CONCORDIA UNIVERSITY

DEPARTMENT OF ECONOMICS

The Behaviour of Extreme Value Theory (EVT) in calculating Value at Risk (VaR) in Small Samples

M.A. RESEARCH PAPER

Student's Name: Mohamad K. Moughnieh
Student ID: 27649762
Date Submitted: 06/02/2017
Supervisor: Dr. Bryan Campbell

Department of Economics
Signature page

This is to certify that the M.A. Research Paper prepared

By: ____ ID#: _________________

Entitled:
____

and submitted in partial fulfillment of the requirements for the degree of
Master of Arts (Economics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Supervisor: ____
Date: _________________________

Examiner: ____
Date: _________________________

Approved by Graduate Program Director: ______________________________
Date: _________________________

Grade Sheet prepared by Graduate Program Assistant (given to Supervisor):
Date: ____________________

Acknowledgement

I would like to express my gratitude to my advisor Dr. Bryan Campbell for his exceptional help and advice during this semester. I am also grateful to the Graduate Program Director Dr. Szilvia Papai and her assistant Elise Melancon for their valuable suggestions and support during the MA program.

Table of Contents

Abstract

In this paper I raise the question “How well does extreme value theory EVT work for calculating value at risk (VaR) in small samples” In an attempt to answer this question, I estimate the mean VaR from 1000 EVT estimates using maximum log likelihood. I Take 3 samples of 400, 600 and 800 from three different distributions: Normal Distribution N (0, 16), Log-Normal Distribution and T-Distribution with degrees of freedom equals 4. For each sample I estimate the EVT values for each distribution using the 99 and 99.9 percentiles.

I found that VaR estimate tends to be more accurate for normal distribution with confidence interval 99% as the sample size increases from 400 to 800. And that for the three different distributions, VaR estimates are positively skewed.

I. Introduction

A significant role of a financial institution is to evaluate market risk which arises from fluctuations in prices of equities, interest rates and exchange rates. One of the most important ways of assessing market risk is to evaluate losses that occur due to a decrease in the portfolio's assets. A commonly used methodology for the estimation of market risk is value at risk (VaR). VaR has become a crucial risk management tool after the Basel Committee at the bank of International Settlements called upon financial institutions to meet capital requirements based on VaR estimates (Abad and Benito, 2012). According to Khindanova, Rachev and Schwartz (2001), VaR is the highest possible loss during a certain period of time at a given confidence interval. In other words, if the daily VaR for a given portfolio is $2 million at the 95% confidence level, this means that a one day loss may exceed $2 million at 5% most of the time.

There are different approaches used to measure VaR. Hull (2015), argues that Historical simulation is the most popular approach for calculating value at risk (VaR) and expected shortfall (ES) for market risk. He defines expected shortfall as the conditional value at risk or expected tail loss that asks: "If things do get bad, what is the expected loss?" unlike value at risk (VaR) that asks "How bad can things get?" He further describes Historical simulation as an approach used to estimate the probability distribution of the change in the value of the current portfolio between today and tomorrow.

It is important to note that VaR doesn't take into account losses beyond a certain quantile in the distribution. However, expected shortfall ES also known as the expectation of returns given that they exceed VaR, takes into account losses beyond quantiles as shown in the figure below.

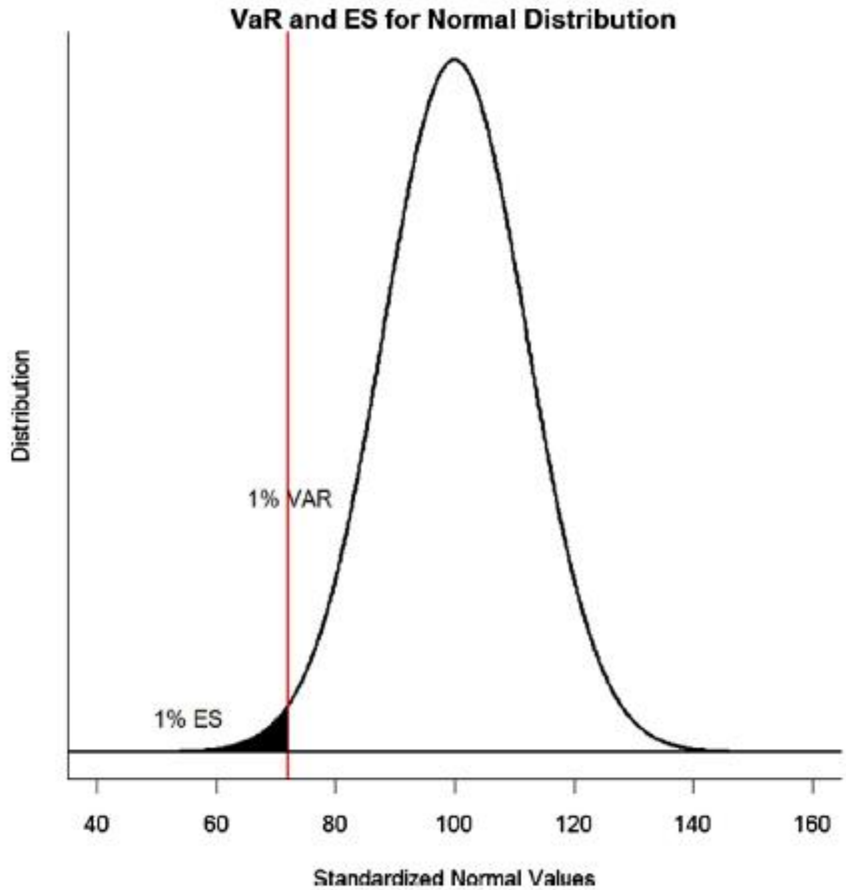


Fig. 2. Value at risk VaR and expected shortfall ES for normal distribution. Adapted from: Allen, D. E., Singh, A. K., & Powell, R. J. (2013). EVT and tail-risk modelling: Evidence from market indices and volatility series. *The North American Journal of Economics and Finance*, *26*, 355-369.

The graph shows 1% VaR for randomly generated normally distributed data. It can be observed that VaR doesn't take into account losses beyond the quantiles unlike expected shortfall as seen in the shaded region.

Yamai and Yoshiba (2002), define VaR as the sufficient capital needed to cover losses from a portfolio over a holding period of fixed days. For example, taking a random variable X with continuous distribution function F that models losses of a financial instrument over a certain period of time.

Thus, VaR is the *p*-th quantile of the distribution *F*

$$VaR_p = F^{-1}(1 - p)$$

where F^{-1} is the inverse of the distribution function *F*

They also explain expected shortfall as the tool used to measure the size of loss exceeding VaR.

$$ES_p = E(X|X > VaR_p)$$

Authors have used different VaR methods to measure portfolio risk. Hendrick (1996) studied three different VaR models: (1) historical simulation approach, (2) equally weighted variance-covariance approach and (3) exponentially weighted variance-covariance approach in an attempt to understand how the various VaR approaches behave and differ among themselves. He found that overall there is no approach that surpasses the other. The results showed that at 95 percent confidence level interval the VaR approaches measure the level of risk accurately however at 99 percent confidence level, the measure is less accurate. Similarly, Cabedo and Moya (2003) estimated VaR for oil risk prices using historical simulation approach, variance covariance approach and historical simulation with ARMA approach that differs from the historical simulation approach, in that it does not directly use the distribution of past returns, instead it uses the distribution of forecasting errors derived from ARMA model. They found that the historical simulation with ARMA approach provides a more accurate measurement than the two other approaches.

Many authors developed mixed approaches and argued that these approaches are effective in VaR estimation. Cvjetković et al (2016), developed an approach that combined historical simulation and extreme value theory to capture heteroscedasticity. They argue that this hybrid approach is suitable for estimating VaR and expected shortfall at higher quantiles. Boudoukh, Richardson and Whitelaw (1998) combined the exponential smoothing approach and historical simulation in one hybrid approach in an attempt to estimating and comparing Var under the three approaches. They concluded that the hybrid approach provides a more accurate VaR estimation than historical simulation and exponential smoothing separately.

