OPTIMIZATION PROJECT I

Team 13

Pratik Gawli (pbg397), Sai Bhargav Tetali (srt2578), Varun Kausika (vsk394), Yingjia Shang (ys23737)

Problem Statement

The Newsvendor Model is a stochastic optimization problem used in inventory management to help determine the optimal order quantity for a product that has uncertain demand. The model is based on the concept of the newsvendor who must decide how many newspapers to order for the next day in order to maximize their profits. The model can be applied to a wide range of products and industries and provides a useful framework for businesses to make informed decisions about their inventory levels and order quantities. In its simplest form, the newsvendor problem consists only of the cost price of printing the newspaper and the selling price. The optimal quantity of papers to be printed will be calculated based on the previous day's demand. The optimal quantity will be determined such that the expected value of profit is maximized.

In this project, we are extending the concept and application of the newsvendor model in a few different ways to make a better approximation of reality and solve different business problems. Briefly stated, our first extension will assume a shortage in newspapers, where we don't have enough to satisfy the demand. In the second extension, we would incorporate a price-sensitive demand situation, where we assume that demand is affected by price linearly with error. The objective is still to determine the optimal price and quantity of newspapers to print that will maximize the expected profit. The optimal price and quantity will depend on the expected demand, the cost of producing each newspaper, the price that customers are willing to pay, and the degree of price sensitivity of the customers.

Methodology

Let's assume we have previous n number of days of demand data, cost of printing the paper as c, the selling price as s, and the optimal quantity as q. The objective in mathematical form would be the following.

Objective:

$$\text{Max. } \frac{1}{n}\Sigma^{n}_{i=1} min(sq - cq,\ sD_i - cq)$$

In the objective, we have a minimum function for each day's profit as the optimal quantity can be less than, equal to or greater than the demand. For each day the profit can be written as the following.

$$P = sq - cq \text{ for } D_i > q$$

Or

$$P = sD_i - cq \text{ for } D_i < q.$$

The objective is not a linear function due to the minimum function and hence we take profit on each day as a variable h and rewrite the objective as below.

$$\text{Max. } \frac{1}{n}\Sigma^{n}_{i=1} h_i$$

The constraints for this problem will be such that the profit value will be satisfying them irrespective of which one is greater out of demand and optimal quantity.

Constraints:

$$h_i \leq sq - cq$$

$$h_i \leq sD_i - cq$$

In the above constraints, if demand is higher than quantity, we only sell the quantity, and the profit will be exactly equal to the first equation, and it will still satisfy the second equation as the profit will be lower than the right-hand side value. In the case where quantity is higher than demand, we sell the full demand, and the profit will be exactly equal to the second equation, and it will still satisfy the first equation as the profit will be lower than the right-hand side. This way we convert a non-linear problem into a linear problem.

Now, the above constraints can be re-written as follows for each day:

$$h_i + (c - s)q \leq 0$$

$$h_i + cq \leq sD_i$$

Rush order and disposal rates:

The Newsvendor problem described above is the simplest form of the problem. An extension to this problem is to have a rushed newspaper printing cost g (g>c) which will apply if we produce a lesser quantity than the demand and try to print the remaining using a rush order. If our optimal quantity is more than the demand, then the extra newspapers need to be disposed of which we are going to assume will be at a cost of t per newspaper. With these extensions, the equation for profit made on each day changes to

$$P = sD_i - cq - g(D_i - q) \text{ for } D_i > q$$

Or

$$P = sD_i - cq - t(q - D_i) \text{ for } D_i < q$$

The objective expression will remain the same as above, which is maximizing the expected profit whereas the constraints will change due to the extra terms in profit calculation.

Constraints:

$$h_i \leq sD_i - cq - g(D_i - q)$$

$$h_i \leq sD_i - cq - t(q - D_i)$$

Where h_i still represents the profit made on day i.

