Stochastic Inventory Control

IE375 Spring 2020 On-Line Slides

Nesim K. Erkip

Meaning of Uncertainty

Suppose that we represent demand, similar to forecasting as

$$D = D_{deterministic} + D_{random}$$

If the random component is small compared to the deterministic component, the models of Chapter 4 will be accurate. If not, randomness must be explicitly accounted for in the model.

In this chapter, assume that demand is a *continuous* or *discrete* random variable with cumulative distribution function F(t) and probability density (mass for discrete) function f(t);

Estimation Problems — Recall Goodness-of-fit tests

Stochastic Inventory Control

- Inventory models we studied so far assume that all components of supply and demand are known with certainty
- In reality several components could be stochastic
 - Demand
 - Quantity & Timing
 - Supply
 - Yield
 - Lead time

Stochastic Inventory Control

- Shortage Costs (*p*): Costs incurred when the demand of a given period can not be satisfied (partially or fully) from stocks, i.e. current inventory level is not sufficient to satisfy the demand not necessarily all voluntarily
- Can be in two forms:
 - Backordering
 - Lost Sales (No backordering)
- In backordeing case, the demand of the customer that could not be satisfied on time is satisfied later at the first opportunity

Components of Inventory Models

- In **backordering** case, the costs are related to the loss of goodwill of customers and the subsequent reluctance to do business with the firm, the cost of delayed revenue (and/or discounts given), extra administrative costs, penalties as termed in a supply contract, etc.
- In **lost sales**, if the customer's demand is not satisfied on time then there is no backordering option. Alternatives are either losing that customer to a competitor or satisfying customer's demand with a priority shipment from chanelles which are not regular. The shortage costs reflect the costs of these alternatives.

Components of Inventory Models

- Continuous Review Models: Inventory can be reviewed continuously and a course of action (e.g. placement of an order of quantity z) can be taken according to the inventory level at a given time instant
- Periodic Review Models: A course of action can be taken according to the inventory level at end (or beginning) of a given time period (e.g. days, months, etc.)
- When an order is placed, there may be a replenishment **lead** time so that the order arrives after τ periods of time

Newsvendor model

Newsvendor Problem

- Perishable products can be carried in inventory for only a very limited time before loosing its value
- We will discuss a single period version of this problem which is also (historically) known as the "newsboy problem" corrected for gender as "newsvendor problem"
- The item being sold can no longer be sold after the period ends
- Unmet demand is lost, excess inventory is discard

Newsvendor Problem

- For example, a daily newspaper being sold at a newsstand can be carried in the inventory only for a single day
- It is outdated at the end of the day, and must be replaced the next day
- The demand of the newspaper is a random variable since it cannot be predicted in advance
- The owner of the newsstand has to choose an order quantity at the start of each day without the knowledge of what the demand of day will be
- If he orders more than the coming day's demand, then the excess papers are wasted (or salvaged for recycling paper)

Perishable Products

- If he orders less than the coming day's demand, then there will be lost profit
- So what is the right quantity to order?
- Some other examples of perishable products:
 - Newspapers, magazines
 - Flowers
 - Seasonal fashion clothing
 - Fresh vegetables & fruits, cake, sushi, etc
 - Computer parts
 - Airline seats, hotel rooms
 - Christmas Tree, Güllaç

Newsvendor Model - Assumptions

- Single perishable product sold in a single period
- Items remaining on hand can be salvaged (sold at clearance for sure)
- No initial inventory
- The decision variable is how much to order, y
- Demand is a random variable but its probability distribution is known
- The objective is to minimize expected total cost where the cost components are
 - c unit cost of purchasing the product
 - s salvage value of remaining units at the end of the period
 - *p* sales price of the product

An Example

- A winter sports store in Ankara is ordering snowboard pants for the upcoming winter season. These pants are imported from Hong Kong and the lead time is usually around 3-4 months. Since the snowboard season in Turkey is usually short around 4 months, the store can order only once before the season, considering the demand for the whole season. Further replenishments within the season are not possible
- Each pant costs the store 90 TL. The sales price during the season is 150 YL. If the pants cannot be sold until the last month of the season, they can be sold at clearance (sales) for 60 TL per unit.
- The demand at the regular season price is not known in advance, but its probability distribution can be determined.