In 2014 the Basel committee on banking supervision advised the financial institutions to shift from using value at risk VaR to using expected shortfall ES when estimating market risk. To implement the new changes, banks started using extreme value theory which looks at what occurs in the tail of distribution (The Economist, 2014). Thus the aim of this paper is to study the behaviour of extreme value theory EVT in the estimation of value at risk VaR in large and small samples.

In this paper I estimate the mean value at risk (VaR) from 1000 EVT estimates. I use a Monte Carlo simulation using three different random samples of 400, 600 and 800. The goal is to assess Var performance and compare its accuracy under each sample. The findings should provide a clear view on how the measurement of VaR is accurate for small sample size. This paper is organized into five main sections: In section 1, I provide an introduction and background information of extreme value theory (EVT) and other approaches that have been used to calculate both value at risk (VaR). In section 2, I provide the reader with a literature review that compiles concrete findings as well as theoretical and methodological contribution to the topic. Section 3 includes the methodology I'll be using and a discussion of the Monte Carlo simulation to be performed. In section 4, I'll present the reader with a profound explanation of the findings. In section 5, I'll give an overall conclusion of the work presented and summarize the main findings and provide suggestions for further research.

II. Literature Review

Extreme value theory (EVT) is a type of statistics that deals with extreme deviations from the median probability distribution. Its main goal is to identify the sample's maximum distribution or the distribution of values above a given threshold. For example given a data set of 1000 insurance claims and we chose the threshold u to be the 95^{th} with 50 points above u.

For each point $\{p_1, p_2, \dots, p_{50}\}$, the excess above u $\{p_1 - u, p_2 - u, \dots, p_{50} - u\}$ is calculated. These are described as random observations from a population with distribution of excesses. Moreover, the Pickands-Balkema-de Haan (PBH) Theorem suggest that for a very large family of distributions and large threshold value u, the distribution of excesses over u can be computed by a generalized Pareto distribution, GPD (Levine, 2009).

According to Dey and Yan (2016), EVT is extensively used in risk management, economics and finance to model events with low probability to occur and estimate their risks. Moreover, according to Ruffino (2011), extreme value theory approach is used to model extreme observations to obtain the distribution of the tails from the original distribution.

In the left side of Figure 1 below, the observations X_2, X_5, X_7 and X_{11} represent the block maxima for four periods of three observations each. On the other hand, the right side represents the extreme observations X_1, X_2, X_7, X_8, X_9 and X_{11} that exceed the threshold u.

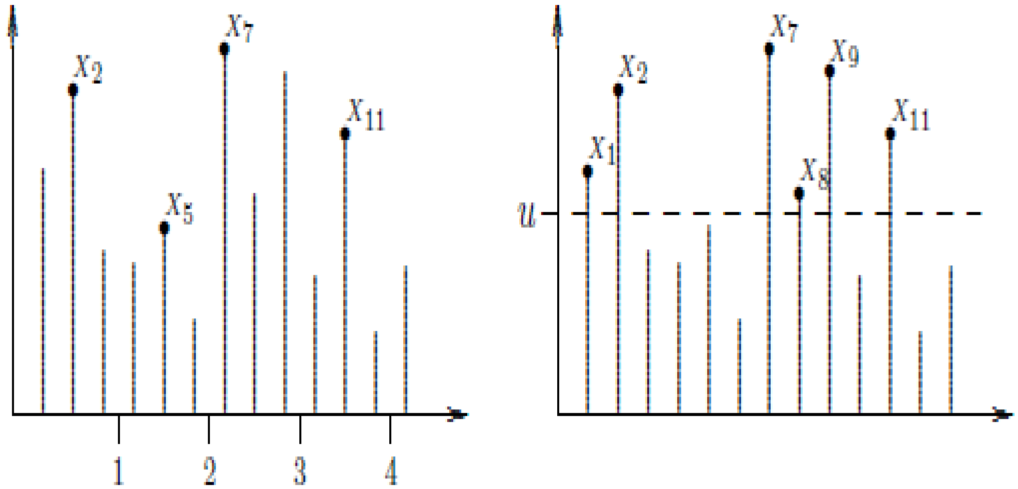


Fig. 2. Block-maxima (left panel) and excesses over a threshold *u* (right panel). Adapted from: Gilli, M. (2006). An application of extreme value theory for measuring financial risk. *Computational Economics*, *27*(2-3), 207-228

According to (Gilli, 2009), extreme value theory is similar to central limit theorem when modelling sum of random variables. They differentiate between two ways of identifying extremes in real data. The first recognizes the maximum assigned to the variable in successive periods such as months or years also known as block maxima that is used to analyze seasonal data such as hydrological data. The second approach considers the realization exceeding a specific threshold.

Moreover, Alen et al (2013), defines extreme value theory as the model that allows capturing of heavy tailed distributions and extreme tail risk. They present two EVT methods; the first method is based on extreme value distribution including the Gumbel, Frechet or Weibull distributions that are generalized as Generalized Extreme Value distribution (GEV) and fitted by block maxima. The second method is based on Generalized Pareto Distribution (GPD) which is known as the Peak Over Threshold (POT). They argue that the POT method uses data exceeding a threshold level and fitted to the GPD as represented in the lemma below:

For a large class of underlying distributions, the excess distribution function F_u can be approximated by GPD for an increasing threshold u.

$$F_u(y) \approx G_{\xi,\sigma}(y), \quad u \to \infty$$

Where $G_{\xi,\sigma}$ is the Generalized Pareto Distribution represented by:

$$G_{\xi,\sigma}(y) = \begin{cases} \left(1 + \frac{\xi}{\sigma} y\right)^{\frac{-1}{\xi}} \\ 1 - e^{\frac{-y}{\sigma}} \end{cases} if\ \xi \neq 0 \text{ and } \xi = 0\ respectively$$

for $y \in [0, (x_F - u)] if\ \xi \geq 0\ and\ y \in \left[0, \frac{-\sigma}{\xi}\right] if\ \xi < 0,$

such that ξ represents the shape parameter and σ is the scale parameter of Generalized Pareto Distribution.

Gilli and Kellizi (2006) focus in their paper on the computation of the tail risk and their associated confidence intervals using extreme value theory. They showed how extreme value theory can be used to model tail risk measures as Value at Risk, expected shortfall and return level, by using daily log returns for different market indices mainly S&P 500 index. They study both the left and the right tail of the return distribution because the left tail represents losses for an investor with a long position on the index, while the right tail represents losses for an investor being short on the index. They conclude that EVT is an effective method for assessing the size of extreme events.

McNeil and Frey (2000) propose a method in modelling the conditional distribution of asset returns against the current volatility and then fitting the GPD on the tails of residuals. Their approach combines pseudo-maximum-likelihood fitting of GARCH models to measure the current volatility and extreme value theory EVT to estimate the tail of the innovation distribution of the GARCH model. In the first stage they use GARCH modelling to estimate the current market volatility. In the second stage, they add market volatility into VaR estimates obtained from the POT model fitted to residuals of a GARCH model. The results show that conditional approach that models the conditional distribution of asset returns against the current volatility background is more effective for value at risk VaR estimation than an unconditional approach.

They argue that expected shortfall estimates using their two staged model is a good alternative measure to estimate in the model. A comparison of estimates for the expected shortfall using their approach and a standard GARCH model with normal innovations show that the innovation distribution should be modelled by a fat tailed distribution using EVT. They also found that that square-root-of-time scaling of one day VaR estimates to obtain VaR estimates for longer time horizons of 5 or 10 days does not perform well for stock market returns and thus they conclude by suggesting a Monte Carlo method based on their fitted models that provides more acceptable results.