Which upon rearranging the terms would be:

$$h_i + (c - g)q \leq (s - g)D_i$$

$$h_i + (c + t)q \leq (s + t)D_i$$

Note that in the code, we formulated this in the same way as we would when price is allowed to affect demand, except we added a constraint to restrict the price to be a constant and solved it linearly. Furthermore, we ignored any residual terms which were included in the objective since in a regression all the residuals sum to zero anyways.

Another way to solve this would be to hardcode the price to be a constant and solve the linear programming problem for the minimal cost.

Price effects on Demand:

The price of a product will always have an effect on its demand. Till now, we have assumed that it doesn't and built our mathematical expressions. Now, let's assume that price has a linear effect and modify our equations accordingly.

With the data that we have on the previous day's demand and price we can fit a linear regression model which will give us linear expression as below:

$$D_i = \beta_0 + \beta_1 s_i + \varepsilon_i$$

Now we need to find an optimal selling price (s) along with the optimal quantity to print which would maximize the expected profit. We use the residuals (ε_i) to vary the demand every day and assume that they are the source of randomness in the data. Hence, the demand on day I would be:

$$D_i = \beta_0 + \beta_1 s + \varepsilon_i$$

Where s = optimal selling price

Substituting the above expression for demand in profit we get the following:

$$P = s(\beta_0 + \beta_1 s_i + \varepsilon_i) - cq - g((\beta_0 + \beta_1 s_i + \varepsilon_i) - q) \text{ for } D_i > q$$

Or

$$P = s(\beta_0 + \beta_1 s_i + \varepsilon_i) - cq - t(q - (\beta_0 + \beta_1 s_i + \varepsilon_i)) \text{ for } D_i < q$$

If we observe the revenue part of the profit, we see that there is a quadratic term in s which cannot be used in constraints or complicates the problem. To deal with this we consider the negative cost as our variable h instead of profit. The negative of cost on each day will look the following:

$$C =- cq - g((\beta_0 + \beta_1 s_i + \varepsilon_i) - q) \text{ for } D_i > q$$

Or

$$C =- cq - t(q - (\beta_0 + \beta_1 s_i + \varepsilon_i)) \text{ for } D_i < q$$

Taking the negative cost as a dummy variable h we again get two constraints for each day due to the uncertainty of whether the optimal quantity will be greater or lesser than demand.

Constraints:

$$h_i \leq - cq - g((\beta_0 + \beta_1 s_i + \varepsilon_i) - q)$$

$$h_i \leq - cq - t(q - (\beta_0 + \beta_1 s_i + \varepsilon_i))$$

The above equations are written based on the logic that on any given day if demand is higher than optimal quantity, then we use the rush order which would be the actual cost and this value will be greater than the value of the cost calculated using the equation for the case and vice versa. Since we are taking the negative costs, the greater than symbol changes to a less than symbol.

Coming to the objective of the problem, we are still maximizing the expected profit. Since, we changed our dummy variables the mathematical expression will be changed to:

Objective:

$$\text{Max. } \frac{1}{n}\Sigma^{n}_{i=1}s(\beta_0 + \beta_1 s + \varepsilon_i) + h_i$$

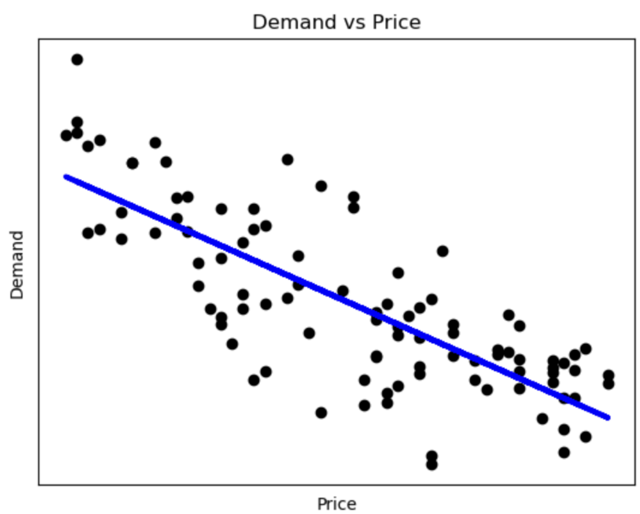
Which upon simplification it can be written as:

$$\text{Max. } \beta_1 s^2 + \beta_0 s + \frac{1}{n}\Sigma^{n}_{i=1}\varepsilon_i s + h_i$$