Newsvendor Model

- Then total profit can be written as
- Profit=
 - +150 * Number of items sold by the store
 - 90 * Number of items purchased by the store
 - + 60 * Number unsold in season and sold in clearance
- Let
 - y = Number purchased by the distributor
 - D = Demand of the bicycle (a random variable)
- Number sold = $\min\{D, y\} = D \max\{0, D y\}$
- Number unsold = $\max\{0, y D\}$

Newsvendor Model

Profit

=150 min
$$\{D, y\}$$
 - 90 y + 60 max $\{0, y - D\}$
=150 D - 150 max $\{0, D - y\}$ - 90 y + 60 max $\{0, y - D\}$

Rewrite

$$150 D - 90 y = 60 D + 90 (D-y)$$
$$= 60 D + 90 \max \{0, D-y\} - 90 \max \{0, y-D\}$$

• Rewrite profit

$$= 60 D - 60 \max \{0, D - y\} - 30 \max \{0, y - D\}$$
$$= (p-c) D - (p-c) \max \{0, D - y\} - (c-s) \max \{0, y - D\}$$

Or cost

$$= 60 \max \{0, D - y\} + 30 \max \{0, y - D\}$$

$$= (p-c) \max \{0, D - y\} + (c-s) \max \{0, y - D\}$$

$$= c_u \max \{0, D - y\} + c_o \max \{0, y - D\}$$

Newsvendor Model Expected Total Cost Function

• The cost incurred when the demand is D:

$$C(D, y) = c_u \max\{0, D - y\} + c_o \max\{0, y - D\}$$

 If D is a discrete random variable with the probability mass function P_D(d)

$$C(y) = E[C(D, y)] = \sum_{d=0}^{\infty} C(d, y) P_D(d)$$

$$= \sum_{d=0}^{\infty} (c_u \max\{0, d-y\} + c_o \max\{0, y-d\}) P_D(d)$$

$$= \sum_{d=0}^{\infty} c_u (d-y) P_D(d) + \sum_{d=0}^{y-1} c_o (y-d) P_D(d)$$

Newsvendor Model Expected Total Cost Function

- Since a discrete probability distribution is hard to find, particularly when the demand ranges over a large number of possible values, the demand is often approximated with a continuous r.v.
- If demand is a continuous r.v. with the density function f(u)

$$C(y) = E[C(D, y)] = \int_{0}^{\infty} C(x, y) f(x) dx$$

$$= \int_{0}^{\infty} (c_u \max\{0, x - y\} + c_o \max\{0, y - x\}) f(x) dx$$

$$= \int_{0}^{\infty} c_u (x - y) f(x) du + \int_{0}^{y} c_o (y - x) f(x) dx$$

How to take the derivative of an integral? Leibnitz Rule

• Derivative of an integral

$$\frac{\partial}{\partial t} \int_{a(t)}^{b(t)} G(x,t) dx$$

$$= \frac{db(t)}{dt} G(b(t),t) - \frac{da(t)}{dt} G(a(t),t)$$

$$+ \int_{a(t)}^{b(t)} \frac{\partial G(x,t)}{\partial t} dx$$

Newsvendor Model - Optimal Policy

• The optimal order quantity y^* that minimizes C(y) is given by

$$F(y^*) = \frac{c_u}{c_u + c_o}$$

- C(y) is a convex function
- p-c: unit cost of under ordering = cost of underage = c_u
- c-s: unit cost of over ordering = cost of overage = c_o

$$F(y^*) = \frac{p - c}{p - s}$$

Newsvendor Model - Optimal Policy

- If *D* is assumed to be a discrete r.v. a similar result is obtained
- *y** is the smallest integer such that

$$F(y^*) \ge \frac{c_u}{c_u + c_o}$$

Marginal cost/revenue interpretation

True only for the given forms of mathematical functions
Optimality condition:

• marginal cost of overage=marginal cost of underage

$$c_u \Pr\{D \ge y^*\} = c_o \Pr\{D \le y^*\}$$

• marginal revenue = marginal cost

$$p \Pr\{D \ge y^*\} + s \Pr\{D \le y^*\} = c$$

Example continued

- Consider the snowboard shop again. If the regular season demand follows a normal distribution with mean 200 and standard deviation 50,
 - How many snowboard pants should the store order before the season to maximize its expected profits?
 - What is the expected profits if the store orders optimally?
 - What is the expected number of customers that will be turned down because the inventory is not available?
 - What is the expected number of snowboard pants that will be sold in clearance sales?

