

Risk Management in Biopharmaceutical Supply Chains

by

Yao Ma

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Engineering - Industrial Engineering and Operations Research

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:
Robert Leachman, Chair
Professor George Shanthikumar
Professor Zuo-Jun Max Shen
Professor Sourav Chatterjee

Spring 2011

Risk Management in Biopharmaceutical Supply Chains

Copyright 2011

by

Yao Ma

Abstract

Risk Management in Biopharmaceutical Supply Chains

by

Yao Ma

Doctor of Philosophy in Engineering - Industrial Engineering and Operations Research

University of California, Berkeley

Robert Leachman, Chair

Biopharmaceutical supply chains present considerable complexity issue for the formulation of optimal plans due to significant uncertainty in the supply chain. The primary goal of biopharmaceutical supply chain planning is to provide reliable supply to patients while coping with various supply chain risks. In chapter 1 first I discuss the key elements and basic characteristics of the biopharmaceutical supply chain . Then I present the major challenges in biopharmaceutical supply chain planning and divide them into two main categories: deterministic complexity problems and stochastic uncertainty problems. In the end of chapter 1 I briefly discuss the most recent work in solving the deterministic complexity problems.

The planning of biopharmaceutical supply chain operations faces risks from various sources. These include customer demand fluctuations, regulatory requirement changes, long quality assurance cycle time, etc. In chapter 2, I review the major risks in biopharmaceutical supply chain and current practice to hedge against these risks. The impact of these risks is evaluated in terms of a cumulative supply and demand perspective. Furthermore I analyze two main risk mitigation tools in supply chain risk management: safety stock and safety time. Then I use simulation to show that safety stock is a preferred approach for risk mitigation in biopharmaceutical supply chains.

In chapter 3, first I focus on stochastic lead time risk and discuss conventional as well as crossover based approaches for safety stock planning. Also I demonstrate the benefit of a proposed approach to safety stock planning with numerical examples and simulation. Then the proposed model is extended to consider batch rejection risk and excursion risk. Batch rejection risk represents the possibility that a batch fails to meet regulatory requirements. The excursion risk reflects potential major disruptions in biopharmaceutical supply chain. Examples of such events include contamination of the production facility, earthquake, etc. These three risks cover most of the major uncertainties in biopharmaceutical supply chain. The model determines the necessary safety stock level to prevent stock outs given these risks as a function of the target service level.

In chapter 4, I discuss the implementation of the proposed model in a multi-echelon biopharmaceutical supply chain. Also I use sensitivity analysis to evaluate the impact of improving key supply chain parameters. Then the model is applied to determine the stock level supplying a regional market of an actual biopharmaceutical supply chain and significant potential savings are demonstrated. In the end I identify a few important potential research problems in biopharmaceutical supply chain management.

To my late father

To my mother

& To my wife

Contents

List of Figures	v
List of Tables	vii
1 Biopharmaceutical Supply Chain Overview	1
1.1 Biopharmaceutical Supply Chain Elements	1
1.2 Biopharmaceutical Supply Chain Characteristics	2
1.3 Challenges in Biopharmaceutical Supply Chain	3
1.4 Optimal Planning in Bioproduction	4
2 Biopharmaceutical Supply Chain Risk Mitigation	8
2.1 Review of Recent Work on Supply Chain Risk Management	8
2.1.1 Supply Chain Risk Management Strategies	8
2.1.2 Demand Risk and Supply Risk	9
2.1.3 Production Planning under Uncertainty	9
2.1.4 Integrate Optimization and Estimation	10
2.1.5 Operational Statistics	10
2.1.6 Robust Linear Programming	11
2.1.7 Hedging Batch Rejection Risk in a Multi Stage Supply Chain Network	12
2.1.8 Stochastic Lead Time Literature Review	13
2.2 Biopharmaceutical Supply Chain Risks	15
2.2.1 Demand Risk	15
2.2.2 Regulatory Uncertainty	15

2.2.3	Process Risk Overview	16
2.2.4	QA Cycle Time Risk	17
2.2.5	Production Process Change Risk – Conditional Release	17
2.3	Current Practice to Hedge against Risk	19
2.4	Safety Stock vs. Safety Time for Biopharmaceutical Supply Chain Risk Mitigation	20
2.4.1	Cumulative Curve perspective of the Biopharmaceutical Supply Chain Risk	20
2.4.2	Safety Stock versus Safety Time in Mitigating Supply Chain Risk	20
2.4.3	Risk Analysis via Simulation - QA Cycle Time’s Impact on Planned Backorders	21
3	Stock Planning in Biopharmaceutical Supply Chain	25
3.1	Safety Stock Level Planning under Stochastic Lead Time	25
3.1.1	Safety Stock Level Planning Based on the Conventional Approach	26
3.1.2	Order Crossover Analysis	27
3.1.3	Safety Stock Level Based on Crossover Analysis	28
3.1.4	Incorporating the Ordering Interval and the Planned Lead Time	34
3.1.5	Safety Stock Savings	38
3.1.6	Numerical Example and Simulation	41
3.1.7	Approximation of the Number of Outstanding Orders	44
3.2	Safety Stock Level Planning under Excursion Risk and Stochastic Lead Time	46
3.2.1	Single Excursion	48
3.2.2	Multiple Excursion Types	54
3.2.3	Numerical Example	56
3.2.4	Simulation	57
3.3	Safety Stock under Stochastic Lead Time and Batch Rejection Risk	60
3.3.1	Model	60
3.3.2	Numerical Example	66
3.3.3	Simulation	67

3.4	Safety Stock under Stochastic Lead Time, Batch Rejection Risk and Excursion Risk	69
3.4.1	Model	73
3.4.2	Numerical Example	76
3.4.3	Simulation	77
4	Safety Stock Planning Implementation in Practice	82
4.1	Supply and Demand Risk in Multiechelon Inventory System	82
4.2	Multiechelon Safety Stock Planning	84
4.2.1	Inventory Adjustment to the Proposed Level	85
4.2.2	Inventory Adjustment to Current FG Level	87
4.3	Sensitivity Analysis	89
4.3.1	Improvement of Cycle Time	90
4.3.2	Improvement of Forecasting Error	91
4.4	Regional Market Case Study	93
4.5	Conclusion and Important Topics in Biopharmaceutical Supply Chain	96
4.5.1	Conclusion	96
4.5.2	Infrastructure and Technology	96
4.5.3	Conditional Release	98
4.5.4	Stock Replenishment Planning - Rotation Cycle	98
	Bibliography	101
A	Calculation of Required Stock Level under Geometric Lead Time and Geometric Excursion Occurrence	105

List of Figures

1.1	Biopharmaceutical Production Line	1
1.2	Product Structure of Planning Engine Formulation	5
2.1	QA Test Flow Chart	18
2.2	Cumulative Curve Perspective of Supply Chain Risk	20
2.3	Mitigate Risk with Safety Stock or Safety Time	21
2.4	Planned Cumulative Supply vs. Simulated Average Cumulative Supply - 50th Percentile	23
3.1	Cumulative Supply, Demand and Gap	29
3.2	Ordering Interval and Average # of Outstanding Orders	37
3.3	Ordering Interval and Required Stock Level	39
3.4	Stock Distribution in the Planning Horizon	42
3.5	Distribution of # of Outstanding Orders	43
3.6	Distribution of Shortage Quantity	44
3.7	Average Inventory Holding and Service Level under Stochastic Lead Time Risk	45
3.8	Binomial Approximation of # of Outstanding Orders	46
3.9	Binomial Approximation of Shortage Quantity	47
3.10	Excursion Model	48
3.11	Stock Distribution in the Planning Horizon	59
3.12	Distribution of # of Outstanding Orders	61
3.13	Distribution of Shortage Quantity	62
3.14	Average Inventory Holding and Service Level under Stochastic Lead Time and Excursion Risk	63

3.15	Stock Distribution in the Planning Horizon	70
3.16	Distribution of # of Outstanding Orders	71
3.17	Distribution of Shortage Quantity	72
3.18	Average Inventory Holding and Service Level under Stochastic Lead Time and Batch Reject Risk	73
3.19	Stock Distribution in the Planning Horizon	80
3.20	Distribution of # of Outstanding Orders	80
3.21	Distribution of Shortage Quantity	81
3.22	Average Inventory Holding and Service Level under Stochastic Lead Time, Batch Reject and Excursion Risk	81
4.1	Multi-echelon Inventory System	85
4.2	Multi-product Inventory System	86
4.3	Cycle Time and Marginal Required Stock Level	91
4.4	Cycle Time and Marginal Dollar Inventory Savings	92
4.5	Forecasting Error and Marginal Required Stock Level	93
4.6	Forecasting Error and Marginal Dollar Inventory Savings	94

List of Tables

2.1	Simulation Settings - QA Cycle Time's Impact on Backorders	22
2.2	Planned Cumulative Supply vs. Simulated Average Cumulative Supply . . .	23
2.3	Average Stock under Different Safety Stock and Safety Time Combination .	24
3.1	Simulation Parameters for Stochastic Lead Time Risk	41
3.2	Simulation Results under Stochastic Lead Time Risk	43
3.3	Simulation Parameters under Stochastic Lead Time and Excursion Risk . . .	57
3.4	Simulation Results under Stochastic Lead Time Risk	60
3.5	Simulation Parameters under Stochastic Lead Time and Batch Reject Risk .	68
3.6	Simulation Results under Stochastic Lead Time and Batch Reject Risk . . .	69
3.7	Simulation Parameters under Stochastic Lead Time, Batch Reject and Excur- sion Risk	78
3.8	Simulation Results under Stochastic Lead Time, Batch Reject and Excursion Risk	79

Acknowledgments

I greatly appreciated everyone who helped me in so many ways during the past five years here at Berkeley. Without any of you, I wouldn't have accomplished what I did.

First and foremost, I would like to express my deepest thankfulness to my advisor Professor Robert Leachman, for supporting me over the years, and for opening an exciting research field for me to explore, to absorb, and to devote. He is always there whenever I am in need, giving me many valuable suggestions on research and other things. His tremendous guidance and feedback leads me to successfully finish my research project.

I'm indebted to Professor Max Shen and George Shanthikumar for their valuable inputs regarding my research and also their constant help over the years. I'm very thankful to my other committee member, Professor Sourav Chatterjee for his support over the years.

I also would like to thank the research support from Bayer HealthCare and all my colleagues at Bayer for their valuable input and support. These include Thomas Panzer, Prasad Saraph, Chris Bogart, Todd Roman, Robert Light, Joerg Dohmeyer, Jeff Bell, Steven Sweetwood and many others.

Finally, I am deeply grateful to my parents, my brother and my wife. A life journey with my wife, Shan Li, is the best thing that ever happened to me. Her love, encouragement and faith in me made my years at Berkeley a wonderful journey.

Chapter 1

Biopharmaceutical Supply Chain Overview

1.1 Biopharmaceutical Supply Chain Elements

This section describes some key elements in a biopharmaceutical supply chain. These major production steps are common in the biotech industry and analysis based on these steps can be applied to most biopharmaceutical production lines.

A typical biopharmaceutical supply chain starts from one of the two kinds of base cell lines: Chinese Hamster Ovary (CHO) cells, or Escherichia coli (E.coli.) cells (Johnston 2009). These cells serve as the growth medium in the fermentation described below.

The production process consists of four major production stages: Expression/Fermentation, Purification, Bulk, Fill, Freeze Dry (FFD) and Packaging. These steps are outlined in Figure 1.1.

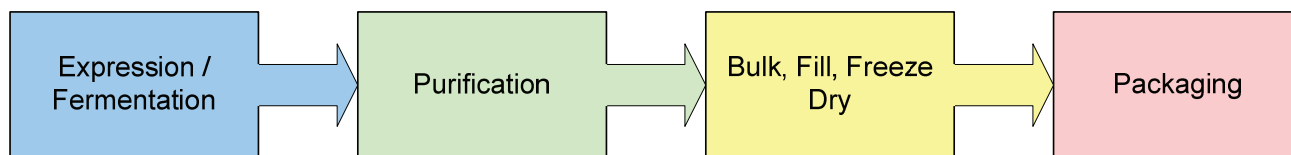


Figure 1.1: Biopharmaceutical Production Line

In **Expression/Fermentation**, a small amount of seed protein is placed into the growth medium and the protein of interest is expressed in large-batch “campaigns” or continuously. The fermentation process is tightly controlled to avoid contamination. The output of this step is called UFTCF.

In **Purification**, the potential contaminants are removed and the protein of interest is recovered. Another goal of this step is to increase material concentration to meet final product requirements. The purification process is highly complex and involves the most uncertainty in the biopharmaceutical production process. For examples, purification may target a certain potency level but the batch potency from production usually varies over a range. Batches outside the regulatory approved range are either discarded or downgraded to lower level product variety. The intermediate product this step outputs is called UFDF.

In **Bulk, Fill, Freeze Dry (“FFD”)**, final product dosage is determined and stabilized to be ready for storage and shipment. Most biopharmaceutical products are required to be stored at 2-8 degrees centigrade. For less stable product, this rule has to be strictly abided to maximize room temperature flexibility for patients. Thus shippers are more restricted and product packaging process is further complicated to accommodate the cold chain shipping. The output of this step is called final container (FC).

In **Packaging**, batches from the previous production step are packaged according to region specific configurations. There are tight regulatory requirements on the elements of packaging kits, including administrative elements, water for injection, direction sheets, etc. This step finishes all the production processes and produces the final goods, (FG).

Due to the complexity of the biopharmaceutical production process, the manufacturing site requires significant capital investment. To maximize the return of investment cost, usually one biological product is manufactured in a single site and then shipped to various regions of the world. As a result of regulatory discrepancies among different regions, sometimes not all products are packaged onsite. Once the final product is packaged, it will be shipped to regional distribution center for final dispatch to consumers.

1.2 Biopharmaceutical Supply Chain Characteristics

Biopharmaceutical production is under stringent regulation from government agencies. In contrast to traditional pharmaceutical products, biopharmaceutical products consist of large molecules instead of chemical compounds. Usually the properties of the biological cannot be fully understood due to technological limitations. As a result, for biopharmaceuticals, not only the final product is subject to quality testing but all intermediate products also need to pass regulatory tests as well.

To avoid cross-contamination, the production is by batch and typically only one batch is present in one facility at one time. At each production stage, there can be multiple alternative production facilities but each facility is subject to regulatory approval.

In order to increase material concentration, it is typical in bioproduction that batches from previous stages are blended to meet product intensity requirement. But this introduces potential risk of mixing good batches with contaminated ones and thus will potentially increase product rejection.

Immediately following each production step is a quality assurance (QA) series of tests. QA ensures all intermediate and final products meet the regulatory requirements. Typical tests include microbiology tests, electrophoresis tests, sterility tests, etc. QA cycle time for different types of intermediate products varies. On average UFDF QA cycle time is the longest.

For a more comprehensive discussion of biotech industry background and the characteristics of biopharmaceutical supply chain, please refer to Johnston (2009).

1.3 Challenges in Biopharmaceutical Supply Chain

Biopharmaceutical supply chains face a unique set of challenges, including demand fluctuations, inventory shelf life expiration, planning complexity, manufacturing uncertainty, etc.

As a life saving industry, the primary goal of biopharmaceutical supply chain planning is to attain a 100% service level. Understanding customer demand as well as its variability plays a critical role. Even for life critical drugs, demand fluctuation can be substantial as a result of industry competition, patient accident/emergency and new markets exploration.

Biopharmaceutical production utilizes relatively recent technology and so far the whole process is not as well understood as traditional pharmaceutical production. Biopharmaceutical production is subject to stringent levels of regulation from authorities. Not only is the drug licensed by authorities, but the whole production process is regulated to ensure an extremely high level of cleanliness. This means any change to the equipment, process, and facility will require new certificates from authorities. Usually the certification process requires months or even longer time, so the planning has to be robust enough to cope with such potential risks. Moreover, regulatory requirements as well as approval timelines vary by country and market, which makes the planning problem further complicated.

Process uncertainty is one of the main differences between traditional pharmaceutical production and bioproduction. Because of this, quality assurance follows immediately after major production steps and the variation of quality assurance cycle time is substantial. For example, the quality assurance time of purification step may vary lot to lot from less than 10 days to as many as 100 days. Production variation is another source of uncertainty. A production step may produce output in a range with some distribution. And some of the outputs may fail to meet regulatory requirements and thus not be qualified for sale.

We can divide the challenges in biopharmaceutical supply chain into planning complexity related issues and planning reliability issues due to uncertainty in the supply chain. The planning complexity is largely due to the complex manufacturing process and stringent regulatory requirements. This complexity also combines with product variety in different markets and various regulatory requirements from authorities throughout the world. The reliability issue is the result of relative nascent technology of biotech manufacturing and a lack of accurate understanding of the process. Thus a lot of variables in biopharmaceutical supply chain are stochastic and subject to substantial variability. In this paper I will focus on tackling

the reliability issue in biopharmaceutical supply chains. But before that I'll review most recent development in optimal planning in bioproduction and build our biopharmaceutical risk management model based on results from optimal planning.

1.4 Optimal Planning in Bioproduction

To address some of the challenges above, Bayer Healthcare provided a research grant to Prof. Leachman of the University of California at Berkeley, working with Prof. Shen and PhD students Rick Johnston and Shan Li. The result of the research project was the biotech Planning Engine, the first optimal planning model in biotech industry. The goal of the planning engine is to build an optimal production planning schedule to meet demand on time while minimizing inventory cost. The planning engine problem statement is quoted as follows:

Produce sufficient but not excessive finished goods to meet final demand while meeting all region-specific regulatory requirements on product, process, equipment, facilities and raw materials. (Leachman, Johnston, Shen, Li 2007)

The core of the planning engine is an event-based time grid formulation (Leachman, 1993, Dessouky & Leachman, 1997). The possible time epochs for batch starts are restricted to integer multiples of the production processing time of each production stage. This time grid model dramatically reduces formulation size and make the optimal planning of such complex biopharmaceutical production computationally feasible.

The formulation is based on a product structure including the following components: inventory type, wire and process route. Inventory type is a classification of input material for the next processing stage. Inventory types are distinguished only to the extent of where the resulting product may be sold. Wire is a valid combination of inventory types blended into a batch input to a processing stage resulting in a specific output inventory type of that stage. Wire enumeration lists every valid mix of input material. All possible facilities and work-in-process types are considered. Associate with each wire is a list of possible facility and processing options. Each option for the wire is called process route. The product structure is illustrated in Figure 1.2.

The primary objective of the formulation is to meet customer demand as much as is feasible by minimizing the backorder cost. The second objective minimizes the discounted number of batches started. The second objective is equivalent to minimize holding cost but the discounted batch start approach is employed in the formulation to reduce problem size.

For the formulation constraints, a set of constraints is enforced for each stage. In the **purification** stage, the first groups of constraints are inventory balance constraints which enforce cumulative supply of UFTCF should be no less than cumulative demand of UFTCF at the purification stage. The second groups of constraints require a fixed amount of material to be allocated to each purification batch start. The third groups of constraints are capacity

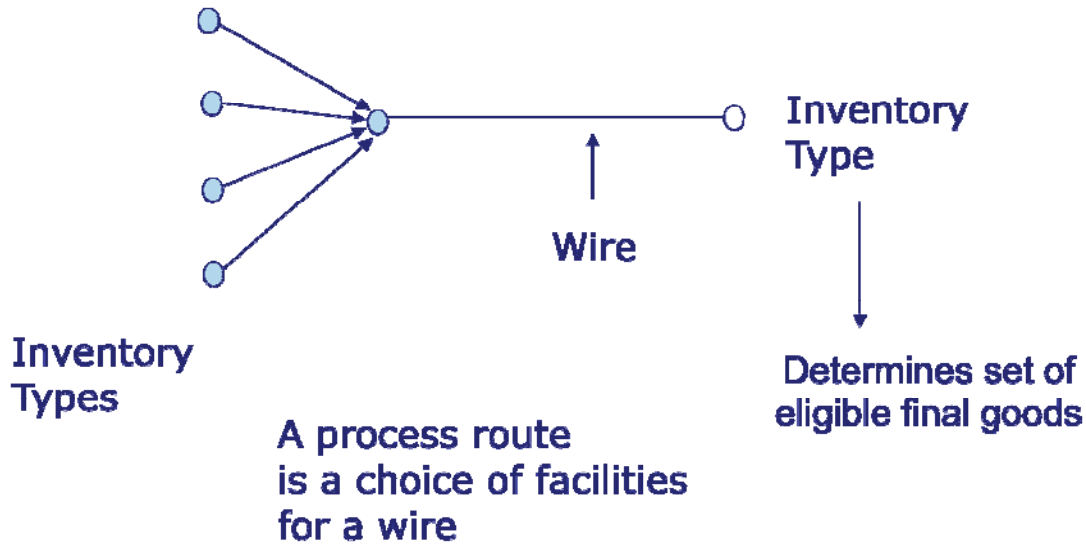


Figure 1.2: Product Structure of Planning Engine Formulation

constraints which enforce at most one batch processed in a facility at each epoch for a batch start. The last group of constraints ensures that the net inventory should be no less than the safety stock over a set of time epochs. To enforce safety stock throughout the planning horizon, it is sufficient to enforce safety stock constraints whenever the cumulative demand or supply changes. But this will increase the complexity of the planning problem. A practical approximation can be set safety stock constraints at a set of time epochs, such as weekly or biweekly.

In the **FFD** stage, similar to purification step, there are inventory balance constraints, safety stock constraints and capacity constraints. In addition to that, a group of constraints is enforced to reflect freeze-drying machine characteristics, such as the lower and upper bound on the amount of material that can enter production, etc. **Packaging** stage constraints also include inventory balance constraints, safety stock constraints, capacity constraints and packaging production characteristic constraints. For **demand** constraints, cumulative demand must be less than sum of inventory, work in process, new production and backorders. Reproduced from Leachman (2008), following are the key decision variables defined in the formulation.

- $X_r(t)$: binary variable indicating if we schedule a batch on process route r at time t or not
- $Z_r(t)$: batch start quantity
- $Y_{mr}(t)$: quantity of inventory type m allocated to form batch quantity $Z_r(t)$
- $b_m(t)$: backorder of product m at time t (defined for t in the demand time grid)

- $I_m(t)$: initial quantity of inventory type m at time t
- $W_m(t)$: cumulative supply of inventory type m projected to be released for follow-on use at time t from current work-in-process (WIP).

We use FFD stage as an example to illustrate the type of constraints in our formulation. First set of constraints are for inventory balance constraints to make sure that cumulative supply is more the cumulative demand.

$$\sum_{\tau \leq t} \sum_{s \in S_{\bar{n}}^2(\tau)} Y_{\bar{n}s}^2(\tau) \leq I_{\bar{n}}^2 + W_{\bar{n}}^2(t) + \sum_{r \in \bar{n}} \sum_{\tau + c_r \leq t} \theta \alpha_r (1 - \beta_r) \delta_r X_r^1(\tau)$$

$$\forall t \in \bigcup_{r \in \bar{n}} T_r^2, \forall \bar{n} \in V^2$$

The second set of constraints make sure mass conservation of allocation and batch start quantity.

$$\sum_{\bar{n} \in s} Y_{\bar{n}s}^2(t) = \sum_{r \in s} Z_r^2(t), \forall t \in \bigcup_{r \in s} T_r^2, \forall s \in S^2(t)$$

The third set of constraints enforce only one batch is allowed in one facility at one time.

$$\sum_{r \in \bar{f}} X_r^2(t) \leq 1, \forall t \in T_{\bar{f}}, \bar{f} \in F^2$$

The fourth set of constraints keep certain level of safety stock for this stage.

$$\lambda_q \leq \sum_{\bar{n} \in H_q^2} \left(W_{\bar{n}}^2(t) + I_{\bar{n}}^2 + \sum_{r \in \bar{n}} \sum_{\tau + c_r \leq t} \theta \alpha_r (1 - \beta_r) \delta_r X_r^1(\tau) - \sum_{\tau \leq t} \sum_{s \in S_{\bar{n}}^2(\tau)} Y_{\bar{n}s}^2(\tau) \right)$$

$$\forall t \in \bigcup_{\bar{n} \in H_q^2} \bigcup_{r \in \bar{n}} T_r^2, q = 1, \dots, Q^2$$

For the demand constraint, cumulative demand at each demand epoch must be met either by initial inventory, WIP, new production, or by backorder.

$$I_{\bar{n}}^4(0) + W_{\bar{n}}^4(t) + \sum_{r \in \bar{n}} \sum_{\tau + c_r \leq t} \alpha_r (1 - \beta_r) \delta_r Z_r^3(\tau) - b_{\bar{n}}(t) \geq D_{\bar{n}}(t) \quad \forall t \in T_{\bar{n}}, \forall \bar{n} \in V^4$$

The Formulation's objective minimizes backorder cost and discounted total batch starts.

$$\text{Minimize } \sum_{m \in V^4} \sum_{t \in T_m} b_m(t) + \sum_{i=1}^3 \sum_{r \in R^i} \sum_{t \in T_r^i} \frac{1}{(1+\gamma)^t} X_r^i(t)$$

This optimal planning model effectively frames the deterministic complexity of the bio-pharmaceutical supply chain planning. First the time grid based approximation of batch

starting time reduces the number of integer variables significantly while at the same time give the formulation enough flexibility to explore for cost saving batch start plans. Secondly the model also addresses the complex problem of product mixing under regulatory constraints. By utilizing a wiring diagram which accounts for the complete batch history, production starts are merged together as much as possible subject to regulation requirements. Thirdly the model allows for batch specific processing characteristics including processing time, yield and reject rate etc.

Based on the core MIP formulation a planning platform called “Planning Engine” was developed. The planning engine platform imports data such as inventory status, work in process status, batch production history, batch production quality result, quality control results, regulatory approvals and customer demand into planning engine database. After integrating imported data as well as user updates to the data, the formulation is built dynamically and solved via CPLEX MIP solver in a reasonable time. Next based on the optimal solution, the planning engine generates optimal planning reports including batch starts report, batch allocation report, mass balance report, etc. Planning engine also incorporates scenario based planning which allows users to ask “what-if” types of questions and assess the impact of potential incidents such as demand disruption, facility shutdown, regulatory approval, etc.

A key feature of the platform is that wiring diagram and process routes are generated by the software at real time based on the current and projected regulatory approvals. The planning engine allows users to enter regulatory restrictions and approvals via a user interface. Based on current regulatory constraints, planning engine generates a list of valid path from raw material to finished goods. The results are then fed into formulation building module of the implementation.

Currently the planning engine platform has entered production environment and is being used by Bayer supply chain planners on a daily basis. The success of planning engine project drives the need for another important research topic in the biopharmaceutical supply chain management: risk management. In particular, analysis is required to determine the most appropriate safety stock levels to use in planning calculations.

Chapter 2

Biopharmaceutical Supply Chain Risk Mitigation

2.1 Review of Recent Work on Supply Chain Risk Management

This section reviews recent literature in supply chain risk management strategies, production planning under uncertainty as well as the approaches of integrating parameter estimation with model optimization.

2.1.1 Supply Chain Risk Management Strategies

One of the biggest obstacles in biopharmaceutical supply chain planning is information sharing and transparency among different departments. For example, a quality assurance department is responsible for testing intermediate and final products. And quality assurance cycle time takes up more than 90% of the total production time. One might assume that the quality assurance progress is shared with supply chain planners so they can make better plan based on quality assurance progress. But unfortunately this is usually not the case and only the results of quality assurance are reported to supply chain planners. This information barrier creates great uncertainty for supply chain planners and usually results in more safety stock holdings. As pointed out by Christopher and Lee (2004), the key to mitigate supply chain risk is improving information sharing among supply chain members. If members have no visibility over the supply chain, its managers will employ buffers, excess capacity and/or slack in lead times to hedge against the uncertainty as well as the lack of confidence in the information. In the quality assurance case, safety stocks serve as the buffer against invisibility.

To identify and hedge major risks in the biopharmaceutical supply chain, Juttner, Peck and Christopher (2004) proposed a general framework. First the risk source is assessed.

Then the consequences of such risk exposure are evaluated. The next step is to identify the supply chain strategy that drives such risk. Finally risk mitigating strategies are suggested. Focused specifically on pharmaceutical industry, Shah (2004) considered the key issues in pharmaceutical supply chain optimization and points out that as the R&D productivity is declining, effective patent life is shortening and market competition is increasing, the pharmaceutical supply chain is no longer just a tool to ensure market supply. Now more and more companies within the pharmaceutical industry are revisiting their supply chains to identify ways of extracting benefits from them.

2.1.2 Demand Risk and Supply Risk

Snyder and Shen (2006) studied the multi-echelon supply chain under two types of uncertainty: supply uncertainty and demand uncertainty. They show that the optimal strategy for dealing with supply uncertainty is in many cases the exact opposite of the optimal strategy for demand uncertainty. And in a practical setting with both demand and supply uncertainty, the optimal strategy must account for the interaction between supply and demand uncertainty. Snyder and Shen (2006) used simulation of a number of studies to demonstrate that the optimal strategy is different for a simple multi-echelon supply chain under demand or supply uncertainty.

2.1.3 Production Planning under Uncertainty

Stefansson, Jesson and Shah (2006) proposed an integrated multi-scale optimization model and solution method for the planning and scheduling of a pharmaceutical production process under uncertain demand. Their approach is hieratically structured and at the top level a campaign plan for the long term planning is optimized. At the middle level the campaign plan as well as the allocation of orders within the campaign are optimized. At the lowest level is the optimization of detailed scheduling of the production tasks. Yano (1987) considered the problem of determining optimal planned lead time in a serial production line where the processing time may be stochastic. The author presents a solution procedure for the two stage serial production system. For the N-stage serial production system the proposed solution procedure provides insight into the character of the optimal solution. Buzacott and Shanthikumar (1994) compared safety stock with safety time in MRP controlled production system under a single stage manufacturing framework. Their results showed that when the future required shipment lead time can be accurately predicted, safety time is usually preferable to safety stock. Otherwise, safety stock is more robust in coping with fluctuations in customer requirement lead time changes. Tang (1995) proposed a discrete time multi-stage production system with stage output rate uncertainty and customer demand uncertainty. He approximated the complex production rule with a linear “restoration” based production rule and shows that this rule leads the system to steady state with closed form mean and variance for both production level and inventory level. Chang (1985) explored the interchangeability of safety stock and safety lead time as buffering techniques for the uncertainty in the demand

and supply of components from lower levels. He shows safety stock is preferred as a buffering technique over the safety lead time and replacing the slack in lead time with safety stock can benefit the overall planning and scheduling of the MRP system. Denardo and Lee (1995) studied a serial production line with uncertain demand, processing time, yield, rework probability and reliability. The uncertainty they consider is linear in the sense that the mean of yield, scrap, etc is a linear function of the workload. A linear discrete time rule for production control is constructed and showed that the system tends to steady state conditions based on this rule.

2.1.4 Integrate Optimization and Estimation

Traditional approaches in risk management usually separate parameter estimation and policy optimization. First the unknown parameters are estimated from historical data and then one plugs the estimations into an optimization formula to find out the optimal policy. As showed by Liyanage and Shanthikumar (2004), this approach may lead to a suboptimal policy when compared to integrate estimation and optimization. Instead, they propose the operational statistics approach that combines estimation and optimization.

Another way for integrating uncertainty with optimization is robust linear programming. In robust linear programming, the problem data is considered uncertain and the uncertainty can be incorporated into a new formulation by converting the uncertain linear programming into a convex nonlinear program. In addition, risk terms can be incorporated into the objective yet still preserve convexity.

2.1.5 Operational Statistics

Traditional approaches for the inventory control problem with unknown demand distribution will first estimate the parameter with historical data and then maximize the expected profit based on the parameter estimation.

Liyanage and Shanthikumar (2004) introduced the idea of operational statistics to find out the optimal inventory control policy. Using the notation from Liyanage and Shanthikumar (2004), consider the classical newsvendor problem. Cost per item is c per unit and selling price is s per unit. For simplicity, salvage value of excess inventory is zero. Historical demand is assumed to be i.i.d. with exponential distribution F_D with unknown mean θ . We observe the demand in the last n periods and need to find the optimal order quantity.

Define $\hat{X}(z) = S(D_1, D_2, \dots, D_n, z)$ to be the order quantity estimated from the data with the optimization parameter z . Let $\eta(z)$ denote the expected profit for the order quantity $\hat{X}(z)$. Consider the class of statistics where $\hat{X}(z) = z\bar{D}$, $z \geq 0$. Then

$$\begin{aligned} \eta(z) &= E[\phi(z\bar{D}, \theta)] \\ &= E[s\theta(1 - \exp\{-\frac{z\bar{D}}{\theta}\}) - cz\bar{D}] \\ &= s\theta(1 - (\frac{n}{n+z})^n) - cz\theta, \quad z \geq 0 \end{aligned}$$

Optimizing over z we can find the optimal order quantity is

$$\hat{X}(z^*) = n \left(\left(\frac{s}{c} \right)^{\frac{1}{n+1}} - 1 \right) \bar{D}$$

Note that for this joint estimation and optimization approach, the optimal order quantity is $\bar{D} \ln \left(\frac{s}{c} \right)$ and it is within the class of statistics $\hat{X}(z)$. Clearly, operational statistics approach outperforms the standard sequential estimation and optimization approach.

Based on this result, Chu, Shanthikumar and Shen (2007) introduced a Bayesian analysis to find the optimal operational statistics and showed that Bayesian analysis leads to the optimal operational statistics.

2.1.6 Robust Linear Programming

Robust linear programming considers problems with uncertain data. We employ notation from Ben-Tal and Nemirovski (1995a). Consider a linear program

$$\min \{c^T x | Ax \geq b\}$$

Here the objective coefficient c are certain and but the data (A, b) are only known to lie within a set U . Then the robust version of the linear program can be written as

$$\min \{c^T x | Ax \geq b \ \forall (A, b) \in U\}$$

Note here the constraints $Ax \geq b$ are called hard constraints because any x must satisfy whatever instance of the data (A, b) within the set U .

If columns A_i of the constraint matrix are known to belong to a given convex set, then the case is call ‘‘column-wise’’ uncertainty. For example, if the constraints $Ax \leq b$, $x \geq 0$ have column-wise uncertainty, the constraints can be written as

$$\sum_{i=1}^n x_i A_i \leq b, \ x \geq 0, \ \forall A_i \in K_i$$

It can be shown that above set of constraints is equivalent to the worse case within the convex set K_i

$$A^* x \leq b, \ x \geq 0, \ a_{ij}^* = \sup_{a_i \in K_i} (a_i)_j$$

So in the case of column-wise uncertainty, the robust linear program is equivalent to a new linear program.

