Optimization Project 1

Team Members:

- Karthick Ramasubramanian *(kr33733)*
- Raymond Li *(rll2497)*
- Rohitashwa Chakraborty *(rc47878)*
- Soumya Nayak *(sn25829)*

Introduction

The goal of this project is to solve a more sophisticated (and more realistic) version of the newsvendor problem. This project involved identifying an optimal quantity to print given a fixed price using linear programming and then identifying an optimal quantity and an optimal price using stochastic programming. It was also the goal of this project to determine whether these estimates vary a lot if the training data was changed using bootstrapping. All the results of our investigations are detailed in this report along with supporting graphs and visualisations.

Regression and Residuals (Q1)

To begin, we wanted to fit a model that would help us predict demand at various price points. In order to do this, we fit a linear regression model on past data points. From this linear model, we are able to see that the equation to determine demand, given a certain price is:

$$Demand = 1924.7175 - 1367.7125x - \epsilon,$$

where x represents a given price and ϵ represents a random error term, which will be represented by the residuals of the data. Later on, we will be using these residuals to generate many demand values for a single price point to help us find the optimal quantity to produce for our product as well as the price to sell it at.

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	1924.7175	111.334	17.288	0.000	1703.750	2145.685
price	-1367.7125	108.379	-12.620	0.000	-1582.816	-1152.609

Demand Estimation (Q2)

By using the fitted demand equation above and its residuals, we were able to generate 99 estimates of demand when the price, p, is equal to 1. We were able to generate these demand values using the code snippet below:

Assume price = 1, c=0.5, g=0.75, and t=0.15

Use residuals to generate demand data

```
predicted_demand = ols_model.predict(np.ones(X.shape))
predicted_demand += residual
predicted_demand[:20]
```

[25] ✓ 0.2s

```
array([351.38562621, 579.52024662, 472.21963007, 448.93724855,
       673.7489942 , 453.13462041, 430.10324468, 480.8744971 ,
       554.19737186, 720.84312138, 449.58299807, 262.93724855,
       617.83400387, 653.19737186, 569.89675532, 623.96862428,
       224.26012331, 385.06275145, 544.70850097, 690.89675532])
```

With these generated demand values, it will allow us to use linear programming to optimize the quantity of our product to produce.

Linear Programming (Q3)

Now we have to find the optimal quantity to produce when price is fixed at 1. Profit h for a given q can be written as:

$$\max_q \ \text{Profit}_i = \frac{1}{N}\sum_{i=1}^{N}\left(pD_i - qc - g(D_i - q)^+ - t(q - D_i)^+\right)$$

where:

- $(x)^+ = \max(x, 0)$
- D_i is the demand on Day ***i***
- q is the **Quantity** produced initially

- p is the **Price** (Sale Price)
- c is the **Cost** of production (Cost Price)
- g is the **Expedited Cost** of production
- t is the **Disposal** Cost

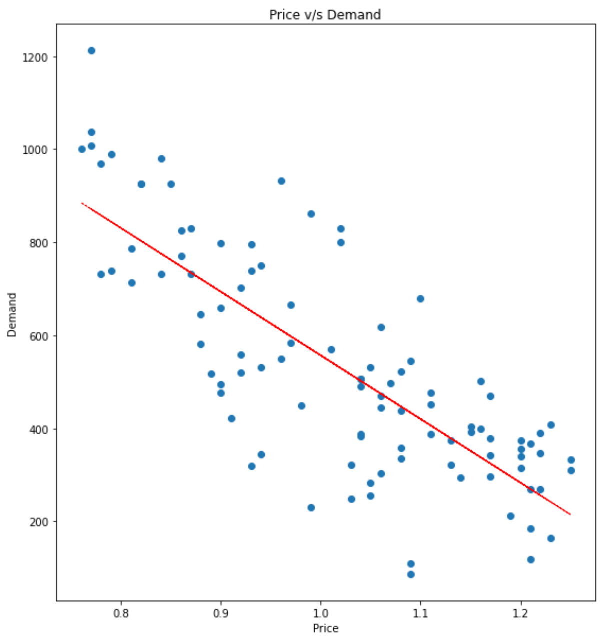
This is an NLP. We can reformulate it into LP of the form:

