We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Single-Period Capacity and Demand Allocation Decision Making under Uncertainty

Sangdo Choi

Abstract

The newsvendor model deals with a single-period capacity allocation problem under uncertainty. The real world examples include perishable products (e.g., fish, vegetable), holiday-related products (e.g., Easter, Christmas, Halloween), seasonal products (e.g., fashion), and promotional products. This section addresses three newsvendor models: traditional newsvendor, inverse newsvendor, and sequential newsvendor models. The main decision under the traditional newsvendor setting is capacity allocation (i.e., how much to order), whereas the main decision under the inverse newsvendor setting is demand allocation (i.e., how many customers to be served) under the fixed capacity. This section demonstrates how to compare profit maximization approach to customer-oriented approach under the traditional newsvendor. The inverse newsvendor applies to revenue management for the hospitality industry. The sequential newsvendor model determines the optimal sequence when the number of customers to be served (determined by the inverse newsvendor model) is given. Normal distribution is considered for analytical solution and numerical studies. In addition, a discrete distribution is considered for numerical studies.

Keywords: Capacity Allocation, Demand Allocation, Newsvendor, Inverse Newsvendor, Sequential Newsvendor

1. Introduction

How can an operations manager make a one-time decision that covers a fixed future period if the manager cannot adjust the decision afterwards? A typical approach to this question is the single-period newsvendor model [1–4]. Suppose the operations manager herein is a newsvendor who would like to maximize profit or to satisfy a probability of not running out of newspaper. The newsvendor must place and receive an order before the start of each day to put on the newspaper stand. All left-over newspapers will be salvaged through paper-collection companies after the day, because nobody is interested in out-dated newspapers. The newsvendor is supposed to know all demand history and is able to forecast demand distribution properly, but not exact demand quantity on certain date. The newsvendor will not be able to match supply with demand exactly, unless the newsvendor is lucky [1], because only one demand scenario is realized during the selling period, i.e., $P(demand\ quantity = order\ quantity) = 0$ for a continuous demand distribution. Similar examples include grocery products (e.g., fish, vegetable), holiday-related

products (e.g., Christmas, Easter, Halloween), seasonal products (e.g., fashion), and promotional products (e.g., T-shirts for a championship basketball or football game) [3]. These products also have a single selling period and will be deeply discounted after the selling season. If the newsvendor orders too much, left-over (overage) inventory is salvaged or steeply discounted. Otherwise, the newsvendor will forgo net profit because of lack of inventory (underage). Erlebacher et al. [5, 6] address the multi-item newsvendor model for inventory optimization problem with a capacity constraint. The newsvendor would like to keep balance between overage and underage, depending on the importance of two opposing directions.

The main decision variable for the traditional newsvendor is how many orders to be placed, which is a capacity allocation problem. Inversely, the newsvendor can also make decision on demand size to take full advantage of capacity [7–11]. If the newsvendor allocates too many demand, the resource is over-utilized (overusage). Otherwise, the resource is under-utilized (underusage). The inverse newsvendor would like to keep balance between overusage and underusage. A sequential newsvendor can make sequencing decision, when the demand size is determined by the inverse newsvendor [12]. Each customer is assigned to a slot in a sequence and the expected service start time should be scheduled. The sequential newsvendor would like to keep balance between earliness and lateness.

This chapter is organized as follows. Section 2 explains the traditional newsvendor model. Section 3 addresses the inverse newsvendor model. Section 4 addresses the sequential newsvendor model. Section 5 concludes this chapter.

2. The traditional newsvendor model

The traditional newsvendor is supposed to deal with inventory control. The newsvendor has to decide the order quantity to maximize the expected profit. If the newsvendor orders one less than the desired quantity, the newsvendor will forgo unit net profit owing to the lost sales. If the newsvendor order one more than the desired quantity, the newsvendor will loose unit net loss owing to the left-over inventory. Through the marginal analysis, the critical fractile determines the desired (or optimal) quantity and is regarded as the customer service level (CSL) [1–4].

On the contrary, the newsvendor may be interested in improving CSL than maximizing internal profit. For example, the newsvendor might want to make 90 percent sure of not running out of inventory, even though the critical fractile to maximize the expected profit is 0.7. The newsvendor would expand the market size in the long run while sacrificing the short-term maximum profit.

2.1 Mathematical model and solution approach

Let p be price; c order cost; s salvage, respectively. Demand D has mean of μ and standard deviation σ. Our decision variable is order quantity, q. The objective function is to maximize the expected profit. The profit function $\pi(q)$ is defined as follows:

$$\max_q \pi(q) = pE[\min(q, D)] - cq + sE[(D - q)^+], \tag{1}$$

where $\min(q, D)$ is the realized sales out of demand and $(D - q)^+$ is left-over inventory, respectively. The profit maximization problem $\pi(q)$ reduces to the equivalent problem $\tilde{\pi}(q)$ to minimize the expected sum of underage and overage as follows:

$$\min_{q} \tilde{\pi}(q) = c^o E\left[(q - D)^+\right] + c^u E\left[(D - q)^+\right], \qquad (2)$$

where $c^u = p - c$, net profit and $c^o = c - s$, net loss, respectively. The optimal solution to either $\pi(q)$ or $\tilde{\pi}(q), q^*$ can be obtained by first and second necessary conditions or through marginal analysis as follows [1, 2, 4]:

$$F(q^*) = \frac{c^u}{c^u + c^o}, \qquad (3)$$

where $F(\cdot)$ is the cumulative distribution function of demand D. In addition, $F(q^*)$ is the probability that you are able to cover all demand up to q^*, CSL for order quantity, q^*.

