The Impact of Bimodal Distribution in Ocean Transportation Transit Time on Logistics Costs: An Empirical & Theoretical Analysis

by

Lita Das

Bachelor of Engineering (Honors) in Mechanical Engineering
The Birla Institute of Technology & Science (2009)

Master of Engineering in Mechanical Engineering
Cornell University (2010)

Submitted to the Engineering Systems Division
in partial fulfillment of the requirements for the degree of

Master of Science in Engineering Systems

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2013

© Massachusetts Institute of Technology 2013. All rights reserved.

Author ................................................................. Engineering Systems Division

Certified by .................. Dr. Chris Caplice
Senior Lecturer, Engineering Systems Division
Executive Director, Center for Transportation and Logistics
Thesis Supervisor

Accepted by .............. Olivier L. de Weck
Professor of Engineering Systems and Aeronautics and Astronautics
Chair, Engineering Systems Division Education Committee
The Impact of Bimodal Distribution in Ocean Transportation Transit Time on Logistics Costs: An Empirical & Theoretical Analysis

by

Lita Das

Submitted to the Engineering Systems Division on June 6, 2013, in partial fulfillment of the requirements for the degree of Master of Science in Engineering Systems

Abstract

As ocean shipments have increased alongside globalization, transit time uncertainty has increased as well. This problem was observed to have variable levels of impacts on logistics cost and safety stock levels. This thesis examines the effects of bimodality in transit time distributions—in particular, the cost of ignoring bimodality. One method common in practice is to completely ignore variability. On the other hand, a popular theoretical method to account for transit time variability is to assume that demand over transit time is normally distributed. Which is, in many cases, false. To display the incorrectness of such assumptions, the paper will compare the two approaches to empirical analysis on bimodal transit time distributions.

Thesis Advisor: Dr. Chris Caplice
Title: Senior Lecturer, Engineering Systems Division & Executive Director, Center for Transportation and Logistics
Acknowledgements

This thesis would not have been possible without the help, support and guidance of so many people that I have come across in the past couple years of my journey at MIT. It has been extremely inspiring to see how every one at MIT dives into complex problems and come up with amazing solutions that affect the world!

The biggest motivator for the work presented in this thesis was my advisor, Dr. Chris Caplice. I am thankful to him for giving me the opportunity to work on such an interesting problem. His guidance and supervision was very valuable throughout the course of the project. He was instrumental in simplifying the ever-so complicated parts of the problem at hand. Time and again he encouraged me to think on different levels-as a researcher as well as a practitioner in the industry.

I am extremely grateful to Dr. Basak Kalkanci, former postdoctoral fellow at Center for Transportation and Logistics and currently Assistant Professor at Georgia Institute of Technology, for her guidance throughout the evolution of this project. This project certainly would not have been at the stage that it is now without her very insightful comments and suggestions. Thank you Basak, for being patient with me while I was trying to understand and discuss different aspects of the problem.

My parents and my sister back in India have always been a huge source of love, encouragement and strength. This certainly played a major role in my journey here at MIT and in the completion of this thesis.

I would also like to thank my officemates and friends in Engineering Systems Division for all the interesting discussions that I have had about the project and otherwise which undoubtedly helped me hone my research skills. Also, thanks to Beth, our graduate administrator for making life here easier. Finally, thank you Colin for keeping me motivated and inspired throughout the past 9 months we have known each other!
Contents

ABSTRACT ................................................................................................................................................ 3

ACKNOWLEDGEMENTS ........................................................................................................................ 5

LIST OF TABLES ................................................................................................................................... 11

CHAPTER 1. INTRODUCTION...................................................................................................... 13

1.1 MOTIVATION..................................................................................................................... 13

1.3 THESIS OVERVIEW ........................................................................................................... 19

1.4 STRUCTURE OF THE THESIS ........................................................................................ 21

CHAPTER 2. LITERATURE REVIEW ............................................................................................ 23

2.1 SUMMARY OF THE REVIEW ........................................................................................... 32

CHAPTER 3. UNDERSTANDING THE DATA ............................................................................. 33

3.1 SOME BASIC CHARACTERISTICS OF PORTS IN THE DATA ........................................ 35

3.1.1 Initial variability analysis of the lanes of the dataset obtained from the manufacturing firm
 .................................................................................................................................35

3.1.2 Geographical information of the origin port location in terms of non-unimodality in the lanes ...........................................................................................................................................35

3.2 SUMMARY OF THE CHAPTER........................................................................................ 43

CHAPTER 4. ANALYSIS....................................................................................................................... 45

4.1. CHARACTERIZING TRANSIT TIME DISTRIBUTION.............................................................. 45

4.1.1 Test of Unimodality- Hartigan’s Dip Test ...........................................................................46

4.1.2 Evidence of non-unimodality in transit time distribution .............................................. 47

4.2. SIMULATING CASES OF BIMODALITY AND ITS EFFECTS ............................................. 48

4.2.2 Creating bimodal distribution from mixture of two normal distributions .................. 52

4.2.3. Derivation of the mean and standard deviation of the mixture distribution ............... 55

4.2.4. Level of Bimodality ........................................................................................................56

4.3. SIMULATION MODEL USED FOR CALCULATION OF SAFETY STOCK ....................... 56

7
4.3.1 Safety stock Analysis .............................................................................................................................. 57
4.4. Simulation Model used for calculation of logistics cost ................................................................. 57

CHAPTER 5. RESULTS ................................................................................................................................................. 65

5.1 Effect of distribution of transit time on safety stock values ................................................. 65

5.1.1. Effect of increasing the level of bimodality-Case II vs. Case III ........................................ 65
5.1.2 Using available data to analyze the effect of non-unimodality in safety stock values ........... 69
5.1.3. Effect of increasing the level of bimodality-Case I vs. Case III ........................................ 71

5.2 Effect of variability (including bimodality) in transit time distribution on logistics cost .......................................................... 72

5.2.1 Comparing Case I & Case II ................................................................................................. 73
5.2.2 Comparing Case I & Case III ......................................................................................... 78
5.2.3 Comparing Case II & Case III ..................................................................................... 82

5.3 Summary of Results .................................................................................................................. 93

CHAPTER 6. CONCLUSION .......................................................................................................................... 95

6.1 Further Research ......................................................................................................................... 97

APPENDIX .............................................................................................................................................. 99

(I) Detailed derivation for mean and variance of demand over transit time .......................................................... 99

(II) Detailed derivation for total cost as proposed by Hadley-Whitin (1963) .......................................................... 101

BIBLIOGRAPHY .................................................................................................................................. 103
List of figures

1-1. End-to-End transportation sequence on a time space diagram ................................. 14
1-2 Shape approximations of transit time distribution under 3 different approaches .......................................................................................................................... 18
3-1. Average and Coefficient of Variation of the end-to-end transportation of the lanes in the data set from the manufacturing firm ........................................................................ 36
3-2. Average and Coefficient of Variation corresponding to the lanes (manufacturer) .................................................................................................................. 38
3-3. Average and Coefficient of Variation corresponding to the lanes (retailer) ........................................................................................................................ 41
4-1(a). Convex and Concave regions of a unimodal distribution .................................. 46
4-1(b). Convex and Concave regions of a bimodal distribution ..................................... 47
4-2. Bimodal and Long Right tailed distribution indicated in distribution of transit time .................................................................................................................. 48
4-3. An example of a bimodal distribution created by mixing two normal distributions ............................................................................................................. 54
5-1. Percentage change in safety stock using Hadley-Whitin(HW) equation versus actual distribution(bimodal) of transit time ........................................................................ 66
5-2. Resultant bimodal distribution and the normal approximation for the cases when level of bimodality is 0 and 4 (fig A and B) ........................................................................ 67
5-3. Resultant bimodal mixture and normal approximation of the mixture distribution........................................................................................................68

5-4. Percentage change in safety stock using deterministic value versus actual distribution (bimodal) of transit time ........................................................................................................71

5-5. Difference between logistics cost of Case I and Case II........................................................................................................74

5-6. Difference between logistics cost of Case I and Case III........................................................................................................79

5-7. Difference between logistics cost of Case II and Case III........................................................................................................83

5-8. Change of difference across increasing CR for a given level of bimodality........85

5-9. Bimodal distribution and normal approximation for low and high level of bimodality ........................................................................................................86

5-10. Resultant distribution with different mixture rates starting with 0.1 on top left to 0.9 on the bottom right..........................................................90
List of tables

3-1. Global distributions of origin port locations for freight into US.................43
4-1. Example across different cases for transit time value approximations...........54
4-2. Illustration of the simulation model for calculation of logistics cost ...............59
5-1 Simulation Results of effect of non-unimodality on safety stock values.............70
5-2. Difference between logistics cost of Case I and Case II in percentage.............77
5-3. Difference between logistics cost of Case I and Case III in percentage.............80
5-4. Difference between logistics cost of Case II and Case III in percentage............83
5-5 Results of effect on logistics cost with changing mixture rates...................89
Chapter 1. Introduction

1.1 Motivation

Maritime transportation traffic has seen tremendous growth in recent years due to a surge in globalization and international trade. Other factors that led to the growth of ocean freight transportation can be attributed to technical improvements of the ships and efficiency at the terminal ports. Moreover, the introduction of containerization made large shipment movements possible. Economies of scale enable ocean transportation to be a low cost mode. Therefore it is not surprising to see that container shipments between North America, Asia, and Europe have more than tripled in the last 15 years, rising from just over 15 million twenty foot equivalent units (TEU) in 1995 to 48 million in 2010, Rodrigue (2013).

The higher trade volumes due to increased ocean traffic coupled with increase in the number of different shipping locations used by firms have made the conducting and coordinating of shipping operations a very complex task. The total end-to-end transportation of a shipment from the point of origin (typically the manufacturer) to the final destination (usually a distribution center) is made up of a series of individual and often independently managed activities. The transportation transit time is the cumulative effect of all of these individual movements.

The sequence of movements in the end-to-end transportation is captured by six time stamps corresponding to the arrival and departure of the shipments at
different locations. These time stamps are: time of departure from the origin (TS1), time of arrival at the origin port (TS2), time of departure at the origin port (TS3), time of arrival at the destination port (TS4), time of departure from the destination port (TS5) and time of arrival at the final destination (TS6). The intervals between these six time stamps constitute five different location segments. These segments are Origin-to-Origin Port (OOP), Origin Port dwell (OPD), Port-to-Port (PTP), Destination Port dwell (DPD) and Destination port-to-final Destination (DPF). A diagrammatic representation of the network consisting of the time stamps and the segments is shown below on a time space diagram.

![Fig 1-1 End-to-End transportation sequence on a time space diagram](image)

Companies shipping goods across the globe are concerned with the total end-to-end transit time for a shipment, as well as the unreliability in the shipments, variability, and shape of the distribution itself. Variability in transit time is observed across all the segments of the end-to-end transportation. Most well recognized information
sources such as Drewry Shipping Consultants Ltd and Lloyds Maritime Intelligence only report the port-to-port component of a global shipment for various trade lanes. This misses for other segments of the movement. However, shippers to analyze the effect of transit time variability while making inventory management decisions should use end-to-end transit times of ocean transportation in global supply chains. It is because most of the variability occurs in origin-to-origin port region.

A number of issues can impact global ocean transit times: port congestion, bad weather, slow steaming, labor strikes at ports and sailing longer distances to avoid pirating to name a few. For instance slow steaming, which was introduced during the recession in 2009, helped cut costs on fuel consumption by reducing ship speeds. A 10% reduction in the speed lowers the engine power by 27% and results in a 19% reduction of fuel consumption, Kloch (2013). Slow steaming will naturally increase the average transportation transit time.

One major effect of increase in variability of transit time is increase in schedule unreliability in terms of its uncertainty. In fact Drewry Maritime Research analysts indicate that only 56% of shipments arrived at selected ports on time in the second quarter of 2011. The corresponding number for the first quarter of 2011 was only 51%. Uncertain schedule reliability causes problems on both the ends of the supply chain that is the manufacturer and the retailer side. This unreliability can lead to unavailability of products on the retailers’ shelves and hence customer dissatisfaction.
On the other hand for the manufacturer unreliability in shipments causes them to observe lower production rates due to increased time in the decision making process from the retailers side. In fact uncertain schedule reliability was identified as a major problem for manufacturers in separate surveys carried out in 2011 by the US Federal Maritime Commission and logistics firm BDP International. As discussed in a recent news article, Bloomberg (2012), only 63.7% of containers were on time in the first 20 weeks of 2012 versus 65.9% a year earlier, according to INTTRA, a US e-commerce platform that handles 525,000 shipments a week. Higher fuel costs, increased competition and lower revenues are affecting the service quality of container shipping companies and leading some to even shed service on certain trade lanes.