On the other hand, if the row a_i of the constraint matrix is known to belong to a convex set, the robust linear program is not necessary equivalent to a linear program. Consider following linear program

$$\min_x \{c^T x | a_i^T x \leq b_i, \ i = 1, \dots, m\}$$

where a_i belongs to the ellipsoid $\varepsilon_i = \{\hat{a}_i + R_i u : \|u\|_2 \leq 1\}$.

There are two common approaches to handle row-wise uncertainty: deterministic model and stochastic model. The deterministic model finds the intersection of all the half space constraints within ellipsoid ε , i.e.

$$b_i \geq \max_{a_i \in \varepsilon_i} a_i^T x = \hat{a}_i^T x + \max_{\|u\|_2 \leq 1} x^T R_i u = \hat{a}_i^T x + \|R_i^T x\|_2$$

So the robust linear program with row-wise uncertainty is equivalent to

$$\min_x \{c^T x \mid \hat{a}_i^T x + \|R_i^T x\|_2 \leq b_i, i = 1, \dots, m\}$$

This form is called a second-order cone optimization problem. So using a deterministic model a robust linear program with row-wise uncertainty is equivalent to a second-order cone programming (SOCP) problem.

The stochastic model approach ensures the uncertain constraints are met with some probability. We rewrite the problem as

$$\begin{aligned} \min \quad & c^T x \\ \text{st.} \quad & \text{prob}(a_i^T x \leq b_i) \geq \eta, i = 1, \dots, m \end{aligned}$$

If we assume a_i is Gaussian with mean \bar{a}_i and variance Σ_i . Then

$$\text{prob}(a_i^T x \leq b_i) = \Phi \left(\frac{b_i - \bar{a}_i^T x}{\|\Sigma_i^{1/2} x\|_2} \right)$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$.

So under Gaussian assumption, the stochastic model formulation can be written as

$$\begin{aligned} \min \quad & c^T x \\ \text{st.} \quad & \bar{a}_i^T x + \Phi^{-1}(\eta) \|\Sigma_i^{1/2} x\|_2 \leq b_i, i = 1, \dots, m \end{aligned}$$

When $\eta \geq 0.5, \Phi^{-1}(\eta) \geq 0$ the norm coefficient has positive coefficient and again the formulation is a second order cone program and can be solved efficiently.

2.1.7 Hedging Batch Rejection Risk in a Multi Stage Supply Chain Network

Johnston (2009) considered the problem of setting safety stock levels for a multi-stage supply chain network with an existing “production plan” of batch starts, batch assignments and demand dates. The existing production plan is generated from standard Manufacturing Requirement Planning software and can meet all demands if no failure occurs in the supply

chain network. The major risk he considers is batch rejection risk and the goal is to determine the safety stock level or additional starts required to make sure the “production plan” can be executed with a certain confidence level in case of batch rejections.

The problem is formalized as a problem of inventory placement in a supply chain network with production uncertainty. He introduces a probabilistic “set covering” constraint to guarantee that there are sufficient safety stocks to hedge against stockouts within certain confidence interval. Each instance of possible combination of batch rejections is modeled as a scenario. Each scenario is assigned a certain probability mass. The objective of the formulation is to minimize the weighted holding cost of safety stocks. The weight depends on the location of the stock. Formulation constraints first make sure that selected scenarios covered by the safety stock reach a certain total probability mass target. Also in a scenario that is chosen, formulation constraints make sure that the demands for that scenario is met. The inventory placement problem is then decomposed into master and sub-problems. The master problem is analogous to a set covering problem and the max flow sub-problem then determines that for a particular set of production failures if the safety stock would be sufficient to meet all the demand. Johnston also proposes an algorithm to solve the master and sub-problems to optimality. The algorithm is illustrated in a 3-stage serial production chain.

Johnston also discusses several extensions of the inventory placement problem. His first extension is to allow partial batch failures where it allows part of the batch to be rejected instead of the whole batch. In practice, these events do occur in biopharmaceutical supply chains. To reduce the computational complexity of scenario enumeration, Johnston also discusses Monte Carlo based scenario generation approach and network decoupling. In the end he also extends the basic model by allowing stochastic production lead times.

2.1.8 Stochastic Lead Time Literature Review

Liberatore (1979) considered one of the simplest stochastic lead time inventory models by extending the classical EOQ model. He assumes there is a stochastic lead time for ordering and the demands are not interchangeable. By formulating the expected total cost with planned lead time and ordering quantity as decision variables, a unique global optimal inventory policy can be found.

Zipkin (1985) considered the modeling of a batch production facility with stochastic demand and lead time. He proposes a model by combining inventory theory and queueing theory. By assuming no order crossing under stochastic lead time, the production process is modeled as a M/M/1 queue. The sojourn time of the queue is interpreted as the production lead time. Mean and variance of the lead time demand distribution can be derived from mean and variance of the queue sojourn time. He then formulates a convex optimization problem by minimizing long run average inventory holding cost and backorder penalty cost.

Eppen and Kipp Martin(1988) investigated the problem of setting safety stock when both the demand in a period and the lead time are stochastic. In the case the parameters of the demand and lead time distributions are known, by utilizing two simple examples they

argue that the standard procedure of assuming normally distributed lead time demand can result in incongruous behavior. Moreover, they point out that the correct interpretation is that the density of lead time demand is a convex combination of the normal density. For the case of unknown parameters for the demand and lead time distribution, they determine the variance of forecast error over the lead time without the normality assumption and use the calculated variance to set the safety stock level.

Song(1994) assessed the effect of lead time uncertainty in a simple stochastic inventory model. By assuming the demand follows a compound Poisson process and no order crossover, Song shows that stochastic larger lead time results in higher optimal base-stock level while more variable lead time leads to higher long-run average cost for fixed based-stock policy. Moreover, she presents that the impact of lead time variability on optimal policy depends on the inventory cost structure. In particular, a more variable lead time requires a higher optimal base-stock level if and only if the unit penalty (holding) cost rate is high.

He, Xu, Ord and Hayya (1996) evaluated the impact of order crossover on an inventory system with constant demand and iid stochastic time and show that the multi-cycle approach by considering order crossover can lead to a better inventory policy. Since the exact analysis of order crossover is intractable, they derive the cost savings lower bound by considering only the pairwise order crossover between the reference order and other orders. With simulation of exponential and uniform lead time they show that the inventory policy derived from the approximation is very close to the optimal.

Robinson, Bradley and Thomas (2001) evaluated the effects of order crossover under a base stock system. They present an iterative algorithm to computer the distribution of the number of outstanding orders. They define the inventory shortfall as the amount of inventory on order but hasn't arrived yet. By quoting the result first given by Zalkind(1978), they showed in an alternative proof that the variance of shortfall is less than the variance of lead time demand. Thus inventory policies based on lead time demand lead to higher inventory holding costs. The importance of using inventory shortfall in determining the order-up-to point is further demonstrated by several numerical examples.

Robinson etc. (2001) showed that for an integer valued discrete time lead time distribution, we have $\text{Var} [L^e] \leq \text{Var} [L]$. This result is proved first for the special case where the lead-time distribution has positive probability masses on at most two adjacent integer points; in that case, equality holds. Then for a general discrete distribution, an iterative approach is applied to construct a modified probability distribution until reaching the special case distribution. During the course of the construction, under the modified probability distribution the gap between the variance of the lead time and the variance of the number of orders outstanding keeps decreasing. As the gap of the final two-point distribution equals to zero, all the previous gaps are less than or equal to zero.

Bradley and Robinson (2005) explored the case of order crossover under stochastic lead time. Based on the assumption of periodic review and independent lead times, they propose an approximation of the inventory shortfall distribution and evaluate the performance of this approximation in the base-stock policy inventory system.

Hayya, Bagchi, Kim and Sun (2008) proposed the term “effective lead time” to describe the time between n th order placement and n th order arrival. Due to the effect of order crossover, the effective lead time has smaller variance than the lead time. They review related literature on stochastic lead time and classify them into three categories: ignoring order crossover, allow small order crossover and acknowledging order crossovers. To demonstrate the analytical complexity of crossover modeling, they analyze the effective lead time in two and three period models. At the end they discuss the possibility of modeling effective lead time as time series and use historical records for forecasting.

2.2 Biopharmaceutical Supply Chain Risks

In supply chain management, the term “risk” usually refers to the uncertainty from a certain source within the supply chain. So the first task for risk management is to identify sources of risks within the supply chain. In biopharmaceutical supply chains, the risk source can be the variation of customer demand or its forecasting error, uncertain output, regulatory uncertainty including approval timeline as well as regulatory requirement changes, the possibility of introducing a product line extension, etc.

2.2.1 Demand Risk

Demand uncertainty is one of the key factors in the biopharmaceutical supply chain. Demand of final goods for biopharmaceutical products can have substantial variability for companies with large pipelines and product varieties. Demand variations are due to a number of factors. First, identification of new patients drives the demand up. In mature market, this increase can be relatively low. But in emerging market, the rate of demand increase is usually significant. For example, Bayer Healthcare recorded a sales increase of 38% in China in 2007 while in the first quarter of 2008, the sales increase reached 65% (BHC World June 2008). Secondly, market competition also leads to demand fluctuation. For life-saving drugs, if a firm was not able to deliver customer orders on time, customers may buy products from its competitors and thus increase competitors’ demand. Thirdly, product line extension may change future demand patterns. Introduction of a new product variety in the product family may affect the demand of other varieties in the same family. As for biopharmaceuticals, introduction of higher potency product may reduce the demand for its lower potency counterpart.

2.2.2 Regulatory Uncertainty

One reason for variability in biopharmaceutical supply chains is that downstream batches are more specific to target markets and exposed to different regulatory requirements in different regions. For example, some region/country’s authorities require that in addition

to in-factory testing, samples of finished biological products need to be sent to the labs in that region and pass all the necessary tests before sale to patients. The testing labs are not transparent to the supply chain managers and thus create substantial difficulties for accurate planning. On the other hand some regions only require in-factory testing. On average, the in house testing duration is 10 days less than regional testing. Moreover sometimes this regulatory uncertainty is coupled with process change and will create a even bigger challenge for supply chain planners. A robust planning model needs to take into account this regulatory discrepancy among different regions and markets.

2.2.3 Process Risk Overview

Process risk refers to the uncertainty associated with the process performance parameters of the biopharmaceutical production. At each production stage, the supply chain is facing several types of potential risk based on the performance parameters; these include uncertainty in yield, variable processing time and output uncertainty. Furthermore, as a relatively new industry, there is continuous process improvement within the biopharmaceutical manufacturing process. As a result of process changes, there is risk of losing process licensure from the regulatory authorities.

Process yield is a key factor in measuring the performance of a manufacturing stage. In batch production planning, process yield is an important parameter for the accuracy of the planning model. If the yield parameter is too small, the production plan is exposed to potential backorder. On the other hand, underestimated yield leads to excess intermediate stage inventory. In biopharmaceutical production, key stages are biological processes and turn out to have substantial variation in yields.

Another important process parameter is the stage processing time. In biopharmaceutical manufacturing, each production stage is immediately followed by quality assurance to ensure the batch meets all the regulatory requirements. At each stage, production may take a few days while quality assurance may take a few weeks or even longer to finish. Another complexity associated with QA cycle time is that its distribution is highly irregular. This is due to the nature of the biological process and the QA testing procedures. After a batch is produced, if it passes all the tests, it will be released in an anticipated time. But if it failed a certain test, additional tests will be done to investigate the batch to determine if it meets the regulatory requirements. As each batch can have significant market value in terms of final goods, the batch won't be discarded until full investigation determines the batch won't be qualified for sale. So in biopharmaceutical manufacturing, it is the case that some batches are finally released after spending 200+ days in QA stage and then released as available inventory for the next stage.

At the quality assurance stage there are a series of tests performed in each batch. A strategy called "Conditional Release" has been introduced in the industry to reduce the quality assurance cycle time. The program divides QA testing into two phases. Each phase includes a number of tests. Batches are released to the next stage once phase I tests are

passed. Batches from the following stage won't be released until they pass phase II tests. For example, fermented batches are released to the purification stage once they pass a subset of QA tests for fermentation. Purified batches won't be released until all fermentation and purification tests are passed. This conditional release program certainly reduces the QA cycle time but it also increases the risk of mixing contaminated fermentation batches with good fermentation batches. As a result, purified batches are subject to higher reject risk. To mitigate this risk, we may need to increase purified material safety stock level. Thus conditional release is actually reducing safety time but increasing safety stock level. If the supply chain is already holding sufficient safety stock to face rejected batch risk, conditional release may turn out to be a good strategy to reduce production cycle time. The overall cost or benefit of conditional release depends on the distribution of batch rejection, the distribution of QA cycle times and also batch mixing.

Besides yield and QA time uncertainty, biopharmaceutical production also faces output variation. The output product type not only depends on the input material but also on the output batch characteristics of the process. Take purification stage as an example. The output batch intensity follows a distribution rather than fixed within certain range. For a specific target intensity product, only batches in certain range may be allowed to use as source batch for the next stage. Batches outside this range will either have to be downgraded to lower intensity level product or discarded.

2.2.4 QA Cycle Time Risk

As mentioned in section 1.1, all the intermediate products in biopharmaceutical production require QA testing to ensure that regulatory requirements are met. At each stage, tests are organized in the following manner. First a set of initial tests are conducted. If the batch passes these tests, it will be released. Batches released in this scenario have relatively short cycle time. But if the initial tests fail, further tests will be done until the batch is confirmed to fail the regulatory requirement. This testing process explains the long tails in the QA cycle time distribution. It is common for a batch to experience over one hundred tests before release as finished goods. Figure 2.1 illustrates the quality assurance test flow.

2.2.5 Production Process Change Risk – Conditional Release

The goal of the conditional release program is to reduce the QA cycle time of fermented material. A conditional release program separates fermented material testing into two phases:

Phase 1 release occurs when a batch of fermented material passes a minimum set of regulatory requirements. These fermented batches will then enter purification stage to produce purified material. However, in parallel with purification production and quality assurance, testing on a sample from the batch continues until the batch passes all required tests.

Phase 2 release occurs when the sample from a batch of fermented material passes all

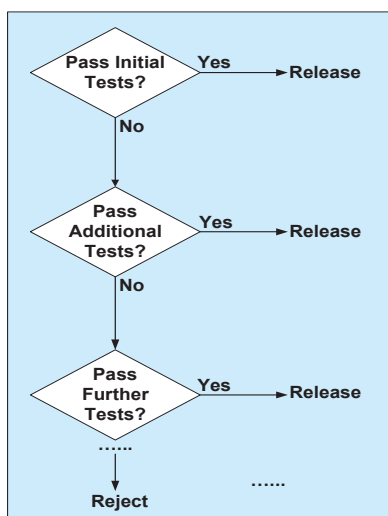


Figure 2.1: QA Test Flow Chart

required safety regulations. This may take up to 90 days after Phase 1 release has occurred and the UFTCF batch may have already been distributed into creation of purified lots. This poses a risk because if Phase 2 results in a reject, all subsequent purified lots using the rejected batch must be blocked.

The conditional release reduced the fermented QA cycle time as only a minimum set of tests are required. And in parallel with phase 2 fermentation tests, purified batches are produced and enter into purification QA stage. Thus the overall production time is reduced as a result of the conditional release program.

But the downside of the conditional release is the increasing risk of purified batches being blocked. Moreover, the source of purified QA cycle time variability is not just from purified QA tests, but also from fermentation QA tests as a result of this program. A case study of the conditional release program at one manufacturing line shows that fermentation QA cycle time is reduced by more than 50% percent after the implementation of the program. But the downstream stage incurs a significant increase in the average and standard deviation of QA cycle time.

After the purification step, UFDF lots enter the FFD step to be filled into vials and then become FC. FC batches will be packaged into FG and ready for shipment. Compared to fermentation and purification steps, FFD and packaging steps are relatively easier technologies and the processes are better understood. So we would expect shorter QA cycle time and less variability for FC and FG. Historical performance data shows that FC QA cycle time is around 30 days while FG QA cycle time is around 20 days. But historical data analysis shows that the QA cycle time variability is still significant for FC and FG.

One reason for the variability is that the FC and FG batches are closer to target markets

in our supply chain and exposed to different regulatory requirements in different regions. For example, European authorities require that in addition to in-factory testing, samples of FG need to be sent to regional labs in Europe and pass all the necessary tests before sale to patients. In contrast, in US the FDA only requires in-factory testing. So on average, US FG QA cycle time is 10 days less than European FG QA cycle time.

2.3 Current Practice to Hedge against Risk

Currently in order to hedge against the various risks mentioned above, biopharmaceutical companies employ a “days-on-hand” approach. The target safety stock level at the critical production stage is set to the supply meeting a certain number of days of demand. For example, at the purification stage, the UFDF safety stock is set to ensure supply up to 6 month of downstream production. The level of days on hand holding is based on planners’ qualitative understanding of the supply chain and years of experience. The level of stock holding is divided into two main part: pipeline stock holding and strategic holding. Pipeline holding level is based on a current estimate of the cycle time of the supply chain. Potential risk of long cycle time and batch reject are also considered when deriving the pipeline holding level. Strategic holding targets major risks in the supply chain including facility shutdown due to contamination, earthquake, factory shutdown, etc. Although the these are really rare events, strategic holding is a significant portion of the total holding level.

This “days-on-hand” approach has a number of worth-noting shortfalls and may lead to excess stock holding. First, this approach doesn’t quantify a number of important risk issues in biopharmaceutical supply chain. These risks include batch reject, quality assurance cycle time, demand variation, etc. Second, due to its qualitative nature, improvement in the supply chain cycle time or yield won’t be reflected in the days-on-hand inventory level in the short run. Thus the supply chain is exposed to process change risk. Third, this approach does not plan safety stock levels according to the requirement from downstream product varieties. Especially in the later production stages, as batches are more and more market region specific, this “days-on-hand” policy will either have reliability issues or incur excess inventory cost as a result of conservative safety stock level.

As an example, one particular biopharmaceutical manufacturing site employs a highly conservative safety stock level. Its purification step is built to stock and purification stage builds up enough UFDF safety stock to supply six months of downstream requirements. The reason is that purification is the most complex stage in the supply chain and exposed to very long cycle time and various risks. FFD and packaging steps are built to order to make sure that inventory holding is maintained at certain target level. As a result of this policy, currently the supply chain management department is holding a significant amount of inventory.

2.4 Safety Stock vs. Safety Time for Biopharmaceutical Supply Chain Risk Mitigation

2.4.1 Cumulative Curve perspective of the Biopharmaceutical Supply Chain Risk

We can summarize the impact of supply chain risk by evaluating its impact on cumulative supply and demand curve. Market demand fluctuations move the cumulative demand curve up or down. Long QA cycle time delays the output of the production so it will move the cumulative supply curve forward in time. Yield loss moves cumulative supply curve downward and process output uncertainty moves cumulative supply up for some product types while others down. So one way to mitigate these risks is by moving the cumulative in the reverse direction. For example, to mitigate the risk of long QA cycle time, we can put slack in lead time and this will move the cumulative supply curve backward in time. Figure 2.2 illustrates this idea.

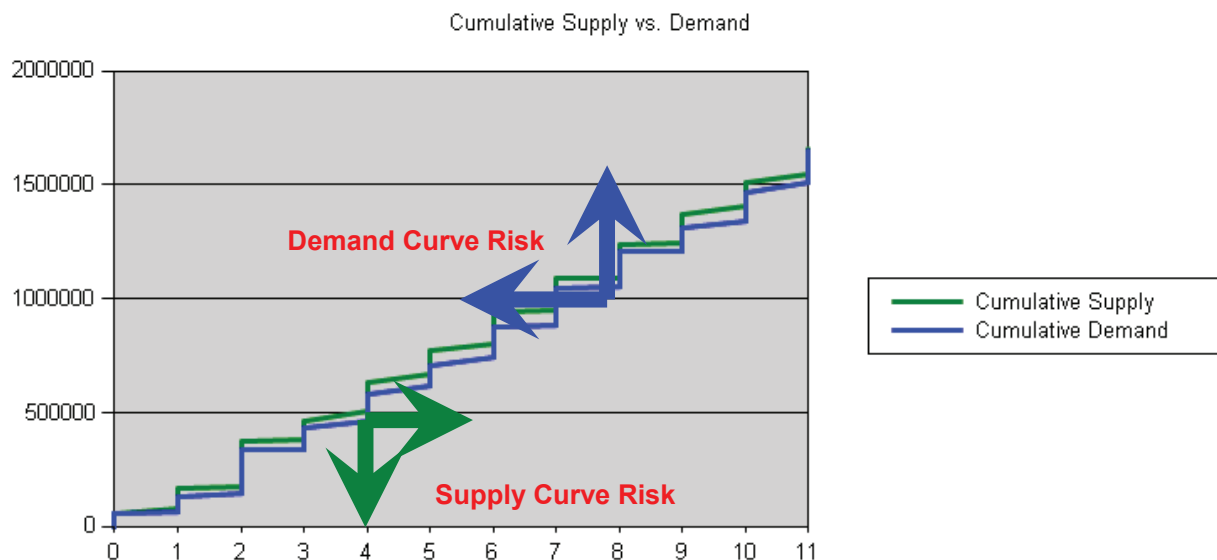


Figure 2.2: Cumulative Curve Perspective of Supply Chain Risk

2.4.2 Safety Stock versus Safety Time in Mitigating Supply Chain Risk

By comparing cumulative supply and cumulative demand curves, the backorder quantity or inventory quantity is just the distance between the two. To guarantee a 100% service level, we want to make sure cumulative supply curve is always above the cumulative demand curve.

To maintain a certain safety level, we can keep a certain distance between the cumulative supply curve and the cumulative demand curve.

There are two approaches to hedge against risks: safety stock and safety time. With safety stock, we can maintain extra inventory in the system and with safety time, we put slack in the planned lead time. In a complicated supply chain network, it is an open problem to decide which one is a better tool to effectively hedge against risks without excess inventory.

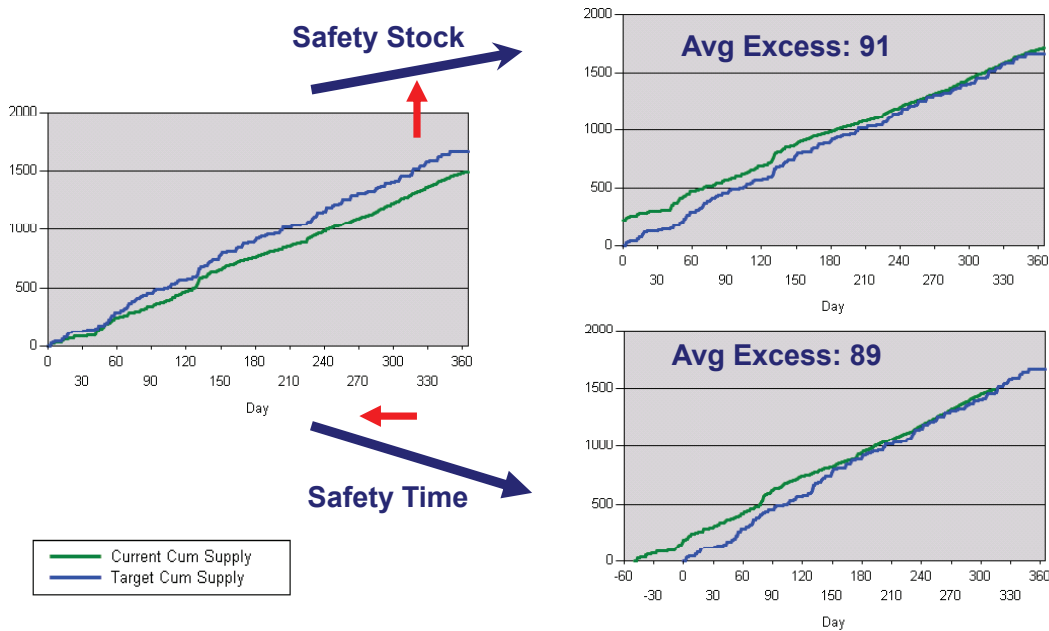


Figure 2.3: Mitigate Risk with Safety Stock or Safety Time

Figure 2.3 illustrates a numerical example of mitigating risk with safety stock or safety time. Initially, the current cumulative supply curve is behind the target cumulative supply curve. With safety stock approach, we can identify the maximum gap between current and target cumulative curve and use safety stock to cover the gap. This approach results in 91 units of average excess inventory. By using safety time, we can start production earlier in the horizon which is equivalent to move the current cumulative supply curve to the left until it's above the target curve. This approach results in average excess inventory of 89 units.

2.4.3 Risk Analysis via Simulation - QA Cycle Time's Impact on Planned Backorders

In this section we evaluate the impact of setting a planned QA cycle time vs. the backorder quantity. Although the planning model will be less likely to backorder when we increase QA cycle time estimation for our planning, it is unclear what level of impact we will have

Percentile	50th	60th	70th	80th
UFDF QA Time	28	31	37	55
FC QA Time	25	31	34	39

Table 2.1: Simulation Settings - QA Cycle Time's Impact on Backorders

regarding the service level. A simulation model is developed to evaluate such impact. The simulation setup is following:

1. Set purification and FFD QA times to 50th, 60th, 70th, 80th, 90th value of 2007 QA actual cycle time distributions
2. Generate random demand over a one year span based on recent historical trends. In order to compare the amount of backorders, the random demand is generated such that it exceeds the currently available surge capacity. So in all the QA settings, we will have a certain amount of backorders. We want to make sure that our supply meets demand and the curve generated in this step is the cumulative planned supply.
3. For fair comparison, we make all work-in-process come out at the same time.
4. Simulate random QA release following the actual historical distribution of year 2007. When QA release date is after planned release date, the production start with the corresponding batch is delayed until the actual release date.
5. Compare the planned and simulated cumulative supply curve to evaluate possible back-order quantity.

The QA cycle time settings corresponding to 50th, 60th, 70th, 80th, and 90th percentile are shown in Table 2.1. We measure the impact of QA cycle time by calculating the average backorder quantity of the generated plan based on the cumulative planned supply curve and cumulative actual supply curve.

Figure 2.4 is the simulation result for a production plan with QA cycle time set to 50th percentile of historical distribution. The green curve is the cumulative planned supply curve and the blue curve is the cumulative actual supply curve. Note that the actual supply curve is 1000 iteration average. The X axis is our planning horizon and the Y axis is the quantity. We can see that by setting QA cycle time to the historical average, actual supply is way behind planned supply.

Table 2.2 illustrates the robustness of different QA cycle time settings. The first column is the percentile of choice for our QA cycle time setting. Second and third columns are corresponding purification and FFD QA cycle times. The fourth column is the average gap between cumulative actual supply and cumulative planned supply, i.e. the backorder quantity. The fifth column is the standard deviation of the gap. The last column is the average gap plus two standard deviation of the gap. The quantity in the last column indicates

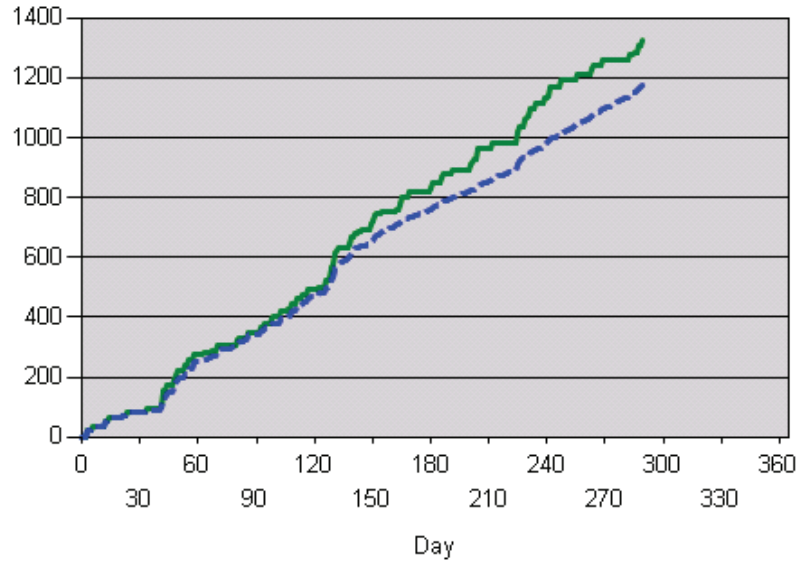


Figure 2.4: Planned Cumulative Supply vs. Simulated Average Cumulative Supply - 50th Percentile

QA Setting	Purif. QA Time	FFD QA Time	Avg Gap	Avg STD	Gap + 2STD
50th	28	25	65	13	91
60th	31	31	54	13	80
70th	37	34	45	12	69
80th	55	39	23	8	39
90th	73	56	8	5	18

Table 2.2: Planned Cumulative Supply vs. Simulated Average Cumulative Supply

how much additional stock holdings are necessary to hedge against stock out risk. As we increase our QA setting percentile, the average and standard deviation of the gap between cumulative planned supply and cumulative actual supply gets smaller and smaller. This is the effect of safety time we put into the model.

Table 2.2 illustrates the effect of safety time on the backorder quantity. We can consider the percentile as the parameter for our safety time approach against supply chain risks. To avoid stock outs, we also need safety stock to make sure cum actual supply is above cum planned supply. Thus we can consider the last column of Table 2.2 as the parameter for our safety stock approach against supply chain risks.

To compare the overall performance of different combinations of safety time and safety stock approach, we need to compare the average stock holding. Table 2.3 calculates the average stock holding for different combinations of safety stock and safety time. The first column is the percentile for QA cycle setting. The second and third columns are QA cycle

QA Setting	UFDF QA	FC QA	Avg FG Gap + 2STD	Avg FC Holding	FC+FG
50th	28	25	91	57	148
60th	31	31	80	77	157
70th	37	34	69	84	153
80th	55	39	39	108	147
90th	73	56	18	172	190

Table 2.3: Average Stock under Different Safety Stock and Safety Time Combination

time settings for UFDF and FC. The fourth column is the additional stock holding to avoid backorders. The fifth column is the average FC holding. The last column is the total average stock holding. Note that as we increase our QA cycle time percentile, we need less and less additional stock FG stock holding. But at the same time we have more and more average FC holding. This is because longer safety time means for those releases earlier than planned, we won't start production until planned date. So longer safety time incurs holding cost on QA material released earlier than planned.

The decreasing additional FG stock holding and increasing FC average holding illustrates the tradeoff between safety time and safety stock. But in terms of total stock holding, we can see that longer safety time incurs more much average total inventory holding. So in biopharmaceutical supply chain, safety stock is a more cost effective approach to hedge against supply chain risks. One justification of this behavior is that in biopharmaceutical planning we usually mix batches from previous stage to produce batch for the next stage. If we choose safety time to hedge against risks, any delay for one of the input batches will cause additional waiting of other input batches. Although we may have less possibility of waiting when we set QA cycle time percentile to 90%, any extreme QA release from an input batch will cause significant additional holding of stock. On the other hand with safety stock we are able to pool the risks of the same product variety into one stock pile and remove dependency among input batches.

Chapter 3

Stock Planning in Biopharmaceutical Supply Chain

3.1 Safety Stock Level Planning under Stochastic Lead Time

In this section we will analyze the safety stock level in biopharmaceutical supply chains when there is stochastic lead time. The inventory system we are considering is a base stock system in which per period demand is i.i.d. and an order with i.i.d. lead time is placed every period. We assume the lead time distribution and demand distribution are independent of each other. In order to maintain a certain base stock level, the ordering amount is equal to realized demand in the previous period. The goal is to determine the required stock level to guarantee a target service level.

First we will look at the conventional approach for safety stock planning and point out its drawback when applying to safety stock planning in biopharmaceutical supply chain. Next starting from an order crossover numerical example we will discuss our proposed approach based on crossover analysis. To make the model available for practical use, we will incorporate order interval and planned lead time. In the end we will calculate safety stock savings and use Monte Carlo simulation to evaluate the performance of the model.

Throughout our analysis in this section and subsequent sections, we will calculate the required stock level at the beginning of our planning horizon. There are two main parts for the required stock level: pipeline stock and safety stock. Pipeline stock is the result of lead time and is a function of average lead time. Safety stock is for hedging against any possible variation in lead time and demand. In the perspective of cumulative curve described in Section 2.4.1, pipeline stock is the average gap between cumulative demand and cumulative supply. Safety stock is expressed as a multiple of the standard deviation of the gap between cumulative demand and cumulative supply.

3.1.1 Safety Stock Level Planning Based on the Conventional Approach

The typical approach in the literature assumes there is no order crossover or the probability of such event is small enough to be ignored. A justification of this approach is that orders are not interchangeable so we cannot use early arrival orders to meet backorders. Based on this assumption the multiple-period model can be reduced to a single period one. To find out how much safety stock is needed to prevent backorders, we look at the lead time demand distribution which is the amount of demand arriving during the ordering lead time. The lead time demand is the convolution between lead time distribution and demand distribution. Its exact distribution is hard to characterize but we can calculate its mean and variance by conditioning.

Let L be the lead time distribution with mean μ_L and variance σ_L^2 . Let D be the demand for a single period with mean μ_D and variance σ_D^2 . Let LTD be the lead time demand. We assume that the lead time is independent of the period demand. Then we can calculate the mean and variance of lead time demand by conditioning on the lead time

$$\begin{aligned}
 \mathbb{E}[\text{LTD}] &= \mathbb{E}[\mathbb{E}[\text{LTD}|L]] \\
 &= \mathbb{E}[\mathbb{E}[\sum_{i=1}^L D_i]] \\
 &= \mathbb{E}[L\mu_D] \\
 &= \mu_D\mathbb{E}[L] \\
 &= \mu_D\mu_L
 \end{aligned}$$

Using the conditional variance formula $\text{Var}(X) = \text{Var}(\mathbb{E}[X|Y]) + \mathbb{E}[\text{Var}(X|Y)]$, we can find the variance of the lead time demand

$$\begin{aligned}
 \text{Var}[\text{LTD}] &= \mathbb{E}[\text{Var}[\text{LTD}|L]] + \text{Var}[\mathbb{E}[\text{LTD}|L]] \\
 &= \mathbb{E}[\text{Var}[\sum_{i=1}^L D_i|L]] + \text{Var}[\mathbb{E}[\sum_{i=1}^L D_i|L]] \\
 &= \mathbb{E}[\sum_{i=1}^L \text{Var}[D_i]] + \text{Var}[\sum_{i=1}^L \mathbb{E}[D_i]] \\
 &= \mathbb{E}[L\sigma_D^2] + \text{Var}[L\mu_D] \\
 &= \mu_L\sigma_D^2 + \mu_D^2\sigma_L^2
 \end{aligned}$$

To prevent backorders, let the base stock level B be the mean lead time demand to make sure on average we have enough inventory to meet demand during ordering lead time

$$B = \mu_D\mu_L$$

In addition to the base stock level, we need a safety stock level SS to cover the random fluctuation of lead time and per period demand. Let k be a service factor that measures the service level we want to guarantee. Given the variance of the lead time demand, we set the safety stock level to be

$$SS_L = k\sqrt{\text{Var} [\text{LTD}]} = k\sqrt{\mu_L\sigma_D^2 + \mu_D^2\sigma_L^2}$$

The required stock level is sum of pipeline stock and safety stock

$$S = \mu_D\mu_L + k\sqrt{\mu_L\sigma_D^2 + \mu_D^2\sigma_L^2}$$

Given that the lead time demand distribution is approximately normal then $k = 2$ yields over 97.5% probability that there are no back orders.