2.2 Numerical example of discrete demand

The newsvendor is supposed to sell Christmas trees between Halloween and Christmas Eve, this year. Suppose that the newsvendor has such a long sales history to build a reasonable demand forecast. **Table 1** shows the demand forecast based upon the historical data.

The newsvendor has to place and receive an order before Halloween, which is supposed to be the first day of selling season. The newsvendor sets the selling price to $25 per unit and promises to pay $10 per unit to a farmer. A local mulch firm will collect left-over trees for $3 per unit to cut them into small pieces for mulch after Christmas. Note that the underage penalty $c^u = 25 - 10 = 15$ per unit and the overage penalty $c^o = 10 - 3 = 7$ per unit. The newsvendor tends to order more than the average 260, which is close to median, because $c^u > c^o$, i.e., the newsvendor wants to avoid underage rather than overage. The critical fractile is $\frac{15}{15+7} = 0.68$. The optimal order quantity should be 300 because of $F(250) < 0.68 < F(300)$. However, if the newsvendor sets CSL to 90 percent, the order quantity should be 350 because of $F(350) > 0.9$, of which profit is lower than the profit of the optimal order quantity 300.

Table 2 provides the expected profit of three order quantities: 250, 300, and 350. Order quantity of 300 is (at least) a local optimum. Note that the profit function $\pi(q)$ is convex function, i.e., increasing-then-deceasing [1, 2, 4]. If the newsvendor would compute the expected profit for all other order quantities, the newsvendor can recognize that order quantity of 300 is global optimal. If the newsvendor orders too much (e.g., 350), salvages are larger than the optimal

Demand quantity	Probability	Cumulative probability
100	0.03	0.03
150	0.07	0.10
200	0.10	0.20
250	0.25	0.45
300	0.30	0.75
350	0.20	0.95
400	0.05	1.00

Table 1.
Demand forecast with probability and cumulative probability for Christmas tree.

Demand	Probability	Revenue	Cost	Salvage	Profit
Order = 250					Exp. Profit = 3,387
100	0.10	100 × 25	250 × 10	150 × 3	450
150	0.10	150 × 25	250 × 10	100 × 3	1,550
200	0.15	200 × 25	250 × 10	50 × 3	2,650
250	0.15	250 × 25	250 × 10	0 × 3	3,750
300	0.25	250 × 25	250 × 10	0 × 3	3,750
350	0.15	250 × 25	250 × 10	0 × 3	3,750
400	0.10	250 × 25	250 × 10	0 × 3	3,750
Order = 300					Exp. Profit = 3,642
100	0.10	100 × 25	300 × 10	200 × 3	100
150	0.10	150 × 25	300 × 10	150 × 3	1,200
200	0.15	200 × 25	300 × 10	100 × 3	2,300
250	0.15	250 × 25	300 × 10	50 × 3	3,400
300	0.25	300 × 25	300 × 10	0 × 3	4,500
350	0.15	300 × 25	300 × 10	0 × 3	4,500
400	0.10	300 × 25	300 × 10	0 × 3	4,500
Order = 350					Exp. Profit = 3,567
100	0.10	100 × 25	350 × 10	250 × 3	−250
150	0.10	150 × 25	350 × 10	200 × 3	850
200	0.15	200 × 25	350 × 10	150 × 3	1,950
250	0.15	250 × 25	350 × 10	100 × 3	3,050
300	0.25	300 × 25	350 × 10	50 × 3	4,150
350	0.15	350 × 25	350 × 10	0 × 3	5,250
400	0.10	350 × 25	350 × 10	0 × 3	5,250

Table 2.
Expected profit for three order quantities: 250, 300, and 350.

quantity and revenues are also larger than the optimal quantity. However, larger ordering cost affects more on the expected profit. The expected profit of 350 is lower than the maximum. If the newsvendor orders too little (e.g., 250), the newsvendor can save salvages compared to the optimal quantity and revenue is not large.

2.3 Numerical example of normally distributed demand

Now take into account a continuous demand distribution. Suppose that the demand distribution is normally distributed with mean of 275 and standard deviation of 50. It is hard to compute the revenue and salvage for each order, because there are infinite scenarios of order quantity. The newsvendor can compute the expected profit, starting from the expected lost sales, which is expressed as follows:

$$E\left[(D - q)^+\right] = \sigma L(z), \tag{4}$$

Step	Item	Profit-based	CSL-oriented
1	Critical Fractile (or CSL)	0.68	0.9
2	Order Quantity	298.6	339.1
3	Expected Lost Sales	7.4	5.0
4	Expected Sales	267.6	270.0
5	Expected Left-over	31.0	69.1
6	Expected Profit	3,797.5	3,566.0

Table 3.
Expected profit for two approaches: Profit maximization vs. CSL-oriented.

where $L(z) = \phi(z) - z(1 - \Phi(z))$ [2]. Note that $\phi(z)$ is normal probability distribution and $\Phi(z)$ is cumulative distribution, respectively.

