<< 2023-01-19 | 2023-01-21 >>

> *Decision is a risk rooted in the courage of being free.*
> *— Paul Tillich*

> *Lectures:*
>
> - *#博士生资格考试资料*

1. Introduction

1.1 The Evolution of Supply Chain Theory

> *ried out may reduce their cost. Supply chains used to be viewed,at least by some managers, as "necessary evils." As a result, the mindset for su*
> *show annotation*

> *kouts,or other degraded service. By the end of the last century, however, the purpose of the supply chain had begunto change as some firms discovered that supply chains could be a source of competitiveadvantage, rather than simply a cost driver. For example, Dell demonstrated*
> *show annotation*

> *ications of supplychain theory. The final chapter of this book is devoted to exploring how the tools of supplychain theory are used in a few of these application areas—electricity systems, health care,and public sector operations. 1.2 DEFINITIONS AND SCOPEThe ter*
> *show annotation*

1.2 Definiton and Scope

> *of supply chains have evolved. Perhaps the mostauthoritative definition comes from the Council of Supply Chain Management Professionals(CSCMP), who define supply chain management as follows*

> :Supply chain management encompa
> show annotation

> ply chain management as follows: Supply chain management encompasses the planning and management of all activitiesinvolved in sourcing and procurement, conversion, and all logistics management activi-ties. Importantly, it also includes coordination and collaboration with channel partners,which can be suppliers, intermediaries, third party service providers, and customers. Inessence, supply chain management integrates supply and demand management withinand across companies . (Council of Supply Chain Manag
> show annotation

> anagement Professionals 2018a). These practices include a huge range of tasks, suchas forecasting, production planning, inventory management, warehouse location, supplierselection, procurement, and shipping. Mathematical models have been d
> show annotation

> twork.Figure 1.2 Supply "chain." The terms "logistics" and "logistics management" are closely related to "supply chainmanagement," and it can be difficult to draw a clear distinction. Some companies use "lo-gistics"
> show annotation

> the physical movement of goods; "supply chain management" includeslogistics, as well as nonmovement activities such as inventory management and procure-ment. For other companies, "logistics
> show annotation

> ed more or less interchangeably. Supply chains are often represented graphically as a schematic network that illustratesthe relationships between its elements . (See Figure 1.1.) Each vertica
> show annotation

> ts, etc.) is called an echelon. A location in the network is referred toas a stage or node. The links between stages represent some type of flow—typically,

> *theflow of goods, but sometimes the flow of information or money . The portion of the supplychain*
> [show annotation]

> *e flow of information or money. The portion of the supplychain from which products originate (the left-hand portion in Figure 1.1) is referred to asupstream, while the demand end is referred to as downstream. Actually, the phrase "supply cha*
> [show annotation]

> *chelon has only a single stage. But today's supply chains more closelyresemble the complex network in Figure 1.1; each echelon may have dozens, hundreds,or even thousands of nodes. (Nevertheless, we will often st*
> [show annotation]

> *t of the supply chainas a whole. The ideal supply chain management model would globally optimize every aspect of thesupply chain, but such a model is impossible both because of the difficulties in modelingsome aspects of the supply chain mathematically and because the resulting model wouldbe too large and complex to solve . Instead, supply chain models t*
> [show annotation]

1.3 Levels of Decision-making in Supply Chain Management

> *AKING IN SUPPLY CHAIN MANAGEMENT It is convenient to think about three levels of supply chain management decisions: strategic,tactical, and operational. • Strategic aspects of the suppl*
> [show annotation]

Three level of supply chain management decisions:

- strategic
- tactical

- operational

2. Forecasting and Demad Modeling

2.1 Introduction

> *DEMAND MODELING2.1 INTRODUCTION Demand forecasting is one of the most fundamental tasks that a business must perform .It can be a significant source*
> *show annotation*

Advantage:

- Improving customer service levels and by reducing costs related to supply–demand mismatches.
  Disadvantage
- biased or otherwise inaccurate forecasting results in inferior decisions and thus undermines business performance.

> *and lower revenue (Ziobro 2016). The goal of the forecasting models discussed in this chapter is to estimate the quantityof a product or service that consumers will purchase . Most classical forecasting tec*
> *show annotation*

> *e that consumers will purchase. Most classical forecasting techniquesinvolve time-series methods that require substantial historical data. Some of these methodsare designed for demands that are stable over time . Others can handle demands that*
> *show annotation*

> *ds to be stable and predictable. However, products today have shorter and shorter life cycles, in part driven by rapidtechnology upgrades for high-tech products. As a result, firms have much less historicaldata available to use for forecasting, and any trends that may be evident in historical datamay be unreliable for predicting the future .5Fundamentals of Supply Chain T*
> *show annotation*

- Large quantities of historical data available.

> *inSections 2.2 and 2.3. Next, i n Section 2.4, we discuss more recent approaches to forecastingdemand using machine learning when we have large quantities of historical data available .In Sections 2.5–2.8, we discuss*
> [show annotation]
>
> *Large quantities of historical data available.*

- Inadequate historical data

> *s of historical data available.I n Sections 2.5–2.8, we discuss several methods that can be used to predict demands for newproducts or products that do not have much historical data. To distinguish these methodsfro*
> [show annotation]

> *not have much historical data. To distinguish these methodsfrom classical time-series–based methods, we call them demand modeling techniques. The methods that we discuss in t*
> [show annotation]

Demand modeling techniques:

- Quantitative

> *this chapter are quantitative. They all involve mathe-matical models with parameters that must be calibrated. In contrast, some popular metho*
> [show annotation]

- Some popular methods

> *meters that must be calibrated. In contrast, some popular methodsfor forecasting demand with little or no historical data, such as the Delphi method, rely onexperts' qualitative assessments or questionnaires to develop forecasts. Demand processes may exhibit var*
> [show annotation]

Demand processes may exhibit various forms of nonstationarity over time. These include the following:

- *Trends*: Demand consistently increases or decreases over time.
- *Seasonality*: Demand shows peaks and valleys at consistent intervals.
- *Product life cycles*: Demand goes through phases of rapid growth, maturity, and decline.

> *id growth, maturity, anddecline. Moreover, demands exhibit random error—variations that cannot be explained or predicted—and this randomness is typically superimposed on any underlying nonstationarity .2.2 CLASSICAL DEMAND FORECASTIN*
> *show annotation*

superimpose: *vt.* 使重叠,使叠加

> *SICAL DEMAND FORECASTING METHODS Classical forecasting methods use prior demand history to generate a forecast. Some ofthe methods, such as movi*
> *show annotation*

> *(single) exponential smoothing, assume that pastpatterns of demand will continue into the future, that is, no trend is present. As a result,these techniques are*
> *show annotation*

As a result, these techniques are best used for mature products with a large amount of historical data

> *large amount of historical data. On the other hand, regression analysis and double and triple exponential smoothing canaccount for a trend or other pattern in the data. We discuss each of these method*
> *show annotation*

2.2.1 Moving Average

> *period t −1.2.2.1 Moving Average* *The moving average method calculates the average amount of demand over a given intervalof time and uses this average to predict the future demand.* *As a result, moving averagefore*
> show annotation

The definition of Moving Average:

$$D_t = I + \epsilon_t$$

where I is the mean or "base" demand and ϵ is a random error term.

> *d and t is a random error term.* *A moving average forecast of order N uses the most recent observed demands.* *Theforecast for the demand in p*
> show annotation

> *0 32.90 48.9012 8.98 32.98 22.78* *That is, the forecast is simply the arithmetic mean of the previous N observations. This isknown as a simple moving average forecast of order N.* *A generalization of the simple m*
> show annotation

2.2.2 Exponential Smoothing

> *.80. 2.2.2 Exponential Smoothing* *Exponential smoothing is a technique that uses a weighted average of all past data as thebasis for the forecast.* *It gives more weight to recent*
> show annotation

Assumption:

- *Single exponential smoothing*: the demand process is *stationary*;
- *Double exponential smoothing*: there is a trend;
- *Triple exponential smoothing*: account for trends and seasonality.

> *or both trends and seasonality.* *These methods all requireuser-specified parameters that determine the relative weights placed on recent and olderobservations when predicting the demand, trend, and seasonality* *.*

> These three weights arecalled,
> show annotation

> demand, trend, and seasonality. These three weights arecalled, respectively, the smoothing factor, the trend factor, and the seasonality factor . Wediscuss each of these three
> show annotation

1. Single Exponential Smoothing

> 0αiDt−i−1,where αi = α(1 − α)i. The single exponential smoothing forecast includes all pastobservations, but since αi < αj for i > j, th
> show annotation

> α)i = 1by (C.50) in Appendix C. These weights can be approximated with an exponential function $f(i) = \alpha e^{-\alpha i}$. This is why this method is call
> show annotation

2. Double Exponential Smoothing

> .2 Double Exponential Smoothing Double exponential smoothing can beused to forecast demands with a linear trend. Such demands can be modeled as
> show annotation

> 2.8)can be explained similarly: It places a weight of β on the most recent estimate of the slope(obtained by taking the difference between the two most recent base signals) and a weight of $1 - \beta$ on the previous estimate . Note that, if the trend is dow
> show annotation

3. Triple Exponential Smoothing

> .3 Triple Exponential Smoothing Triple exponential smoothing can be usedto forecast demands that exhibit both trend and seasonality. Seasonality means that thedeman
> show annotation

Seasonality: the demand series has a pattern that repeats every N periods for some fixed N.

> *seasonal factor one season ago. The idea behind smoothing with trend and seasonality is basically to "de-trend" and "de-seasonalize" the time series by separating the base signal from the trend and seasonalityeffects . The method uses three smoothin*
> *show annotation*

> *s ago) using weighting factor γ. Initializing triple exponential smoothing is a bit trickier than for single or double expo-nential smoothing. To do so, we usually need at le*
> *show annotation*

> *double expo-nential smoothing. To do so, we usually need at least two entire seasons' worth of data($2N$ periods), which will be used for the initialization phase. One common method is toinitialize the slope as S2N = 1N(DN+1 −D1N + DN+2 −D2N +*
> *show annotation*

> *+2 −D2N + ···+ D2N −DNN). (2.14) In other words, we take the per-period increase in demand between periods 1 and $N+1$, and the per-period increase between periods 2 and $N+2$, and so on; and then we take the average over those N values. To initialize the seasonal fact*
> *show annotation*

2.2.3 Linear Regression

> *25.90. 2.2.3 Linear Regression Historical data can also be used to forecast demands by determining a cause–effect rela-tionship between some independent variables and the demand . For instance, the demandfor sa*
> *show annotation*

> *rice and a givenset of features. In linear regression, the model specification assumes that the dependent variable, Y , is alinear combination of the*

> independent variables. *For example, in simple linear r*
> show annotation

> *stimate theparameters β0 and β1.* To build a regression model, we need historical data points—observations of both theindependent variable(s) and the dependent variable. *Let (x1,y1),(x2,y2),..., (xn,yn*
> show annotation

2.3 Forecast Accuracy

2.3.1 MAD, MSE and MAPE

Measurements:

- MAD: Mean Absolute Deviation
- MSE: Mean Squared Error
- MAPE: Mean Absolute Percentage Error

$$\mathrm{MAD} = \frac{1}{n}\sum_{t=1}^{n}|e_t|$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{t=1}^{n}e_t^2$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{t-i}^{n}\left|\frac{e_t}{D_t}\right| \times 100$$

> *ean and variance, respectively.* If the mean of the forecast error, μe,equals 0, we say the forecasting method is unbiased: It does not produce forecasts that aresystematically either too low or too high. *However, even an unbiased forec*
> show annotation

> *denominator of the coefficient.* MAD is sometimes preferred to MSE in real applicationsbecause it avoids the calculation of squaring, though modern spreadsheet and statisticspackages can compute either performance measure easily. *When the forecast errors arenor*
> show annotation

2.4 Machine Learning in Demand Forecasting

2.4.1 Introduction

> *nWe are in the age of big data. The huge volume of data generated every day, the high velocityof data creation, and the large variety of sources all make today's business informationenvironment different than it was only a decade ago. Using data intelligently is key*
> *show annotation*

> *hniques for demand forecasting. Compared with classical forecastingmethods such as the time series methods discussed in Section 2.2, machine learning modelsoften significantly increase prediction accuracy. 2.4.2 Machine LearningIn general*
> *show annotation*

2.4.2 Machine Learning

> *4.2 Machine LearningIn general, machine learning (ML) refers to a set of algorithms that can learn from andmake predictions about data . These algorithms take data as*
> *show annotation*

> *veloped rapidly in recent years. Both techniques fall into the overall field of data science, which covers a wider range oftopics, including database design and data visualization techniques. One category of ML algorithms is*
> *show annotation*

> *Regression is a simpleexample. In contrast, unsupervised learning explores relationships and structures withinthe data without any known "ground truth" labels or outputs. For example, if we wish toparti*
> *show annotation*

> *her market information (inputs). Common supervised learning methods include linear regression (and its nonlinear ex-tensions), kernel methods,*

> tree-based models, support vector machines (SVMs), and neuralnetworks . Graphical models involving hid
> show annotation

> nes (SVMs), and neuralnetworks. Graphical models involving hidden Markov models (or, in their simplest form,mixture models) and Markov random fields also receive considerable attention. In thefollowing subsections, we
> show annotation

1. Linear regression
2. Tree-based models

> the outputs are cate-gorical). The trees used for these two types of problems are referred to as regression treesand classification trees, respectively. In demand forecasting, regression trees have receivedmore attention because of their simplicity and interpretability .A regression tree divides the s
> show annotation

> , similar to linear regression. However, in practice, the number of possible partitionsmay be too large to enumerate. Therefore, it is common to use a binary splitting methodcalled recursive partitioning, which generates two regions from the original region at eachiteration. For the purposes of prediction,
> show annotation
>
> 递归分割方法

> gh variance of the forecast, so researchershave developed methods that combine several trees to enhance the prediction performance.These include random forests, bagging, and boosting. Tree-based models are used widel
> show annotation

