Erasmus School of Economics

Bachelor Thesis

Extrapolating the Empirical Distribution

Using extreme value theory to improve demand forecasting in inventory management [1]

Author:
Rex Wang
(343490)

Academic Supervisor:
Prof.Dr.Ir. R. Dekker

Co-reader:
Dr. A.J. Koning

June 29, 2013

[1]Special thanks to Prof.Dr. Laurens de Haan for his guidance and help on the magnificent extreme value theory.

Abstract

The empirical method is one of the best-performing demand forecasting methods in inventory decision making for products with intermittent demands. However, this method has difficulties obtaining high fill rates since the used corresponding empirical distribution cannot handle extreme values that has not been observed. To improve the empirical method on this issue, this thesis first briefly introduces the extreme value theory (EVT) which provides a solid theoretical basis and framework for tail estimation and extrapolation, and then proposes three extrapolation methods. Verified by simulation studies, the extrapolation method applying EVT works the best among the proposed methods and successfully extrapolates the empirical distribution in the tail area.

Contents

Chapter 1

Introduction

A fundamental aspect of inventory management is accurate demand forecasting. Without an estimate of the future customer demand it is impossible to plan the levels of inventories that will be required to offer customers a good level of service.

Inventory control for parts with infrequent demands requires a separate discipline since their demand is often characterized as being intermittent or irregular with a large proportion of zero and sudden high values. Such items include service (spare) parts and high-priced capital goods, e.g. heavy machinery, and are often described as 'slow-moving'. For the companies making, using, or maintaining this kind of items, it is essential to determine the right amount of stocks in order to avoid high inventory costs and penalties in case of availability.

Quite some methods have been developed to forecast the demand and improve inventory control for spare parts. Basten et al. (2012) is one of the few studies which compared several spare part demand forecasting methods using real data from three companies. One of the best-performing forecasting method appears to be the *empirical method* introduced by (Porras and Dekker, 2008). This method samples the lead time demand (LTD) from the daily demands using a moving window and makes use of the corresponding empirical distribution to determine the important inventory control parameter, the re-order point.

Let $(X_1, \ldots, X_n)$ be an obtained sample LTD of size n, the corresponding empirical distribution F_n is defined as the following,

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{X_i \leq x\}}.$$

To determine a re-order point s for a specified service level of fill rate β, one chooses the smallest s satisfying:

$$\beta \leq 1 - \frac{ES(s)}{Q},$$

where Q is a pre-determined lot size, and $ES(s)$ is the expected units short for a given re-order point s, which is defined as follows:

$$ES(s) = \mathbb{E}\left[(X - s)\,\mathbf{1}_{\{X>s\}}\right] = \int_s^\infty (x - s)\,\mathrm{d}F(x).^1$$

As Basten et al. (2012) points out, using the empirical distribution F_n to determine re-order points has difficulties obtaining high fill rates, since it cannot take care of values greater than the sample maximum, in other words, it fails to evaluate $ES(s)$ for $s > X_{n,n}$.

To tackle this issue we need to extrapolate the empirical distribution in the tail area.

This bachelor-thesis therefore investigates *how the empirical distribution can be extrapolated in the tail area in order to improve the empirical method?*

Methodology of our research consists mainly of two parts. In the first part, we propose and discuss extrapolation methods by applying the extreme value theory (EVT), since EVT provides a solid theoretical basis and framework for tail estimation and extrapolation. Moreover, the concerned quantity, expected shortage $ES(s)$, is similar to the expected shortfall $\mathbb{E}(X|X > \mathrm{VaR}_q)$ in financial risk management, where VaR_q is the q-th quantile of the distribution. To accurately estimate the expected shortfall for large VaR_q's, one often applies EVT in finance. Obviously, we can also apply EVT in our context to estimate the expected shortage. The second part of our methodology is to use simulation studies to assess the performance of the extrapolation methods. Based on the simulation results, we wish to determine an extrapolation method that effectively captures the tail behavior of the lead time demands and successfully extrapolates the empirical distribution. The best performing extrapolation method can be used to improve the empirical method.

The set-up of this thesis is as follows. Chapter 2 introduces EVT briefly and discusses some useful application of the theory. Using the idea of EVT, we propose in Chapter 3 several tail extrapolation methods to model the tail of LTD distributions. In Chapter 4 we conduct the simulation study where the performance of these extrapolation methods will be assessed and discussed. Finally, conclusions and limitations are given in Chapter 5.

[1] In this thesis, unless otherwise mentioned, all the integrals are Lebesgue-Stieltjes integration. In this way, we do not need to distinguish discrete and continuous distributions.

Chapter 2

Extreme Value Theory (EVT)

Extreme value theory (EVT) provides a framework to formalize the study of behavior in the tail of a distribution. EVT allows us to use extreme observations to measure the density in the tail. This measure can be extrapolated to parts of the distribution that are yet to be observed in the empirical data.

EVT was pioneered by Fréchet (1927) and Fisher and Tippett (1928). They have developed the one-dimensional probabilistic EVT. The asymptotic theory was then unified and extended by Gnedenko (1943). The statistical theory was initiated by Pickands (1975). Since the 1980s the contours of the relevant statistical theory started taking shape. And nowadays, EVT has already been applied in various fields, from finance to flood-control.

The approach to EVT in this thesis follows most closely de Haan and Ferreira (2006) because of its accessibility, thoroughness and self-containedness. All the theorems and results used in this chapter are stated and proved in this book.

This chapter first briefly introduces EVT in a way such that we will just have enough theoretic background to apply it. Then we will discuss two important applications of EVT and the relevant estimators. In the end of this chapter, two examples will be considered to demonstrate the power of EVT.

2.1 General Theory

The idea of EVT is basically to restrict the behavior of the distribution function in the tail to resemble a limited class of functions that can be fitted to the tail of the distribution.

Consider first the Generalized Pareto distribution (GPD). The cumulative distribution function of GPD is defined by

$$G_\gamma(x) = \begin{cases} 1 - (1+\gamma x)^{-1/\gamma} & \text{for } \gamma \neq 0, \\ 1 - e^{-x} & \text{for } \gamma = 0, \end{cases}$$

where the support is $x \geq 0$ when $\gamma \geq 0$ and $0 \leq x \leq -1/\gamma$ when $\gamma < 0$.

This distribution is generalized in the sense tht it subsumes certain other distributions under a common parametric form. γ is the important *shape* parameter of the distribution. If $\gamma > 0$, G_γ is a reparametrized version of the ordinary Pareto distribution; $\gamma = 0$ corresponds to the exponential distribution and $\gamma < 0$ is known as Pareto type II distribution.

The following limit theorem is one of the key results in EVT and explains the importance of the GPD.

Theorem 1. *(de Haan and Ferreira, 2006, Theorem 1.2.1) The distribution function F is in the domain of attraction of the extreme value distribution GPD if and only if there exists some positive function f such that*

$$\lim_{s \uparrow x^*} \frac{1 - F(s + xf(s))}{1 - F(s)} = 1 - G_\gamma(x) \tag{2.1}$$

for all x for which with $1 + \gamma x > 0$, where $x^* := \sup\{x : F(x) < 1\}$. γ is called the extreme value index. For $\gamma > 0$, x^* is infinite; for $\gamma < 0$, x^* is finite; for $\gamma = 0$, x^* can be finite or infinite.

For this thesis it is sufficient to know that all the common distributions of statistics belong to the domain of attraction of GPD (e.g. normal, lognormal, uniform, beta, exponential etc.). In fact, it is not easy to find distribution functions that do not belong to this domain of attraction. In practice, this condition cannot be checked since we do not know the tail. But this is a common feature in statistics, e.g., for estimating the mean one has to assume it exists and for assessing the accuracy one usually assumes the existence of the second moment.

There is an alternative formulation of this theorem in terms of a function U which is the *left-continuous inverse* of $1/(1 - F)$, i.e.,

$$U(t) := \inf\{y : \ 1/(1 - F(y)) \geq t\} = \inf\{y : \ F(y) \geq 1 - 1/t\}. \tag{2.2}$$

This function plays a vital role in extreme value theory which will become clear in the next section. The following theorem assures the reformulation of (2.1) in terms of U.

Theorem 2. *(de Haan and Ferreira, 2006, Theorem 1.1.6)*
For $\gamma \in \mathbb{R}$, the following statement is equivalent to (2.1):
There is a positive function α such that for $x > 0$,

$$\lim_{s \to \infty} \frac{U(sx) - U(s)}{\alpha(s)} = \frac{x^\gamma - 1}{\gamma}. \tag{2.3}$$

where for $\gamma = 0$ the right-hand side is interpreted as $\log(x)$.
Moreover, this equivalence also holds with

$$f(s) = \alpha\left(\frac{1}{1 - F(s)}\right). \tag{2.4}$$

2.2 Applications of EVT

This section discusses the two most important applications of EVT.

Let X be a random variable with distribution function F which belongs to the domain of attraction of GPD. Theorem 1 implies loosely speaking that from some high threshold s onward (i.e. $X > s$) the distribution function can be written approximately as

$$1 - F(x) \approx (1 - F(s)) \left\{ 1 - G_\gamma \left(\frac{x - s}{f(s)} \right) \right\}, \quad x > s.$$

which is a parametric family of distribution tails. One can expect this approximation to hold for intermediate and extreme order statistics.