1. Compute the critical fractile, or CSL.

2. Compute the associated quantity with CSL, $norm.inv(CSL, \mu, \sigma)$.

3. Compute the expected lost sales, $\sigma L(z)$.

4. Compute the expected sales: = expected demand - expected lost sales.

5. Compute the expected left-over: = order quantity - expected sales.

6. Compute the expected profit: = $c^u \times$ expected sales - $c^o \times$ expected left-over.

The newsvendor can take two perspectives: internal profit maximization vs. higher CSL. **Table 3** shows computational steps to get the expected profits of both profit-based and CSL-oriented approaches, respectively. For profit-based approach, the critical fractile is computed and its associated order quantity is determined accordingly. The expected profit is $3,797.5. For CSL-oriented approach, the newsvendor is supposed to determine the desired CSL first. Suppose that the newsvendor would like to guarantee 90% probability of not running out, i.e., 90% of demand will be covered by the order quantity. Because of higher CSL, the order quantity is far larger than the optimal order quantity; lower expected lost sales; larger left-over. Henceforth, the expected profit is lower. The newsvendor can choose either order quantity based on your strategic direction.

3. The inverse newsvendor model

The inverse newsvendor model applies to revenue management, which deals with fixed capacity and has to determine demand allocation [7, 8, 10, 11]. Airline industry uses quantity (i.e., number of seats) for capacity, whereas hospital may use time unit for capacity. Time-based inverse newsvendor model can be addressed for time-sensitive service industries such as hospital and law-firm.

The inverse newsvendor can take into account both identical and non-identical service durations. When the inverse newsvendor takes into account all identical service durations, the decision reduces to the number of allocation, i.e., how many customers will be assigned. When the inverse newsvendor takes into account

heterogeneous service durations, the decision reduces to setting priority problem. Who should be allocated first and who can be next on? [9].

3.1 Mathematical model for identical service durations

Let h be the given and fixed capacity in hour. Each customer requires service duration, T which follows normal distribution with mean of μ and standard deviation σ. Assume that all customers are homogeneous, i.e., they have the same mean and standard deviation. The inverse newsvendor has to decide the number of customers to be served, x to minimize the sum of expected overusage and underusage. Consider the unit overusage penalty, c^g and unit underusage penalty, c^ℓ. The objective function $\rho(x)$ is defined as follows:

$$\min_x \rho(x) = c^g E\left[\left(\sum_{k=1}^{x} T_k - h\right)^+\right] + c^\ell E\left[\left(h - \sum_{k=1}^{x} T_k\right)^+\right]. \tag{5}$$

$\sum_{k=1}^{x} T_k$ also follows normal distribution with mean of $x\mu$ and variance of $x\sigma^2$. Let $z = \frac{h-x\mu}{\sqrt{x}\sigma}$. Overusage $E\left[\left(\sum_{k=1}^{x} T_k - h\right)^+\right]$ and underusage $E\left[\left(h - \sum_{k=1}^{x} T_k\right)^+\right]$ are defined as follows [12]:

$$E\left[\left(\sum_{k=1}^{x} T_k - h\right)^+\right] = (\phi(z) - z(1 - \Phi(z)))\,\sigma, \tag{6}$$

$$E\left[\left(h - \sum_{k=1}^{x} T_k\right)^+\right] = (\phi(z) + z\Phi(z))\,\sigma. \tag{7}$$

Figure 1 depicts a graphical representation of an inverse newsvendor problem with $\mu = 2$, $\sigma = 0.8$, and $h = 9$. The optimal solution to (5), x^* is defined as follows [9]:

$$x^* = \lfloor \hat{x} \rfloor \text{ or } \lceil \hat{x} \rceil, \tag{8}$$

$$\text{where } \hat{x} = \left(\frac{-z\sigma + \sqrt{z^2\sigma^2 + 4\mu h}}{2\mu}\right)^2. \tag{9}$$

$$F_N\left(h; x^*, x^{*2}\right) = \frac{c^g}{c^\ell + c^g} = \Phi(z), \tag{10}$$

For the case of **Figure 1**, the optimal allocation can be either 4 or 5 by visualization and analytical solution, (8) and (9).