Newsvendor Problem Normally Distributed Demand

• When the demand is Normally distributed then $F(y^*)$ is computed using tables

$$F(y) = \int_{-\infty}^{(y-\mu)/\sigma} \phi(t)dt$$

- where $\phi(t)$ is the standard normal density
- If demand is Normal with mean μ and standard deviation σ then it can be shown that

$$F(y) = \Phi\left(\frac{y-\mu}{\sigma}\right)$$

Values of F(z) function are available in tabulated forms

Revenue management: an implementation of the Newsboy model

• Revenue management (aka yield management or demand management), is a business practice that can basically be described as a way for business to maximize expected revenue, and thereby expected profits, by selling their products to the right customer at the right price at the right time. Essentially, getting the most out of your supply by targeting segmented micro markets to maximize expected revenue.

Example

- December 25, flight TK107 leaving Ankara for İstanbul at 6:30 am.
- Aircraft assigned is an Airbus 310-304 with 210 seats (28 business class 182 economy)
- Two types of customers for the economy cabin
 - Business travelers: Book their tickets late. Charge their tickets to their companies, therefore less sensitive to price. Need flexibility.
 - Leisure travelers: Book well in advance. Pay for their own tickets. Sensitive to price. Do not need flexibility.
- Two fares
 - Full fare = 150 TL targeting business travelers. Can change the date or return the ticket without any charge. Can book any time as long as there is space
 - Discount fare = 90 TL targeting leisure travelers.
 Changes or returns with penalty. Should book at least 3 weeks in advance

Example



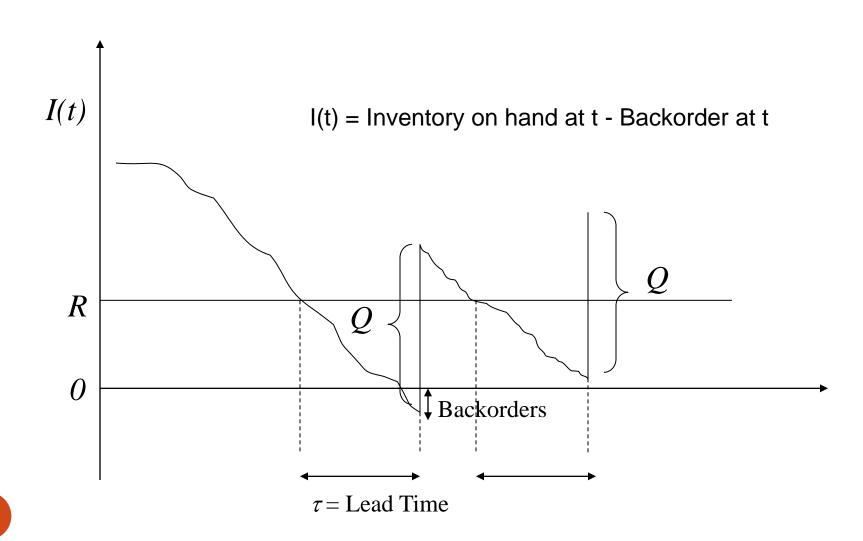
- How many discount tickets should we sell? Or when should we stop selling discount tickets?
- Trade-offs
 - If too many discount tickets are sold, then the airline foregoes the extra revenue from the full fare customers
 - If too few discount tickets are sold, then there may be empty seats at the flight
- Trivial case: deterministic demand
 - D_{discount}=300, D_{full-fare}=100
 - Sell only 82 discount seats (B*=82) and then close the discount class for sale, protect 100 seats for full fare customers