As a practical numerical example, let's assume we are planning the required stock level for purified material. Purification stage is the most technically difficult stage in biopharmaceutical manufacturing. Let's assume the ordering lead time has a Geometric distribution with mean 56 days. Daily demand for the purification stage has a mean of 8 mmu and standard deviation of 1.6 mmu. Safety factor is 3. Then conventional approach gives us a total required stock level of 1793 mmu. Here mmu is a common quantity measure for biological products in the biotech industry.

3.1.2 Order Crossover Analysis

In this section we will use a numerical example to illustrate the crossover effects among orders. In biopharmaceutical supply chains, usually there is significant lead time variation for each production stage. It is not uncommon that the standard deviation of the lead time is comparable to the mean. So in this example we choose an exponential distribution to illustrate this effect.

Let assume that order 1 is placed at the beginning of period 1 and order 2 is placed at the beginning of period $T+1$ where T is the ordering interval. The ordering lead time is exponentially distributed with rate λ . Let X_i be ordering lead time of order i . Then X_1 is the arrival time of order 1 and $T+X_2$ is the arrival time of order 2. The probability that order 2 arrives before order 1 is

$$\begin{aligned} P(X_1 > T + X_2) &= P(X_1 - X_2 > T) \\ &= \int_0^\infty \int_{x_2+T}^\infty \lambda e^{-\lambda x_1} \lambda e^{-\lambda x_2} dx_1 dx_2 \\ &= \frac{1}{2} e^{-\lambda T} \int_0^\infty 2\lambda e^{-2\lambda x_2} dx_2 \\ &= \frac{e^{-\lambda T}}{2} \end{aligned}$$

When the lead time is exponentially distributed with mean 30 and ordering interval is 3, the probability of order 2 arrives before order 1 is 0.45. If we apply safety stock policy based on lead time demand then we exclude almost 50% possibility of using early arrival orders to make up backorders. In the next section we will quantify the crossover effect and derive the required stock level with crossover taken into account.

3.1.3 Safety Stock Level Based on Crossover Analysis

In this section we will present a crossover based safety stock calculation. We'll look at the calculation in two different perspectives: cumulative curve and effective lead time. In the cumulative curve perspective, we try to model the gap between cumulative supply and cumulative demand and then derive the required stock level. In the effective lead time perspective, we explore the effect of crossover by replacing lead time with effective lead time and then apply traditional lead time demand approach. The cumulative curve perspective gives us a general framework to incorporate more risks in subsequent modeling. The effective lead time perspective can be seen as a generalization of the lead time demand model.

Cumulative Curve Based Safety Stock Level

Let d_i be the demand quantity in the i th period and $D_i = \sum_{j=-\infty}^i d_j$ be the total demand quantity up until i th period. We call D_i the cumulative demand curve. Let s_i be the supply quantity in the i th period and $S_i = \sum_{j=-\infty}^i s_j$ be the total supply quantity up until i th period. Then S_i is the cumulative supply curve. Note that due to the existence of stochastic lead time, batch reject, yield and excursion risk in our supply chain system, supply quantity per period should be interpreted as actual released quantity rather than production start quantity in that period. For the same reason cumulative supply curve is the cumulative actual released quantity.

To guarantee a 100% service level in a biopharmaceutical supply chain, we want to make sure that the cumulative supply curve is above cumulative demand curve almost surely throughout our planning horizon, i.e.

$$\forall i, P(S_i \geq D_i) = 1 \Leftrightarrow \forall i, P(S_i - D_i \geq 0) = 1$$

The notation $S_i - D_i$ refers to the gap between cumulative supply and cumulative demand. One way to derive safety stock level is by modeling cumulative supply curve and cumulative demand curve separately and then find out their gap to calculate necessary stock holdings. For example (Zipkin 1986), the demand arrival can be modeled as compound Poisson process and cumulative supply curve can be modeled as departure process of a queueing system. This approach usually assumes certain order of demand arrival, finished product release and noninterchangibility among orders.

Our approach is based on the modeling of the gap between cumulative supply and cumulative demand directly. The advantage of this approach is first that we allow order

crossover as well as interchangeability among compatible product types. Second we relax any condition on the order departure process and thus explore the benefit of order crossover. Figure 3.1 illustrates the concept of cumulative supply, cumulative demand and the gap between the two.

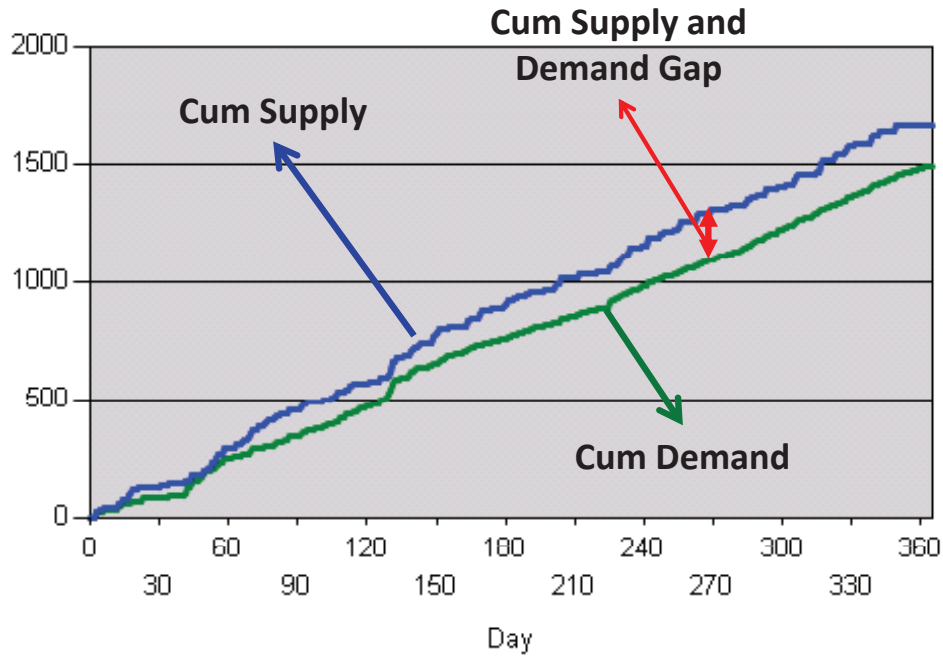


Figure 3.1: Cumulative Supply, Demand and Gap

We will start our analysis from a base stock system and then extend to more complex models. Assume there is i.i.d. customer demand per period and the order quantity placed per period is equal to the demand of the previous period. Each order placed takes i.i.d. stochastic lead time to be released. We further assume that customer order quantity is independent of order lead time. We will use following notation:

N : Number of outstanding orders. An outstanding order is defined as an order placed but not arrived as planned

μ_N : Average number of outstanding orders

σ_N : Standard deviation of the number of outstanding orders

SF: Shortfall. Shortfall is defined as the gap between cumulative supply and cumulative demand

F_i : Probability that order lead time is less than i days

T : Production order interval

L : Order lead time

μ_L : Average order lead time

σ_L : Standard deviation of order lead time

D_i : Order quantity in i^{th} period

μ_D : Average order quantity in one period

σ_D : Standard deviation of order quantity in one period

k : Safety factor

S : Required stock level to guarantee target service level

If there is no stock holding in this system, the cumulative supply curve for sure falls below cumulative demand curve due to the stochastic lead time. The gap between cumulative supply and demand curve is a random quantity and we will derive our required stock level based on the statistical properties of the gap.

To compute the gap, first we calculate the number of orders placed but not yet arrived. Then we express the gap between the cumulative supply and demand in terms of the number of outstanding orders. The analysis of this basic model is based on Robinson etc.(2001). A similar approach dates back to Zalkind (1976). As in Robinson etc. (2001), we will call the gap “shortfall” to denote the shortage quantity of cumulative supply behind cumulative demand.

Let 1_A be the indicator function which is 1 when the condition A is satisfied and 0 otherwise. We can find the number of outstanding orders by counting all the orders placed in the past but not yet arrived

$$N = \sum_{i=0}^{\infty} 1_{\{L_{t-i} > i\}}$$

Given the properties of an indicator function

$$E[1_A] = P(A)$$

$$\text{Var}[1_A] = P(A)(1 - P(A))$$

Then we can calculate the mean and variance of the number of outstanding orders

$$\begin{aligned} E[N] &= \sum_{i=0}^{\infty} P(L_{t-i} > i) \\ &= \sum_{i=0}^{\infty} (1 - F_i) \\ &= \mu_L = \mu_N \end{aligned}$$

$$\begin{aligned}\text{Var} [N] &= \sum_{i=0}^{\infty} F_i (1-F_i) \\ &= \sigma_N^2\end{aligned}$$

Note that the average number of outstanding orders is equal to the mean lead time. Also note that $\sigma_N^2 \neq \sigma_L^2$. Actually as shown by Robinson etc. (2001) and Zalkind (1976), we have the property that $\sigma_N^2 \leq \sigma_L^2$.

We then can compute the mean shortfall SF by conditioning on the number of outstanding orders. It is not surprising that the mean shortfall is just the product of the mean lead time and the mean per-period order quantity. Note that the mean shortfall is the same as mean lead time demand.

$$\begin{aligned}\mathbb{E}[\text{SF}] &= \mathbb{E}[\mathbb{E}[\text{SF}|N]] \\ &= \mathbb{E}[\mathbb{E} \left[\sum_{i=1}^N D_i \right]] \\ &= \mathbb{E}[N\mathbb{E}[D_i]] \\ &= \mathbb{E}[N\mu_D] \\ &= \mu_D\mu_N \\ &= \mu_D\mu_L\end{aligned}$$

The variance of SF can be calculated using the conditional variance formula.

$$\begin{aligned}\text{Var} (\text{SF}) &= \text{Var} [\mathbb{E}[\text{SF}|N]] + \mathbb{E}[\text{Var} [\text{SF}|N]] \\ &= \text{Var} [N\mu_D] + \mathbb{E}[N\sigma_D^2] \\ &= \mu_D^2\sigma_N^2 + \mu_N\sigma_D^2 \\ &= \mu_D^2 \sum_{i=1}^{\infty} F_i(1-F_i) + \mu_L\sigma_D^2\end{aligned}$$

To guarantee a target service level, we need the required stock level to cover both the average shortfall and its possible random fluctuation.

The pipeline stock level is

$$B = \mu_D\mu_L$$

The safety stock level is

$$\text{SS} = k \sqrt{\mu_D^2 \sum_{i=1}^{\infty} F_i(1-F_i) + \mu_L\sigma_D^2}$$

The required stock level in this case is:

$$\begin{aligned} S &= \mathbb{E}[\text{SF}] + k\sqrt{\text{Var}(\text{SF})} \\ &= \mu_D\mu_L + k\sqrt{\mu_D^2 \sum_{i=1}^{\infty} F_i(1-F_i) + \mu_L\sigma_D^2} \end{aligned}$$

Let's define the average stock holding as the average slack between cumulative supply and cumulative demand. Then in this model if there is no stock holding the average gap between cumulative supply and cumulative demand is just the average shortfall. Then with the required stock holding, the average inventory level is just the safety stock plus the pipeline stock since on average the safety stock will not be consumed. Let S_A be the average inventory holding. Then we have

$$S_A = \frac{1}{2}\mu_D\mu_L + k\sqrt{\mu_D^2 \sum_{i=1}^{\infty} F_i(1-F_i) + \mu_L\sigma_D^2}$$

Safety Stock Level with Effective Lead Time Demand

In this section we will derive the same result as in the previous section but from the perspective of effective "lead time demand".

The no-order-crossover assumption in the lead time demand analysis section eliminates the effect of interchangeability among different orders. Although no order crossover assumption gives us analytical convenience, in practice it will lead to excess inventory since it prevents orders arriving earlier to make up for orders arriving late. To assess the effect of interchangeability, assume n subsequent orders placed at time D_i with order arrival at time $A_i, i = 1, \dots, n$. The lead time for order i is $A_i - D_i$. When considering the effect of interchangeability, the actual lead time for order i is $A_{(i)} - D_i$, where $A_{(i)}$ is the order statistic for A_i . In other words, actual lead time is the time between i th order placement and i th order arrival. Hayya, etc. (2008) propose the term "effective lead time" to describe this time interval. To incorporate the effect of interchangeability, we can use the mean and variance of effective lead time to calculate the mean and variance of lead time demand and then find the necessary safety stock to guarantee a certain service level. Let L^e be the effective lead

time and L_i^e be the i th effective lead time. Then with the law of large numbers, we have

$$\begin{aligned}
\mathbb{E}[L^e] &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n A_{(i)} - D_i \\
&= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n A_i - D_i \\
&= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n L_i \\
&= \mathbb{E}[L]
\end{aligned}$$

To find out the variance of L^e , we use the result that L^e and N are equivalent, which is indicated by Hayya, etc. (2008). Intuitively if L^e is the actual time between order arrival and departure, then in the case of one order placed per period, the number of outstanding orders should be equivalent to the length of the effective lead time. Given that L^e and N are equivalent, we can calculate the variance of effective lead time as the variance of the number of outstanding orders

$$\begin{aligned}
\text{Var} [L^e] &= \text{Var} [N] \\
&= \sum_{i=0}^{\infty} \text{Var} [\delta(L_{t-i} > i)] \\
&= \sum_{i=0}^{\infty} F_i(1 - F_i)
\end{aligned}$$

To determine the level of inventory necessary to prevent backorders, we can use a similar approach as in the previous section except we use effective lead time. Let LTD^e be the demand during the effective lead time. Then conditioning on the effective lead time, the mean effective lead time demand can be calculated as

$$\begin{aligned}
\mathbb{E}[LTD^e] &= \mathbb{E}[\mathbb{E}[LTD^e | L^e]] \\
&= \mathbb{E}[\mathbb{E}[\sum_{i=1}^{L^e} D_i]] \\
&= \mathbb{E}[L^e * \mu_D] \\
&= \mu_D \mathbb{E}[L^e] \\
&= \mu_D \mu_{L^e} \\
&= \mu_D \mu_L
\end{aligned}$$

Conditioning on the effective lead time, we can calculate the variance of effective lead time demand

$$\begin{aligned}
\text{Var} [\text{LTD}^e] &= \mathbb{E}[\text{Var} [\text{LTD}^e|\text{L}^e]] + \text{Var} [\mathbb{E}[\text{LTD}^e|\text{L}^e]] \\
&= \mathbb{E}[\text{Var} [\sum_{i=1}^{\text{L}^e} D_i|\text{L}^e]] + \text{Var} [\mathbb{E}[\sum_{i=1}^{\text{L}^e} D_i|\text{L}^e]] \\
&= \mathbb{E}[\sum_{i=1}^{\text{L}^e} \text{Var} [D_i]] + \text{Var} [\sum_{i=1}^{\text{L}^e} \mathbb{E}[D_i]] \\
&= \mathbb{E}[\text{L}^e * \sigma_D^2] + \text{Var} [\text{L}^e * \mu_D] \\
&= \mu_L^e \sigma_D^2 + \mu_D^2 \sigma_{L^e}^2 \\
&= \mu_L \sigma_D^2 + \mu_D^2 \sum_{i=1}^{\infty} F_i(1-F_i)
\end{aligned}$$

To prevent backorders, let the pipeline stock level be the mean effective lead time demand

$$B = \mu_D \mu_L$$

Also we need a safety stock level SS to cover the random fluctuations of lead time and per period demand. Let k be a service factor that measures the service level we want to guarantee. Given the variance of the effective lead time demand, we set the safety stock level to be

$$\begin{aligned}
\text{SS}_{L^e} &= k \sqrt{\text{Var} [\text{LTD}^e]} \\
&= k \sqrt{\mu_L \sigma_D^2 + \mu_D^2 \sum_{i=0}^{\infty} F_i(1-F_i)}
\end{aligned}$$

The required stock level is

$$\begin{aligned}
S &= \mathbb{E} [\text{LTD}^e] + k \sqrt{\text{Var} (\text{LTD}^e)} \\
&= \mu_D \mu_L + k \sqrt{\mu_L \sigma_D^2 + \mu_D^2 \sum_{i=0}^{\infty} F_i(1-F_i)}
\end{aligned}$$

3.1.4 Incorporating the Ordering Interval and the Planned Lead Time

In this section we will extend our basic model to make it applicable to biopharmaceutical supply chain planning. The two additional parameters we will incorporate are the ordering interval and the planned lead time.

In the previous section we derive the required stock level given the case that an order is placed every period. In typical biopharmaceutical supply chains, all production is batch processing and each facility allows only one batch to be resident at a time. There is a substantial setup cost associated with each batch start. So supply managers combine production of the same product type as much as possible and it is often the case that a certain product is batch produced every few days. So we need to incorporate this order interval in our model.

Also in our basic model we assume we don't have any prior information regarding the demand quantity until the order is actually placed. In actual biopharmaceutical supply chain planning, future customer orders are collected from each region and combined together by sales and operational planning (S&OP) staff. These future orders are in the form of forecast demand which spans over the planning horizon. The demand planning horizon can be from a few months to more than a year. Based on the demand forecast, supply chain planners are able to start production certain periods prior to the order arrival. We call this time planned lead time and will incorporate it into our model as well.

Let T be the ordering interval. To simplify notation, we assume that the planned lead time is a multiple of the ordering interval. Let lT denote the planned lead time, where l is an integer.

We start our calculation from the number of outstanding orders. Also we assume the time reference point is right after the most recent order. When there is an ordering lead time, the number of outstanding orders is the difference between the number of late orders and the number of early orders. Late orders are those that arrived later than planned lead time l while early orders arrive earlier than planned. Rearranging terms, we can see that the effect of planned lead time is just a constant shift when compared to the case of no planned lead time.

$$\begin{aligned}
N &= -\sum_{i=0}^l \delta(L_{t-iT} \leq iT) + \sum_{i=l+1}^{\infty} \delta(L_{t-iT} > iT) \\
&= -\sum_{i=0}^l (1 - \delta(L_{t-iT} \geq iT)) + \sum_{i=l+1}^{\infty} \delta(L_{t-iT} > iT) \\
&= -l + \sum_{i=0}^{\infty} \delta(L_{t-i} > iT)
\end{aligned}$$

The mean and variance of the number of outstanding orders are as follows:

$$\begin{aligned}
E[N] &= -l + \sum_{i=0}^{\infty} P(L_{t-iT} > iT) \\
&= \sum_{i=0}^{\infty} (1 - F_{iT}) - l
\end{aligned}$$

$$\text{Var} [N] = \sum_{i=0}^{\infty} F_{iT}(1-F_{iT})$$

Let SF be the shortfall quantity. Let D_T be the the convolution of i.i.d demand over T periods. Mean SF can be calculated as

$$\begin{aligned} \mathbb{E}(\text{SF}) &= \mathbb{E}[\mathbb{E}[\text{SF}|N]] = \mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^N D_T\right]\right] \\ &= \mathbb{E}[N\mathbb{E}[D_T]] \\ &= \mathbb{E}[NT\mu_D] \\ &= T\mu_D\mu_N \\ &= T\mu_D\left(\sum_{i=0}^{\infty} (1-F_{iT}) - 1\right) \end{aligned}$$

The variance of SF is given by

$$\begin{aligned} \text{Var} (\text{SF}) &= \text{Var} [\mathbb{E}[\text{SF}|N]] + \mathbb{E}[\text{Var} [\text{SF}|N]] \\ &= \text{Var} [\mathbb{E} [ND_T]] + \mathbb{E}[\text{Var} \left[\sum_{i=1}^N D_T\right]] \\ &= \text{Var} [NT\mu_D] + \mathbb{E}[NT\sigma_D^2] \\ &= \mu_D^2 T^2 \sigma_N^2 + \mu_N T \sigma_D^2 \\ &= \mu_D^2 T^2 \sum_{i=0}^{\infty} F_{iT}(1-F_{iT}) + \left(\sum_{i=0}^{\infty} (1-F_{iT}) - 1\right) T \sigma_D^2 \end{aligned}$$

If the lead time L is geometric with parameter p and we assume there is no planned lead time, then

$$\begin{aligned} \mathbb{E} [N] &= \sum_{i=0}^{\infty} (1-F_{iT}) \\ &= \frac{1}{1 - (1 - p)^T} \end{aligned}$$

$$\begin{aligned}
\text{Var} [N] &= \sum_{i=0}^{\infty} F_{iT} (1 - F_{iT}) \\
&= \frac{(1-p)^{T+1} + p(1-p)^T}{1 - (1-p)^{2T}} \\
&= \frac{(1-p)^T}{1 - (1-p)^{2T}}
\end{aligned}$$

When T is large enough, the expected number of outstanding orders approaches to unity. And in this case we can not use conditional expectation to calculate the shortfall variance. In this case, the probability of order crossover is very small and the problem is then amenable to the conventional approach based on lead time demand calculation. So we can say that no crossover based lead time demand is a special case of our model when we let order interval go to infinity.

Figure 3.2 illustrates the relationship between the average number of outstanding orders and the ordering interval. The lead time distribution is geometric with mean 10. The demand rate is 1 per day with standard deviation of 0.2. When $T = 1$, the number of outstanding orders is equal to the mean lead time. When $T = 25$, the number of outstanding orders is slightly more than 1. This is because when the ordering interval is large enough, the possibility of crossover is very small and on average there is only one order outstanding.

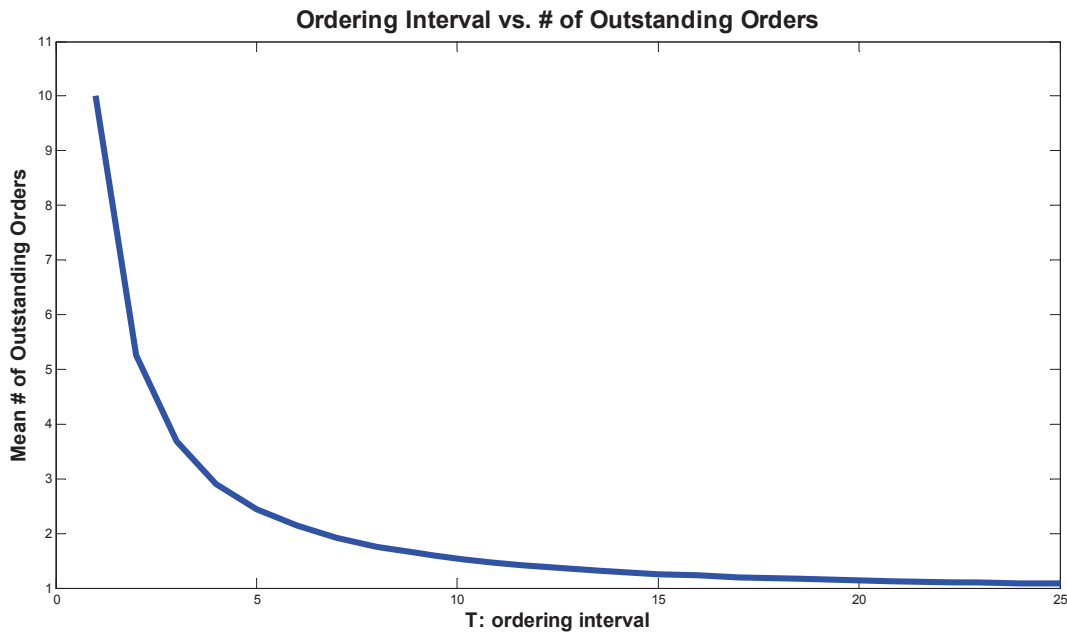


Figure 3.2: Ordering Interval and Average # of Outstanding Orders

Given the mean and variance of the number of outstanding orders, we can calculate mean and variance of the shortfall:

$$\begin{aligned}\mathbb{E}(\text{SF}_t) &= T\mu_D\mu_N \\ &= \frac{T\mu_D}{1 - (1 - p)^T}\end{aligned}$$

$$\begin{aligned}\text{Var}(\text{SF}_t) &= \mu_D^2 T^2 \sigma_N^2 + \mu_N T \sigma_D^2 \\ &= \mu_D^2 T^2 \frac{(1 - p)^T}{1 - (1 - p)^{2T}} + \frac{1}{1 - (1 - p)^T} T \sigma_D^2\end{aligned}$$

Based on the required stock holding, the average inventory holding S_A reveals the safety stock level:

$$S_A = \frac{1}{2} \frac{T\mu_D}{1 - (1 - p)^T} + k \sqrt{\mu_D^2 T^2 \frac{(1 - p)^T}{1 - (1 - p)^{2T}} + \frac{1}{1 - (1 - p)^T} T \sigma_D^2}$$

As we increase our ordering interval we are saving setup cost for order placement. But at the same time we need to hold more and more stock to cope with the uncertainty during the larger and larger ordering interval. Figure 3.3 shows the required stock level as a function of the ordering interval. All the parameters are the same as in the previous figure. The horizontal line indicates the level of required stock derived from lead time demand analysis. First note that required stock level is an increasing function of the ordering interval. This reflects the tradeoff between setup cost and holding cost. Second, note that we reach the level of lead time demand analysis when the ordering interval is around 12. This indicates that our crossover analysis based inventory policy can save us more than 90% of the setup cost given that we have same setup cost for large and small batches.

Let's examine the same practical numerical example as in Section 3.1.1. The ordering lead time is Geometric distributed with mean 56 days. Daily demand mean is 8 mmu and standard deviation is 1.6 mmu. Safety factor is 3. Also let's assume that we have a purification batch start every 3 days. Then based on the formulas in this section, the required stock level is 663 mmu. Recall in Section 3.1.1, required stock level derived from conventional approach is 1793 mmu. The percentage savings of required stock is 63%.

3.1.5 Safety Stock Savings

By incorporating order crossover and interchangeability, we explored the effect of early order arrivals making up for late ones. As shown in the previous section, the lead time and the effective lead time have the same mean but different variance. And we expect that the effective lead time has a smaller variance than the lead time. Robinson etc. (2001) shows

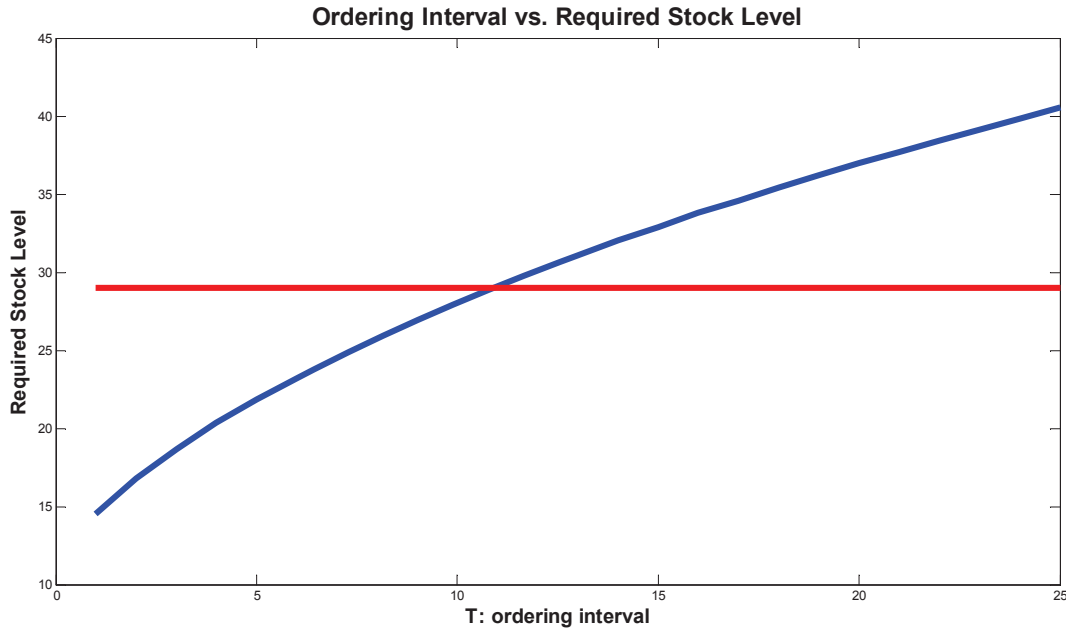


Figure 3.3: Ordering Interval and Required Stock Level

that for an integer valued discrete time lead time distribution, we have $\text{Var} [L^e] \leq \text{Var} [L]$. This result is proved first for the special case where the lead-time distribution has positive probability masses on at most two adjacent integer points, in which case the equality holds. Then for a general discrete distribution, an iterative approach is applied to construct a modified probability distribution until reaching the special case distribution. During the course of the construction, under the modified probability distribution the gap between the variance of the lead time and the variance of the number of orders outstanding keeps decreasing. As the gap of the final two-point distribution equals to zero, all previous gaps are less than or equals to zero.

The required stock level based on lead time demand is

$$S_L = \mu_D \mu_L + k \sqrt{\mu_L \sigma_D^2 + \mu_D^2 \sigma_L^2}$$

The required stock level based on crossover analysis is

$$S_C = \mu_D \mu_L + k \sqrt{\mu_D^2 \sum_{i=1}^{\infty} F_i (1 - F_i) + \mu_L \sigma_D^2}$$

Base on all the results so far, we can calculate the savings of required stock level by considering order crossover and interchangeability when there is no planned lead time and

the ordering interval as

$$S_L - S_C = k\sqrt{\mu_L\sigma_D^2 + \mu_D^2\sigma_L^2} - k\sqrt{\mu_L\sigma_D^2 + \mu_D^2 \sum_{i=0}^{\infty} F_i(1-F_i)}$$

where k is the service factor.

And the ratio of savings in required stock is

$$\begin{aligned} \frac{SS_L - SS_C}{SS_L} &= \frac{k\sqrt{\mu_L\sigma_D^2 + \mu_D^2\sigma_L^2} - k\sqrt{\mu_L\sigma_D^2 + \mu_D^2 \sum_{i=0}^{\infty} F_i(1-F_i)}}{k\sqrt{\mu_L\sigma_D^2 + \mu_D^2\sigma_L^2}} \\ &= \frac{\sqrt{\mu_L\sigma_D^2 + \mu_D^2\sigma_L^2} - \sqrt{\mu_L\sigma_D^2 + \mu_D^2 \sum_{i=0}^{\infty} F_i(1-F_i)}}{\sqrt{\mu_L\sigma_D^2 + \mu_D^2\sigma_L^2}} \end{aligned}$$

For the special case where we have deterministic demand (i.e. $\sigma_D = 0$), the ratio of safety stock savings is

$$\frac{SS_L - SS_C}{SS_L} = \frac{\sigma_L - \sqrt{\sum_{i=0}^{\infty} F_i(1-F_i)}}{\sigma_L}$$

Note in this case the stock savings only depends on the standard deviation of the lead time and the standard deviation of the effective lead time.

If the lead time distribution is geometric with parameter p , its variance is $\frac{1-p}{p^2}$, and the effective lead time has variance

$$\begin{aligned} \text{Var}(L^e) &= \sum_{i=1}^{\infty} F_i(1-F_i) \\ &= \sum_{i=1}^{\infty} (1 - (1-p)^i)(1-p)^i \\ &= \frac{(1-p)}{p(2-p)} \end{aligned}$$

The corresponding safety stock savings ratio is

$$\begin{aligned} \frac{SS_L - SS_{L^e}}{SS_L} &= \frac{\sqrt{\frac{1-p}{p^2}} - \sqrt{\frac{(1-p)}{p(2-p)}}}{\sqrt{\frac{1-p}{p^2}}} \\ &= 1 - \sqrt{\frac{p}{2-p}} \end{aligned}$$

When the lead time distribution is geometric with mean of 10 days, then $p = \frac{1}{11}$. By using effective lead time, in this case we can save almost 80% on safety stock.

Parameter	Value	Description
μ_D	10	Mean Demand
σ_D	2	Demand Std
μ_L	30	Mean Lead Time
H	750	Planning Horizon
T	5	Ordering Interval

Table 3.1: Simulation Parameters for Stochastic Lead Time Risk

3.1.6 Numerical Example and Simulation

In this section we use Monte Carlo simulation to study the performance of a base stock system under safety stock policy derived from crossover analysis. We assume per period demand is normal distributed with mean 10 and standard deviation 2. Ordering lead time is Geometric distributed with mean of 30 days. Simulation horizon is 750 days. An order is place every 5 days. The size of each order is 50. Also assume there is no planned ordering lead time. Table 3.1.7 summarizes simulation parameters.

Given the parameters in Table 3.1.7, we can calculate the required stock level as a function of safety factor k . Use formula from Section 3.1.4, we can get

$$\begin{aligned}\mathbb{E}(\text{SF}) &= T\mu_D \frac{1}{1 - (1 - p)^T} \\ &= 281\end{aligned}$$

$$\begin{aligned}\text{Var}(\text{SF}) &= \mu_D^2 T^2 \frac{(1 - p)^T}{1 - (1 - p)^{2T}} + \frac{1}{1 - (1 - p)^T} T \sigma_D^2 \\ &= 7709\end{aligned}$$

Then the required stock level as a function of safety factor k is

$$\begin{aligned}S(k) &= \mathbb{E}(\text{SF}_t) + k\sqrt{\text{Var}(\text{SF}_t)} \\ &= 281 + 88k\end{aligned}$$

To evaluate the performance of our stock policy, at the beginning of the simulation we set the inventory level to $S(k)$ given choice of safety factor k . Then first we simulate demand quantity day by day. Every T days an order is placed and actual lead time for this order is randomly generated from the lead time distribution. At the end of each day we calculate cumulative supply and cumulative demand. Cumulative supply is the total quantity arrived so far. The quantity depends on the actual lead time of each order. Also we add up demand to date to get cumulative demand.