3.2 Numerical example for identical service durations

Consider an operating room (OR) with 8 or 9 hour capacity. When each patient requires 2 hour service durations on average, how many patients would be assigned in the OR daily? Overusage penalties would cover overtime pay to the attending surgeon(s), nurses, anesthesiologist, and other staff. Underusage penalties would cover opportunity cost when the OR is under-utilized, but be hard to measure. The inverse newsvendor can determine the optimal demand size if the newsvendor knows parameters of service duration and two penalties. Numerical studies show
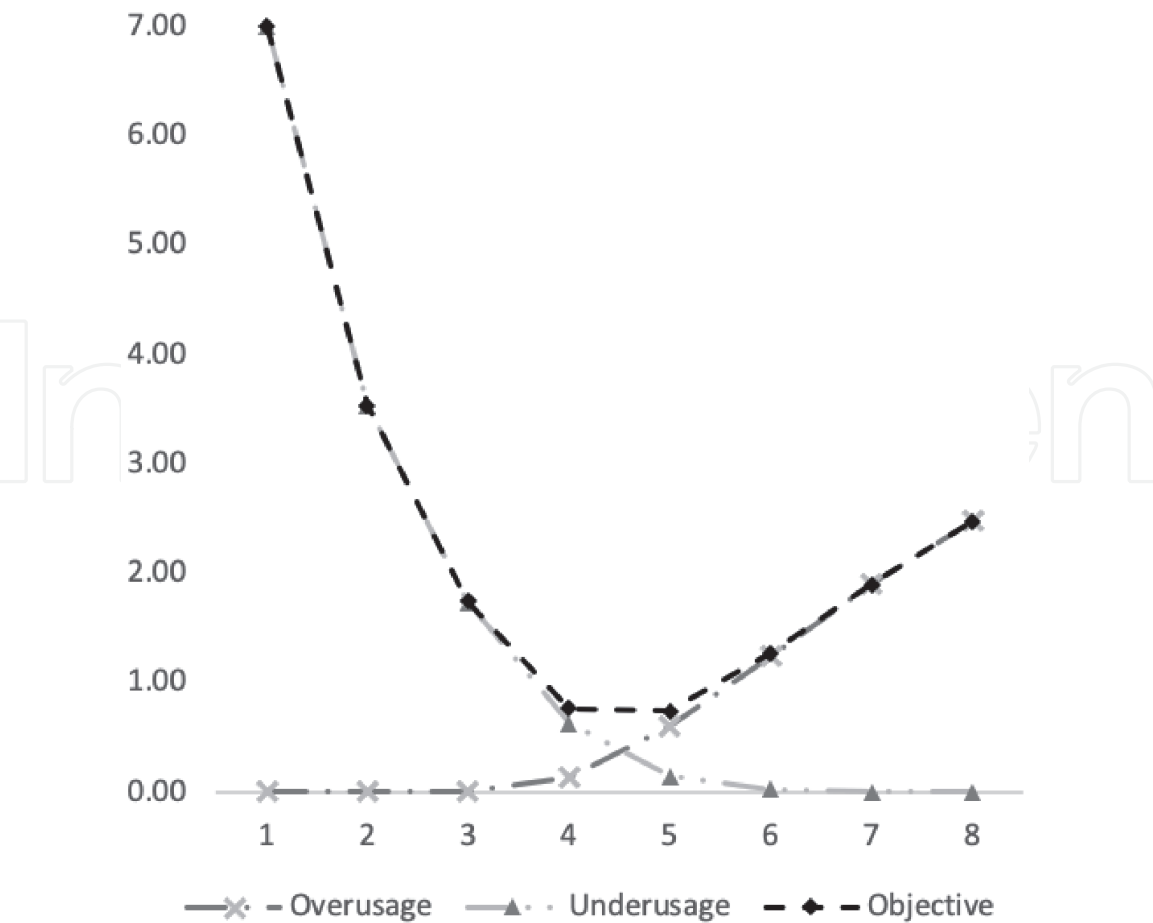
Figure 1.
Graphical representation of an inverse newsvendor model. The objective function for the case of $\mu = 2, \sigma = 0.8$, and $h = 9$. The overusage is an ever-increasing function of x, whereas the underusage is an ever-decreasing function of x. Hence, the objective function is a decrease-then-increase function.

Scenario	1	2	3	4	5	6	7	8
μ	2	2	2	2	3	3	3	3
σ	0.2	0.2	0.8	0.8	0.3	0.3	1.2	1.2
h	8	8	8	8	9	9	9	9
$c^g : c^\ell$	0.1:0.9	0.9:0.1	0.1:0.9	0.9:0.1	0.1:0.9	0.9:0.1	0.1:0.9	0.9:0.1
x	4.26	3.75	5.17	3.10	3.23	2.79	4.03	2.23
$\rho(\lfloor x \rfloor)$	0.16	0.20	0.32	0.25	0.21	0.30	0.42	0.33
$\rho(\lceil x \rceil)$	0.20	0.16	0.41	0.64	0.30	0.21	0.61	0.83
x^*	4	4	5	3	3	3	4	2

Table 4.
Patient allocation under different cost ratios, service durations and capacity.

impact of cost ratio and parameters on patient allocation. Consider the following data set in **Table 4**.

Table 4 summarizes numerical studies with varying mean, standard deviation, cost ratios, and capacity values. The inverse newsvendor can take into account two capacity levels: 8 or 9. Hospital may operate 8 hours each day or 9 hours if the inverse newsvendor expects high possibility of overtime. Each surgery duration requires 2 or 3 hours. Take into account two levels of standard deviation for each service duration. Two extremely different cost ratios are considered.

The ratio of capacity to the mean service duration, $\frac{h}{\mu}$ can be a base scenario. Actual allocation can be the base scenario, one more allocation, or one less allocation from the base scenario. For example, scenarios 1–4 have the ratio of 4 and scenarios 5–8 have the ratio of 3. When the inverse newsvendor has a non-integer value of ratio, the newsvendor can use either floor or ceiling value of the ratio. Actual allocations are 3, 4, or 5 for scenarios 1–4; 2, 3, or 4 for scenarios 5–8, respectively.