A Stochastic Continuous Review Model and (R, Q) Policy

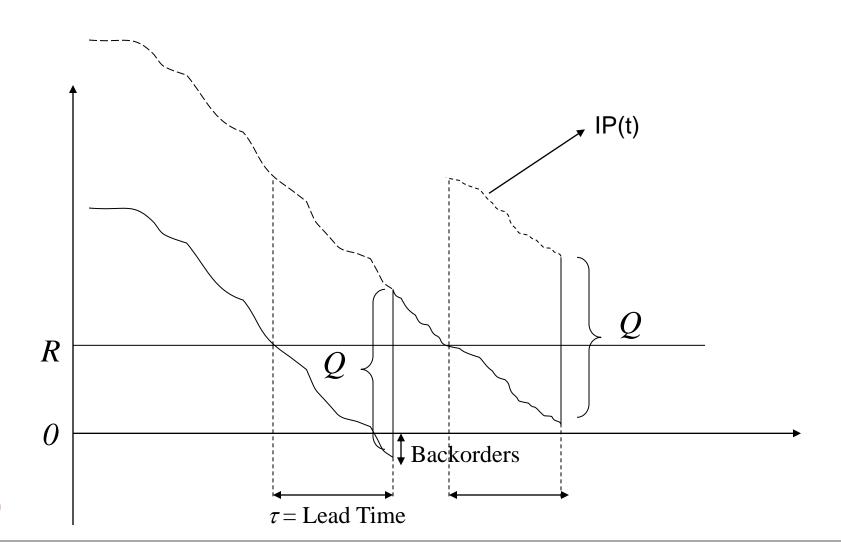
A Stochastic Continuous Review Model (R, Q) Policy

- In the basic EOQ model, the lead time τ does not have a significance as long as an order is placed τ units of time before the cycle ends (inventory drops to zero)
- When the demand is stochastic, replenishment lead time becomes very important because the realized demand during the lead time and the amount of inventory we had at the time of order placement determine the likelihood of shortages

- We will assume that the demand D during lead time is a random variable and has a given pdf denoted by f(x) and a cdf denoted by F(x)
- (*R*, *Q*) is a continuous review inventory control policy suitable for stochastic demand environments
- When the level of inventory on hand drops to R units, then an order of Q units is placed and this order arrives after τ units of time
- R is called the reorder level, and Q is called the order quantity.



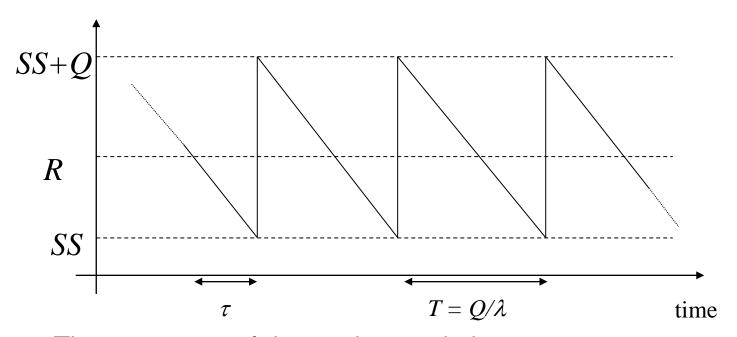
- Inventory level, I(t), alone is not sufficient to properly operate the policy
 - Inventory level I(t) is below R until (and possibly after) the replenishment order arrived
 - Demand may necessitate other replenishment orders to be placed before a specific replenishment order is received
- The way around this problem is to define a new quantity called the Inventory Position which is the inventory on hand (IOH) + the orders already placed but not yet arrived (OO) backorders (BO).
- IP(t) = IOH(t) + OO(t) BO(t)
- Then, an order is placed when the IP hits the reorder level, R.



(R, Q) Policy Assumptions

- R and Q are the decision variables of the system
- The demand during lead time is uncertain, but we know the probability distribution of demand
- If a stock-out occurs before the order is received, the excess demand is not satisfied and are filled once the order arrives
- Cost components are the fixed ordering cost *K*, unit variable ordering cost *c*, unit holding cost *h* per unit time, and unit shortage cost *p* per unit time

(R, Q) Policy Expected Inventory Levels



- λ : The mean rate of demand per unit time
- SS: Safety Stock (the expected inventory level before the replenishment)
- The inventory level varies between SS and SS+Q
- $SS = R \lambda \tau$