Let $X_1, X_2, \ldots$ be i.i.d. random variables with distribution function F, and F_n the corresponding empirical distribution function. Let us apply the last approximation with $s := X_{n-k,n}$, where we choose $k = k(n) \to \infty$, $k/n \to 0$, as $n \to \infty$. Then

$$1 - F(x) \approx (1 - F(X_{n-k,n})) \left\{ 1 - G_\gamma \left(\frac{x - X_{n-k,k}}{f(X_{n-k,n})} \right) \right\}$$

and, since $1-F(X_{n-k,n}) \approx 1-F_n(X_{n-k,n}) = k/n$, and by relation (2.4) $f(X_{n-k,n}) = \alpha(1/(1 - F(X_{n-k,n}))) \approx \alpha(n/k)$, we have

$$1 - F(x) \approx \frac{k}{n} \left\{ 1 - G_\gamma \left(\frac{x - X_{n-k,k}}{\alpha(\frac{n}{k})} \right) \right\}. \tag{2.5}$$

This approximation is valid for any x larger than $X_{n-k,n}$ and can be used even for $x > X_{n,n}$, which is outside the range of the observation. This is in fact the basis for applications of extreme value theory.

Next we consider the second application of EVT applying Theorem 2. Relation (2.3) leads to the following approximation:

$$U(x) \approx U(s) + \alpha(s) \frac{(\frac{x}{s})^\gamma - 1}{\gamma}, \qquad x > s.$$

This approximation in useful when one wants to estimate a quantile $F^{\leftarrow}(1-p) = U(1/p)$ with p very small, since this quantile is then related to a much lower quantile $U(s) = F^{\leftarrow}(1 - 1/s)$, which can be estimated by an intermediate order statistic. Hence we choose $s := n/k$ with $k = k(n) \to \infty$, $k/n \to 0$, $n \to \infty$. Then for small p,

$$U\left(\frac{1}{p}\right) \approx U\left(\frac{n}{k}\right) + \alpha\left(\frac{n}{k}\right) \frac{(\frac{k}{np})^\gamma - 1}{\gamma}. \tag{2.6}$$

According to (2.2), the quantity $U(n/k)$ can be estimated by the intermediate order statistic $X_{n-k,n}$. Again this approximation will be used not only for $1/p < n$, but also for extrapolation outside the sample.

An immediate result of relation (2.6) is to approximate the endpoint x^*. If the underlying distribution F belongs to the domain of attraction of GPD for some negative γ, the endpoint x^* is by Theorem 1 finite. An estimator of x^* can be motivated from relation (2.6). Recall $x^* := \inf\{x : F(x) < 1\}$, let $p \to 0$ in (2.6), we obtain

$$U(\infty) \approx U\left(\frac{n}{k}\right) - \frac{\alpha(\frac{n}{k})}{\gamma}.$$

Hence the following estimator for the endpoint is proposed

$$\hat{x}^* := X_{n-k,n} - \frac{\alpha(\frac{n}{k})}{\gamma}. \tag{2.7}$$

In order to make approximation (2.5) and (2.6) applicable, we need to estimate γ, the function α at point n/k which will be discussed in the next section.

2.3 The moment estimators

Various estimators for the extreme value index γ have been introduced and developed in the field of EVT. The most commonly used estimators are discussed in de Haan and Ferreira (2006), such as the Hill (1975), Pickands (1975), Maximum-Likelihood (MLE), moment, probability weighted moment (PWM) (Hosking and Wallis, 1987) and negative Hill (Falk, 1995).

The Hill estimator is consistent only for positive values of γ, the MLE is defined for $\gamma > -1/2$, the PWM is consistent for $\gamma < 1$, and the negative Hill for $\gamma < -1/2$. Only the Pickands and moment estimators are defined and consistent for all real values of γ. However, for a large range of values of γ, the Pickands estimator has larger asymptotic variance than the others. Moreover, the implementation of the moment estimator is quite convenient comparing with the other estimators. Therefore, we will only consider the moment estimators for the rest of this thesis.

For $j = 1, 2$, define

$$M_n^{(j)} := \frac{1}{k}\sum_{i=0}^{k-1}\left(\log X_{n-i,n} - \log X_{n-k,n}\right)^j.$$

The moment estimator for γ is defined as follows

$$\hat{\gamma} := M_n^{(1)} + 1 - \frac{1}{2}\left(1 - \frac{\left(M_n^{(1)}\right)^2}{M_n^{(2)}}\right)^{-1}.$$

Related to this moment estimator, we introduce an estimator for the scale function α.

Define

$$\hat{\gamma}_{-} := 1 - \frac{1}{2}\left(1 - \frac{\left(M_n^{(1)}\right)^2}{M_n^{(2)}}\right)^{-1}.$$

The estimator for the scale function is define as

$$\hat{\alpha}\left(\frac{n}{k}\right) := X_{n-k,n} M_n^{(1)} \left(1 - \hat{\gamma}_{-}\right).$$

This estimator has been proved to be consistent and asymptotically normally distributed.

There is still one issue left to be considered, i.e., the choice of k. It is theoretically proved for $k(n) \to \infty$, and $k(n)/n \to 0$ as $n \to \infty$, we have consistency of the estimators $\hat{\gamma}$ and $\hat{\alpha}$. However, it remains a complex problem to choose the k with respect to the sample size n optimally. There is trade-off between choosing k too small or too large. On the one hand, if k is too small, there are few observations which will result in large variance. On the other hand, if k is chosen too large, we involve "non-extreme" observations, which will impose bias in the estimators. A practical solution is to choose the k using a diagram of estimates, which means that we compute and plot $\hat{\gamma}$ for $k = 1, \ldots, n-1$, and the choose the k at which the values of $\hat{\gamma}$ is stabilized. For example, Figure 2.1 exhibits the diagram of estimates for a sample with size 50 of folded Slash distributed random variables. (How to generate folded Slash distributed random variables is explained in Appendix B.) In this case, we will choose $k = 25$.

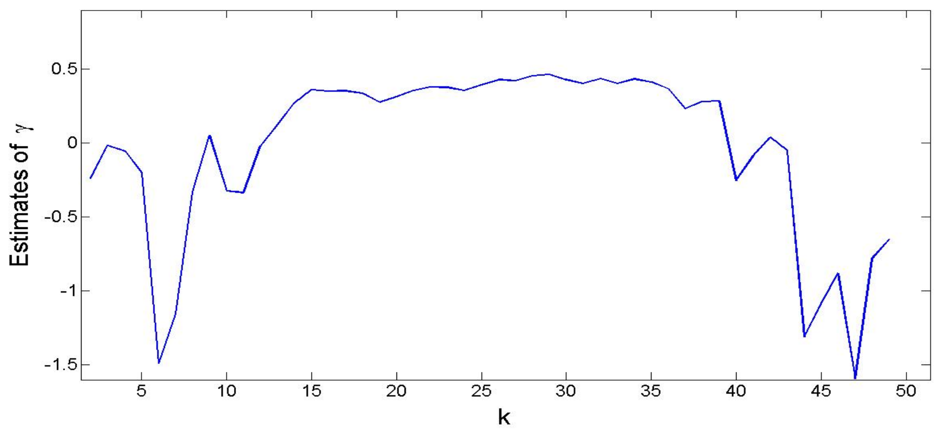


Figure 2.1: Diagram of the estimates of γ with respect to the choice of k

Using these estimators, relations (2.5), (2.6) and (2.7) can be henceforth applied to estimate tail probabilities, quantiles and endpoints respectively.

2.4 Examples

This section contains two examples of applications of EVT.

The first example demonstrates how EVT can be applied to estimate the tail probability. Let $X \sim \text{Weibull}(\theta, \lambda)$. The tail probability can be easily calculated by using the corresponding CDF F. For a large c, we have

$$\mathbb{P}(X > c) = 1 - F(x; \theta, \lambda) = \exp\left(-\left(\frac{c}{\lambda}\right)^{\theta}\right).$$

However, in practice, we do not know the underlying distribution most of the times, but only have a sample of random variables from the distribution. Instead of using any other advanced statistical techniques, simply applying relation (2.5) already provides a good estimator of the tail probability.

To illustrate the accuracy of this estimator, we perform simulation for the Weibull distribution with different parameters k and λ, and difference sample sizes. The simulation is conducted as follows:

1. draw a sample of size $n = 50$ from the Weibull distribution with $\theta = 5$ and $\lambda = 1/2$;
2. estimate $\hat{\gamma}$ and $\hat{\alpha}(\frac{n}{k})$ for different k, and choose the optimal k by means of diagram of estimates of γ;
3. apply relation (2.5) to approximate $\mathbb{P}(X > c)$ for different values of c.
4. Repeat step 1-3 for a large number of times and take the average of the tail probability estimates.
5. Repeat step 1-4 for different pairs of (θ, λ) and sample size n.

Table 2.1 exhibits the simulation results of the estimation. Obviously, the larger the sample, the more accurate the tail probability estimates become. Moreover, the further the point locates in the tail area, the more accurate the estimation become. Although EVT is a asymptotic theory, the estimates appear to be already close to the theoretical value with a small sample of size 50.

The second example shows how EVT can be used as the quantile function of a distribution. Let $Y \sim \text{Uniform}(a, b)$, we try to estimate the value x_p for which $\mathbb{P}(X \leq x_p) = 1 - p$. Theoretically we determine x_p for a given p as follows:

$$x_p = F^{\leftarrow}(1 - p; a, b) = a + (1 - p)(a - b)\mathbf{1}_{[0,1]}(p).$$

In addition, as the uniform distributions have finite endpoints, the estimates of the endpoints will be computed as well. Using relation (2.6), we perform the same procedure as the first example to estimate the quantiles for different samples sizes, different parameters and different values of p.