When $c^g > c^l$ (i.e., overusage is more penalized than underusage), the inverse newsvendor tends to allocate less patients (than the base) to avoid overusage penalty. On the contrary, when $c^l > c^g$ (i.e., underusage is more penalized than overusage), the inverse newsvendor tends to allocate more patients (than the base) to avoid underusage penalty. Allocating one more patient or one less patient would affect a lot on the objective function. As a matter of fact, allocating more (less) patients means ONE more (less) patient than the base scenario.

Variance may amplify impact of cost-ratio, which means there must exist interactive effect between variance and cost-ratio. When $c^g < c^l$ (e.g., scenarios 1, 3, 5 and 7), the larger variance, the more allocated patients. When $c^g > c^l$ (e.g., scenarios 2, 4, 6, and 8), the larger variance, the less allocated patients. For lower variance examples (scenarios 1, 2, 5, and 6), cost-ratio would not affect on allocation much.

3.3 Mathematical model for non-identical service durations

Suppose that there are N customers, of which index is $i = 1, 2, \cdots, N \in \mathcal{I}$, respectively and that individual service time T_i of customer i has mean of μ_i and standard deviation σ_i. New decision variable x_i is a binary variable, 1 if customer i is served, 0 otherwise. The number of customers to be served is $\sum_i x_i$. The total service time is defined as $\sum_i x_i T_i$. The inverse newsvendor problem with non-identical service durations can be represented as follows:

$$\min_{x_i, i \in \mathcal{I}} c^g E\left[\left(\sum_{i=1}^N x_i T_i - h\right)^+\right] + c^\ell E\left[\left(h - \sum_{i=1}^N x_i T_i\right)^+\right]. \tag{11}$$

The inverse newsvendor should evaluate $2^N - 1$ possible combinations to find the optimal number of customers to be served. To find the optimal solution based on numerical evaluation of (11), the inverse newsvendor can reformulate it using Stochastic Programming with discrete scenarios $\omega \in \Omega$. The inverse newsvendor can adopt the sample average approximation (SAA) approach to get a close approximation [13]. Let T_i^ω be the service time for customer i under scenario ω; u^ω underusage under scenario ω; o^ω overusage under scenario ω; p^ω probability of scenario ω, respectively. SAA formulation is given as follows:

$$\min \qquad c^l u^\omega p^\omega + c^g o^\omega p^\omega \tag{12}$$

$$s.t. \quad \sum_{i=1}^N x_i T_i^\omega + u^\omega \qquad \geq h, \omega \in \Omega \tag{13}$$

$$\sum_{i=1}^N x_i T_i^\omega \qquad - o^\omega \leq h, \omega \in \Omega \tag{14}$$

The inverse newsvendor can get the optimal solution of the SAA approach [9]. However, it is hard to derive a certain (intuitive) rule for the optimal allocation of customers. A heuristic to get a near-optimal solution in a reasonable time limit is

prescribed: smallest-variance (SV) first, which is close to the optimal solution [9]. The heuristic is based on the discussion that partial expected values are associated with variability rather than central location measure such as mean or median.

Take advantage of the results from the case of identical service durations from Subsection 3.1. Suppose that n customers are about to be served. Let $\bar{\mu}$ be the sample average service time for n customers; $\bar{\sigma}$ the standard deviation of the sample average service times for n customers, respectively. If n is equal to the solution of (8) and (9) with $\bar{\mu}$ and $\bar{\sigma}$, the inverse newsvendor can stop adding customers to be served. The detail procedure of the heuristic with the SV selection rule is described as follows [9]:

- **Initialization.** Let $\mathcal{A} = \mathcal{A}^* = \{\}, \mathcal{N} = \{1, 2, \cdots, N\}$, and $Z_{opt} = \infty$.

- **Step 1.** Select i with the smallest variance. Remove i from $\mathcal{N}$ and add i to $\mathcal{A}$.

- **Step 2.** Compute the sample mean $\bar{\mu}$ and sample standard deviation $\bar{\sigma}$ of the set $\mathcal{A}$.

- **Step 3.** Plug $\bar{\mu}$ and $\bar{\sigma}$ into (8) and (9) to compute the optimal number of customers to be served, say x^*.

- **Step 4.** Compute the objective function value, say Z_{curr}. If $Z_{curr} < Z_{opt}$, let $\mathcal{A}^* \leftarrow \mathcal{A}$ and $Z_{opt} \leftarrow Z_{curr}$.

- **Step 5.** If $x^* \leq |\mathcal{A}^*|$ and $\mathcal{N} \neq \{\}$, go to Step 1. Otherwise, go to Step 6.

- **Step 6.** Let $\mathcal{A}^*$ be the set of optimally assigned customers and Z_{opt} be the heuristic results.

The inverse newsvendor can show how the SV heuristic works with the following example. The inverse newsvendor can use cost ratio of 0.5:0.5; 120 min blocks without loss of generality. **Table 5** shows all parameter values of ten customers: μ and σ. The SV heuristic will select customers as the following order: $10 \rightarrow 7 \rightarrow 9 \rightarrow 6 \rightarrow 8 \rightarrow \cdots \rightarrow 2$.

The followings are detail steps resulted from the SV selection rule.