Variability in transit time and unreliability in delivery of shipments in ocean transport ultimately affect the shippers’ operational performance. The uncertainty in transit time can be modeled as a probability distribution and considered explicitly in the calculation of an optimal inventory policy for a shipper. However, our review of common practices in industry and conversations with companies and academic experts suggest that most companies do not necessarily consider transport time variability or unreliability in their inventory planning. In fact, most planning systems such as SAP and Oracle are configured to only consider deterministic transit times. Other systems, such as SAP APO, use a simplified approach to address variability: standard deviation of the transit time is calculated and used in the classical Hadley and Whitin (HW) (1963). The details of the normal
approximation used in HW formula for total cost are discussed later in this thesis in section 4.2.1. However, calculations done using this formula usually assume that the transit times are normally distributed. For purposes of this thesis, it is assumed that any reference to the HW formula also considers a normal approximation to the transit time distribution.

Our analysis of around 125,000 container movements from 4 major US shippers that occurred in 2011 and 2012 has shown that ocean transport transit times are rarely normal and are often found to be bimodal. This transit time distribution has not been well examined in the literature. There could be a number of reasons behind it. One possible reason is the fact that bimodality occurs more frequently in ocean versus other transportation modes. Hence bimodality in distributions has gone unobserved if analysis is done for modes other than that of ocean. Secondly, identifying different modes and hence bimodality in transit time distributions requires more complex analytical techniques than in the case of unimodal distributions. Moreover, it was also observed that certain transportation lanes that have a bimodal transit time are usually low frequency but high impact in terms of the volume of shipment carried. As a result it is possible that the analysts do not concern themselves with the effect of bimodality because they only consider the fact that they are low frequency lanes but fail to incorporate the high volume shipments corresponding to the lanes.
1.2 Setting the stage

As indicated above transit time distributions can be dealt within an inventory management system in any of three ways. The first is treating the transit time distribution as a deterministic value and completely ignore variability. The second is approximating it to a normal distribution of transit time. The third way is using the actual bimodal distribution. Figure 1-2 shows the shapes of the distributions under the three approaches for an actually occurring bimodal distribution of transit time.

Fig 1-2 Shape approximations of transit time distribution under 3 different approaches

It can be observed that the frequency of an instance of transit time value is very different in all the three cases because of the shape of the distribution.

In order to understand the effect of such approximations let us consider a simple case where transit time for an item is 5 days 50% of the time and 7 days for rest of the 50%. Also, let demand be normally distributed with a mean of 20 units and a variance of 5 (units)^2. It is also required that the probability of stock out during
transit time should be less than 5%. Under this situation the safety stock inventory carried under different cases are:

(i) Deterministic transit time -> 8.98 units  
(ii) Normal Approximation -> 34 units  
(iii) Actual distribution (Bimodal) -> 58.9 units

The actual distribution gives the optimal amount of safety stock inventory level that should be carried and the approximations highly underestimate this amount of stock. This in turn could lead to higher back orders and hence large back-ordering cost. This thesis will therefore try to investigate the extent of effect of incorrect assumptions in transit time distributions to make inventory decisions.

1.3 Thesis Overview

To recap, transit time distribution can often be non-unimodal. This can dramatically affect inventory management decisions of shippers, in the form of safety stock levels, ordering quantities, and ordering frequency. In summary, these non-unimodal distributions can lead to increased logistics cost.

The aim of this thesis then is to show the extent to which these distributions are impactful and under what circumstances are they impactful. This is done in a four-step process.

The first step in accomplishing this is establishing how we can identify when the distribution of transit time is not unimodal. This will be done via a statistical test of
unimodality called the Hartigan's dip test. This test will be used to prove or disprove whether distribution of transit time is unimodal.

Moving on, in the second step we will investigate how frequently the occurrence of non-unimodality in transit time distribution happens. Is it rare? Is it in extraneous circumstances? We will establish the regularity, and look into the circumstances surrounding that regularity. Does it happen in certain trade lanes?

Having identified when a distribution is not unimodal, and the regularity of that occurrence, the third step will involve digging into the impact on costs of this occurrence. In order to do this we will use the simplest case of non-unimodality—the case of two modes in the distribution or bimodality. The effect of bimodality on logistics cost is observed under three scenarios, which are called Case I, II & III in this thesis.

- **Case I** corresponds to the case of completely ignoring variability and only using the average transit time when setting inventory levels. This is the most common policy used in practice. So, in other words the safety stock levels, reorder points and order quantities are set based on a time invariant single number assigned to the transit time value.

- **Case II** on the other hand corresponds to the case when the inventory management decisions are calculated using the Hadley-Whitin formula assuming normal distribution of demand over transit time.
• *Case III* denotes optimizing inventory decisions using the actual transit time distribution, which is bimodal.

Finally, in the fourth step we will provide insights on the conditions when it makes sense to use the different cases outlined in step 3. These conditions will include the characteristics of the trade lane and service level decided by the shippers.

**1.4 Structure of the thesis**

The remainder of the thesis is organized as following. Chapter 2 provides a review of how the problem of variability of transit time is addressed in the literature. A section of the review specifically deals with the issue of assumption of normality in transit time demand when in reality it is not. Chapter 3 introduces the data used for characterizing variability in transit time. Methodology adopted for identifying non-unimodality in transit time distribution and Monte Carlo simulation, which is used to calculate the safety stock levels, and logistics cost is described in chapter 4. This chapter also describes the structure for comparative analysis done in order to understand the effect of bimodality on the above parameters of inventory management decisions. Chapter 5 then describes and discusses the results and the implications derived from the analysis. Finally, chapter 6 concludes with future research that could be done in order to dig deeper into the problem of bimodality in transit time distribution in maritime transportation.
Chapter 2. Literature Review

Numerous papers have addressed the issue of uncertainty in transit time and its impact on inventory management. While some papers make their case by deriving the distribution of demand over transit time from a given distribution of demand and transit time others suggest different distributions of demand of transit time.

The review is divided into three segments. The first segment briefly outlines the papers, which account for the effects of demand over transit time distribution on inventory decisions. The second segment deals with a few seminal papers that studied the effects of stochastic nature of transit time on the policies. Finally the third segment will illustrate the papers that discuss the robustness of the assumption of normality of demand over transit time when it is not the case.

Segment 1: Effect of Shape of transit time distribution on inventory management decisions

The effect of shape of transit time distribution on inventory management decisions has been studied in terms of coefficient of variation (CV), skewness and kurtosis. Lau & Zaki (1982) work with (R, Q) inventory policy. A (R, Q) system here refers to a policy of inventory management where Q quantity of units are ordered every time the inventory level goes below the reorder point which is R units. It is shown that when (R, Q) systems take into account the stock out risks, the shape of the demand over transit time distribution becomes extremely important. Stock out risks are the
risks associated with inventory going below a certain desired level.

In fact, a change in kurtosis can significantly change \( Q \) and \( R \). For instance at a service level of 80\%, their calculations suggest a safety stock level and hence reorder point that is 15 times larger in the case of skewness and kurtosis values at (1,8) as compared to that of a combination of (1,3). The CV of the distributions is kept constant.

The paper also discusses that although the values of stock outs vary significantly with changing shapes of transit time demand distribution, the corresponding shortage cost component is very small as compared to the magnitude of the total cost. However it is analyzed that even for small values of backordering cost, the annual shortage cost component of the total cost function become comparable in magnitude.

Tadikamalla (1984) studies and compares five different distributions with the same mean and standard deviation for comparing inventory management decisions related to a \( (R, Q) \) policy. The analyzed distributions constitute both symmetric (normal & logistic) and asymmetric (lognormal, gamma and weibull). The differences are evaluated by calculating the expected number of back orders, the reorder levels for a given service level, and the order quantities. It is concluded that for a low CV, the values of the parameters do not differ significantly across the various distributions. However, for large CV's the normal and the logistic distributions are not found to be suitable to model the transit time distribution.
because in this case the probability that it can take negative values grows largely. The asymmetric distributions in this case show a pronounced difference in the values of the parameters in the extreme tails. Finally, the paper claims that in the case when the distribution of demand over transit time is unknown, lognormal distribution is more appealing because of its computational simplicity based on the computational techniques.

Naddor (1978) also shows the effect of different distributions of demand, transit time and shortage costs on inventory decisions. The optimal decisions and costs are evaluated for two sets of distributions. The first set consists of the Poisson, beta, uniform and 2-point distributions while the other set is comprised of beta, negative-binomial, beta and 2-point distributions. For each of the two sets of distributions and two levels of transit time (0 or 3 days) and service levels (0.9 or 0.99), four different demand distributions are constructed. The difference between the parameter values of the policy and the optimal costs are more pronounced with large values of CV's as compared to lower values. However, such large CV's are only seen in extreme cases and that they are not usually observed in practice.

Segment 2: Working with stochastic nature of transit time while making inventory planning decisions

Kaplan (1970) and Ehrhardt (1984) draw their insights from a finite horizon dynamic programming inventory model created with the assumptions of no cross overs of orders and that transit time is independent of the number and size of
outstanding orders. The model is done for a single item whose demand is independent and identically distributed over time. Kaplan (1970) discusses how the optimal inventory policy changes with respect to two options related to ordering cost: either fixed, or proportional to order size. The paper uses a multidimensional minimization problem to calculate the optimal ordering policies. The model is set up to minimize the total cost which is a sum of convex expected holding and shortage costs, linear ordering cost and a fixed setup cost. Ehrhardt (1984) on the other hand denotes that the parameters of the dynamic model used by Kaplan (1970) were not related to the marginal transit time distribution and that sufficient conditions were not laid for optimality of myopic policies. Hence the paper tries to establish conditions for optimality of myopic base stock policies.

Liberatore (1979) develops an optimization model whose objective is to minimize the total expected cost per unit time in the presence of stochasticity in transit time. The model is shown to be similar to the standard EOQ model by considering the special case of deterministic transit time in the above-mentioned optimization model. However, no closed form solutions are given for the order quantity and reorder point.

While the approaches discussed in the papers above is very different from the one discussed in this thesis, one major take away is that stochasticity of transit time can impact inventory policies. While they look at optimal parameters of an ordering policy, on the other hand this thesis deals with looking into cost implications of
incorrect inventory policy, which is a result of misassumption of the distribution of transit time. Moreover, this thesis looks into continuous distribution of transit time as compared to discrete random variable assignment that have been used in the papers discussed in the above segment.

_Segment 3: Effect of incorrectly assuming normality in demand over transit time distribution_

Literature suggests variable levels of effects of assumption of normality in transit time demand distribution. For instance while Eppen and Martin (1988) discuss that using a normal approximation can lead to erroneous inventory management decisions, on the other hand Tyworth and O' Neill (1997) state that normal approximation is actually robust. However, Lau and Lau (2003) go on to show that normal approximations are not robust even for distributions with low CV's. It is possible that such a difference originates from the fact that Tyworth and O'Neill (1997) derive their conclusions from empirical data obtained for products from seven different industries. Hence the results probably cannot be generalized to all industries and modes of transportation used.

While stating that the demand over transit time is popularly modeled as a normally distributed random variable, Eppen and Martin (1988) demonstrate the potential errors in such an assumption. They mention that normality approximation is justified based on the central limit theorem and suggest that the confusion might be a result of a notion that demand over transit time is a convex combination of normal
random variables which may not necessarily be normally distributed. The errors are illustrated by calculation of reorder points and safety stock levels that are obtained by the normality assumption, which are found to be inconsistent with the expected probability of stock out. They use an exponential smoothing forecast model for calculation of demand in each period and assume that transit time is estimated from historical data available.

The paper effectively demonstrates the fallacy in incorrectly assuming normality in transit time through a simple example, which is described here. 2 cases are used to illustrate the errors in normality assumption when it is not true. Under a given demand that is normally distributed \( \sim N(100,10) \) units/day.

Case (a): Transit time is deterministic and is 4 days

Case (b): Transit time is 2 days 50% of the time and 4 days for rest of the time

Now, we know that the reorder point \( R \) is given by:

\[
R = E(X_{DoLT}) + k \sigma_{DoLT}
\]

Where, \( k \) = required service level, assumed to be at 95%. In other words the probability of stock out is less than equal to 5%.

\( X_{DoLT} \) & \( \sigma_{DoLT} \) are the mean and the standard deviation of demand over transit time.

