Project 3: Optimization

Emily Lagnese   Mari-Laura Peña   Luke Hartfield   Riju Hariharan

Fall 2025

Contents

1 Introduction

The newsvendor problem is a classic framework in operations management used to determine how much of a product to produce when demand is uncertain. At its core, the model compares the costs of over-producing and under-producing relative to a random demand realization. Because unsold units are wasted but unmet demand represents lost sales, the decision maker must balance these opposing risks to maximize expected profit. This framework is widely used in publishing, retail, manufacturing, and any setting where production occurs before demand is known.

Modeling demand uncertainty explicitly is essential because relying on point forecasts alone can lead to systematically suboptimal production decisions. Real demand varies from day to day, and ignoring this randomness makes it impossible to quantify the financial trade-offs between printing too much or too little. Incorporating randomness allows us to estimate expected profit more accurately, evaluate the sensitivity of decisions to variation in the data, and understand how different operational policies (like rush-printing or disposal fees) affect optimal behavior.

This project explores several extensions of the standard newsvendor model to better reflect realistic printing operations. First, we incorporate rush-order costs for unmet demand and disposal costs for excess production. Second, we model demand as a function of price and jointly optimize both price and quantity, leading to a quadratic or quadratically constrained program. Finally, we use bootstrap resampling to assess how stable the optimal decisions are across alternative demand samples. Together, these components allow us to compare the traditional model to a more sophisticated one and evaluate whether improved operational decisions can increase profitability for the publishing company.

2 Standard Newsvendor Model

Original profit function

In the standard newsvendor model with fixed unit selling price p, unit production cost c, random demand D_i in scenario i, and order quantity q, the per-scenario profit is

$$\pi_i(q) = p \min\{q, D_i\} - cq.$$

Why it is nonlinear

This objective is nonlinear in q because of the term $\min\{q, D_i\}$. The minimum operator makes the profit piecewise linear and kinked at $q = D_i$, which prevents us from directly using a linear programming solver. To obtain an LP, we introduce auxiliary variables and additional linear constraints that replicate the behavior of the minimum without using nonlinear expressions explicitly.

Linear reformulation using dummy variables

A common reformulation introduces a sales variable s_i for each scenario, representing the number of units actually sold:

$$0 \le s_i \le q, \qquad 0 \le s_i \le D_i.$$

At the optimum we have $s_i = \min\{q, D_i\}$, and the profit can be written as a linear function

$$\pi_i = p\, s_i - cq.$$

The objective is then to maximize the average profit $\frac{1}{n}\sum_{i=1}^{n} \pi_i$ subject to the linear constraints on q and s_i.

Assumptions

Key assumptions of this baseline model include:

- The selling price p and unit production cost c are known and constant.
- Demand scenarios D_i are exogenous and represent the empirical distribution of demand.
- Unsold units have no salvage value and unmet demand cannot be backordered or rushed.
- The decision maker chooses a single order quantity q before demand is realized.

3 Extension 1: Rush Costs and Disposal Costs

Modified objective

In the basic newsvendor model we only had a single production cost c, so profit in scenario i was essentially revenue from sales minus cq. In this extension we add two more realistic costs: a rush (quick-printing) cost r for copies printed after demand is realized, and a disposal cost s for unsold copies. With price fixed at p and demand D_i, the profit in scenario i becomes

$$\pi_i(q) = D_i - cq - g(D_i - q)^+ - t(q - D_i)^+,$$

where the last two terms penalize shortages and surpluses. Our objective is to choose q to maximize the expected (average) profit over all demand scenarios.

Explanation of $(\cdot)^+$

The notation $(x)^+ = \max\{x, 0\}$ denotes the positive part of x. In this model we use

$$(D_i - q)^+ \quad \text{for shortage (how much demand exceeds the order quantity)}$$

and

$$(q - D_i)^+ \quad \text{for surplus (how much we over-produce).}$$

Rush and disposal costs are applied only when these quantities are positive.

Linearization of the new cost structure

Because $(\cdot)^+$ is not linear, we replace each of these terms by auxiliary variables. For each scenario i we create:

$$\begin{aligned} d_i^+ &\geq 0 \quad \text{for shortage,} \\ e_i &\geq 0 \quad \text{for surplus,} \end{aligned}$$

and link them with simple linear constraints:

$$\begin{aligned} d_i^+ &\geq D_i - q, \\ e_i &\geq q - D_i. \end{aligned}$$

At the optimum,

$$d_i^+ = (D_i - q)^+, \qquad e_i = (q - D_i)^+,$$

so total cost becomes linear in q, d_i^+, and e_i.

LP formulation

Putting everything together, the Part 3 model is a linear program with decision variable (order quantity) q and scenario-specific auxiliary variables d_i^+ and e_i. For each scenario we compute profit

$$\pi_i = D_i - cq - gd_i^+ - te_i,$$

and the objective is to maximize the average profit

$$\max \quad \frac{1}{n}\sum_{i=1}^{n} \pi_i$$

subject to the linear shortage/surplus constraints and non-negativity of d_i^+, e_i and any other LP variables.