> velopregression trees to predict stock-keeping unit (SKU) sales for a European grocery retailer
> show annotation

> *n. 库存量单位*

3. Support vector machines

> *2.4.2.3 Support vector machines SVMs are designed to partition the space ofinput variables into two regions, i.e., to make a binary prediction about a given outputbased on which region a given input vector falls into. The partition is accomplished b*
> *show annotation*

> *nd make a predictionaccordingly. SVMs can be generalized to allow nonlinearities by mapping the input space into ahigh-dimensional space using kernel functions. In essence, this allows the reg*
> *show annotation*

> *ear, i.e., is not a hyperplane. Popular choices ofkernel functions include polynomials and radial basis functions (RBFs) .Since SVMs can be used to make*
> *show annotation*

4. Neural networks

> *blem.) 2.4.2.4 Neural Networks A neural network consists of several nodes, also calledneurons, arranged into layers. The first layer of nodes represe*
> *show annotation*

> *a (possibly nonlinear) function. The key challenge in fitting a neural network model is the determination of the weights α_{0m} and α_m. This is usually done using some sort of algorithm that modifies the weights as the network "learns" right and wrong answers. The most common such algorithm*
> *show annotation*

> *es the network harder to train. Suchdeep neural networks have led to huge advances in machine learning, with great successesnot only in classification and prediction problems such as image processing and demandforecasting,*

> but also, when coupled with reinforcement learning (RL), in solving decisionproblems such as those in board games ; one famous example is Google D
> show annotation

2.5 Demand Modeling Techniques

> closely as possible. To do so, they need tounderstand the life cycles and demand dynamics of their products .One of the authors has worked w
> show annotation

> make them work, unsuccessfully. It turns outthat classical forecasting techniques did not work well with the company's highly variable,short-life-cycle products, so the firm introduced products at the wrong times in the wrongquantities. The forecasting team's reaction
> show annotation

2.6 Bass Diffusion Model

> oducts.)2.6 BASS DIFFUSION MODEL The sales patterns of new products typically go through three phases: rapid growth, maturity,and decline . The Bass diffusion model (Bass
> show annotation

> or television sets in the 1960s. The premise of the Bass model is that customers can be classified into innovatorsand imitators. Innovators (or early adopters)
> show annotation

> increases, and then decreases. Thegoal of the Bass model is to characterize this behavior in an effort to forecast the demand .It mathematically characterizes
> show annotation

> se who have not yet adopted it. Moreover, it attempts topredict two important dimensions of a forecast: how many customers will eventually

> adoptthe new product, and when they will adopt . Knowing the timing of adoption
> show annotation

> duct, and when they will adopt. Knowing the timing of adoptions is importantas it can guide the firm to smartly utilize resources in marketing the new product. Ouranalysis of this model is ba
> show annotation

2.6.1 The Model

It assumes that $P(t)$, the probability that a given buyer makes an initial purchase at time t given that she has not yet made a purchase, is a linear function of the number of previous buyers, that is,

$$P(t) = p + \frac{q}{m} D(t)$$

where $D(t)$ is the cumulative demand by time t. q and m represent the coefficient of imitation and the market size.

This equation include two influence factors:

- coefficient of innovation, denoted p, which is a constant, independent of how many other customers have adopted the innovation before time t.
- the "contagion" effect between the innovators and the imitators, denoted $\frac{q}{m} D(t)$, which is proportional to the number of customers who have already adopted by time t.

> (t) in terms of its derivative. Our preference would be to have a closed-form expressionfor D(t). Fortunately, this is possible: Theorem 2.1D(t) =m 1 −e−(p+q)t1
> show annotation

> uct and willdecline thereafter. In summary, by varying the values of p and q, we can represent manydifferent patterns of demand diffusion . EXAMPLE

> 2.10The bookstore man
> show annotation

2.6.2 Discrete Time Version

> odel.2.6.2 Discrete-Time Version A discrete-time version of the Bass model is available. In this case, dt represents the
> show annotation

2.6.3 Parameter Estimation

> DEL 292.6.3 Parameter Estimation The Bass model is heavily driven by the parameters m, p, and q. In this section, we brieflydisc
> show annotation

> .51)p = am (2.52)q = −mc. (2.53) However, because the Bass model is typically used for new products, in most caseshistorical data are not available to estimate the parameters. Instead, m is typically esti-ma
> show annotation

> ble to estimate the parameters. Instead, m is typically esti-mated qualitatively, using judgment or intuition from management about the size of themarket, market research, or the Delphi method . In some markets these estimate
> show annotation
>
> 德尔菲法：专家调查法

> dailments (Lilien et al. 2007). The parameters p and q tend to be relatively consistent withina given industry, so these can often be estimated from the diffusion patterns of similarproducts. Lilien and Rangaswamy (1998) pr
> show annotation

2.6.4 Extensions

> verviews of these applications. The original model has also been extendedin a number of ways. Ho et al. (2002) provide a join
> show annotation

2.7 Learning Indicator Approach

> s.2.7 LEADING INDICATOR APPROACH Product life cycles are becoming shorter and shorter, so it is difficult to obtain enoughhistorical data to forecast demands accurately. One idea that has proven to wor
> show annotation

- Solution: leading indicators

> to forecast demands accurately. One idea that has proven to work well insuch situations is the use of leading indicators—products that can be used to predict thedemands of other, later products because the two products share a similar demand pattern .This approach was introduced by
> show annotation

> llular phones, or grocery items. The idea is first to group the products into clusters so that all of the products within a clustershare similar attributes. There are several ways to perfo
> show annotation

> of theproducts in the cluster. Even though all of the products are on the market simultaneously,the lag provides enough time so that supply chain planning for the products in the clustercan take place based on the forecasts provided by the leading indicator. Of course, correctlyidentifying the leading indicator is critical. Wu et al. (2006) suggest the fol
> show annotation

2.8 Discrete Choice Models

2.8.1 Introduction to Discrete Choice

> Introduction to Discrete Choice In economics, discrete choice models involve choices between two or more discrete alterna-tives . For example, a customer choose
> show annotation

> ce is usually more challenging.) The idea behind discrete choice models is to build a statistical model that predictsthe choice made by an individual based on the individual’s own attributes as well as theattributes of the available choices. For example, a student’s choice
> show annotation

> oice as the dependent variable, choice models are atthe aggregate (population) level and assume that each decision-maker’s preferences arecaptured implicitly by that model. At first, it may seem that discr
> show annotation

> y useful for forecasting demand. Discrete choice models take many forms, including binary and multinomial logit, binaryand multinomial probit, and conditional logit. However, there are several feat
> show annotation

> the decision-maker must choose. For a discrete choice model, the setof alternatives in the choice set must be mutually exclusive, exhaustive, and finite. Thefirst two requirements mean
> show annotation
>
> 互斥、全面、有限的

> I; this utility is denoted Uni. Discrete choice models usuallyassume that the decision-maker is a utility maximizer. That is, he will choose alternative iif and only if $U_{ni} > U_{nj}$ for all $j \in I, j \neq i$ .If we know the utility values U
> show annotation

2.8.2 The Multinomial Logit Model

> .8.2 The Multinomial Logit Model Next we derive the multinomial logit model. (Refer to McFadden (1974) or Train (2009)for further details of the derivation.) "Multinomial" means that there are multiple optionsfrom which the decision-maker chooses . (In contrast, binomial models
> show annotation

3. Deterministic Inventory Models

3.1 Introduction to Inventory Modeling

> er in increments of those units. These are all reasons that firms plan to hold inventory. In addition, firms may holdunplanned inventory—for example, inventory of products that have become obsolete soonerthan expected .Firms may hold inventory of goo
> show annotation

> me may be uncertain, and so on. In fact, although we tend todiscuss inventory models as though the firm is buying a product from an outside supplier,most inventory models apply equally well to production systems, in which case we aredeciding ho
> show annotation

> 1.2 Classifying Inventory Models Mathematical inventory models can be classified along a number of different dimensions: • Demand. Is demand deterministi
> show annotation

Mathematical inventory models can be classified along a number of different dimensions:

- Demand
- Lead time
- Review type
- Planning horizon
- Stockout type

- Ensuring good service
- Fixed cost
- Perishability

> *after which they can't be sold). Like all mathematical models, inventory models must balance two competing factors—realism and tractability . In many cases, it is more accu*
> *show annotation*

> *ven real-life factor.3.1.3 Costs The goal of most inventory models is to minimize the cost (or maximize the profit) of theinventory system. Four types of costs are most co*
> *show annotation*

Four types of costs:

- Holding cost
- Fixed cost
- Purchase cost
- Stockout cost

> *DETERMINISTIC INVENTORY MODELS• Stockout cost. This is the cost of not having sufficient inventory to meet demand, alsocalled the penalty cost or stockout penalty , and is denoted by p. If excess*
> *show annotation*

> *ory Level and Inventory Position There are several measures that we use to assess the amount of inventory in the system atany given time. On-hand inventory (OH) refers to*
> *show annotation*

There are several measures that we use to assess the amount of inventory in the system at any given time.

- On-hand inventory (OH) refers to the number of units that are actually available at the stocking location.

- Backorders (BO) represent demands that have occurred but have not been satisfied. Generally, it's not possible for the on-hand inventory and the backorders to be positive at the same time
- The inventory level (IL) is equal to the on-hand inventory minus backorders
- inventory position (IP), which equals the inventory level plus items on order

> *UANTITY PROBLEM 51review model, the economic order quantity (EOQ) model, perhaps the oldest and best-known mathematical inventory model (Section 3.2), and some of its extensions; and thena periodic-review model , the Wagner–Whitin model (Secti*
> *show annotation*

> *els are considered in Chapter 6. The models discussed in this chapter are sometimes known as economic lot size problems. In fact, there is some inconsist*
> *show annotation*
>
> *经济批量问题*

3.2 Continuous Review: The Economic Order Quantity Problem

3.2.1 Problem Statement

> *Y PROBLEM3.2.1 Problem Statement The economic order quantity (EOQ) problem is one of the oldest and most fundamentalinventory models; it was first introduced by Harris (1913). The goal is to determine theoptimal amount to order each time an order is placed to minimize the average cost per year. (We'll express everything per ye*
> *show annotation*

> *placed, and the process repeats. Any optimal solution for the EOQ model has two important properties: • Zero-inventory ordering (ZIO)*
> *show annotation*

Any optimal solution for the EOQ model has two important properties:

1. Zero-inventory ordering (ZIO) property;
2. Constant order sizes.

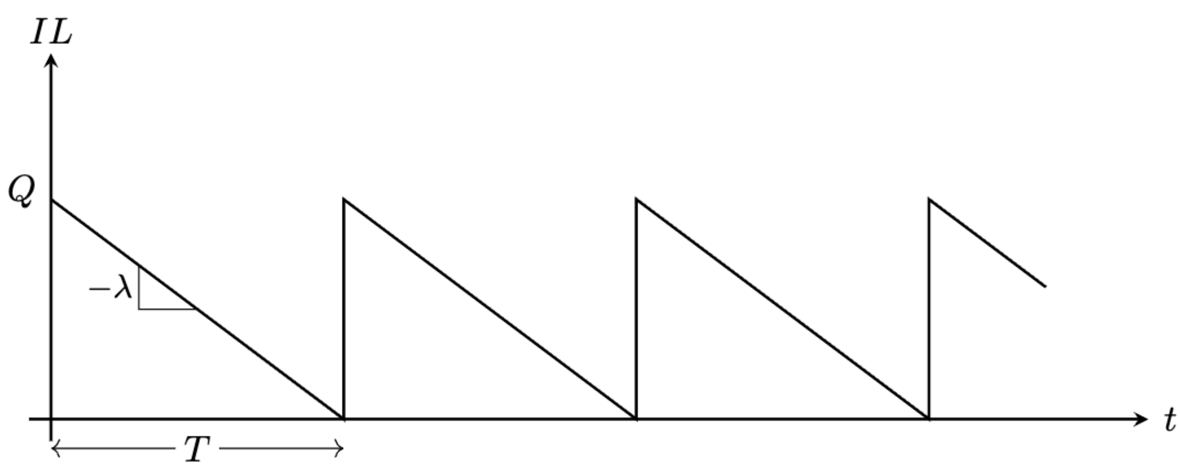


Figure 3.2 EOQ inventory level curve.

T is the *cycle length*. meaning the amount of time between orders, and it relates to the order quantity Q and λ by the equation:

$$T = \frac{Q}{\lambda}$$

3.2.2 Cost Function

> *ationT = Qλ .3.2.2 Cost Function We want to find the optimal Q to minimize the average annual cost. (We say "average"annual cost sin*
> *show annotation*

> *ng very tiny order quantities. T he key trade-off is between fixed cost and holdingcost: If we use a large Q, we'll place fewer orders and hold more inventory (small fixedcost but large holding cost), whereas if we use a small Q, we'll place more orders and holdless inventory (large fixed cost but small holding cost). The strategy for solving the EOQ*
> *show annotation*

- *Order cost* per year

$$\frac{K}{T} = \frac{K\lambda}{Q}$$

- Average Annual *Holding Cost*

$$\frac{hQ}{2}$$

where K is a fixed cost per order, h is an inventory holding cost per unit per year. c is purchase cost per unit ordered.