		$\mathbb{P}(X > c)$	Approximation by (2.5)		
			$n = 50$	$n = 500$	$n = 5000$
$\theta = 5$, $\lambda = 1/2$	$c = 25$	0.1069	0.1363	0.1104	0.1103
	$c = 35$	0.0710	0.0750	0.0729	0.0705
	$c = 50$	0.0423	0.0511	0.0423	0.0423
$\theta = 8$, $\lambda = 1/3$	$c = 25$	0.2318	0.1881	0.1832	0.1989
	$c = 35$	0.1948	0.2133	0.2021	0.2004
	$c = 50$	0.1585	0.1649	0.1564	0.1579
	$c = 80$	0.1160	0.0992	0.1082	0.1160
	$c = 100$	0.0982	0.0856	0.0945	0.0985
$\theta = 10$, $\lambda = 1/4$	$c = 25$	0.2844	0.2668	0.2833	0.2873
	$c = 35$	0.2547	0.2401	0.2462	0.2520
	$c = 50$	0.2242	0.1982	0.2243	0.2242
	$c = 80$	0.1860	0.1582	0.1854	0.1859
	$c = 500$	0.07	0.0675	0.0688	0.685
	$c = 800$	0.0503	0.0540	0.467	0.485

Table 2.1: Simulation results of tail probability estimations

		$F^{\leftarrow}(1 - p; a, b)$	Approximation by (2.6)		
			$n = 50$	$n = 500$	$n = 5000$
$a = 10$, $b = 50$	$p = 0.01$	49.6	48.0507	49.4834	49.5165
	$p = 0.05$	48	46.5003	47.9105	47.8995
	$p = 0.10$	46	44.5153	45.8139	46.0432
	$p = 0.2$	42	41.1714	42.0302	42.0086
	x^*	50	49.2922	49.9657	49.9789
$a = 75$, $b = 100$	$p = 0.01$	99.75	99.3899	99.7482	99.7493
	$p = 0.05$	98.75	98.3312	98.7284	98.7398
	$p = 0.10$	97.5	97.4532	97.4897	97.4880
	$p = 0.2$	95	94.8104	94.9464	95.0110
	x^*	100	99.9968	100.0102	99.9991

Table 2.2: Simulation results of quantile estimations

The results are shown in Table 2.2.

Similar conclusions can be drawn for this example. The more observations there are available, the better the EVT quantile (and endpoints) estimates become. Nevertheless, EVT already provides reasonably accurate estimates with a small sample (with 50 observations).

Chapter 3

Tail Extrapolation for Empirical distribution

This chapter explains three tail extrapolation methods that are meant to improve the empirical method proposed by Porras and Dekker (2008).

Recall that in inventory decision making, one needs to determine inventory control parameters, such as the re-order points. The empirical method is used to estimate the distribution of LTD and consequently determine the re-order points. This method samples the LTD from the daily demands using a moving window and makes use of the corresponding empirical distribution.

Let $(X_1, \ldots, X_n)$ be an obtained sample of LTD of size n. From the corresponding empirical distribution F_n, we can determine the probability mass function $\widehat{f}(x)$,

$$\widehat{f}(x) = \mathrm{d}\, F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{X_i = x\}}. \tag{3.1}$$

The re-order point s for a specified fill rate β is chosen to be the smallest s satisfying:

$$\beta \leq 1 - \frac{ES(s)}{Q}, \tag{3.2}$$

where Q is a pre-determined lot size and $ES(s)$ is evaluated as follows:

$$ES(s) = \sum_{x|x>s} (x - s)\widehat{f}(x).$$

As Basten et al. (2012) points out, this method has difficulties obtaining high fill rates, since it fails to evaluate $ES(s)$ for $s > X_{n,n}$. To tackle this issue, we propose the following three tail extrapolation methods.

3.1 Tail extrapolation using EVT

The first tail extrapolation method applies a few important results of EVT assuming that the underlying distribution of LTD belongs to the domain of attraction of GPD.

We first use the estimators discussed in section 2.3 to estimate all the parameters needed to apply the theorems, such as the shape parameter γ and the scale function $\alpha(\frac{n}{k})$. Note that the choice of a proper k can be determined by means of a diagram of estimates.

Then Relation (2.5) suggests that

$$\mathbb{P}(X < X_{n,n}) = F(X_{n,n}) \approx 1 - \frac{k}{n} \cdot G_{\hat{\gamma}} \left(\frac{X_{n,n} - X_{n-k,n}}{\hat{\alpha}(\frac{n}{k})} \right) := \widehat{F}(X_{n,n}). \tag{3.3}$$

If $\widehat{F}(X_{n,n}) = 1$, we will just continue using the empirical method. $\widehat{F}(X_{n,n}) < 1$ means that there may be greater values than the sample maximum $X_{n,n}$ and extrapolation is necessary.

The extrapolation starts with scaling down the probability mass function $\widehat{f}(x)$ with the value $\widehat{F}(X_{n,n})$, i.e., $\forall x \leq X_{n,n}$,

$$\widehat{f}(x) = \frac{\widehat{F}(X_{n,n})}{n} \sum_{i=1}^{n} \mathbf{1}_{\{X_i = x\}}. \tag{3.4}$$

Then we can evaluate the expected shortage $ES(s)$ as follows:

$$\widehat{ES}(s) = \sum_{x | s < x < X_{n-k,n}} (x - s)\widehat{f}(x) + \widehat{ES}_{tail}\left(\max(s, X_{n-k,n})\right), \tag{3.5}$$

where

$$\widehat{ES}_{tail}(s) = \begin{cases} \frac{k}{n} \cdot \frac{\hat{\alpha}(\frac{n}{k})}{1-\hat{\gamma}} \left(1 + \hat{\gamma}\left(\frac{s - X_{n-k,n}}{\hat{\alpha}(\frac{n}{k})}\right)\right)^{1-1/\hat{\gamma}} & 0 < \hat{\gamma} < 1, \\ \frac{k}{n} \cdot \hat{\alpha}\left(\frac{n}{k}\right) \exp\left(\frac{-s + X_{n-k,n}}{\hat{\alpha}(\frac{n}{k})}\right) & \hat{\gamma} = 0, \\ \frac{k}{n} \cdot \frac{\hat{\alpha}(\frac{n}{k})}{1-\hat{\gamma}} \left\{ \left(1 + \hat{\gamma}\left(\frac{s - X_{n-k,n}}{\hat{\alpha}(\frac{n}{k})}\right)\right)^{1-1/\hat{\gamma}} \right. & \\ \left. - \left(1 + \hat{\gamma}\left(\frac{\hat{x}^* - X_{n-k,n}}{\hat{\alpha}(\frac{n}{k})}\right)\right)^{1-1/\hat{\gamma}} \right\} & \hat{\gamma} < 0. \end{cases} \tag{3.6}$$

The derivation of $\widehat{ES}_{tail}$ is presented in Appendix A.1.
Note if $\hat{\gamma} < 0$, the endpoint x^* is supposed to be finite and should be estimated using (2.7). However, for $\hat{\gamma} \geq 1$, the expected shortage goes to infinity, and we would just use the original empirical method.

Furthermore, one of the nice features of the empirical distribution is to capture all the special characteristics of the sample, e.g., a sample of LTD exclusively contains values that are multiples of 2. To retain this feature during extrapolation, we choose the step size of extrapolation ε as the smallest positive difference between the sample values, i.e.,

$$\varepsilon := \min\{X_{i+1,n} - X_{i,n} > 0 : \; 1 \leq i \leq n-1\}. \tag{3.7}$$

Hence, for a target fill rate β, the re-order point s will be chosen to be the smallest s, such that (3.2) holds, where $ES(s)$ is evaluated by (3.5) and (3.6). Starting from an initial value of s, we will keep increasing s with step size ε until (3.2) holds.

3.2 Tail extrapolation using exponential distribution

The second tail extrapolation method makes use of the tail of exponential distributions. The exponential distribution has a heavier tail than the normal distribution, but a lighter tail than the heavy-tailed distributions like the Cauchy distribution. However, many of the common distributions belong to the class of distributions with simple exponential tails (Jones, 2008), for example the Gamma distribution and Weibull distributions. Therefore we decide to propose this method using the tail of exponential distributions.

The extrapolation procedure is similar to the first one.

Assuming that the LTD is exponentially distributed, a Maximum-Likelihood estimators of the mean $\hat{\mu}$ will be estimated based on the sample. The estimator $\hat{\mu}$ is used for the following estimator to estimate the tail probability $\mathbb{P}(X < X_{n,n})$:

$$\widehat{F}(X_{n,n}) := 1 - \exp\left(-\frac{x}{\hat{\mu}}\right). \tag{3.8}$$

Note that we will always have $\widehat{F}(X_{n,n}) < 1$ due to the Maximum-Likelihood estimators $\hat{\mu}$. The probability mass function $\widehat{f}$ will therefore always be scaled down with the value $\widehat{F}(X_{n,n})$ using (3.4).

The expected shortage $ES(s)$ in (3.2) will be evaluated as follows:

$$\widehat{ES}(s) = \sum_{x | s < x < X_{n,n}} (x-s)\widehat{f}(x) + \widehat{ES}_{tail}\left(\max(s, X_{n,n})\right), \tag{3.9}$$

where

$$\widehat{ES}_{tail}(s) = \hat{\mu} \exp\left(-\frac{s}{\hat{\mu}}\right). \tag{3.10}$$

The derivation of (3.10) is presented in Appendix A.2.

Furthermore, this method also makes use of the step size ε determined by (3.7), in order to capture the special characteristics of the sample.

Hence, to determine re-order points given a target fill rate, we begin with an initial value of s, say $X_{1,n}$. We keep increasing s each time by step size ε, until (3.2) holds.

3.3 Tail extrapolation using normal distribution

The third tail extrapolation method makes use of the tail of normal distributions, since normal distributions are very popular in statistics, and are often used to fit real-valued random variables whose distributions are not known (Casella and Berger, 2001).

The extrapolation procedure is very similar to the other methods.

Given the sample of LTD, we first determine the Maximum-Likelihood estimators $\hat{\mu}$ and $\hat{\sigma}$ assuming that the LTD is normally distributed. Then we can use the following estimator to estimate the tail probability $\mathbb{P}(X < X_{n,n})$:

$$\widehat{F}(X_{n,n}) := \Phi\left(\frac{X_{n,n} - \hat{\mu}}{\hat{\sigma}}\right). \tag{3.11}$$

Due the Maximum-Likelihood estimators, $\widehat{F}(X_{n,n})$ will always be less than 1. Hence the probability mass function $\widehat{f}$ will always be scaled down with the value $\widehat{F}(X_{n,n})$ using (3.4).