(R, Q) Policy Cost Components

- Holding Cost: This cost is estimated from the average inventory curve
- Average inventory carried per unit time

$$SS + Q/2$$

Therefore total holding cost per unit time is

$$h(R - \lambda \tau + Q/2)$$

• Note that this cost is approximate because there may be shortages in a given cycle and a holding cost must not be charged for the quantity backordered

(R, Q) Policy Cost Components

- The expected demand during a cycle time T is $\lambda \cdot T$. On the other hand, every cycle Q units of inventory enter to the system. Therefore, on the average $Q = \lambda \cdot T$ and $T = Q/\lambda$
- Setup cost per unit time is $K / T = \lambda \cdot K / Q$
- Shortage Cost: Note that shortages can occur only during lead times
- Moreover, shortages will occur only if the demand during lead time exceeds *R* units

(R, Q) Policy Cost Components

• Therefore, the expected number of shortages that occur in one cycle

$$n(R) = E\left[\max(D - R, 0)\right] = \int_{R}^{\infty} (x - R)f(x)dx$$

- where D is the demand during lead time and f(x) is its pdf
- The total expected shortage costs per unit time is $p \cdot n(R) / T = p \cdot \lambda \cdot n(R) / Q$

(R, Q) Policy Cost Components

- Proportional ordering cost: Over a long period of time, number of units that enter the system and that leave the system must be same whatever the control policy parameters are
- Any feasible policy will replenish inventory at the rate of demand, on the average
- Therefore, the total ordering costs per unit time must be independent of Q
- We can also see this from the expression

$$c \cdot Q / T = c \cdot \lambda \cdot Q / Q = \lambda c$$

(R, Q) Policy Objective Function

• ETC(R, Q) = Expected total average costs of operation under this policy per unit time

$$ETC(R,Q) = h\left(\frac{Q}{2} + R - \lambda\tau\right) + \frac{\lambda K}{Q} + p\frac{\lambda n(R)}{Q}$$

 The objective is to find R* and Q* values that minimize the function ETC(R, Q)

(R, Q) Policy Optimality Conditions

• By taking the first order derivatives of ETC(R, Q) with respect to R and Q, we can show that R* and Q* values can be found by solving the following set of equations:

$$Q = \sqrt{\frac{2\lambda(K + pn(R))}{h}} \tag{1}$$

$$1 - F(R) = \frac{Qh}{p\lambda} \tag{2}$$

(R, Q) Policy Derivation of the Optimality Conditions

$$ETC(R,Q) = h\left(\frac{Q}{2} + R - \lambda \tau\right) + \frac{\lambda K}{Q} + p\frac{\lambda n(R)}{Q}$$

$$\frac{\partial ETC(R,Q)}{\partial Q} = \frac{h}{2} - \frac{\lambda K}{Q^2} - \frac{p\lambda n(R)}{Q^2} = 0$$

$$\frac{\partial ETC(R,Q)}{\partial R} = h + \frac{p\lambda n'(R)}{Q} = 0$$

$$n(R) = \int_{R}^{\infty} (x - R) f(x) dx \Rightarrow n'(R) = -(1 - F(R))$$

By substituting the expression for n'(R) into the first order conditions, we obtain the optimality conditions

(R, Q) Policy Procedure to find R* and Q*

$$Q = \sqrt{\frac{2\lambda(K + pn(R))}{h}} \quad (1) \qquad 1 - F(R) = \frac{Qh}{p\lambda} \quad (2)$$

- 1. $Q_0 = \sqrt{(2\lambda K/h)}$ (EOQ formula)
- 2. Solve equation (2) for R_0 by using Q_0
- 3. i = 1
- 4. Solve equation (1) for Q_i by using R_{i-1}
- 5. Solve equation (2) for R_i by using Q_i
- 6. If $R_i = R_{i-1}$ and $Q_i = Q_{i-1}$ (or close enough) then let $R^* = R_i$, $Q^* = Q_i$ and stop
- 7. Else let i = i + 1 and go to step 4.