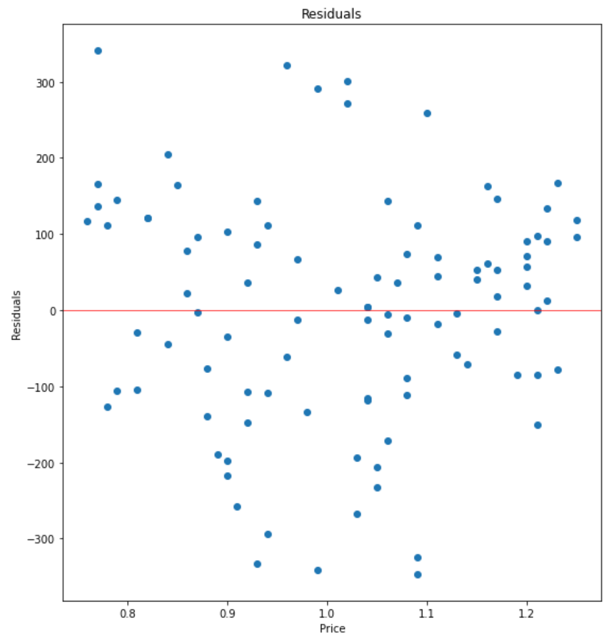
$$\max_{q,h_1,h_2,\ldots,h_N} \quad \text{Profit}_i = \frac{1}{N}\sum_{i=1}^{N} h_i$$

subject to:

$$h + (c-g)q < (p-g)D_i$$
$$h + (c+t)q < (p+t)D_i$$
$$h > -\infty$$
$$q > 0$$

On solving this, the optimal quantity comes out to be 472 newspapers with an expected profit of $231.48 per day.





Stochastic Programming (Q4)

While the above example is a reasonable good approximation for the real world, We know from basic economics that Demand is not an arbitrary quantity, but instead, it is inversely proportional to the price of a commodity. Therefore, the natural progression to newspaper vendor problem is to understand the effect of price on the demand and the number of newspapers to print.

Thus, we now add another extension to our model by making the demand, a function of price. When we do this, our objective function becomes a quadratic expression in ‘p’. We solve this using stochastic programming and quadratic optimization.

$$\max_{q} \ \text{Profit}_i = \frac{1}{N}\sum_{i=1}^{N}(pD_i - qc - g(D_i - q)^+ - t(q - D_i)^+)$$

where:

- $(x)^+ = \max(x, 0)$
- D_i is the demand on Day ***i***
- q is the **Quantity** produced initially
- p is the **Price** (Sale Price)
- c is the **Cost** of production (Cost Price)
- g is the **Expedited Cost** of production
- t is the **Disposal** Cost

Given: $D_i = \beta_0 + \beta_1 p + \epsilon_i$

Therefore, the Equation $pD_i - qc - g(D_i - q)^+ - t(q - D_i)^+$ can be rewritten as:

$$\Rightarrow p(\beta_0 + \beta_1 p + \epsilon_i) - qc - g((\beta_0 + \beta_1 p + \epsilon_i) - q)^+ - t(q - (\beta_0 + \beta_1 p + \epsilon_i))^+$$
$$\Rightarrow p(\beta_0 + \epsilon_i) - [qc + g((\beta_0 + \beta_1 p + \epsilon_i) - q)^+ + t(q - (\beta_0 + \beta_1 p + \epsilon_i))^+] + (p\beta_1 p)$$
$$\Rightarrow p(\beta_0 + \epsilon_i) - h_i + (p\beta_1 p)$$

where

$$h_i = qc + g((\beta_0 + \beta_1 p + \epsilon_i) - q)^+ + t(q - (\beta_0 + \beta_1 p + \epsilon_i))^+$$

upon rewriting, we get:

$$\max_{q,p,h_1,h_2,\ldots,h_N} \ \text{Profit}_i = \frac{1}{N}\sum_{i=1}^{N} p(\beta_0 + \epsilon_i) - h_i + (p\beta_1 p)$$

subject to:

$$1) h_i + (g - c)q - p\beta_1 g > g(\beta_0 + \epsilon_i)$$

$$2) h_i - (t+c)q + p\beta_1 t > -t(\beta_0 + \epsilon_i)$$

$$3) h > -\infty$$

$$4) q > 0$$

Note: here $p\beta_1 p$ is the Quadratic part

This system of equation allows us to optimise for the optimal price at which to sell and the optimal quantity to print at the price, thereby optimising profits.

