89] Wingo, D. R.: *The left-truncated Weibull distribution: theory and computation*. In. Statistical Papers, Springer, 1989

DOI: 10.2195/lj_Proc_fu_en_201210_01
URN: urn:nbn:de:0009-14-34525

[FM76] Finkelstein, J. M.: *Confidence bounds on the parameters of the Weibull process*. In: Technometrics, American Society for Quality Control, 1976, p. 18, 115-117

[GD60] Greenwood, J. Arthur; Durand, David: *Aids for Fitting the Gamma Distribution by Maximum Likelihood*. In: Technometrics, Vol. 2, No. 1, American Society for Quality Control, 1960, p. 55-65

[GD85] Gibbons, J. D.: *Nonparametric Methods for Quantitative Analysis*. Second Edition, American Sciences Press, 1985

[GH98] Gross, Donald; Harris, Carl M.: *Fundamentals of queueing theory*. Third edition, John Wiley & Sons, 1998

[HT91] Takagi, Hideaki: *Queueing Analysis*. In: Vol. 1 Vacation and Priority Systems, North-Holland, 1991

[KSR08] Kliche, D. V.; Smith, P. L.; Johnson, R. W.: *L-moment estimators as applied to gamma drop size distributions*. In: Journal of Applied Meteorology and Climatology, 47, 2008, p. 3117-3130

[MO82] Moore, R. J.: *Derivatives of the incomplete gamma integral*. In: Appl. Statist. 31, 1982, p. 330-335

[MRB91] McEwen, Robert P.; Parresol, Bernard R.: *Moment expressions and summary statistics for the complete and truncated weibull distribution*. In: Communications in Statistics - Theory and Methods, 20:4, Marcel Dekker, 1991, p. 1361-1372

[PHB93] Papadopoulos, H. T.; Heavey, C.; Browne, J.: *Queueing Theory in Manufacturing Systems Analysis and Design*. Springer, 1993

[RC81] Cooper, Robert B.: *Introduction to queueing theory*. Second edition, North Holland, 1981

[RKS09] Johnson, Roger W.; Kliche, D.; Smith, P. L.: *Maximum likelihood estimation from a left-truncated distribution*. 34th Conference on Radar Meteorology, 2009