Using the set of equations derived in Appendix (I) for the demand over transit time parameters, we get,

Case 1: When transit time is deterministic and is 4 days

\[
X_{DoLT} = (100)(4) = 400; \sigma_{DoLT} = (10)(\sqrt{4}) = 20
\]
Therefore, \( R = 4(100) + 1.645(20) = 433 \text{ units} \)

Case 2: Transit time is 2 days 50% of the time and 4 days for rest of the time
\[ X_{DolT} = (100)(3) = 300; \sigma_{DolT} = \sqrt{(3)(10)^2 + (100)^2(1)} = 101.5 \]

Therefore, \( R = 3(100) + 1.645(101.5) = 467 \text{ units} \)

This example shows that reducing transit time by half with a probability of 0.5 in turn increases the reorder point. This indicates that more number of units is required to ensure a probability of stock out of less than 5% when the transit time is reduced by half 50% of the time as compared to when the transit time is consistently 4 units. This being clearly incorrect, implies that assumption of normality in transit time, when in reality it is not, can lead to faulty inventory management decisions that can potentially affect logistics cost.

Lau and Lau (2003) use a Heuts, van Lieshout and Baken's (HLB) cost differential criterion to evaluate the performance difference between the use of normal approximation and the actual distribution to model transit time distribution. HLB cost differential is defined as a percentage of the relative difference between actual total cost that is incurred due to a normal approximation and the cost corresponding to reorder point and order quantity parameters obtained by implementation of the actual transit time demand distribution. The shape of this distribution is shown to only affect the holding and the backordering cost parts of the total cost function. This has been used in the analysis shown in this thesis when evaluating effect on logistics cost. The distribution used for demand over transit
time is beta distribution with same mean and standard deviation but with varying skewness and kurtosis.

In specific relation to this thesis, Tyworth and O’Neill (1997) & Bookbinder and Lorhdahl (1989) discuss bimodally distributed demand over transit time. The former use a symmetrical bimodal distribution (50% of products on time and the rest delayed). However, it is different from our observation of transit time data, which suggests different bimodal distribution characteristics. Assuming such a bimodal distribution similar to the paper does not encompass all possible kinds of bimodal distributions (e.g. symmetric and unsymmetric) as seen in the data used for this thesis.

On the other hand, Bookbinder and Lorhdahl (1989) discuss a different approach—“distribution-free approach” to tackling the problem of transit time demand distributions that are not normal. They propose a bootstrap method of sampling with replacement to estimate the transit time distribution that is finally used to determine the reorder point. Bootstrap method is a statistical procedure used for resampling from a given set of data, which could be done with replacement such that a particular data point can occur more than once in a sample created. This procedure is repeated for three service levels. The comparison between bootstrap procedure and normal approximation method is then used to evaluate the performance of both the methods in terms of total cost. It is interesting to see the discussion related to bimodally distributed transit time distributions under three
categories of skewness- symmetric, positively and negatively. The bimodal distribution is modeled as a mixture of normal distributions, which is similar to the method adopted in the thesis. However, the CV of the individual normal distributions used to create the mixture is restricted to values less than 0.2. The paper concludes that bootstrap method is more accurate in modeling demand over transit time that is not normal for non-standard distributions like uniform or bimodal distributions (negatively or symmetric). Another important conclusion conforming to previous studies states that the effectiveness of the bootstrap method as compared to normal approximation increases with increasing CV of the transit time demand. On the other hand it is also concluded that for standard shape distributions with a positive skewness (CV>=1), bootstrap estimate is as good as normal approximation method.

Similar to papers above, Bagchi, Hayya and Chu (1986) also show the impact of ignoring variability in transit time on safety stock levels and hence logistics cost with the help of a case study. This paper asserts that normal approximation for demand over transit time when in reality it is otherwise can lead to significant errors. One primary reason cited by the paper is that the true distribution for demand over transit time is usually skewed to the right, which is similar to our observation in the paper. The paper also discusses the possible reasons for use of a normality assumption. One of the reasons discussed is familiarity and extensive tabulation of normal distributions.
2.1 Summary of the review

It was ascertained that the distribution of demand over transit time can have effects on inventory management decisions in terms of the reorder quantities, safety stock levels desired and hence the total logistics cost. It is derived that distribution of demand over transit time is preferably approximated to a normal distribution because of familiarity with it as compared to other distributions. However, the review suggests that the effect of incorrectly assuming normality in demand over transit time distribution is not consistent. While on one hand it was shown that normal approximations are robust for low coefficient of variations (<0.45), on the other it was disproven even at much lower CV value of 0.2.

Also, bimodality has been observed to be occurring in transit time distributions in maritime transportation data that is available with us. However, such distributions have not been concentrated upon. This thesis aims to study the combined effect of bimodality in transit time distributions and a range of service levels. Moreover, this thesis throws lights on another aspect of variability in transit time pertaining to industry- the issue of using a deterministic value for transit time as a common practice for making inventory management decisions. We will observe different regions of comparison based service levels and the extent to which one is worse off by choosing a deterministic value for transit time or a normal approximation or the actual bimodal distribution. Finally, the empirical analysis used here is not specific to any industry.
Chapter 3. Understanding the Data

The entire dataset consists of around 125,000-container shipments from retailers, manufacturers and freight forwarders. As discussed earlier, we identified five different segments of ocean transportation pertaining to a one data record. As a quick reminder these segments include origin-to-origin port, origin port dwell, origin port to destination port, destination port dwell and finally destination port to final destination. While a few datasets contained information for all the above-mentioned segments, many of the shippers had limited information on the first two segments because many shippers do not take ownership of their shipment until it is ready to leave the origin port. The data sets were used to observe any occurrences of bi-(or multi)-modality in the transit time distribution. Transit time data was consolidated for every unique origin port - destination port pair. These unique pairs are called trade lanes in this thesis.

Initial work involved cleaning the dataset so as to exclude very low volume lanes. Specifically all lanes that carried less than 10 shipments over the year 2010-2011 were excluded from the data set. We removed records with inconsistent time stamps, locations and other fields.

There are a number of reasons for occurrence of bimodality on transit time distributions some of which are discussed below.
• Delays such as bumping of ocean freight at the origin port or weekend delays in offloading freight at the destination port could create bimodality. When cargo is bumped, it means that it misses the scheduled vessel and its transportation is delayed until the next available ship.

• A common characteristic among many (but not all) lanes, from the data available with us, that are not unimodal is that the shipments on the same lane are operated by different carriers. Bimodality on a trade lane occurs because of different mean transit times and the difference in the performance of different carriers on the lane.

• Some companies handle both high-value and low-value products. Hence they choose to use different service levels of shipping that have different transit times on the same trade lane. For instance, the company could decide to prioritize the high value product and hence bump the low value.

• Finally, a switch to slow steaming can also result in bimodality in transit time distribution. This is possible because the carriers now have the added flexibility to adjust their speeds at their will. In other words they can choose to speed up or slow down to fulfill the requirement of faster transit times or lower fuel costs. This range of speeds can thus account for cases of bimodality (or non unimodality) in the transit time distribution.
3.1 Some Basic Characteristics of Ports in the data

In order to understand the basic characteristics of transit time in the data available, we will investigate on two levels (i) the variability in the different segments of the retailer and manufacturer data (ii) evidence of existence of non-unimodality in transit time distributions.

3.1.1 Initial variability analysis of the lanes of the dataset obtained from the manufacturing firm

The total transit time along a lane is the sum of transit time of the five individual segments that has been discussed before. Therefore in order to understand variability in the total transit time it is necessary to break down the total transit time variability into variability across the different segments. This is done with the help of scatter plots to indicate the mean and variability in terms of coefficient of variation (CV) of the lanes. CV of the lanes is defined as the ratio of the standard deviation to the mean of the distribution. The figure below contains plots, which have the mean of the transit time of the segment pertaining to the lane on the X-axis while the corresponding CV is shown in the Y-axis. Variability in transportation time is measured in terms of CV. The 3 different shapes on the plot correspond to different levels of coefficient of variation that spread between 0 and 2 units.

At the outset it can be observed from figure 3-1 that the maximum CV of the total transit time in end-to-end transportation is 0.35. If the shipper decides to base inventory management decisions only on this plot then he/she would probably
reason that the dominant parameter that effects decision making process for inventory management is the mean of the transit time rather than the CV.

However, breaking down the transportation sequence into its segments tells a different story. It indicates the segments of very high variability, which could be different from the one with higher mean values in the total transit time. It also helps the shipper identify the exact segments that need improvement to curb large amounts of variability and hence could result in different policies for inventory management in terms of selection of carrier or travel route of the shipment or even stocking policies.

Fig 3-1. Average and Coefficient of Variation of the end-to-end transportation of the lanes in the data set from the manufacturing firm
The plots for individual segments are grouped together in figure 3-2. On comparison between the origin part (which includes segments from origin to origin port & origin port dwell) and the destination part (which includes segments destination port dwell & destination port to final destination) it can be said the former is more variable than the latter.

Few reasons for this could include inferior loading and unloading operations probably due to lower automation of the operations in the origin side as compared to the destination side. It could also be because of a larger congestion at the origin ports as compared to destination port sides thus leading to lower destination port dwell times. Or it could also be a result of the fact that the shipper takes responsibility of the shipment at the destination side as compared to the fact that the freight forwarder is responsible for the shipment at the origin side and hence could be handling contracts from various shippers.

It can be seen that port-to-port segment have the least variability amongst all the 5 segments. Origin port dwell segment on the other hand is shown to be the most variable segment of all.
Fig 3-2. The graphs below indicate the Average and Coefficient of Variation corresponding to the lanes from the graph above. [Namely, i) Origin to origin port ii) Origin port dwell iii) Port to Port iv) Destination port dwell v) Destination port to final destination]

LEGEND

- $0 \leq CV \leq 1.25$
- $1.25 < CV \leq 2$
- $CV > 2$

Average transit time on a trade lane
Average transit time on a trade lane
1. Average transit time on a trade lane

2. Average transit time on a trade lane
The retailer data on the other hand does not contain information about the segments from origin to origin port and the port dwell. Hence the plots for the retailer data contain variability information about the port-to-port segment and the combined information of destination port dwell and destination port to final destination.

*Fig 3-3. The graphs below indicate the Average and Coefficient of Variation for the retailer. [Namely i) Port-to-Port ii) Destination port dwell & Destination port to final destination]*
As, in the case of the manufacturer data, the port-to-port segment for the retailer data is less variable than the destination part segment. It is also true that latter part of the segment is more variable for the retailer as compared to that of the manufacturer.

3.1.2 Geographical information of the origin port location in terms of non-unimodality in the lanes

The origin port locations of the container information from various types of companies are listed below. Table 3-1 also contains information about the proportion lanes in the particular region that are not unimodally distributed. The results shown in the table below combine information from the manufacturer, retailer and freight forwarder for all imports into US.
Table 3-1 Global distributions of origin port locations for freight into US

<table>
<thead>
<tr>
<th>Origin Port Locations</th>
<th>Total no. Of lanes</th>
<th>Not Unimodal lanes</th>
<th>% not unimodal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asia</td>
<td>289</td>
<td>45</td>
<td>16</td>
</tr>
<tr>
<td>EU</td>
<td>33</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>Americas</td>
<td>17</td>
<td>8</td>
<td>47</td>
</tr>
<tr>
<td>Africa</td>
<td>21</td>
<td>3</td>
<td>14</td>
</tr>
</tbody>
</table>

It was also observed that the cumulative percentage of volume of shipments carried by the non-unimodal lanes constitute about 80% of the total volume, 74% of which originate from Asia. There could be a number of reasons for this observation. One reason probably pertains to the fact that a majority of imports into US come from Asia. Moreover, supply chains from Asia to US are relatively longer than the ones corresponding to other importing regions. This implies that the total transit time is longer and has a greater probability of variability, which could show in bimodality, as compared to the chains from originating in other regions.

3.2 Summary of the chapter

It was observed that transit time distribution in maritime transportation is variable and this variability, measured in form of coefficient of variation, can range between 0 and values even greater than 2 units (which means large variations). This variability should be investigated on individual segments instead of the total transit..
time. The segments in the origin part are found to be relatively more variable than that of the segments in the destination.

The number of non-unimodal lanes was inspected on trade lanes for combined retailer, manufacturer and freight forwarder data. It was observed that even though the number of lanes that are non-unimodal are a small percentage of the total number of lanes, the volume of shipments carried on these lanes are large in magnitude compared to the total amount. Hence it is important that we investigate the effect of non-unimodality on inventory management decisions in form of logistics cost.
Chapter 4. Analysis

Each transaction in the dataset has timestamps associated with major milestones during its travel. These include origin/destination information, and the ocean carriers responsible for the transportation. This information is used to characterize the transit time distribution for an origin-destination pair. The method adopted to analyze the effect of non-unimodality in transit time distribution involves identifying whether the distribution is unimodal or not, followed by simulation to understand the effect of uncertainty on logistics cost.