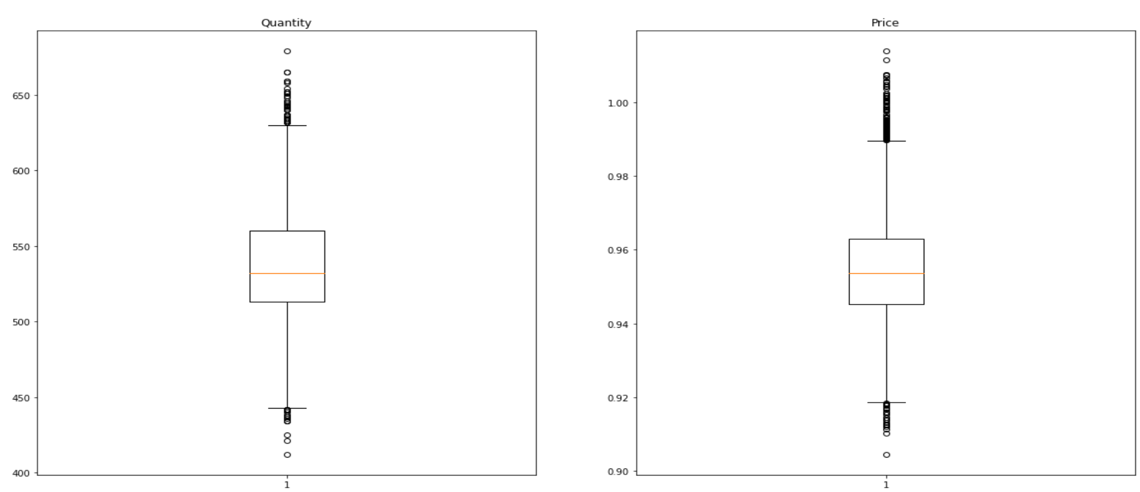
In conclusion, when demand is expressed as a function of price, we get slightly different estimates. The optimal price is about 95 cents and the optimal quantity to be produced is 535 newspapers. Making our expected profits approximately $234.42 dollars.

Bootstrapping (Q6 and Q7)

For the next step we are bootstrapping to check how sensitive the optimal price and quantity are to the dataset. Since our dataset is small, only 99 data points, bootstrapping allows us to make sure that our previous results are not due to overfitting. It also helps us to find confidence intervals.

We run the model, fit Betas via QP method to the bootstrap dataset.

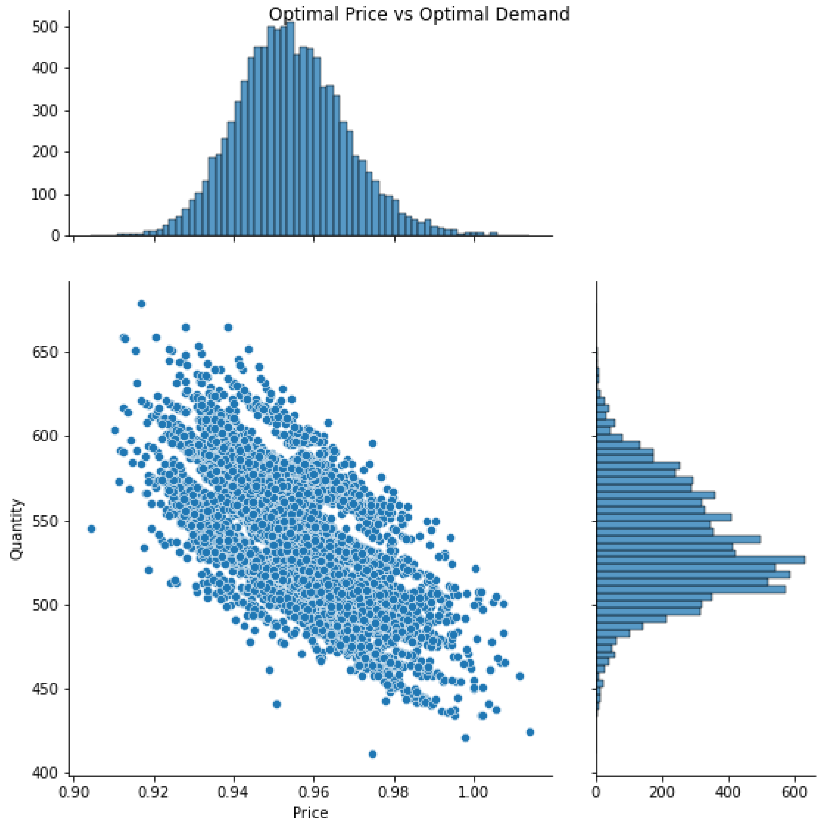
<u>Boxplots for Price and Optimal Quantity:</u>