One important note to make here is that the sum residuals will be close to zero which is the property of linear regression. If we decide on a selling price and just solve for the optimal price, we will go back to the previously explained problem type where profits were taken as dummy variables. With the representation of demand in terms of the price here we can continue to take negative costs as h variables to keep the constraints linear and solve the problem as a linear program that would give the same result as the one obtained from following the previous method.

Analysis Results

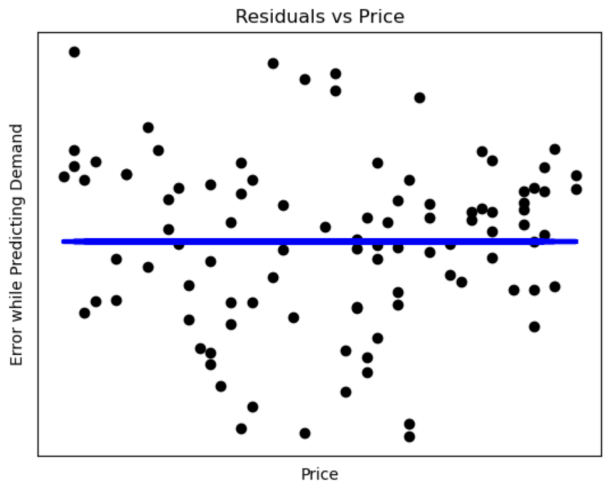
Given the data on price and demand, we first fitted a linear regression model to the dataset to observe the trend in the data:



The linear model we get from fitting the regression is:

Demand = -1367.71*price + 1924.72

Combining the equation and the graph, we find out that price has a negative effect on the demand of newspaper. As the price goes up, the demand for newspapers goes down. This is further proven by the negative coefficient of the model. To ensure the model validity, we have also plotted the residuals to ensure homoscedasticity:



By checking for homoscedasticity through residual plots, we can assess the validity of statistical inference and the accuracy of the predicted values. As we see from the graph, the plot shows a random scatter of points with no discernible patterns, which means the variance of the residuals is constant across all levels of the fitted values.

As the next step, we set c = 0.5, g = 0.75, and t = 0.15. Using the residuals, we generate the demand data using price = 1. The brief overview of the transformed data is as follows:

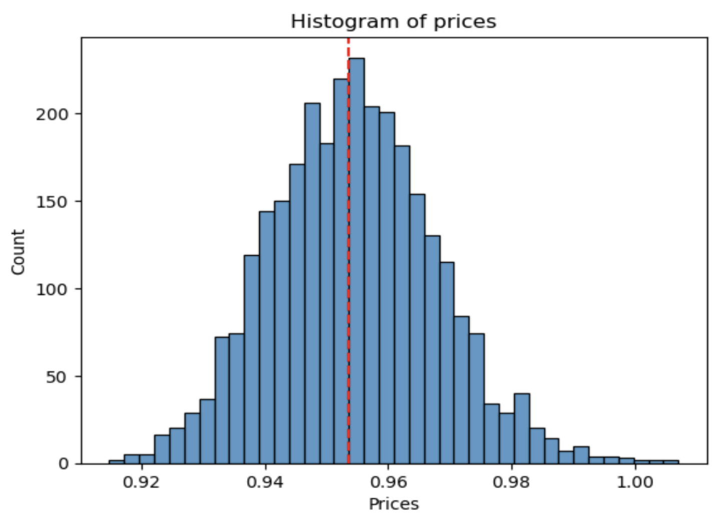
	price	demand	predicted	residuals	gen_demand
0	1.05	283	-3.536025e-14	-205.619393	351.385626
1	0.86	771	-1.245151e-13	22.515227	579.520247
2	1.21	185	3.971753e-14	-84.785389	472.219630
3	0.94	531	-8.697623e-14	-108.067771	448.937249
4	0.76	1002	-1.714387e-13	116.743975	673.748994