Other researchers looked at extreme value theory sample size. Wada et al (2016), suggest an approach to run to extreme value estimation for small samples of size at most 10 of observations with shortened values, or high measurement uncertainty. The approach is called the likelihood-weighted method (LWM), which is a straightforward Bayesian approach to extreme value estimation for small samples of poor quality. It includes Bayesian inference for the group generalised Pareto or generalised extreme value likelihood and uniform prior distributions for parameters. They argue that this approach provides reasonable estimates of extreme value model parameters and their uncertainties to estimate return values from small samples of poor quality.

VaR estimates using extreme value theory was used to analyze 10 Asian financial markets in a study by Carvalhal and Mendes (2003). They investigated VaR estimates and compared the results with normal and empirical estimates. They found that the extreme value method of estimating Var is more conservative approach than the historical ones when determining capital requirements. Zikovic and Aktan (2009) examined the behaviour of conditional and unconditional extreme value theory and hybrid historical simulation and other models using 95, 99 and 99.5 percent confidence level. They concluded that extreme value theory and hybrid historical simulation provide an accurate estimate of risk while other models underestimate the level of risk.

Yamai and Yoshiba (2004) concluded that when the distribution has a fat tail, the estimation errors of expected shortfall are greater than that of VaR. Thus increasing sample size is the best alternative to reduce errors. They found that expected short fall (ES) provides a better estimation with larger samples. They suggested that both expected shortfall and VaR both complement each other to provide better risk measurements. Moreover, Emmer, Kratz, & Tasche (2015) compared VaR, expected shortfall and expectiles, and determined that expectiles is a better alternative to both VaR and ES. They found that backtesting for ES is less straight forward than back testing for VaR. Chan and Gray (2006) argued that expected value theory is an effective and convenient approach to predict VaR in electricity market. They found that extreme value theory compared to historical simulation based approach and other parametric models, EVT performed better in forecasting out of sample VaR.

Furthermore, Mladenovic, Miletic, and Miletic (2012) came to the conclusion that the extreme value theory is slightly better than the GARCH model regarding VaR estimation. They suggested that both approaches can be used to measure market risk effectively. Extreme value theory was used in Onour (2010) to estimate and compare VaR and expected shortfall values of GCC stock markets to those of S&P500 by applying Generalized Pareto Distribution model (GPD). His results shows that the GCC market show higher risk than S&P 500. In an attempt to study the nonlinear estimation of the tails of distribution in emerging markets, Gencay and Selcuk (2004), investigated the performance of VaR models in these markets. They found that VaR estimates using extreme value theory (EVT) is more accurate at higher quantiles than VaR using historical simulation.

III. Methodology

Take 3 samples of 400, 600 and 800 from three different distributions:

1- Normal Distribution N (0, 16);

2- Log-Normal Distribution;

3- T-Distribution with degrees of freedom equals 4.

As we have seen, according to Gnedenko (1943), EVT shows that the tails of different probability distributions have similar properties. He suggests that $F(v)$ is the cumulative distribution function for a variable v assumed the loss on a portfolio over a certain period of time and u is a value of v on the right tail of the distribution.

- The probability that v lies between u and $u + y$ is $F(u + y) - F(u)$.
- The probability that v is greater than u is $1 - F(u)$.
- The probability that v lies between u and $u + y$ conditional on $v > u$ is $F_u(y)$ where,

$$F_u(y) = \frac{F(u+y) - F(u)}{1 - F(u)} \,. \qquad (1)$$

It is the cumulative probability distribution for which v exceeds u given that it exceeds u.

His results conclude that for distributions $F(v)$, the distribution of $F_u(y)$converges to a generalized Pareto distribution. The generalized Pareto (cumulative) distribution is;

$$G_{\xi,\beta}(y) = 1 - \left[1 + \xi\frac{y}{\beta}\right]^{-\frac{1}{\xi}} \,. \qquad (2)$$

The distribution's parameters β and ξ are estimated from the data. The parameter ξ is the shape parameter and determines the heaviness of the tail of the distribution and the parameter β is a scale parameter where β and ξ are estimated using maximum likelihood equation.

The probability density function $g_{\xi,\beta}(y)$ of the cumulative distribution in equation (2) is calculated by differentiating $G_{\xi,\beta}(y)$ with respect to y. We get

$$g_{\xi,\beta}(y) = \frac{1}{\beta}\left(1 + \frac{\xi y}{\beta}\right)^{-\frac{1}{\xi}-1} . \qquad (3)$$

Following Hull (2015), I choose a value of u close to the 95th percentile point of the sample data for each distribution. I then rank the observations from each sample such that $v > u$. The likelihood function after maximizing its logarithm is:

$$ln\left[\frac{1}{\beta}\left(1 + \frac{\xi(v_i - u)}{\beta}\right)^{-\frac{1}{\xi}-1}\right] . \qquad (4)$$

For each sample, I calculate the value at risk using the following equation

$$VaR = u + \frac{\beta}{\xi}\left\{\left[\frac{n}{n_u}(1-q)\right]^{-\xi} - 1\right\} . \qquad (5)$$

For each sample, VaRs are estimated using a 99 and 99.9 percentile. After a 1000 series of runs, 1000 distribution of portfolio values are observed and will be used to assess Value at Risk. For each distribution with different sample, I calculate the mean VaR from the 1000 VaR estimates. For the normal distribution, calculated mean from each sample is then compared with the true VaR values from the normal-distribution table. For the log normal distribution, calculated mean from each sample is then compared with the true VaR values from the log normal-distribution table. For the t-distribution, calculated mean from each sample is then compared with the true VaR values from the t-distribution table. In the following section I analyze the results from the Monte Carlo simulations and compare the mean VaRs with the estimated ones.

IV. Analysis of Results

4.1 Normal Distribution using 99 percentile

	Sample 400	**Sample 600**	**Sample 800**
Estimated mean VaR	2.3560	2.3463	2.3366
True VaR (Z-Table)	2.3300	2.3300	2.3300
Amount of loss/gain	$(26,000)	$(16,300)	$(6,600)
Error Percentage	1.12%	0.70%	0.28%
Standard Deviation	0.1718	0.1319	0.1170
Minimum	1.7794	1.9710	2.0268
Maximum	2.9957	2.7401	2.6799
Median	2.3524	2.3416	2.3288
Skewness	0.20	0.07	0.14

The table above shows the estimated and true VaR values along with the amount of gain/loss for a normal distribution using 99 percentile for 400, 600 and 800 samples. It is noticed that for the 3 samples the estimated VaR is greater than the true VaR.

Mean estimated VaR is $2.356 million at the 99% confidence level for sample 400, this means it is expected that a one day loss will exceed $2.356 million 1% of the time. However, it is observed from the true VaR that the one day loss will exceed $2.33 million 1% of the time. Thus there is a 26,000 gain which accounts for negative loss. On the other hand, for sample of 800 the daily estimated VaR is $2.3366 million at the 99% confidence level, this means it is expected that a one day loss will exceed $2.3310 million 1% of the time. However, it is observed from the true VaR that the one day loss will exceed $2.33 million 1% of the time. Thus there is a $6,600 gain which accounts for negative loss. It is noticed that when the sample increases to 800, the percentage error changes from 1.12% to 0.28% and the VaR becomes more accurate. Hence, for a normal distribution with 99 confidence interval, as the sample size increase, VaR accuracy increases. Moreover, skewness estimates for the three samples are showing positively skewed losses, indicating that tail of the loss distribution is longer to the right of the centre. Both the skewness estimates reported in the table above and the histograms in Appendix A suggest that the distribution is moderately skewed right.   .