When we reach the end of the horizon, we collect the simulation statistics for this simulation instance. Two key statistics are service level and average inventory holding. Service level is the percentage of periods that cumulative supply is above cumulative demand. In biopharmaceutical supply chain planning, the primary goal is to make sure that the service level is very close to 100%, for example 99.9%. In practice it's up to the supply chain manager to set a service level target given the inventory cost constraints. The main metric for the overall performance of a supply chain is the average inventory. It is the average inventory position over the planning horizon. Average inventory holding multiplied by unit value gives us the total cost for inventory holding. The unit value may vary depending on the specific stage of the product. Usually for biopharmaceuticals the closer to the final market, the more expensive it is. Ideally we want 100% service while maintaining minimum average inventory.

After collecting simulation statistics, we calculate the mean and standard deviation of service level and also the mean and standard deviation of inventory holding. Figure 3.4 shows the inventory holding distribution over the planning horizon for a typical iteration. We can see that there is no stock out in the planning horizon. The lowest stock level in the horizon is around 50 while at the end of the horizon the holding is around 300.

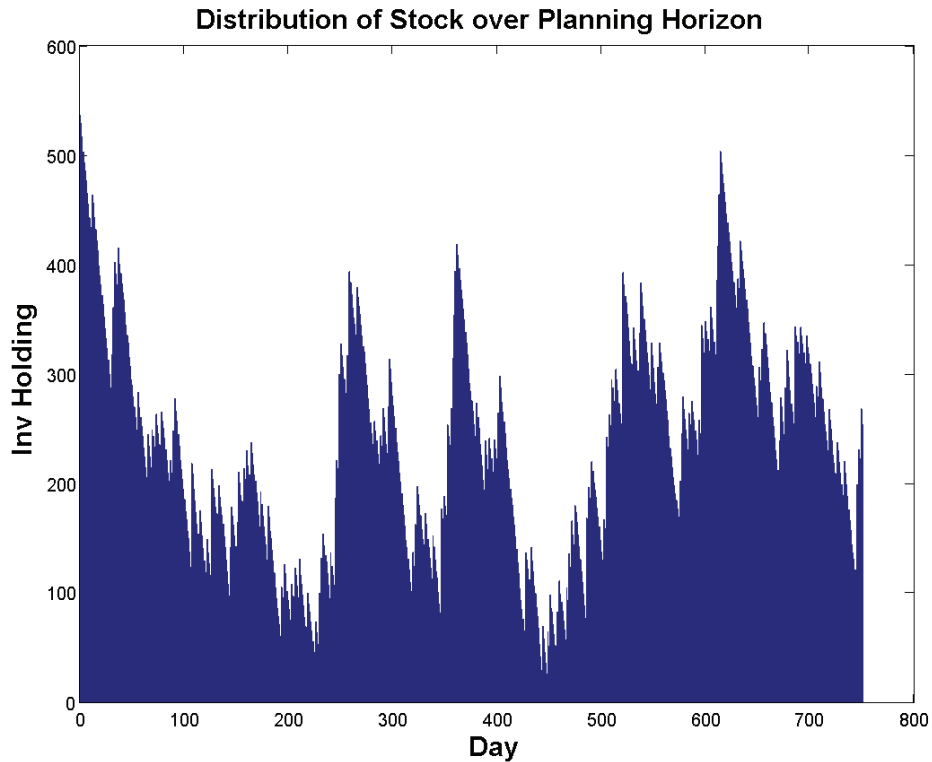


Figure 3.4: Stock Distribution in the Planning Horizon

From the simulation result we can see that setting service factor $k=2$ gives us more than 97% service level and an average inventory of 357. On the other hand, $k=0$ gives

	$k = 0$		$k = 0.5$		$k = 1$		$k=1.5$		$k=2$	
	B=280.7									
	SS = 0		SS = 43.9		SS = 87.8		SS = 131.6		SS = 175.5	
	Mean	Stde	Mean	Stde	Mean	Stde	Mean	Stde	Mean	Stde
Service Level	55.5%	0.1%	72.9%	0.1%	86.0%	0.1%	93.7%	0.1%	97.6%	0.10%
Inv Holding	43.4	0.1	71.6	0.2	107.1	0.2	145.9	0.2	188.2	0.3

Table 3.2: Simulation Results under Stochastic Lead Time Risk

us only around 50% service level but with much less inventory. With simulation we can also collect statistics regarding the distribution of the outstanding stock level and shortage quantity. Figure 3.5 is the distribution of the number of outstanding orders. Figure 3.6 shows the histogram of the shortage quantity. For both results the number of iterations is 10,000. Notice the outstanding orders histogram is more regular and unimodal.

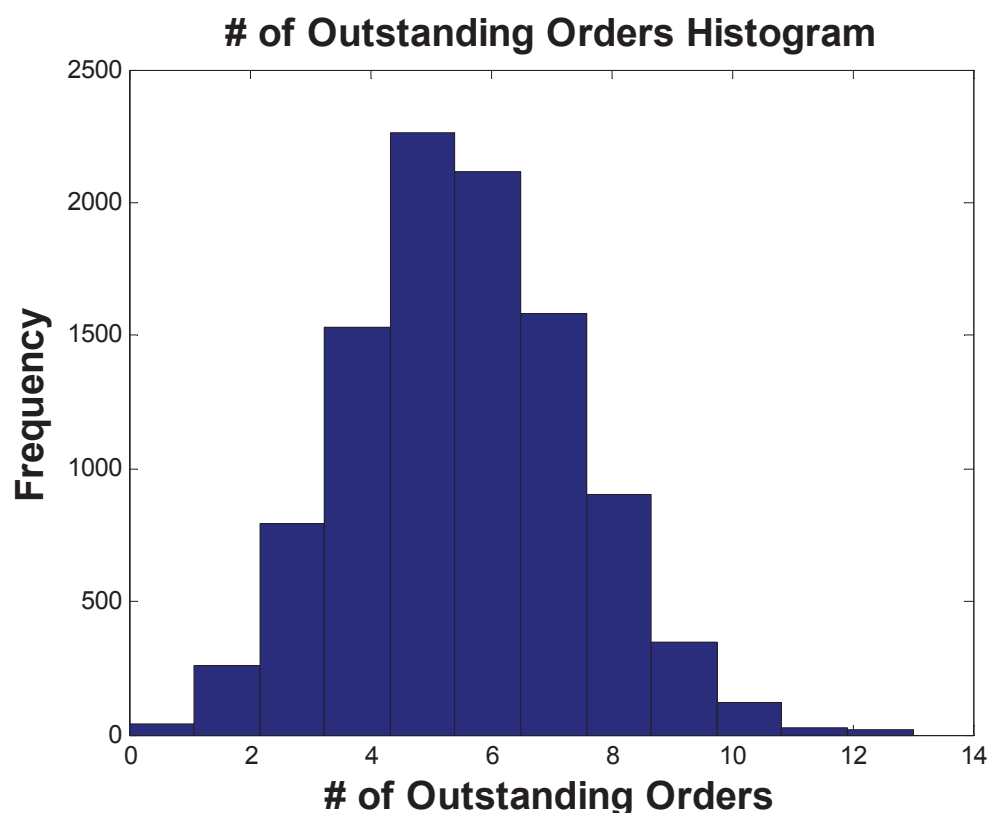


Figure 3.5: Distribution of # of Outstanding Orders

In biopharmaceutical supply chains, an important factor of performance measure is average inventory. Figure 3.7 summarizes the simulation results by comparing average inventory and service level. We can see that service level is an increasing concave function of average

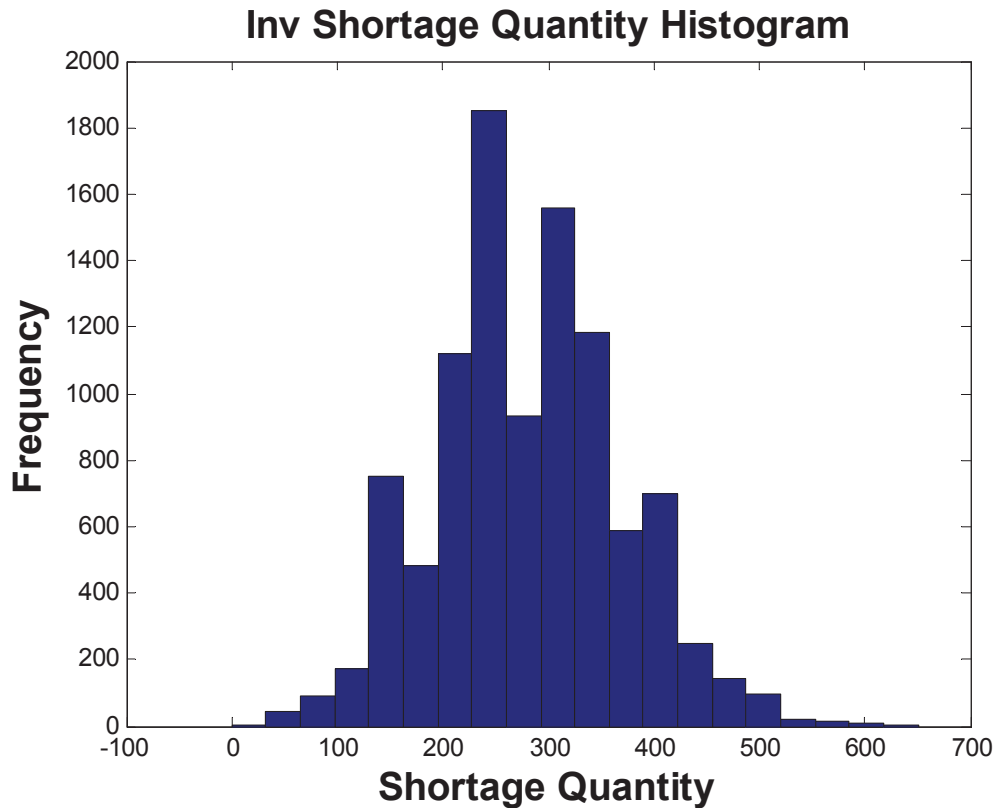


Figure 3.6: Distribution of Shortage Quantity

inventory.

3.1.7 Approximation of the Number of Outstanding Orders

During the derivation of the safety stock level an important assumption we made was that the shortage quantity is asymptotically normal. We match first the two moments of the shortage quantity to the normal distribution and then derive the required safety stock level given the target service level. By comparing the distribution of the number of outstanding orders and shortage quantity, we can see that the outstanding orders quantity histogram is more regular. This observation leads to an approximation of the number of outstanding orders as follows.

Let N be the number of outstanding orders. We approximate the distribution of N by matching its first two moments to a binomial distribution. Let μ_N and σ_N be the mean and standard deviation of the number of outstanding orders. Denote the parameters for the binomial distribution as n and p . Then our approximation assumes

$$\mu_N = np \tag{3.1}$$

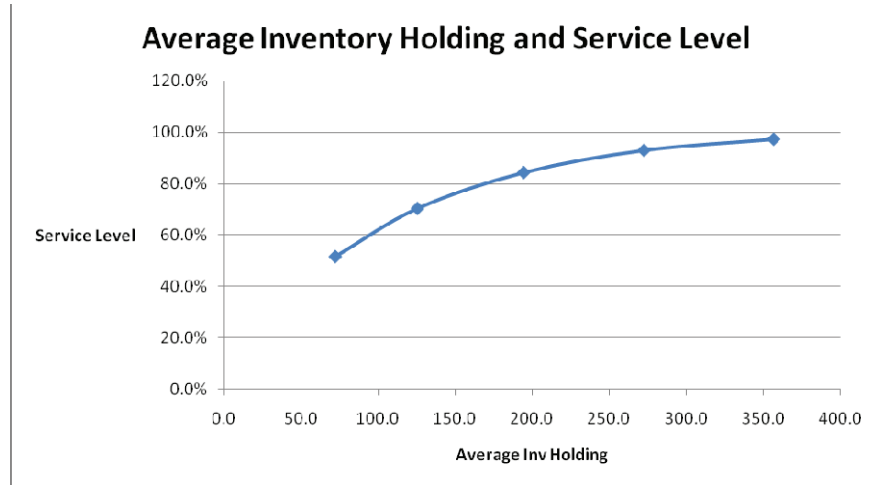


Figure 3.7: Average Inventory Holding and Service Level under Stochastic Lead Time Risk

and

$$\sigma_N^2 = np(1-p) \quad (3.2)$$

Solving for n and p with equations 3.1 and 3.2, we obtain the estimates

$$\hat{p} = 1 - \frac{\sigma_N^2}{\mu_N} \quad (3.3)$$

and

$$\hat{n} = \frac{\mu_N}{1 - \frac{\sigma_N^2}{\mu_N}} \quad (3.4)$$

Figure 3.8 illustrates the distribution of outstanding orders from simulation and its binomial approximation. The red bar is the distribution of number of outstanding orders from simulation. The number of simulation iterations is 10,000. The blue bar is its binomial approximation and the parameters of binomial is estimated from Equation 3.3 and 3.4. Notice at the right side tail, binomial approximation underestimates the density but the error is quite small.

Figure 3.9 compares the exact distribution of the number of outstanding orders with its binomial approximation. The blue bar represents the distribution of shortage quantity generated from the binomial approximation. The red bar distribution is collected from Monte Carlo simulation where simulation parameters are the same as in Table . The total number of simulation iterations is 10,000. Note that when the shortage quantity is less than 400, the approximation doesn't fit simulation very well while the right tail has a much smaller approximation error. But as we want to maintain the highest level of service we are more interested in right tail of the distribution. The mean shortage plus three standard deviations of the shortage is around 544.

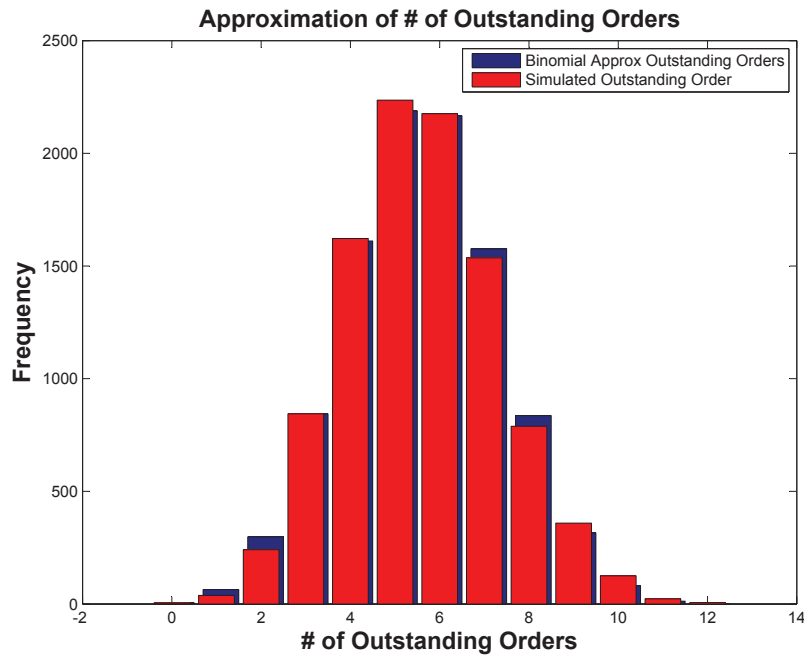


Figure 3.8: Binomial Approximation of # of Outstanding Orders

3.2 Safety Stock Level Planning under Excursion Risk and Stochastic Lead Time

In Chapter 3.1 we discussed safety stock level planning under an i.i.d stochastic lead time distribution. In practice the i.i.d assumption is often violated. A defective machine may cause processing orders to have a longer processing time. Contamination or quality testing may lead to longer cycle time and batch rejection. Thus we need to relax the i.i.d assumption to allow correlated lead times and often this means we need additional safety stock to cope with this kind of risk. In this section we will consider a special case of correlated lead time called excursion. The term “excursion” is borrowed from semiconductor industry and it is defined as any major deviation in process that is sustained until detection. In our discussion, when excursion happens, it means all the orders placed during the excursion fail to arrive. In practice this could be a machine breakdown or a major contamination of the facility

To model the excursion, we assume there is a stochastic time until excursion occurrence. The distribution of time to excursion occurrence is known in advance. After the excursion happens, it takes a deterministic time to discover the excursion and recover from it. For example, when a contamination happens in fermentation, we won’t detect the excursion until a number of required tests are performed over the fermentation batch. Next technicians will need to isolate the fermentation facility and then clean up the contaminants. In our excursion model we assume that the total time of excursion detection and recovery is known in advance. In practice this can be estimated by process engineers. During the excursion, we

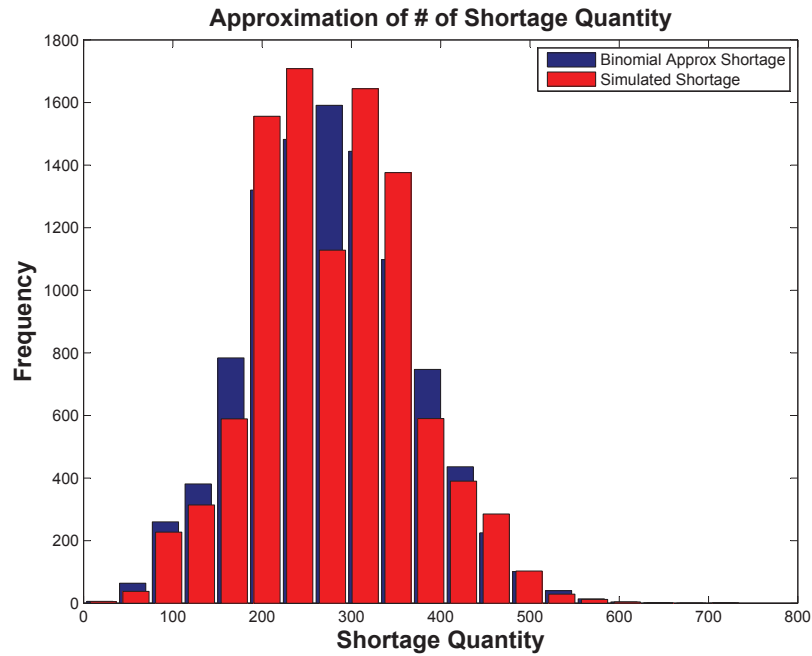


Figure 3.9: Binomial Approximation of Shortage Quantity

assume that all orders placed are discarded for quality purposes. This is typical practice in a biopharmaceutical supply chain to ensure the quality of the final product. Once the process is recovered from excursion, we assume that the lead time follows the same distribution as pre-excursion. Figure 3.10 illustrates the general framework of our excursion modeling.

We use following notation in this section.

N – Number of outstanding orders. An outstanding order is defined as an order placed but not arrived as planned

SF – Shortfall. Shortfall is defined as the gap between cumulative supply and cumulative demand

F_i – $F_i = P(L \leq i)$, probability that order lead time is less than i days

L – Order lead time

μ_L – Average order lead time

σ_L – Standard deviation of order lead time

D_i – Order quantity in i^{th} period

μ_D – Average order quantity in one period

σ_D – Standard deviation of order quantity in one period

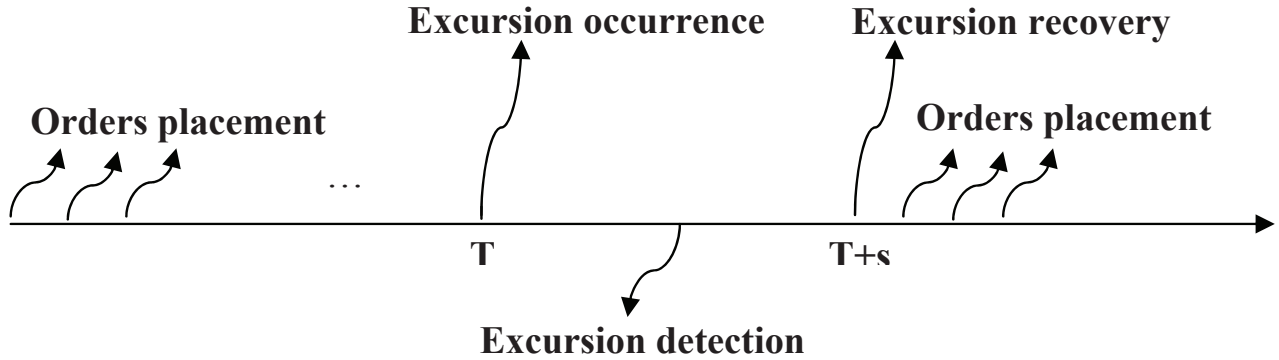


Figure 3.10: Excursion Model

T – Time to excursion occurrence

$G_i - G_i = P(T \leq i)$, probability that time to excursion is less than i days

s – Time from excursion occurrence to excursion recovery

In section 3.2.1 we consider the case when there is only one excursion. The main idea is to calculate the shortage of at the time of excursion as well as at the time when there is no excursion. The required stock level need to cover the maximum of the two. In section 3.2.2 we extend the model in section 3.2.1 and consider the case when there are multiple independent excursions.

3.2.1 Single Excursion

There are two major risk of concern in this model. One is the uncertainty in the ordering lead time and the other is the excursion risk. The goal of safety stock planning is to guarantee a certain service level under these two risks.

Suppose the time horizon starts at zero. Given an excursion happens at time t , the number of outstanding orders is

$$N_t = \sum_{i=1}^t \delta(L_{t-i} > i)$$

Its mean and variance are

$$\mathbb{E}[N_t] = \sum_{i=1}^t (1 - F_i)$$

$$Var [N_t] = \sum_{i=1}^t F_i(1 - F_i)$$

So the mean and variance of shortfall at the time excursion is

$$\begin{aligned}\mathbb{E}[SF_t] &= \mathbb{E}[N]\mathbb{E}[D] \\ &= \mu_D \sum_{i=1}^t (1-F_i)\end{aligned}$$

$$\begin{aligned}\text{Var} [SF_t] &= \mu_N \sigma_D^2 + \mu_D^2 \sigma_N^2 \\ &= \sum_{i=1}^t (1-F_i) \sigma_D^2 + \mu_D^2 \sum_{i=1}^t F_i (1-F_i)\end{aligned}$$

All the orders placed between time t and time $t+s$ will not arrive in the future because of the excursion. So during time t and $t+s$, the number of outstanding orders has two parts: outstanding order placed before time t and orders placed during time t and $t+s$. At time $t+s$, the number of outstanding orders is

$$N_{t+s} = \sum_{i=1}^t \delta(L_{t-i} > s+i) + s$$

Its expectation is

$$\mathbb{E}[N_{t+s}] = \sum_{i=1}^t (1-F_{s+i}) + s$$

And variance is

$$\text{Var} [N_{t+s}] = \sum_{i=1}^t F_{s+i} (1-F_{s+i})$$

The mean shortfall at time $t+s$ is then

$$\begin{aligned}\mathbb{E}[SF_{t+s}] &= \mathbb{E}[N_{t+s}]\mathbb{E}[D] \\ &= \left(\sum_{i=1}^t (1-F_{s+i}) + s \right) \mu_D \\ &= s\mu_D + \mu_D \sum_{i=1}^t (1-F_{s+i})\end{aligned}$$

The variance of the shortfall at time $t+s$ is then

$$\begin{aligned}\text{Var} [SF_{t+s}] &= \mu_N \sigma_D^2 + \mu_D^2 \sigma_N^2 \\ &= \left(\sum_{i=1}^t (1-F_{s+i}) + s \right) \sigma_D^2 + \mu_D^2 \sum_{i=1}^t F_{s+i} (1-F_{s+i})\end{aligned}$$

Up to now we have calculated the mean and variance of shortfall conditioned on the assumption that the excursion happens at time t . Summing over all possible time of excursion occurrence, we can get the expected inventory shortfall at the time of recovery. Let SF_R denote the inventory shortfall at the time of excursion recovery. Let T be the time of excursion occurrence and G be the cumulative distribution of the excursion occurrence time distribution. The expected shortfall at the time of excursion recovery is

$$\begin{aligned}
\mathbb{E}[SF_R] &= \mathbb{E}[\mathbb{E}[SF_R|T=t]] \\
&= \mathbb{E}\left[\left(\sum_{i=1}^T (1-F_{s+i})+s\right)\mu_D\right] \\
&= \sum_{t=0}^{\infty} P(T=t) \left(\sum_{i=1}^t (1-F_{s+i})+s\right)\mu_D \\
&= s\mu_D + \mu_D \sum_{t=0}^{\infty} \sum_{i=1}^t (1-F_{s+i})P(T=t) \\
&= s\mu_D + \mu_D \sum_{i=1}^{\infty} \sum_{t=i}^{\infty} (1-F_{s+i})P(T=t) \\
&= s\mu_D + \mu_D \sum_{i=1}^{\infty} (1-F_{s+i}) \sum_{t=i}^{\infty} P(T=t) \\
&= s\mu_D + \mu_D \sum_{i=1}^{\infty} (1-F_{s+i})(1-G_{i-1})
\end{aligned}$$

Using the conditional variance formula, we can calculate the variance of the inventory shortfall at the time of recovery

$$\begin{aligned}
Var [SF_R] &= Var (\mathbb{E}[SF_R|T]) + \mathbb{E}[Var (SF_R|T)] \\
&= Var (s\mu_D + \mu_D \sum_{i=1}^T (1-F_{s+i})) + \mathbb{E}\left[\left(\sum_{i=1}^T (1-F_{s+i})+s\right)\sigma_D^2 + \mu_D^2 \sum_{i=1}^T F_{s+i}(1-F_{s+i})\right] \\
&= \mu_D^2 Var \left(\sum_{i=1}^T (1-F_{s+i})\right) + \mathbb{E}\left[\left(\sum_{i=1}^T (1-F_{s+i})+s\right)\sigma_D^2 + \mu_D^2 \sum_{i=1}^T F_{s+i}(1-F_{s+i})\right] \\
&= (1) + (2)
\end{aligned}$$

The second term in above equation can be simplified as follows

$$\begin{aligned}
(2) &= \mathbb{E}\left[\left(\sum_{i=1}^T (1-F_{s+i})+s\right)\sigma_D^2 + \mu_D^2 \sum_{i=1}^T F_{s+i}(1-F_{s+i})\right] \\
&= s\sigma_D^2 + \sigma_D^2 \sum_{t=1}^{\infty} P(T=t) \sum_{i=1}^t (1-F_{s+i}) + \mu_D^2 \sum_{t=1}^{\infty} P(T=t) \sum_{i=1}^t F_{s+i}(1-F_{s+i})
\end{aligned}$$

Notice

$$\begin{aligned}
\sum_{t=1}^{\infty} P(T=t) \sum_{i=1}^t (1-F_{s+i}) &= \sum_{t=1}^{\infty} \sum_{i=1}^t P(T=t) (1-F_{s+i}) \\
&= \sum_{i=1}^{\infty} (1-F_{s+i}) \sum_{t=i}^{\infty} P(T=t) \\
&= \sum_{i=1}^{\infty} (1-F_{s+i}) (1-G_{i-1})
\end{aligned}$$

And similarly

$$\sum_{t=1}^{\infty} P(T=t) \sum_{i=1}^t F_{s+i}(1-F_{s+i}) = \sum_{i=1}^{\infty} F_{s+i}(1-F_{s+i})(1-G_{i-1})$$

So

$$(2) = s\sigma_D^2 + \sigma_D^2 \sum_{i=1}^{\infty} (1-F_{s+i}) (1-G_{i-1}) + \mu_D^2 \sum_{i=1}^{\infty} F_{s+i}(1-F_{s+i})(1-G_{i-1})$$

Then we get a simplified form of $Var [SF_R]$

$$\begin{aligned}
Var [SF_R] &= \mu_D^2 Var \left(\sum_{i=1}^T (1-F_{s+i}) \right) + r\sigma_D^2 + \sigma_D^2 \sum_{i=1}^{\infty} (1-F_{s+i})(1-G_{i-1}) \\
&+ \mu_D^2 \sum_{i=1}^{\infty} F_{s+i}(1-F_{s+i})(1-G_{i-1}) \\
&= \mu_D^2 \left(Var \left(\sum_{i=1}^T (1-F_{s+i}) \right) + \sum_{i=1}^{\infty} F_{s+i}(1-F_{s+i})(1-G_{i-1}) \right) \\
&+ \sigma_D^2 \left(s + \sum_{i=1}^{\infty} (1-F_{s+i})(1-G_{i-1}) \right)
\end{aligned}$$

If the lead time is geometrically distributed with parameter p and the excursion occurrence time is geometrically distributed with parameter q , then $F_i = 1 - (1-p)^i$ and $G_j = 1 - (1-q)^j$. Then mean shortfall at the time of excursion recovery is

$$\begin{aligned}
\mathbb{E}[SF_R] &= s\mu_D + \mu_D \sum_{i=1}^{\infty} (1-p)^{s+i} (1-q)^{i-1} \\
&= s\mu_D + \mu_D \frac{(1-p)^{s+1}}{1 - (1-p)(1-q)}
\end{aligned}$$

The variance of shortfall at the time of excursion recovery is

$$\begin{aligned}
\text{Var} \left(\sum_{i=1}^T (1-F_{s+i}) \right) &= \text{Var} \left(\sum_{i=1}^T (1-p)^{s+i} \right) \\
&= \text{Var} \left(\frac{(1-p)^{s+1}(1-p)^T}{1-(1-p)} \right) \\
&= \frac{(1-p)^{2(s+1)}}{p^2} \text{Var} \left((1-p)^T \right)
\end{aligned}$$

To calculate $\text{Var} \left((1-p)^T \right)$, first we need to calculate $\mathbb{E}[(1-p)^T]$ and $\mathbb{E}[(1-p)^{2T}]$.

$$\begin{aligned}
\mathbb{E}[(1-p)^T] &= \sum_{t=1}^{\infty} P(T=t)(1-p)^t \\
&= \sum_{t=1}^{\infty} q(1-q)^{t-1}(1-p)^t \\
&= \frac{q}{1-q} \sum_{t=1}^{\infty} ((1-p)(1-q))^t \\
&= \frac{q(1-p)}{1-(1-p)(1-q)}
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}[(1-p)^{2T}] &= \sum_{t=1}^{\infty} P(T=t)(1-p)^{2t} \\
&= \sum_{t=1}^{\infty} q(1-q)^{t-1}(1-p)^{2t} \\
&= \frac{q}{1-q} \sum_{t=1}^{\infty} (1-q)^t(1-p)^{2t} \\
&= \frac{q(1-p)^2}{1-(1-p)^2(1-q)}
\end{aligned}$$

Then we can calculate $\text{Var} \left((1-p)^T \right)$

$$\begin{aligned}
\text{Var} \left[(1-p)^T \right] &= \mathbb{E}[(1-p)^{2T}] - (\mathbb{E}[(1-p)^T])^2 \\
&= \frac{q(1-p)^2}{1-(1-p)^2(1-q)} - \left(\frac{q(1-p)}{1-(1-p)(1-q)} \right)^2 \\
&= \frac{p^2(1-p)^2q(1-q)}{(1-(1-p)^2(1-q))(1-(1-p)(1-q))^2}
\end{aligned}$$

Thus

$$\begin{aligned} Var \left[\sum_{i=1}^T (1-F_{s+i}) \right] &= \frac{(1-p)^{2(s+1)}}{p^2} \frac{p^2(1-p)^2 q(1-q)}{\left(1 - (1-p)^2(1-q)\right) \left(1 - (1-p)(1-q)\right)^2} \\ &= \frac{(1-p)^{2s+4} q(1-q)}{\left(1 - (1-p)^2(1-q)\right) \left(1 - (1-p)(1-q)\right)^2} \end{aligned}$$

Then in the case of a geometric lead time distribution and a geometric excursion time distribution, the variance of inventory shortfall can be simplified as

$$\begin{aligned} Var [SF_R] &= \mu_D^2 \left(Var \left(\sum_{i=1}^T (1-F_{s+i}) \right) + \sum_{i=1}^{\infty} F_{s+i}(1-F_{s+i})(1-G_{i-1}) \right) \\ &+ \sigma_D^2 \left(r + \sum_{i=1}^{\infty} (1-F_{s+i})(1-G_{i-1}) \right) \\ &= \mu_D^2 \left(\frac{(1-p)^{2s+4} q(1-q)}{\left(1 - (1-p)^2(1-q)\right) \left(1 - (1-p)(1-q)\right)^2} + \frac{(1-p)^{s+1}}{1 - (1-p)(1-q)} - \frac{(1-p)^{2s+2}}{1 - (1-p)^2(1-q)} \right) \\ &+ \sigma_D^2 \left(s + \frac{(1-p)^{s+1}}{1 - (1-p)(1-q)} \right) \end{aligned}$$

To plan safety stock level under excursion risk and stochastic lead time, we need stock to cover two parts of risk: shortfall when excursion happens and shortfall due to lead time.

When there is no excursion, from Section 3.1.4 we know that required stock level is

$$S_{ne} = \mu_D \mu_L + k \sqrt{\mu_D^2 \sum_{i=1}^{\infty} F_i(1-F_i) + \mu_L \sigma_D^2}$$

And the required stock level at the time of excursion recovery is

$$\begin{aligned} S_e &= \mathbb{E}[SF_R] + k \sqrt{Var [SF_R]} \\ &= s\mu_D + \mu_D \frac{(1-p)^{s+1}}{1 - (1-p)(1-q)} + k \sqrt{Var [SF_R]} \end{aligned}$$

where

$$\begin{aligned}
Var [SF_R] &= \mu_D^2 \left(Var \left(\sum_{i=1}^T (1-F_{s+i}) \right) + \sum_{i=1}^{\infty} F_{s+i} (1-F_{s+i})(1-G_{i-1}) \right) \\
&+ \sigma_D^2 \left(r + \sum_{i=1}^{\infty} (1-F_{s+i})(1-G_{i-1}) \right) \\
&= \mu_D^2 \left(\frac{(1-p)^{2s+4} q (1-q)}{(1-(1-p)^2(1-q))(1-(1-p)(1-q))^2} + \frac{(1-p)^{s+1}}{1-(1-p)(1-q)} - \frac{(1-p)^{2s+2}}{1-(1-p)^2(1-q)} \right) \\
&+ \sigma_D^2 \left(s + \frac{(1-p)^{s+1}}{1-(1-p)(1-q)} \right)
\end{aligned}$$

The overall required stock level is

$$S = \max(S_{ne}, S_e)$$

3.2.2 Multiple Excursion Types

A biopharmaceutical supply chain is usually exposed to more than one type of excursion risk. For example, a facility is exposed to material contamination risk and earthquake risk at the same time. To better analyze the potential excursion risk, supply chain and quality control managers classify the potential risks according to their severity and the risk source. Risks can be categorized into multiple levels from the most severe one to the least significant. For example, earthquake risk may place the whole manufacturing site in danger and suspend the production for a long period of time. On the other hand, contamination risk usually happens within a facility and is much easier to be isolated. Based on the source of risk, excursions can be categorized into environment, equipment, procedure deviations validation related etc.