> *isKT = KλQ . (3.1)Holding Cost: The average inventory level in a cycle is $Q/2$, so the average amount ofinventory per year is $Q/2$ ·1 year = $Q/2$. (Another way to think about th*
> [show annotation]

- *Total Cost*

$$g(Q) = \frac{K\lambda}{Q} + \frac{hQ}{2}$$

3.2.3 Optimal Solution

> *igure 3.3.3.2.3 Optimal Solution The optimal Q can be obtained by taking the derivative of $g(Q)$ and setting it to 0: dg(Q)dQ = −KλQ2 + h2 = 0=⇒ Q2 =*
> [show annotation]

$$\begin{aligned} \frac{dg(Q)}{dQ} &= -\frac{K\lambda}{Q^2} + \frac{h}{2} = 0 \\ &\Longrightarrow Q^2 = \frac{2K\lambda}{h} \\ &\Longrightarrow Q^* = \sqrt{\frac{2K\lambda}{h}} \end{aligned}$$

Q^* is the *economic order quantity*. Then, the optimal total cost is $g(Q^*)$ is:

$$g(Q^*) = \sqrt{2K\lambda h}$$

> moreinventory. (And vice versa.) Another way to see that the fixed and holding costs are equal in the optimal solution isto note that the product of the two terms in (3.3) i sKλQ · hQ2 = Kλh2 ,a constant. I
> show annotation

> ot true for many otherproblems. The ability to express g(Q≣) in closed form allows us to learn about structuralproperties of the EOQ and related models, such as the power-of-two policies discussed inSection 3.3, as well as to embed the EOQ into other, richer models, such as the locationmodel with risk pooling (LMRP) in Section 12.2 .The optimal EOQ solution and it
> show annotation

> (3.7)Using Theorem 3.1, we can make some statements about how the solution changes asthe parameters change: • As h increases, Q≣ decreases,
> show annotation

> m prefers small demand, however. Remember that the EOQ only reflects costs, not revenues; the increased cost of large λwould be outweighed by the increased revenue . EXAMPLE 3.1Joe's Corner Store
> show annotation

3.2.4 Sensitivity Analysis

> 304. 3.2.4 Sensitivity Analysis Suppose the firm did not want to order Q^* exactly. For example, it might need to order in multiples of $10(Q = 10n)$, or it might want to order every month ($T = 1/12$). How muchmore expensive is a sub
> show annotation

Theorem: Suppose Q^* is the *optimal order quantity* in the EOQ model, then for any $Q > 0$:

$$\frac{g(Q)}{g(Q^*)} = \frac{1}{2}\left(\frac{Q^*}{Q} + \frac{Q}{Q^*}\right)$$

> = Q≣/2), the error is also 1.25. Theorem 3.2 ignores the per-unit cost c. If we include the annual cost cλin the numeratorand denominator of (3.8), then the

> *percentage increase in cost would be even smaller (andthe expressions would not si*
> [show annotation]

3.2.5 Order Lead Times

> *sWe assumed the lead time is 0. What if the lead time was positive—say, L years? Theoptimal solution doesn't change—we just place our order L years before it's needed. Forexample, if L = 1 month = 1/*
> [show annotation]

> *ethe inventory level reaches 0. It's generally more convenient to express this in terms ofthe reorder point (r). When the inventory level reache*
> [show annotation]

So how do we compute r? It should be equal to the amount of product demanded during the lead time, or

$$r = \lambda L$$

3.3 Power-of-two Policies

> *, for example, every √10 weeks? In this section, we discusspower-of-two policies, in which the order interval is required to be a power-of-two multipleof some base period. The base period may be any time*
> [show annotation]

> *he base period is a day (say), t hen the power-of-two restriction says that orders can beplaced every 1 day, or every 2 days, or every 4 days, or every 8 days, and so on , or every 1/2day, or every 1/4*
> [show annotation]

> *nvolving base periods like √10. Wealready know that the EOQ model is relatively insensitive to deviations from the optimalsolution from Theorem 3.2. Our goal is to determine exactly how much more expensive apower-of-two policy is than the optimal policy .Power-of-two policies have anot*
> [show annotation]

> *own inventory planning easier. T he problem of finding optimalorder intervals in this setting is one version of a problem known as the one warehouse,multiretailer (OWMR) problem . The optimal policy for the OWM*
> show annotation

3.3.2 Error Bound

> *isfying (3.14).3.3.2 Error Bound Theorem 3.3 If ˆT is the optimal power-of-t*
> show annotation

Theorem 3.3 If $\hat{T}$ is the optimal power-of-two order interval and T^* is the optimal (not necessarily power-of-two) order interval, then

$$\frac{f(\hat{T})}{f(T^*)} \leq \frac{3}{2\sqrt{2}} \approx 1.06$$

> *, thenf(ˆT)f(T) ≤ 32√2 ≈1.06.I n other words, the cost of the optimal power-of-two policy is no more than 6% greaterthan the cost of the optimal (non-power-of-two) policy . This holds for any choice of t*
> show annotation

Hints for proof: the optimal power-of-two order interval $\hat{T}$ must be in the interval $\left[\frac{1}{\sqrt{2}}T^*, \sqrt{2}T^*\right]$. Since we don't know precisely where $\hat{T}$ falls in the range, so it is only a *worst-case* bound that occurs on the endpoints of the range.

3.4 The EOQ With Quantity Discounts

> *THE EOQ WITH QUANTITY DISCOUNTS It is common for suppliers to offer discounts based on the quantity ordered . The larger theorder, the lower*
> show annotation

> *n bulk, you pay less per unit.) The specificstructure for the discounts can take many forms, but two types are most common: all-unitsdiscounts and incremental discount s. Both discount structures use*
> show annotation

3.4.1 All-Units Discounts

> *le 3.5.3.4.1 All-Units Discounts We can no longer ignore the purchase cost as we did in (3.3). In fact, not only do we needto include the purchase cost itself, but we must also account for the fact that the holdingcost typically depends on the purchase cost , as discussed in Section 3.1.3.*
> show annotation

> *its quantity discount structure. Suppose we knew that the optimal order quantity lies in region j. Then we would simplyneed to find the Q that minimizes the EOQ cost function for region j: gj(Q) = cjλ + KλQ + icjQ2 . (3.1*
> show annotation

$$g_j(Q) = c_j\lambda + \frac{K\lambda}{Q} + \frac{ic_jQ}{2}$$

Its minimizer is given by:

$$Q_j^* = \sqrt{\frac{2K\lambda}{ic_j}}$$

3.4.2 Incremental Discounts

> *60. 3.4.2 Incremental Discounts We now turn our attention to incremental discounts. The total cost function for region j is given by gj(Q) = c(Q)Q λ + KλQ + ic(Q)Q Q*
> show annotation

$$g_j(Q) = \frac{c(Q)}{Q}\lambda + \frac{K\lambda}{Q} + \frac{i\frac{c(Q)}{Q}Q}{2}$$

where

$$c(Q) = \sum_{i=0}^{j} c_i(b_{i+1} - b_i) + c_j(Q - b_j)$$

Then, we can rewrite $g_j(Q)$ as

$$
\begin{aligned}
g_j(Q) =& \frac{1}{Q}\left[\sum_{i=0}^{j-1} c_i\left(b_{i+1}-b_i\right)-c_j b_j\right]\lambda + c_j\lambda + \frac{K\lambda}{Q} \\
&+ \frac{i}{2}\left[\sum_{i=0}^{j-1} c_i\left(b_{i+1}-b_i\right)-c_j b_j\right] + \frac{ic_jQ}{2} \\
=& c_j\lambda + \frac{i\bar{c}_j}{2} + \frac{(K+\bar{c}_j)\lambda}{Q} + \frac{ic_jQ}{2}
\end{aligned}
$$

where

$$
\bar{c}_j = \sum_{i=0}^{j-1} c_i\left(b_{i+1}-b_i\right)-c_j b_j
$$

so the minimizer is given by

$$
Q^* = \sqrt{\frac{2(K+\bar{c}_j)\lambda}{ic_j}}
$$

with cost

$$
g_j(Q^*) = c_j\lambda + \frac{i\bar{c}_j}{2} + \sqrt{2(K+\bar{c}_j)\lambda i c_j}
$$

3.4.3 Modified All-Units Discounts

> *4.3 Modified All-Units Discounts* *All-units discounts are somewhat problematic because, for order quantities Q just to theleft of breakpoint j, it is cheaper to order bj than to order Q* *, even though Q < bj. Forexample*
> *show annotation*

> *le, ship 390 kg, declare 400 kg.* *This structure is sometimes known as the modified all-units discount structure.* *Its c(Q)curve is displayed in F*
> *show annotation*

> *e c(Q) curve; see Figure 3.8(b).* *A special case of the modified all-units discount structure is the carload discount struc-ture, in which the bj are equally spaced and cj is the same for all j.* *This structure arisesfrom rail*
> *show annotation*

> d all-units discounts structure. Unfortunately, modified all-units discount structures are much more difficult to analyzethan the discount structures discussed above. (See, for example, Chan et al.
> show annotation

3.5 The EOQ with Planned Backorders

> hat backorders are not allowed. In this section, we discuss avariant of the EOQ problem in which backorders are allowed. Since demand is determin-68
> DET
> show annotation

Let p be the backorder penalty per item per year, and let x be the fraction of demand that is backordered. Both Q and x are decision variables. The holding cost is charged based on on-hand inventory; the average on-hand inventory is given by:

$$\frac{1}{2}Q(1-x)\cdot\frac{(1-x)T}{T}=\frac{Q(1-x)^2}{2}$$

Similarly, the average backorder level is:

$$\frac{Qx^2}{2}$$

There, the total average cost per year in the EOQB is given by:

$$g(Q,x)=\frac{hQ(1-x)^2}{2}+\frac{pQx^2}{2}+\frac{K\lambda}{Q}$$

> is a function of both Q and x. Therefore, to minimize it, we need to take partialderivatives with respect to both variables and set them equal to 0.
> ∂g∂x = −hQ(1 −x) + pQx = 0 (3.24
> show annotation

$$\frac{\partial g}{\partial x}=-hQ(1-x)+pQx=0$$
$$\frac{\partial g}{\partial Q}=\frac{h(1-x)^2}{2}+\frac{px^2}{2}-\frac{K\lambda}{Q^2}=0$$

for the first equation, we have:

$$\begin{aligned}-hQ(1-x)+pQx&=0\\ \iff h(1-x)&=px\\ \iff x^*&=\frac{h}{h+p}\end{aligned}$$

Then, plug x^*, we have

$$Q^*=\sqrt{\frac{2K\lambda}{h(1-x)^2+px^2}}$$
$$Q^*=\sqrt{\frac{2K\lambda(p+h)}{hp}}$$

Then

$$g(Q^*,x^*)=\sqrt{\frac{2K\lambda hp}{h+p}}$$

> *.28)g(Q≣,x≣) =√2Kλhph + p (3.29) How do the optimal solution and cost in Theorem 3.5 compare to the analogous quantitiesfrom the EOQ model? First, comparing (3.29) and (3.*
>
> show annotation

As $p\to\infty$, Q^* approaches the optimal EOQ order quantity, x^* approaches 0, and the optimal cost approaches the EOQ optimal cost.

> *tantial savings inholding cost. As we continue to increase the number of backorders, the marginal savingsin holding cost decreases and the marginal increase in backorder cost increases . At somepoint, the marginal cos*
>
> show annotation

3.6 The Economic Production Quantity Model

> *ONOMIC PRODUCTION QUANTITY MODEL In a manufacturing environment, the amount of time required to produce a batch of itemsusually depends on how large the batch is—producing more items requires more time. TheEOQ*

> model cannot handle this
> show annotation

> ardless of the order quantity. I n other words,the EOQ assumes that the production rate is infinite—an arbitrary number of items can beproduced in a fixed amount of time . This assumption may be reasona
> show annotation

> ng the inventory at a rate ofλ. Let $\rho = \lambda/\mu$ be the utilization ratio, which indicates the portion of time the system isactive. Q is now interpreted as a production batch size rather than an order quantity. The process is depicted in Figur
> show annotation

> ch cycle, is $\rho T(\mu - \lambda) = (1 - \rho)Q$. The fixed cost per year is still $K\lambda/Q$, as in the EOQ model, since $T = Q/\lambda$. Theaverage inventory level is (
> show annotation

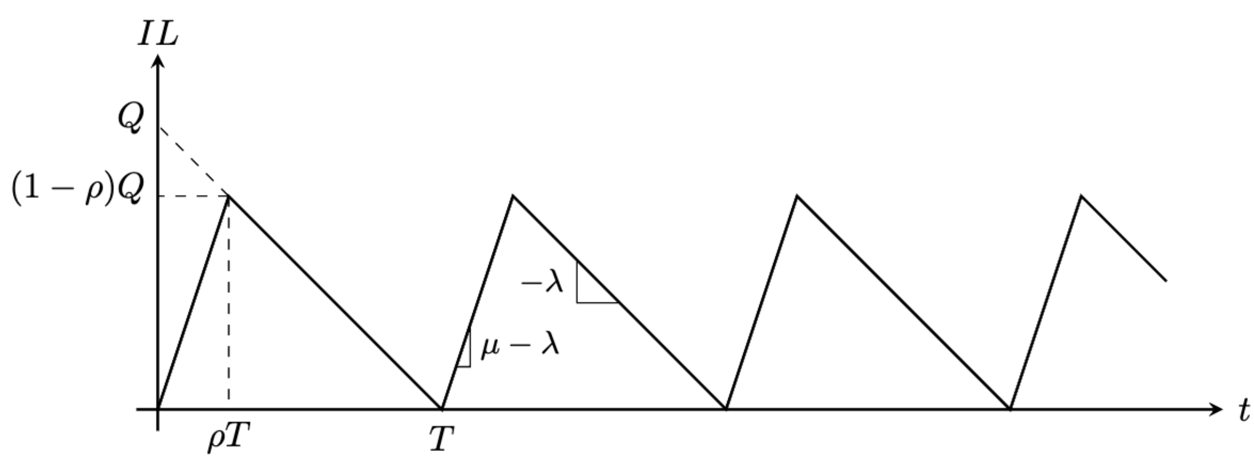


Figure 3.11 EPQ inventory level curve.