Then we can evaluate the expected shortage $ES(s)$ as follows:

$$\widehat{ES}(s) = \sum_{x|s<x<X_{n,n}} (x - s)\widehat{f}(x) + \widehat{ES}_{tail}\left(\max(s, X_{n,n})\right), \tag{3.12}$$

where

$$\widehat{ES}_{tail}(s) = \frac{\hat{\sigma}}{\sqrt{2\pi}} \exp\left(-\frac{(s-\hat{\mu})^2}{2\hat{\sigma}^2}\right) + (\hat{\mu} - s)\left(1 - \Phi\left(\frac{s-\hat{\mu}}{\hat{\sigma}}\right)\right). \tag{3.13}$$

The derivation of (3.13) is presented in Appendix A.3.

Last but not the least, this extrapolation method also makes use of the step size ε determined by (3.7), when determining the re-order point. An initial value of s will be increased with this step size until (3.2) holds, where $ES(s)$ is of course evaluated by (3.12) and (3.13).

3.4 General algorithm

It is noteworthy that the three extrapolation methods discussed above follow a very similar procedure which can be summarized by Algorithm 1.

Algorithm 1 Tail Extrapolation for Empirical distribution

Require: X : the sample of lead time demands
β : the target fill rate

Output: s : reorder point satisfying condition (3.2)

construct the empirical mass function $\widehat{f}$ of the sample X, see (3.1)
estimate all the parameters needed to apply the extrapolation method
estimate $\mathbb{P}(X \leq X_{n,n})$ using estimator $\widehat{F}(X_{n,n})$ (3.3), (3.8) or (3.11)
if $\widehat{F}(X_{n,n}) < 1$ **then**
    scale $\widehat{f}$ down with $\widehat{F}(X_{n,n})$ by (3.4)
    determine the step size ε by (3.7)
    define $\widehat{ES}(s)$ by (3.5), (3.9) or (3.12)
    initialize S e.g. set $s = X_{1,n}$
    while $1 - \widehat{ES}(s)/\overline{X} < \beta$ **do**
        $s = s + \varepsilon$
    end while
else
    use the original empirical method
end if

3.5 Simulations

In this section, we perform a simulation study in order to assess the performance of these discussed tail extrapolation methods.

Let X denote the LTD. For different target fill rate β and various distributions of X, we first determine the theoretic re-order point s^* by solving the following equation for s numerically:

$$1 - \frac{\int_s^\infty (x - s)\,\mathrm{d}F(x)}{Q} = \beta,$$

where F is the distribution function of X and Q is chosen to be 1.
Then we repeat the following steps for a large number of times:

1. generate a sample of X with size N according to the specified distribution;

2. determine the re-order points by means of the empirical distribution and the three discussed tail extrapolation methods respectively.

Calculate the average of the re-order points for different methods.

This simulation is performed for folded normal distribution with location parameter 6 and scale parameter 5 and folded non-standard t distribution with 2 degrees of freedom and location parameter 6. We make use of folded distribution because of the non-negative product demands in inventory management. Note this folded t distribution has a much heavier tail than the folded normal distribution. The density functions and sample simulation of these two folded distributions are presented in Appendix B.

By these two distributions, we can show how the extrapolation methods perform with respect to the empirical distribution on light-tailed as well as heavy-tailed distributions.

The results of this simulation are shown in Table 3.1 and 3.2.

In the case of the light-tailed folded normal distribution, EVT extrapolation method can yield re-order points which quickly converge from above to the optimal s^*. Even with a small sample size of 10, the re-orders are already very close to s^*. Extrapolation methods using exponential and normal tails appear to be overestimating overall, which will result in redundant high inventory costs. With small sample sizes (< 50), the empirical distribution yields re-order points that are significantly lower than s^* and therefore will have the difficulty to achieve the fill rate targets, especially for high fill rates.

In the case of the heavy-tailed folded t distribution, although none of the method works well with small sample sizes, EVT extrapolation still outperforms the other methods. The re-order points determined by EVT method converge to s^* much faster than the empirical distribution. The tails of exponential and normal distributions seem to be too light to fit the heavy tails like this t distribution.

tfr(β)	s^*	sample size	s determined by means of			
			EVT tail	EXP tail	Normal tail	Empirical
0.90	14.3158	$N = 10$	15.3830	28.4430	20.5430	13.3430
		$N = 20$	15.0843	27.5843	20.1843	14.0804
		$N = 50$	15.1419	28.3219	21.0219	15.4178
		$N = 100$	14.4888	27.7088	20.6688	14.8173
0.95	15.6921	$N = 10$	16.1324	32.1324	21.3924	13.4475
		$N = 20$	16.0324	32.4004	21.4204	14.4089
		$N = 50$	16.0056	32.1056	22.5856	16.5912
		$N = 100$	15.9391	32.7991	22.7791	16.5673
0.99	18.5034	$N = 10$	18.2137	38.6537	24.4937	13.3124
		$N = 20$	19.0852	41.2652	25.9252	15.5480
		$N = 50$	18.7996	42.8596	25.8396	17.3837
		$N = 100$	18.7902	43.2702	26.5102	18.0480

Table 3.1: Simulation results of X is folded normally distributed with location parameter 6 and scale parameter 5.

In conclusion, EVT extrapolation method performs as expected better than the other two extrapolation methods and the empirical distribution. To verify whether EVT extrapolation method indeed improves the empirical method for inventory control, we will present an extensive simulation study of an inventory control system in the next chapter.

tfr(β)	s^*	sample size	s determined by means of			
			EVT tail	EXP tail	Normal tail	Empirical
0.95	21.6333	$N = 10$	16.1324	32.1324	21.3924	13.4475
		$N = 20$	17.6820	28.9220	13.6420	11.6678
		$N = 50$	20.8853	30.9653	15.7253	15.3164
		$N = 100$	21.6984	30.8584	15.6984	16.7761
0.97	34.3646	$N = 10$	17.0450	28.9450	14.3050	10.7901
		$N = 20$	21.9048	31.5448	15.4848	13.5296
		$N = 50$	25.2633	34.7633	16.9033	15.5948
		$N = 100$	29.2727	35.9727	19.6727	19.9879
		$N = 500$	35.0703	40.01403	28.53403	30.4852
0.99	100.3537	$N = 10$	15.0976	28.3376	13.8376	9.5258
		$N = 20$	18.9780	32.0180	15.4380	11.7195
		$N = 50$	32.7535	37.3335	17.9535	16.4519
		$N = 100$	68.4903	49.2703	30.3103	29.7746
		$N = 500$	81.9582	57.1382	46.1782	49.1690
		$N = 1000$	101.9768	62.3368	56.09680	62.3753

Table 3.2: Simulation results of X is folded non-standard t distributed with 2 degree of freedom and location parameter 6.

Chapter 4

Simulation study

To compare the performance of the tail extrapolation methods in the context of our original inventory problem, we perform an extensive simulation study of an inventory control system. This chapter consists of two parts. In the first part, the simulation set-up will be explained. In the second part, the simulation results will be shown and discussed.

4.1 Simulation set-up

We apply an (s, nQ) inventory control policy with a daily review period and backordering, where the duration of replenishment lead time L is assumed to be constant. This policy means that if the inventory position IP drops to s or below, where

$$\text{IP} = \text{stock on hand} + \text{outstanding orders - back orders}, \tag{4.1}$$

an order is placed of size nQ where

$$n = \min\{n \in \mathbb{N} : \text{IP} + nQ > s\}, \tag{4.2}$$

and $Q \in \mathbb{N}$ is the minimum order quantity (MOQ). Simulation of this inventory system is performed for two types of demands, frequent demand and intermittent demand.

4.1.1 Type I: frequent demand

For the first type of frequent product demands, a non-negative random number will be generated as the daily demand according to a specified probability distribution in the beginning of each day. If this daily demand exceeds the current stock level, the excess amount will be regarded as back orders, and the total back orders will be updated with this amount.

To obtain a sample of LTD for further analysis, we apply the empirical method proposed by Porras and Dekker (2008). Let D be the vector containing all the daily demands up to this day. A window of size L will be placed over D, which means that the demand on day 1 to day L is summed up to obtain the first lead time demand. The window moves one period (day) at a time until the end of the vector D is reached. For example, if the vector D has 100 elements, the sample of the lead time demands will have a size of $100 - L + 1$. This sample of LTD will then be used to determine the re-order points s for a given target fill rate β, where we apply the methods described in Chapter 3. Recall the re-order point s is chosen to be the smallest s satisfying:

$$\beta \leq 1 - \frac{ES(s)}{Q}.$$

In this simulation study, Q, or the MOQ, is chosen to be the maximum of 1 and the smallest positive difference between the sample values of LTD, in other words, the step size ε in (3.7). That is,

$$Q := \max(1, \varepsilon).$$

The re-order point is updated at the beginning of each review period, thus each day. Then the total inventory holding costs will be updated which is the sum of the current total holding costs and the current stock level. If there is an order arriving, the stock level will be updated and the back orders will be delivered. At the end of each day, an order will be placed if the inventory position drops at s or below.

Each simulation run considers a period of N days. After a warm-up period of $10 + L$ days, the starting stock is set at $s + Q$. In the end of each simulation run, statistics like the total inventory holding costs and achieved fill rate are collected for further analysis. The procedure of one simulation run can be summarized by Algorithm 2.

4.1.2 Type II: intermittent demand

For the second type of intermittent demand, we only adapt some slight change in the simulation model of Type I demand.

To take the intermittent feature of demand into account, we first generate Poisson distributed random numbers which indicate the number of days between each two consecutive non-zero demands and these demands will be generated according to the specified probability distribution. The rest of the procedures remains the same.

