40.012 MANUFACTURING AND SERVICE OPERATIONS

Assignment 3

Michael Hoon

1006617

June 2, 2024

Question 1

(a)

Since we are working with discrete demand here, we will employ the optimal policy newsvendor model for discrete demand. We define the underage cost c_u as the cost per unit of unsatisfied demand, and the overage cost c_o to be the cost per unit of positive inventory remaining at the end of the project. Thus, $c_u = \$10 - \$5 = \$5$ (lost profit from not selling a drink that could have been sold), and $c_o = \$5 - \$2 = \$3$ (difference between the cost of producing a drink and the salvage value of an unsold drink).

In the continuous model, the optimal solution is the value of decision variable Q that makes the distribution equal to the critical ratio, defined as

$$F(Q^*) = \frac{c_u}{c_u + c_o}$$
$$= \frac{5}{5 + 3}$$
$$= 0.625$$

In the discrete case however, the critical ratio will generally fall between two values of $F(Q)$, and the optimal solution procedure is to local the critical ratio between two values of $F(Q)$ and choose the Q corresponding to the **higher value** (i.e. we need to find the smallest order quantity for which the cumulative probability equals or exceeds the critical ratio). The table of cumulative probability distribution is given below:

Piney Drinks	30	40	50	60	**70**	80	90
PDF	0.05	0.12	0.20	0.24	0.17	0.14	0.08
CDF	0.05	0.17	0.37	0.61	**0.78**	0.92	1.00

Table 1: Probability Distribution Tables

From Table 1, we can see that the smallest order quantity where the CDF exceeds the critical ratio $F(Q^*) = 0.625$ is 0.78, which corresponds to 70 drinks. Thus, Piney Pineapple Drinks should plan on taking **70 drinks to the beach each day.**

(b)

Since we are now working with a normally-distributed demand with mean 60 and variance 81, demand is now a random variable with distribution $D \sim \mathcal{N}(60, 81)$. Using the same critical ratio as before, we obtain a standardized value of $Z = 0.32$ from the Normal Distribution table. The optimal Q^* is then obtained by scaling:

$$Q^* = \sigma Z + \mu$$
$$= \sqrt{81} \times 0.32 + 60$$
$$= 62.88$$
$$\approx \boxed{63}$$

Thus, they should plan on taking **63 drinks to the beach every day** assuming a normally-distributed demand.

(c)

If it is equally likely that they get orders between 30 and 90 units on Sundays, we are working with a Uniformly Distributed Demand, with $D \sim U(30, 90)$. The probability p of each discrete unit being sold is the same, where $p = \frac{1}{7}$:

Piney Drinks	30	40	50	60	**70**	80	90
PDF	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$
CDF	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{3}{7}$	$\frac{4}{7}$	$\frac{5}{7}$	$\frac{6}{7}$	1

Table 2: Uniform Probability Distribution for Sunday

From the table, the smallest order quantity where the CDF exceeds the critical ratio $F(Q^*) = 0.625$ is $\frac{5}{7} = 0.714$, which corresponds to 70 drinks. Thus, Piney Pineapple Drinks should plan on taking **70 drinks to the beach each day.**

Question 2

(a)

Considering the usual assumptions of the lot-size reorder point system, we define the following parameters:

1. Proportional Order Cost $c = \$10$ per unit

2. Lead time $\tau = 6$ months

3. Setup Cost $K = \$50$ per order

4. Holding Cost $h = Ic = 0.20 \times \$10 = \$2$ per unit

5. Stock-out Cost $p = \$25$ per unit

The demand is uniformly distributed from 270 to 330, so we have that $D \sim U(270, 330)$ for 6 months. We assume that the mean rate of demand is thus $\lambda = 50 \times 12 = 600$ units per year. Defining $G(Q, R)$ as the expected average annual cost of holding, setup, and shortages, we have:

$$G(Q, R) = h \left(\frac{Q}{2} + R - \lambda\tau \right) + \frac{K\lambda}{Q} + p \left(\frac{\lambda n(R)}{Q} \right) \tag{1}$$