- Initial Step. $\mathcal{A} = \mathcal{A}^* = \{\}; \mathcal{N} = \{1, 2, 3, \cdots, 10\}; Z_{opt} = \infty$

- Iteration 1. $\mathcal{A} = \{10\}; \bar{\mu} = 10.3; \bar{\sigma} = 1.80; x^* = 12 > |\mathcal{A}^*| = 1; Z_{curr} = 54.87; Z_{opt} = 54.87; \mathcal{A}^* = \{10\}$

- Iteration 2. $\mathcal{A} = \{10, 7\}; \bar{\mu} = 18.79; \bar{\sigma} = 3.07; x^* = 6 > |\mathcal{A}^*| = 2; Z_{curr} = 41.21; Z_{opt} = 41.21; \mathcal{A}^* = \{10, 7\}$

- Iteration 3. $\mathcal{A} = \{10, 7, 9\}; \bar{\mu} = 16.38; \bar{\sigma} = 4.09; x^* = 7 > |\mathcal{A}^*| = 3; Z_{curr} = 35.43; Z_{opt} = 35.53; \mathcal{A}^* = \{10, 7, 9\}$

Customer	1	2	3	4	5	6	7	8	9	10
μ_i	18.5	27.9	24.9	28.5	26.8	27.5	27.3	19.8	11.6	10.3
σ_i	6.5	21.7	17.3	7.5	18.8	5.7	3.9	6.4	5.6	1.8

Table 5.
Customer service duration information.

- Iteration 4. $\mathcal{A} = \{10, 7, 9, 6\}; \bar{\mu} = 19.16; \bar{\sigma} = 4.55; x^* = 6 > |\mathcal{A}^*| = 4; Z_{curr} = 21.67; Z_{opt} = 21.67; \mathcal{A}^* = \{10, 7, 9, 6\}$

- Iteration 5. $\mathcal{A} = \{10, 7, 9, 6, 8\}; \bar{\mu} = 19.3; \bar{\sigma} = 4.97; x^* = 6 > |\mathcal{A}^*| = 5; Z_{curr} = 11.82; Z_{opt} = 11.82; \mathcal{A}^* = \{10, 7, 9, 6, 8\}$

- Iteration 6. $\mathcal{A} = \{10, 7, 9, 6, 8, 1\}; \bar{\mu} = 19.16; \bar{\sigma} = 5.27; x^* = 6 = |\mathcal{A}^*|; Z_{curr} = 5.53; Z_{opt} = 5.53; \mathcal{A}^* = \{10, 7, 9, 6, 8, 1\}$

- Iteration 7. $\mathcal{A} = \{10, 7, 9, 6, 8, 1, 4\}; \bar{\mu} = 20.51; \bar{\sigma} = 5.64; x^* = 6 < |\mathcal{A}^*| = 7; Z_{curr} = 12.15; Z_{opt} = 5.53; \mathcal{A}^* = \{10, 7, 9, 6, 8, 1\}$. Stop.

4. The sequential newsvendor model

A sequential newsvendor has to determine the sequence of assigned customers and their arrival times, when the newsvendor already knows the total number of customers to be served in a fixed duration [12]. Once the sequential newsvendor determines the sequence, the arrival time of each customer can be the cumulative expected service time of all prior customers without loss of generality. Basically, this is a block scheduling problem that determines the starting times of blocks.

The sequential newsvendor model applies to time-sensitive service industries as the inverse newsvendor. The inverse newsvendor may decide the strategic level decision, whereas the sequential newsvendor decides the tactical level decision and relies on the inverse newsvendor decision. The newsvendor may use identical service durations for the strategic decision (e.g., capacity size, demand size) and non-identical durations for the tactical decision (e.g., setting priority).

4.1 Mathematical model and solution approach

Suppose that the sequential newsvendor has to serve $|I|$ customers (or customer groups) and that each customer $i \in I$ requires different service duration T_i, of which mean is μ_i and its standard deviation is σ_i. The sequential newsvendor has to determine its sequence and starting time of each patient i.

Use map $\Delta : I \to K$ to represent a set of sequences (or permutations), each of which $\delta \in \Delta$ assigns each customer to one and only one sequence position, hence $|K| = |I|$. Use subscripts $[k]$ for k^{th} block sequence position and i for customer to avoid potential confusion.

Decision variables must prescribe planned block durations and block sequence, δ. The sequential newsvendor determines the planned end time of block in the kth position, given a sequence δ, prescribed by $y_{[k]}^\delta$. The planned end time of the block corresponds to the end of service durations and is important in deciding the number of hours that the server will be required to work. Define $B_{[k]}^\delta := T_{[1]}^\delta + \cdots + T_{[k]}^\delta$ as the random end time to complete all services assigned to blocks $[1]$ through $[k]$ and compare it with the decision variable $y_{[k]}^\delta$.

Assume that one service begins as soon as the previous one ends [14–16]. This assumption appears to be reasonable because each customer can be prepared well in advance of his/her scheduled start time and successive services within each block is likely to be performed by the same server so that s/he would be available as well. However, expediting efforts is required if a planned service ends earlier than the planned end time. The sequential newsvendor can penalize the earliness.