Although sometimes different risks can be correlated, usually significant risks in a biopharmaceutical supply chain are independent of each other. In practice, it is a valid assumption that earthquake risk is independent of contamination risk. Such independence among different risk gives us an opportunity to do risk pooling and save on required stock holding to hedge against stockouts.

In this section we extend the model in section 5.1 by considering the case when there are multiple independent excursion risks. For representational simplicity, we assume all excursion times are geometrically distributed. Let $T_j, j=1, \dots, J$ denote the time excursion j happens and assume T_j is geometrically distributed with parameter q_j . Let s_j be the time to recovery for excursion j . We also assume that the excursion parameters q_j are small. The possibility of two overlapping excursions is very small and can be ignored. Let T be the time of first excursion in our horizon i.e., $T = \min_j T_j$.

First we show that T is approximately geometrically distributed with parameter $\sum_{j=1}^J q_j$. We can calculate the cumulative probability of T as

$$\begin{aligned}
P(T \leq t) &= P(\min(T_1, \dots, T_J) \leq t) \\
&= 1 - P(\min(T_1, \dots, T_J) > t) \\
&= 1 - P(T_1 > t, \dots, T_J > t) \\
&= 1 - \prod_{j=1}^J P(T_j > t) \\
&= 1 - \prod_{j=1}^J (1 - q_j)^t
\end{aligned}$$

Since q_j is small, we have $(1 - q_j)^t \approx 1 - tq_j$. Then

$$\begin{aligned}
P(T \leq t) &\approx 1 - \prod_{j=1}^J (1 - tq_j) \\
&\approx 1 - \left(1 - t \sum_{j=1}^J q_j\right) \\
&\approx 1 - \left(1 - \sum_{j=1}^J q_j\right)^t
\end{aligned}$$

Comparing to the cumulative distribution of Geometric random variable $P(T_j \leq t) = 1 - (1 - q_j)^t$, we conclude that T is approximately Geometric with parameter $\sum_{j=1}^J q_j$.

As a geometric distribution is the discrete counterpart of an exponential distribution, we can calculate the probability that excursion j happens first is $P(T = T_j) = \frac{q_j}{\sum_{j=1}^J q_j}$. Then under the assumption that multiple excursions may happen but not at the same time, the mean inventory shortfall at the time of recovery from excursion when lead time has a geometric distribution with parameter p is given by

$$\begin{aligned}
\mathbb{E}[SF] &= \sum_{j=1}^J \frac{q_j}{q} (s_j \mu_D + \mu_D \frac{(1-p)^{s_j+1}}{1 - (1-p)(1-q_j)}) \\
&= \frac{\mu_D}{q} \sum_{j=1}^J s_j q_j + \frac{\mu_D}{q} \sum_{j=1}^J \frac{q_j (1-p)^{s_j+1}}{1 - (1-p)(1-q_j)}
\end{aligned}$$

Note that we let $q = \sum_{j=1}^J q_j$.

And the variance is

$$\begin{aligned}
Var [SF] &= Var \left[\sum_{j=1}^J \frac{q_j}{q} SF_j \right] \\
&= \sum_{j=1}^J \left(\frac{q_j}{q} \right)^2 Var [SF_j]
\end{aligned}$$

where $Var [SF_j]$ is as follows:

$$Var [SF_j] = \mu_D^2 \left(\frac{(1-p)^{2s_j+5}}{p^2(p^2-3p+3)(p-2)^2} + \frac{(1-p)^{s_j+1}}{1-(1-p)(1-q_j)} + \frac{(1-p)^{2s_j+2}}{1-(1-p)^2(1-q_j)} \right) + \sigma_D^2 \left(s_j + \frac{(1-p)^{s_j+1}}{1-(1-p)(1-q_j)} \right)$$

3.2.3 Numerical Example

Again we use a typical purification stage of the biopharmaceutical supply chain as our numerical example. We assume the excursion has frequency of occurrence once of every 5 years and the duration of excursion (from occurrence to recovery) is 2 months. We start purification every 3 days and the lead time is geometrically distributed with mean of 56 days. Daily demand for the purification is 8 mmu and standard deviation is 1.6 mmu. And we choose $k = 3$ as the safety factor. We can use the formula in Section 3.2.1 to calculate the mean and variance of the inventory shortfall at the time of excursion recovery.

$$\begin{aligned} \mathbb{E}[SF_R] &= s\mu_D + \mu_D \sum_{i=1}^{\infty} (1-F_{s+i})(1-G_{i-1}) \\ &= 213 \end{aligned} \quad (3.5)$$

$$\begin{aligned} Var [SF_R] &= \mu_D^2 \left(\frac{(1-p)^{2s+4}q(1-q)}{(1-(1-p)^2(1-q))(1-(1-p)(1-q))^2} + \frac{(1-p)^{s+1}}{1-(1-p)(1-q)} - \frac{(1-p)^{2s+2}}{1-(1-p)^2(1-q)} \right) \\ &+ \sigma_D^2 \left(s + \frac{(1-p)^{s+1}}{1-(1-p)(1-q)} \right) \\ &= 298 \end{aligned}$$

When there is no excursion, from Section 3.1.4 we know that required stock level is

$$\begin{aligned} S_{ne} &= \mu_D \mu_L + k \sqrt{\mu_D^2 \sum_{i=1}^{\infty} F_i(1-F_i) + \mu_L \sigma_D^2} \\ &= 664 \end{aligned}$$

And the required stock level at the time of excursion recovery is

$$\begin{aligned} S_e &= \mathbb{E}[SF_R] + k \sqrt{Var [SF_R]} \\ &= 1106 \end{aligned}$$

Parameter	Value	Description
μ_D	10	Mean Demand
σ_D	2	Demand Std
μ_L	30	Mean Lead Time
H	750	Planning Horizon
T	5	Ordering Interval
y	5	Excursion Frequency (Years)
s	60	Excursion Duration (Days)

Table 3.3: Simulation Parameters under Stochastic Lead Time and Excursion Risk

The overall required stock level is

$$\begin{aligned} S &= \max(S_{ne}, S_e) \\ &= 1106 \end{aligned}$$

So the required stock level to hedge against stochastic lead time risk and excursion risk is 1106 mmu. When there is no excursion risk, the required stock level is 663 mmu. Note that when we apply the lead-time-demand-based approach without considering any order crossover, the required stock level is 1793 mmu. This indicates that lead time demand analysis leads to excess inventory holding even comparable to the case when there is excursion risk taken into account.

3.2.4 Simulation

In this section we use Monte Carlo simulation to study the performance of a base stock system under the the safety stock policy derived from crossover analysis. The supply chain system is exposed to stochastic lead time risk and excursion risk. We assume per period demand is normally distributed with mean 10 and standard deviation 2. The ordering lead time has a geometrically distribution with mean of 30 days. Simulation horizon is 750 days. An order is place every 5 days. The size of each order is 50. Also assume there is no planned ordering lead time. For the excursion parameters, we assume the excursion frequency is once every 5 years and the excursion duration is 60 days. Table 3.1.7 summarizes the simulation parameters.

Given the parameters in Table 3.3, we can calculate the required stock level as a function of safety factor k . Using formula from Section 3.2.1, we get

$$\mathbb{E}[SF_R] = 220$$

In practice, to ensure we have enough stock when an excursion hits, we can enforce $\mathbb{E}[SF_R] \geq s\mu_D$. We want to make sure we have enough inventory to last the period of excursion duration. In this case we want to have $\mathbb{E}[SF_R] \geq 600$.

$$\text{Var}(\text{SF}_R) = 307$$

Then the required stock level as a function of the safety factor k is

$$\begin{aligned} S(k) &= \mathbb{E}(\text{SF}_R) + k\sqrt{\text{Var}(\text{SF}_R)} \\ &= 600 + 307k \end{aligned}$$

To evaluate the performance of our stock policy, at the beginning of the simulation we set the inventory level to $S(k)$ given a choice of the safety factor k . Then first we simulate demand quantity day by day. Every T days an order is placed and the actual lead time for this order is randomly generated from the lead time distribution. At the end of each day we calculate cumulative supply and cumulative demand. Cumulative supply is the total quantity arrived so far. The quantity depends on the actual lead time of each order. Also we add up demand to date to get cumulative demand.

To simulate excursion, at the beginning of each iteration, first we sample a time to excursion from the excursion distribution. If this time to excursion is longer than our planning horizon, then all calculation regarding cumulative supply is the same as before. But if the excursion happens within our planning horizon, all the orders placed during the excursion duration are considered contaminated and thrown away. Since sometimes excursion recovery time could be longer than our planning horizon, in this case our planning horizon is extended until excursion recovery.

When we reach the end of the horizon, we collect the simulation statistics for this simulation instance. Two key statistics are service level and average inventory. The service level is the percentage of periods that cumulative supply is above cumulative demand. In biopharmaceutical supply chain planning, the primary goal is to make sure that the service level is very close to 100%, for example 99.9%. In practice it's up to the supply chain manager to set a service level target given the inventory cost constraints. The main metric for the overall performance of a supply chain is the average inventory. It is the average inventory position over the planning horizon. Average inventory holding multiplied by unit value gives us the total cost for inventory holding. The unit value may vary depending on the specific stage of the product. Usually for biopharmaceuticals the closer to the final market, the more expensive it is. Ideally we want 100% service while maintaining minimum average inventory.

After collecting simulation statistics, we calculate the mean and standard deviation of service level and also the mean and standard deviation of inventory. Figure 3.11 shows the inventory distribution over the planning horizon for a typical iteration. We can see that there is no stock out in the planning horizon. An Excursion occurs on day 126. The excursion causes the big continuing drop of inventory from day 150 until 250. The lowest stock level in the horizon is around 100 while at the end of the horizon the holding is around 400.

From the simulation result we can see that setting a service factor $k=2$ gives us more than 100% service level and an average inventory holding of 656.9. On the other hand,

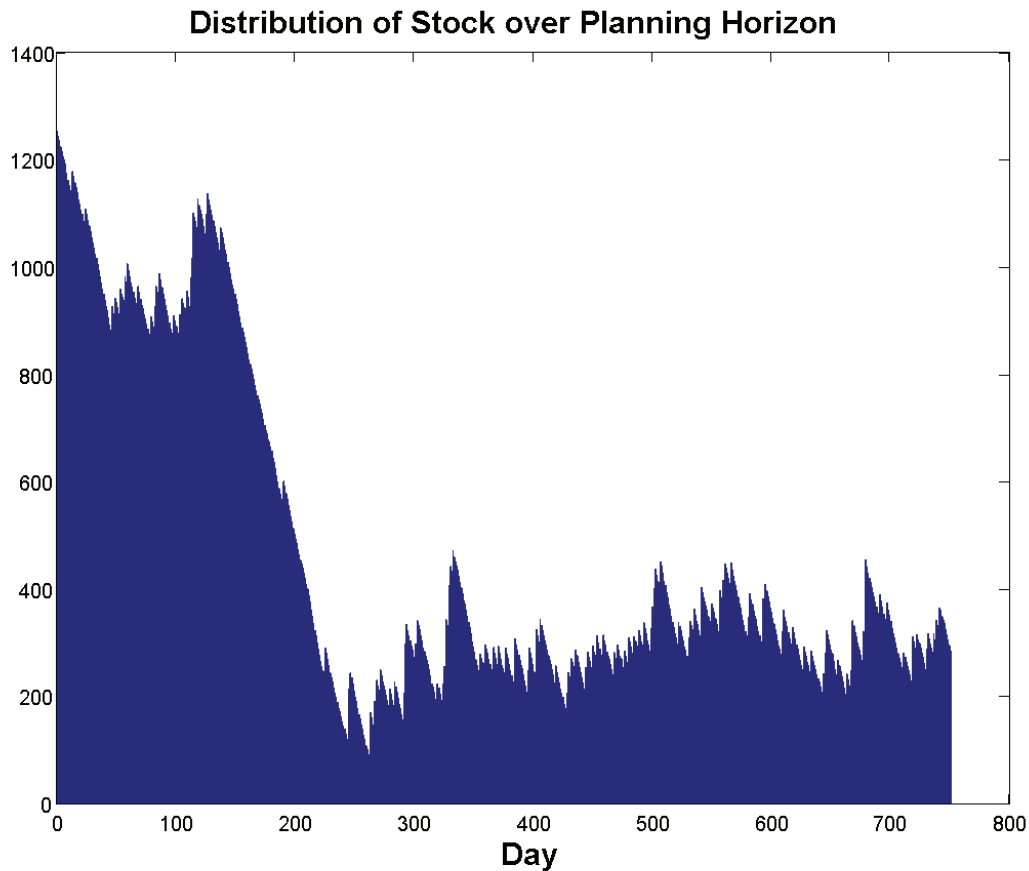


Figure 3.11: Stock Distribution in the Planning Horizon

$k=0$ gives us only around 50% service level but with much less inventory holding. With simulation we can also collect statistics regarding the distribution of the outstanding stock level and shortage quantity. Figure 3.12 is the distribution of the number of outstanding orders. Figure 3.13 shows the histogram of the shortage quantity. For both results the number of iterations is 4,000.

In biopharmaceutical supply chains, an important factor of performance measure is average inventory holding. Figure 3.14 summarizes the simulation results by comparing average inventory holding and service level. We can see that service level is an increasing concave function of average inventory holding.

	k = 0		k = 0.5		k = 1		k=1.5		k=2	
B	600									
SS	0		153.6		307.2		460.8		614.4	
	Mean	Stde	Mean	Stde	Mean	Stde	Mean	Stde	Mean	Stde
Service Level	53.1%	0.8%	59.9%	0.7%	83.8%	0.4%	98.8%	0.1%	100.0%	0.0%
Inv Holding	172.2	2.6	261.5	3.8	360.8	4.5	510.8	4.8	656.9	4.6

Table 3.4: Simulation Results under Stochastic Lead Time Risk

3.3 Safety Stock under Stochastic Lead Time and Batch Rejection Risk

In section 4 we discussed the calculation of the required stock level to hedge against stochastic lead time risk and excursion risk. In this section will consider a supply chain system under batch rejection risk. Rejection is a common risk factor in supply chain management. In biopharmaceutical supply chains, since the manufacturing is batch based, the rejection is usually in terms of an entire batch. The most common reason for batch rejection is quality assurance failure. Because of the high value of each batch, in order to recover the most value from each batch, partial batch reject does occur in the quality assurance process.

Compared to traditional pharmaceutical manufacturing, the biopharmaceutical manufacturing process is less well understood and the exact root cause of batch rejection is unknown. Thus each batch rejection is considered independent of others unless there is a major excursion causing all batches to be rejected. Thus in our model we will assume each batch has independent reject rate. And batch rejection is usually not constant over time. This could be due to the learning curve of the biological process. Also any adoption of new technique may introduce increased reject rate which may reduce over time. In a typical biopharmaceutical manufacturing process, the reject rate may vary from 20% to less than 2%.

When there is batch rejection risk, each batch start has a potential risk of being rejected. In this case, even if each batch arrived perfectly on time; the cumulative supply could still fall below cumulative demand due to random batch rejection. And the gap between the two increases over time. So in our batch reject related model, we will introduce a planning horizon and our required stock level will be a function of our planning horizon. This planning horizon can be considered as the safety stock replenishment cycle and the required stock level can be selected to guarantee that we won't stock out in the planning horizon.

3.3.1 Model

In this section we will derive the safety stock level in a single stage system under stochastic lead time and random batch rejection. We assume that we have i.i.d. demand per period

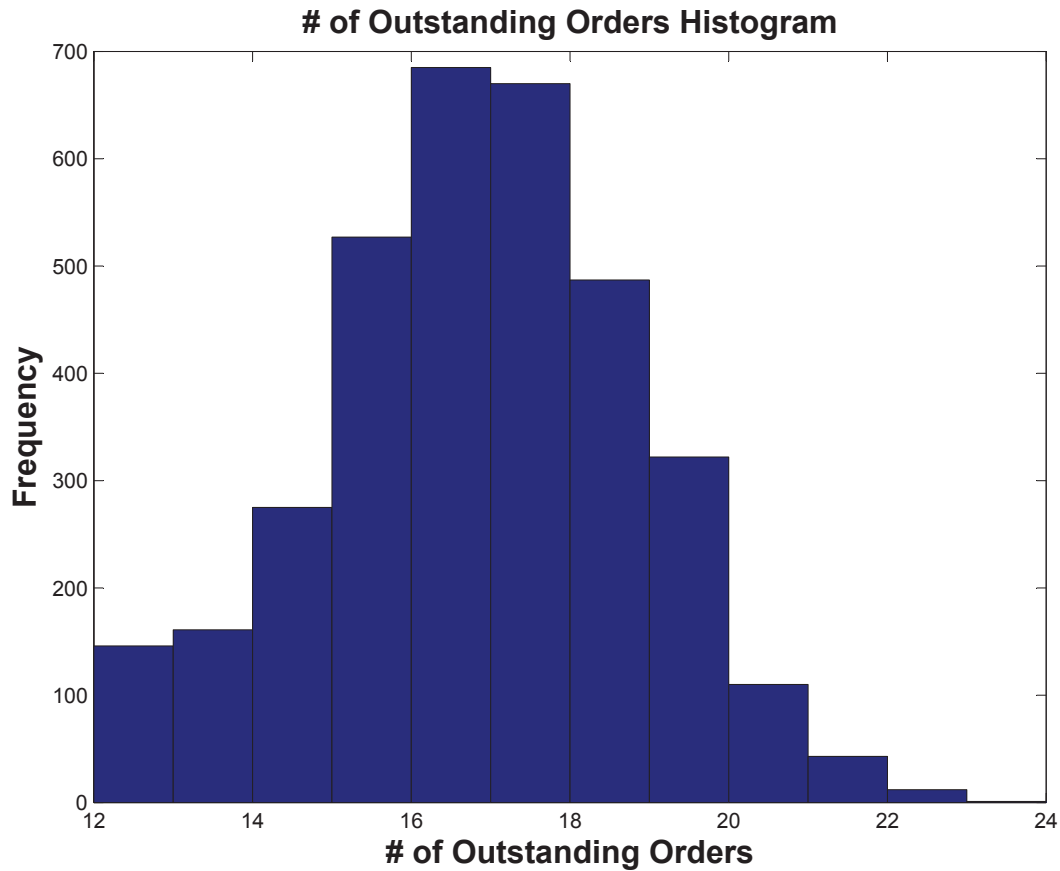


Figure 3.12: Distribution of # of Outstanding Orders

with mean μ_D and standard deviation σ_D . An order of one batch is placed every T periods to meet the demand. The ordering batch size is the sum of the previous T periods' total demand. Each order takes a stochastic lead time to arrive and then meets the demand. And each order also has probability r of being rejected and the batch rejection is i.i.d. We further assume that batch rejection is independent of the ordering lead time.

We use the following notion in this section.

N – Number of outstanding orders. An outstanding order is defined as an order placed but not arrived as planned

SF – Shortfall. Shortfall is defined as the gap between cumulative supply and cumulative demand

F_i – Probability that order lead time is less than i days, $F_i = P(L \leq i)$

T – Production order interval

L – Order lead time

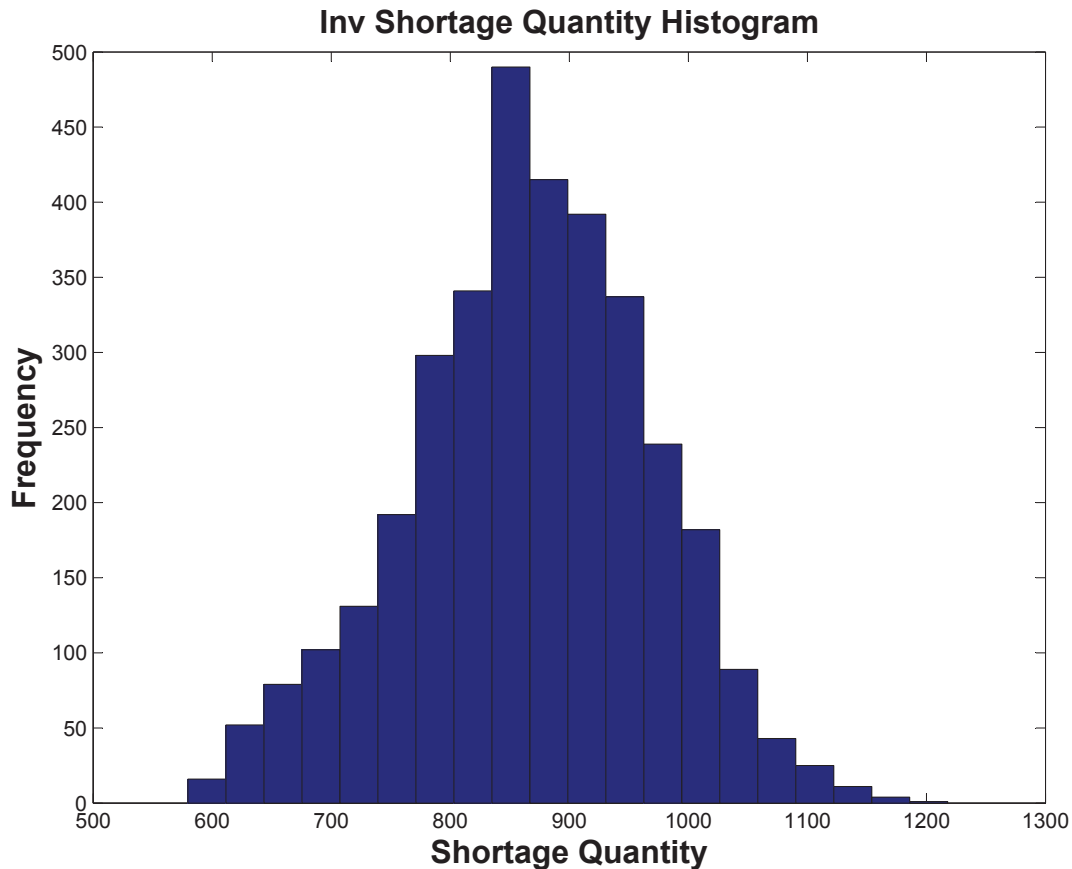


Figure 3.13: Distribution of Shortage Quantity

μ_L – Average order lead time

σ_L – Standard deviation of order lead time

D_i – Order quantity in i^{th} period

μ_D – Average order quantity in one period

σ_D – Standard deviation of order quantity in one period

R_i – Equal to 1 if order placed at time i is rejected, 0 otherwise

r – Probability of order being rejected

H – Planning horizon, some multiple of ordering interval

l – Planned lead time, some multiple of ordering interval

Given stochastic lead time and batch rejection risk, the number of outstanding orders now has two parts: orders arrived later than planned less orders arrived earlier than planned lead time. Order arrival later than planned lead time include those orders with late arrival

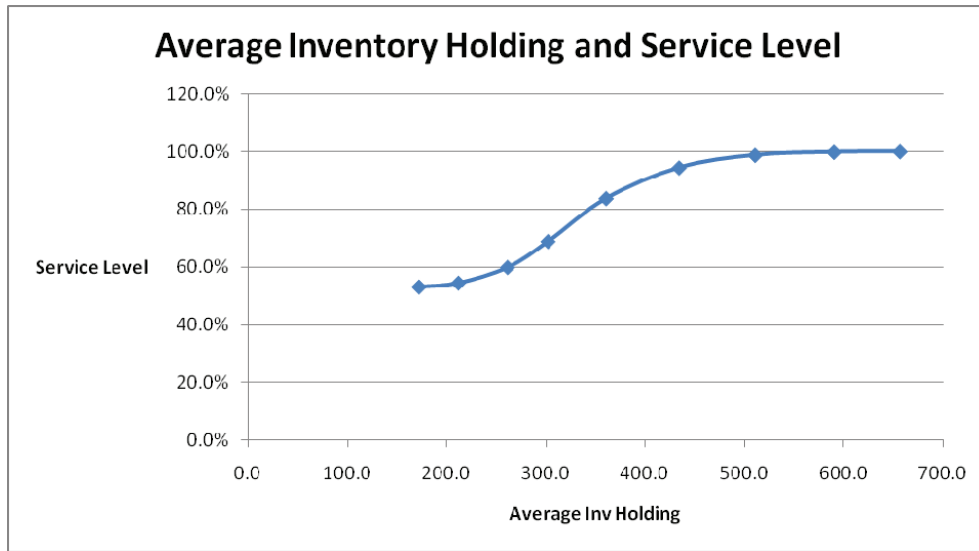


Figure 3.14: Average Inventory Holding and Service Level under Stochastic Lead Time and Excursion Risk

as well as earlier arrival but rejected. Orders arriving earlier than the planned lead time are those arriving early and not rejected. The number of outstanding orders at the end of our planning horizon is

$$\begin{aligned}
 N &= - \sum_{i=1}^l \delta(L_{t-iT} \leq iT) (1 - \delta(R_{iT})) + \sum_{i=l+1}^H \delta(L_{t-iT} > iT) + \sum_{i=l+1}^H \delta(L_{t-iT} \leq iT) \delta(R_{iT}) \\
 &= - \sum_{i=1}^l \delta(L_{t-iT} \leq iT) (1 - \delta(R_{iT})) + H - l + \sum_{i=l+1}^H (\delta(R_{iT}) - 1) \delta(L_{t-iT} \leq iT) \\
 &= H - l - \int_{i=1}^H \delta(L_{t-iT} \leq iT) (1 - \delta(R_{iT}))
 \end{aligned}$$

To interpret the above formula, we have a total of H orders in our planning horizon. The planned lead time reduces the number of outstanding orders by 1. Also those arrivals within our planning horizon also decrease the total number of outstanding orders.

$$\begin{aligned}
 \mathbb{E}[N] &= H - l - E\left[\sum_{i=1}^H \delta(L_{t-iT} \leq iT) (1 - \delta(R_{iT}))\right] \\
 &= H - l - \sum_{i=1}^H P(L_{t-iT} \leq iT) (1 - P(R_{iT})) \\
 &= H - l - (1 - r) \sum_{i=1}^H F_{iT}
 \end{aligned}$$

To calculate the variance of the number of outstanding orders, we need following formula.

Given X is independent of Y , we have

$$\begin{aligned}
\text{Var}[XY] &= E[X^2Y^2] - (E[XY])^2 \\
&= E[X^2] E[Y^2] - (E[X] E[Y])^2 \\
&= (\text{Var}[X] + (E[X])^2)(\text{Var}[Y] + (E[Y])^2) - (E[X])^2(E[Y])^2 \\
&= \text{Var}[X] \text{Var}[Y] + \text{Var}[X] (E[Y])^2 + (E[X])^2 \text{Var}[Y]
\end{aligned}$$

First we calculate

$$\begin{aligned}
\text{Var}[\delta(L_{t-iT} \leq iT)(1 - \delta(R_{iT}))] &= \text{Var}[\delta(L_{t-iT} \leq iT)] \text{Var}[1 - \delta(R_{iT})] \\
&\quad + \text{Var}[\delta(L_{t-iT} \leq iT)] (E[1 - \delta(R_{iT})])^2 \\
&\quad + (E[\delta(L_{t-iT} \leq iT)])^2 \text{Var}[1 - \delta(R_{iT})] \\
&= F_{iT}(1 - F_{iT})r(1 - r) + F_{iT}(1 - F_{iT})(1 - r)^2 + F_{iT}^2r(1 - r) \\
&= F_{iT}(1 - r)(1 - (1 - r)F_{iT})
\end{aligned}$$

Then the variance of the number of outstanding orders is

$$\begin{aligned}
\text{Var}[N] &= \text{Var}\left[\sum_{i=1}^H \delta(L_{t-iT} \leq iT)(1 - \delta(R_{iT}))\right] \\
&= \sum_{i=1}^H F_{iT}(1 - r)(1 - (1 - r)F_{iT})
\end{aligned}$$

Next let's calculate the mean and variance of the inventory shortfall quantity by conditioning on the number of outstanding orders:

$$\begin{aligned}
\mathbb{E}(SF) &= \mathbb{E}[\mathbb{E}[SF_t|N]] \\
&= \mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^N D_T\right]\right] \\
&= \mathbb{E}[N\mathbb{E}[D_T]] \\
&= \mathbb{E}[NT\mu_D] \\
&= T\mu_D\mu_N \\
&= T\mu_D(H - l - \sum_{i=1}^H F_{iT}(1 - r))
\end{aligned}$$

The variance of SF_t is

$$\begin{aligned}
Var(SF) &= Var[\mathbb{E}[SF_t|N]] + \mathbb{E}[Var[SF_t|N]] \\
&= Var\left[\mathbb{E}\left[\sum_{i=1}^N D_T\right]\right] + \mathbb{E}\left[Var\left[\sum_{i=1}^N D_T\right]\right] \\
&= Var[NT\mu_D] + \mathbb{E}[NT\sigma_D^2] \\
&= \mu_D^2 T^2 \sigma_N^2 + \mu_N T \sigma_D^2 \\
&= \mu_D^2 T^2 \left(\sum_{i=1}^H F_{iT}(1-r)(1-(1-r)F_{iT})\right) \\
&\quad + (H-l - \sum_{i=1}^H F_{iT}(1-r))T\sigma_D^2
\end{aligned}$$

The required stock holding S is the sum of the average shortfall and a multiple of the standard deviation of the shortfall. H is the stock replenishment cycle in our model. At the beginning of our replenishment cycle, we need to maintain inventory holding of S to hedge against risks during the replenishment cycle H and reach target service level. Within this replenishment cycle H , our stocking holding on average goes down by the amount of expected shortfall. Let S_A be the average stock holding during the replenishment cycle. Then we have

$$S_A = \frac{1}{2}\mathbb{E}(SF) + k\sqrt{Var(SF)}$$

When the lead time distribution is Geometric with parameter p , we have

$$F_i = 1 - (1-p)^i, \quad i = 1, 2, \dots$$

Then the expected number of outstanding orders is:

$$\begin{aligned}
\mathbb{E}[N] &= H - l - \sum_{i=1}^H F_{iT}(1-r) \\
&= H - l - (1-r) \sum_{i=1}^H (1 - (1-p)^{iT}) \\
&= rH - l + (1-r) \frac{(1-p)^T (1 - (1-p)^{HT})}{1 - (1-p)^T}
\end{aligned}$$

The variance of the number of outstanding order is

$$\begin{aligned}
Var [N] &= \sum_{i=1}^H F_{iT}(1-r)(1-(1-r)F_{iT}) \\
&= \sum_{i=1}^H (1-(1-p)^{iT})(1-r) \left(1-(1-r)(1-(1-p)^{iT})\right) \\
&= H(r-r^2) + (1-3r+2r^2) \frac{(1-p)^T(1-(1-p)^{HT})}{1-(1-p)^T} - (1-r)^2 \frac{(1-p)^{2T}(1-(1-p)^{2HT})}{1-(1-p)^{2T}}
\end{aligned}$$

The mean and variance of the shortfall quantity are

$$\begin{aligned}
\mathbb{E}(SF) &= T\mu_D\mu_N \\
&= T\mu_D(rH-1+(1-r)\frac{(1-p)^T(1-(1-p)^{HT})}{1-(1-p)^T})
\end{aligned}$$

$$\begin{aligned}
Var(SF) &= \mu_D^2 T^2 \sigma_N^2 + \mu_N T \sigma_D^2 \\
&= \mu_D^2 T^2 \left(H(r-r^2) + (1-3r+2r^2) \frac{(1-p)^T(1-(1-p)^{HT})}{1-(1-p)^T} \right) \\
&\quad - \mu_D^2 T^2 \left((1-r)^2 \frac{(1-p)^{2T}(1-(1-p)^{2HT})}{1-(1-p)^{2T}} \right) \\
&\quad + \left(rH-1+(1-r)\frac{(1-p)^T(1-(1-p)^{HT})}{1-(1-p)^T} \right) T \sigma_D^2
\end{aligned}$$

3.3.2 Numerical Example

In this section we will discuss a practical numerical example to illustrate the calculation of the required stock level in a biopharmaceutical supply chain. For comparison purpose, we use the same numerical example as in Section 3.1.6 and Section 3.2.3.

We consider the purification stage in biopharmaceutical manufacturing as our numerical example. Ordering lead time is assumed to be geometrically distributed with mean 56 days. Daily demand for the purification stage is 8 mmu and standard deviation is 1.6 mmu. To make sure we have very close to 100% service level, safety factor is set to 3. And we place a purification order every 3 days. Each order has probability of 1% being rejected. Then

based on the formula in section 5.1 we can calculate the mean and standard deviation of the shortfall and also the required stock level as follows.

$$\begin{aligned}\mathbb{E}(SF_t) &= T\mu_D(rH - 1 + (1 - r)\frac{(1 - p)^T (1 - (1 - p)^{HT})}{1 - (1 - p)^T}) \\ &= 464\end{aligned}$$

$$\begin{aligned}Var(SF_t) &= \mu_D^2 T^2 \left(H(r - r^2) + (1 - 3r + 2r^2) \frac{(1 - p)^T (1 - (1 - p)^{HT})}{1 - (1 - p)^T} \right) \\ &\quad - \mu_D^2 T^2 \left((1 - r)^2 \frac{(1 - p)^{2T} (1 - (1 - p)^{2HT})}{1 - (1 - p)^{2T}} \right) \\ &\quad + \left(rH - 1 + (1 - r) \frac{(1 - p)^T (1 - (1 - p)^{HT})}{1 - (1 - p)^T} \right) T\sigma_D^2 \\ &= 6024\end{aligned}$$

$$\begin{aligned}S &= \mathbb{E}(SF_t) + k\sqrt{Var(SF_t)} \\ &= 464 + 78k\end{aligned}$$

Given the required stock at the beginning of our planning horizon, the average stock holding over the planning horizon is

$$\begin{aligned}S_A &= \frac{1}{2}\mathbb{E}(SF_t) + k\sqrt{Var(SF_t)} \\ &= 232 + 78k\end{aligned}$$

The required stock level is 697 mmu. Recall that the required stock level derived from lead time demand based approach is 1793 mmu. When there is no batch rejection, the required stock level derived from crossover analysis is 663 mmu. When there is excursion risk, the required stock level is 813 mmu.

3.3.3 Simulation

In this section we will use Monte Carlo simulation to evaluate the performance of the required stock level derived from our calculation. The key performance measure is service level throughout our planning horizon. Also we want to avoid excess stock holdings so we will look at average stock holding as well.

Parameter	Value	Description
μ_D	10	Mean Demand
σ_D	2	Demand Std
r	5%	Batch Reject Rate
μ_L	30	Mean Lead Time
H	750	Planning Horizon
T	5	Ordering Interval

Table 3.5: Simulation Parameters under Stochastic Lead Time and Batch Reject Risk

We assume per period demand is normally distributed with mean 10 and standard deviation 2. The ordering lead time is geometric distributed with mean of 30 days. The simulation horizon is 750 days. An order is placed every 5 days. The size of each order is 50. Also assume there is no planned ordering lead time. The batch reject rate is 5%. Table 3.5 summarizes the simulation parameters.