Where $n(R)$ is the expected nunmber of stock-outs in the newsvendor model:

$$n(R) = \mathbb{E}(\max(D - R, 0)) = \int_{R}^{\infty} (x - R)f(x)\,\mathrm{d}x \tag{2}$$

Since we have a uniformly distributed demand D, the PDF of the annual demand $f(x)$ is given by:

$$f(x) = \begin{cases} \dfrac{1}{330 - 270} = \dfrac{1}{60}, & 270 \le x \le 330 \\ 0 & \text{otherwise} \end{cases}$$

and the CDF $F(x)$ is given by:

$$F(x) = \begin{cases} 0 & x < 270 \\ \dfrac{x - 270}{60} & 270 \le x \le 330 \\ 1 & x > 330 \end{cases}$$

The objective now is to find the optimal order quantity Q^* and R^* that minimises $G(Q, R)$. Employing the iterative algorithm, we have that

$$\begin{aligned} Q_0 = EOQ &= \sqrt{\frac{2K\lambda}{h}} \\ &= \sqrt{\frac{2 \times 50 \times 600}{2}} \\ &= 100\sqrt{3} \\ &= 173.20508 \end{aligned}$$

Next, we need to find R_0 from the equation

$$1 - F(R) = \frac{Qh}{p\lambda} \tag{3}$$

Using $Q_0 = 100\sqrt{3}$, we have:

$$1 - F(R_0) = \frac{100\sqrt{3} \times 2}{25 \times 600}$$

$$R_0 = 60 \times \left(1 - \frac{100\sqrt{3} \times 2}{25 \times 600}\right) + 270$$

$$= 328.614359$$

To find $n(R_0)$, we use Equation :

$$n(R_0) = \int_{R_0}^{\infty} (x - R_0) f(x) \, dx$$

$$= \int_{R_0}^{330} \frac{x - R_0}{60} \, dx$$

$$= 0.016$$

with this, we can find Q_1 with the equation:

$$Q = \sqrt{\frac{2\lambda[K + pn(R)]}{h}} \tag{4}$$

$$Q_1 = \sqrt{\frac{2 \times 600 \times [50 + 25 \times 0.016]}{2}}$$

$$= 12\sqrt{210}$$

$$= 173.89652$$

for R_1, we repeat with Equation 3

$$1 - F(R_1) = \frac{Q_1 h}{p\lambda}$$

$$R_1 = 60 \times \left(1 - \frac{21\sqrt{210} \times 2}{25 \times 600}\right) + 270$$

$$= 328.60882$$

Comparing the obtained values, we have

$$Q_0 = 173.20508 \qquad\qquad R_0 = 328.614359$$
$$Q_1 = 173.89652 \qquad\qquad R_1 = 328.60882$$

we see that the values are very similar, thus we can conclude that the algorithm has converged, and the optimal order quantity Q^* and reorder point R^* is $(Q_1, R_1) \approx (174, 329)$.

(b)

We consider a Type I service level in a (Q, R) system, where probability of not stocking out during the order arrival lead-time is given by α. Since α completely determines R^*, we can decouple the computations of Q^* and R^*. Since this is the proportion of order cycles without stockouts, we have $P(D \leq R) = F(R)$. To find α, we just set $F(R) = \alpha$ as follows:

$$
\begin{aligned}
\alpha &= F(R_1) \\
&= \frac{329 - 270}{60} \\
&= \frac{59}{60} \\
&\approx \boxed{0.983}
\end{aligned}
$$

The probability of not stocking out during the order arrival lead-time is thus around 98.3%.

(c)

We now consider a Type II service, where the proportion of demand that is met from stock is defined as β (Fill Rate). Since $n(R)/Q$ is the average fraction of demands that stock out each cycle, then we have that

$$
\frac{n(R)}{Q} = 1 - \beta \tag{5}
$$

Using Equation 5, we have

$$
\begin{aligned}
\beta &= 1 - \frac{n(R_1)}{Q_1} \\
&= 1 - \frac{\int_{329}^{330} \left(\frac{x - 329}{60} \right) \, \mathrm{d}x}{174} \\
&= 0.999952107
\end{aligned}
$$

Thus, around 99.995% of the demand is met from stock.