4.1.3 Set-up

For both types of demand, we have performed the simulation with a constant lead time of 10 days and a simulation period of 70 days. Such a short simulation

Algorithm 2 Simulation inventory control

Require: L : the lead time
Q : the minimum order quantity
N : the number of simulation runs
β : the target fill rate

Output: H : total inventory holding cost
FR : the achieved fill rate

for $t = 1 \to 10 + L$ **do**
    generate a daily demand D_t according to a specified distribution
end for
determine the reorder point s using all the past demands D ▷ use Chapter 3
set current inventory level x to $s + Q$
for $t = 10 + L + 1 \to N$ **do**
    update $H = H + x$
    if order arrival on day $t =$ **true** **then**
        update inventory level
        fulfill and update back orders (if there is any)
    end if
    generate a daily demand D_t according to a specified distribution
    update inventory level x
    if stockout occurs **then**
        update back orders
        update total back orders
    end if
    calculate IP by (4.1)
    if new order should be placed **then**
        determine n to place an order of quantity nQ (4.2)
        update order lists
    end if
    update the re-order point s using all the past demand up to t ▷ use Chapter 3
end for
return FR = 1 − total back orders/total demands

period is chosen since the lack of data is a typical problem in practice when one deals with products having intermittent demands. We want stay with the practice as far as possible. Moreover, the simulation has been replicated for 5000 times for each method and different daily demand distributions. The considered daily demand distributions are:

1. Folded Normal distribution with location parameter 2 and scale parameter 1.
2. Folded t distribution with 2 degrees of freedom and location parameter 2
3. Folded Slash distributed with location parameter 2 and scale parameter 1.

The choice of folded distributions is again because of the non-negative product demands.

Furthermore, it is worth to mention how we choose the parameter k when applying the tail extrapolation using EVT. Due to the large number of replicates of the simulation, manually choosing the parameter k by means of a diagram of estimates each time when we determine the re-order point is not realistic. Therefore, we have pre-specified to choose k as follows:

	sample size N						
	≤ 15	$16-20$	$21-25$	$26-30$	$31-40$	$41-50$	$51-60$
$k:$	$\lceil 0.8N \rceil$	$\lceil 0.75N \rceil$	$\lceil 0.7N \rceil$	$\lceil 0.65N \rceil$	$\lceil 0.6N \rceil$	$\lceil 0.5N \rceil$	$\lceil 0.4N \rceil$

It is important to keep this in mind since it implies that we are not applying EVT optimally.

4.2 Results

In all cases, we compare the performance of the extrapolation methods by establishing trade-off curves between inventory holding costs and achieved fill rates, and the achieved fill rates against target fill rates. Sub-figures (a) show the the average total holding costs as a function of the achieved fill rate, while sub-figures (b) show the the achieved fill rates as a function of the target fill rate. This gives better insights than comparing the holding costs for each achieved fill rate. Besides, it is very hard to get exactly the same achieved fill rate for all methods. Note that we have only considered high target fill rates, since our purpose is to tackle the issue that the empirical method has difficulties obtaining high fill rates.

Figure 4.1 and 4.2 show the simulation results for the folded normally distributed daily demands for both types demands respectively. Figure 4.3 and 4.4 show the results for the folded t distribution, and Figure 4.5 and 4.6 exhibit the results for the folded Slash distribution. It is noteworthy that the performance of the methods does not significantly differ between Type I and Type II demands in the sense that if a method does not perform well for one of the type demands, it does not work for the other Type demands either.

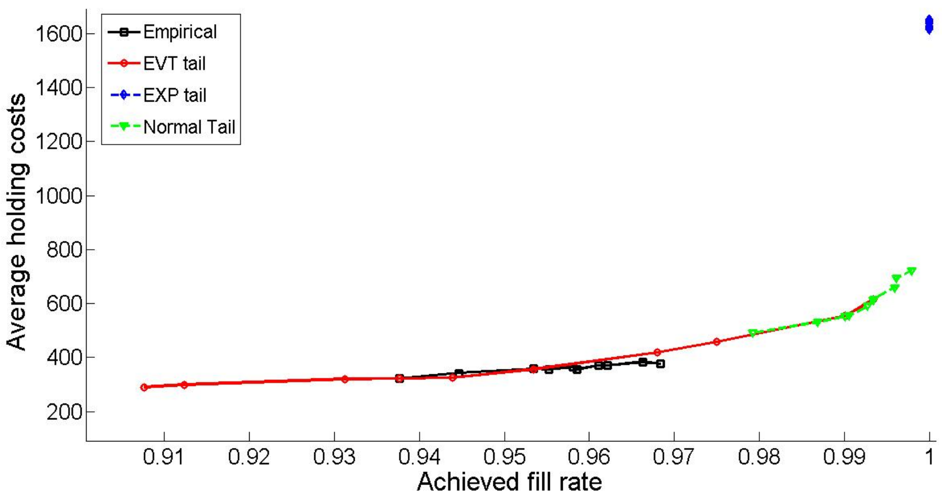
The empirical distribution has indeed difficulties obtaining high fill rates in all cases, even for the light-tailed folded normal distribution, not to mention the other two distributions with heavier tails.

The extrapolation method using exponential tails appear to overestimate for the folded normal distribution as well as the folded t distribution. The fitted exponential distributions have probably too heavy tails. However, the tails of the fitted exponential distributions seem to be too light for the heavy-tailed folded Slash distribution.

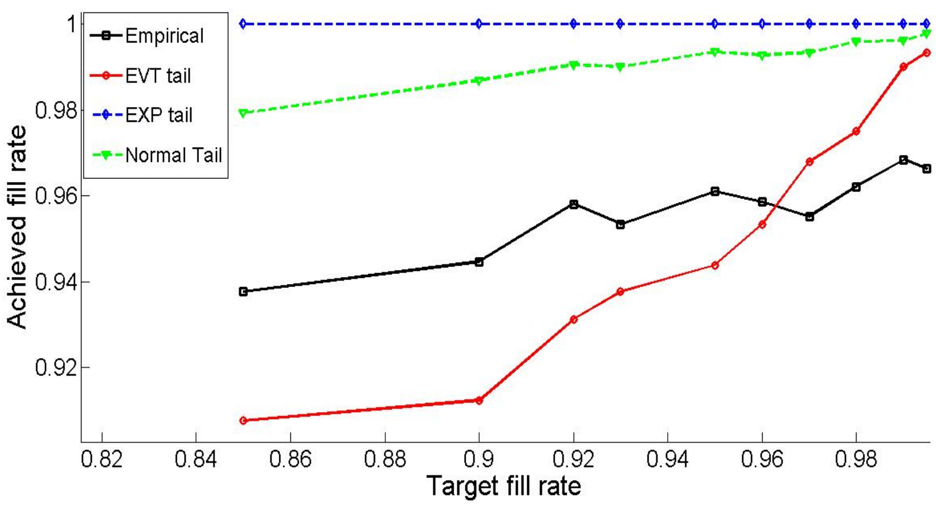
The extrapolation method using normal tails does not perform well either. In the cases of folded normal and t distribution, this method yields achieved fill rates that are much higher than the target fill rates for relatively low targets (≤ 0.92). However, for really high targets (> 0.96), the achieved fill rates are far below the target. In the cases of the heavy-tailed folded Slash distribution, the fitted normal distribution is also too light to obtain high fill rates.

In all cases, the extrapolation method applying EVT outperforms the other two extrapolation methods and performs much better than the empirical distribution. For the folded normal and t distribution, this method yields achieved fill rates that are very close to the target fill rates: for relatively low targets (≤ 0.93), the achieved target fill rates are only slightly higher than the their target; for high targets (≥ 0.95), the achieved target fill rates are just below their target. In the cases of the heavy-tailed folded Slash distribution, although the EVT extrapolation method works better than the other methods, it also has difficulties to obtain high fill rates. This can be explained by the fact that it is unrealistic to capture the tail behavior of heavy-tailed distributions with such few data. After the warm-up period in each simulation run, we start only with a sample of size 11 to predict the LTD. One cannot expect an accurate demand forecast with this sample size if the demand follows a heavy-tailed distribution. Even if the method starts to work better as the sample size increases, the overall achieved fill rate remains low. Besides, recall that we are not applying this method optimally due to the pre-specified parameter k.

In summary, the simulation results have demonstrated that the extrapolation method applying EVT performs the best among the proposed methods and indeed can improve the empirical method in inventory decision making for products with both frequent and intermittent demands.