Given the parameters in Table 3.5, we can calculate the required stock level as a function of the safety factor k . Using formula from section 5.1, we get

$$\begin{aligned}\mathbb{E}(SF_t) &= T\mu_D(rH - 1 + (1 - r)) \frac{(1 - p)^T (1 - (1 - p)^{HT})}{1 - (1 - p)^T} \\ &= 642\end{aligned}$$

$$\begin{aligned}\text{Var}(SF_t) &= \mu_D^2 T^2 \left(H(r - r^2) + (1 - 3r + 2r^2) \frac{(1 - p)^T (1 - (1 - p)^{HT})}{1 - (1 - p)^T} \right) \\ &\quad - \mu_D^2 T^2 \left((1 - r)^2 \frac{(1 - p)^{2T} (1 - (1 - p)^{2HT})}{1 - (1 - p)^{2T}} \right) \\ &\quad + \left(rH - 1 + (1 - r) \frac{(1 - p)^T (1 - (1 - p)^{HT})}{1 - (1 - p)^T} \right) T\sigma_D^2 \\ &= 2425\end{aligned}$$

Then the required stock level as a function of safety factor k is

$$\begin{aligned}S &= \mathbb{E}(SF_t) + k\sqrt{\text{Var}(SF_t)} \\ &= 642 + 156k\end{aligned}$$

We follow the same simulation algorithm as in Section 3.2.3. To evaluate the performance of our stock policy, at the beginning of the simulation we set the inventory level to

	k = 0		k = 0.5		k = 1		k=1.5		k=2	
	B=641.6									
	SS = 0		SS = 77.9		SS = 155.7		SS = 233.6		SS = 311.5	
	Mean	Stde	Mean	Stde	Mean	Stde	Mean	Stde	Mean	Stde
Service Level	87.2%	0.7%	93.8%	0.5%	97.5%	0.3%	99.1%	0.2%	99.7%	0.1%
Inv Holding	215.1	3.1	281.2	3.3	356.3	3.5	434.9	3.8	508.1	3.9

Table 3.6: Simulation Results under Stochastic Lead Time and Batch Reject Risk

according a given choice of the safety factor. Then first we simulate demand quantity day by day. Every T days an order is placed and the actual lead time for this order is randomly generated from the lead time distribution. To simulation batch rejection, every time we place an order we generate a Bernoulli random variable to determine if the batch is being rejected or not. At the end of each day we calculate cumulative supply and cumulative demand. Cumulative supply is the total quantity arrived so far. The quantity depends on the actual lead time of each order. Also we add up demand to date to get cumulative demand. At the end of the simulation we evaluate the service level and average inventory holding at the end of our simulation. Ideally we want very close to 100% service level and low average inventory holding.

After collecting simulation statistics, we calculate the mean and standard deviation of service level and also the mean and standard deviation of inventory holding.

From the simulation result we can see that setting service factor to 2 gives us more than 99% service level and an average inventory holding of 508. On the other hand, zero service factor gives us only around 87% service level but with much less inventory holding. In biopharmaceutical supply chain, an important factor of performance measure is average inventory holding. Figure 3.18 summarizes the simulation results by comparing average inventory holding and service level. We can see that service level is an increasing concave function of average inventory holding.

3.4 Safety Stock under Stochastic Lead Time, Batch Rejection Risk and Excursion Risk

In this section we consider a single stage supply chain system exposed to stochastic lead time risk, batch rejection risk and excursion risk. In biopharmaceutical supply chains, the production ordering lead time is usually stochastic and subject to significant variation. In production, the reason for stochastic lead time is quality assurance uncertainty. In downstream stages of the supply chain, the uncertainty in lead time is due to shipping delays, paperwork handling and quality assurance as well. Batch rejection may occur as a result of quality assurance. As biological products, biopharmaceutical intermediate and final products are much less stable and more complex than traditional pharmaceuticals. To ensure safety

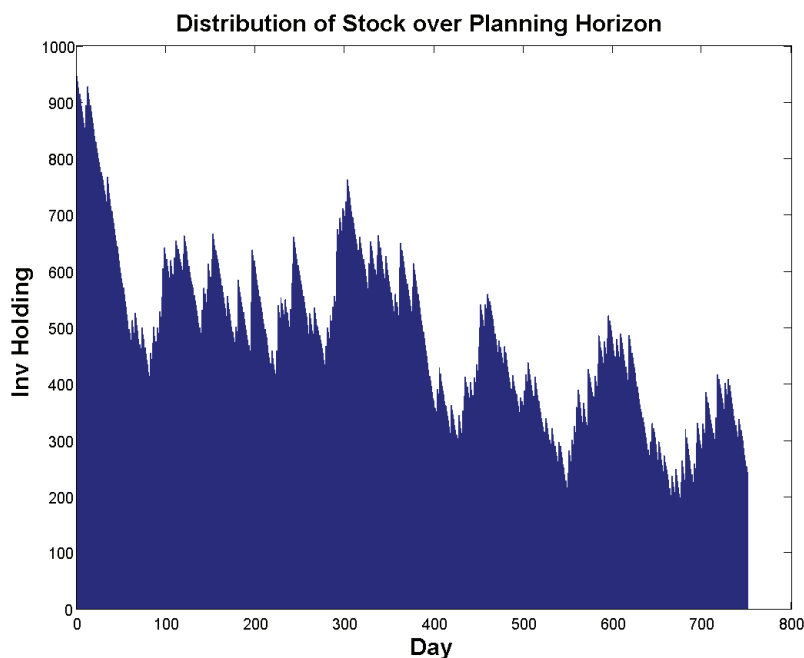


Figure 3.15: Stock Distribution in the Planning Horizon

of the product, batches are rejected when they fail to meet regulatory requirements or are suspected to be contaminated. Also as a relatively new technology, the biopharmaceutical manufacturing process is not well understood. It's not uncommon that some major incident happens and causes the a significant amount of batches being rejected due to contamination. In this paper we call this kind of supply chain disruption an excursion. Stochastic lead time risk, batch rejection risk and excursion risk are three major risks in biopharmaceutical supply chains and our goal is to derive the required stock holding to hedge against these three risks within our planning horizon.

To find out the necessary but not excess required stock holding, our general approach is to model the gap between cumulative supply and cumulative demand curve given there is zero stock holding at the beginning of our planning horizon. Then our calculated required stock holding we keep at the beginning of our planning horizon will move cumulative supply curve upward and make sure the cumulative supply curve is above cumulative demand curve close to 100% of the time.

Compared to the models in previous chapters, one thing we impose is the planning horizon. Our required stock holdings only prepares for possible risks within the planning horizon. For example, the required stock holding only covers possible batch rejection within the planning horizon. If we were preparing for all possible batch rejects in the infinite horizon, the required stock holding goes to infinity. We can also think of the planning horizon as our safety stock replenishment cycle. Our required stock holding tells us how much safety stock we need to avoid stock outs in the current cycle. At the beginning of next cycle, we need

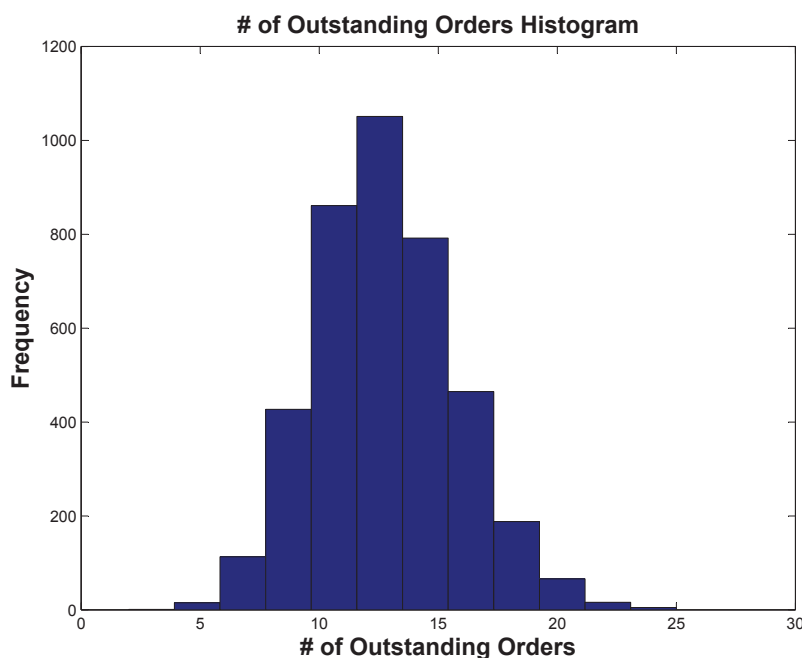


Figure 3.16: Distribution of # of Outstanding Orders

certain safety stock replenishment mechanism to raise the stock holding to that level again.

In the excursion analysis we use the expectation method to calculate the expected gap between cumulative supply and cumulative demand. When we impose a planning horizon, the required stock holding is a function of the probability the excursion happens in our planning horizon. Although this is theoretically correct, in practice it may lead to very small stock holding when the excursion probability is small. In this case, the derived model is unacceptable to supply chain managers due to the significant implied risk from the low stock holding level. For example, if we are calculating required stock level to hedge against earthquake risk, the once-in-a-hundred-year earthquake probability may give us extremely small required holding. But due to the significance of the potential damage of the earthquake to the supply chain, supply chain managers usually would like to hold much more stock than that. And the level of holding depends on the severity of such an excursion or the time taken to recover from excursion. So to make our model practically applicable, we will make sure that our stock holding is more than enough to cover the total demand from the excursion occurrence to the excursion recovery.

In this section first we derive the model to calculate the required stock level. Then we use a practical example to illustrate the application of the model in biopharmaceutical supply chain. Next we use Monte Carlo simulation to evaluate the performance of the required stock level. In the end we will discuss the impact of the model parameters.

We use following notion in this section.

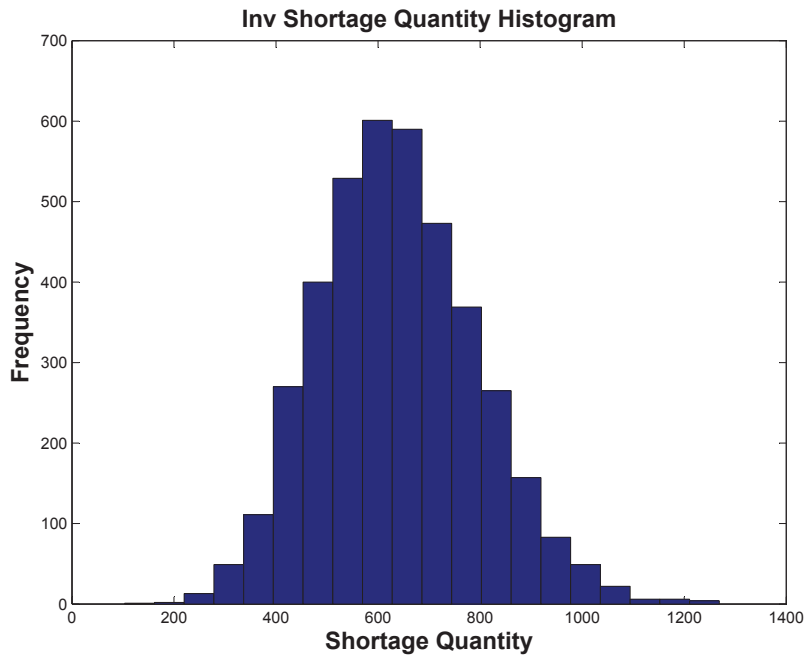


Figure 3.17: Distribution of Shortage Quantity

N – Number of outstanding orders. An outstanding order is defined as an order placed but not arrived as planned

SF – Shortfall. Shortfall is defined as the gap between cumulative supply and cumulative demand

F_i – Probability that order lead time is less than i days, $F_i = P(L \leq i)$

L – Order lead time

μ_L – Average order lead time

σ_L – Standard deviation of order lead time

D_i – Order quantity in i^{th} period

μ_D – Average order quantity in one period

σ_D – Standard deviation of order quantity in one period

R_i – Indicator variable, equals to one if order placed at time i is rejected and 0 otherwise

r – Probability of order being rejected

T – Time to excursion occurrence

G_i – $G_i = P(T \leq i)$, probability that time to excursion is less than i days

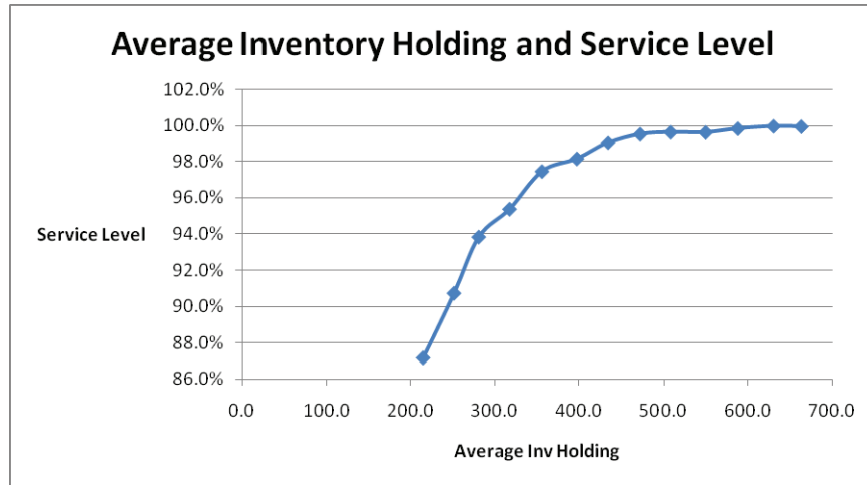


Figure 3.18: Average Inventory Holding and Service Level under Stochastic Lead Time and Batch Reject Risk

s – Time from excursion occurrence to excursion recovery

H – Planning horizon, multiples of ordering interval

3.4.1 Model

Suppose the time horizon starts at zero. Given an excursion happens at time t , it takes a total time of s to discover the excursion and recover from it. At time $t+s$, the number of outstanding orders is sum of three parts: the number of orders placed but not yet arrived, the number of orders arrived but rejected due to batch rejection and the number of orders being rejected due to excursion.

$$\begin{aligned}
 N_{t+s} &= \sum_{i=1}^t \delta(L_{t-i} > s+i) + \sum_{i=1}^t \delta(L_{t-i} \leq s+i) R_{t-i} + s \\
 &= s + \sum_{i=1}^t (1 - \delta(L_{t-i} \leq s+i)) + R_{t-i} \delta(L_{t-i} \leq s+i) \\
 &= s + \sum_{i=1}^t (1 - \delta(L_{t-i} \leq s+i)) (1 - R_{t-i})
 \end{aligned}$$

The expected number of outstanding orders is

$$\begin{aligned}\mathbb{E}[N_{t+s}] &= \mathbb{E}\left[t + s - \sum_{i=1}^t \delta(L_{t-i} \leq s+i) (1-R_{t-i})\right] \\ &= s + \sum_{i=1}^t (1 - (1-r)F_{s+i})\end{aligned}$$

The variance of the number of outstanding orders is

$$\begin{aligned}\text{Var} [N_{t+s}] &= \text{Var} \left[s + \sum_{i=0}^t (1 - \delta(L_{t-i} \leq s+i) (1-R_{t-i})) \right] \\ &= \sum_{i=0}^t \text{Var} [\delta(L_{t-i} \leq s+i) (1-R_{t-i})]\end{aligned}$$

To simplify the formula for variance of the number of outstanding orders we apply following property. Given X is independent of Y, then

$$\begin{aligned}\text{Var} [XY] &= E[X^2Y^2] - (E[XY])^2 \\ &= \text{Var} [X] \text{Var} [Y] + \text{Var} [X] (E[Y])^2 + (E[X])^2 \text{Var} [Y]\end{aligned}$$

The summation term in the number of outstanding orders is

$$\begin{aligned}\text{Var} [\delta(L_{t-i} \leq s+i) (1-R_{t-i})] &= \text{Var} [\delta(L_{t-i} \leq s+i)] \text{Var} [1-R_{t-i}] + \text{Var} [\delta(L_{t-i} \leq s+i)] (E[1-R_{t-i}])^2 \\ &\quad + (E[\delta(L_{t-i} \leq s+i)])^2 \text{Var} [1-R_{t-i}] \\ &= F_{s+i} (1 - F_{s+i}) r (1-r) + F_{s+i} (1 - F_{s+i}) (1-r)^2 + F_{s+i}^2 r (1-r) \\ &= (1-r)F_{s+i} (1 - (1-r)F_{s+i})\end{aligned}$$

So the variance of the number of outstanding orders is

$$\begin{aligned}\text{Var} [N_{t+s}] &= \sum_{i=1}^t \text{Var} [\delta(L_{t-i} \leq s+i) (1-R_{t-i})] \\ &= \sum_{i=1}^t (1-r)F_{s+i} (1 - (1-r)F_{s+i})\end{aligned}$$

In our batch production system the gap between cumulative supply and cumulative demand depends on the number of outstanding orders. Since the order quantity of each order is random, use condition on the number of outstanding orders to calculate the expected shortfall.

$$\mathbb{E}(SF_{t+s}) = \mathbb{E}[\mathbb{E}[SF_{t+s} | N]] = \mu_D \left(s + \sum_{i=1}^t (1 - (1-r)F_{s+i}) \right)$$

Using the conditional variance formula, we can calculate variance of the shortfall as

$$\begin{aligned} \text{Var} [SF_{t+s}] &= \mu_N \sigma_D^2 + \mu_D^2 \sigma_N^2 \\ &= \left(s + \sum_{i=1}^t (1 - (1-r)F_{s+i}) \right) \sigma_D^2 + \mu_D^2 \sum_{i=1}^t (1-r)F_{s+i} (1 - (1-r)F_{s+i}) \end{aligned}$$

We have calculated the expected the gap between cumulative supply and cumulative demand and its variance as well. Summing over all possible times of excursion occurrence, we can get the expected inventory shortfall at the time of excursion recovery. Let SF_R denote the inventory shortfall at the time of excursion recovery. Note that T is the time of excursion occurrence and G is the cumulative distribution of the excursion occurrence time distribution and H is the planning horizon. We want to make sure our required stock level can hedge against any risk within the planning horizon. But SF_R represents the required stock level at the time of excursion recovery given that an excursion happens, regardless of our planning horizon. Let's define $SF_{RH} \triangleq SF_R I_{\{T \leq H\}}$ where $I_{\{\cdot\}}$ is the indicator function. It is the shortfall at the time of the recovery when the excursion happens within our planning horizon. Our goal is to calculate the mean and variance of SF_{RH} . The mean SF_{RH} is

$$\begin{aligned} \mathbb{E}[SF_{RH}] &= \mathbb{E}[\mathbb{E}[SF_{RH} | T]] \\ &= \mathbb{E} \left[\mu_D \left(s + \sum_{i=1}^T (1 - (1-r)F_{s+i}) \right) I_{\{T \leq H\}} \right] \\ &= \sum_{t=1}^H P(T=t) \left(\mu_D \left(s + \sum_{i=1}^t (1 - (1-r)F_{s+i}) \right) \right) \\ &= s\mu_D G_H + \mu_D \sum_{t=1}^H \sum_{i=1}^t (1 - (1-r)F_{s+i}) P(T=t) \\ &= s\mu_D G_H + \mu_D \sum_{i=1}^H \sum_{t=i}^H (1 - (1-r)F_{s+i}) P(T=t) \\ &= s\mu_D G_H + \mu_D \sum_{i=1}^H (1 - (1-r)F_{s+i}) \sum_{t=i}^H P(T=t) \\ &= s\mu_D G_H + \mu_D \sum_{i=1}^H (1 - (1-r)F_{s+i}) (G_H - G_{i-1}) \end{aligned}$$

Note that G_H is the probability the excursion happens within our planning horizon. $s\mu_D G_H$ is the expected rejected quantity due to the excursion. When we have low probability of an excursion, $s\mu_D G_H$ will be very small.

To calculate the variance of the inventory shortfall at the time of recovery, first we calculate $\mathbb{E}[SF_{RH}^2]$, then calculate $\text{Var}[SF_{RH}]$.

$$\begin{aligned}
E [SF_{RH}^2|T] &= E[E[(SF_{RH}^2|T)|N]] \\
&= E \left[E \left[\left(\sum_{i=1}^N D_i \right)^2 \right] \right] \\
&= E \left[E \left[\sum_{i=1}^N D_i^2 + \sum_{i<j} 2D_i D_j \right] \right] \\
&= E [NE[D_i^2] + N(N-1)E[D_i D_j]] \\
&= E [N(\sigma_D^2 + \mu_D^2) + N(N-1)\mu_D^2] \\
&= \sigma_D^2 E[N] + \mu_D^2 E[N^2] \\
&= \sigma_D^2 E[N] + \mu_D^2 (E[N])^2 + \mu_D^2 Var(N) \\
&= \sigma_D^2 \left(s + \sum_{i=1}^T [1 - (1-r)F_{s+i}] \right) + \mu_D^2 \left(s + \sum_{i=1}^T [1 - (1-r)F_{s+i}] \right)^2 \\
&\quad + \mu_D^2 \sum_{i=1}^T ((1-r)F_{s+i} [1 - (1-r)F_{s+i}])
\end{aligned}$$

$$\begin{aligned}
E [SF_{RH}^2] &= E[E[SF_{RH}^2|T]] \\
&= \sum_{t=1}^H (\sigma_D^2 \left(s + \sum_{i=1}^T [1 - (1-r)F_{s+i}] \right) + \mu_D^2 \left(s + \sum_{i=1}^T [1 - (1-r)F_{s+i}] \right)^2 \\
&\quad + \mu_D^2 \sum_{i=1}^T ((1-r)F_{s+i} [1 - (1-r)F_{s+i}])) P(T=t)
\end{aligned}$$

Then we can calculate the variance of SF_{RH} as

$$Var [SF_{RH}] = E [SF_{RH}^2] - (E [SF_{RH}])^2$$

If the lead time is geometrically distributed with parameter p and the excursion occurrence time is geometric distributed with parameter q , then we have cumulative distribution function for the lead time and excursion time: $F_i = 1 - (1-p)^i$ and $G_j = 1 - (1-q)^j$. Based on these we can calculate the closed form of the required stock level at the beginning of our planning horizon. Detailed calculation in this case is presented in Appendix A.

3.4.2 Numerical Example

In this section we will discuss a practical numerical example to illustrate the calculation of required stock level in a biopharmaceutical supply chain. For comparison purpose, we use

the same numerical example as in sections 3.1, 3.2, 3.3.

As a practical example we consider the purification stage in biopharmaceutical manufacturing. The model is very general and can be easily applied to other stages of the biopharmaceutical supply chain. We assume the ordering lead time is geometrically distributed with mean 56 days. Daily demand for the purification stage is 8 mmu and standard deviation is 1.6 mmu. To make sure we have very close to 100% service level, the safety factor is set to 3. And we place a purification order every 3 days. Each order has a probability of 1% of being rejected. Also we assume that the purification stage is exposed to facility contamination risk which can occur every 5 years and the total time for excursion detection and recovery is 60 days. The planning horizon is one year. Then based on the formula in Section 3.4.1 we can calculate mean and standard deviation of the shortfall and also the required stock level:

$$\mathbb{E}(SF_{RH}) = 113$$

$$Std(SF_{RH}) = 243$$

$$\begin{aligned} S &= \mathbb{E}(SF_{RH}) + k\sqrt{Std(SF_{RH})} \\ &= 113 + 243k \end{aligned}$$

When $k = 3$, the required stock level is 842 mmu. Recall that the required stock level derived from the lead time demand based approach is 1793 mmu. Note that the average demand during the excursion is $8 \times 60 = 480$ mmu but from our calculation the average shortfall is just 113 mmu. This is due to the low probability that excursion happens within our planning horizon. In practice holding such low stock level is unacceptable due to the significant impact of the excursion. If the supply chain is exposed to earthquake risk and it takes around 3 months to recover from the earthquake damage, then supply chain manager usually prefers to hold at least 3 months of strategic stock to hedge against such a catastrophic event. So to make our model more applicable to practice, we can make sure that the estimate of the average gap between cumulative supply and cumulative demand is at least the size of the excursion. Then we have

$$\begin{aligned} \mathbb{E}(\widetilde{SF}_{RH}) &= \max(\mathbb{E}(SF_{RH}), s\mu_D) \\ &= 480 \end{aligned}$$

The required stock level in this case is 1209 mmu.

3.4.3 Simulation

In this section we will use Monte Carlo simulation to evaluate the performance of the required stock level derived from our calculation. The key performance measure is service level

Parameter	Value	Description
μ_D	10	Mean Demand
σ_D	2	Demand Standard Deviation
r	5%	Batch Reject Rate
μ_L	30	Mean Lead Time
y	5	Excursion Frequency (Years)
s	60	Excursion Duration (Days)
H	750	Planning Horizon
T	5	Ordering Interval
N	4000	# of Simulation Iterations

Table 3.7: Simulation Parameters under Stochastic Lead Time, Batch Reject and Excursion Risk

throughout our planning horizon. Also we want to avoid excess stock holdings so we will look at average stock holding as well. In our experiment we simulate three different risks in the biopharmaceutical supply chain: stochastic lead time risk, batch rejection risk and excursion risk. Stochastic lead time is simulated as geometric random variable. Batch rejection is simulated as a Bernoulli random variable. Excursion occurrence time is geometrically distributed.

We assume per period demand is normally distributed with mean 10 and standard deviation 2. The ordering lead time is geometrically distributed with mean of 30 days. The simulation horizon is 750 days. An order is place every 5 days. The size of each order is 50. Also assume there is no planned ordering lead time. The batch reject rate is 5%. Besides stochastic lead time and batch rejection risk, the supply chain system is exposed to the risk of an excursion which can happen once every 5 years and the time from excursion occurrence to excursion recovery is 60 days. Table 3.7 summarizes the simulation parameters.

Given the simulation parameters, we can calculate the required stock level as a function of the safety factor k . Using formula from Section 3.4.1, we get

$$\begin{aligned}\mathbb{E}(SF_{RH}) &= s\mu_D G_H + \mu_D \sum_{i=1}^H (1 - (1-r)^i) (G_H - G_{i-1}) \\ &= 279\end{aligned}$$

$$Std(SF_t) = 427$$

Then the required stock level as a function of the safety factor k is

$$\begin{aligned}S &= \mathbb{E}(SF_t) + k\sqrt{Var(SF_t)} \\ &= 279 + 427k\end{aligned}$$

	k = 0		k = 0.5		k = 1		k=1.5		k=2	
	B=279									
	SS = 0		SS = 213.5		SS = 427.1		SS = 640.6		SS = 854.2	
	Mean	Stde	Mean	Stde	Mean	Stde	Mean	Stde	Mean	Stde
Service Level	18.8%	0.2%	65.3%	0.3%	94.3%	0.2%	99.7%	0.0%	100.0%	0.0%
Inv Holding	16.1	0.1	103.8	0.6	283.5	1.1	489.3	1.2	706.7	1.2

Table 3.8: Simulation Results under Stochastic Lead Time, Batch Reject and Excursion Risk

To evaluate the performance of our stock policy, at the beginning of the simulation we set the inventory level to the required stock level given the choice of the safety factor. In each simulation iteration, we first simulate daily demand with normal distribution. Then the excursion time is simulated based on a geometric distribution. Every time we place an order, we simulate the ordering lead time. If excursion happens within our planning horizon, all orders during the excursion duration are rejected. Based on the simulation sample path, at the end of each day we calculate cumulative supply and cumulative demand. Cumulative supply is the total quantity arrived so far. The quantity depends on the actual lead time of each order. Also we add up demand to date to get cumulative demand. At the end of the simulation we evaluate the service level and average inventory holding. After collecting simulation statistics, we calculate the mean and standard deviation of the service level and also the mean and standard deviation of inventory holding. Since we are more concerned about the performance of the supply chain when the excursion happens, when calculating simulation statistics, only simulation iterations with excursion occurrence within the planning horizon will be counted.

From the simulation result we can see that setting the service factor to 2 gives us more than 99% service level and an average inventory holding of 508. On the other hand, zero service factor gives us only around 19% service level but with much less inventory holding. In biopharmaceutical supply chain, an important factor of performance measure is average inventory holding. Figure 3.22 summarizes the simulation results by comparing average inventory holding and service level. We can see that service level is an increasing concave function of average inventory holding.

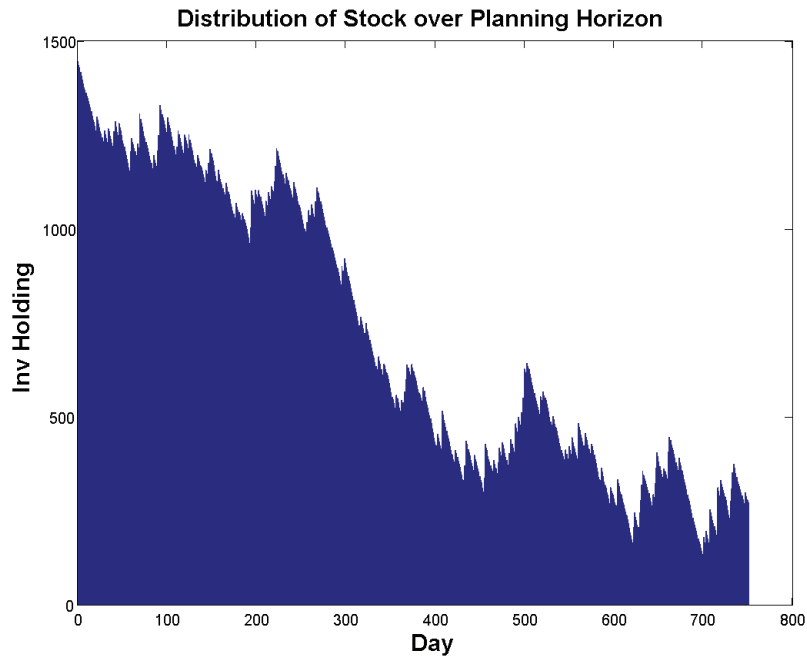


Figure 3.19: Stock Distribution in the Planning Horizon

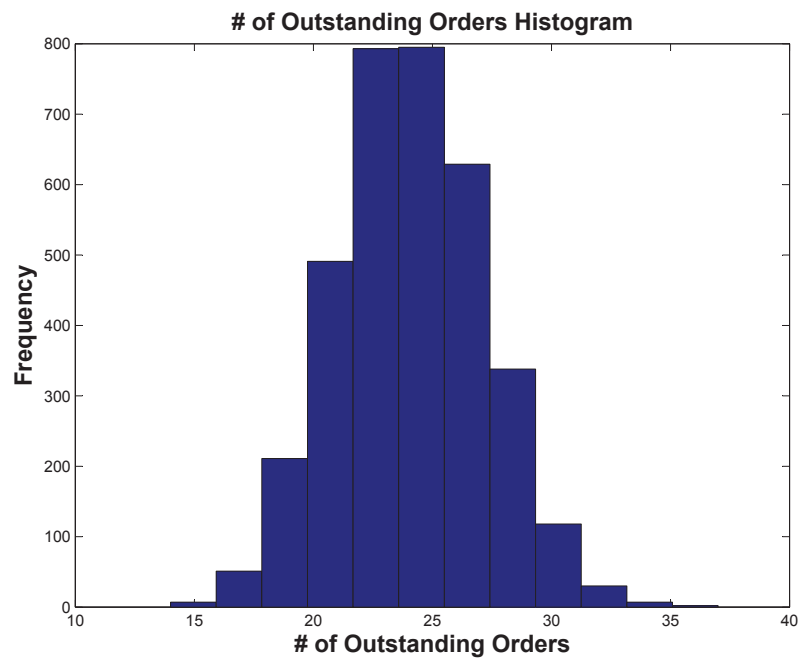


Figure 3.20: Distribution of # of Outstanding Orders

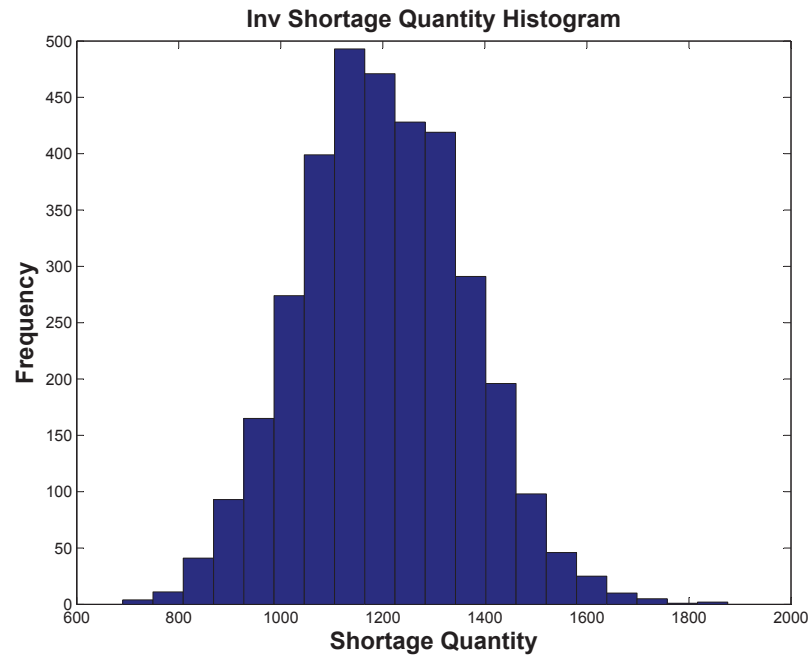


Figure 3.21: Distribution of Shortage Quantity

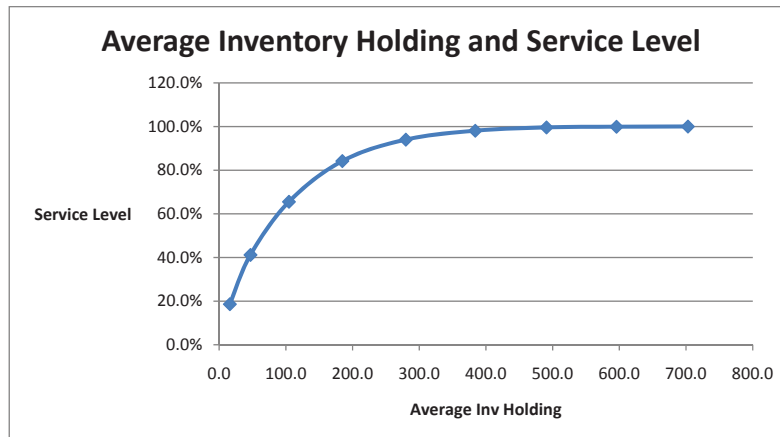


Figure 3.22: Average Inventory Holding and Service Level under Stochastic Lead Time, Batch Reject and Excursion Risk

Chapter 4

Safety Stock Planning Implementation in Practice

Up to this point we have analyzed safety stock levels for a single manufacturing stage. Biopharmaceutical supply chains are multi-stage. We begin with a literature review of risk management in multi-echelon inventory system and then extend the model developed in chapter 3 to multi-stage settings.