4.1. Characterizing transit time distribution

A hypothesis test is designed to determine if the transit time distribution in the empirical data is unimodal in nature. The null hypothesis is that the distribution is unimodal with the alternative hypothesis is it has more than one mode. The null hypothesis is rejected at a significance level of 5%. The hypothesis test was done using bootstrapping with replacement for a thousand samples. A test of unimodality was used on each sample to obtain a dip value. The hypothesis is accepted (or rejected) at the above significance level if the number of samples that pass the test of unimodality or are found to be unimodal is more (or less) than 50 out of the 1000 samples.
4.1.1 Test of Unimodality- Hartigan's Dip Test

Hartigan and Hartigan (1985) introduced a statistical metric called "dip" that can be used to measure the departure of a given distribution from unimodality. This test calculates a dip statistic, which is defined as the maximum difference between the empirical distribution and the unimodal distribution that minimizes the maximum difference. The dip statistic evaluates the departure of a given distribution from the best fitting unimodal distribution. In case of a unimodal distribution $f$ with a cumulative distribution function $F$ and mode $m$, it is known that $F$ is convex on the interval $(-1,m)$ and concave on the interval $(m,1)$. In other words at the right hand side of the mode, the density is non-increasing with a non-positive derivative and behaves in an opposite manner on the left hand side of the mode. Fig. 4-1(A) illustrates the regions in a unimodal distribution. When the distribution is unimodal the dip statistic is zero.

*Fig4-1 (a) Convex and Concave regions of a unimodal distribution*
However, if the empirical distribution has more than one mode the cumulative distribution has several regions of convexity and concavity.

*Fig 4-1 (b) Multiple Convex and Concave regions of a bimodal distribution*

This empirical distribution is then modified at different levels of "dip" until it forms a unimodal distribution. The larger the dip needed, the larger the departure from unimodality. This test is done on the transit time values of lanes on container level data available with us. The lanes as indicated earlier are identified as unique origin destination pairs for all the dataset present with us.

4.1.2 Evidence of non-unimodality in transit time distribution

It was observed that non-unimodality (or distributions with multiple modes in their transit time distribution) is prevalent across different kinds of shippers. However, they occur at different levels for retailer, manufacturer and freight forwarder data. In case of the retailer, non-unimodal distributions occurred in only 2-4% of origin-destination lanes but account for 12% of shipment volume. On the other hand, for
the manufacturer, the corresponding number averaged for 22% of lanes accounting for 75% of shipment volume. For the freight forwarder, non-unimodality was observed for 24% of the lanes, which is equivalent to 85% in shipment volume. A common norm across many trade lanes indicated that the transit times are heavy right-tailed. Examples of histograms of empirical distributions that have multiple modes and long right tail are shown below:

*Fig4-2 Bimodal and Long Right Tailed distribution indicated in distribution of transit time (in days)*

The above observations indicate that even though non-unimodality is observed in a relatively small percentage of lanes, it is not advisable to ignore them. The reason lies in the fact that these lanes account for very high volume lanes that could severely impact logistics cost.

4.2. Simulating Cases of Bimodality and its Effects

In order to evaluate the effects of non-unimodality in distribution of transit time, a comparative analysis method is applied on three different cases (I, II, III), which are described in the introduction chapter.
4.2.1. Variability in transit time distribution and demonstration of the use of normal approximation in Hadley Whitin formula.

For a variable demand and transit time distribution let,

- \( \text{DoLT} = \text{Demand over transit time, a continuous random variable} \)
- \( E[\text{DoLT}] = \text{Mean of demand over transit time} \)
- \( \sigma_{\text{DoLT}} = \text{Standard Deviation of demand over transit time} \)
- \( D = \text{Demand per unit time period (eg. Units/day), a continuous random variable} \)
- \( LT = \text{Length of transit time in time periods (eg. days)} \)
- \( E[LT] = \text{Mean transit time} \)
- \( \sigma^2_{LT} = \text{Variance of transit time} \)
- \( E[D] = \text{Mean demand during one time period} \)
- \( \sigma^2_{D} = \text{Variance of demand during 1 period} \)
- \( k = \text{service level, which is the level at which the probability of demand is always less than the quantity ordered} \)

Under the assumption that observed demand and transit time are independent and that demand is uncorrelated between the transit time periods the mean and variance of demand over transit time is given by,

\[
E[X_{\text{DoLT}}] = E[LT]E[D] \\
\sigma^2_{\text{DoLT}} = E[LT]\sigma^2_{D} + (E[D])^2\sigma^2_{LT}
\]
Performing a random sum of random numbers derives the above equations. The random numbers are the demands observed and the random sum is the sum of the random demands over the transit time period. Detailed derivation for the above equations is shown in Appendix (I).

The normality assumption in case of Hadley Whitin formula, which is discussed in section 4-2 of Hadley and Whitin (1963) is demonstrated below. The average annual variable cost is given by the sum of purchase cost, ordering cost, holding cost and back order cost.

Let us introduce a new set of notations as below:

- \( A = \text{ordering cost} \)
- \( Q = \text{order quantity} \)
- \( R = \text{reorder point} \)
- \( v = \text{unit cost of the item which is independent of } Q \)
- \( h = \text{per unit inventory carrying charge measured in per unit time and per unit dollar quantity held} \)
- \( b = \text{per unit back ordering cost} \)
- \( s = \text{safety stock} \)
- \( \text{DoLT} = \text{demand over transit time} \)
- \( g(\text{DoLT}) = \text{marginal distribution of demand over transit time} \)
- \( G(\text{DoLT}) = \text{cumulative distribution of demand over transit time} \)
- \( E[\text{DoLT}] = \mu \text{ and } \sigma_{\text{DoLT}} = \sigma \)
\[ D = \text{average annual demand} \]

\[ TC = \text{Total Cost} \]

Then,
\[
TC = vD + \frac{AD}{Q} + hv\left[\frac{Q}{2} + R - \mu\right] + \frac{bD}{Q} \left[\int_{R}^{\infty} xg(DoLT) - R \ast G(DoLT)\right]
\]

The four terms in the total cost equation are purchase cost, ordering cost, holding cost and back ordering cost where,

- **Purchase cost**: Given by the product of per unit cost and the number of units ordered
- **Ordering cost**: Cost of ordering \( Q \) units when the average demand is \( D \)
- **Holding cost**: Given by product of annual holding cost and average on hand inventory
- **Back-ordering cost**: Given by the product of backordering cost per unit and number of back orders

The detailed derivation of the equation is shown in Appendix (II).

If \( g(DoLT) \) is assumed to be normally distributed such that \( g(x) \sim N(x; \mu, \sigma) \) then

\[
\int_{R}^{\infty} xg(x) = \int_{R}^{\infty} x \ N(x; \mu, \sigma) = \int_{R}^{\infty} \frac{x}{\sigma} \phi\left(\frac{x - \mu}{\sigma}\right) dx
\]

On substituting \( \frac{x - \mu}{\sigma} = y \) we get,

\[
\int_{R}^{\infty} xg(x) = \sigma \int_{\frac{R - \mu}{\sigma}}^{\infty} y\phi(y)dy + \mu \int_{\frac{R - \mu}{\sigma}}^{\infty} \phi(y)dy
\]
Now, we denote \( \phi \left( \frac{R - \mu}{\sigma} \right) = Service \ level \). This is defined as a level such that the demand over transit time is always below the inventory carried. Let \( k = \phi^{-1}(Service \ Level) \). As shown in Appendix (II) safety stock (SS) in this system is \( R - \mu \). Therefore, \( SS = k\sigma_{DOLT} \) and the reorder point (R) is given by:

\[
R = E(X_{DOLT}) + k\sigma_{DOLT}
\]

Substituting values for \( E(X_{DOLT}) \) and \( \sigma_{DOLT} \) we obtain that:

\[
R = E[LT]E[D] + k \sqrt{(E[LT]\sigma_D^2 + (E[D])^2\sigma_{LT}^2)}
\]

The above equation for the reorder point is derived when demand over transit time is approximated to be a normal distribution. This approximation has been frequently used in theory because normal distributions are more tractable than other distributions. For the purposes of this thesis any mention of Hadley Whitin formula should be considered to be equivalent to normal approximation of demand over transit time.

### 4.2.2 Creating bimodal distribution from mixture of two normal distributions

In order to simulate cases of non-unimodality in transit time data the simplest case of multimodality -bimodality is assumed. Bimodality can be simulated as a mixture
of two normal distributions. Such applications of mixing two normal distributions to create multimodal distributions has been seen in many fields like economics, finance, astronomy. In the field of inventory management, Bookbinder and Lordahl (1989) create bimodal distributions of demand over transit time by mixing two normal distributions.

A bimodal distribution is simulated at a certain mixture rate $\pi$ such that any point in the resultant distribution lies in the first normal distribution with a probability $\pi$ and in the second distribution with a probability $1- \pi$.

The probability density function of the resulting mixture distribution of the transit time is obtained as a linear combination of two normal distributions such that:

\[
\begin{align*}
    f(x) &= \pi f_1(x) + (1 - \pi)f_2(x) \quad \ldots \text{the PDF} \\
    F(x) &= \pi F_1(x) + (1 - \pi)F_2(x) \quad \ldots \text{the CDF}
\end{align*}
\]

Where,

\[
    f_i(x) \text{ has a mean } \mu_i \text{ and std. dev } \sigma_i
\]

\[
    \text{and } 0 \leq \pi \leq 1
\]

$\pi$ is defined as the mixture rate such that a given transit time value lies in the first normal distribution with a probability $\pi$ and with a probability $1- \pi$ in the second normal distribution. An example of such a mixture distribution is as follows.
Let us consider two normal distributions with the following parameters and a mixture rate of 0.5 associated with the first distribution i.e. \( \tau = 0.7 \). Fig. 4-2 shows the corresponding bimodal distribution of transit time.

**Table 4-1 Parameters of two normal distributions used to create a bimodal distribution**

<table>
<thead>
<tr>
<th></th>
<th>Normal 1</th>
<th>Normal 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>27</td>
<td>17</td>
</tr>
<tr>
<td>SD</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

**Fig 4-3. An example of a bimodal distribution created by mixing two normal distributions**
4.2.3. Derivation of the mean and standard deviation of the mixture distribution

We know that

\[ f(x) = \pi f_1(x) + (1 - \pi)f_2(x) \]

Where

\[ f_i(x) \text{ has a mean } \mu_i \text{ and std. dev } \sigma_i. \]

\[ 0 \leq \pi \leq 1. \]

Also let the resulting distribution bimodal distribution has a mean \( \mu \) and standard deviation \( \sigma \).

Now, we also know that,

\[ E[x^k] = \pi_1 E_1[x^k] + \pi_2 E_2[x^k] \]

\[ Variance = E[x^2] - (E[x])^2 \]

Substituting \( k = 1 \) implies

\[ \mu = \pi \mu_1 + (1 - \pi) \mu_2 \]

Similarly, \( E[x^2] = \pi (\mu_1^2 + \sigma_1^2) + (1 - \pi)(\mu_2^2 + \sigma_2^2) \)

Therefore,

\[ \sigma^2 = \pi (\mu_1^2 + \sigma_1^2) + (1 - \pi)(\mu_2^2 + \sigma_2^2) - (\pi \mu_1 + (1 - \pi) \mu_2)^2 \]

For instance for the above mentioned example mixture distribution:

Mean = \((0.7 \times 27) + (1 - 0.7) \times 17\) = 25.4

Standard Deviation =

\((0.7(27^2 + 2^2) + (1 - 0.7)(17^2 + 2^2) - (0.7 \times 27 + (1 - 0.7) \times 17)^2)^{1/2}\) = 5.851

55
4.2.4. Level of Bimodality

In order to understand the effects of multimodality through bimodal distributions, a concept called level of bimodality is introduced. For the purposes of this thesis the term level of bimodality refers to the normalized difference between the two means of the normal distributions used to create the mixture bimodal distribution. Dividing the difference between the two means by the mean of the resultant bimodal distribution gives us the level of bimodality for the distribution. For instance in the above example the level of bimodality is \((22 - 17)/25.4 = 0.2\). Different cases of bimodal distributions are created based on the difference between the two means of the normal distributions or levels. It is done by changing the mean of one of the distributions while keeping the mean of the other constant.

4.3. Simulation Model used for calculation of safety stock

A part of the analysis deals with understanding the effects of normal approximation on the amount of safety stock required to account for the variability of the demand over transit time. Simply put, the variability of transit time impacts a company's safety-stock -- the chance of a delay in shipping translates into a need to hold more inventory.