- The order interval

$$T = \frac{Q}{\lambda}$$

- Active interval

$$\frac{Q}{\mu} = \rho T$$

- Maximum inventory level

$$\rho T(\mu - \lambda) = (1 - \rho)Q$$

- Fixed cost per year: $\frac{K\lambda}{Q}$
- Average inventory level: $\frac{(1-\rho)Q}{2}$
- Average annual holding cost: $\frac{h(1-\rho)Q}{2}$
- Total annual cost

$$g(Q) = \frac{K\lambda}{Q} + \frac{h(1-\rho)Q}{2}$$

> *sg(Q) = KλQ + h(1 −ρ)Q2 . (3.30) We could find the Q that minimizes this cost function by differentiating, as we did for theEOQ, but it is simpler to recognize*
> *show annotation*

3.7 Periodic Review: The Wagner-Whitin Model

3.7.1 Problem Statement

> *TIN MODEL3.7.1 Problem Statement We now shift our attention to a periodic-review model known as the Wagner–Whitinmodel (Wagner and Whitin 1958). Simil*
> *show annotation*

> *58). Similar to the EOQ model, t he Wagner–Whitin modelassumes that the demand is deterministic, there is a fixed cost to place an order, and stock-outs are not allowed . The objective is to choose ord*
> *show annotation*

> *ties to minimize the total cost. However, unlike the EOQ model, the Wagner–Whitin model allows the demand to changeover time—to be different in each period . This model is sometimes referr*

> show annotation
>
> dynamic economic lot-sizing (DEL) model

> o be different in each period. T his model is sometimes referred to as thedynamic economic lot-sizing (DEL) model or the uncapacitated lot-sizing (ULS) model. Because of the fixed cost, it ma
> show annotation

> f events 2 and 3 were reversed.) We first formulate this model as a mixed-integer optimization problem (MIP). We willthen discuss a dynamic programming (DP) algorithm for solving it. 3.7.2 MIP FormulationOur formula
> show annotation

3.7.2 MIP Formulation

> tory at the end of each period. Constraints (3.34)are the inventory-balance constraints: They say that the ending inventory in period t isequal to the starting inventory, plus the new units ordered, minus the demand. Constraints(3.35) prohibit qt f
> show annotation

> constraints on the y variables. This problem can be interpreted as a simple supply chain network design problem (to bemore precise, an arc desig
> show annotation

> gn problem; see Section 8.7.2). It can be solved as an MIP, but it ismore common to solve it using DP or as a shortest path problem, as we discuss in the nextsection . See Pochet and Wolsey (1995, 2
> show annotation

3.7.3 Dynamic Programming Algorithm

> depends on the following result: Theorem 3.7 Every optimal solution to the Wagner–Whitin model has the ZIO property;that is, it is optimal to place orders only in time periods in which the initial inventory iszero .Proof.

> *Suppose (for a contradic*
> [show annotation]

> *ur 1991,Wagelmans et al. 1992). Despite the efficiency of this algorithm, a number of heuristicshave been introduced and are still popular in practice. These include Silver–Meal, partperiod balancing, least unit cost, and other heuristics (Silver et al. 1998). One expla*
> [show annotation]

4. Stochastic Inventory Models: Periodic Review

> *DIC REVIEW4.1 INVENTORY POLICIES In this chapter and the next, we will consider inventory models in which the demand isstochastic . A key concept in these chapter*
> [show annotation]

> *which the demand isstochastic. A key concept in these chapters will be that of a policy. A policy is a simple rulethat provides a solution to the inventory problem. For example, consider a periodi*
> [show annotation]

> *l more closely in Section 4.4.) One could imagine severalpossible policies for this system. Here are a few: 1. Every R periods, place an ord*
> [show annotation]

1. Every R periods, place an order for Q units.
2. Whenever the inventory position falls to s, order Q units.
3. Whenever the inventory position falls to s, place an order of sufficient size to bring the inventory position to S.
4. Place an order whose size is equal to the first two digits of last night's lottery number. Then, wait a number of periods equal to the last two digits of the lottery number before placing another order.

> *icy depends on its parameters.1 Forexample, policy 1 sounds reasonable, but only if we choose good values for R and Q. It is often possible (and always*
> show annotation

> *smallest possible expected cost. When using policies, then, inventory optimization really has two parts: Choosing theform of the optimal policy and choosing the optimal parameters for that policy. Sometimeswe can't solve one of t*
> show annotation

> *d approximately optimal values. Similarly, for some problems,no one even knows the form of the optimal policy, so we simply choose a policy that seemsplausible .We'll consider periodic-review*
> show annotation

> *scuss the lost-sales assumption. Before continuing, we introduce two important concepts in stochastic inventory theory:cycle stock and safety stock. Cycle stock (or working invento*
> show annotation

- *Cycle stock* (or working inventory) is the inventory that is intended to meet the expected demand.
- *Safety stock* is extra inventory that's kept on hand to buffer against uncertainty.

> *cle and safety stock components. We'll see later that the cycle stock depends on the mean of the demand distribution, whilethe safety stock depends on the standard deviation. 1 We don't mean the inputs to th*
> show annotation

4.2 Demand Processes

> *PROCESSES 894.2 DEMAND PROCESSES In real life, customers tend to arrive at a retailer at random, discrete points in time. Similarly,(some) retailers plac*
> show annotation

, and so on up thesupply chain. One way to model these demands is using a Poisson process, which describesrandom arrivals to a system over time. If each customer may demand mor
show annotation

ion withmean μ and variance σ2. Sometimes, the normal distribution is used as an approximationfor the Poisson distribution, in which case $\mu = \sigma^2$ since the Poisson variance equals itsmean. (This approximation is especially accurate when the mean is large.)In the continuous-review case,
show annotation

period is normally distributed. One drawback to using the normal distribution is that any normal random variable willsometimes have negative realizations, even though the demands that we aim to model arenonnegative . If the demand mean is much gre
show annotation

l that we can simply ignore it. This suggeststhat the normal distribution is appropriate as a model for the demand only if $\mu \gg \sigma$ — say,if $\mu > 4\sigma$. If this condition fails to hold
show annotation

4.3 Periodic Review With Zero Fixed Costs: Base-stock Policies

FIXED COSTS: BASE-STOCKPOLICIES For the remainder of this chapter, we focus on periodic-review models . The time horizonconsists of T
show annotation

ant throughout the time horizon. We will model the time value of money by discounting future periods using a discountfactor $\gamma \in (0, 1]$. That is, $1 spent (or received
show annotation

> , then there is no discounting. For the single-period and finite-horizonproblems, our objective will be to minimize the total expected discounted cost over thehorizon . However, the total cost over a
> show annotation

4.3.1 Base-Stock Policies

> m follows a base-stock policy.2 A base-stock policy works as follows: In each time period, we observe the current inventory position and then place an order whose size is sufficient to bring the inventory position up to S . (We sometimes say we "order up
> show annotation

4.3.2 Single Period: The Newsvendor Problem

> uct during a singletime period. Single-period models are most often applied to perishable products, whichinclude (as you might expect) products such as eggs and flowers that may spoil, but alsoproducts that lose their value after a certain date, such as newspapers, high-tech de
> show annotation

> gle-period model presented here. This model is one of the most fundamental stochastic inventory models, and many ofthe models discussed subsequently in this book use it as a starting point. It is often referredto as the n
> show annotation

> and a standard deviation of 8. If the newsvendor has unsold newspapers left at the endof the day, he cannot sell them the next day, but he can sell them back to the publisherfor $0.12 (called the salvage value). The question is: How many newsp
> show annotation

> apers each day—buthow many more? The inventory carried by the newsvendor can be decomposed into two components:cycle stock and

> safety stock . As noted in Section 4.1, cycle
> show annotation

> eutral term “newsvendorproblem.” As previously noted, the newsvendor model applies to perishable goods. In particular,it applies to goods that perish before the next ordering opportunity. Many perishable goodshave a she
> show annotation

> fit and loss-of-goodwill costs. The holding cost is the cost per unitof positive ending inventory, while the stockout cost is the cost per unit of negative endinginventory. The costs h and p are sometimes referred to as overage and underage costs, respectively (and some authors d
> show annotation

> on’tplay a role in the analysis. We will refer to the model discussed here as the implicit formulation of the newsvendorproblem since the costs and revenues are not modeled explicitly but instead are accountedfor in the holding and stockout costs h and p . (In contrast, see the explicit
> show annotation

> e discrete demand distributions. Our goal is to determine the base-stock level S to minimize the expected cost in thesingle period. The strategy for solving this pr
> show annotation

> X −x)−] =∫ x0(x −y)f(y)dy. (4.5) These functions are known as the loss function and the complementary loss function,3respectively. They can be defined for any pro
> show annotation

Let $I(S,d) = (S-d)^+$, and $B(S,d) = (d-S)^-$ be the on-hand inventory and backorders, respectively. At the end of period if the firm orders up to S and sees a demand of d units. The cost for an observed demand od d is:

$$\begin{aligned} g(S,d) &= hI(S,d) + pB(S,d) \\ &= h(S-d)^+ + p(d-S)^- \end{aligned}$$

Since the demand is *stochastic*, however, we must have an expectation over D. Let $I(S) = \mathbb{E}[I(S,d)] = \mathbb{E}[B(S,D)]$ be the *expected* on-hand inventory and backorders if the firm orders up to S. Then

$$
\begin{aligned}
g(S) &= hI(S) + pB(S) \\
&= h\mathbb{E}\left[(S-D)^+\right] + p\mathbb{E}\left[(D-S)^+\right] \\
&= h\int_0^\infty (S-d)^+ f(d)\,\mathrm{d}\,d + p\int_0^\infty (d-S)^+ f(d)\,\mathrm{d}\,d \\
&= h\int_0^S (S-d) f(d)\,\mathrm{d}\,d + p\int_S^\infty (d-S) f(d)\,\mathrm{d}\,d
\end{aligned}
$$

let

$$
\begin{aligned}
n(x) &= \mathbb{E}\left[(X-x)^+\right] = \int_x^\infty (y-x) f(y)\,\mathrm{d}\,y \\
\bar{n}(x) &= \mathbb{E}\left[(X-x)^-\right] = \int_0^x (x-y) f(y)\,\mathrm{d}\,y
\end{aligned}
$$

Then:

$$
g(S) = h\bar{n}(S) + pn(S)
$$

Since

$$
\begin{aligned}
\bar{n}(S) &= \int_0^S (S-d) f(d)\,\mathrm{d}\,d \\
&= \int_0^{+\infty} (S-d) f(d)\,\mathrm{d}\,d - \int_S^{+\infty} (S-d) f(d)\,\mathrm{d}\,d \\
&= S - \mu + \int_S^{+\infty} (d-S) f(d)\,\mathrm{d}\,d \\
&= S - \mu + n(S)
\end{aligned}
$$

where μ is the *mean value* of $f(x)$. So we can simplify the expression of $g(S)$:

$$
\begin{aligned}
g(S) &= h[S - \mu + n(S)] + pn(S) \\
&= h(S-\mu) + (h+p)n(S)
\end{aligned}
$$

These functions are known as the *loss function* and the *complementary loss function*, respectively. Here we can calculate the derivatives of the above two equations according to the [Leibniz's Rule](#):

$$n'(x) = (y-x)f(y)\Big|_{y=x} + \int_x^{+\infty} \frac{\partial(y-x)f(y)}{\partial x} \mathrm{d}\, y$$

$$= 0 + \int_x^{+\infty} (-1)f(y)\, \mathrm{d}\, y$$

$$= F(x) - F(+\infty)$$

$$= F(x) - 1$$

Similarly,

$$\bar{n}'(x) = F(x)$$

So we have $n'' = \bar{n}'' = f(x) > 0$, so both $n(\cdot)$ and $\bar{n}(\cdot)$ are convex. To minimize $g(S)$, thus, we set its first derivative to 0,

$$\begin{aligned} \frac{\mathrm{d}\, g(S)}{\mathrm{d}\, S} &= h + (h+p)[F(x) - 1] = (h+p)F(S) - p = 0 \\ F(S) &= \frac{p}{h+p} \\ S &= F^{-1}\left(\frac{p}{p+h}\right) \end{aligned}$$

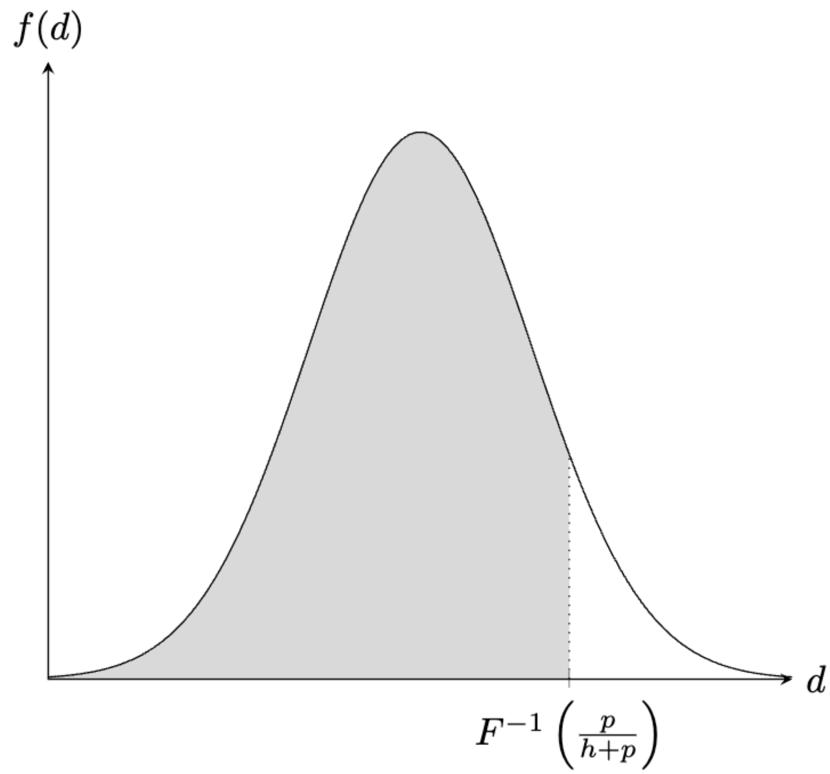


Figure 4.1 Optimal solution to newsvendor problem plotted on demand distribution.