Question 3

Question 4

Question 5

Multi-Echelon Inventory Management (MEIM) systems are crucial in optimising inventory levels (across multiple stages) of a supply chain to minimise overall costs while meeting customer demand efficiently. Instead of studying a classic paper on MEIM, we will conduct a short literature review on **more recent transformative and novel research** being done in this domain.

The paper titled "Multi-echelon inventory optimization using deep reinforcement learning" [1] published in the Central European Journal of Operations Research, studies the applicability of a **Deep Reinforcement Learning (DRL)** approach to different multi-echelon inventory systems, with the objective of **minimizing the holding and backorder costs**. Recent advancements in IT and data exchange have also allowed for more sophisticated, multi-echelon inventory optimization (MEIO) techniques. However, these optimisation models can quickly become very complex, and rely heavily on assumptions and simplifications.

A more promising method introduced by the authors involves using Reinforcement Learning (RL), which was originally developed to solve sequential decision making problems in dynamic environments (including Robotics). Due to recent advances in Deep Learning, RL is often combined with neural networks to form DRLs. The authors review existing literature on MEIM and the application of RL here, noting that quite some research has already been done. They highlight the complexities involved in coordinating inventory policies across multiple locations and how DRL, particularly **Proximal Policy Optimization (PPO)**, can address these challenges.

The study focuses on three specific network structures: linear, divergent, and general (based on real-life manufacturer). The problem is modelled via a **Markov Decision Process (MDP)** where all decisions are made centrally by a single decision maker. The PPO algorithm is implemented to solve the MDPs in an approximate manner, and handle the continuous action space inherent in inventory management problems. The steps include state space transformation, action space transformation, and reward function transformation to tailor the PPO for inventory management tasks. The state space includes variables like inventory levels and demand rates, while the action space pertains to order quantities. The reward function is designed to penalize holding and backorder costs, driving the optimization process towards cost-effective inventory management (via hyperparameter tuning through computing a loss function). The results are as follows: with a **Linear Network**, the PPO algorithm achieved an average improvement of 16.4% over benchmark solutions, the **Divergent Network** with an 11.3% improvement, and **General Network** with a 6.6% improvement. This demonstrates the effectiveness of the PPO algorithm in **reducing costs** and improving efficiency across different network structures, and how DRL can **adapt to various complexities and dynamics in a MEIM system**.

The author's contribution to this research was threefold. A DRL method was applied to a continuous action space (not yet widely used) to a MEIM system, creating **better scalability** to larger problems. Second, they investigated a general inventory system, which has received less attention in literature due to its complexity and intractability, and showed that the PPO algorithm can **improve inventory policies**. Lastly, they demonstrated the flexibility of the PPO algorithm by applying it to a variety of cases **without major modifications** and **without any parameter tuning**. They also noted that more research should be done on the **explainability of the method** and DRL in general, as it is a key component of implementing such methods in practice.
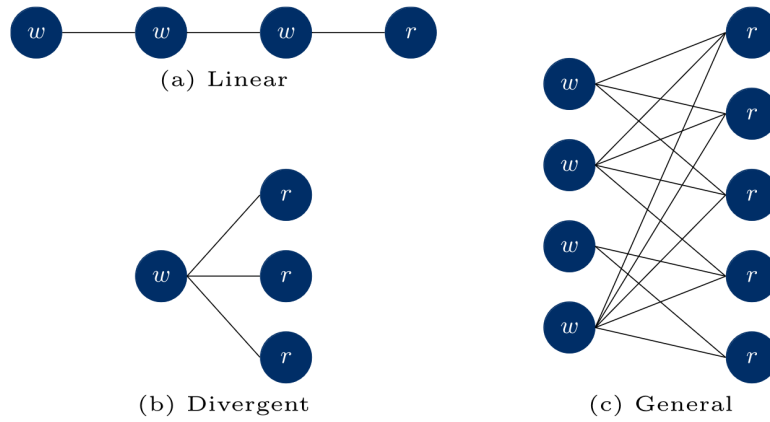
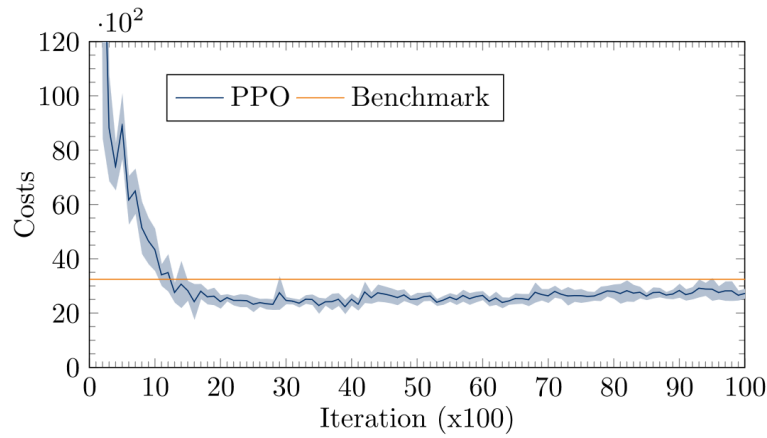Figure 1: Three Types of Network Structures, with Warehouses w and Retailers r