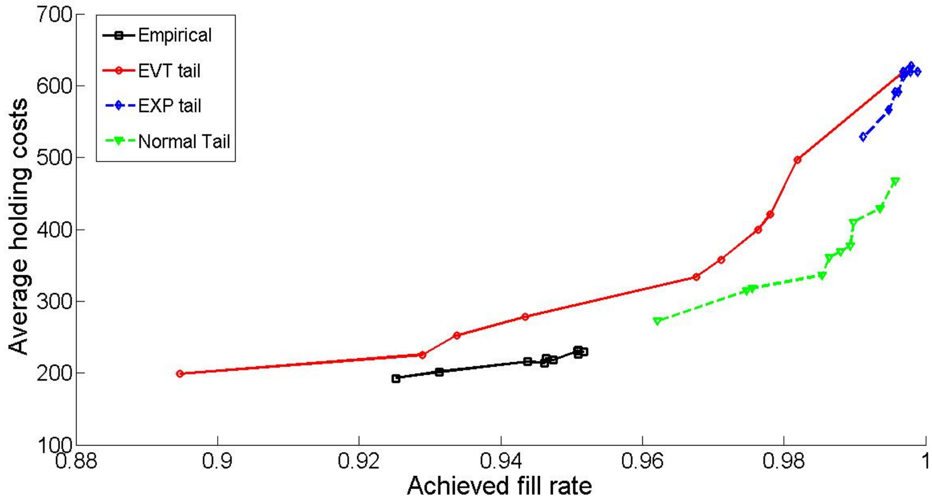


(a)

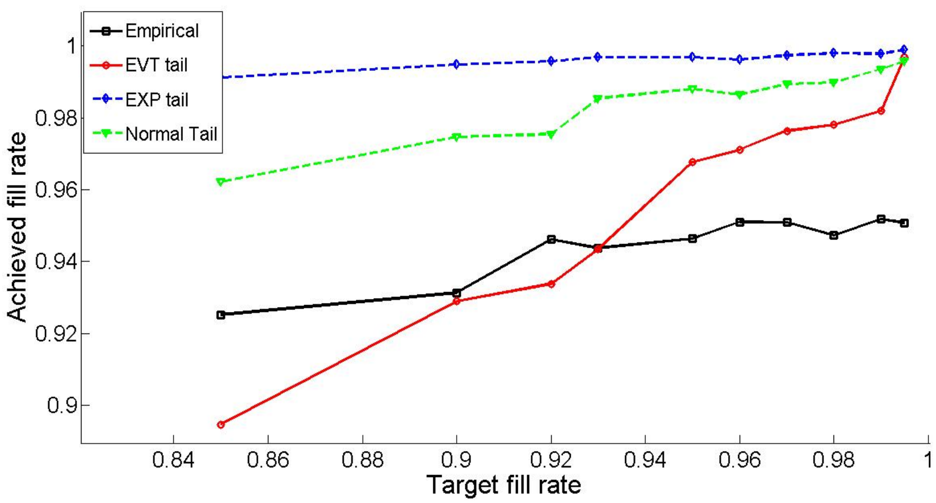


(b)

Figure 4.1: Type I Simulation results for folded normally distributed daily demands with location parameter 2 and scale parameter 1.

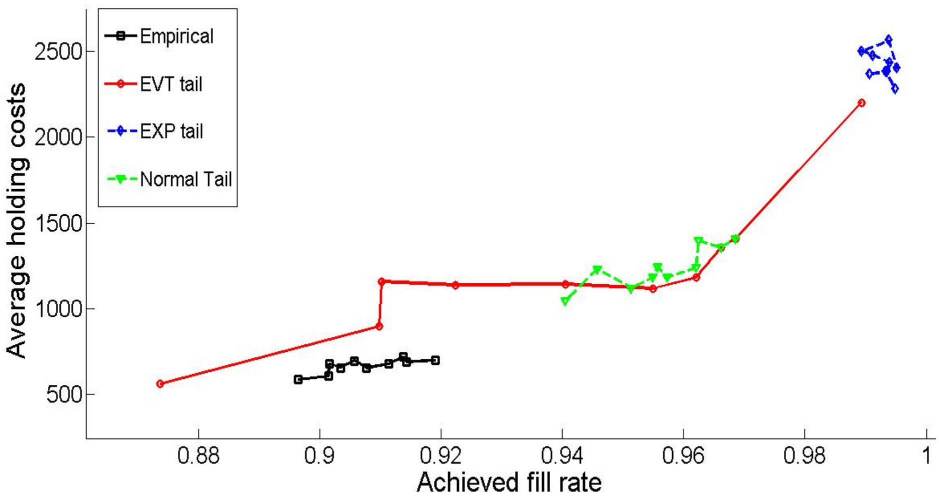


(a)

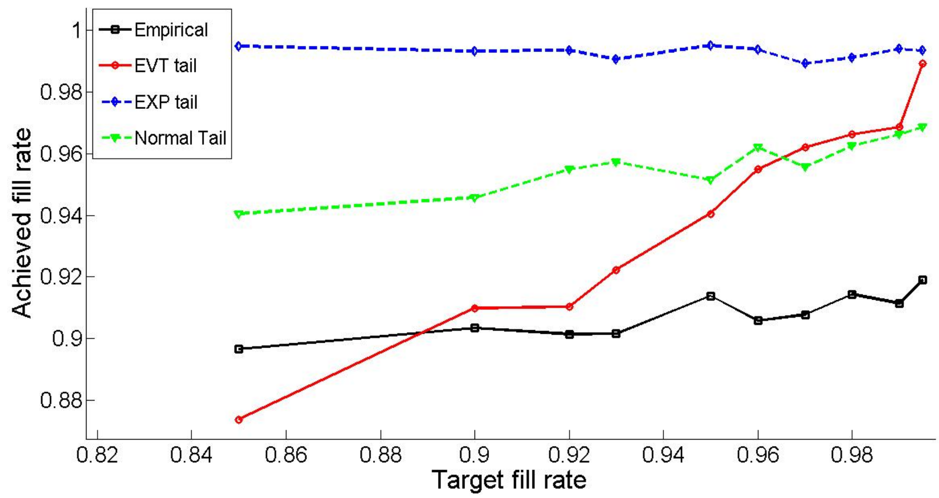


(b)

Figure 4.2: Type II Simulation results for folded normally distributed daily demands with location parameter 2 and scale parameter 1.

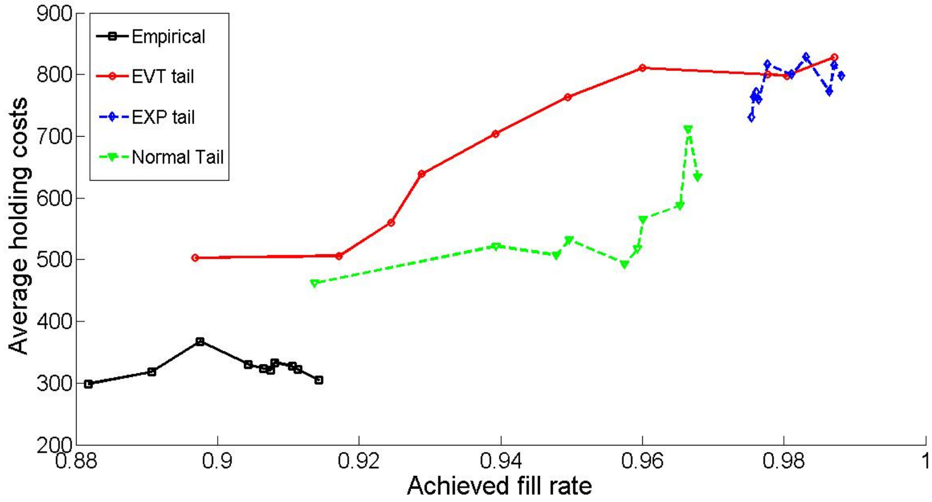


(a)

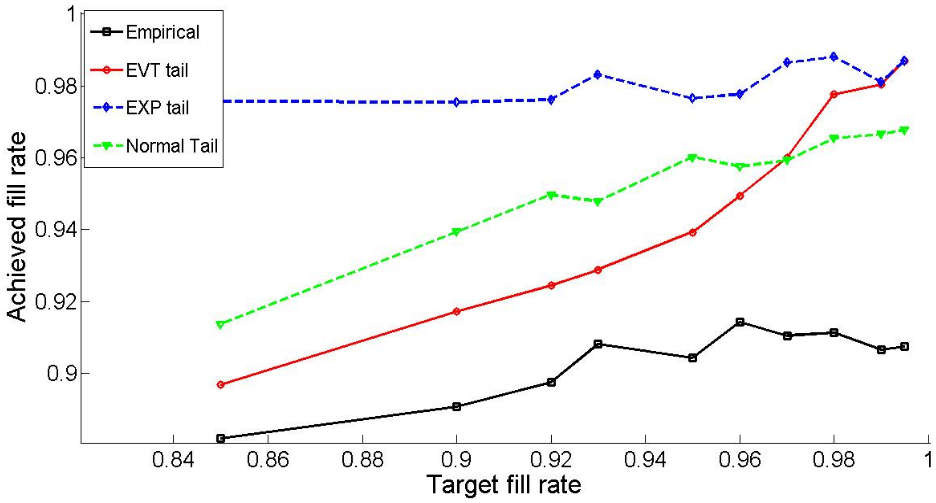


(b)

Figure 4.3: Type I Simulation results for folded t distributed daily demands with 2 degrees of freedom and location parameter 2.

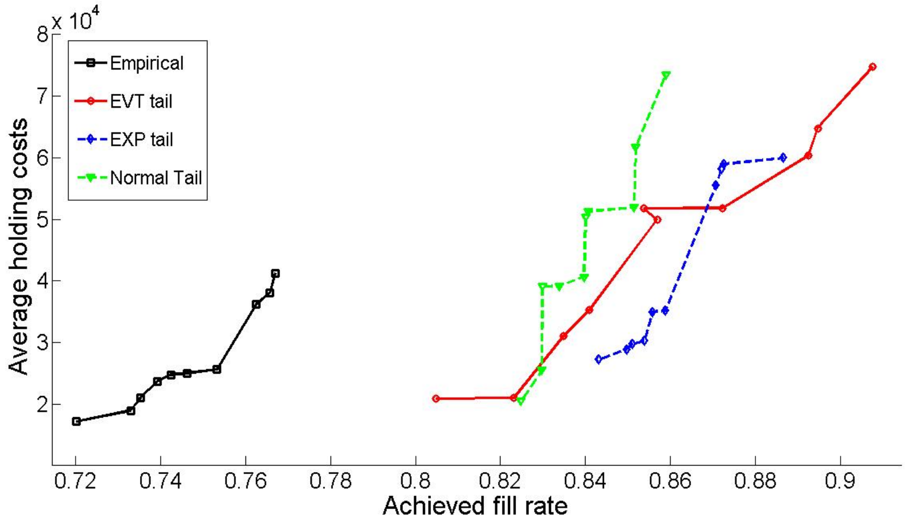


(a)