Note：直观来理解，就是最优解 S^* 符合 $P[d > S^*] = \frac{p}{h+p}$, 即右边表示售罄是的损失 p, 左边表示持货成本 h。S^* 会随着 h 增加而减少，随着 p 增加而增加。

> *theorem (whichwe've now proven). Theorem 4.1 The optimal base-stock level for a single-period model with no fixed costs(the newsvendor model) is given by S= = F−1(ph + p).94 STOCHASTIC*
> show annotation

> *plotted on demand distribution. p/(h + p) is known as the critical ratio (or critical fractile). It is implicit in a result byAr*
> show annotation

> *. 4.3.2.4 Explicit Formulation The formulation given in Sections 4.3.2.2–4.3.2.3 in-terprets hand pas the overage and underage costs, respectively—the cost per unit of havingtoo much or too little inventory. The actual cost and revenue par*
> show annotation

> (h + r −v) ̄n(S) + pn(S). (4.19) Sometimes, this is instead formulated as a profit maximization problem in which we maximize $\pi(S) \equiv -g(S)$. 96 STOCHASTIC INVENTORY MODELS:
> show annotation

> 1(p + r −ch + p + r −v). (4.20) We can translate this to the implicit version of the problem by determining the overageand underage costs (which we'll denote by h′ and p′
> show annotation

Actually, we can translate this explicit formulation to the implicit version by determining the overage and underage costs (which we'll denote by h' and p', respectively).

- $h' = h + c - v$: For each unit of excess inventory, we incur a holding cost of h, and we paid c for the extra unit but earn v as a salvage value.
- $p' = p + r - c$: For each stockout, we incur a penalty of p in addition to the lost profit $r - c$.

$$S^* = F^{-1}\left(\frac{p'}{h'+p'}\right) = F^{-1}\left(\frac{p+r-c}{h+c-v+p+r-c}\right) = F^{-1}\left(\frac{p+r-c}{h+p+r-v}\right)$$

where r is the revenue earned per unit sold, c is the cost per unit purchased, and v is the salvage value earned for each unit of excess inventory.

> ve constant (see Problem 4.15).I t is perfectly acceptable to set any of the cost or revenue parameters to 0 if they arenegligible or should not be included in the model .One word of caution: Avoid mixi
> show annotation

> ph + p⇐⇒ S䷀ = μ + σΦ−1(ph + p). If we let α = p/(h + p), we have S䷀ = μ + zασ. (4.24)The first te
> show annotation

If we let $\alpha = \frac{p}{h+p}$, we have

$$S^* = \mu + z_\alpha \sigma$$

> alreadyhas too much inventory. But should the firm order any units? By the convexity of g(S),the answer is no: It would be better to leave the inventory level where it is . Therefore, theoptimal order qu
> show annotation

> ays to do this (see Chapter 2); one of the simplest is to use a moving average(Section 2.2.1) to estimate μand what we might call a moving standard deviation to estimateσ in period t :ˆμt = 1Nt−1∑i=t−Ndt ˆσt =√√√√
> 1
> show annotation

>)Δg(S)Figure 4.3 g(S) and Δg(S). 4.3.2.8 Discrete Demand Distributions Suppose now that D is discrete. In thiscase, (4.3) becomes g(S) = hS∑d=0(S −d)f(d) + p∞∑d=S
> show annotation

4.3.3 Finite Horizon

> roblem 4.16.4.3.3 Finite Horizon Now consider a multiple-period problem consisting of a finite number of periods, T.Suppose we are at the beginning of period t . Do we need to know the history
> show annotation

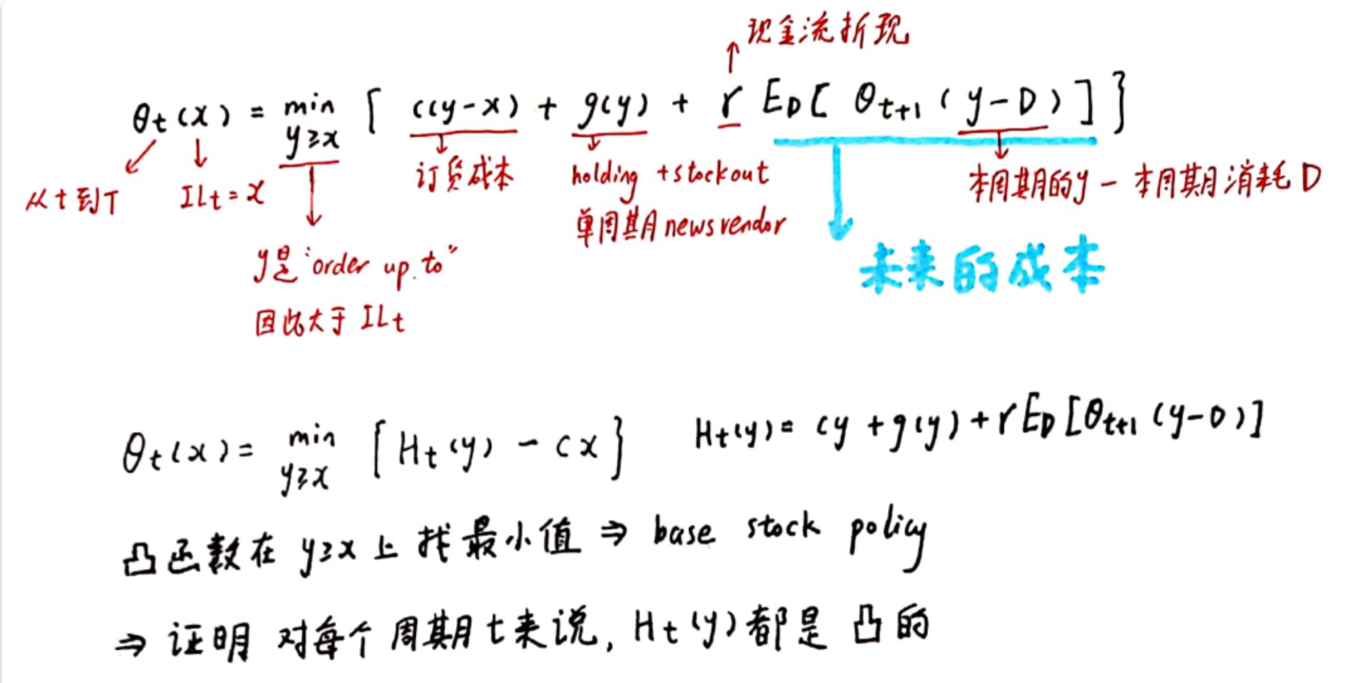
Let $\theta_t(x)$ be the optimal expected cost in periods $t, t+1, \cdots, T$ if we begin period t with an inventory level x. Then, we have

$$\theta_t(x) = \min_{y \geq x} \{c(y-x) + g(y) + \gamma \mathbb{E}_D[\theta_{t+1}(y-D)]\}$$

where

$$g(y) = h\int_0^y (y-d)f(d)\,\mathrm{d}\,d + p\int_y^\infty (d-y)f(d)\,\mathrm{d}\,d = h\bar{n}(y) + pn(y)$$

The interpretation of each term can be explained as follows:



Note: figure from [【供应链理论基础】零固定成本的周期性盘点条件下基本库存策略（Base-Stock Policy）最优性的证明]()

> *depended only on the period, t.* *First consider what happens at the end of the time horizon. Presumably, on-handunits and backorders must be treated differently now that the horizon has ended than theywould be during the horizon* *. The terminal cost function, de*
>
> *show annotation*

4.3.4 Infinite Horizon

> *osts is the case inwhich T = ∞.* *This problem is sometimes referred to as the infinite-horizon newsvendormodel.* *If the number of periods is infi*
>
> *show annotation*

> *It certainly will be if γ = 1.)* *An alternate objective is to minimizethe expected cost per period. The former case is known as the discounted-cost criterion,while the latter is known as the average-cost criterion* *. We'll consider the average-cos*
>
> *show annotation*

Under the average cost criterion, we assume $\gamma = 1$. The expected cost in a given period if we use back-stock level S is given by:

$$g(S) = h\int_0^S (S-d)f(d)\,\mathrm{d}\,d + p\int_S^\infty (d-S)f(d)\,\mathrm{d}\,d = h\bar{n}(y) + pn(y)$$

Now suppose $\gamma < 1$, consider the discounted-cost criterion. The optimal base-stock level is the same in every period, and it is given by

$$S^* = F^{-1}\left(\frac{p-(1-\gamma)c}{h+p}\right)$$

Then, if demand is normally distributed, then after modifying to account for γ, the results would be

$$S^* = \mu + \sigma\Phi^{-1}\left(\frac{p-(1-\gamma)c}{h+p}\right) = \mu + z_\alpha\sigma$$

where $\alpha = \left(\frac{p-(1-\gamma)c}{h+p}\right)$.

> *4.1—is optimal, inevery period! In formulating (4.38), we glossed over two potentially problematic issues. First, whydidn't we account for*
> *show annotation*

Two problematic issues:

1. Why didn't we account for the purchase cost c,
2. Why didn't we account for the cost in future periods?

4.4 Periodic Review With Nonzero Fixed Costs: (s, S) Polices

4.4.1 (s, S) Policies

> *s,S) POLICIES4.4.1 (s,S)Policies We now consider the more general case in which the fixed cost K may be nonzero. If $K \neq 0$, it may no longer make sense to order in every period, since each order incurs a cost. PERIODIC REVIEW WITH NONZERO FIX*
> *show annotation*

> FIXED COSTS: (s,S) POLICIES 115 Instead, the firm should order only when the inventory position becomes sufficiently low. In particular, we will assume in
> show annotation

> inventory position up to S. Bot Both s and S are constants, and $s \leq S$.The quantity s is known as the reorder point and S as the order-up-to level . The reorderpoint and order-up-
> show annotation

> olicy, as we doin this section; the optimality of (s,S) policies for multiperiod problems was not provenuntil Scarf's (1960) paper .(s,S) polices are closely relat
> show annotation

> each period can only be 0 or 1. We will discuss how to determine the optimal s and S for the single-period, finite-horizon, and infinite-horizon cases separately , just as we did in Section 4.3
> show annotation

> 3 for the zero-fixed-cost case. Actually, the single-period case is not nearly as useful for the $K > 0$ caseas it is for the $K = 0$ case . This is because single-period
> show annotation

4.4.2 Single Period

> rt of the (single) period is x. For given s and S, theordering rule is: If $x \leq s$, order $S - x$; otherwise, order 0. Once we order (or don't), weinc
> show annotation

一旦订购（或不订购），就会产生持有和缺货成本，就像在零固定成本模型中一样，只是基础库存水平被 S（如果我们订购）和 x（如果我们不订购）取代。因此，该期间的总预期成本（作为 s 和 S 的函数）由下式给出

$$g(s,S) = \begin{cases} K + g(S), & \text{if } x \leq s \\ g(x), & \text{if } x > s \end{cases}$$

> ed costs, we areassuming c = 0.) Optimizing $g(s, S)$ over s and S is actually quite easy (Karlin 1958b): We already knowfrom Theorem 4.1 that $F - 1(p/(h + p))$ minimizes $g(S)$, so our aim should be to order u
> show annotation

4.4.3 Finite horizon

> ue if x > s.4.4.3 Finite Horizon The finite-horizon model with nonzero fixed costs can be solved using a straightforwardmodification of the DP model for the zero-fixed-cost case from Section 4.3.3. Just as before,θt(x) represents
> show annotation

> (and act optimally thereafter). Now $\theta_t(x)$ must account for the fixed cost in period t (if any), as well as the purchase cost and expected holding and stockout costs in period t, and the expected future costs, as in the $K = 0$ model. In particular ,θt(x) = miny≥x{Kδ(y −x) + c(y −
> show annotation

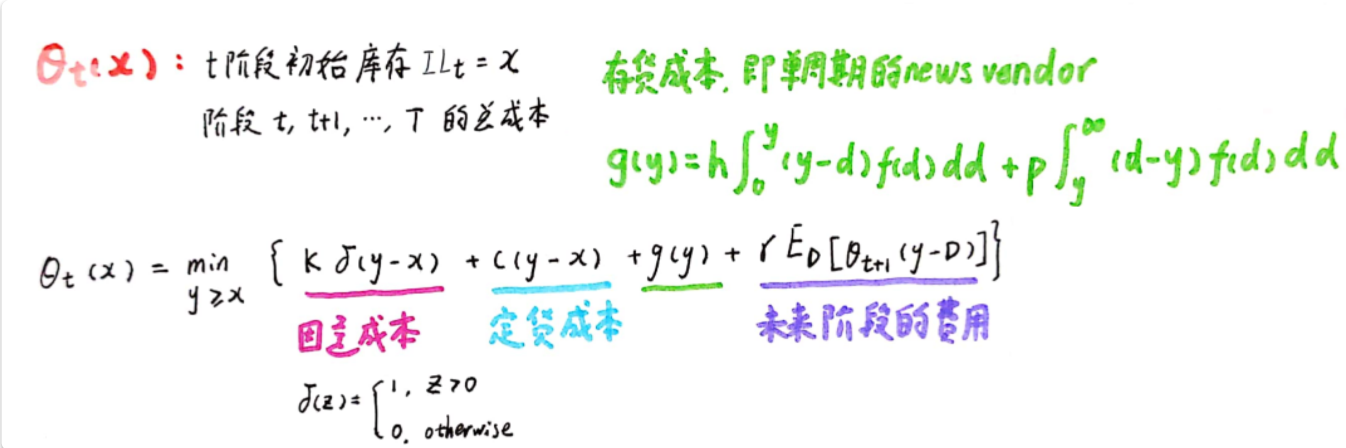
Now $\theta_t(x)$ must account for the fixed cost in period t (if any), as well as the purchase cost and expected holding and stockout costs in period t, and the expected future costs, as in the $K = 0$ model. In particular,

$$\theta_t(x) = \min_{y \geq x} \{K\delta(y - x) + c(y - x) + g(y) + \gamma \mathbb{E}_D[\theta_{t+1}(y - D)]\}$$

where

$$\delta(z) = \begin{cases} 1, & \text{if } z > 0 \\ 0, & \text{otherwise} \end{cases}$$