4.2 Log Normal Distribution using 99 percentile

	Sample 400	**Sample 600**	**Sample 800**
Estimated mean VaR	10.4460	10.4320	10.3970
True VaR	10.2779	10.2779	10.2779
Amount of loss/gain	$(168,100)	$(154,100)	$(119,100)
Error Percentage	1.64%	1.5%	1.16%
Standard Deviation	1.693	1.593	1.557
Minimum	6.494	6.745	6.884
Maximum	17.841	16.664	16.364
Median	10.290	10.250	10.258
Skewness	0.49	0.61	0.55

The table above shows the estimated and true VaR values along with the amount of gain/loss for a log normal distribution using 99 percentile for 400, 600 and 800 samples. It is noticed that for the 3 samples the estimated VaR is greater than the true VaR.

Daily estimated VaR for sample 400 is $10.4460 million at the 99% confidence level. This means it is expected that a one day loss will exceed $10.4460 million 1% of the time. However, it is observed from the true VaR that the one day loss will exceed $10.2779 million 1% of the time. Thus there is a $168,100 gain which accounts for negative loss. On the other hand, for sample of 800 the daily estimated VaR is $10.3970 million at the 99% confidence level, this means it is expected that a one day loss will exceed $10.3970 million 1% of the time. However, it is observed from the true VaR that the one day loss will exceed $10.2779 million 1% of the time.

Thus there is a 119,100 gain which accounts for negative loss. It is noticed that when the sample increases to 800, the percentage error changes from 1.64% for sample 400 to 1.5% for sample 600 to 1.16% for sample 800, and the VaR becomes more accurate. Hence, for a log normal distribution with 99 confidence interval, as the sample size increase, VaR accuracy increases and standard deviation decreases. Furthermore, skewness estimates for the three samples are showing positively skewed losses, indicating that tail of the loss distribution is longer to the right of the centre. Both the skewness estimates reported in the table above and the histograms in Appendix B suggest that the distribution is moderately skewed right.

4.3 T-Distribution using 99 percentile

	Sample 400	Sample 600	Sample 800
Estimated mean VaR	3.8465	3.8325	3.8319
True VaR (T-Table)	3.7470	3.7470	3.7470
Amount of loss/gain	$(99,500)	$(85,500)	$(84,900)
Error Percentage	2.66%	2.28%	2.26%
Standard Deviation	0.5210	0.4830	0.4921
Minimum	2.5189	2.4273	2.6358
Maximum	5.9472	6.0204	6.1934
Median	3.7864	3.7641	3.7745
Skewness	0.73	0.82	0.81

The table above shows the estimated and true VaR values along with the amount of gain/loss for a t- distribution using 99 percentile for 400, 600 and 800 samples. It is noticed that for the 3 samples the estimated VaR is greater than the true VaR.

Daily estimated VaR for sample 400 is $3.8465 million at the 99% confidence level. This means it is expected that a one day loss will exceed $3.8465 million 1% of the time. However, it is observed from the true VaR that the one day loss will exceed $3.7470 million 1% of the time. Thus there is a $99,500 gain which accounts for negative loss. On the other hand, for sample of 800 the daily estimated VaR is $3.8319 million at the 99% confidence level, this means it is expected that a one day loss will exceed $3.8319 million 1% of the time. However, it is observed from the true VaR that the one day loss will exceed $3.7470 million 1% of the time. Thus there is a 84,900 gain which accounts for negative loss. It is noticed that when the sample increases to 800, the percentage error changes from 2.66% for sample 400 to 2.28% for sample 600 to 2.26% for sample 800, and the VaR becomes more accurate. Hence, for a t-distribution with 99 confidence interval, as the sample size increase, VaR accuracy increases and standard deviation decreases from 0.5210 to 0.4921. Furthermore, skewness estimates for the three samples are showing positively skewed losses, indicating that tail of the loss distribution is longer to the right of the centre. Both the skewness estimates reported in the table above and the histograms in Appendix C suggest that the distribution is moderately skewed right.

4.4 Normal Distribution using 99.9 percentile

	Sample 400	**Sample 600**	**Sample 800**
Estimated mean VaR	2.9733	2.9875	3.0218
True VaR (Z-Table)	3.1000	3.1000	3.1000
Amount of loss/gain	$126,700	$112,500	$78,200
Error Percentage	4.08%	3.63%	2.52%
Standard Deviation	0.7895	0.3144	0.2652
Minimum	2.1265	2.2096	2.1506
Maximum	24.3817	5.2497	4.2030
Median	2.8918	2.9679	3.0163
Skewness	20.17	0.85	0.52

This table presents the estimated and true VaR values along with the amount of gain/loss for a normal distribution using 99.9 percentile for 400, 600 and 800 samples. It is noticed that for the 3 samples the estimated VaR is less than the true VaR.

The mean estimated VaR for sample 400 is $2.9733 million at the 99.9% confidence level, which means it is expected that a one day loss will exceed $2.9733 million 0.1% of the time. However, it is observed from the true VaR that the one day loss will exceed $3.1000 million 1% of the time. Thus there is a $126,700 expected loss that decreases to $112,500 for a sample of 600. In addition, for sample of 800 the daily estimated VaR is $3.0218 million at the 99.9% confidence level, this means it is expected that a one day loss will exceed $3.0218 million 0.1% of the time. However, it is observed from the true VaR that the one day loss will exceed $3.1000 million 0.1% of the time. Thus there is $78,200 loss.

It is noticed that when the sample increases to 800, the percentage error decreased from 4.08% to 2.52% and the VaR becomes more accurate. Hence, for a normal distribution with 99.9 confidence interval, as the sample size increase, VaR accuracy increases and standard deviation decreases. In addition, skewness estimates for the three samples are showing positively skewed losses, indicating that tail of the loss distribution is longer to the right of the centre. Surprisingly, for sample 400 the distribution is extremely rightly skewed which then decreases from 20.17 to 0.52 in sample 800. Both the skewness estimates reported in the table above and the histograms in Appendix D suggest that the distribution is skewed right.

4.5 Log Normal Distribution using 99.9 percentile

	Sample 400	**Sample 600**	**Sample 800**
Estimated mean VaR	21.6690	21.6370	21.6420
True VaR	22.1980	22.1980	22.1980
Amount of loss/gain	$529,000	$561,000	$556,000
Error Percentage	2.38%	2.53%	2.50%
Standard Deviation	10.624	10.580	10.564
Minimum	8.243	8.246	8.018
Maximum	98.955	99.369	98.906
Median	18.829	18.724	18.728
Skewness	2.45	2.46	2.46

The table above shows the estimated and true VaR values along with the amount of gain/loss for a log normal distribution using 99.9 percentile for 400, 600 and 800 samples. It is noticed that for the 3 samples the estimated VaR is less than the true VaR.

Daily estimated VaR for sample 400 is $21.6690 million at the 99.9% confidence level. This means it is expected that a one day loss will exceed $21.6690 million 0.1% of the time. However, it is observed from the true VaR that the one day loss will exceed $22.1980 million 0.1% of the time. Thus there is a $529,000 loss. On the other hand, for sample of 800 the daily estimated VaR is $21.6420 million at the 99.9% confidence level, this means it is expected that a one day loss will exceed $21.6420 million 0.1% of the time. However, it is observed from the true VaR that the one day loss will exceed $22.1980 million 0.1% of the time. Thus there is a 556,000 loss. It is noticed that when the sample increases to 800, the percentage error changes from 2.38% for sample 400 to 2.53% for sample 600 to 2.50% for sample 800, and the VaR becomes more accurate. Hence, for a log normal distribution with 99.9 confidence interval, as the sample size increase, VaR accuracy increases and standard deviation decreases from 10.624 to 10.564. Furthermore, skewness estimates for the three samples are showing positively skewed losses, indicating that tail of the loss distribution is longer to the right of the centre. Both the skewness estimates reported in the table above and the histograms in Appendix E suggest that the distribution is moderately skewed right.