The simulation assumes that transit time and demand are independent of each other. It is also assumed that we observe a normally distributed demand with a mean of 5 and standard deviation of 1 unit. The simulation is set up to generate a random value between 0 and 1 corresponding to the cumulative distribution function in order to select a value of transit time. It is done by discretizing the
transit time values such that \( t = F(t) - F(t-1) \), where \( F(t) \) is the CDF of the mixture distribution and \( t \) is the transit time value. The distribution of demand is used to generate values, which are then compounded to obtain demand over transit time. A customer (cycle) service level of 95\% is used to obtain the safety stock value for the compound distribution of demand over transit time. The resulting value is the average of 10000 runs of Monte Carlo simulation.

4.3.1 Safety stock Analysis

The value of safety stock obtained from the actual distribution using the simulation process described above is compared against the corresponding values we get from using Hadley-Whitin equation with normal assumptions. The parameter used to compare the two values is Change in Safety Stock (CSS). CSS is defined as the difference between safety stock from Hadley Whitin equation and that from the bimodal distribution from the simulation. Percentage change in safety stock (PSS) is given by the ratio of change in safety stock to that of the value calculated by using Hadley Whitin equation. The results are analyzed by changing the difference between the two normal distributions with a constant mixture rate and constant standard deviation.

4.4. Simulation Model used for calculation of logistics cost

The second part of the analysis is to understand the effect of distribution of transit time on inventory management decisions includes evaluation of cost under three different cases as discussed above. Inventory calculations involve simulating...
demand over transit time that is observed in reality and then comparing the
inventory levels calculated separately for each of the cases I, II & III to obtain the
costs that the firm incurs in case of stocking excess or not enough of the product.
The analysis of inventory cost is based on a given critical ratio (that in effect is the
ratio of under-stocking to over-stocking cost).

We will discuss the simulation on two levels (1) Simulation structure with inventory
policy (2) Derivation of cost model used in the simulation.

4.4.1 Simulation structure with inventory policy used

The inventory policy used is such that:

1. Order up to level – If the ending inventory level goes below R by x units we
   order R-x units

2. Complete back ordering – Any demand that is not fulfilled in a time period is
   backordered and is satisfied in the next time period.

3. Frequency of ordering – Ordering is done at the end of every unit time
   period.

Order up to level when the inventory management system uses the actual
distribution is the optimum amount calculated based on the service level (critical
ratio) that is being observed. The order up to level when the system uses normal
approximation uses Hadley Whitin formula, which derived in section 4.2.1. The case
of deterministic transit time uses an order up to level, which is the expected value of
demand over transit time. So, the amount ordered at the end of every time period to reach the respective order up to levels is the same for all the three cases. Therefore the ordering and purchase cost is the same across all the three cases. Hence the net difference in the cost is given by the under stocking and over stocking costs.

This is illustrated in the following example. Let us assume a the following order up to levels and demand such that:

Table 4-2 Illustration of the simulation model for calculation of logistics cost

<table>
<thead>
<tr>
<th>Case of inventory management system</th>
<th>Order up to level (= Starting Inventory)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Distribution</td>
<td>20</td>
</tr>
<tr>
<td>Normal Approximation</td>
<td>25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Demand</th>
<th>Actual Distribution</th>
<th>Normal Approximation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Amount left after satisfying demand</td>
<td>Amount to be ordered in the next period</td>
</tr>
<tr>
<td>10</td>
<td>(20-10) = 10</td>
<td>(20-10) = 10</td>
</tr>
<tr>
<td>12</td>
<td>(20-12) = 8</td>
<td>(20-8) = 12</td>
</tr>
<tr>
<td>25</td>
<td>(20-25) = -5</td>
<td>(20-(-5)) = 25</td>
</tr>
</tbody>
</table>

Where,

\[ \text{Amount left after satisfying demand (Ending inventory level)} \]

\[ = \text{Inventory level} - \text{Demand for the time period} \]

\[ \text{Amount to be ordered in the next period} \]

\[ = \text{Order up to level} - \text{Ending inventory level} \]
4.4.2 Derivation of cost model

A stationary infinite horizon inventory model is considered, in which the optimal base stock is calculated from the critical ratio. The critical ratio discussed is considered to be equivalent to service level targeted by the firm. The derivation of the cost model used for the evaluation is adapted from the cost structure derived by Zipkin (2000) when transit time is a random variable. This cost structure is demonstrated below.

List of Notation:

- \( \gamma \) = Discount cost rate \( 0 < \gamma \leq 1 \) on fixed ordering cost
- \( L(t) \) = Transit time that randomly changes over \( t \)
- \( h(t) \) = Inventory holding cost rate at time \( t \)
- \( b(t) \) = Backorder penalty cost rate at time \( t \)
- \( x(t) \) = Inventory position at time \( t \) before ordering
- \( y(t) \) = Inventory position at time \( t \) after ordering
- \( C(t, x(t)) \) = Inventory holding or backorder cost at time \( t \)
- \( C(t, y) \) = Expected inventory backorder cost
- \( [a]^+ \) = Maximum of \( (a, 0) \)
- \( D(t) \) = Demand at time \( t \)
- \( D[t, t + L) \) = Transit time demand starting at \( t \)
- \( T \) = Time horizon which could be finite or infinite
- \( z(t) \) = Order size at time \( t \)
- \( CSL \) = Customer (Cycle) Service Level
a) Model for a time period of length say T

Now given that the order placed at time t will arrive at some future time denoted by \( t + L(t) \). The decision making process that helps in the formulation of the cost model is comprised of two steps. At time \( t < T \),

Step 1: Net inventory is observed which is equal to \( x(t) \)

Step 2: An order size \( z(t) \) is decided

Under the assumptions that:

1. No cross overs, which means that, orders arrive in the same sequence that they were issued. Mathematically, if \( t \) is the time in which the order was issued then
   \[
   t + L(t) \leq (t + 1) + L(t + 1)
   \]

2. Transit time \( L(t) \) is independent of demand

3. Stationary transit time would imply that \( L(t) \) has the same distribution over time and is denoted by random variable \( L \)

From Zipkin (2000, p. 409), the expected inventory back order cost \( C(t, y) \) after the order is placed in step 2 can be written as

\[
C(t, y) = E[y^T \hat{C}(t + L, y - D[t, t + L])]
\]

where,

\[
\hat{C}(t, x) = h(t)[x - D]^+ + b(t)[D - x]^+
\]

From definition inventory position observed just after ordering \( x(t+1) \)

\[
x(t + 1) = y - D[t, t + L]
\]
Given stationary transit time and infinite horizon the cost function becomes

\[ C(y) = E[yL\hat{\theta}(y - D)] \]

Now if we assume an average ordering cost (i.e. no discounting of cost or \( \gamma = 1 \))

\[ C(y) = E[\hat{\theta}(y - D)] \]

And by definition we know that for any inventory level \( i \)

\[ \hat{\theta}(i) = h[i - D]^+ + b[D - i]^+ \]

b) Optimal ordering quantity for a single or unit time period model-Newsboy Vendor Problem

Now in a single or unit period of time, the quantity that maximizes profit or minimizes the total cost for a firm the newsboy or the newsvendor problem gives. The details of the model are as follows.

Let \( Q \) be the quantity that is ordered and since this is a single period model \( z(1) = Q = x(1) = y(1) \).

Therefore the total cost of ordering \( Q \) when a demand \( r \) is observed is given by:

Total cost of ordering \( Q \) units = Cost of overstocking + Cost of understocking

Therefore,

\[
C(Q) = h \int_0^Q (Q - r)f(r)dr + b \int_Q^{\infty} (r - Q)f(r)dr
\]

\[
\frac{dC(Q)}{dQ} = hF(Q) - b(1 - F(Q))
\]
Now, at optimal $Q^*$, $\frac{dC(Q)}{dq} = 0$ that is setting the first order condition to 0. This also means the expected marginal profit from ordering one unit less is equal to the expected marginal cost of ordering an extra unit.

Therefore,

$$hF(Q^*) = b(1 - F(Q^*))$$

Rearranging would give that

$$F(Q^*) = \frac{b}{b + h} = CSL$$

But by definition we know that $F(Q^*) = \text{Customer service level (CSL)}$ or the level at which the probability of demand is always less than the quantity ordered $Q$.

Coming back to infinite horizon problem, the general cost equation becomes

$$C(y) = E[h[y - D]^+ + b[D - y]^+]$$

Dividing and subtracting a term $(b+h)$ gives

$$= (b + h)\left\{ \frac{h}{b + h} (E[y - D]^+) + \frac{b}{b + h} (E[D - y]^+) \right\}$$

$$= (b + h)\{(1 - CSL) * (E[y - D]^+) + CSL * (E[D - y]^+)\}$$

**Effective cost** = \{(1 - CSL) * (E[y - D]^+) + CSL * (E[D - y]^+)\}

For a given value of $(b+h)$ we can have various values for CSL. However, since CSL is a probability value it can only range anywhere between from $[0,1]$. This enables us to bind the calculations for the cost for values of CSL in the range of $[0,1]$. Therefore we do not have to deal with infinite options for $(h,b)$ values. This makes the analysis simpler.
A stochastic inventory model and simulations for the calculation of the cost and safety stock level capture the impact of variability caused due to bimodality in transit times. For non-unimodal lanes, we model the transit time as a bimodal distribution for tractability and for consistency with our data (as most of the lanes that were not unimodal tended to have bimodal transit times). As discussed in this section, mixing two normal distributions creates the bimodal transit time distributions.
Chapter 5. Results

5.1 Effect of distribution of transit time on safety stock values

The effect of transit time distribution on the safety stock values are observed with respect to changing levels of bimodality while keeping the standard deviations constant and mixture rate at 0.5. Then, real data was used to evaluate the effect on four example lanes selected so as to represent different types of non-unimodal lanes—the very long-tailed and bimodal lanes.

The trend was observed to be decreasing and then increasing with respect to the percentage change in safety stock values over increasing levels of bimodality. The regions of negative values of safety stock imply that the normal approximation in Hadley Whitin formula underestimates the value of safety stock as compared to that of the actual distribution. For the positive values, the normal approximation overestimates.

5.1.1. Effect of increasing the level of bimodality-Case II vs. Case III

This section explores trends in percentage change in safety stock with increasing levels of bimodality between the case when the inventory management system approximates transit time distribution to normal and that when it uses the actual distribution. The results are graphically shown below in Fig.5-1. The increasing level of bimodality is obtained by increasing the mean of one of the normal
distributions in the mixture by 2 units at every step. The mean of the other distribution is kept constant.

*Fig 5-1. Percentage change in safety stock using Hadley-Whitin (HW) equation versus actual distribution (bimodal) of transit time*

The plot above shows level of bimodality in the x-axis and the percentage change in safety stock on the y-axis. Hadley Whitin formula with normal approximation underestimates the amount of safety stock when the level of bimodality is less than 0.45 and then later over estimates for the rest of the instances of levels of bimodality.

Three main regions can be noted from this plot:

1. *Region A-B:* This initial dip corresponds to a transition from a nearly null difference in the percentage change in safety stock (A) to a difference of a larger magnitude of about 9% (B).
Point A is very close to zero because the difference between the means of the two normal distributions at this point is also zero. Therefore, the resultant of the mixture of the two normal distributions is actually equal to a normal distribution as shown in fig. 5-2(A).

Whereas, the change of a larger magnitude in B is because of a larger increment of the difference between the means and hence level of bimodality. Graphically, the resultant bimodal distribution goes from 5-2(A) to 5-2(B) when the change happens from 0 difference to the first drop.

*Fig 5-2.* Resultant bimodal distribution and the normal approximation for the cases when level of bimodality is 0 and 4 (fig A and B)

2. **Point C:** This is a point when Hadley Whitin value crosses the axis indicating a second point (besides A), which corresponds to an insignificant percentage change in the safety stock.
3. *Region C-D:* Also indicated in the graph is that the Hadley Whitin equation with normal approximation method overestimates more and more as the difference between the means increases.

At this point it is important to discuss the reason why we observe a behavior of an initial underestimation and then overestimation by the normality assumption. In the extreme case when the level of bimodality goes to 0.9 the resultant bimodal distribution and the normal approximation of the distribution looks as indicated in the Fig 5-3.

*Fig 5-3. Resultant bimodal mixture and normal approximation of the mixture distribution.*

![Resultant bimodal mixture and normal approximation of the mixture distribution.](image)

Fig. 5-3 shows that the normal distribution assumes much higher values of transit time with a greater probability as compared to that of the bimodal mixture (e.g., the mean value of 31 for the distribution). It also means that increasing levels of bimodality are equivalent to increasing coefficient of variations. Hence, normality
assumption in Hadley Whitin (HW) equation overestimates safety stock values as compared to the case of the bimodal distribution.