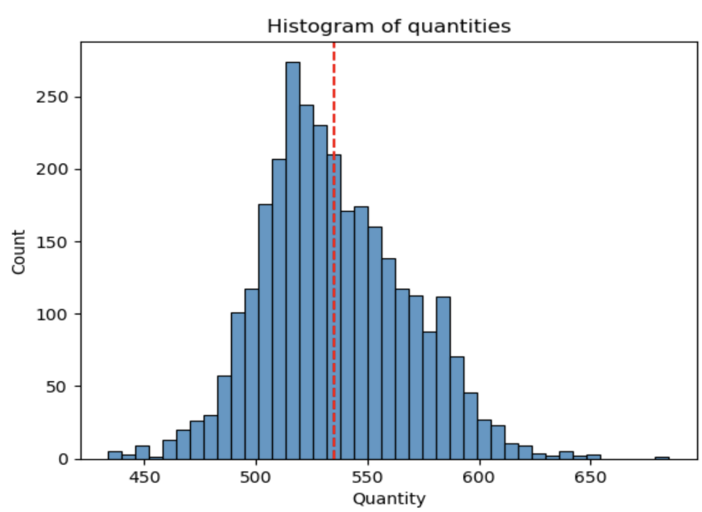
Then, we solve for the optimal quantity to produce when p = 1, forming it as linear programming. To be more specific, we solved this as integer programming because we constrained the quantity of newspapers to be an integer to produce a more accurate level of profit. Our resulting optimal quantity is 472, and the corresponding profit is approximately $231.48.

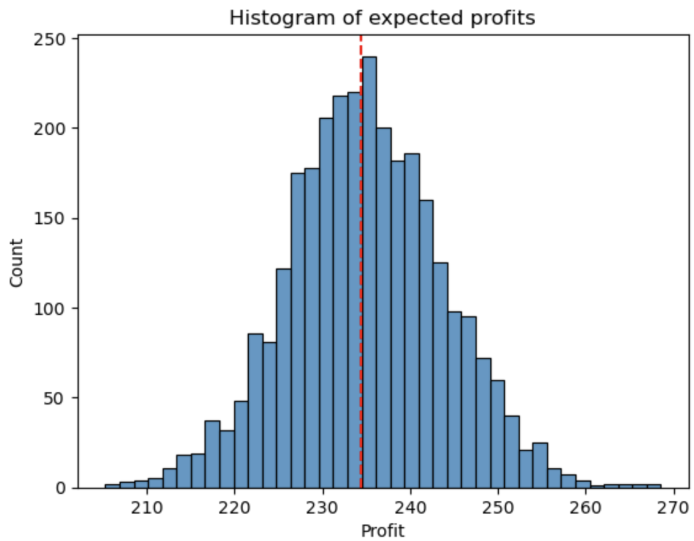
We thus implement the price and demand impact into our optimization problem and solve it as a quadratic programming problem. We still constrained the quantity of newspapers to be an integer. The resulting optimal quantity is 535, and the corresponding will be $234.41. Additionally, the optimal price level would decrease to approximately $0.95 per newspaper. Comparing this with the result we get from not integrating the effect of price and demand, we can observe that the optimal level of quantity boosts by a moderate amount, but the corresponding profit does not differ significantly. This is probably due to a slight decrease in price. As we decrease the price level, the demand for the newspaper also increases.