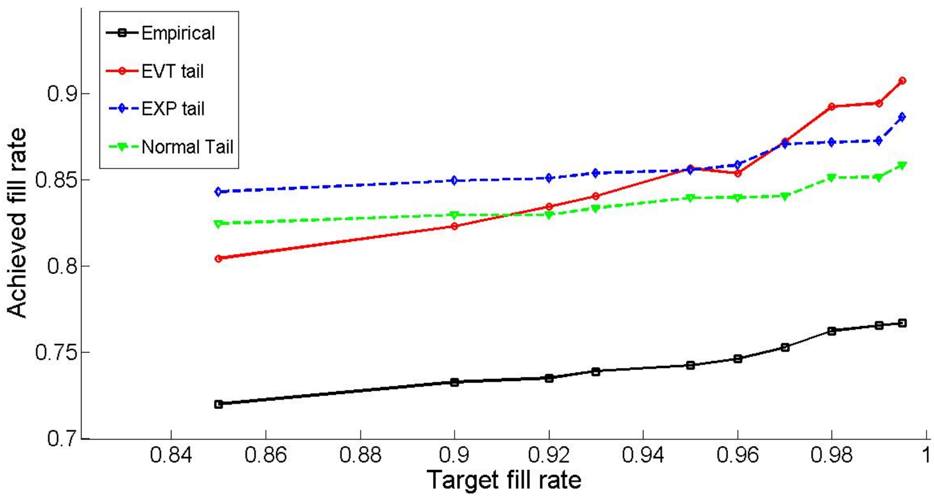


(b)

Figure 4.4: Type II Simulation results for folded t distributed daily demands with 2 degrees of freedom and location parameter 2.

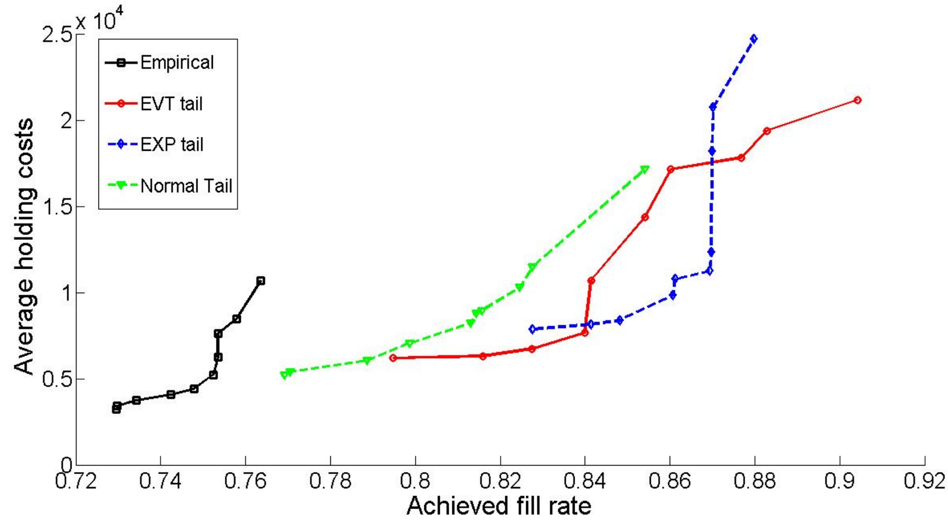


(a)

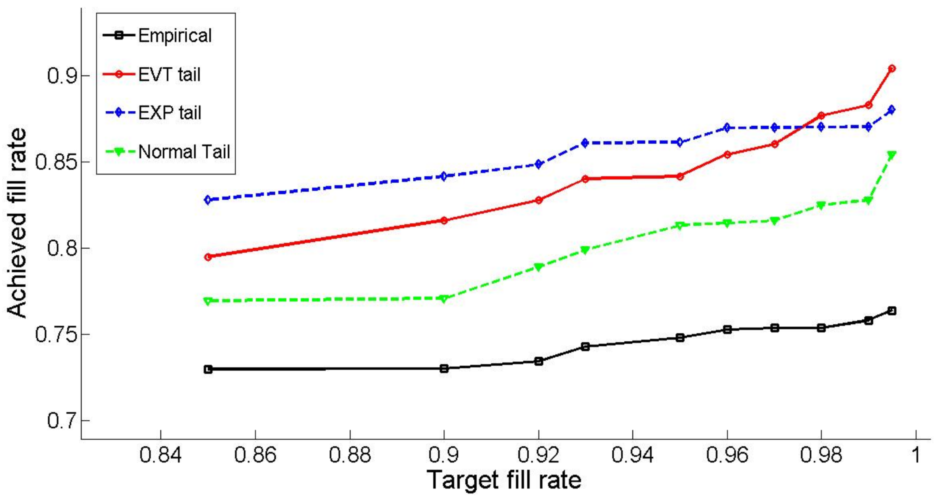


(b)

Figure 4.5: Type I Simulation results for folded Slash distributed daily demands with location parameter 2 and scale parameter 1.



(a)



(b)

Figure 4.6: Type II Simulation results for folded Slash distributed daily demands with location parameter 2 and scale parameter 1.

Chapter 5

Conclusions and limitations

In this thesis, we have introduced and applied the extreme value theory (EVT) to present a tail extrapolation method which can be used to improve the empirical method that is proposed by Porras and Dekker (2008) in the inventory management for products with intermittent demands. This tail extrapolation method is compared with two alternative extrapolation methods based on exponential distribution and normal distribution tails.

The performance of these tail extrapolations methods are assessed by re-order points simulations in Section 3.5 and the extensive inventory control system simulations in Chapter 4. All the simulation results have demonstrated that the EVT tail extrapolation method outperforms the other two methods and successfully extrapolates the empirical distribution. The results of the inventory control system simulations showed.

It is also noteworthy that although EVT is an asymptotic theory, our extrapolation method applying EVT still performs well in almost all the cases, even for a very small sample of LTD (with a sample size of 10). This is a good news for us, since the lack of data is a typical problem in practice when one deals with products having intermittent demands. If the EVT extrapolation method is applied in practice, it will yield better results than the current empirical method with only few data. With the lapse of time, more data can be collected, this extrapolation method will certainly perform even better.

In conclusion, the extrapolation method applying EVT successfully extrapolates the empirical distribution and therefore can be used to improve the empirical method for inventory control of products with intermittent demands.

There are a few limitations in our research. The simulation study in Chapter 4 has not demonstrated the optimal performance of the EVT extrapolation method due to the choice of k. In practice, one should determine k manually by means of the diagram of estimates. In addition, it would be more convincing if one can assess the performance of the proposed extrapolation methods by means of real industry data. Moreover, the proposed extrapolation methods assume that the LTD follows a particular probability distribution and neglect other possibilities, e.g. the LTD may follow a stochastic process. Last but not least, this thesis only considers a constant lead time while in practice the lead time is often stochastic. One should consider bootstrapping methods to handle stochastic lead times. Nevertheless, the extrapolation method applying EVT can contribute to the bootstrapping methods. Note that we can condition the LTD on the lead times, and for each fixed lead time, we can use the empirical distribution applying the EVT extrapolation method.

Bibliography

Basten, R., van Wingerden, E., Dekker, R., and Rustenburg, W. (2012). More grip on inventory control through improved forecasting: A comparative study at three companies. Working paper, Eindhoven University of Technology.

Casella, G. and Berger, R. L. (2001). Duxbury Press.

de Haan, L. and Ferreira, A. (2006). Springer Series in Operations Research and Financial Engineering. Springer.

Falk, M. (1995). Some best estimators for distributions with finite endpoint. *Statistics*, pages 115–125.

Fisher, R. and Tippett, L. (1928). Limiting forms of the frequency distribution of the largest and smallest member of a sample. *Proc. Cambridge Philosophical Society*, pages 180–190.

Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum. *Annales de la Société Polonaise de Mathématique*, pages 93–116.

Gnedenko, B. (1943). Sur la distributiion limite du terme maximum d'une série aléatoire. *Annals of Mathematics*, pages 423–453.

Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, pages 331–341.

Hosking, J. and Wallis, J. (1987). Parameter and quantile estimation for the generalized pareto distribution. *Technometrics*, pages 339–349.

Jones, M. C. (2008). On a class of distributions with simple exponential tails. *Statistica Sinica*, 18(3):1101–1110.

Leone, F. C., Nelson, L. S., and Nottingham, R. B. (1961). The folded normal distribution. *Technometrics*, pages 543–550.

Pickands, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics*, pages 119–131.

Porras, E. and Dekker, R. (2008). An inventory control system for spare parts at a refinery: An empirical comparison of different re-order point methods. *European Journal of Operational Research*, 184:101–132.

Psarakis, S. and Panaretos, J. (1990). The folded t distribution. *Communications in Statistics: A Theory and Methods*, (7):2717–2734.

Rogers, W. H. and Tukey, J. W. (1972). Understanding some long-tailed symmetrical distributions. *Statistica Neerlandica*, pages 211–226.

Appendix A

Derivation section 2

A.1 Tail extrapolation using EVT

To calculate the expected shortage given the threshold s, we first rewrite $ES(s)$ as

$$\begin{aligned} ES(s) &= \mathbb{E}\left[(X-s)\mathbf{1}_{\{X>s\}}\right] \\ &= \int_s^{x^*} (x-s)\,\mathrm{d}F(x) \\ &= \int_s^{x^*}\int_s^{x} \mathrm{d}u\,\mathrm{d}F(x) \\ &\overset{\text{(Fubini)}}{=} \int_s^{x^*}\int_u^{x^*} \mathrm{d}F(x)\,\mathrm{d}u \\ &= \int_s^{x^*} (1-F(u))\,\mathrm{d}u, \end{aligned} \tag{A.1}$$

where F is the distribution function of X and $x^* := \sup\{x : F(x) < 1\}$.