Note : Since here fixed cost is not zero, so add the fixed costs term $K\delta(y - x)$. It also consists of the general equation of periodic review, as follows (figure from Link):



> each startinginventory level x. However, just as before, we would rather have a simple policy to follow, rather than having to specify $y_t(x)$ for every t and x . And, just as before, this is a
> show annotation

> oint at which we stop ordering. In particular, for period t, there are values S_t and s_t such that for $x \leq s_t$, we have $y_t(x) = S_t$, and for $x > s_t$, we have $y_t(x) = x$. In other words, these curves each depict an (s, S) policy. We will prove in Section 4.5.2.
> show annotation

4.4.4 infinite Horizon

> 0: yt(x).4.4.4 Infinite Horizon Recall that the infinite-horizon model with no fixed costs (Section 4.3.4) is as simple asthe single-period model (Section 4.3.2). Unfortunately, this is not true
> show annotation

> -period model (Section 4.3.2). U nfortunately, this is not true in the fixed-costcase. The infinite-horizon model is more difficult than its single-period or finite-horizoncounterparts . To analyze it, we will need a
> show annotation

> ll need a bit of renewal theory. A renewal process is a random variable $N(t)$ that counts the number of “renewals”that have occurred by time t, where the amount of time between the $(n-1)st$ renewal and the nth renewals is a

> random variable X_n . The Xn are independent and ide
> show annotation

4.5 Policy Optimility

> (optimal) 4.5 POLICY OPTIMALITY Now that we know how to find the optimal S for a base-stock policy (Section 4.3) and the optimal s and S for an (s, S) policy (Section 4.4), we prove that those policy types are in fact optimal for their respective problems . In a way this is a lot to ask—
> show annotation

> In a way this is a lot to ask—w e are trying toshow that no other policy, of any type, using any parameters, can outperform our chosenpolicy type (provided we choose the optimal parameters) in the long run . Fortunately, wedo not need to
> show annotation

> imal policy has the desiredform. We will first consider the zero-fixed-cost case, then the fixed-cost case, in both casesconsidering single-period, finite-horizon, and infinite-horizon cases separately . We willuse the same assumption
> show annotation

> ion as in Section 4.4, as well. We continue to assumethat the cost and demand parameters are stationary, but the results below still hold if thesevary from period to period (deterministically). Let's focus for a minute on fini
> show annotation

Recall the optimal cost in periods $t, \cdots, T$ if we begin period t with an inventory level of x, can be calculated recursively as:

$$\theta_t(x) = \min_{y \geq x} \{K\delta(y-x) + c(y-x) + g(y) + \gamma \mathbb{E}_D[\theta_{t+1}(y-D)]\}$$

> entory level xin each period t. Our goal throughout this section will be to use the structure of (4.81) toshow that the optimal actions follow the policies we

> *have conjectured are optimal* .4.5.1 Zero Fixed Costs: Base-St
> show annotation

4.5.1 Zero Fixed Costs: Base-Stock Policies

> *Fixed Costs: Base-Stock Policies We first consider the case in which K = 0 and prove that—regardless of the horizonlength—a base-stock policy is always optimal . These results date back to Kar*
> show annotation

Single Period:

we'll consider the special case in which $T = 1$ and $K = 0$. We'll also assume that the terminal cost function (i.e., $\theta_{t+1}(x) = 0$, see Section 4.3.3) is equal to 0. Then, the optimal cost reduces to:

$$\theta(x) = \min_{y \geq x} \{c(y - x) + g(y)\}$$

We cam rewrite $\theta(x)$ as

$$\theta(x) = \min_{y \geq x} \{H(y) - cx\}$$

where

$$H(y) = cy + g(y)$$

Since we are calculating $\theta(x)$ for fixed x, we see that the optimal decision can be found by minimizing $H(y)$ over $y \geq x$, that is, starting at $y = x$, we want to minimize $H(y)$ looking only "to the right" of x.

If $H(y)$ is convex, we can have:

- If $x < S$, then the optimal strategy is to set $y = S$
- if $x \geq S$, the optimal strategy is to do nothing, to set $y = x$. In other words, the optimal policy is a *base-stock policy*.

> *a base-stock policy is optimal. And H(y) is convex because g(y) is convex,so we have now sketched the proof of the following theorem. SyH(y)(a) H(y)*

> convex; base-stoc
> show annotation

> cal shapes of the function H(y). Theorem 4.10 A base-stock policy is optimal for the single-period problem with no fixedcosts. What if H(y) is nonconvex? (This
> show annotation

Finite Horizon:

> IC REVIEW4.5.1.2 Finite Horizon It was simple to prove that H(y) is convex, and therefore thata base-stock policy is optimal, for the single-period problem . Our main goal in this sectionw
> show annotation

4.5.2 Nonzero Fixed Costs: (s, S) Policies

> nzero Fixed Costs: (s,S)Policies We now allow K 6= 0 and prove that an (s,S) policy is optimal. We will present formalproofs fo
> show annotation

4.6 Lost Salse

> (See Zheng(1991).)4.6 LOST SALES Throughout this chapter, we have assumed that unmet demands are backordered. In thissection, we assume instea
> show annotation

> re backordered. In thissection, we assume instead that they are lost. The distinction is only important when T > 1.(When T = 1, unmet demands can only be lost.) 4.6.1 Zero Lead TimeIn this sect
> show annotation

4.6.1 Zero Lead Time

> ssume that the lead time L = 0. First consider the case in whichK = 0. In the finite-horizon model, the DP recursion (4.36) changes only slightly :θt(x) =

> miny≥x{c(y −x) + g(y) +
> show annotation

$$\theta_t(x) = \min_{y \geq x} \left\{ c(y-x) + g(y) + \gamma \mathbb{E}_D[\theta_{t+1}(y-D)]^+ \right\}$$

Note: the positive part of $y - D$ to reflect the fact that the inventory level cannot become *negative*.

> *hout modification.LOST SALES 137 A base-stock policy is still optimal for the infinite-horizon model. Under the average-cost criterion (γ = 1) with lost sales, it is no longer true that we or*
> show annotation

5. Stochastic Inventory Models: Continuous Review

5.1 (r, Q) Policies

> *TINUOUS REVIEW5.1 (r,Q) POLICIES In this chapter, we consider a setting similar to the economic order quantity (EOQ) model(Section 3.2) but with stochastic demand. The mean demand per year is λ.*
> show annotation

> *The mean demand per year is λ. The inventory position is monitored continuously, and orders may be placed at any time. There is a deterministic lead time $L(\geq 0)$. Unmet demands are backordered .If the demand has a continuous*
> show annotation

> *. Unmet demands are backordered. If the demand has a continuous distribution, then the inventory level decreases smoothlybut randomly over time, with rate λ, as in Figure 5.1. (Think of liq*
> show annotation

(r, Q) **Policy**:

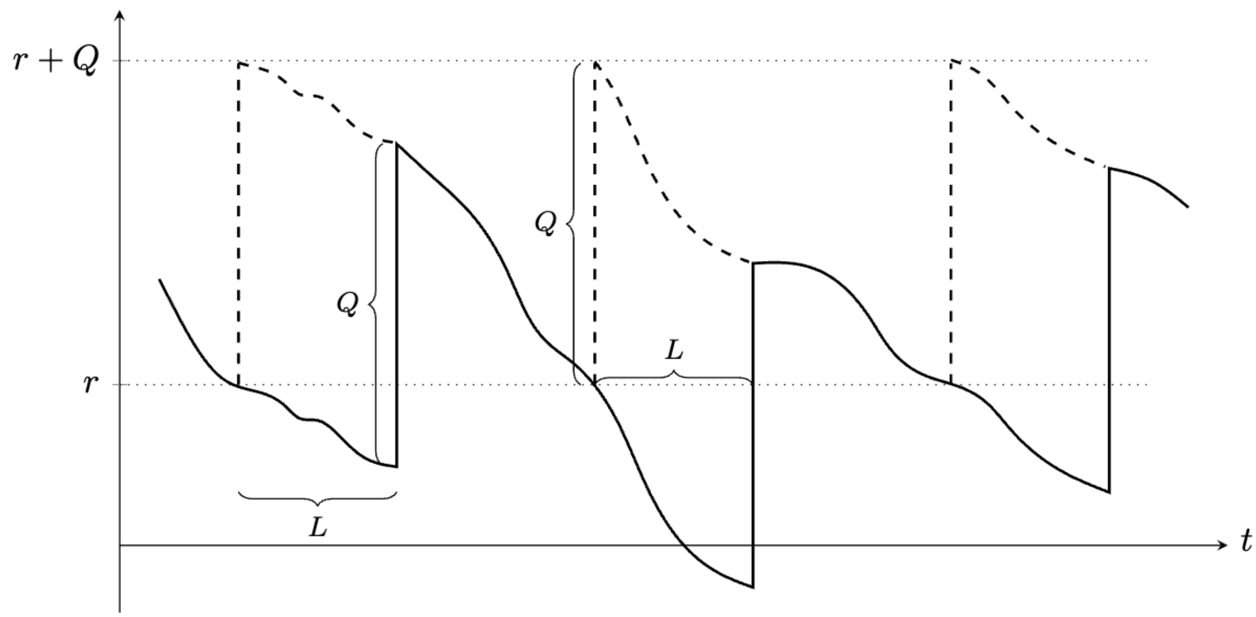


Figure 5.1 Inventory level (solid line) and inventory position (dashed line) under (r, Q) policy.

> sson process, as in Section 5.5. We'll assume the firm follows an (r,Q) policy: When the inventory position reaches acertain point (call it r), we place an order of size Q. L years later, the order arrives . In theintervening time, the in
>
> show annotation

> and, or wemay have stocked out. Note that the inventory level (solid line in Figure 5.1) and inventoryposition (dashed line) differ from each other during lead times but coincide otherwise. An(r,Q) policy is known to be o
>
> show annotation

Differing from the EOQ model, which has a single decision variable Q, the (r, Q) Policy has two decision variables: Q (the order quantity, sometimes called the *batch size*) and r (the *reorder point*)

> er (r,Q) policy.reorder point). Our goal is to determine the optimal r and Q to minimize the expected costper year. In a continuous-review setting,
>
> show annotation

> or for periodic-review systems, since in either case theinventory position may fall strictly below the reorder point before a replenishment order isplaced. In

> *this chapter, we will focus f*
> [show annotation]

5.2 Exact (r, Q) Problem With Continuous Demand Distribution

> *ND DISTRIBUTIONIn this section, we introduce an exact model for systems with continuous demand distribu-tions. We first formulate the expected cost function and then derive optimality conditionsfor it .We continue to consider the usu*
> [show annotation]

The usual costs:

- Fixed cost: $K \geq 0$
- Purchase cost: $c \geq 0$
- Holding cost: $h > 0$
- Stockout cost: $p > 0$
- D: the lead-time demand (*运输期间*的需求), a random variable with mean μ and variance σ^2, pdf $f(d)$ and cdf $F(d)$

5.2.1 Expected Cost Function

> *ime.5.2.1 Expected Cost Function Our first step is to derive an exact expression for the expected cost as a function of r and Q. We place orders, on average, eve*
> [show annotation]

First, we place orders, on average every Q/λ years. There the expected fixed cost is given by $K\lambda/Q$.

If the inventory position at time t is given by $IP(t)$, then the inventory level at time $t + L$ is given by:

$$IL(t+L) = IP(t) - (t, t+L]$$

As in the periodic-review case, we can drop the time indices in steady state and write:

$$IL = IP - D$$

where D is the lead-time demand.

> *of stochasticlead-time settings. Once we determine the distribution of IP, the (unconditional) expected inventory costthen follows from the law of total expectation. In particular, let ̄g(x) be the*
> show annotation

Let $\bar{g}(x)$ be the rate at which the inventory cost accrues when $IL = x$:

$$\bar{g}(x) = hx^+ + px^-$$

Note: $g(\cdot)$ is a rate because the inventory level is changing continuously over time, given in units of money per year. Then the expected inventory cost per year is:

$$\begin{aligned} \mathbb{E}[\text{ inventory cost }] &= \mathbb{E}_{IL}[\bar{g}(IL)] \\ &= \mathbb{E}_{IP}\left[\mathbb{E}_{IL|IP}[\bar{g}(IL)]\right] \\ &= \mathbb{E}_{IP}\left[\mathbb{E}_{D}[\bar{g}(IP - D)]\right] \\ &= \mathbb{E}_{IP}[g(IP)] \end{aligned}$$

where

$$g(y) = h\mathbb{E}[(y - D)^+] + p\mathbb{E}[(D - y)^+]$$

$g(y)$ is the rate at which the expected inventory cost accrues at time $t + L$ when the inventory position at time t equals y.