Although the plot in Fig. 5-1 shows the result of one set of simulation experiment, running the same for different values of demand yields the same shape for the plot. However, the level below which the results of HW equation transition from negative to positive (point C) is different for different simulations. Therefore, while the trend can be ascertained, the characteristics of the point at which the normality assumption transitions from underestimation to overestimation needs to be researched further.

5.1.2 Using available data to analyze the effect of non-unimodality in safety stock values

Following the above analysis, real data was used to gauge the effect of normal approximation on safety stock levels calculated for a given service level. For initial analysis, four lanes were selected from amongst the ones that did not pass the Hartigan’s dip test of unimodality and are hence not normally distributed.

These four lanes were selected so as to compare a range of different shapes occurring in non-unimodal lanes—very long tails and bimodal lanes. As of now, a large change in safety stock is defined as any change, which is greater than 5%. The results of simulation for the four lanes along with their histograms are tabulated.
below. Also, in the columns of the table are the mean and the standard deviation of transit time on the trade lanes corresponding to the histograms.

Table 5-1 Simulation Results of effect of non-unimodality on safety stock values

<table>
<thead>
<tr>
<th>Histogram</th>
<th>Mean</th>
<th>Stdev</th>
<th>Case III</th>
<th>Case II</th>
<th>Change</th>
<th>% change</th>
<th>HW estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>110A</td>
<td>12.7</td>
<td>2.0</td>
<td>18.39</td>
<td>17.47</td>
<td>-0.92</td>
<td>-5.25</td>
<td>Under</td>
</tr>
<tr>
<td>1051A</td>
<td>14.6</td>
<td>2.2</td>
<td>13.54</td>
<td>19.72</td>
<td>6.18</td>
<td>31.34</td>
<td>Over</td>
</tr>
<tr>
<td>1052A</td>
<td>21.0</td>
<td>2.7</td>
<td>19.93</td>
<td>23.55</td>
<td>3.62</td>
<td>15.37</td>
<td>Over</td>
</tr>
<tr>
<td>1045A</td>
<td>25.4</td>
<td>3.4</td>
<td>29.73</td>
<td>29.36</td>
<td>-0.37</td>
<td>-1.26</td>
<td>Almost same</td>
</tr>
</tbody>
</table>

As indicated in the table above, there can be large deviations of the values of safety stock due to normal approximation of the transit time distribution. Both of the cases of large deviations (highlighted in bold) are caused when Hadley Whitin formula
with normal approximation overestimates the safety stock levels. The largest deviation was caused in the right tailed distribution lane followed by a bimodal distributed pattern in the transit time.

However, the last case was a bit of surprise. A small bump towards the right is expected to cause a good amount of deviation between the normal approximation of the HW equation and using the actual distribution but the results suggest otherwise. This needs to be investigated further. The above results suggest that there might be a possible relation between the position of the mean and the median in the distribution of transit time and change in percentage change in safety stock.

5.1.3. Effect of increasing the level of bimodality-Case I vs. Case III

The percentage difference between the safety stock values is always underestimated in Case I when the inventory system is designed to completely ignore variability as compared to Case III. This is shown in the plot in fig. 5-4 below.

*Fig 5-4. Percentage change in safety stock using deterministic value versus actual distribution (bimodal) of transit time*
The minimum value by which ignoring variability underestimates the safety stock level is around -50%. The under estimation becomes worse with increasing levels of bimodality.

Case I underestimates the amount of safety stock because it completely ignores variability in lead-time implying no uncertainty in the transit time. Hence the safety stock inventory carried is lower. This estimation becomes worse with higher levels of bimodality because variability in the actual distribution increases along with it. Hence the consequences of using an inventory management system that ignores variability instead of the actual distribution become worse.

5.2 Effect of variability (including bimodality) in transit time distribution on logistics cost

As a recap, the effect of transit time distribution on logistics cost is compared over three cases (I, II & III) which represent the scenarios of completely ignoring variability, assuming a normal approximation with Hadley Whitin formula and using the actual bimodal distribution respectively. The change in costs was observed with respect to two shapes of bimodality – symmetric and non-symmetric or skewed. For symmetric distribution of transit time, the effect was compared with respect to changing levels of bimodality and service level or critical ratio.
Changing the mixture rate $\tau$ for a specific level of bimodality and thus realizing different skewness and kurtosis values create the skewed distribution. The results for skewed distribution highlight the initial results. Further research is needed to uncover more trends with respect to the shape parameters- skewness and kurtosis.

5.2.1 Comparing Case I & Case II

Generally, we observe that ignoring variability in transit time could have grave impacts on the inventory cost. However, digging deeper reveals that this statement is not true under all values of critical ratios and levels of bimodality. Based on the results, it is recommended that shippers:

- may choose to ignore variability and avoid updating their inventory management system to account for variability for (1) low service levels (critical ratio $\leq 0.6$) for all levels of bimodality and (2) intermediate service levels (critical ratio between 0.6 and 0.8) but only for low levels of bimodality
- should consider updating their inventory management system to account for variability in order to allow for approximating the distribution to normal for the remaining combinations of service levels and levels of bimodality
- should absolutely update the inventory management system for very high service levels (critical ratio $\geq 0.8$) under all levels of bimodality.
The results of the relative difference between the costs of the two cases are plotted below in Fig. 5-5. Relative difference is obtained by dividing the difference of the costs, between the cases when the inventory management system is designed to assume a constant number for transit time which is equal to the mean of the distribution (Case I) and when it approximates the distribution to normal (Case II), by the cost obtained in the latter (i.e. Case II).

*Fig 5-5. Difference between logistics cost of Case I (NV) and Case II (HW)*

In order to understand in terms of positive and negative values, if the relative difference is negative it implies that the cost obtained by ignoring variability (Case I) is more expensive than considering a normal approximation of the transit time distribution (Case II).
A clear demarcation is observed across different levels of bimodality and critical ratios. The first range that corresponds to positive values of relative difference is seen in two regions:

- For low levels of critical ratios (<0.55) and over all levels of bimodality and,
- For intermediate critical ratios but only for low levels of bimodality.

The magnitude of the values of this range is only between 0 to 1%. So, the difference is very low and could probably also be attributed to variations due to simulation. However, it can probably be said that inventory managers can be indifferent between the two cases II and I for low levels of bimodality and low levels of critical ratios. In fact, some might even argue in favor of ignoring variability in transit time distribution for the scenarios mentioned above even when it results in a slightly larger cost.

The reason behind negligible differences between the two cases is the result of low critical ratios. For lower values of critical ratios, which in effect are the ratio of underage to overage cost, the underage cost is relatively lower than that of overage cost. This implies that the model penalizes the cost function more if the excess inventory is carried than required than under stocking. So, it forces Hadley Whitin simulation model to stock less to avoid high overage cost. This probably results in a bit higher logistics cost due to a lot of backorders than the case of deterministic transit time value.
The second range covers values whose maximum of the absolute of the relative difference is less than 10%. The relative difference is negative implying that it is more expensive to ignore variability. This encompasses values that range from medium to high values of mean of the resultant distribution or equivalently, levels of bimodality and for intermediate values of critical ratios that lie in the 0.6 to 0.85 ranges. The minimum and the maximum values of the absolute value of this difference are 1% and 8% respectively.

The regions when such a range of relative difference is seen correspond to situations of increasing variability which is also denoted by increasing levels of bimodality. This also means that the coefficient of variation increases as compared to the distributions covered by the previous range (between 0 and 1%). It hence makes sense that we observe worse effects of approximating the transit time distribution to a deterministic value as compared to a normal approximation when there is a larger variability in transit time.

The third range corresponds to situations that are most affected by ignoring variability in transit time with respect to the normal approximation. Numerically, this range covers all relative differences whose absolute value is greater than 10%. In fact, while the minimum value amongst all these situations is 11%, the maximum value of this difference can go as high as 395%. The extremely high differences (considered to be any value larger than 100%) correspond to the situations of very high critical ratios (>=0.85) and very high levels of bimodality. This makes sense
because with higher variability the cost structure penalizes Case I assumption of ignoring variability to a much larger extent and hence it becomes increasingly worse (or more expensive) than Case II.

The difference in percentage corresponding to the plot in Fig.5-5 is tabulated in Table 5-2. Each cell in the table corresponds to the relative change in cost for the value of critical ratio shown on the column of the table and the level of bimodality along with the mean of the resultant bimodal distribution indicated in the rows.

**Table 5-2. Difference between logistics cost of Case I and Case II in percentage**

<table>
<thead>
<tr>
<th>Level of Bimodality</th>
<th>Mean of Residual Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.90</td>
<td>31</td>
</tr>
<tr>
<td>0.87</td>
<td>30</td>
</tr>
<tr>
<td>0.83</td>
<td>29</td>
</tr>
<tr>
<td>0.79</td>
<td>28</td>
</tr>
<tr>
<td>0.74</td>
<td>27</td>
</tr>
<tr>
<td>0.69</td>
<td>26</td>
</tr>
<tr>
<td>0.64</td>
<td>25</td>
</tr>
<tr>
<td>0.58</td>
<td>24</td>
</tr>
<tr>
<td>0.52</td>
<td>23</td>
</tr>
<tr>
<td>0.45</td>
<td>22</td>
</tr>
<tr>
<td>0.38</td>
<td>21</td>
</tr>
<tr>
<td>0.30</td>
<td>20</td>
</tr>
<tr>
<td>0.21</td>
<td>19</td>
</tr>
<tr>
<td>0.11</td>
<td>18</td>
</tr>
<tr>
<td>0.00</td>
<td>17</td>
</tr>
</tbody>
</table>

Hence a value of magnitude of 0 or 1% in the table implies that the difference in costs between ignoring variability and using a normal approximation for a transit time distribution is negligible or very small. This also means that, for the critical
ratios and levels of bimodality that correspond to such small values of relative difference, it is not worth investing in an inventory management system that is enhanced to use a normal approximation as compared to a simpler case of using a deterministic value for transit time.

The implications are opposite when the differences take very large values as shown in red cells. For the critical ratios and levels of bimodality corresponding to such large values it might be in the interest of the inventory managers to invest in enhancing the inventory management system to use a normal approximation if currently it completely ignores variability.

5.2.2 Comparing Case I & Case III

The plot in Fig.5-6 shows the relative cost difference between the two cases over different critical ratios and levels of bimodality. As in the previous comparison, relative difference is obtained by dividing the difference between the costs obtained in Case I & Case III, in which the inventory management system uses the actual distribution, by the cost obtained in the latter (i.e. Case III).

Based on the results the recommendations for the shipper are similar to that seen in the comparison from the previous section (between Case I & II). The only difference is that the magnitude of relative differences becomes worse here as compared to that in the previous section. It makes sense because using the actual distribution makes the effect of ignoring variability worse.
Fig 5-6 below plots the difference across six different ranges as mentioned in the legend alongside. The differences are all negative implying that Case I is always more expensive than Case III except when it is minimal as indicated by zero. Moreover, the entire range corresponds to minimum and maximum values of 1% and 578%.

*Fig 5-6. Difference between logistics cost of Case I (NV) and Case III (AD)*

Where \[ Abs(Diff) = \frac{NV - AD}{AD} \]

**LEGEND**

- $0\% \leq Abs(Diff) \leq 15\%$
- $15\% < Abs(Diff) \leq 30\%$
- $30\% < Abs(Diff) \leq 50\%$
- $50\% < Abs(Diff) \leq 80\%$
- $80\% < Abs(Diff) \leq 100\%$
- $Abs(Diff) > 100\%$
Table 5-3. Difference between logistics cost of Case I and Case III in percentage

<table>
<thead>
<tr>
<th>Critical Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>0.90</td>
</tr>
<tr>
<td>0.87</td>
</tr>
<tr>
<td>0.83</td>
</tr>
<tr>
<td>0.79</td>
</tr>
<tr>
<td>0.74</td>
</tr>
<tr>
<td>0.69</td>
</tr>
<tr>
<td>0.64</td>
</tr>
<tr>
<td>0.59</td>
</tr>
<tr>
<td>0.52</td>
</tr>
<tr>
<td>0.45</td>
</tr>
<tr>
<td>0.38</td>
</tr>
<tr>
<td>0.30</td>
</tr>
<tr>
<td>0.21</td>
</tr>
<tr>
<td>0.11</td>
</tr>
<tr>
<td>0.00</td>
</tr>
</tbody>
</table>

It is easy to see the demarcation in the difference between cases I and III with different levels of bimodality and critical ratios. The first range corresponds to an absolute difference of less than 15% which occur:

- For low levels of critical ratios (<0.55) and over all levels of bimodality and,
- For intermediate critical ratios (between 0.6 and 0.7) but only for low levels of bimodality.