4.6 T-Distribution with degrees of freedom equal 4 using 99.9 percentile

	Sample 400	**Sample 600**	**Sample 800**
Estimated mean VaR	7.0844	7.0930	7.0906
True VaR (T-Table)	7.1730	7.1730	7.1730
Amount of loss/gain	$88,600	$80,000	$82,400
Error Percentage	1.24%	1.11%	1.14%
Standard Deviation	3.0557	3.0388	3.0331
Minimum	3.2310	3.2943	3.2874
Maximum	27.8328	27.9708	28.0676
Median	6.3554	6.3640	6.3617
Skewness	2.63	2.68	2.65

The table above shows the estimated and true VaR values along with the amount of gain/loss for a t- distribution using 99.9 percentile for 400, 600 and 800 samples. It is noticed that for the 3 samples the estimated VaR is less than the true VaR.

Daily estimated VaR for sample 400 is $7.0844 million at the 99.9% confidence level. This means it is expected that a one day loss will exceed $7.0844 million 0.1% of the time. However, it is observed from the true VaR that the one day loss will exceed $7.1730 million 0.1% of the time. Thus there is a $88,600 loss. On the other hand, for sample of 800 the daily estimated VaR is $7.0906 million at the 99.9% confidence level, this means it is expected that a one day loss will exceed $7.0906 million 0.1% of the time. However, it is observed from the true VaR that the one day loss will exceed $7.1730 million 0.1% of the time. Thus there is an 82,400 loss.

It is noticed that when the sample increases to 800, the percentage error changes from 1.24% for sample 400 to 1.11% for sample 600 to 1.14% for sample 800, and the VaR becomes more accurate. Hence, for a t- distribution with 99.9 confidence interval, as the sample size increase, VaR accuracy increases and standard deviation decreases from 3.0557 to 3.0331. Furthermore, skewness estimates for the three samples are showing positively skewed losses, indicating that tail of the loss distribution is longer to the right of the centre. Both the skewness estimates reported in the table above and the histograms in Appendix F suggest that the distribution is moderately skewed right, and as the sample increase skewness increases.

V. Conclusion

An extreme value theory maximum likelihood method is used to estimate value at risk VaR on different sample sizes and confidence intervals using normal, lognormal and t-distributions. The study reveals that for confidence interval of 99%, as the sample size increases from 400 to 600 to 800 the accuracy of VaR increases and VaR estimate is more accurate under normal distribution. The study also shows VaR estimate tends to be more accurate for T-Distribution with confidence interval 99.9% as the sample size increase from 400 to 800. And that for the three different distributions, VaR estimates are positively skewed.

For future research it is recommended to investigate VaR performance by using actual data on a specific historical window. Moreover, estimating value at risk using more than one approach provides a better comparison of VaR estimations and presents a better understanding of which approach measures VaR's accuracy effectively. In addition, expected shortfall ES is another tool that measures risk and should be embedded in Monte Carlo simulations and taken into consideration since it's a better measure of risk than VaR as it takes into account the whole of the tail of the distribution.

Bibliography

A.P. (2014, April 9). What Usain Bolt can teach banks about financial risk *The Economist*, Retrieved from:http://www.economist.com/blogs/economist-explains/2014/04/economist-explains-4

Abad, P., & Benito, S. (2013). A detailed comparison of value at risk estimates. *Mathematics and Computers in Simulation*, *94*, 258-276.

Allen, D. E., Singh, A. K., & Powell, R. J. (2013). EVT and tail-risk modelling: Evidence from market indices and volatility series. *The North American Journal of Economics and Finance*, *26*, 355-369.

Bensalah, Y. (2000). *Steps in applying extreme value theory to finance: a review*. Bank of Canada.

Benstock, D., & Cegla, F. (2016). Sample selection for extreme value analysis of inspection data collected from corroded surfaces. *Corrosion Science*, *103*, 206-214.

Boudoukh, J., Richardson, M., & Whitelaw, R. (1998). The best of both worlds. Risk, 11(5), 64-67.

Cabedo, J. D., & Moya, I. (2003). Estimating oil price 'Value at Risk'using the historical simulation approach. *Energy Economics*, *25*(3), 239-253.

Carvalhal, A., & Mendes, B. V. (2003). Value-at-risk and extreme returns in Asian stock markets. *International Journal of Business*, *8*(1).

Chan, K. F., & Gray, P. (2006). Using extreme value theory to measure value-at-risk for daily electricity spot prices. *International Journal of Forecasting*, *22*(2), 283-300.

Dey, D., & Yan, J. (2016). *Extreme Value Modeling and Risk Analysis : Methods and Applications*. Boca Raton, FL: Chapman and Hall/CRC.

Emmer, S., Kratz, M., & Tasche, D. (2015). What is the best risk measure in practice? A comparison of standard measures. *Journal of Risk*, *18*(2), 31-60.

Gencay, R., & Selcuk, F. (2004). Extreme value theory and Value-at-Risk: Relative performance in emerging markets. *International Journal of Forecasting*, *20*(2), 287-303.

Gilli, M. (2006). An application of extreme value theory for measuring financial risk. *Computational Economics*, *27*(2-3), 207-228

Hendricks, D. (1996). Evaluation of value-at-risk models using historical data (digest summary). *Economic Policy Review Federal Reserve Bank of New York*, *2*(1), 39-67.

Hull, J. (2015). *Risk Management and Financial Institutions*, John Wiley & Sons.

Khindanova, I., Rachev, S., & Schwartz, E. (2001). Stable modeling of value at risk. *Mathematical and Computer Modelling*, *34*(9), 1223-1259.

Levine, D. (2009). Modeling tail behavior with extreme value theory. Risk Management, 17, 14-18.

McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of empirical finance*, *7*(3), 271-300.

Mladenović, Z., Miletić, M., & Miletić, S. (2012). Value at Risk in European Emerging Economies: Empirical Assessment of Financial Crisis Period. In*Scientific Conference: from Global Crisis to Economic Growth: which way to take* (pp. 20-09).

Onour, I. A. (2010). Extreme risk and fat-tails distribution model: Empirical analysis. *Journal of Money, Investment and Banking*, (13).

Radivojević, N., Cvjetković, M., & Stepanov, S. (2016). The new hybrid value at risk approach based on the extreme value theory. *Estudios de Economia*, *43*(1).

Rufino, C. C. (2011). Empirical Comparison of Extreme Value Theory Vis-À-Vis Other Methods of VAR Estimation Using ASEAN 3 Exchange Rates.*DLSU Business & Economics Review*.