> *e itsoptimizer, given by (4.17). It remains to determine the distribution of IP. By the definition of an (r, Q) policy, we know that IP takes values only in $[r, r + Q]$. It turns out that IP has a very*
> show annotation

在一些简单例子中，IP 可以是简单的分布，如均匀分布。因此，我们可以用如下积分形式计算 *期望损失*:

$$\mathbb{E}[\text{inventory cost}] = \frac{1}{Q}\int_r^{r+Q} g(y)\,\mathrm{d}\,y$$

Combining the expected inventory cost and the expected fixed cost $K\lambda/Q$, we have the expected total cost per year:

$$g(r, Q) = \frac{K\lambda + \int_r^{r+Q} g(y)\,\mathrm{d}\,y}{Q}$$

> *eng (1992) proves the following: Lemma 5.1 g(r,Q) is jointly convex in r and Q. Proof. LetI(r,Q) = 1Q∫ r+QrE[(y*
> [show annotation]

> *r,Q) is proven byZipkin (1986a). In what follows, we use the expected cost expression (5.7) to derive optimality conditionsfor r and Q by first fixing Q and finding the optimal corresponding r, and then optimizingover Q . Although these conditions tell*
> [show annotation]

5.2.2 Optimality Conditions

> *) be the optimal r for agiven Q. Lemma 5.2 For any $Q > 0, r = r(Q)$ if and only if g(r) = g(r + Q). (5.9)Proof. Fol*
> [show annotation]

Lemma For any $Q > 0, r = r(Q)$ if and only if

$$g(r) = g(r + Q)$$

> *r, due to the convexity of g(y). The motivation behind this result is that, during one replenishment cycle, we need topass through all of the inventory positions in [r,r + Q] , and we spend an equal amount o*
> [show annotation]

Theorem (r, Q) minimize $g(r, Q)$ if and only if:

$$g(r, Q) = g(r + Q) = g(r)$$

5.3 Approximations for (r, Q) Problem With Continuous Distribution

5.3.1 Expected-Inventory-Level Approximation

> *ed-Inventory-Level Approximation The first approximation we discuss is probably the best known and most widely coveredapproximation to find r and Q. (Unfortunately, it is also one o*
> *show annotation*

> *nts byHadley and Whitin (1963). We call this the expected-inventory-level (EIL) approximation,for reasons that will become clear shortly .The approach relies on the foll*
> *show annotation*

> *that will become clear shortly. The approach relies on the following two simplifying assumptions to make the modeltractable: • Simplifying Assumption 1 (SA1)*
> *show annotation*

Two assumptions:

1. Simplifying Assumption 1 (SA1): Incur holding costs at a rate of $h \cdot IL$ per year, where IL is the inventory level, whether IL is positive or negative.
2. Simplifying Assumption 2 (SA2): The stockout cost is charged once per unit of unmet demand, not per year.

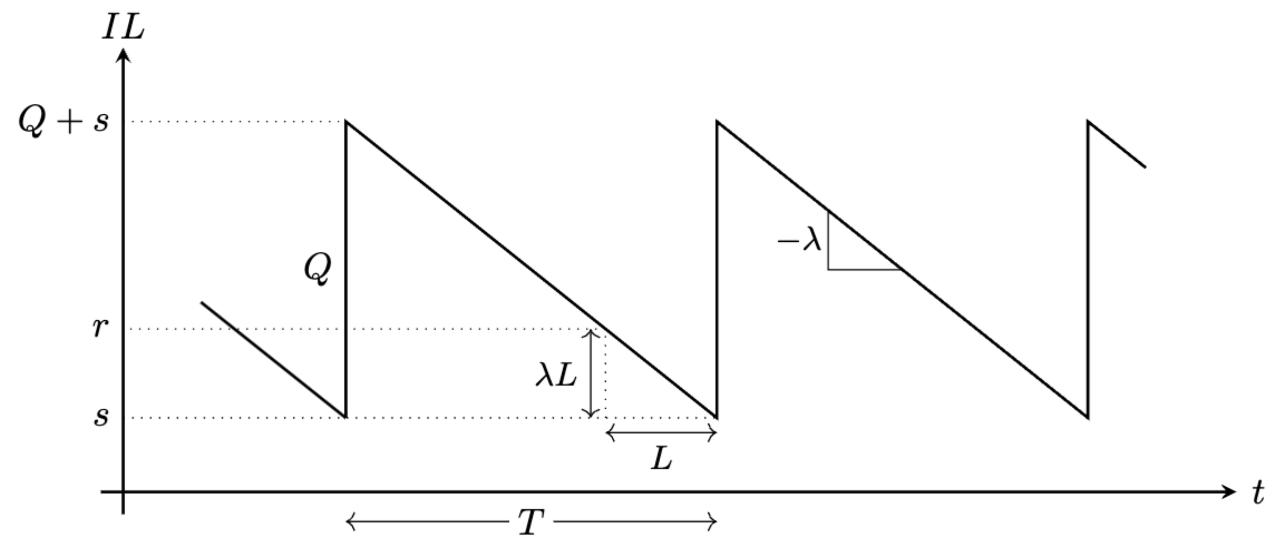


Figure 5.3 Expected inventory curve for (r, Q) policy.

The costs:

- The expected on-hand inventory when the order arrives:

$$s = r - \lambda L$$

- The average inventory level:

$$s + \frac{Q}{2} = r - \lambda L + \frac{Q}{2}$$

1. By *SA1*, the *expected holding cost* per year is:

$$h\left(r - \lambda L + \frac{Q}{2}\right)$$

Note: this expression is only approximate, since that we are calculating the expected holding cost as $h \cdot \mathbb{E}[IL]^+$ (provided that $\mathbb{E}[IL] > 0$).

> *L+], and the two are not equal. That is why we refer to this as the "expected-inventory-level" approximation. The problem is more difficult without SA1 because of the nonlinearity introduced by the $[\cdot]^+$ operator. Aspreviously noted, the EIL app*
>
> show annotation

2. *Fixed cost* : $\frac{K\lambda}{Q}$

- *Stockout cost* :

$$\mathbb{E}[(D-r)^+] = \int_r^\infty (d-r)f(d)\,\mathrm{d}\,d = n(r)$$

> *=∫ ∞r(d −r)f(d)dd = n(r), (5.14)* *where $n(r)$ is the loss function for the lead-time demand distribution* . *(See Section 4.3.2.2or Section*
> *show annotation*

The expected number of stockout per year is

$$\frac{n(r)}{\mathbb{E}[T]} = \frac{\lambda n(r)}{Q}$$

3. then by SA2, the *expected stockout cost* per year is simply:

$$\frac{p\lambda n(r)}{Q}$$

> *r year is simplypλn(r)Q . (5.15)* *Note that we are assuming that r > 0, which is a reasonable assumption in practice* . *(Thereason we make simplifying*
> *show annotation*

The total cost per year:

$$g(r,Q) = h\left(r - \lambda L + \frac{Q}{2}\right) + \frac{K\lambda}{Q} + \frac{p\lambda n(r)}{Q}$$

Solution:

- For Q :

$$
\begin{aligned}
\frac{\partial g}{\partial Q} = \frac{h}{2} - \frac{K\lambda}{Q^2} - \frac{p\lambda n(r)}{Q^2} &= 0 \\
\iff \quad \frac{1}{Q^2}[K\lambda + p\lambda n(r)] &= \frac{h}{2} \\
\iff \quad Q^2 &= \frac{2[K\lambda + p\lambda n(r)]}{h} \\
Q &= \sqrt{\frac{2\lambda[K + pn(r)]}{h}}
\end{aligned}
$$

- For r :

$$
\begin{aligned}
\frac{\partial g}{\partial r} = h + \frac{p\lambda n'(r)}{Q} &= 0 \\
\iff \quad h + \frac{p\lambda(F(r) - 1)}{Q} &= 0 \\
r = F^{-1}\left(1 - \frac{Qh}{p\lambda}\right)
\end{aligned}
$$

> *5)), sor = F−1(1 − Qhpλ). (5.18) Now we have two equations with two unknowns, but these equations cannot be solvedin closed form. The approach given in Algorithm*
> *show annotation*

> *cannot be solvedin closed form. The approach given in Algorithm 5.1 first sets Q equal to the EOQ quantity,i.e., ignoring the demand randomness. It then proceeds iteratively, s*
> *show annotation*

> *le 5.1. 5.3.1.3 Service Levels One major limitation of (r,Q) policies as formulated aboveis that p is very hard to estimate . But there is a close relations*
> *show annotation*

5.3.2 EOQB Approximation

> *139.1. 5.3.2 EOQB Approximation There are important connections between the EOQ problem with planned backorders (EOQB; Section 3.5) and (r,Q) po*
> *show annotation*

5.3.3 EOQ + SS Approximation

> *5.2. 5.3.3 EOQ+SS Approximation* *Another common approximation for r and Q is to convert the inventory-cost parametersinto a service level and then to use the approach described in Section 5.3.1.3 for type-1service level constraints* *. In particular,Q =√2KλhAPPROXIM*
> show annotation

5.4 Exact (r, Q) Problem With Continuous Distribution: Properties of Optimal r and Q

> *ON:PROPERTIES OF OPTIMAL r AND Q* *We now return to the exact model from Section 5.2. We have two main goals in thissection* *. First, we will analyze the pro*
> show annotation

1. We will analyze the properties of optimal solutions (and their costs) for (r, Q) policies, by deriving *optimality conditions* for r and Q and then providing *properties* of the resulting optimal solutions.
2. We will compare (r, Q) policies to the EOQB model and prove.

> *to the EOQBmodel and prove that,* *if the EOQB model is used as a heuristic for optimizing r and Q, asdiscussed in Section 5.3.2, the resulting error has a fixed bound.* *We do this by treating theEOQB*
> show annotation

Let $G(Q)$ equal the expected cost per year as a function of Q, assuming r is set optimally for that Q,

$$G(Q) = g(r(Q), Q)$$

Let $H(Q)$ be the value of $g(y)$ at $y = r(Q)$, or $r(Q) + Q$

$$H(Q) = g(r(Q)) = g(r(Q) + Q)$$

Then we have

$$G(Q)=\frac{K\lambda+\int_0^Q H(y)\,\mathrm{d}\,y}{Q}$$

6. Multi-echelon Inventory Models

> *.1 INTRODUCTIONIn this chapter, we study inventory optimization models for multiechelon (or multistage)systems with shipments made among the stages . There are two common ways to i*
> show annotation

6.1 Introduction

There are two common ways to interpret the stages or nodes in the multiechelon system:

1. Stages represent locations in a supply chain network, and links among the stages represent physical shipments of goods.
2. Stages represent processes that the product must undergo during manufacturing, assembly, and/or distribution.

> *hipments of goods. For example, the stages in Figure 6.1(a)may represent the following physical locations: a supplier in China, a factory inCalifornia, a warehouse in Chicago, and a retailer in Detroit (respectively). 2. Stages represent processes th*
> show annotation

> *ns betweensteps in the process. For example, the stages in Figure 6.1(a) may represent thefollowing processes: manufacturing, assembly, testing, and packaging. These fourfunctions may take place in four different locations or all within the same building —it is largely irrelevant from t*
> show annotation

8. Facility Location Models

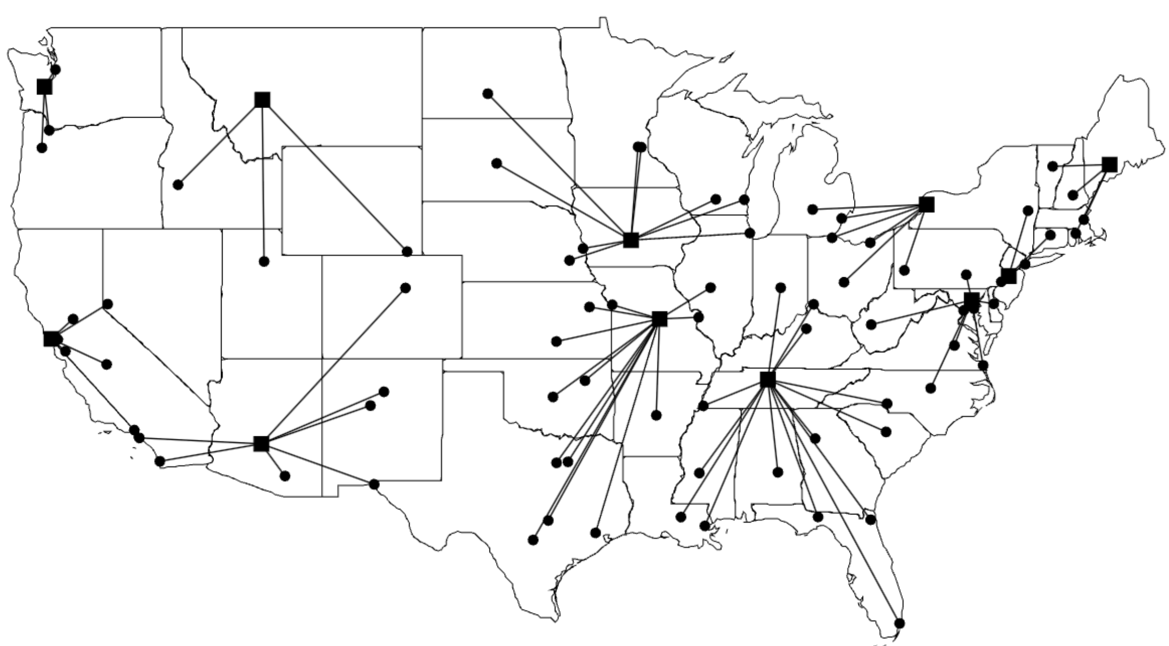
8.1 Introduction

> *LOCATION MODELS8.1 INTRODUCTION* One of the major strategic decisions faced by firms is the number and locations of factories,warehouses, retailers, or other physical facilities *. This is the purview of a large*
> [show annotation]

> *as facility location problems.* The key trade-off in most facility locationproblems is between the facility cost and customer service. *If we open a lot of facilities(*
> [show annotation]

- More facilities

> *lity cost and customer service.* If we open a lot of facilities(Figure 8.1(a)), we incur high facility costs (to build and maintain them), but we can providegood service since most customers are close to a facility. *On the other hand, if we openfe*
> [show annotation]

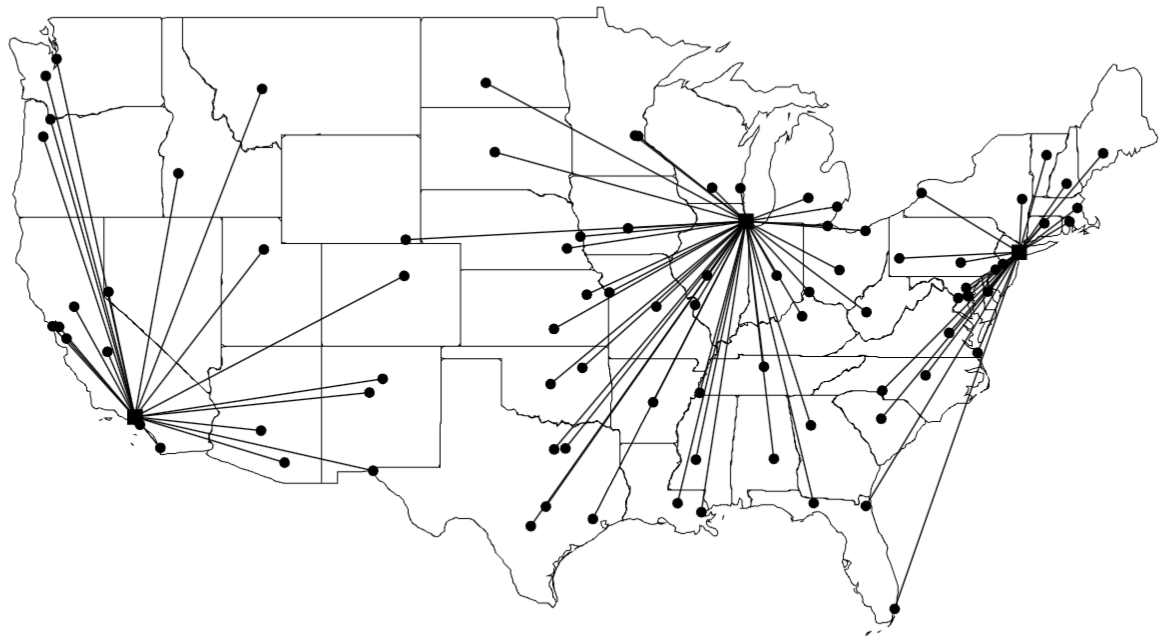


(a) Many facilities open.