The second and the third range, which correspond to values between 15 and 50%, can be termed as moderate range differences. These values correspond to intermediate critical ratios (0.65<=CR<=0.9). The plot indicates that these ranges
progressively transition from lower critical ratios and high levels of bimodality to higher critical ratios and all levels of bimodality.

The fourth and fifth ranges corresponding to values between 50 and 100% behave similar to the trend seen in the second and the third range of progressive increase over critical ratios and levels of bimodality. They can be referred to as high ranges. They are observed to be occurring in CR's ranging between 0.75 and 0.9. For ranges between 0.75 and 0.8 the high differences occur for high levels of bimodality. However, these difference ranges occur for lower levels of bimodality in case of higher critical ratios i.e. between 0.8 and 0.9.

The final range, which could be noted as very high range differences correspond to values that are greater than 100%. As expected they occur at regions that correspond to high critical ratios and high levels of bimodality. However, for very high CR of 0.95 they occur over all levels of bimodality. The highest difference can go up to a value such that ignoring variability can be about 6 times worse than the case when the actual distribution of transit time is used to make inventory management decisions.

In summary, it can be reiterated that it becomes more and more expensive to ignore variability in transit time with high critical ratios and high levels of bimodality. In fact small difference range here that corresponds to values that are less than 15% could also be considered high for industry that transact in millions of dollars in
ocean transportation logistics. As expected the value of differences are larger in the comparison between Case I & III as compared to that between Case I & II. It is so because case III corresponds to the actual distribution that is used in the simulation.

5.2.3 Comparing Case II & Case III

As in the previous cases, the plot below shows the relative difference between the two cases. As usual, the relative difference is obtained by dividing the difference between the costs in the two cases (II & III) by the cost obtained in the latter (i.e. Case III).

In order to understand in terms of positive and negative values, if the relative difference is negative it implies that the cost obtained by approximating the transit time distribution to normal (Case II) is larger than the value obtained by considering the actual transit time distribution.

Figure 5-7 plots the difference across three different ranges as mentioned in the legend alongside. The table in 5-4 shows the breakdown of the percentage values for different combinations of critical ratios and level of bimodality corresponding to the plot in Fig.5-7.

We observe three main regions based on values of critical ratios and levels of bimodality. A general observation from the plot shows that unlike the differences between the combinations of cases discussed above, it is not true that the magnitude
of differences increases progressively with increasing critical ratios and levels of bimodality.

**Fig 5-7. Difference between logistics cost of Case II (HW) and Case III (AD)**

**Table 5-4. Difference between logistics cost of Case II and Case III in percentage**

<table>
<thead>
<tr>
<th>Level of Bimodality</th>
<th>Mean of Resistant Distribution</th>
<th>Critical Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.5</td>
<td>0.55</td>
</tr>
<tr>
<td>0.90</td>
<td>31</td>
<td>0%</td>
</tr>
<tr>
<td>0.87</td>
<td>30</td>
<td>-1%</td>
</tr>
<tr>
<td>0.83</td>
<td>29</td>
<td>-1%</td>
</tr>
<tr>
<td>0.79</td>
<td>28</td>
<td>-1%</td>
</tr>
<tr>
<td>0.74</td>
<td>27</td>
<td>0%</td>
</tr>
<tr>
<td>0.69</td>
<td>26</td>
<td>0%</td>
</tr>
<tr>
<td>0.64</td>
<td>25</td>
<td>0%</td>
</tr>
<tr>
<td>0.58</td>
<td>24</td>
<td>-1%</td>
</tr>
<tr>
<td>0.52</td>
<td>23</td>
<td>-1%</td>
</tr>
<tr>
<td>0.45</td>
<td>22</td>
<td>0%</td>
</tr>
<tr>
<td>0.38</td>
<td>21</td>
<td>0%</td>
</tr>
<tr>
<td>0.30</td>
<td>20</td>
<td>2%</td>
</tr>
<tr>
<td>0.21</td>
<td>19</td>
<td>-1%</td>
</tr>
<tr>
<td>0.11</td>
<td>18</td>
<td>-3%</td>
</tr>
<tr>
<td>0.00</td>
<td>17</td>
<td>-1%</td>
</tr>
</tbody>
</table>
The plot in Fig 5-7 corresponds to three ranges as indicated. In the first range the values of differences go to maximum of 5%. This range can be termed as small difference range. This range of differences corresponds to different pockets of regions on the plot based on critical ratios and levels of bimodality.

- For low critical ratios and low levels of bimodality. However, it was also seen that at very low CR (<=0.5) the differences are low across all levels of bimodality.
- Over all critical ratios for levels of bimodality that are equivalent to the mean of the resultant distribution of less than 21 units corresponding to a level of bimodality of less than 0.38 units.

The second range corresponds to moderate levels of differences that are between 5 and 10%. This range again occurs along different regions. The regions in this range correspond to critical ratios and levels of bimodality that are greater than that in the range of small differences (<5%). They again occur for intermediate CR's (around 0.8) for higher levels of bimodality as compared to that in the previous range.

The final range of high differences occurs in the middle regions of the plot that correspond to the intermediate values of CR and high levels of bimodality. They are also observed for very high levels of bimodality and very high CR (=0.95).
It was observed that the differences peak in terms of magnitude at intermediate critical ratios and finally in the case of large CR of 0.95 for high levels of bimodality. This also means that the difference shows an increasing pattern up to a peak followed by decreasing pattern and finally an increasing pattern at very high CR's and high levels of bimodality.

This V-shaped trend is indicated in Fig.5-8 below. The trend can be easily seen when the difference for levels of bimodality between intermediate and high levels i.e. 21 and 31 units of mean of the resultant distribution. It seems like it is very expensive for the above regions in the areas of CR's that range between 0.6 and 0.8.

*Fig5- 8. Change of difference across increasing CR for a given level of bimodality*

At low critical ratios like 0.5 implying that overage cost is equal to underage cost, the cost due to normal approximation in HW equation and optimal should be close since optimal stocking quantity is close to average of demand over transit time. As
the critical ratio increases, HW stocks more to account for the growing underage cost and so the difference between the logistics cost increases. However, the change in critical ratio also changes the cost structure such that the difference between the underage and the overage cost changes. As critical ratio increases even further, the stocking quantity in HW simulation increases but it also penalizes very little for overstocking (and penalized a lot for under stocking). Hence the difference in costs goes down.

There is another aspect to the difference observed. The V-shaped trend is observed for higher levels of bimodality whereas the shape is flatter (or the peak is much smaller) for low levels of bimodality. This can be understood from the figure 5-9 below. Clearly, the normal approximation gets worse with higher levels of bimodality. The difference between the stocking quantities and hence the logistics cost for low levels of bimodality is smaller than that of high levels. It is because the difference in the frequencies of a given occurrence of transit time is larger between bimodal and corresponding normal approximation for high levels of bimodality as compared to the lower levels.

**Fig5- 9 Bimodal distribution and normal approximation for**

*Low Level of bimodality*  
*High Level of Bimodality*
5.2.3.1 Effect on skewed bimodal distributions

From the previous section, it is observed that the difference between the costs in the case when the inventory management system uses the actual distribution vs the corresponding normal approximation case follows a V-shaped trend across increasing critical ratios. So, this section will aim at unraveling any hidden trends pertaining to this critical ratio across different mixture rates.

The mixture rates used for the analysis range from 0.1 to 0.9 with an increment of 0.1 units at each stage. The effect of mixture rate is studied for critical ratio of 0.7 because the maximum difference between the cases was observed at this ratio. The simulation was done for one instance out of the fifteen resultant means of the mixture distributions used in the previous sections. The instance chosen was 23, which is the resultant of mixture of normal distributions with mean 29 and 17 for a mixture rate of 0.5.

At this point it might also be useful to remember that the mixture rate of \( \pi \), such that \( \pi < 1 \), implies that a particular instance of transit time lies in the first normal distribution (larger mean) with probability \( \pi \) or else lies in the second normal distribution with a probability \( (1 - \pi) \), used to create the mixture bimodal distribution.

Changing the mixture rate from 0.5 to any other value makes the bimodal distribution skewed. The skewness and kurtosis are the parameters of the
distribution that shed light on the shape. They can be understood as the spread from
the mean and height of the peak of the distribution respectively. These parameters
are used to understand the effect of mixture rate on the difference between the
logistics cost.

For a bimodal distribution setting used in this thesis, a mixture rate of 0.5 would
result in a skewness of 0 because a particular instance of transit time has an equal
probability of belonging to either of the two distributions used to create the mixture
distribution. Hence the spread of the values of transit time around the mean is same
on both sides of the mean. However, for different combinations of mixture rates and
parameters of the two normal distributions used for analysis, skewness can be
positive, negative or even zero while kurtosis could take either a positive or a
negative value.

Skewness and kurtosis of a distribution are mathematically defined as the third and
the fourth moments about the mean or

\[
\text{Skewness} = E\left[\left(\frac{X - \mu}{\sigma}\right)^3\right]
\]

\[
\text{Kurtosis} = E\left[\left(\frac{X - \mu}{\sigma}\right)^4\right]
\]

For a mixture of two normal distributions that are used to create a bimodal
distribution, the skewness and kurtosis are derived to be:

\[
\text{Skewness} = \frac{1}{\sigma^3}\left[\pi(\mu_1 - \mu)(3\sigma_1^2 + (\mu_1 - \mu)^2) + (1 - \pi)(\mu_2 - \mu)(3\sigma_2^2 + (\mu_2 - \mu)^2)\right]
\]
Kurtosis =
\[
\frac{1}{\sigma^4} \left[ \pi \left( 3\sigma_1^4 + 6(\mu_1 - \mu)^2\sigma_1^2 + (\mu_1 - \mu)^4 \right) + (1 - \pi) \left( 3\sigma_2^4 + 6(\mu_2 - \mu)^2\sigma_2^2 + (\mu_2 - \mu)^4 \right) \right]
\]

Where

\[ \pi = \text{mixture rate}, \text{and} \]

\[ \mu & \sigma \text{ are the mean and standard deviation respectively of the resultant bimodal mixture} \]

which are derived in section 4.2.1

The results of the simulation are tabulated below in table 5-5. The columns of the table indicate the mixture rate and the corresponding parameters of the resultant bimodal mixture namely the mean, standard deviation, skewness and kurtosis. The last column indicates the relative difference between Cases II &III in percentage, for the given mixture rate.

<table>
<thead>
<tr>
<th>Mixture rate</th>
<th>Mean</th>
<th>Std Dev</th>
<th>CV</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>% Difference between cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>18.2</td>
<td>4.12</td>
<td>0.226</td>
<td>1.78</td>
<td>5.99</td>
<td>-5.4%</td>
</tr>
<tr>
<td>0.2</td>
<td>19.4</td>
<td>5.20</td>
<td>0.268</td>
<td>1.18</td>
<td>3.18</td>
<td>-7.0%</td>
</tr>
<tr>
<td>0.3</td>
<td>20.6</td>
<td>5.85</td>
<td>0.284</td>
<td>0.72</td>
<td>2.03</td>
<td>0.0%</td>
</tr>
<tr>
<td>0.4</td>
<td>21.8</td>
<td>6.21</td>
<td>0.285</td>
<td>0.35</td>
<td>1.53</td>
<td>-6.8%</td>
</tr>
<tr>
<td>0.5</td>
<td>23.0</td>
<td>6.33</td>
<td>0.275</td>
<td>0.00</td>
<td>1.38</td>
<td>-11.3%</td>
</tr>
<tr>
<td>0.6</td>
<td>24.2</td>
<td>6.21</td>
<td>0.257</td>
<td>-0.35</td>
<td>1.53</td>
<td>-9.9%</td>
</tr>
</tbody>
</table>
The table indicates that mixture rates from 0.1 to 0.4 have the same absolute value of skewness and kurtosis as the ones with rates from 0.6-0.9, however with opposite signs for the former value. This is understood graphically from the figure shown below in Fig 5-10. For instance, the third and the seventh graphs corresponding to mixture rates of 0.3 and 0.7 indicate the smaller bumps of the same frequency levels on the opposite sides of the mean of the resultant distribution. This therefore indicates same absolute value but opposite signs of skewness.