As the next step, we are interested in how sensitive the optimal price and quantity are to our dataset. So we took a bootstrap sample of the original dataset and fit the new betas to the new bootstrapped sample by repeating the quadratic programming section. This time, we find the optimal quantity to be 544, and the corresponding profit is $240.71. Additionally, the optimal price has decreased to $0.94. As we can see, both the quantity and profit have increased a little bit while the price decreased a bit. But overall, the level of quantity and profit stays within a stable range.

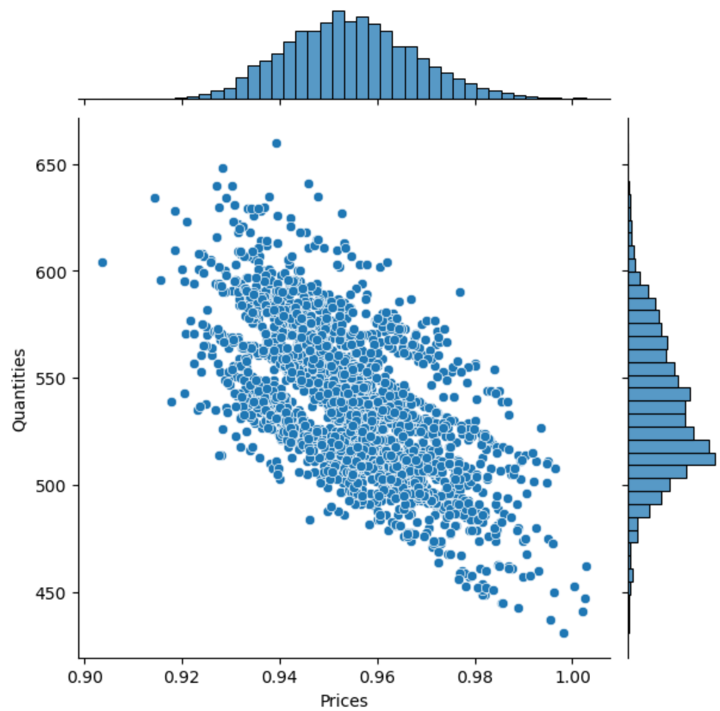
To further assess the sensitivity of price and quantity level to the dataset, we have repeated the process of bootstrapping to observe the distribution of the bootstrapped results. We have repeated the bootstrap 3000 times, fitting new betas to the newly simulated data each time to find the optimal price and quantities. To visualize the distribution of newly generated results, we have included different graphs to illustrate our findings.







The three graphs display the distribution of frequency that each quantity, price, or profit level that occurs during bootstrap. In the price histogram, we can observe that most of the optimal prices level are within the range of approximately 0.94 to 0.96. This also verifies the price levels we can from the previous two quadratic programming results (0.94 and 0.95). The red line represents the price level we got from solving the Linear Programming problem, and the value falls within this range. However, we find out the optimal quantity level differs a lot from our previous results. On the second histogram, we added a red line to represent the optimal quantity we get when we were solving the linear programming part (without price and demand adjustment). The distribution of the most frequent quantity focuses on the range of 500 to 550. Thus, we can summarize that, after implementing the price and demand effect, the quantity is likely to go up, with a slight adjustment in the decrease of price levels. In the expected profit histogram, we also added the red line representing the optimal profit we found during linear programming. As we can see, the range of expected profit focuses on around \$230 to \$240, and the result we got from LP is included in this range, as well as the results we got from quadratic programming (\$234.41 and 240.71). So we can conclude that most of the time, the optimal profit would be within this range.