4.1 Supply and Demand Risk in Multiechelon Inventory System

Clark and Scarf (1960) was one of the first papers investigating optimal policies in a multi-echelon inventory setting. They considered a multi-echelon model with stochastic demand and deterministic lead time. Based on the idea of echelon inventory and general assumptions regarding the purchasing cost, holding cost and shortage cost, they proposed an approach to find the optimal policy recursively by decomposing the two echelon problem into two separate single echelon problem. The optimal policy was a function of echelon inventory. Federgruen and Zipkin (1984) extended Clark and Scarf (1960) results to the infinite horizon case using discounted and average cost function. In the case of normal demands, they proposed simplified calculation for optimal policy by finding a closed form for the average penalty cost function.

Simulation and numerical procedures also have been proposed to evaluate inventory policy in a multi-echelon setting. Clark (1960) discussed the use of simulation to evaluate the performance of a multiechelon inventory system based on a practical example. Three different situations for low level echelons were discussed based on whether the item is repairable in the repair depot. The measures of the performance focused on stock distribution which was the stock level as a function of time and supply expenditure. Arkin etc. (1989) considers the computational complexity issue in uncapacitated multi-echelon production planning

problems. They show that polynomial time algorithms are available for single item serial and assembly systems. But for more complex structures, the time complexity is NP-complete. Houtum and Zijm(1991) applied numerical techniques in evaluating stochastic multi-echelon production system. Their model assumed stochastic stationary demand and deterministic lead time in each stage. The numerical routines are utilized to analyze the incomplete convolutions of distribution functions after decomposing multi-echelon problems into single echelon one. Glasserman and Tayur (1995) developed simulation based methods to estimate the sensitivities of inventory costs with respect to policy parameters. With simulation they were able to consider more complex multi-echelon inventory systems. They developed estimators of performance measures' derivatives with respect to the base-stock levels and applied them in the optimization for finding the optimal inventory policy.

Clark (1971) did a nonmathematical survey of related research in multi-echelon inventory. He identified two popular structures of multi-echelon inventory systems which are serial and parallel. The multi-echelon inventory problems are categorized into several types including deterministic or stochastic, single or multi product, stationary or nonstationary. The main approaches for stochastic models are expected cost, stationary process analysis based on queueing theory and dynamic programming. He also discussed allocation models and contingency support models which are closely related to the work in this dissertation. The goal of allocation model is to distribute a given amount of system stock to various activities in a multi-activity system to satisfy a cost or performance objective. In our multi-echelon biopharmaceutical supply chain safety stock planning, the objective is to allocate sufficient safety stock in different stages of the supply chain to ensure close to 100% service level but also make sure average inventory holding is minimized. The contingency support models Clark surveyed are similar to excursion modeling in a biopharmaceutical supply chain. In addition to the modeling of normal supply systems, the contingency modeling aims to find the best strategy in case of postulated contingency operations.

A significant number of papers focus on the analysis of base-depot inventory systems. Sherbrooke (1967) discussed the redistribution of stock in a base-depot inventory system under compound Poisson demand. He proposed a five stage approximation scheme to find the optimal policy which minimizes the expected cost. To find the optimal solution, the calculation needs to be done for each level of depot stock. The stock redistribution is determined with marginal cost analysis which is the maximum decrease in expected backorders per unit cost. Simon (1969) considered a two-echelon base-depot inventory model for repairable items. The stage lead time was assumed to be deterministic and demand was generated from Poisson process. He calculated the exact expression for the stationary distribution of stock on hand in each base as well as the depot. Graves (1985) addressed the problem of determining the inventory level in a multi-echelon inventory system for repairable items. He considered base-depot structure, compound demand process and deterministic lead time. He proposed a model to compute the mean and variance of the depot backorder level recursively. Then the base stock level is approximated by fitting its first two moments to that of a negative binomial distribution. Based on the approximation, the stockage level at each base can be determined as a function of the fill rate.

Due to the complexity of multi-echelon inventory systems, a number of approximation schemes are proposed to find the near optimal policy. Schwarz and Schrage (1974) proposed the optimal and near optimal approximations for a multi-echelon assembly system under continuous review with constant demand in an infinite planning horizon. Their approach is based on the idea of echelon inventory and the optimal solution can be found by a branch and bound routine. De Bodt and Graves (1985) considered a multi-stage inventory system with stochastic and stationary demand. The analysis starts from the two echelon case and is based on the idea of joint replenishment where for each order placed by upstream stage there is an accompanying order by downstream stage. The cost expression is approximated as a function of reordering policy. Then they extended the model to M-echelon case and also proposed heuristics to determine order quantity multipliers. Rong etc. (2008) studied a continuous review multi-echelon distribution network under Poisson demand. They developed two heuristics to approximate the base-stock level in the systems. Compared to the approach by Graves (1985) they employed a bottom-up base to depot approach to approximate the base and depot stock levels.

Supply and demand uncertainty may have quite different impacts on the supply chain and thus may require different management strategies. Snyder and Shen (2006) explored the demand and supply uncertainty issue in multi-echelon supply chains. The demand uncertainty they considered was stochastic iid demand and the supply uncertainty is modeling by supply disruption driven by a two state Markov chain. They applied simulation to show that the two types of uncertainty require quite different optimal strategies in terms of ordering frequency, inventory placement, etc.

4.2 Multiechelon Safety Stock Planning

In this section we will discuss safety stock planning in a multi-echelon single product system. We focus our discussion in a two-echelon case and the model can be easily extended to cases with more echelons. In our two echelon setting, we denote the upstream echelon as stage 2 and the downstream echelon as stage 1. Each stage is exposed to stochastic lead time risk, batchion reject and excursion risk. There is an inventory holding point right after each stage. Figure 4.1 shows the basic structure of a two-echelon inventory system. Rectangle represent the stages and triangles are the inventory positions. We assume that the output of stage 1 meets demand directly. S_1 and S_2 are the inventory holding levels for the output of stage 1 and stage 2 respectively.

Now let's consider the inventory holding for the output of stage 1. In the case when there is sufficient inventory holding at the output of stage 2, any order placed at stage one will not have any delay due to backordered input material incurred by stage 2. Then the inventory holding s_1 only needs to cover risks occurring in stage 1. The risks may include stochastic lead time, batch rejection and excursion. This should be the lower bound of s_1 . Another extreme case is when $s_2 = 0$ so there is no inventory holding of stage 2 output. So any order placed in stage 1 triggers an order placed in stage 2. In this case the inventory

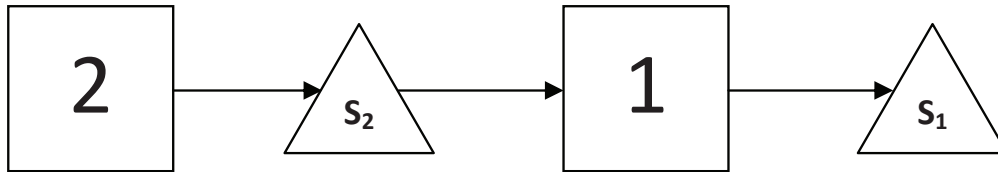


Figure 4.1: Multiechelon Inventory System

holding s_1 not only needs to cover risks in stage 1 but also risks in stage 2 as well.

Recall when there is only stochastic lead time risk, the required stock level for a single stage system is

$$S = T\mu_D \left(\sum_{i=0}^{\infty} (1-F_{iT}) \right) + k \sqrt{\mu_D^2 T^2 \sum_{i=0}^{\infty} F_{iT}(1-F_{iT}) + \left(\sum_{i=0}^{\infty} (1-F_{iT}) \right) T\sigma_D^2}$$

Assume the cumulative distribution function of stage 1 cycle time is F^1 and the cumulative distribution function of stage 1 and 2 cycle time is F^{12} . Then we have following bound for the required stock level of stage 1 s_1

$$S_1 \geq T\mu_D \left(\sum_{i=0}^{\infty} (1-F^1_{iT}) \right) + k \sqrt{\mu_D^2 T^2 \sum_{i=0}^{\infty} F^1_{iT}(1-F^1_{iT}) + \left(\sum_{i=0}^{\infty} (1-F^1_{iT}) \right) T\sigma_D^2}$$

$$S_1 \leq T\mu_D \left(\sum_{i=0}^{\infty} (1-F^{12}_{iT}) \right) + k \sqrt{\mu_D^2 T^2 \sum_{i=0}^{\infty} F^{12}_{iT}(1-F^{12}_{iT}) + \left(\sum_{i=0}^{\infty} (1-F^{12}_{iT}) \right) T\sigma_D^2}$$

To implement the model in practice, we can have two different approaches. The first is try to adjust the inventory level in each stage to the target level proposed by our model. In this way the multi-echelon problem is decomposed into single stage ones. The second approach is keeping track of excess holdings in downstream stages and then reducing inventory in upstream stages. The reduction should depend on the excess we have in downstream stages as well as the cycle time of those stages.

4.2.1 Inventory Adjustment to the Proposed Level

In this section we propose a procedure to adjust the current inventory level to the target level proposed by our model. We start our adjustment recursively from downstream stage to the upstream stage. Note that at each stage we can have either excess inventory or shortage.

When there is excess inventory at current stage, we reduce the demand for the previous stage based on a multiplier parameter. This multiplier parameter determines how fast excess is reduced to our proposed level. A conservative policy would be choosing a relatively big multiplier to gradually reduce the excess. On the other hand, when there is shortage at the current stage, we choose a small multiplier to request replenishment from upstream quickly. Also in this particular implementation, we separate demand uncertainty from supply uncertainty by letting the final packaging stage absorb all the demand uncertainties. This kind of policy is typical in biopharmaceutical supply chains which require close to 100% service level. This means the final stage should never stocks out. For packaging stage, instead of passing the exact customer upstream we can just pass the average demand upstream. During the safety stock replenishment we would be able to replenish the packaging stage inventory back to the target level.

Figure 4.2 illustrates the inventory system we use to demonstrate our adjustment process. Here FFD stage fills product for different countries and each country has their own specific packaging requirement. An example of this situation would be filling of 1000IU product which may go to US, Europe or Canada markets.

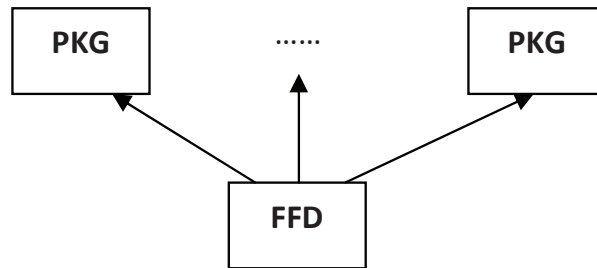


Figure 4.2: Multi-product Inventory System

Following are the notation for this section:

$\tilde{\mu}_{D_i}$ – Demand per period at production stage i adjusted by the inventory holding of previous stage

F_j^i – Probability lead time is less than or equal to j at production stage i

$\mu_{D_{il}}$ – Demand per period of product l at production stage i

$\sigma_{D_{il}}$ – Demand standard deviation per period of product l at production stage i

T_{il} – Order interval of product l at production stage i

k_{il} – Safety factor of product l at production stage i

E_{il} – Excess inventory holding of product l at production stage i

m_{il} – Safety stock replenishment horizon multiplier (in terms of production interval) of product l at production stage i . m_{il} is determined by the available capacity to replenish stock

S_{il} – Target inventory holding at production stage i

\tilde{S}_i – Actual Inventory stock holding at production stage i

Let 1 denote the index for the packaging stage and 2 for the FFD stage. The target inventory level for product l at the packaging stage is

$$S_{1l} = T_{1l}\mu_{D_{1l}} \sum_{j=1}^{\infty} (1 - F_{jT_{1l}}^1) + k_{1l} \sqrt{\sum_{j=1}^{\infty} (1 - F_{jT_{1l}}^1) T_{1l} \sigma_{D_{1l}}^2 + \mu_{D_{1l}}^2 T_{1l}^2 \sum_{j=1}^{\infty} F_{jT_{1l}}^1 (1 - F_{jT_{1l}}^1)}$$

The excess or shortage to target at the packaging stage is

$$E_{1l} = \tilde{S}_{1l} - S_{1l}$$

We adjust the demand for the FFD stage according to the excess or shortage at the packaging stage. When there is shortage, we choose m_{1l} to be relatively small to speed up the replenishment for the shortage.

$$\tilde{\mu}_{D_{2l}} = \mu_{D_{1l}} - \frac{E_{1l}}{m_{1l}T_{1l}}$$

Combine the demand for different products and then we have total demand for the FFD stage

$$\tilde{\mu}_{D_2} = \sum_{l=1}^n \tilde{\mu}_{D_{2l}}$$

Then the target inventory level for the FFD stage is

$$S_2 = \tilde{\mu}_{D_2}\mu_{L_2} + k_2\mu_{D_2} \sqrt{\sum_{j=1}^{\infty} F_j^2 (1 - F_j^2)}$$

Note here we assume all the demand uncertainty is absorbed in the packaging stage so we didn't consider any demand uncertainty in above formula.

4.2.2 Inventory Adjustment to Current FG Level

A typical biopharmaceutical supply chain requires close to 100% service level. To achieve this, usually excess inventory is held in downstream stages which are closer to the finished goods. When implementing our model in practice, instead of opting for method proposed in the previous section, supply chain managers may choose to hold excess inventory in finished goods as a precautionary measure. The advantage is that excess finished goods holding is not exposed to supply risks. We need to account for such excess downstream inventory when planning the holding for upstream stages.

The general idea of our proposed approach is based on the interchangeability between safety stock and safety time. When there is excess holding in finished goods, we can overcome

longer cycle time in packaging. By converting the excess finished goods into cycle time we can allow for longer lead time for stages prior to packaging.

Let's consider the inventory system illustrated in Figure 4.1. For simplicity we only consider the stochastic lead time risk. Then based on our proposed model, the required stock level for the output of stage 1 is

$$S_1 = T_1 \mu_D \left(\sum_{i=0}^{\infty} (1 - F_{iT_1}^1) \right) + k \sqrt{\mu_D^2 T_1^2 \sum_{i=0}^{\infty} F_{iT_1}^1 (1 - F_{iT_1}^1) + \left(\sum_{i=0}^{\infty} (1 - F_{iT_1}^1) \right) T_1 \sigma_D^2}$$

Assume current holding for the output of stage 1 is I_1 . And we have excess holding, i.e. $I_1 > S_1$. The lead time for stage 1 is L_1 . We would like to find the lead time \widetilde{L}_1 which corresponds to our current inventory holding. A natural candidate is $\widetilde{L}_1 = (1 + \alpha)L_1$ where $\alpha \geq 0$. Note we have

$$\begin{aligned} \widetilde{F}_{iT_1}^1 &= P(\widetilde{L}_1 \leq iT_1) \\ &= P((1 + \alpha)L_1 \leq iT_1) \\ &= P(L_1 \leq \frac{iT_1}{1 + \alpha}) \\ &= F_{\frac{iT_1}{1 + \alpha}}^1 \end{aligned}$$

We want to find α such that

$$\begin{aligned} I_1 &= T_1 \mu_D \left(\sum_{i=0}^{\infty} (1 - \widetilde{F}_{iT_1}^1) \right) + k \sqrt{\mu_D^2 T_1^2 \sum_{i=0}^{\infty} \widetilde{F}_{iT_1}^1 (1 - \widetilde{F}_{iT_1}^1) + \left(\sum_{i=0}^{\infty} (1 - \widetilde{F}_{iT_1}^1) \right) T_1 \sigma_D^2} \\ &= T_1 \mu_D \left(\sum_{i=0}^{\infty} (1 - F_{\frac{iT_1}{1 + \alpha}}^1) \right) + k \sqrt{\mu_D^2 T_1^2 \sum_{i=0}^{\infty} F_{\frac{iT_1}{1 + \alpha}}^1 (1 - F_{\frac{iT_1}{1 + \alpha}}^1) + \left(\sum_{i=0}^{\infty} (1 - F_{\frac{iT_1}{1 + \alpha}}^1) \right) T_1 \sigma_D^2} \\ &= \mathcal{S}_1(\alpha) \end{aligned}$$

Depending on the distribution of lead time, α can be found either by solving the above equation analytically or using numerical procedures. Then we have

$$\alpha = \mathcal{S}_1^{-1}(I_1)$$

Since we are keeping to the current inventory level at packaging, we will adjust FC inventory to our target level based on downstream FG holding. We account for FG excess by allowing less lead time for the FFD stage. Let \widetilde{L}_2 be the lead time of stage 2 induced by

the current FG holding. We can decompose the overall lead time of FFD and packaging:

$$\begin{aligned}
 L &= L_2 + L_1 \\
 &= L_2 - \alpha L_1 + (1 + \alpha)L_1 \\
 &= (L_2 - \alpha L_1) + \widetilde{L}_1 \\
 &= \widetilde{L}_2 + \widetilde{L}_1
 \end{aligned}$$

So we have $\widetilde{L}_2 = L_2 - \alpha L_1$. Let $\widetilde{F}_{iT_2}^2$ be the cumulative distribution function for \widetilde{L}_2 . Assuming that T_2 is a multiple of T_1 , then the target required stock level for stage 2 is

$$S_2 = T_2 \mu_D \left(\sum_{i=0}^{\infty} (1 - \widetilde{F}_{iT_2}^2) \right) + k \sqrt{\mu_D^2 T_2^2 \sum_{i=0}^{\infty} \widetilde{F}_{iT_2}^2 (1 - \widetilde{F}_{iT_2}^2) + \left(\sum_{i=0}^{\infty} (1 - \widetilde{F}_{iT_2}^2) \right) T_2 \sigma_D^2}$$

Current inventory holding I_2 may be greater than our target level S_2 . We can either adjust the current level to our target level using the approach discussed in the previous section or keep the current holding and apply the method in this section to the upstream purification stage.

4.3 Sensitivity Analysis

In this section we will discuss the impact of different supply chain parameters on the level of safety stock holdings. This analysis is not only of theoretical interest but more importantly of great practical implication. Biopharmaceutical supply is a complex interconnected system and the level of safety stock depends on the performance of each major component of the supply chains. Quality assurance department of the biopharmaceutical supply chain is responsible for testing all intermediate and final product. Since the actual production time is fixed and much smaller than the quality assurance cycle time, the performance of this department determines the overall cycle time of the supply chain. In addition to that, the quality assurance department also manages batch rejection. The sensitivity analysis of lead time or batch reject against safety stock holding can give us a quantitative measure regarding the benefit of the marginal improvement of the quality assurance department. This analysis can also help us make economic decisions for investing in resources for the quality assurance department.

In our model the demand mean and standard deviation can have two practical interpretations. First it may represent the demand uncertainty in our supply chain. Customer demands may fluctuate over time due to various reasons including market growth, customers switching to competitors, etc. Second it may be the result of forecasting uncertainty. In biopharmaceutical supply chain, all the batch starts are based on future demand forecasting. The sales and marketing department is responsible for demand forecasting and planning.

Forecasting error is an important measure of demand forecasting performance. In the past the performance of demand forecasting is solely determined by the level of forecasting error. Based on the sensitivity analysis we will be able to evaluate dollar savings of inventory holding from unit improvement of forecasting error.

To simplify notation, in this section we will restrict our analysis to the case of stochastic lead time risk only. Also we assume the stochastic lead time is geometrically distributed. When there are batch rejection and excursion risks, analysis can be derived in a similar manner. First let's review the main results of safety stock planning when there is stochastic lead time risk. When the ordering interval T is 1 and the planned lead time is zero, the mean shortfall is

$$\begin{aligned}\mathbb{E}(\text{SF}_t) &= \mu_D \mu_N \\ &= \frac{\mu_D}{p}\end{aligned}$$

The variance of shortfall is

$$\begin{aligned}\text{Var}(\text{SF}_t) &= \mu_D^2 \sigma_N^2 + \mu_N \sigma_D^2 \\ &= \mu_D^2 \frac{1-p}{1-(1-p)^2} + \frac{1}{p} \sigma_D^2\end{aligned}$$

The required stock level is

$$S_1 = \frac{\mu_D}{p} + k \sqrt{\mu_D^2 \frac{1-p}{1-(1-p)^2} + \frac{1}{p} \sigma_D^2} \quad (4.1)$$

4.3.1 Improvement of Cycle Time

Let's assume that with additional investments in resources in the quality assurance department we will be able to reduce the mean overall cycle time but the distribution remains the same. First let's express the required stock level in terms mean lead time. Note we have

$$\mu_L = \frac{1}{p}$$

Plug in $p = \frac{1}{\mu_L}$ into Equation 4.2 and we get

$$S = \mu_D \mu_L + k \sqrt{\mu_D^2 \frac{\mu_L(\mu_L - 1)}{2\mu_L - 1} + \mu_L \sigma_D^2}$$

Then the sensitivity of the safety stock level to the mean cycle time is

$$\frac{\partial S}{\partial \mu_L} = \mu_D + k \frac{\mu_D^2 \frac{2\mu_L^2 - 2\mu_L + 1}{(2\mu_L - 1)^2} + \sigma_D^2}{2 \sqrt{\mu_D^2 \frac{\mu_L(\mu_L - 1)}{2\mu_L - 1} + \mu_L \sigma_D^2}}$$

Let's use a numerical example to illustrate the usage of this analysis. We use practical numbers from the purification stage as our example. Assume average daily demand is 8 mmu with standard deviation of 1.6 mmu. Lead time is geometrically distributed. The safety factor is 3. Figure 4.3 illustrates the relationship between cycle time and marginal change of required stock level. We can see that in this example the absolute change of required stock level decreases as the cycle time increases. The curve is decreasing convex and converges to mean demand as cycle time goes to infinity. Figure 4.4 shows the marginal dollar savings in inventory by assuming each mmu has a market value of \$600,000 and financing cost is 18%.

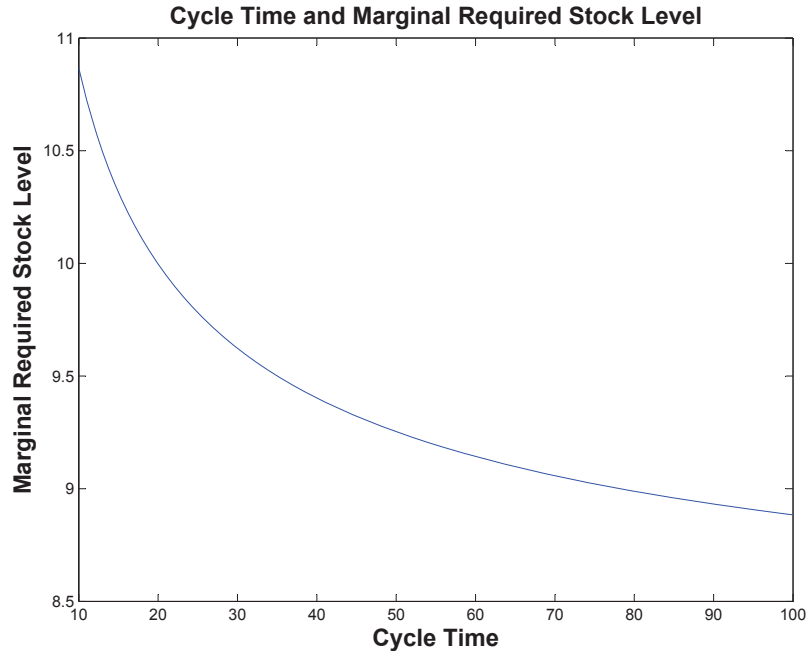


Figure 4.3: Cycle Time and Marginal Required Stock Level

4.3.2 Improvement of Forecasting Error

In this section we will discuss the sensitivity of required stock level to the forecasting error. The sales and marketing department is responsible for planning future demand and coming up with accurate forecasts. Based on the forecast, the operational planning department plans batch starts accordingly. Usually the forecast has a certain range of forecasting error due to incomplete knowledge of the market. By investing more resources in the forecasting process, the sales and marketing department is able to reduce the forecasting error. Examples of such investment include more accurate data collection, market survey, promotion, better IT system, etc. Eventually better forecasting means more accurate planning for operational planning and less inventory holding to hedge against possible stock out.

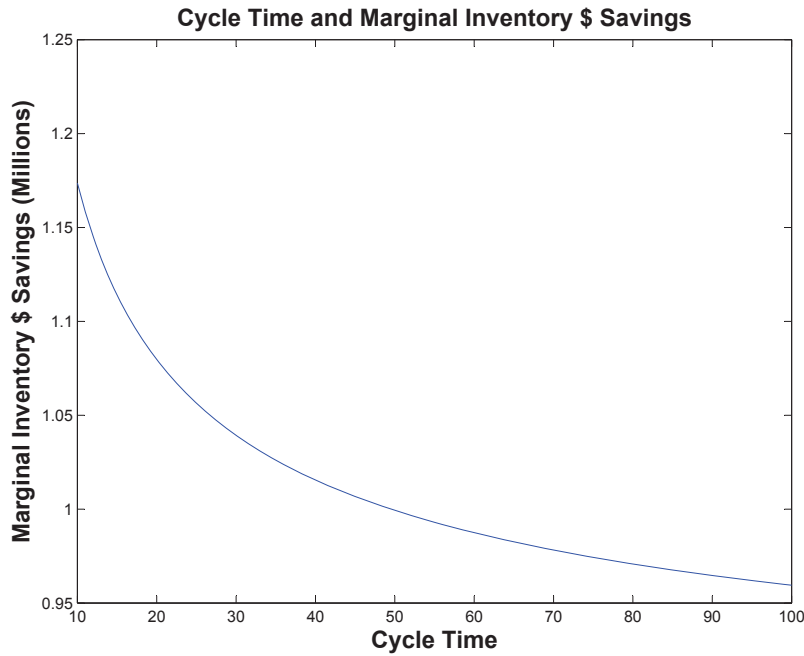


Figure 4.4: Cycle Time and Marginal Dollar Inventory Savings

Our analysis aims to find the relationship between forecasting error and inventory holding. This analysis will help companies evaluate whether or not to invest in improving forecasting accuracy as well as the financial benefit of such investment.

First let's assume the forecasting error is ρ and the unit is the number of percentage points. Then we have

$$\sigma_D = \frac{\rho \mu_D}{100}$$

The required stock level can be rewritten as

$$S = \frac{\mu_D}{p} + k \sqrt{\mu_D^2 \frac{1-p}{1-(1-p)^2} + \frac{1}{p} \frac{\mu_D^2 \rho^2}{10^4}}$$

Taking a derivative with respect to the forecasting error ρ ,

$$\frac{\partial S}{\partial \rho} = k \frac{\frac{1}{p} \mu_D^2 \frac{\rho}{10^4}}{\sqrt{\mu_D^2 \frac{\mu_L(\mu_L-1)}{2\mu_L-1} + \frac{1}{p} \frac{\mu_D^2 \rho^2}{10^4}}}$$

Let's use the same numerical example as in the previous section. Average daily demand is 8 mmu and forecasting error is 20%. Lead time is geometrically distributed. The safety

factor is 3. Figure 4.5 illustrates the relationship between forecasting error and marginal change of required stock level. We can see that in this example the absolute change of required stock level increases as the forecasting error increases. The curve is increasing concave. Figure 4.6 shows the marginal dollar savings in inventory by assuming each mmu has a market value of \$600,000 and financing cost is 18%.

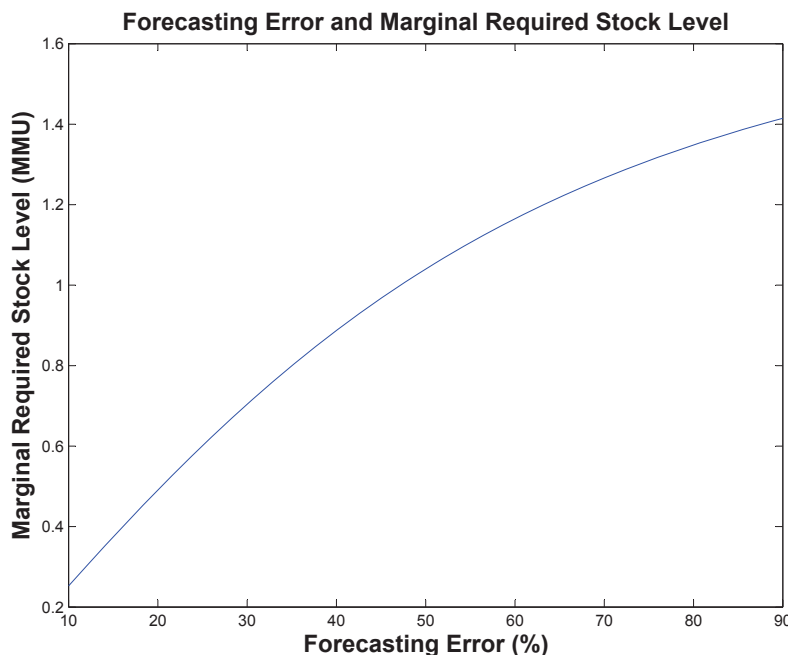


Figure 4.5: Forecasting Error and Marginal Required Stock Level

4.4 Regional Market Case Study

In this section we discuss the implementation of the proposed model for determining intermediate and finished goods stocks supplying a regional market of a biopharmaceutical product. There are five major global markets (in the order of sales) for the Kogenate product: Europe, US, Canada, Japan and rest of the world. The Japan market represents around 8% of total global sales and around 130 MMU per year in terms sales quantity. There are five product varieties based on potency sold in this market. Due to regulatory requirements, the Japan market requires a special fermentation material for its products. Although this fermentation material is allowed for sales in other global market regions, this is avoided to prevent possible stockouts in the Japan market. So the supply chain for the Japan market is relatively isolated from other markets after the fermentation stage. Due to its relative isolation and small sales volume, Japan market is a good place to start putting our model into practical use.

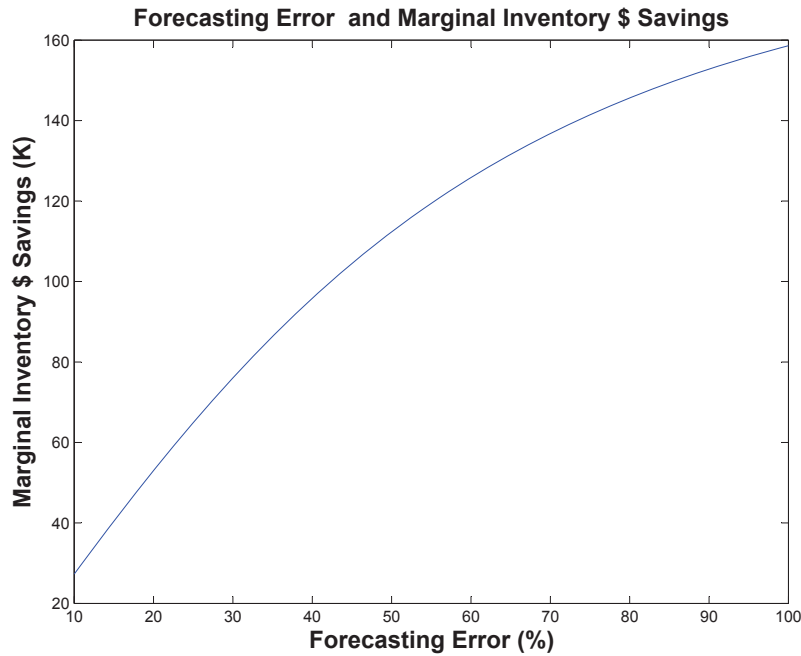


Figure 4.6: Forecasting Error and Marginal Dollar Inventory Savings

All Japan market products go through the same stages as we discussed before: fermentation, purification, FFD and packaging. Fermentation, purification and FFD stages are processed in a large manufacturing site supplying all regions. But due to the regulatory requirement, FC materials are packaged in a regional packaging facility in Japan. So FFD stage lead time not only includes stage production time, quality assurance cycle time but also the shipping time from the main manufacturing site to Japan. And the Japan regional center holds FC and FG materials. The goal of this case study is to plan the safety stock levels of FC and FG materials for the Japan regional center.

The safety stock planning model and simulation in chapter 3.4 was originally implemented in Matlab. In order to make the model available to supply chain planners, a VBA module was developed so that the model is accessible via Microsoft Excel. The inputs to the Excel spreadsheet are:

- Production start interval: this indicates how often the production batch starts are scheduled in this stage. For example, purification production starts occurs every 3 days.
- Production order cycle time: based on historical data in this particular implementation the overall cycle time for each stage is assume to be geometrically distributed. So given the average cycle time we are able to specify its complete distribution.
- Customer demand (mean and standard deviation): per period customer order volume

mean and standard deviation.

- Reject rate: probability of each batch being rejected.
- Excursion frequency: this value specifies how often an excursion occurs. In this model we assume the time to excursion is geometrically distributed. Given the excursion frequency we will be able to find the corresponding parameter for the Geometric distribution. This parameter is hard to estimate from historical data since excursions are very rare events. The appropriate value for this parameter can be based on supply chain planners' experience and evaluation of potential excursion risks.
- Excursion recovery time: this is how long it takes to discover the excursion plus the duration to fix the excursion. The time to discover the excursion is usually determined by the production process. For example, excursions in fermentation can be discovered after a certain number of tests in the quality assurance department of the fermented batch. The time to fix the excursion depends on the process as well as resources available. For example, the time to fix an excursion in one FFD facility could be determined by how much capacity is available in other FFD facilities and how soon we can shift the production to those facilities.
- Safety factor: the service level of the proposed safety stock level is determined by this parameter. Given the target service level we can use simulation to determine the appropriate value for this parameter.

Based on above parameters, the Excel VBA module calculates following two outputs:

- Pipeline stock level: this equals to the average shortfall in our proposed formula. This value is a function of the stage lead time, reject rate and excursion. On average we need at least this amount to avoid stock out in our planning horizon.
- Safety stock level: this is equal to safety factor times standard deviation of the shortfall in our proposed formula. This value gives us additional buffer protection against possible fluctuations in demand, lead time, batch rejection and excursion.