From the boxplot we observe that optimal quantity varies from 520 to 560 mostly and price between $0.945 to $0.965. The previous results are outside the IQR range.

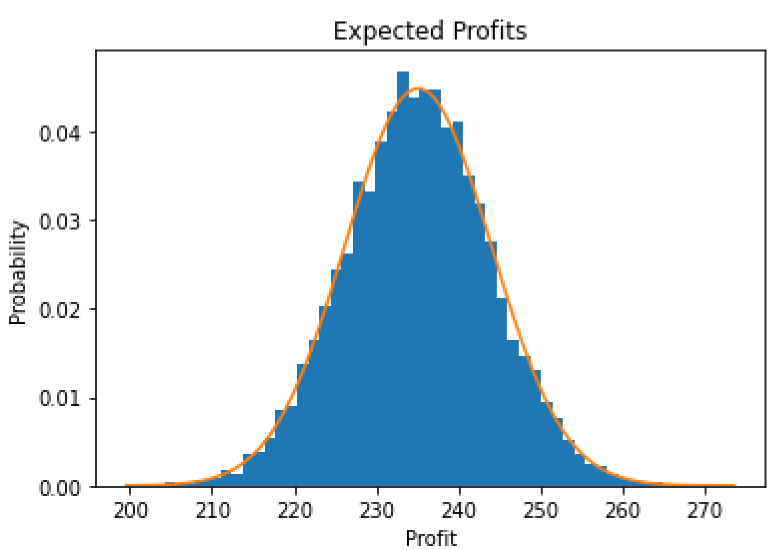
Box plots show that medians for price and optimal quantity are $0.94273 and 597, respectively.

Jointplot for Price vs Optimal Quantity:



As expected we observe a between price and optimal quantity with the Demand Curve.

Histogram for Expected Profit:



Expected Profits generated over 1000 iterations of bootstrapping appear to be normally distributed with a mean of $235.9 which is not so different from $231 obtained before.

Conclusion

Linear programming determines that for a fixed selling price of $1, the optimal number of newspapers to print would be 472. This would yield an expected profit of $231.48 per day.

Extending this model by allowing the demand to be influenced by price (Demand is a function of price), thereby making the price unknown, offers us more flexibility in understanding the optimal pricing and quantities to print to maximise profits. We leverage Stochastic Programming to determine that the optimal quantity the vendor should print is 535 and the price it should set is $0.95.

To check for the sensitivity of our model, we take several bootstrapped samples (≈10,000 iterations) of the training data and find the optimal price, quantity and expected profits for each.

Our observation for a density plot of expected profit approximates to a normal distribution (95% Confidence Interval between $237.15 (± $17.69)), which is line up with the central limit theorem. The optimal price and quantity plots also approximate to a normal distribution with a 95% Confidence interval between $0.95 (± $0.03) and 536 (± 64.08) respectively. This reveals that our estimates are not very sensitive.

In conclusion, We can conclude that Stochastic Programming is a powerful tool to solve complex problems such as estimating future price and quantity of a product that would maximise expected profit utilising historical data.