<table>
<thead>
<tr>
<th>Mixture Rate</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25.4</td>
<td>26.6</td>
<td>27.8</td>
</tr>
<tr>
<td></td>
<td>5.85</td>
<td>5.20</td>
<td>4.12</td>
</tr>
<tr>
<td></td>
<td>0.230</td>
<td>0.195</td>
<td>0.148</td>
</tr>
<tr>
<td></td>
<td>-0.72</td>
<td>-1.18</td>
<td>-1.78</td>
</tr>
<tr>
<td></td>
<td>2.03</td>
<td>3.18</td>
<td>5.99</td>
</tr>
<tr>
<td></td>
<td>-6.3%</td>
<td>-4.0%</td>
<td>-1.9%</td>
</tr>
</tbody>
</table>

*Fig 5-10. Resultant distribution with different mixture rates starting with 0.1 on top left to 0.9 on the bottom right with an increment of 0.1 units in the rate with each figure*
It is observed that the relative percentage difference between the costs for the two cases shows an increasing and then a decreasing pattern with increasing mixture rates. A major takeaway from the values in the table indicate that the difference in the value of logistics cost is probably a function of all the four parameters of the actual bimodal distribution: mean, standard deviation, skewness and kurtosis.

To understand the implication let us consider the following observations. The cost difference between case II and III show a larger magnitude for the mixture rate of 0.1 as compared to that of 0.9. A mixture rate of 0.9 indicates that the probability that a particular transit time value lies in the distribution with the larger mean is 9 times (and hence a negative skew) as that of lying in the one with a smaller mean. Hence initial guess would suggest that the difference between the costs should be
larger in the case when the mixture rate is 0.9 as compared to that of a rate of 0.1. However, it turns out that the intuition is not right. It might be because of the fact that a mixture rate of 0.1 results in a larger coefficient of variation of the resultant distribution as compared to that of a rate of 0.9. Such a pattern is also seen for mixture rate pairs of 0.2 & 0.8.

It was also observed that there is very negligible difference for a mixture rate of 0.3 while a much larger magnitude of difference for a rate of 0.7. A similar pattern was observed for the pairs corresponding to mixture rates of 0.4 & 0.6. Even though the former has a larger CV, they differ with respect to the sign of the skew. As compared to the previous pair of 0.9 & 0.1, the smaller bumps for the pairs of 0.3 & 0.7 are much larger in comparison to the former pair. Alternatively, the parameters that explain the shape of the distribution (specifically the skew) probably dominate the results for the percentage difference in cost in this case as compared to the difference in the CV which could explain differences in costs in the previous example case.

Such a pattern suggests that the difference in the costs is explained by collectively examining all the three parameters associated with the resultant mixture distribution namely the Coefficient of Variation (or mean and standard deviation, Skewness and Kurtosis.)
5.3 Summary of Results

As expected, it was observed that enhancing the inventory management system to use the actual transit time distribution for inventory management decisions would result in calculation of cheaper options as compared to the other cases. However, the magnitude of savings achieved by adding this capability to the current inventory system can differ based on the desired service levels (or critical ratios) and level of bimodality in the distribution.

On comparing the case of approximating the distribution to normal and the one where variability is ignored, there exist a few regions (low service levels and low levels of bimodality) such that the latter performs at least as good as normal approximation.

A V-shaped curve was observed for the plot between the differences of the cost in assuming normality and using the actual distribution. This suggests that the maximum difference between the costs occurs at an intermediate critical ratio of 0.75 units.

Finally, the analysis of effect of shape parameters suggests that logistics cost can be explained as a cumulative effect of skewness and kurtosis values of the distribution along with its coefficient of variation. However, further investigation is needed to develop theory behind this statement.
Chapter 6. Conclusion

In this thesis, we investigated the effect of bimodally distributed transit time on logistics cost. Multimodality was observed in ocean transit time distribution in the dataset available with us, which includes data from retailers, freight forwarders and manufacturing firms. Cumulatively, the percentage of lanes that were not unimodal was 17% of the total lanes but they constitute large shipment volumes summing up to about 80%.

Often, the industry chooses to ignore variability in transit time many a times due to limited built in capabilities of the inventory management systems. Also, a common practice in literature involves modeling variability in transit time as a normally distributed random variable when in reality it is not. Considering the above two findings, the analysis to understand the effect of bimodality in transit time was done by comparing three ways of approaching variability in transit time. Case I is the scenario of completely ignoring inherent variability; Case II is the scenario where distribution of transit time is incorrectly assumed to be normal and Case III is when the inventory management system uses the actual distribution.

The analysis was done by simulating a number of instances of bimodality, which was then used to quantify the differences between the three cases by using Monte Carlo sampling. Additionally, in order to examine the effect on logistics cost, service level targets (or critical ratios) and mixture rates were varied.
It was observed that it makes sense to use inventory management systems, which do not have the added capabilities to account for variability, for low service levels (critical ratio < 0.55) under all levels of bimodality. This is also true for intermediate service levels (<=0.7) but only for low levels of bimodality.

However, the shippers should consider updating their inventory management systems in situations of high customer service levels desired and high levels of bimodality in transit time distribution. As expected, the system that uses the actual distribution to make inventory management decisions give the best results in terms of cost.

The case when the distribution is approximated to normal vs. the one that uses actual distribution differ the most in terms of cost at a critical ratio of 0.7, for high levels of bimodality. The shape of the relative difference between the costs hence takes a V-shaped trend.

Logistics cost also changes with respect to different mixture rates used to create the bimodal distributions. It was observed that cost is a cumulative effect of four parameters of the transit time distribution- mean, standard deviation, skewness and kurtosis values.

Finally with respect to safety stock, there is not an easy fix to the problem of incorrect assumption of normality in the transit time distribution when in reality it
is bimodal. There are clear regions of overestimation and underestimation of the amount of safety stock that should be carried.

6.1 Further Research

We see the following potential extensions of this thesis research. First it would be interesting to analyze the theoretical bounds of the relative difference in logistics costs for the three cases discussed as a function of parameters of the transit time distribution and the service levels targeted.

Secondly, it might also be helpful to work on an algorithmic model to facilitate the process of incorporating variability into currently available inventory management systems.

Next, it would be worth investigating the occurrence of cases of higher levels of multimodality (besides bimodality) and their effects on logistics cost. This thesis was only able to identify non-unimodality in transit time distributions. Bimodality was used because it is the simplest case of multimodality.

Further, research could also be conducted to understand if effects of bimodal distribution on logistics cost imitate any other unimodal distributions that might have an optimum inventory policy associated with it in available literature.
Finally, frequency and impact of bimodally distributed transit time lanes warrant a thorough investigation into the reasons that lead to such a phenomenon of large differences in logistics cost in the three cases. Hence it might also be helpful to conduct a sensitivity analysis in order to formally identify the factors that cause bimodality in transit time distributions. This would enable the shippers to predict if the lane is going to be bimodal given a certain set of trade lane characteristics.
Appendix

(I) Detailed derivation for mean and variance of demand over transit time

The mean and the variance of demand over transit time is given by:

\[ E[X_{D_{oLT}}] = E[L_T]E[D] \]

\[ \sigma^2_{D_{oLT}} = E[L_T] \sigma^2_D + (E[D])^2 \sigma^2_{L_T} \]

Derivation:

We know that,

\[ D_{oLT} = \sum_{i=1}^{L_T} d_1 + d_2 + d_3 \ldots \ldots + d_{LT} \]

An observation of demand has two components – stochastic and a deterministic. Hence each occurrence \( d_i \) is given by:

\[ d_i = E[d_i] + \tilde{d}_i \]

Such that the stochastic component \( \tilde{d}_i \) has a mean of 0 and a variance of \( \sigma^2_{\tilde{d}} \). Also, \( E[d_i] = E[D] \). Therefore,

\[ E[X_{D_{oLT}}] = E[d_1 + d_2 + d_3 \ldots \ldots + d_{LT}] \]

\[ = E[(E[d_1] + \tilde{d}_1) + (E[d_2] + \tilde{d}_2) + \ldots + (E[d_{LT}] + \tilde{d}_{LT})] \]

\[ = E[(E[d_1] + E[d_2] + E[d_3] + \ldots + E[d_{LT}]) + E[(\tilde{d}_1 + \tilde{d}_2 + \tilde{d}_3 + \ldots + \tilde{d}_{LT})]] \]

\[ = E[L_T]E[D] + 0 \]

\[ E[X_{D_{oLT}}] = E[L_T]E[D] \]

From the above derivation it is shown that

\[ X_{D_{oLT}} = (E[d_1] + E[d_2] + E[d_3] + \ldots + E[d_{LT}]) + (\tilde{d}_1 + \tilde{d}_2 + \tilde{d}_3 + \ldots + \tilde{d}_{LT}) \]
$$= E[LT \cdot E[D]] + (\tilde{d}_1 + \tilde{d}_2 + \tilde{d}_2 \ldots + \tilde{d}_{LT})$$
$$= (LT \cdot E[D]) + (\tilde{d}_1 + \tilde{d}_2 + \tilde{d}_2 \ldots + \tilde{d}_{LT})$$

Since both the terms in the bracket and LT and demand are independent

$$\sigma^2_{D_{oLT}} = (E[D])^2 \cdot \sigma^2_{LT} + Var((\tilde{d}_1 + \tilde{d}_2 + \tilde{d}_2 \ldots + \tilde{d}_{LT}))$$

Let \( \theta = \tilde{d}_1 + \tilde{d}_2 + \tilde{d}_2 \ldots + \tilde{d}_{LT} \). Therefore,

$$\sigma^2_{\theta} = E[\theta^2] - (E[\theta])^2$$

But \( E[\theta] = E[\tilde{d}_1 + \tilde{d}_2 + \tilde{d}_2 \ldots + \tilde{d}_{LT}] = 0 \) from definition

Therefore, \( \sigma^2_{\theta} = E[\theta^2] = E[(\tilde{d}_1 + \tilde{d}_2 + \tilde{d}_2 \ldots + \tilde{d}_{LT})^2] \)

$$= E[\tilde{d}_1^2 + \tilde{d}_2^2 + \ldots \tilde{d}_{LT}^2 + 2 \Sigma_{i=1}^{LT} \Sigma_{j=i+1}^{LT} \tilde{d}_i \tilde{d}_j]$$

Because demands between time periods are independent and uncorrelated expectation of the second term in the summation goes to 0 because:

$$E[\tilde{d}_i \tilde{d}_j] = E[\tilde{d}_i]E[\tilde{d}_j] = 0$$

Therefore,

$$E[\theta^2] = E[\tilde{d}_1^2 + \tilde{d}_2^2 + \ldots \tilde{d}_{LT}^2]$$

$$= E[LT]E[\tilde{d}_1^2]$$

But \( \sigma^2_{\tilde{d}_i} = \sigma^2_D = E[\tilde{d}_i^2] - E[\tilde{d}_i]^2 = E[\tilde{d}_1^2] - 0. \)

This implies that \( E[\theta^2] = E[LT]E[\tilde{d}_1^2] = E[LT] \sigma^2_D \)

Combining all the above equations yields:

$$\sigma^2_{D_{oLT}} = E[LT] \sigma^2_D + (E[D])^2 \sigma^2_{LT}$$
(II) Detailed derivation for total cost as proposed by Hadley-Whitin (1963)

\[ TC = vD + \frac{AD}{Q} + hv \left[ \frac{Q}{2} + R - \mu \right] + \frac{bD}{Q} \left[ \int_{R}^{\infty} xg(DoLT) - R \ast G(DoLT) \right] \]

Purchase Cost = per unit cost *number of units ordered on average

\[ = v \ast D \]

Safety stock(s) is the expected net inventory at the time of order delivery, which is carried at the retailer site. Hence the expected inventory right after the arrival of an order is equal to \( Q+s \). Therefore for a given cycle the inventory goes from \( Q+s \) at the start to \( s \) at the end. Thus the average inventory on hand is \( \frac{1}{2} (Q + s + s) = \frac{Q}{2} + s \). Also by definition \( s = \int_{0}^{\infty} (R-x)g(x)dx = R - \mu \) because \( \int_{0}^{\infty} xg(x) = \mu \). Where \( x = demand over transit time \)

Holding cost = annual holding cost *average on-hand inventory

\[ = (hv) \ast \left[ \frac{Q}{2} + R - \mu \right] \]

We will have back orders if the demand over transit time is observed to be more than the quantity ordered. Therefore, the expected number of back orders per cycle is given by \( = \int_{R}^{\infty} (x-R)g(x)dx = \int_{R}^{\infty} xg(x) - R \ast G(x) \). The average number of backorders incurred per year is the expected number of backorders per cycle multiplied by the number of cycles occurring annually. Hence the number of back orders annually is \( \frac{D}{Q} \left[ \int_{R}^{\infty} xg(x) - R \ast G(x) \right] \).

Back ordering cost = back ordering cost per unit*number of backorders

\[ = \frac{bD}{Q} \ast \left[ \int_{R}^{\infty} xg(x) - R \ast G(x) \right] \]
Bibliography