Figure 2: Performance of the PPO algorithm in the **linear** inventory system. The blue shade depicts the 95% confidence interval
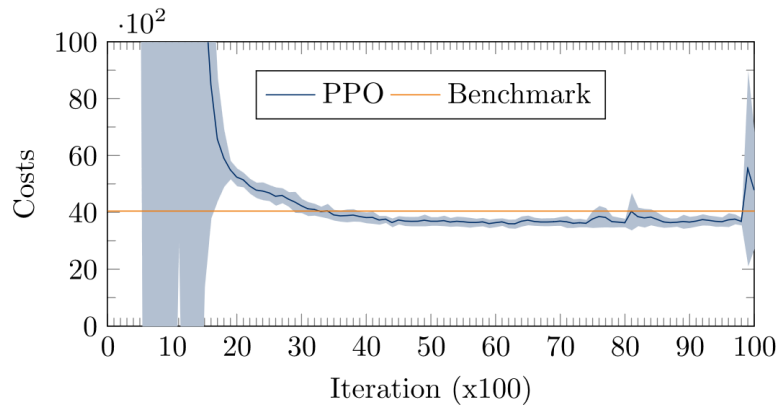

Figure 3: Performance of the PPO algorithm in the **divergent** inventory system. The blue shade depicts the 95% confidence interval
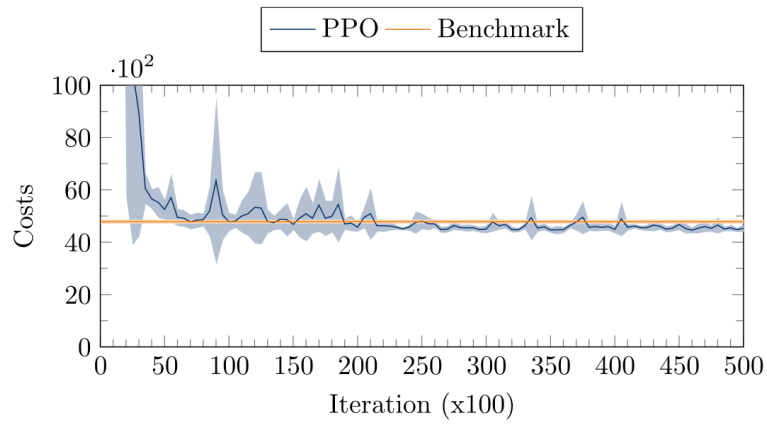
Figure 4: Results of the PPO algorithm on the case of manufacturing Company (**General Network**). The blue shade depicts the 95% confidence interval

References

[1] Kevin Geevers, Lotte van Hezewijk, and Martijn R. K. Mes. Multi-echelon inventory optimization using deep reinforcement learning. *Central European Journal of Operations Research*, July 2023.