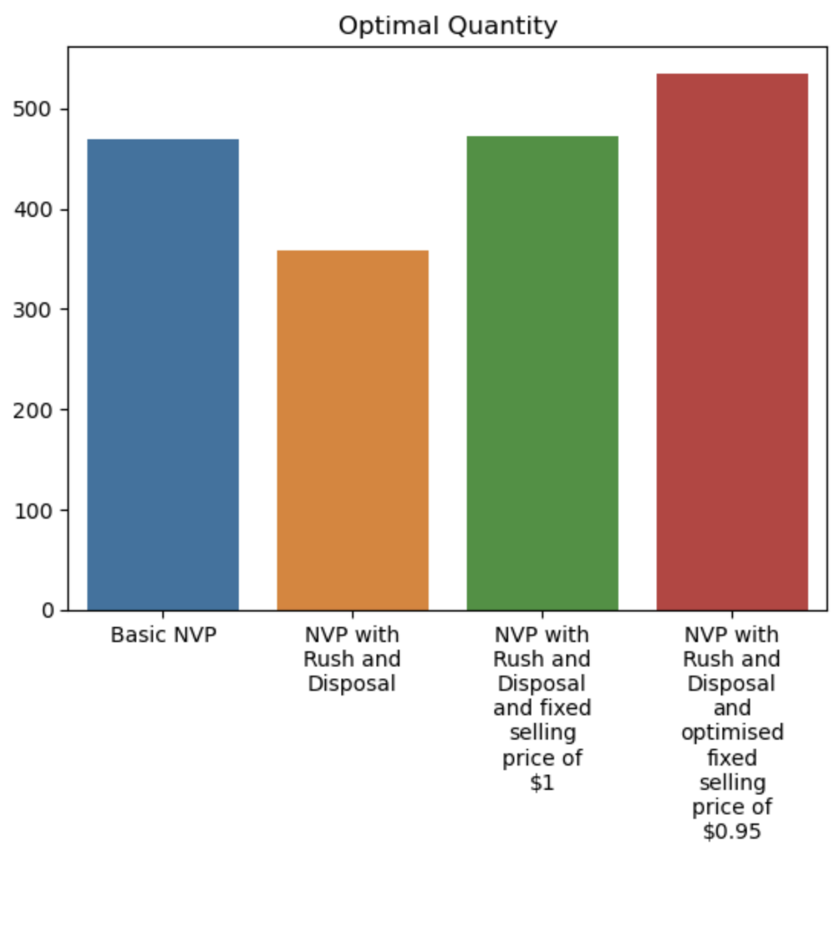
In this graph, we combine the quantity and price histograms and graph the scatterplot to represent the relationship between the two variables. We are able to observe a downward trend. This trend is very similar to the linear regression model we trained at the very beginning of the process. Thus, we can further prove the negative impact that prices have on the demand for newspapers.