(R, Q) Policy with Normally Distributed Lead-Time Demand

• When the demand during lead time is Normally distributed then n(R) is computed using the standardized loss function

$$L(z) = \int_{z}^{\infty} (t - z)\phi(t)dt$$

- If lead time demand is Normal with mean μ and standard deviation σ then it can be shown that

$$n(R) = \sigma L \left(\frac{R - \mu}{\sigma} \right)$$

Values of L(z) function are available in tabulated forms

Example

- Harvey's Specialty Shop sells a special mustard that is purchased from overseas. The mustard costs to the shop \$10 a jar and requires a six-month lead time for replenishment of stock. Harvey's use a 20% annual interest rate to compute holding costs and estimate that if a customer demands the mustard when it is out of stock, the loss-of-goodwill cost is \$25 a jar. Bookkeeping expenses for placing an order amount to about \$50.
- The demand during lead time has a mean of 100 jars and a standard deviation of 25. How should Harvey control the replenishment of the mustard?

Example

- Under Harvey's inventory policy, determine
 - 1. The safety stock.
 - 2. The average annual holding, setup, and penalty costs.
 - 3. The average time between the placement of orders
 - 4. The proportion of order cycles in which no stockout occurs
 - 5. The proportion of demand that is not met.

- It is generally difficult to estimate *p*
- A suitable substitute of estimating *p* values is using service levels
- Instead of specifying a shortage cost, we set targets related to shortages
- Type I service: In this case, we specify the probability of not stocking out during the lead time
- We use α to denote Type I service levels
- Therefore $\alpha = P\{D \le R\} = F(R)$

- \bullet α can be interpreted as the proportion of cycles that we have a stock-out occurrence
- Whether 1 or 100 items are backordered does not affect Type I service level
- Both are considered as a stock-out phenomenon
- Procedure to find control parameters under Type I service level:
 - 1. Determine R that satisfies $F(R) = \alpha$
 - $2. \quad \text{Set } Q = EOQ$
- This is the optimal policy.

- Type II Service Level: measures the proportion of demand that is met from stocks
- $oldsymbol{\beta}$ is used to denote Type II Service Levels
- ullet eta is also called the fill rate for the inventory system.
- This is the measure that makes sense to the managers.
- A manager would be interested in achieving a certain fill rate, not α .

• Note that n(R) / Q is the average fraction of items that are backordered in any cycle, therefore

$$1 - \beta = n(R) / Q$$

- We can use EOQ for the order quantity and then find the reorder level satisfying the Type II service level. (Optimal?)
- Although determining the best policy satisfying a Type I service measure is easier than the one satisfying a Type II service measure, they should not be used instead of each other.

Example

• Consider again Harvey's Specialty Shop. Harvey's feels uncomfortable with the assumption that the stock-out cost is \$25 and decide to use a service level criterion instead. Suppose that they choose to use a 98 percent service objective.

Optimal (R, Q) Policies subject to Type II Service Level Constraint

- Using EOQ formula for order sizes when we have service level targets is only an approximation
- We can derive a procedure to optimize *Q* and *R* values by satisfying our Type II service level target
- Essentially, we will be solving the following optimization problem:

min.
$$ETC(R,Q)$$

s. to $\frac{n(R)}{Q} = 1 - \beta$

Type II Service Level Constraint

• By using equation (2):

$$1 - F(R) = \frac{Qh}{p\lambda} \qquad \Rightarrow p = \frac{Qh}{\lambda(1 - F(R))}$$

Substitute p in equation (1) and obtain:

$$Q = \sqrt{\frac{2\lambda\{K + Qhn(R)/[\lambda(1 - F(R))]\}}{h}}$$

Type II Service Level Constraint

• It can be shown that the positive root of the above equation is

$$Q = \frac{n(R)}{1 - F(R)} + \sqrt{\frac{2\lambda K}{h}} + \left(\frac{n(R)}{1 - F(R)}\right)^{2}$$
(3)

- Equation (3) is called as SOQ (Service Level Order Quantity) Formula
- SOQ is solved simultaneously with

$$n(R) = (1 - \beta)Q \tag{4}$$

Procedure to find R* and Q* with Type II Service Level Constraint

- 1. $Q_0 = \sqrt{(2\lambda K/h)}$ (EOQ formula)
- 2. Solve equation (4) for R_0 by using Q_0
- 3. i = 1
- 4. Solve equation (3) for Q_i by using R_{i-1}
- 5. Solve equation (4) for R_i by using Q_i
- 6. If $R_i = R_{i-1}$ and $Q_i = Q_{i-1}$ then let $R^* = R_i$, $Q^* = Q_i$ and stop
- 7. Else let i = i + 1 and go to step 4