The objective function penalizes the expected earliness $E\left[\left(y_{[k]}^{\delta} - T_{[k]}^{\delta}\right)^{+}\right]$ and expected lateness $E\left[\left(T_{[k]}^{\delta} - y_{[k]}^{\delta}\right)^{+}\right]$ of each block $k \in K$. The sequential newsvendor imposes earliness penalty c^{e} and lateness penalty c^{l}, respectively. The former represents the cost of expediting the start time of the next surgery; and the letter, the cost of delaying the start time of the next surgery. The sequential newsvendor can build a schedule that balances the expected costs of earliness and lateness associated with each block, defining objective function $f_{[k]}\left(y_{[k]}^{\delta}\right), k \in K, \delta \in \Delta$

$$ f_{[k]}\left(y_{[k]}^{\delta}\right) = c^{e} E\left[\left(y_{[k]}^{\delta} - T_{[k]}^{\delta}\right)^{+}\right] + c^{l} E\left[\left(T_{[k]}^{\delta} - y_{[k]}^{\delta}\right)^{+}\right]. \tag{15}$$

The sequential newsvendor has to determine the optimal planned end time $\hat{y}_{[k]}^{\delta}$ of the kth block, $k \in K$ and the optimal block sequence $\hat{\delta}$. **Figure 2** depicts a graphical representation of the sequential newsvendor model. For each sequence $\delta \in \Delta$, the objective function $\sum_{k \in K} f_{[k]}\left(y_{[k]}^{\delta}\right)$ should be minimized. The sequential newsvendor has to find the best solution out of all minimized solutions. The sequential newsvendor problem can be defined as follows:

$$ \min_{\delta \in \Delta} \min_{y_{[k]}^{\delta} : k \in K} \sum_{k \in K} f_{[k]}\left(y_{[k]}^{\delta}\right) \tag{16}$$

$$ s.t. \quad y_{[k-1]}^{\delta} \leq y_{[k]}^{\delta} \quad k = 2, \cdots, |K|, \delta \in \Delta \tag{17}$$

$$ y_{[k]} \geq 0 \quad k \in K, \delta \in \Delta \tag{18}$$

Fix a sequence δ to find the optimal planned end time. Suppress this superscript for the sake of simplicity. (16) is separable with respect to $y_{[k]}$ [12]:

$$ \min \left\{ f_{[1]}\left(y_{[1]}\right) + f_{[2]}\left(y_{[2]}\right) + \cdots + f_{[|K|]}\left(y_{[|K|]}\right) \right\} \tag{19}$$
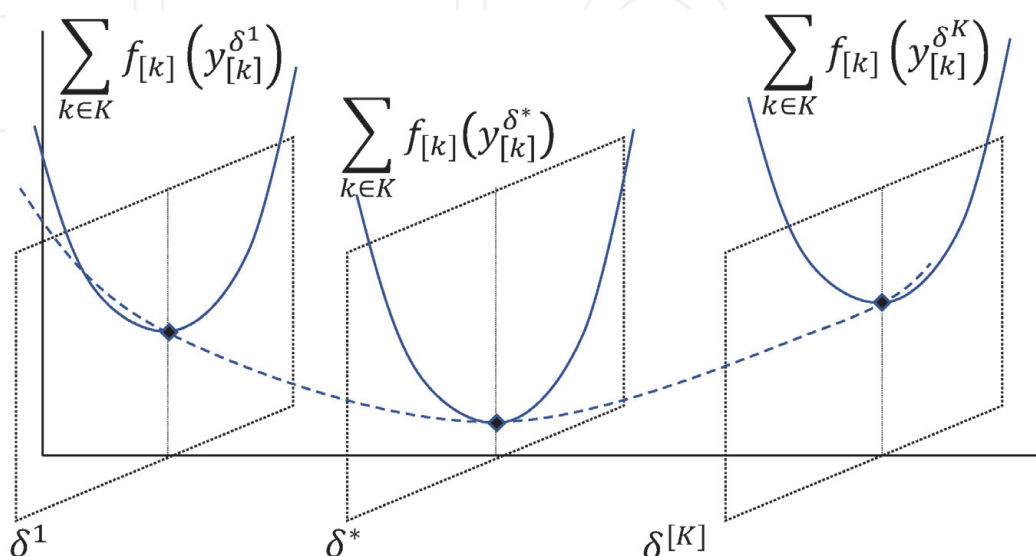

Figure 2.
Graphical representation of a sequential newsvendor model. For each sequence $\delta \in \Delta$, the objective function should be minimized.

$$\equiv \min f_{[1]}\left(y_{[1]}\right) + \min f_{[2]}\left(y_{[2]}\right) + \cdots + \min f_{[|K|]}\left(y_{[|K|]}\right). \tag{20}$$

Let $\overline{\mu}_{[k]} = \mu_{[1]} + \mu_{[2]} + \cdots + \mu_{[k]}$ be the mean of the random end time $B_{[k]}$ and $\overline{\sigma}_{[k]} = \sqrt{\sigma_{[1]}^2 + \sigma_{[2]}^2 + \cdots + \sigma_{[k]}^2}$ be the standard deviation of $B_{[k]}$, respectively. Random end time $B_{[k]}, k \in K$ is also normally distributed as follows:

$$B_{[k]} = T_{[1]} + T_{[2]} + \cdots + T_{[k]} \sim N\left(\overline{\mu}_{[k]}, \overline{\sigma}_{[k]}^2\right) \tag{21}$$

Henceforth, the optimal planned end time $\hat{y}_{[k]}$ of kth block can be obtained as follows [17]:

$$\hat{y}_{[k]} = \overline{\mu}_{[k]} + z\overline{\sigma}_{[k]} \tag{22}$$

$$\Phi(z) = \frac{c^l}{c^e + c^l}. \tag{23}$$

The optimal objective function value for k^{th} block is given as follows [12]:

$$f_{[k]}\left(\hat{y}_{[k]}\right) = \left(c^e + c^l\right)\frac{\overline{\sigma}_{[k]}}{\sqrt{2\pi}}e^{-z^2}. \tag{24}$$

(24) is an increasing function of $\overline{\sigma}_{[k]}$. Hence, smallest-variance first rule is the optimal sequencing rule.