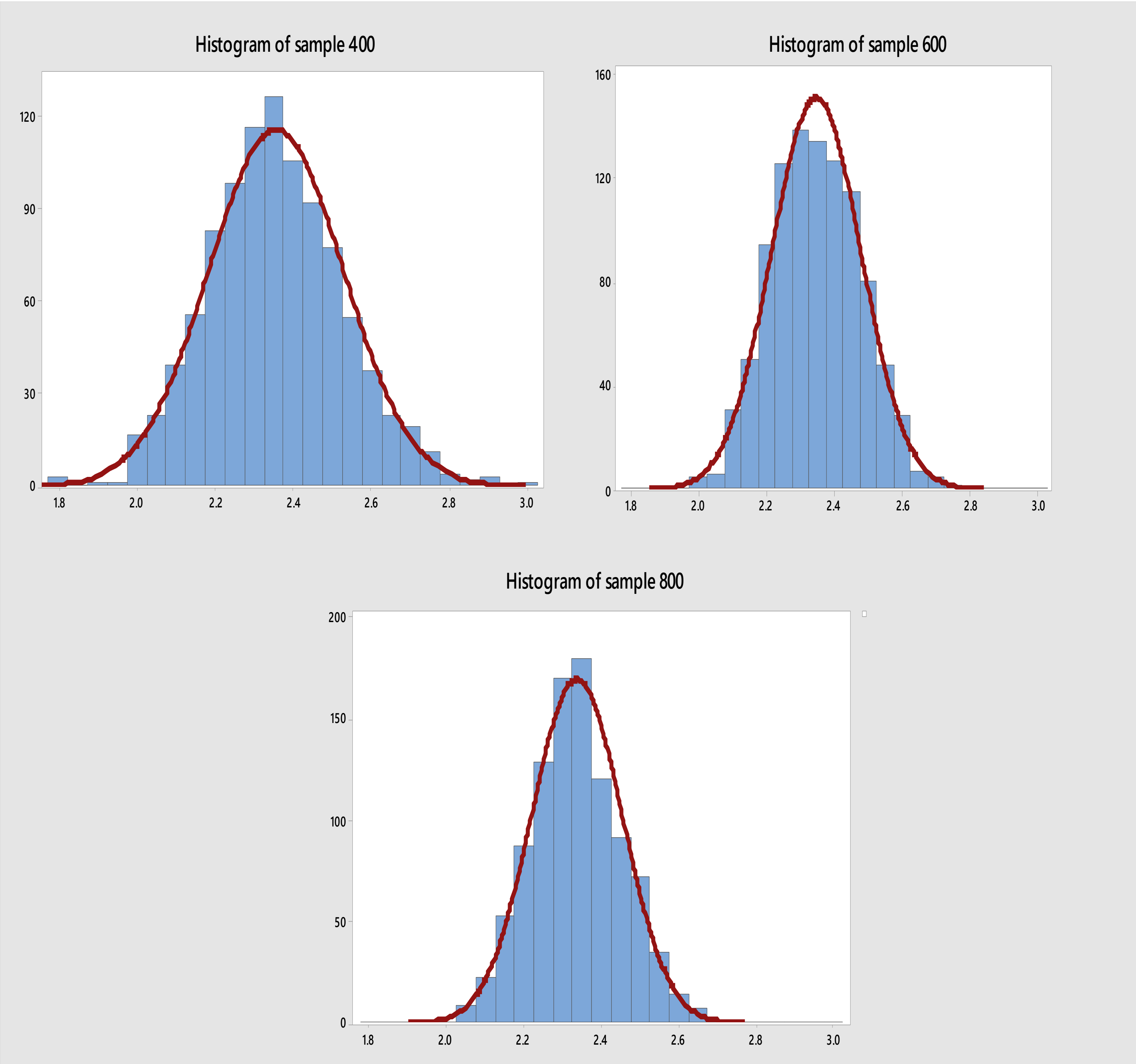
Wada, R., Waseda, T., & Jonathan, P. (2016). Extreme value estimation using the likelihood-weighted method. *Ocean Engineering*, *124*, 241-251.

Yamai, Y., & Yoshiba, T. (2002). Comparative analyses of expected shortfall and value-at-risk (3): their validity under market stress. *Monetary and Economic Studies*, *20*(3), 181-237.

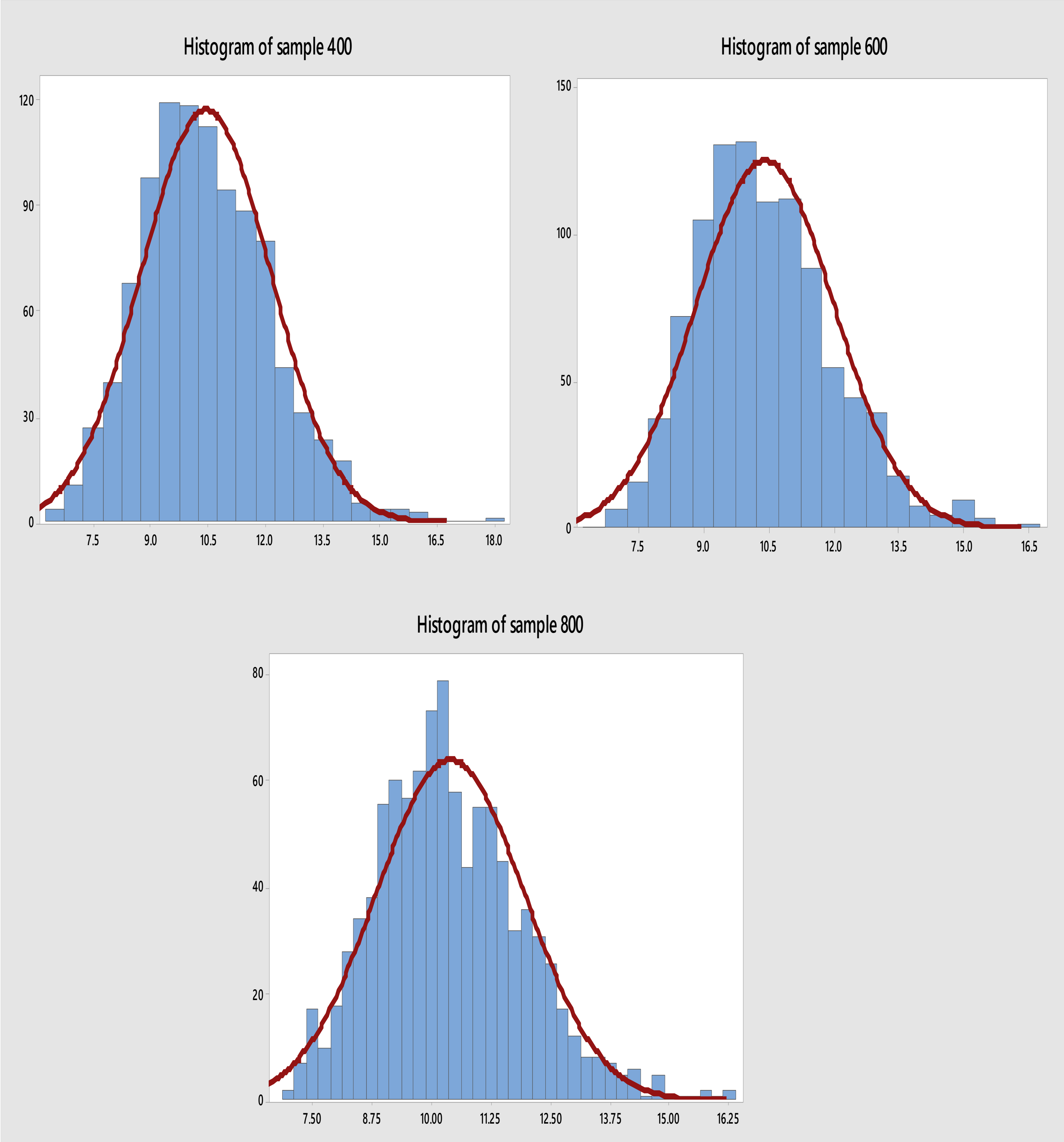
Yamai, Y., & Yoshiba, T. (2005). Value-at-risk versus expected shortfall: A practical perspective. *Journal of Banking & Finance*, *29*(4), 997-1015.

Zikovic, S., & Aktan, B. (2009). Global financial crisis and VaR performance in emerging markets: A case of EU candidate states-Turkey and Croatia.*Zbornik radova Ekonomskog fakulteta u Rijeci, časopis za ekonomsku teoriju i praksu-Proceedings of Rijeka Faculty of Economics, Journal of Economics and Business*, *27*(1), 149-170.