Type 2 Service Level Constraint

$$Q = \frac{n(R)}{1 - F(R)} + \sqrt{\frac{2\lambda K}{h}} + \left(\frac{n(R)}{1 - F(R)}\right)^{2} \qquad n(R) = (1 - \beta)Q$$

- Find the optimal policy for Harvey's Specialty Store if they want to achieve a 98% Type 2 service level.
 - $Q_0 = 100$, $n(R_0) = 2$, $z_0 = 1.02$, $R_0 = 126$
 - $Q_1 = 114$, $n(R_1) = 2.28$, $L(z_1) = 0.0912$, $z_1 = 0.95$, $R_1 = 124$
 - $Q_2 = 114$, Stop! \Rightarrow (Q,R) = (114, 124)
 - Holding+Setup Cost for Q=100, R=126 is \$252/year
 - Holding+Setup Cost for Q=114, R=124 is \$250/year

Imputed Shortage Cost

- Of course in this problem we do not use any shortage cost, but only specify the desired service level.
- Although no shortage cost is specified, the optimal solution of this problem is the same with the optimal solution of the problem with shortage cost for some shortage cost, p.
- That means there is some value of p such that the policy satisfying Type I and Type II constraints also satisfies equations 1&2

Imputed Shortage Cost

- This value of p is known as the imputed shortage cost.
- Once can easily obtain this value using

$$p = \frac{Qh}{\lambda (1 - F(R))}$$

- The imputed shortage cost is a useful way to determine whether the value chosen for the service level is appropriate.
 - $\alpha = .98 \Rightarrow (Q,R) = (100, 151) \Rightarrow p = 50
 - $\beta = .98 \Rightarrow (Q,R) = (114, 124) \Rightarrow p = 6.67

Practical Issues

Forecasting Input

- Mean demand per period
- Forecast error
- Use of MAD
- Normality assumption Fast Moving Items
- In real life you may not be able to obtain the lead time demand directly
- You will obtain demand forecast on a periodic basis, such as monthly.
- Then you need to convert the demand distribution to correspond to the lead-time

Scaling of Lead-Time Demand

- This is easily done, when the demand is normally distributed because sums of independent normal random variables are also normally distributed.
- Thus, the lead-time demand is also a normal random variable.
- Then, all you need is to determine the mean and the standard deviation of lead-time demand
- Let a period's demand have mean λ and standard deviation ν , and let τ be the lead-time in terms of periods.
- Then for the demand during lead-time, the mean is $\lambda \tau$ and the variance is $V^2 \tau$.

$$\mu = \lambda \tau, \ \sigma = \upsilon \sqrt{\tau}, \ c.o.v. = \frac{\upsilon}{\lambda \sqrt{\tau}}$$

• Hence the coefficient of variation of lead-time demand decreases as the lead-time increases

Stochastic Lead Time

• Let lead-time be a random variable with

$$E[\tau] = \mu_{\tau}, \ Var(\tau) = \sigma_{\tau}^{2}$$

So lead-time demand distribution

$$\mu = \lambda \mu_{\tau}$$
, and $\sigma^2 = \mu_{\tau} \upsilon^2 + \lambda^2 \sigma_{\tau}^2$

where μ = expected lead – time demand,

$$\sigma^2$$
 = variance of lead – time demand.

• Normality assumption will work here, as well.

Multi-item systems: ABC Analysis

- ✓ ABC analysis is based on the Pareto Curve. Pareto discovered that the distribution of wealth follows an increasing exponential curve. A similar curve describes the distribution of the value of inventory items in a multi-item system. (See Figure 5-7).
- ✓ The value of a Pareto curve analysis in this context is that one can identify the items accounting for most of the dollar volume of sales. Rough guidelines:
 - ✓ the first 20% of the items account for 80% of the sales, A items
 - ✓ the next 30% of the items account for 15% of the sales, B items, and
 - ✓ the last 50% of the items only account for 5% of the sales, C items.
- ✓ To determine the attention to be given to items

Pareto Curve: Distribution of Inventory by Value