5. Conclusions

Three newsvendor models are addressed to match supply with demand, or vice versa. The traditional newsvendor model can answer how much to order, given that the newsvendor knows demand distribution. The inverse newsvendor model applies to the strategic level decision, e.g., how many customers should be allocated in a fixed capacity. Time-based newsvendor model has been used for service-oriented settings (e.g., operating rooms, law firm). The sequential newsvendor model determines the sequence of the assigned customers by the strategic inverse newsvendor model, and prescribes the corresponding expected arrival times of the customers. The optimal sequence should be a variability-based rule, because the objective function elements involve partial expected values: overage vs. underage, overusage vs. underusage, or earliness vs. lateness. The smallest-variance-first assignment rule is optimal to minimize the expected earliness and lateness when the newsvendor takes into account normally distributed service durations.

All newsvendor models keep balance between surplus (i.e., supply > demand) and deficit (supply < demand), accepting the fact that the newsvendor cannot match supply with demand all the time. Supply chain professionals may face with either case of surplus or deficit, not matched. When a product is highly profitable (or net profit is greater than net loss), the newsvendor tends to order more than the average to avoid the lost sales in the long run. On the contrary, when net loss is greater than net profit, the newsvendor tends to order less than the average in the long run. However, if the newsvendor is myopic, the newsvendor tends to order average demand without respect to cost structure or demand shape, so-called pull-to-center [18]. To avoid pull-to-center bias, supply chain professionals must understand how to get optimal decision considering cost structure and demand parameters for long-term perspective.

Abbreviations

CSL   Customer Service Level
OR   Operating Room
SAA   Sample Average Approximation
SV   Smallest Variance

Author details

Sangdo Choi
Bloomsburg University of Pennsylvania, Bloomsburg, PA, USA

*Address all correspondence to: sangdo.choi@snu.ac.kr

References

[1] Cachon GP, Terwiesch C (2004) *Matching supply with demand: An introduction to operations management* (New York: McGraw-Hill Custom Publishing).

[2] Chopra S, Meindl P (2014) *Supply Chain Management, Strategy, Planning, and Operation* (New York: Pearson).

[3] Jacobs FR, Chase RB (2014) *Operations and Supply Chain Management* (New York: McGraw-Hill Irwin).

[4] Silver EA, Peterson R (1985) *Decision Systems for Inventory Management and Production Planning* (New York: John Wiley & Sons).

[5] Erlebacher SJ (2000) Optimal and heuristic solutions for the multi-item newsvendor problem with a single capacity constraint. Production and Operations Management 9:303 – 318.

[6] Punia S, Singh SP, Madaan JK (2020) From predictive to prescriptive analytics: A data-driven multi-item newsvendor model. *Decision Support Systems* 136:113340.

[7] Bakal IS, Geunes J, Romeijn EE (2008) Market selection decisions for inventory models with price-sensitive demand. *Journal of Global Optimization* 41:633 – 657.

[8] Carr S, Lovejoy W (2000) The inverse newsvendor problem: Choosing an optimal demand portfolio for capacitated resources. *Management Science* 46:912–927.

[9] Choi S, Ketzenberg M (2018) An inverse newsvendor model to set the optimal number of customers in a capacitated environment. *International Journal of Production Economics* 1:1–12.

[10] Merzifonluoglu Y, Geunes J, Romeijn H (2012) The static stochastic knapsack problem with normally distributed item sizes. *Mathematical Programming* 134:459 – 489.

[11] Taaffe K, Geunes J, Romeijn HE (2008) Target market selection and marketing effort under uncertainty: The selective newsvendor. *European Journal of Operational Research* 189:987 – 1003.

[12] Choi S, Wilhelm WE (2014) An approach to optimize master surgical block schedules. *European Journal of Operational Research* 235:138–148.

[13] Birge JR, Louveaux F (1997) Introduction to Stochastic Programming. Springer, New York

[14] Blake J, Donald J (2002) Mount sinai hospital uses integer programming to allocate operating room time. Interface 32:63 – 73.

[15] Denton BT, Miller AJ, Balasubramanian HJ, Huschka TR (2010) Optimal allocation of surgery blocks to operating rooms under uncertainty. Operations Research 58:1 – 15.

[16] Fei H, Chu C, Meskens N (2009) Solving a tactical operating room planning problem by a column-generation-based heuristic procedure with four criteria. Annals of Operations Research 166:91 – 108.

[17] Porteus E (2002) Foundations of Stochastic Inventory Theory. Stanford University Press, Stanford

[18] Schweitzer M, Cachon G (2000) Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. *Management Science* 46:404–420.