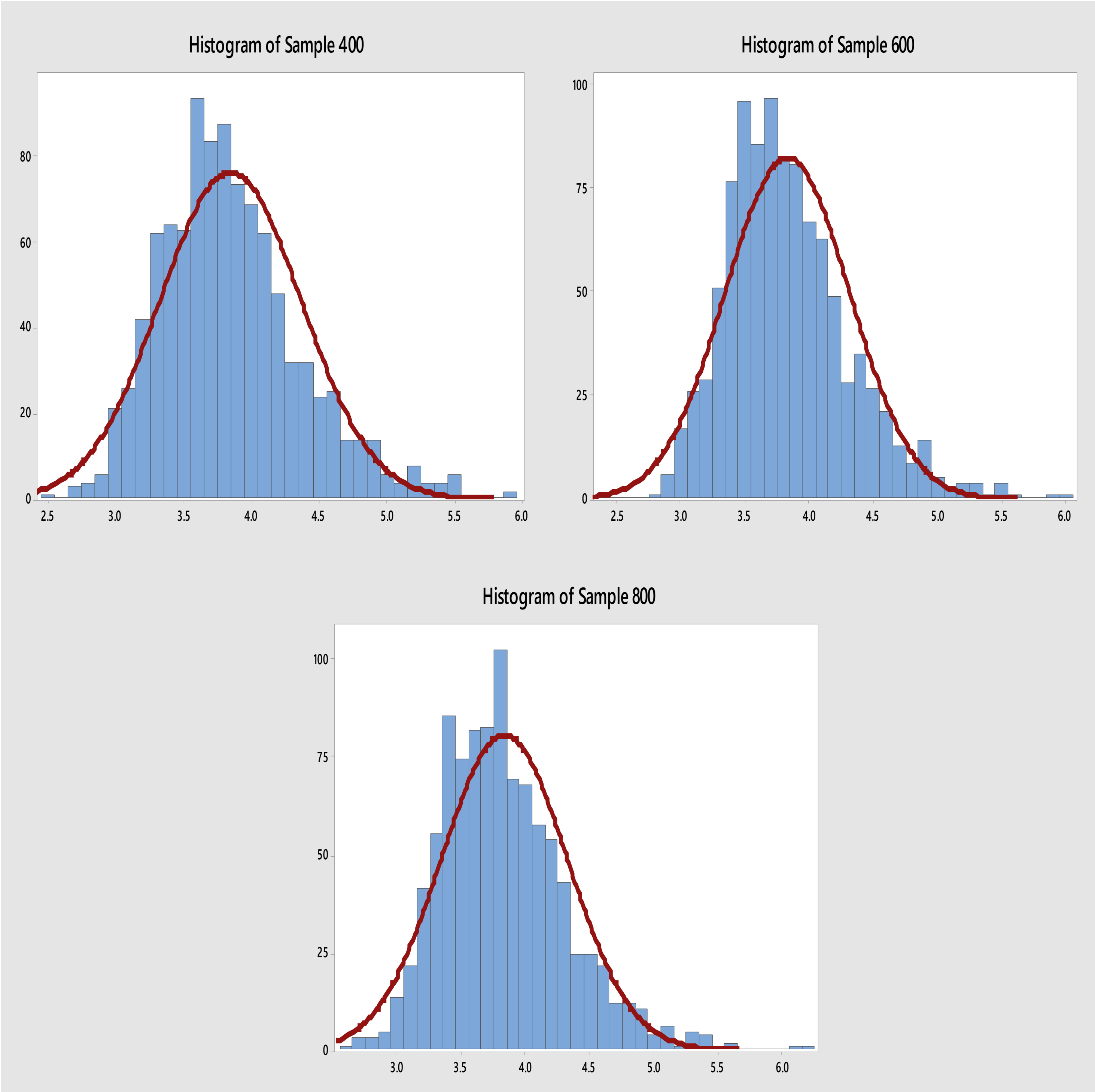
Appendix A: Normal Distribution using 99 percentile



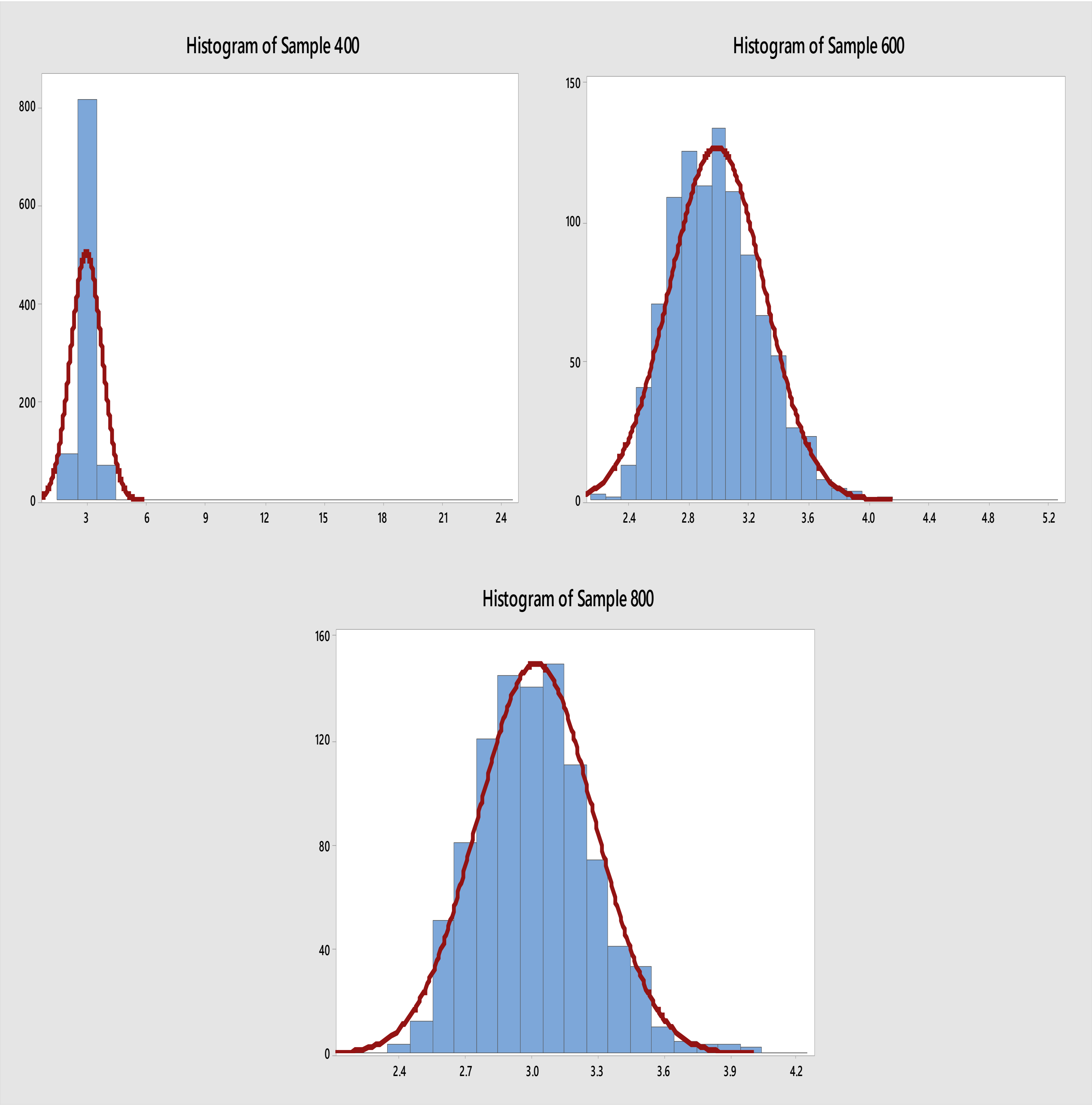
Appendix B: Log Normal Distribution using 99 percentile