To calculate the safety stock level for each product type, parameter values are either estimated from historical data (for lead time and reject rate), projected from future forecasting (for demand mean and standard deviation) or based on current process settings and target service level (for start interval and safety factor).

Currently there is weekly review for FC and FG material in Japan regional center. So to evaluate current inventory holding we took the past 52 week of inventory levels for each product type and calculate the average. Next we compare the proposed safety stock level against the current inventory holding policy.

Another important measure of inventory level in biopharmaceutical supply chain is days on hand or DOH. DOH measures how many days of average demand that current inventory level is going to sustain. DOH can be calculated as follows.

$$\text{DOH} = \frac{\text{Inv Level}}{\text{Average Daily Demand}} \times 365$$

Based on the proposed inventory level and daily demand, average FC holding is 49 DOH and average average FG holding is 38 DOH. Total DOH based on the proposed model is 87 days. Current total FC and FG DOH is 150 days. Under the proposed model the percentage of savings is 42%. To calculate the dollar value savings, let's roughly assume that each MMU of FC and FG material has a market value of \$600,000. Financing and holding cost is 18%. Then under the proposed model, the total dollar savings is \$2.6 million.

4.5 Conclusion and Important Topics in Biopharmaceutical Supply Chain

4.5.1 Conclusion

In this dissertation we study the effect of order crossover in the context of safety stock planning to mitigate biopharmaceutical supply chain risks. Order crossover is evaluated in the cumulative curve perspective and the effective lead time perspective. We also incorporate the ordering interval and planned lead time in our analysis to make the model applicable to the biopharmaceutical supply chains. We further propose a framework to integrate mutually independent stochastic lead time risk, batch rejection risk and excursion risk into our safety stock planning model. The performance of the proposed models are verified via Monte Carlo simulations. We then extend our single stage model into multi-echelon multi-product case. The proposed model is implemented in an actual biopharmaceutical supply chain and demonstrates significant improvement over the existing approach. Our model assumes that there is sufficient excess capacity in the supply chain to replenish safety stocks as required. When there is a capacity bottleneck, we need to employ a rotation cycle type model to plan the multi-product safety stock replenishment when the supply chain is exposed to stochastic lead time, batch reject and excursion risks. This is an open problem in biopharmaceutical supply chain management as discussed in section 4.5.4 below. Other research challenge in biopharmaceutical supply chains concerns the performance evaluation of the conditional release strategy.

4.5.2 Infrastructure and Technology

The basis for successful supply chain management is accurate data input. If we have a perfect model for our supply chain but the input data is far from what's really happening in

this system, then we can hardly rely on the output of the model. To have a good estimate of cycle time we need the historical cycle time distribution. To plan safety stock, we need accurate counting of current inventory across all stages in our supply chain. To estimate batch rejections, we need to collect historical batch rejection statistics.

All these data seems quite natural and easy to secure for a complex biopharmaceutical supply chain. But in practice it is not. The reason is that most of the large corporations nowadays record their supply chain activities and monitor them by running an enterprise resource planning or ERP system whose basic architecture dates back to the 1960s. From industrial engineering point of view, such system acts more like an activity recording system rather than a optimal planning tool. For example, we may imagine that the stage cycle time or lead time is an important parameter in a biopharmaceutical supply chain and we should be able to pull the statistics straight out of the ERP system fairly easily. But in the case study supply chain, the ERP system only records the movement of each material type using a material movement table. This table has two fields: movement type and movement quantity. Movement type indicates the destination of the movement and the movement quantity specifies how many are moved to the target location. Given the complete material movement history, we would be able to calculate the cycle time for each stage. But this calculation is not part of the ERP system and the cycle time statistics are not stored in the database either. To evaluate the cycle time performance, supply chain planners analyze the material movement periodically to get a rough estimate of the cycle time.

This discrepancy between what supply chain planning needs and what's available in the IT system creates difficulties for accurate planning. Another obstacle for better planning is supply chain planners' lack of IT skills. In the biopharmaceutical industry, the planning relies heavily on manual work using Excel spreadsheets. More advanced IT skills such as SQL and script language for data retrieving and processing are beyond the capability of most planners. Using the same example mentioned previously, to find out stage cycle times, we need to calculate based on the historical movements for each batch. This requires SQL and heavy scripting and the exact calculation was not explored until the initiation of a "planning engine" project between carried out by UC Berkeley and Bayer.

The "planning engine" project is one of the first attempts to bring automated planning into the biopharmaceutical industry. Not only the project introduces large scale integer optimization into the biopharmaceutical supply chain management, also it is built upon a software platform not for activities recording but rather for optimal planning and risk analytics. At the same time, supply chain planners are being trained to equip themselves with more advanced IT skills to get accustomed to the new platform. The success of the project demonstrated the importance of planning based IT infrastructure and the necessity of planners' advanced IT skills.

4.5.3 Conditional Release

A critical problem associated with biopharmaceutical supply chains is long cycle time. The cycle time of each stage includes processing time and quality assurance time. The processing time is relatively short, typically a few days. But the quality assurance cycle time varies from a few weeks to more than a year. Each quality assurance on a typical batch consists of a series of quality tests. And these tests don't have to be organized in a sequential manner. To reduce overall quality assurance cycle time, a process strategy called conditional release is employed in the biopharmaceutical supply chain. To illustrate the idea of conditional release, let's assume that we have a two-stage system where stage 2 is the upstream stage and stage 1 is the downstream stage. Conditional release separates the quality assurance testing of stage 2 into two phases. Phase I release occurs when the batch passes a minimum set of tests. The batch will then be released as input to stage 1. To increase material concentration, it's common to blend different stage 2 output batches to produce stage 1 materials. In parallel with stage 1 production and quality assurance, testing on the batch from stage 2 continues with all the remaining tests. And the phase II release occurs when the batch passes all the stage 2 tests. If the batch fails some of the tests, then phase II results will confirm the rejection of the batch.

Because we are overlapping the quality assurance testing of two stages, apparently we can reduce the overall cycle time. But there is potential downside with this approach. Since we release the stage 2 batch before confirming it's valid, it's possible that we are mixing "good" batches with "bad" batches. If phase II tests turn out to mandate rejection of the batch then not only the batch itself has to be discarded but also the stage 1 batch produced from the rejected batch. Essentially this policy sacrifices yield for shorter cycle time.

So an important research question is to evaluate whether or not the conditional release is a good policy. The performance of this policy depends on many variables including batch rejections, stochastic lead time, and batch mixing. If the biopharmaceutical process is still in the early stages of the technological learning curve, batch rejection may be high and conditional release may incur significant yield loss. On the other hand if the batch rejection is under control, conditional release may reduce overall cycle time significantly.

4.5.4 Stock Replenishment Planning - Rotation Cycle

As the goal of biopharmaceutical supply chain management is to maintain close to 100% service level, supply chain planners invest heavily in production capacity and always try to keep a significant redundant capacity in the supply chain. Most biopharmaceutical products are recently developed and under patent protection. Also the market is less competitive compared with tradition pharmaceutical products. Due to these reasons, manufactures usually charge a high premium for the products. Compared to potential cost of lost sales due to backorder, the cost of capacity expansion is much less significant. For example, at one biopharmaceutical production site, around \$350 million investment was made over the past few years to expand capacity. But compared with the \$1.2 billion annual sales worldwide

of the products of this site, the cost of capacity expansion is well justified. In practice it's not surprising to find a biopharmaceutical supply chain running at less than 50% of its total available capacity.

But as patent expiry approaches and more competitors enter the biopharmaceutical market, a decreasing selling price of biopharmaceuticals can be anticipated. Also as more companies join the biopharmaceutical business, the cost of raw material and facility expansion will go up. Due to the decreasing revenue and increasing cost, biopharmaceutical supply chain planners will need to utilize as much existing capacity as possible and at the same time maintain the required stock level proposed in this paper to cope with various risks in the supply chain.

Leachman and Gascon (1988) proposed a heuristic scheduling policy for the multi-item single machine production system with time varying demands. To illustrate their basic idea, following uses the same notation as their original paper. Given N items to replenish and let F_i denote the time to complete the replenishment for item i , $i = 1, \dots, N$. Let RO_i denote the inventory run-out time for item i . Then feasibility of the rotation cycle can be checked by enforcing following for all i

$$F_i \leq RO_{i+1}$$

When capacity is tight, then $F_i > RO_{i+1}$ for some i and the production system has negative slack. The key idea of the Dynamic Cycle Length Heuristic is that when there is negative slack, the cycle lengths of all items needs to be scaled back to eliminate negative slack. Let T^* denote the fundamental cycle length of the economic rotation cycle and let $T_i^* = k_i T^*$ denote the target production cycle length for item i . When there is negative slack in the system, to make the rotation cycle feasible let's assume we scale the fundamental cycle time by α . Then the maximum inventory level of item i is

$$S_i(\alpha) = ss_i + \frac{\alpha T_i^* D_i (P_i - D_i)}{P_i}$$

Assume the item index represents their replenishment order, then replenishment of item i will complete at time

$$F_i(\alpha) = F_{i-1}(\alpha) + c_i + \frac{S_i(\alpha) - [I_i - F_{i-1}(\alpha)D_i - c_i D_i]}{P_i - D_i}$$

We need to find largest nonnegative $0 \leq \alpha \leq 1$ such that

$$F_i(\alpha) \leq RO_{i+1}$$

for all $i = 1, \dots, N$ where the run out time is determined by the current inventory level

$$RO_i = \frac{I_i - ss_i}{D_i} - c_i$$

In biopharmaceutical supply chains the production process is batch production and only one batch is in one facility at one time to avoid any possible cross contamination. Also it is typical that one facility or machine supports the production of multiple product types. So to efficiently utilize capacity and replenish stock in the supply chain, we need to embed an economic rotation cycle into the facility schedule. Our rotation cycle should guarantee sufficient supply to avoid stock outs in all product types. Compared to the model proposed by Leachman and Gascon (1988), biopharmaceutical rotation cycle scheduling poses a few additional challenges. First not only there is stochastic demand but also there is a long and variable lead time. We can not guarantee batches to be released as inventory by certain dates. Second, the process is batch production. So we need to consider the effect of discrete jump of inventory instead of smooth continuous increasing in holding. Third, in addition to the stochastic lead time risk, there is also batch reject and excursion risk we need to take into account when scheduling the rotation cycle. This remains an open research challenge.

Bibliography

- [1] E. Arkin, D. Joneja and R. Roundy, Computational complexity of uncapacitated multi-echelon production planning problems, *Operations Research Letters* 8 (1989), pp. 61C66.
- [2] Bass, F. M. (1969). A New Product Growth Model for Consumer Durables. *Management Science*, 15(5), 215-227.
- [3] Ben-Tal, A., A. Nemirovski (1995a). Robust solutions to uncertain linear problems, Research Report # 6/95, Optimization Laboratory, Faculty of Industrial Engineering and Management, Technion, the Israel Institute of Technology, Technion City, Haifa 32000, Israel. BHC World, June 2008
- [4] J.H. Bookbinder and M. Cakanyildirim, Random lead time and expedited orders in (Q, R) inventory systems, *Eur J Oper Res* 115 (1999), 300C313.
- [5] Boyd SP, Vandenberghe L, *Convex Optimization*, 2004, Cambridge University Press
- [6] Bradley, J.R., Robinson, L.W., 2005, Improved Base-stock approximations for independent stochastic lead times with order crossover, *Manufacturing and Service Operations management*, vol. 7, no. 4, 319-329.
- [7] Buzacott, J.A. and J.G. Shanthikumar (1994), Safety Stock versus Safety Time in MRP Controlled Production Systems, *Management Science*, Vol. 40, pp. 1678-1689.
- [8] Chang C.A., The interchangeability of safety stocks and safety lead time, *Journal of Operations Management* 6 (1) (1985) 35-42.
- [9] Christopher M. and Lee H., Mitigating supply chain risk through improved confidence, *International Journal of Physical Distribution & Logistics Management* 34 (5) (2004), 388-396.
- [10] Chu, L. Y., J. G. Shanthikumar and Z. J. M. Shen (2008), Solving Operational Statistics via a Bayesian Analysis, *Operations research letters*, 36 (2008) 110-116.
- [11] Clark, A. J., "The Use of Simulation to Evaluate a Multiechelon, Dynamic Inventory Model," *Naval Research Logistics Quart.*, 7 (1960), 429-446.
- [12] CLARK, A., "An Informal Survey of Multi-Echelon Inventory Theory," *Naval Res. Logist. Quart.*, 19 (1972), 621-650.

- [13] Clark, A. and H. Scarf, "Optimal Policies for a Multi-Echelon Inventory Problem," *Management Sci.*, 6 (1960), 465-490.
- [14] DEBODT, M., AND S. C. GRAVES. 1985. Continuous Review Policies for a Multi-Echelon Inventory Problem With Stochastic Demand. *Mgmt. Sci.*31, 1286-1299.
- [15] Denardo E.V. and Lee T.Y.S., Managing uncertainty in a serial production line, *Operations Research* 44 (1996) 382-392.
- [16] Dessouky M, Leachman R (1997) Dynamic models of production with multiple operations and general processing times. *Journal of the Operational Research Society* 48: 647-654
- [17] Eppen, G. and R. Martin, Determining Safety Stock in the Presence of Stochastic Lead Time and Demand, *Management Science.*, 34 (1988), 1380-1390.
- [18] Federgruen, A., P. Zipkin. 1984. Computational issues in an infinite horizon, multi-echelon inventory model. *Oper. Res.* 32818-836.
- [19] Gascon, Andre, Robert C. Leachman and Arthur A. De Guia, 1993. "Optimal Planning and Control of Consumer Products Packaging Lines," in *Optimization in Industry*, Tito A. Ciriani and Robert C. Leachman, eds., p. 173-189, John Wiley & Sons, Ltd., Sussex, England.
- [20] Glasserman, P. and Tayur, S., 1995. Sensitivity analysis for base-stock levels in multi-echelon production-inventory systems. *Management Science* 41, pp. 263C281.
- [21] Goncalves, Jose, R. C. Leachman, A. Gascon and Z. Xiong, 1994. "An Heuristic Scheduling Policy for Multi-Item, Multi-Machine Scheduling Systems with Time-Varying, Stochastic Demands," *Management Science*, 40 (11), p. 1455-1468 (November, 1994).
- [22] Graves, S.1985. A multi-echelon inventory model for a repairable item with one-for-one replenishment. *Management Sci.*311247-1256.
- [23] Hayya, J.C., Bagchi, U., Kim, J. G., Sun, D., 2008, On static stochastic order crossover, *Int. J. Production Economics*, 114 (2008) 404-413.
- [24] He, X.X., Xu, S.H., Ord, J.K. and Hayya, J.C., 1998. An inventory model with order crossover. *Operations Research* 46, pp. 112-119.
- [25] van Houtum, G. J. and W. H. M. Zijm, "Computational Procedures for Stochastic Multi-echelon Systems," *International J. Production Economics*, (1991), 223-237.
- [26] Johnston L., Qualifying exam report, 2006, Dept. of Industrial Engineering and Operations Research, University of California at Berkeley
- [27] Johnston L., Hedging against supply-side operational risk in supply chain networks, 2008, Dept. of Industrial Engineering and Operations Research, University of California at Berkeley

- [28] Juttner U., Peck H., and Christopher M., Supply chain risk management: Outlining an agenda for future research. *International Journal of Logistics: Research and Applications* 6(4): 197-210, 2003.
- [29] Langenhoff, L. J. G. and W. H. M. Zijm, "An Analytical Theory of Multi-Echelon Production / Distribution Systems," *Statistica Neerlandica*, 44, 3 (1990), 149-174.
- [30] Leachman, Robert C. and A. Gascon, 1988. "A Heuristic Scheduling Policy for Multi-Item, Single Machine Production Systems with Time-Varying, Stochastic Demands," *Management Science*, 34 (3), p. 377-390 (March, 1988).
- [31] Leachman, Robert C., Z. K. Xiong, A. Gascon and K. Park, 1991. "An Improvement to the Dynamic Cycle Lengths Heuristic for Scheduling the Multi-Item, Single Machine," *Management Science*, 37 (9), p. 1201-1205 (September, 1991).
- [32] Leachman R (1993) Modeling Techniques for Automated Production Planning in the Semiconductor Industry. In *Optimization in Industry*, Ciriani T, Leachman R (eds).
- [33] Leachman R., L. Johnston, Shen Z. J., and Li S., An automated planning engine for biopharmaceutical production, working paper, Dept. of Industrial Engineering and Operations Research, University of California at Berkeley, 2008
- [34] LIBERATORE, M. J., The EOQ Model under Stochastic Lead Times, *Operations Research*, 27 (1979), 391-39.
- [35] Liyanage L. H., J. G. Shanthikumar, A practical inventory control policy using operational statistics, *Operations Research Letters*, 33(4) (2005) 341-348.
- [36] Robinson, L. W., J. R. Bradley, L. J. Thomas. 2001. Consequences of order crossover under order-up-to policies. *Manufacturing Service Oper. Management* 3 175-188.
- [37] Roman T., Improvement of a planning engine for biotech production and a case-study planning analysis for a new product, 2008
- [38] Rong, Ying, Bulut, Zumbul and Snyder, Lawrence V., Heuristics for Base-Stock Levels in Multi-Echelon Distribution Networks (September 18, 2009). Available at SSRN: <http://ssrn.com/abstract=1475469>
- [39] Sahinidis, N. V. (2003). Optimization under uncertainty: State-of-the-art and opportunities. In I. E. Grossmann, & C. M. McDonald (Eds.), *Proceedings of FOCAPO2003*, Coral Springs.
- [40] SCHWARZ, L. B., AND L. SCHRAGE. 1975. Optimal and System Myopic Policies for Multi-Echelon Production/Inventory Assembly Systems. *Mgmt. Sci.* 21, 1285- 1294.
- [41] Shah, N. Process industry supply chains: Advances and challenges. *Computers and Chemical Engineering*. 2005, 29, 1225-1235.

- [42] Shah, N. (2003). Pharmaceutical supply chains: Key issues and strategies for optimisation. In Proceedings of the fourth international conference on foundations of computer-aided process operations (pp. 73-85).
- [43] SHERBROOKE, C. C. 1968. METRIC: A Multi-Echelon Technique for Recoverable Item Control. *Opns. Res.* 16, 122-141.
- [44] SIMON, R. M., "Stationary Properties of a Two Echelon Inventory Model for Low Demand Items," *Oper. Res.*, 19, 3 (1981), 761-777.
- [45] Snyder, L. V. and Z.-J. M. Shen. Supply chain management under the threat of disruptions. *The Bridge* (National Academy of Engineering), 36(4), 39-45, Winter 2006
- [46] Snyder, L. V. and Z.-J. M. Shen. Supply and Demand Uncertainty in Multi-Echelon Supply Chains. 2006
- [47] Song, J., The Effect of Leadtime Uncertainty in a Simple Stochastic Inventory Model, *Management Science*, 40 (1994a), 603-613.
- [48] Stefansson, H., Jensson, P., Shah N, 2006 AIChE Annual Meeting, 2006
- [49] Tang, C.S., 1990. The Impact of uncertainty on a production line. *Management Science* 36, 1518-1532.
- [50] Yano, C.A. (1987), Setting planned lead times in serial production systems with tardiness costs, *Management Science* 33, 95-106.
- [51] Yano C. A., H. L. Lee. 1995. Lot-sizing and random yields: A review. *Operations Research* 43(3) 311C334.
- [52] Zalkind D. (1978) Order level inventory system with independent stochastic lead times. *Management Science*, 24, 1384-1392.
- [53] Zipkin, P. 1986. Models for Design and Control of Stochastic, Multi-Item Batch Production Systems. *Operations Research*, 34, 91-104.

Appendix A

Calculation of Required Stock Level under Geometric Lead Time and Geometric Excursion Occurrence

Here is a list of notation for the calculation.

N – Number of outstanding orders. An outstanding order is defined as an order placed but not arrived as planned

SF – Shortfall. Shortfall is defined as the gap between cumulative supply and cumulative demand

SF_R – Shortfall at the time of excursion recovery

SF_{RH} – Shortfall at the time of excursion recovery when the excursion happens within our planning horizon

F_i – Probability that order lead time is less than i days, $F_i = P(L \leq i) = 1 - (1 - p)^i$

L – Order lead time

μ_L – Average order lead time

σ_L – Standard deviation of order lead time

D_i – Order quantity in i^{th} period

μ_D – Average order quantity in one period

σ_D – Standard deviation of order quantity in one period

R_i – Equal to one if order placed at time i is rejected, 0 otherwise

r – Probability of order being rejected

T – Time to excursion occurrence

$G_i - G_i = P(T \leq i) = 1 - (1 - q)^i$, probability that time to excursion is less than i days

s – Time from excursion occurrence to excursion recovery

H – Planning horizon, multiples of ordering interval

To calculate the required stock level, we need to find the mean and variance of the shortfall. Based on Equation 3.6, we can calculate mean shortfall

$$\begin{aligned}
\mathbb{E}[SF_{RH}] &= s\mu_D \left(1 - (1 - q)^H\right) + \mu_D \sum_{i=1}^H \left(1 - (1 - r) \left(1 - (1 - p)^{s+i}\right)\right) \left((1 - q)^{i-1} - (1 - q)^H\right) \\
&= s\mu_D \left(1 - (1 - q)^H\right) + \mu_D \sum_{i=1}^H \left(r + (1 - r)(1 - p)^{s+i}\right) \left((1 - q)^{i-1} - (1 - q)^H\right) \\
&= s\mu_D \left(1 - (1 - q)^H\right) - \mu_D H r (1 - q)^H - \frac{\mu_D (1 - r) (1 - q)^H (1 - p)^{s+1} (1 - (1 - p)^H)}{p} \\
&\quad + \frac{\mu_D r (1 - (1 - q)^H)}{q} + \frac{\mu_D (1 - r) (1 - p)^{s+1} \left(1 - (1 - p)^H (1 - q)^H\right)}{1 - (1 - p) (1 - q)}
\end{aligned}$$

To calculate $Var[SF_{RH}]$, first we calculate $E[SF_{RH}^2|T]$. Then summing over possible excursion occurrence time we can find $E[SF_{RH}^2|T]$. Given an excursion happens at time T , we can find $E[SF_{RH}^2|T]$:

$$\begin{aligned}
E[SF_{RH}^2|T] &= E[E[(SF_{RH}^2|T)|N]] \\
&= E\left[E\left[\left(\sum_{i=1}^N D_i\right)^2\right]\right] \\
&= E\left[E\left[\sum_{i=1}^N D_i^2 + \sum_{i < j} 2D_i D_j\right]\right] \\
&= E[N E[D_i^2] + N(N-1) E[D_i D_j]] \\
&= E[N(\sigma_D^2 + \mu_D^2) + N(N-1)\mu_D^2] \\
&= \sigma_D^2 E[N] + \mu_D^2 E[N^2] \\
&= \sigma_D^2 E[N] + \mu_D^2 (E[N])^2 + \mu_D^2 Var(N) \\
&= \sigma_D^2 \left(s + \sum_{i=1}^T [1 - (1 - r)F_{s+i}]\right) + \mu_D^2 \left(s + \sum_{i=1}^T [1 - (1 - r)F_{s+i}]\right)^2 \\
&\quad + \mu_D^2 \sum_{i=1}^T ((1 - r)F_{s+i} [1 - (1 - r)F_{s+i}])
\end{aligned}$$

Summing over all possible time of excursion within our planning horizon, then we can calculate $E [SF_{RH}^2]$:

$$\begin{aligned} E [SF_{RH}^2] &= E[E[SF_{RH}^2|T]] \\ &= \sum_{t=1}^H (\sigma_D^2 \left(s + \sum_{i=1}^T [1 - (1-r) F_{s+i}] \right) + \mu_D^2 \left(s + \sum_{i=1}^T [1 - (1-r) F_{s+i}] \right)^2) \\ &\quad + \mu_D^2 \sum_{i=1}^T ((1-r) F_{s+i} [1 - (1-r) F_{s+i}]) P(T=t) \end{aligned}$$

We then split the calculation of $E [SF_{RH}^2]$ into three parts

$$E [SF_{RH}^2] = \sum_{t=1}^H (\sigma_D^2 \left(s + \sum_{i=1}^T [1 - (1-r) F_{s+i}] \right)) P(T=t) \quad (A.1)$$

$$+ \sum_{t=1}^H (\mu_D^2 \left(s + \sum_{i=1}^T [1 - (1-r) F_{s+i}] \right)^2) P(T=t) \quad (A.2)$$

$$+ \sum_{t=1}^H (\mu_D^2 \sum_{i=1}^T ((1-r) F_{s+i} [1 - (1-r) F_{s+i}])) P(T=t) \quad (A.3)$$

The first part of $E [SF_{RH}^2]$ is

$$\begin{aligned} (A.1) &= \sum_{t=1}^H (\sigma_D^2 \left(s + \sum_{i=1}^T [1 - (1-r) F_{s+i}] \right)) P(T=t) \\ &= \sigma_D^2 \sum_{i=1}^H \left(s + \sum_{i=1}^t (1 - (1-r)(1 - (1-p)^{s+i})) \right) (1-q)^{t-1} q \\ &= q\sigma_D^2 \sum_{i=1}^H \left(s + rt + \frac{(1-r)(1-p)^{s+1}(1 - (1-p)^t)}{p} \right) (1-q)^{t-1} \\ &= q\sigma_D^2 \sum_{i=1}^H \left(s + rt + \frac{(1-r)(1-p)^{s+1}}{p} \right) (1-q)^{t-1} - q\sigma_D^2 \sum_{i=1}^H \left(\frac{(1-r)(1-p)^{s+1+t}}{p} \right) (1-q)^{t-1} \\ &= q\sigma_D^2 \left(s + rt + \frac{(1-r)(1-p)^{s+1}}{p} \right) \frac{1 - (1-q)^H}{q} - \frac{q\sigma_D^2(1-r)}{p} \frac{(1-p)^{s+2}(1 - (1-p)^H(1-q)^H)}{1 - (1-p)(1-q)} \end{aligned}$$

The second part of $E [SF_{RH}^2]$ is

$$\begin{aligned}
(A.2) &= \sum_{t=1}^H (\mu_D^2 \left(s + \sum_{i=1}^T [1 - (1-r)F_{s+i}] \right)^2) P(T=t) \\
&= q\mu_D^2 \sum_{t=1}^H \left(s + rt + \frac{1-r}{p} (1-p)^{s+1} (1 - (1-p)^t) \right)^2 (1-q)^{t-1} \\
&= q\mu_D^2 \sum_{t=1}^H \left(s + \frac{(1-r)(1-p)^{s+1}}{p} + rt - \frac{(1-r)(1-p)^{s+1+t}}{p} \right)^2 (1-q)^{t-1} \\
&= q\mu_D^2 \sum_{t=1}^H \left(s + \frac{(1-r)(1-p)^{s+1}}{p} \right)^2 (1-q)^{t-1} \tag{A.4}
\end{aligned}$$

$$+ q\mu_D^2 \sum_{t=1}^H 2 \left(s + \frac{(1-r)(1-p)^{s+1}}{p} \right) \left(rt - \frac{(1-r)(1-p)^{s+1+t}}{p} \right) (1-q)^{t-1} \tag{A.5}$$

$$+ q\mu_D^2 \sum_{t=1}^H \left(rt - \frac{(1-r)(1-p)^{s+1+t}}{p} \right)^2 (1-q)^{t-1} \tag{A.6}$$

The second part of $E [SF_{RH}^2]$ is further decomposed into three parts as follows.

$$\begin{aligned}
(A.4) &= q\mu_D^2 \sum_{t=1}^H \left(s + \frac{(1-r)(1-p)^{s+1}}{p} \right)^2 (1-q)^{t-1} \\
&= \mu_D^2 \left(s + \frac{(1-r)(1-p)^{s+1}}{p} \right)^2 (1 - (1-q)^H)
\end{aligned}$$

$$\begin{aligned}
(A.5) &= q\mu_D^2 \sum_{t=1}^H 2 \left(s + \frac{(1-r)(1-p)^{s+1}}{p} \right) \left(rt - \frac{(1-r)(1-p)^{s+1+t}}{p} \right) (1-q)^{t-1} \\
&= 2q\mu_D^2 \left(s + \frac{(1-r)(1-p)^{s+1}}{p} \right) \left(\frac{rH(H+1)}{2} - \frac{(1-r)(1-p)^{s+2}(1 - (1-p)^H)}{p^2} \right)
\end{aligned}$$

$$\begin{aligned}
(A.6) &= q\mu_D^2 \sum_{t=1}^H \left(rt - \frac{(1-r)(1-p)^{s+1+t}}{p} \right)^2 (1-q)^{t-1} \\
&= q\mu_D^2 \sum_{t=1}^H \left(r^2t^2 - \frac{2rt(1-r)(1-p)^{s+1+t}}{p} + \frac{(1-r)^2(1-p)^{2s+2+2t}}{p^2} \right) (1-q)^{t-1} \\
&= q\mu_D^2 \sum_{t=1}^H (r^2t^2) (1-q)^{t-1} \tag{A.7}
\end{aligned}$$

$$- q\mu_D^2 \sum_{t=1}^H \left(\frac{2rt(1-r)(1-p)^{s+1+t}}{p} \right) (1-q)^{t-1} \tag{A.8}$$

$$+ q\mu_D^2 \sum_{t=1}^H \left(\frac{(1-r)^2(1-p)^{2s+2+2t}}{p^2} \right) (1-q)^{t-1} \tag{A.9}$$

$$\begin{aligned}
(A.7) &= q\mu_D^2 \sum_{t=1}^H (r^2t^2) (1-q)^{t-1} \\
&= r^2q\mu_D^2 \left(\sum_{t=1}^H (t+1)t(1-q)^{t-1} - \sum_{t=1}^H t(1-q)^{t-1} \right) \\
&= r^2q\mu_D^2 \left(\left(\sum_{t=1}^H A^H \right)'' - \left(\sum_{t=1}^H A^{t-1} \right)' \right), A=1-q \\
&= r^2q\mu_D^2 \left(\frac{-2 + (H+2)(H+1)A^H - 2H(H+2)A^{H+1} + H(H+1)A^{H+2}}{(A-1)^3} \right. \\
&\quad \left. - \frac{1 - (H+1)A^H + HA^{H+1}}{(1-A)^2} \right), A=1-q
\end{aligned}$$

$$\begin{aligned}
(A.8) &= q\mu_D^2 \sum_{t=1}^H \left(\frac{2rt(1-r)(1-p)^{s+1+t}}{p} \right) (1-q)^{t-1} \\
&= \frac{2q\mu_D^2 r(1-r)}{p} \sum_{t=1}^H t(1-p)^{s+1+t} (1-q)^{t-1} \\
&= \frac{2q\mu_D^2 r(1-r)}{p} \sum_{t=1}^H t(1-p)^{s+1+t} (1-q)^{t-1} \\
&= \frac{2q\mu_D^2 r(1-r)(1-p)^{s+2}}{p} \left(\sum_{t=1}^H B^t \right), B = (1-p)(1-q) \\
&= \frac{2q\mu_D^2 r(1-r)(1-p)^{s+2}}{p} \frac{1 - (H+1)B^H + HB^{H+1}}{(1-B)^2}, B = (1-p)(1-q)
\end{aligned}$$

$$\begin{aligned}
(A.9) &= q\mu_D^2 \sum_{t=1}^H \frac{(1-r)^2 (1-p)^{2s+2+2t}}{p^2} (1-q)^{t-1} \\
&= \frac{q\mu_D^2 (1-r)^2}{p^2} \sum_{t=1}^H (1-p)^{2s+2+2t} (1-q)^{t-1} \\
&= \frac{q\mu_D^2 (1-r)^2 (1-p)^{2s+4}}{p^2} \frac{\left(1 - (1-p)^{2H} (1-q)^H \right)}{1 - (1-p)^2 (1-q)}
\end{aligned}$$

$$\begin{aligned}
(A.3) &= \sum_{t=1}^H (\mu_D^2 \sum_{i=1}^T ((1-r)F_{s+i} [1 - (1-r)F_{s+i}])) P(T=t) \\
&= \mu_D^2 \sum_{t=1}^H \left(\sum_{i=1}^T \left((1-r)(1 - (1-p)^{s+i}) \left[1 - (1-r)(1 - (1-p)^{s+i}) \right] \right) \right) (1-q)^{t-1} q \\
&= \mu_D^2 q r (1-r) \sum_{t=1}^H t (1-q)^{t-1} \tag{A.10}
\end{aligned}$$

$$+ \mu_D^2 q (1-r) \sum_{t=1}^H \frac{(1-2r)(1-p)^s}{p} (1 - (1-p)^t) (1-q)^{t-1} \tag{A.11}$$

$$- \mu_D^2 q (1-r) \sum_{t=1}^H \frac{(1-r)(1-p)^{2s+2}}{1 - (1-p)^2} (1 - (1-p)^{2t}) (1-q)^{t-1} \tag{A.12}$$

$$\begin{aligned}
(A.10) &= \mu_D^2 q r (1-r) \sum_{t=1}^H t (1-q)^{t-1} \\
&= \mu_D^2 q r (1-r) \left(\sum_{t=1}^H A^t \right)', A=1-q \\
&= \frac{\mu_D^2 r (1-r)}{q} \left(1 - (H+1)(1-q)^H + H(1-q)^{H+1} \right), A=1-q
\end{aligned}$$

$$\begin{aligned}
(A.11) &= \mu_D^2 q (1-r) \sum_{t=1}^H \frac{(1-2r)(1-p)^s}{p} (1 - (1-p)^t) (1-q)^{t-1} \\
&= \frac{\mu_D^2 q (1-r) (1-2r) (1-p)^s}{p} \left(\frac{1 - (1-q)^H}{q} - \frac{(1-p)(1 - (1-p)^H (1-q)^H)}{1 - (1-p)(1-q)} \right)
\end{aligned}$$

$$\begin{aligned}
(A.12) &= \mu_D^2 q (1-r) \sum_{t=1}^H \frac{(1-r)(1-p)^{2s+2}}{1 - (1-p)^2} (1 - (1-p)^{2t}) (1-q)^{t-1} \\
&= \frac{\mu_D^2 q (1-r)^2 (1-p)^{2s+2}}{1 - (1-p)^2} \left(\frac{1 - (1-q)^H}{q} - \frac{(1-p)^2 (1 - (1-p)^{2H} (1-q)^H)}{1 - (1-p)^2 (1-q)} \right)
\end{aligned}$$

Here we have calculated all the parts for a closed form of expression $E [SF_{RH}^2]$. Based on the variance formula, we can thereby calculate $Var [SF_{RH}]$.