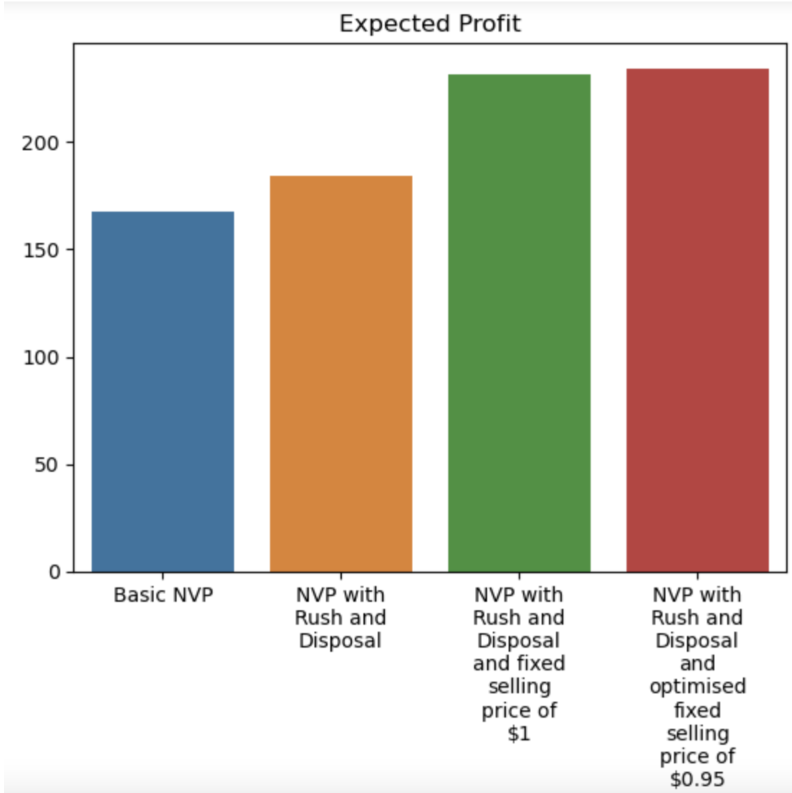
Model Evaluation

To compare the standard newsvendor model, the model with rush and disposal, and the model with optimized price levels, we have obtained the results for both the basic model and the optimized models. In the basic model, we are taking a different selling price on each day assuming that price has no effect on demand. This has been done for the next model as well where we have the options of rush orders and disposal. For the more basic model, since the company has been using it for several years, the ease of use and simplicity would save time for the company to determine an optimal pricing strategy. Since it is just a simple mathematical model that can be used to determine the optimal order quantity for a single-period inventory problem with uncertain demand, it provides a quick and intuitive way to balance the trade-off between stocking too much inventory and risking overstocking, and stocking too little inventory and risking stockouts. It does not require a lot of data or complex calculations, making it accessible to small and medium-sized businesses with limited resources. Additionally, the basic model can be used as a benchmark for evaluating more sophisticated inventory models that incorporate dynamic pricing, like what we are doing with the quadratic programming problem.

However, the basic newsvendor model has limitations in its assumptions, such as the assumption of a single period and no effect of the selling price on demand. These assumptions may not hold in more complex real-world scenarios. For example, if we observe the next model we get a higher expected profit with a lower optimal quantity as now we have the flexibility to produce less due to the availability of rush orders. But this again depends on the cost of rush orders and disposal. Also, as we see from the linear regression model, the demand for newspapers would decrease as the price level increases in the real world. Additionally, the demand may also be affected by many other realistic factors in the real world such as the emergence of media, material shortages, or the demand changes in newspapers. Thus, a more sophisticated optimization newsvendor model can be introduced to the company.

The advantage of a newsvendor model with an optimized price level is that it allows the company to adjust the selling price based on the level of demand as well as the company's remaining inventory, thus leading to higher profit and better inventory management. Additionally, the company can also adjust the price level in response to changes in the market, such as competition, seasonality, or change in demand.





We have included two graphs to illustrate the optimal quantity and expected profit of each newsvendor model. From the optimal quantity graph, we see that while the model with the optimized price level achieves the highest quantity sold, it doesn't have a big difference in quantity from the basic model. However, looking at the expected profit graph, the model with the optimized price level achieves the highest profit, which is approximately $50 higher than the basic model. With the optimized level of prices, this model can capture all the benefits of market fluctuations and provide a more accurate and profitable approach to pricing strategy. Thus, we would conclude that the model with an optimized selling price is the most profitable approach.

One disadvantage of the optimization model is that it can be complex and difficult to implement compared with the basic model. It requires more data, calculations, and sophisticated algorithms. Additionally, it can be more difficult to communicate and understand the customers, who might be used to a fixed price level. Moreover, a dynamic pricing model might be more sensitive to errors in demand forecasting. So the implementation of this model must be very careful and thoughtful when determining the best-optimized pricing strategy.

Conclusion

From the analysis results we get from linear programming and quadratic programming, we believe that using a dynamic newsvendor model would enable the company to regularly adjust its price level to adapt to the most current market trends and conditions. Moreover, the updated model can also help the firm to maximize its profit. However, the increase in profit and change in quantities to be printed is not by a large factor, so one should take into consideration the compute costs for more complex models to be used while making a decision about the same. The obtained profits and quantities to be printed from the bootstrapped data are in line with the Central Limit Theorem. Thus, we would suggest the company adopt the model with the optimized pricing level if only increasing profits is the objective without any budget constraints, and use the standard newsvendor model as a benchmark for determining the pricing strategy.