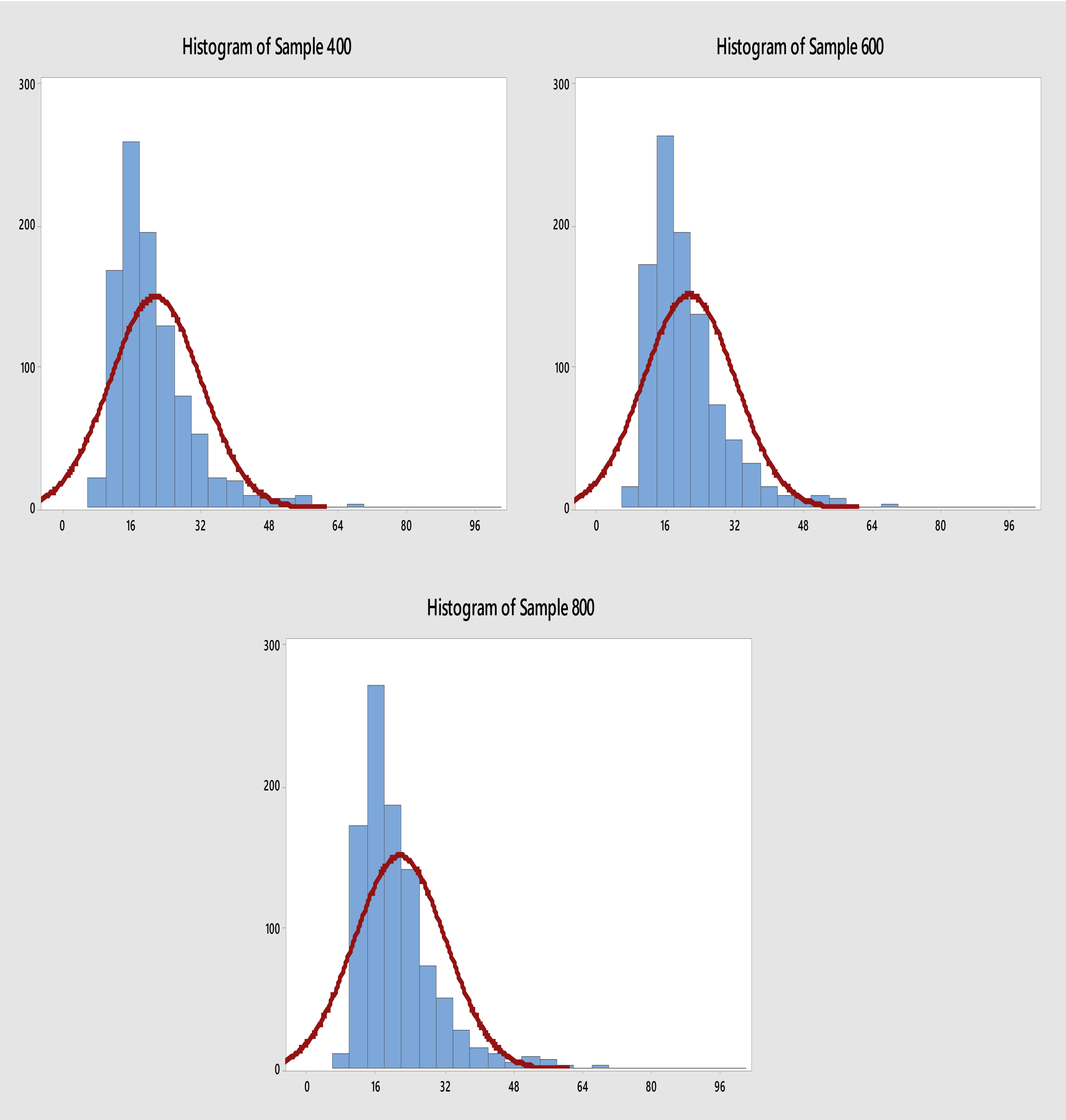
Appendix C: T-Distribution using 99 percentile



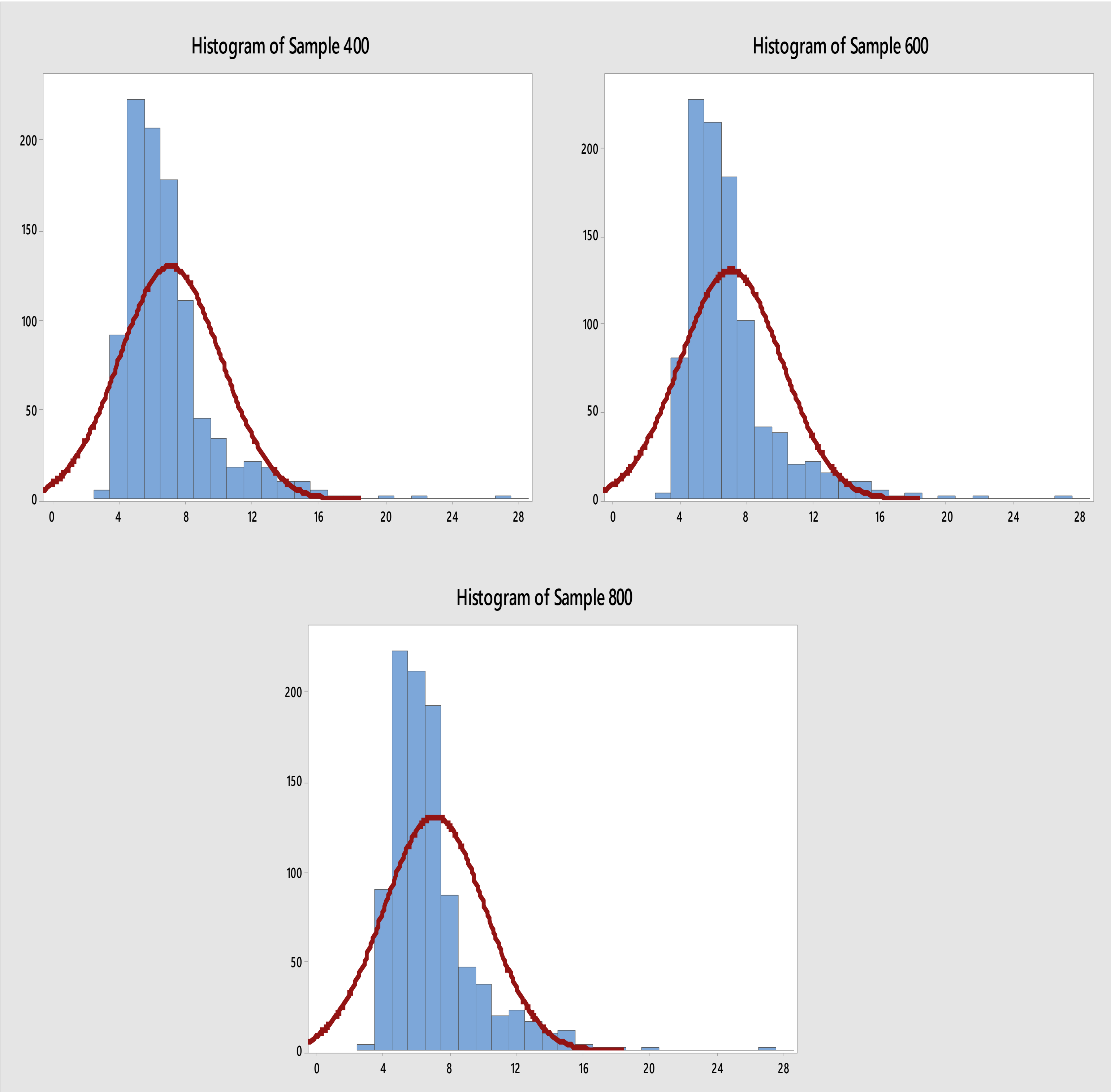
Appendix D: Normal Distribution using 99.9 percentile



Appendix E: Log Normal Distribution using 99.9 percentile



Appendix F: T-Distribution using 99.9 percentile



Appendix G: List of Abbreviations

VaR	Value at Risk
ES	Expected Shortfall
ARMA	Autoregressive Moving Average
EVT	Extreme Value Theory
GPD	Generalized Pareto Distribution
PBH	Pickands-Balkema-de Haan Theorem
GEV	Generalized Extreme Value Distribution
POT	Peak Over Threshold
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
LWM	Likelihood Weighted Method
GCC	Gulf Cooperation Council Countries