Relation (2.5) suggests that we can approximate (A.1) by

$$\frac{k}{n}\int_s^{x^*} \left\{1 - G_\gamma\left(\frac{x - X_{n-k,n}}{\alpha(\frac{n}{k})}\right)\right\}\mathrm{d}x, \quad \forall\, s \geq X_{n-k,n}. \tag{A.2}$$

For $0 < \gamma < 1$, $x^* = \infty$ by Theorem 1 and (A.2) equals to

$$\frac{k}{n}\int_s^{\infty}\left\{1+\gamma\left(\frac{x-X_{n-k,n}}{\alpha(\frac{n}{k})}\right)\right\}^{-1/\gamma}\,\mathrm{d}x$$

$$= \frac{k}{n}\cdot\frac{\alpha(\frac{n}{k})}{1-\gamma}\left\{1+\gamma\left(\frac{x-X_{n-k,n}}{\alpha(\frac{n}{k})}\right)\right\}^{1-1/\gamma}\Bigg|_s^{\infty}$$

$$= \frac{k}{n}\cdot\frac{\alpha(\frac{n}{k})}{1-\gamma}\left(1+\gamma\left(\frac{s-X_{n-k,n}}{\alpha(\frac{n}{k})}\right)\right)^{1-1/\gamma}.$$

For $\gamma = 0$ and assuming $x^* = \infty$, (A.2) equals to

$$\frac{k}{n}\int_s^{\infty}\exp\left(\frac{-x+X_{n-k,n}}{\alpha(\frac{n}{k})}\right)\,\mathrm{d}x$$

$$= \frac{k}{n}\cdot\left(-\alpha\left(\frac{n}{k}\right)\right)\exp\left(\frac{-x+X_{n-k,n}}{\alpha(\frac{n}{k})}\right)\Bigg|_s^{\infty}$$

$$= \frac{k}{n}\cdot\alpha\left(\frac{n}{k}\right)\exp\left(\frac{-s+X_{n-k,n}}{\alpha(\frac{n}{k})}\right).$$

For $\gamma < 0$, x^* is finite by Theorem 1 and (A.2) equals to

$$\frac{k}{n}\int_s^{x^*}\left\{1+\gamma\left(\frac{x-X_{n-k,n}}{\alpha(\frac{n}{k})}\right)\right\}^{-1/\gamma}\,\mathrm{d}x$$

$$= \frac{k}{n}\cdot\frac{\alpha(\frac{n}{k})}{1-\gamma}\left\{1+\gamma\left(\frac{x-X_{n-k,n}}{\alpha(\frac{n}{k})}\right)\right\}^{1-1/\gamma}\Bigg|_s^{x^*}$$

$$= \frac{k}{n}\cdot\frac{\alpha(\frac{n}{k})}{1-\gamma}\left\{\left(1+\gamma\left(\frac{s-X_{n-k,n}}{\alpha(\frac{n}{k})}\right)\right)^{1-1/\gamma} - \left(1+\gamma\left(\frac{x^*-X_{n-k,n}}{\alpha(\frac{n}{k})}\right)\right)^{1-1/\gamma}\right\}.$$

For $\gamma \geq 1$, it is clearly that (A.2) goes to infinite.
In summary, with suitable estimators $\hat{\gamma}$, $\hat{x}^*$ and $\hat{\alpha}(\frac{n}{k})$, we can estimate the

expected shortage for all $s \geq X_{n-k,n}$ as

$$\widehat{ES}_{tail}(s) = \begin{cases} \frac{k}{n} \cdot \frac{\hat{\alpha}(\frac{n}{k})}{1-\hat{\gamma}} \left(1 + \hat{\gamma}\left(\frac{s-X_{n-k,n}}{\hat{\alpha}(\frac{n}{k})}\right)\right)^{1-1/\hat{\gamma}} & 0 < \hat{\gamma} < 1, \\ \frac{k}{n} \cdot \hat{\alpha}\left(\frac{n}{k}\right) \exp\left(\frac{-s+X_{n-k,n}}{\hat{\alpha}(\frac{n}{k})}\right) & \hat{\gamma} = 0, \\ \frac{k}{n} \cdot \frac{\hat{\alpha}(\frac{n}{k})}{1-\hat{\gamma}} \left\{\left(1 + \hat{\gamma}\left(\frac{s-X_{n-k,n}}{\hat{\alpha}(\frac{n}{k})}\right)\right)^{1-1/\hat{\gamma}} \right. & \\ \left. - \left(1 + \hat{\gamma}\left(\frac{\hat{x}^*-X_{n-k,n}}{\hat{\alpha}(\frac{n}{k})}\right)\right)^{1-1/\hat{\gamma}}\right\} & \hat{\gamma} < 0. \end{cases}$$

A.2 Tail fitting by exponential distribution

Suppose $X \sim \mathrm{EXP}(\mu)$, we have

$$\begin{aligned} \int_s^\infty (x-s)f(x)\,\mathrm{d}x &\overset{(A.1)}{=} \int_s^\infty (1-F(x))\,\mathrm{d}x \\ &= \int_s^\infty e^{-\frac{x}{\mu}}\,\mathrm{d}x \\ &= -\mu e^{-\frac{x}{\mu}}\Big|_s^\infty \\ &= \mu e^{-\frac{s}{\mu}}. \end{aligned}$$

Using the Maximum-Likelihood estimator $\hat{\mu}$, we can estimate the expected shortage as

$$\widehat{ES}_{tail}(s) = \hat{\mu}\exp\left(-\frac{s}{\hat{\mu}}\right).$$

A.3 Tail fitting by normal distribution

Suppose $X \sim \mathcal{N}(\mu, \sigma)$, we have

$$\begin{aligned} \int_s^\infty (x-s)f(x)\,\mathrm{d}x &= \int_s^\infty (x-\mu)f(x)\,\mathrm{d}x + \int_s^\infty (\mu-s)f(x)\,\mathrm{d}x \\ &= \int_s^\infty (x-\mu)\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\mathrm{d}x + (\mu-s)\mathbb{P}(X>s) \\ &\overset{(t=x-\mu)}{=} \int_{s-\mu}^\infty t\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{t^2}{2\sigma^2}\right)\mathrm{d}t + (\mu-s)\mathbb{P}\left(Z > \frac{s-\mu}{\sigma}\right) \end{aligned}$$

$$
\begin{aligned}
&= \frac{\sigma}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2\sigma^2}\right)\Bigg|_{s-\mu}^{\infty} + (\mu - s)\left(1 - \Phi\left(\frac{s-\mu}{\sigma}\right)\right) \\
&= \frac{\sigma}{\sqrt{2\pi}} \exp\left(-\frac{(s-\mu)^2}{2\sigma^2}\right) + (\mu - s)\left(1 - \Phi\left(\frac{s-\mu}{\sigma}\right)\right).
\end{aligned}
$$

Using the Maximum-Likelihood estimators $\hat{\mu}$ and $\hat{\sigma}$ of the sample, we can estimate the expected shortage as

$$
\widehat{ES}_{tail}(s) = \frac{\hat{\sigma}}{\sqrt{2\pi}} \exp\left(-\frac{(s-\hat{\mu})^2}{2\hat{\sigma}^2}\right) + (\hat{\mu} - s)\left(1 - \Phi\left(\frac{s-\hat{\mu}}{\hat{\sigma}}\right)\right).
$$

Appendix B

Folded distributions

The folded distributions are often used when the measurement system produces only non-negative measurements. In this Appendix, we present the density function and the random number generation of some folded distributions.

B.1 Folded normal distribution

The folded normal distribution is proposed by Leone et al. (1961).

Let $X \sim \mathcal{N}(\mu, \sigma)$, then its absolute value $Y = |X|$ will have the following density function

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma}\left\{\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) + \exp\left(-\frac{(-y-\mu)^2}{2\sigma^2}\right)\right\}, \quad \forall\, y \geq 0.$$

The random variable Y is said to have location parameter μ and scale parameter σ.

It is very easy to generate a folded normal distributed random variable with location parameter μ and scale parameter σ, that is just taking the absolute value of a generated normally distributed random variable with parameters μ and σ.

B.2 Folded non-standard t distribution

The folded non-standard t distribution is first thoroughly studied by Psarakis and Panaretos (1990).

Let T be a random variable having the non-standard t distribution with ν degrees of freedom and location parameter μ as defined by the density function

$$f_T(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}}\left(1 + \frac{(x-\mu)^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad \forall\, \nu \in \mathbb{N}, x \in \mathbb{R}.$$

Let W be the folded random variable of T, i.e. $W = |T|$. The density function of W is given by

$$f_W(w) = f_T(w) + f_T(-w), \quad \forall\, w \geq 0.$$

To generate a folded non-standard t distributed random variable with ν degrees of freedom and location parameter μ, we follow the following steps:

1. Generate a standard t distributed random number with ν degrees of freedom, say T_1. This can be done by using for example the built-in MATLAB function `trnd` with input ν.

2. Calculate $T_2 = T_1 + \mu$.

3. Taking the absolute value of T_2, we get the desired random variable.

B.3 Folded Slash distribution

The Slash distribution was introduced by Rogers and Tukey (1972). It is a less famous probability distribution, which can be obtained by a standard normal variate divided by an independent standard uniform variate. In other words, if the random variable Z has a normal distribution with zero mean and unit variance, the random variable U has a uniform distribution on $[0, 1]$ and Z and U are statistically independent, then the random variable $X = Z/U$ has a slash distribution. The slash distribution is hence an example of a ratio distribution. Another example of a ratio distribution is the Cauchy distribution.

The tail of the Slash distribution is heavier than the tail of the normal and exponential distribution, but lighter than the tail of the Cauchy distribution. It is typically known to have a long tail.

A Slash distributed Y with location parameter μ and scale parameter σ can be obtained by the ratio

$$Y = \frac{X}{U},$$

where X has a folded normal distribution with location parameter μ and scale parameter σ and U has a uniform distribution on $[0, 1]$, and X and U are independent.

We also use this relation to generate Slash distributed random variables.