- Few facilities

a facility. On the other hand, if we openfew facilities (Figure 8.1(b)), we reduce our facility costs but must travel farther to reachour customers (or they to reach us). Most (but not all) location prob

show annotation



(b) Few facilities open.

customers (or they to reach us). Most (but not all) location problems make two related sets of decisions: (1) where tolocate, and (2) which customers are assigned or allocated to which facilities. Therefore,facility location prob

show annotation

allocated to which facilities. Therefore,facility location problems are also sometimes known as location–allocation problems .A huge range of approaches has

show annotation

ing facility location decisions. These differ in terms of how they model facility costs (for example, some include the costsexplicitly, while others impose a constraint on the number of facilities to be opened) andhow they model customer service (for example, some include a transportation cost, whileothers require all or most facilities to be covered—that is, served by a facility that is withinsome specified distance). Facility location problems

> *come*
> *show annotation*

> *ion as well as later extensions.* *In addition to supply chain facilities such as plants and warehouses, location modelshave been applied to public sector facilities* *such as bus depots and fire sta*
> *show annotation*

> *notreally "facilities" at all.* *In addition, many operations research problems can be formulatedas facility location problems or have subproblems that resemble them* *. Facility locationproblems are*
> *show annotation*

> *th theoretical and applied work.* *In this chapter, we will begin by discussing a classical facility location model, theuncapacitated fixed-charge location problem (UFLP)* *, in Section 8.2. The UFLP and i*
> *show annotation*

> *the UFLP), and in Section 8.4,* *we discuss cover-ing models (including the p-center, set covering, and maximal covering problems)* *. Webriefly discuss a variety of*
> *show annotation*

8.2 The Uncapacitated fixed-charge Location Problem

8.2.1 Problem Statement

> *N PROBLEM8.2.1 Problem Statement* *The uncapacitated fixed-charge location problem (UFLP) chooses facility locations inorder to minimize the total cost of building the facilities and transporting goods fromfacilities to customers.* *The UFLP makes location decisio*
> *show annotation*

> *oreven fire stations and homes.* *Sometimes it's also useful to think of an upstream echelon,again with fixed location(s), that serves the DCs.* *Each*

> *potential DC location has a*
> *show annotation*

- Fixed cost

> *ocation(s), that serves the DCs. Each potential DC location has a fixed cost that represents building (or leasing) thefacility; the fixed cost is independent of*
> *show annotation*

- Transportation cost

> *is270 FACILITY LOCATION MODELSa transportation cost per unit of product shipped from a DC to each customer. There isa single product. The D*
> *show annotation*

- *Objective*

> *s assumption in Section 8.3.1.) The problem is to choosefacility locations to minimize the fixed cost of building facilities plus the transportationcost to transport product from DCs to customers, subject to constraints requiring everycustomer to be served by some open DC .As noted above, the key trade-o*
> *show annotation*

> *er to be served by some open DC. As noted above, the key trade-off in the UFLP is between fixed and transportation costs .If too few facilities are open,*
> *show annotation*

8.2.2 Formulation

Define the following notations:

Sets:

- I : set of customers

- J: set of potential facility locations
  Parameters
- h_i : annual demand of customer $i \in I$
- c_{ij} : cost to transport one unit of demand from facility $j \in J$ to customer $i \in I$
- f_j : fixed annual cost to open a facility at site $j \in J$
  Decision Variables
- x_j : 1 if facility j is opened, 0 otherwise
- y_{ij} : the fraction of customer i's demand that is served by facility j

Note : The transportation costs c_{ij} might be of the form $k \times$ distance for some constant k (if the shipping company charges k per mile per unit) or more arbitrary (for example, based on airline ticket prices, which are not linearly related to distance)

> *linearly related to distance). In the former case,distances may be computed in a number of ways: • Euclidean distance: The distan*
> *show annotation*

Distances:

- *Euclidean distance*:
- *Manhattan* or *rectilinear metric*
- *Great circle*
- *Highway/network*
- *Matrix*

> *based on airline ticket prices. In general, we won't be concerned with how transportation costs are computed—we'll assume they are given to us already as the parameters c_{ij} .The UFLP is formulated as follo*
> *show annotation*

The UFLP is formulated as follows:

$$\min \quad \sum_{j \in J} f_j x_j + \sum_{i \in I} \sum_{j \in J} h_i c_{ij} y_{ij}$$

subject to

$$
\begin{aligned}
\sum_{j \in J} y_{ij} &= 1 && \forall i \in I \\
y_{ij} &\leq x_j && \forall i \in I, \forall j \in J \\
x_j &\in \{0, 1\} && \forall j \in J \\
y_{ij} &\geq 0 && \forall i \in I, \forall j \in J
\end{aligned}
$$

Note : in the discussion that follows, we'll use z^* to denote the optimal objective value of (UFLP).

> *anne (1964) and Balinski(1965). The objective function (8.3) computes the total (fixed plus transportation) cost. Inthe discussion that follows,*
> *show annotation*

> *together ensure that 0 ≤yij ≤1. In fact, it is always optimalto assign each customer solely to its nearest open facility. (Why?) Therefore, there alwaysex*
> *show annotation*

> *objective value is no greater. It is important to understand that theIPs have the same optimal objective value, but the LPs have different values—one providesa weaker LP bound than the other. The UFLP is NP-hard (Garey and J*
> *show annotation*

> *exact algorithmsand heuristics. Some of the earliest exact algorithms involve simply solving the IP usingbranch-and-bound . Today, this would mean solving*
> *show annotation*

> *o solve problems ofmodest size. Therefore, a number of other optimal approaches were developed. Two ofthese—Lagrangian relaxation and a dual-ascent method called DUALOC—are discussedin Sections 8.2.3 and 8.2.4, respectivel y. Many other IP techniques, suc*
> *show annotation*

8.2.3 Lagrangian Relation

> Relaxation8.2.3.1 Introduction One of the methods that has proven to be most effective forthe UFLP and other location problems is Lagrangian relaxation, a standard technique forinteger programming (as well as other types of opti
> show annotation

> ARGE LOCATION PROBLEM 273behind Lagrangian relaxation is to remove a set of constraints to create a problem that'seasier to solve than the original. But instead of just removing th
> show annotation
>
> Actually, it is to construct a convex function, then let is derivative to zero.

> ethem in the objective function by adding a term that penalizes solutions for violating theconstraints . This process gives a lower bou
> show annotation

> on the optimal objective value. When the upper and lower bounds are close (say, within1%), we know that the feasible solution we have found is close to optima l.For more details on Lagrangian
> show annotation

> are relaxed.8.2.3.2 Relaxation We relax constraints (8.4), removing them from the problem andadding a penalty term to the objective function: ∑i∈Iλi⎛⎝⎜⎜1 −∑j∈Jyij⎞⎠⎟⎟The λi are c
> show annotation

$$\sum_{j \in J} \lambda_i (1 - y_{ij})$$

The λ_i are called *Lagrange multipliers*.

> present the vector of λi values. For now, assume λ is fixed. Relaxing constraints (8.4) gives us the following problem,known as the Lagrangian subproblem: (UFLP-LRλ) minimize ∑j∈Jfjxj + ∑
> show annotation

The *Lagrangian subproblem* :

$$
\begin{aligned}
\min \sum_{j\in J} f_j x_j + \sum_{i\in I}\sum_{j\in J} h_i c_{ij} y_{ij} + \sum_{j\in J} \lambda_i (1 - y_{ij}) \\
= \sum_{j\in J} f_j x_j + \sum_{i\in I}\sum_{j\in J} (h_i c_{ij} - \lambda_i) y_{ij} + \sum_{j\in J} \lambda_i
\end{aligned}
$$

subject to

$$
\begin{aligned}
y_{ij} \le x_j \qquad & \forall i \in I, \forall j \in J \\
x_j \in \{0, 1\} \qquad & \forall j \in J \\
y_{ij} \ge 0 \qquad & \forall i \in I, \forall j \in J
\end{aligned}
$$

> *.How can we solve this problem?* *It turns out that the problem is quite easy to solveby inspection—we don't need to use an IP solver or any sort of complicated algorithm.* *274 FACILITY LOCATION MODELSSupp*
>
> show annotation

> *ginal problem:zLR(λ) ≤z≣. (8.16)* *The point of Lagrangian relaxation is not to generate feasible solutions, since the solutionsto (UFLP-LRλ) will generally be infeasible for (UFLP). Instead, the point is to generategood (i.e., high) lower bounds in order to prove that a feasible solution we've foundsome other way is good.* *For example, if we've found a f*
>
> show annotation

8.2.4 The DUALOC Algorithm

8.3 Other Minisum Models

> *. 2006).8.3 OTHER MINISUM MODELS* *The UFLP is an example of a minisum location problem. Minisum models are so calledbecause their objective is to minimize a sum of the costs or distances between customers andtheir assigned facilities* *(as well as possibly other term*
>
> show annotation

> *ch as fixed costs). In contrast,* *covering location problems are more concerned with the maximum distance, with the goalof ensuring that most or all customers are located close to their assigned facilities* *.296 FACILITY*

> *LOCATION MODELSAt*
> [show annotation]

> *eneralizing, it can be said that* *minisum models are more commonlyapplied in the private sector, in which profits and costs are the dominant concerns, andcovering models are most commonly applied in the public sector, in which service, fairness,and equity are more important* *. For further discussion of this*
> [show annotation]

8.3.1 The Capacitated Fixed-Charge Location Problem

> *ion in many practical settings.* *The UFLP can be easilymodified to account for capacity restrictions; the resulting problem (not surprisingly) iscalled the capacitated fixed-charge location problem, or CFLP.* *Suppose vj is the maximumdemand*
> [show annotation]

8.4 Covering Models

> *many other guidelines, says that* *fire departments should have the objective of arriving to afire within 4 minutes of receiving a call* *(National Fire Protection Associ*
> [show annotation]

> *isum models can helpmuch with, s* *ince the optimal solutions to those problems may assign some customers tovery distant facilities if it is cost effective to do so* *. Instead, we need to use the no*
> [show annotation]

> *it is cost effective to do so.* *Instead, we need to use the notion ofcoverage, which indicates whether a given customer is within a prespecified distance, orcoverage radius, of an open facility.* *For example, Figure 8.11 shows t*
> [show annotation]

> om the equator.In this section, *we discuss three seminal facility location models that use coverage todetermine the quality of the solution* . The first, the set covering lo
> show annotation

1. Set covering location problem (SCLP): locates the minimum number of facilities to cover every demand node
2. *maximal covering location problem* (MCLP): covers as many demands as possible while locating a fixed number of facilities.
3. *p-center problem*: locates a fixed number of facilities to minimize the maximum distance from a demand node to its nearest open facility.

8.4.1 The Set Covering Location Problem

> Covering Location Problem (SCLP) *In the set covering location problem (SCLP), we are required to cover every demand node;the objective is to do so with the fewest possible number of facilities.* The SCLP was firstformulated in
> show annotation

Parameters:

- $a_{ij} = 1$: if facility $j \in J$ can cover customer $i \in I$ (if it is open), 0 otherwise.

> ∈I (if it is open), 0 otherwise *The coverage parameter a_{ij} can be derived from a distance or cost parameter such as c_{ij} in the UFLP,* for example:aij ={1, if cij ≤r0
> show annotation

Then, the SCLP can be formulated as follows:

$$\min \sum_{j \in J} x_j$$

subject to

$$\sum_{j \in J} a_{ij} x_j \geq 1 \qquad \forall i \in I$$

$$x_j \in \{0, 1\} \qquad \forall j \in J$$

> *82) are integrality constraints. Sometimes we wish to minimize the total fixed cost of the opened facilities, rather than the total number, in which case the following objective function is appropriate :minimize ∑j∈Jfjxj. (8.83)The SC*
> [show annotation]

$$\min \sum_{j \in J} f_j x_j$$

8.4.2 The Maximal Covering Location Problem

> *han 88.7%, as we will see below. The maximal covering location problem (MCLP) seeks to maximize the total number ofdemands covered subject to a limit on the number of open facilities. It was introduced byChurch and*
> [show annotation]