4 Extension 2: Price-Dependent Demand

Linear regression setup

To model how demand responds to price, we first estimate a simple linear regression of demand on price using the historical data:

$$D_i = \beta_0 + \beta_1 p_i + \varepsilon_i,$$

where D_i is the observed demand when the price is p_i, and ε_i is the random error. Fitting this regression gives us estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ that describe the average price–demand relationship in the data.

Interpretation of $\hat{\beta}_0$ and $\hat{\beta}_1$

In this specification:

- $\hat{\beta}_0$ is the intercept: the baseline demand level when price is zero.
- $\hat{\beta}_1$ is the slope: the change in expected demand for a one-unit increase in price.

Since our estimate $\hat{\beta}_1 < 0$, increasing price reduces expected demand, while decreasing price raises expected demand. The fitted demand curve is

$$\widehat{D}(p) = \hat{\beta}_0 + \hat{\beta}_1 p.$$

Using residuals to simulate demand

The residuals from the regression capture the randomness around this line:

$$\hat{\varepsilon}_i = D_i - \big(\hat{\beta}_0 + \hat{\beta}_1 p_i\big).$$

For any candidate price p, we generate a set of demand scenarios by plugging p into the regression and adding each residual:

$$D_i(p) = \hat{\beta}_0 + \hat{\beta}_1 p + \hat{\varepsilon}_i.$$

This treats the $\hat{\varepsilon}_i$ as sampled shocks and yields a realistic empirical distribution of demand at that price, which we then use in the newsvendor optimization.

Quadratic nature of price–demand interaction

The profit expression from the fixed-price newsvendor model still applies: revenue from selling demand, minus regular production, quick-printing, and disposal costs. However, once demand is written as a linear function of price, the revenue term becomes p times that linear expression. Substituting $D_i(p)$ into profit introduces terms proportional to p^2:

$$\pi_i(p, q) = pD_i(p) - cq - g(D_i(p) - q)^+ - t(q - D_i(p))^+,$$

so expected profit becomes a quadratic function of price and quantity. The constraints remain linear, but the quadratic objective means the problem is now a Quadratic Program (QP).

QCP vs. QP formulation strategy

Following the project description, profit can be rewritten as "Revenue − Cost" and a variable h_i is introduced capturing (negative) cost in each scenario:

$$h_i = -\text{Cost}_i(p, q).$$

Maximizing expected profit is equivalent to minimizing expected cost or, equivalently, maximizing the average of the z_i:

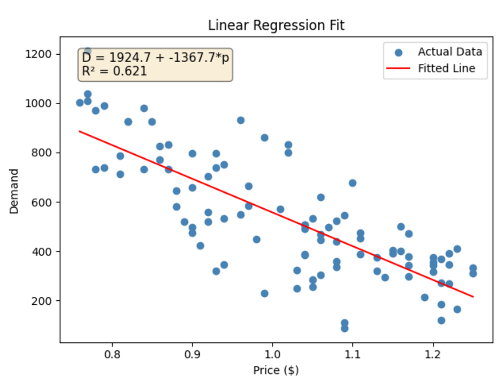
$$\max_{p,q,h} \quad \frac{1}{n} \sum_{i=1}^{n} h_i.$$

The constraints define h_i via linear inequalities that reflect shortage and surplus costs. Because only the objective contains the quadratic term in p and all constraints are linear, the resulting model is a QP (or equivalently a convex QCP with a quadratic objective and linear constraints), which can be solved in Gurobi.

5 Results

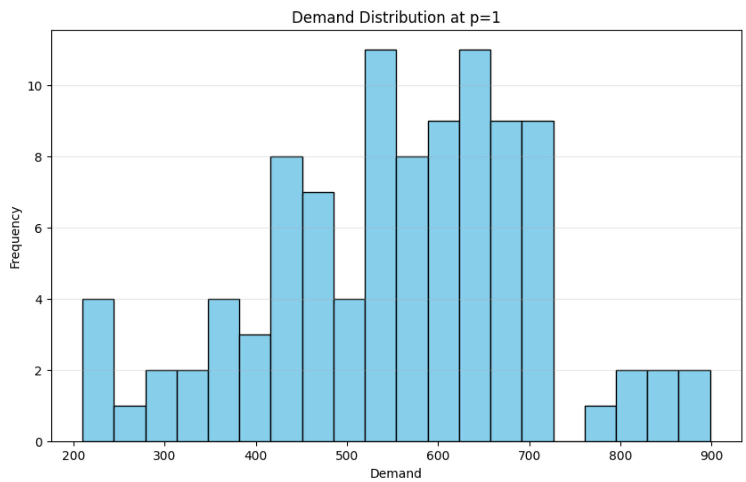
Part 1

The scatterplot shows the observed price–demand pairs along with the fitted linear regression line. The estimated demand function is $\hat{D} = 1924.7 - 1367.7 \cdot p$, which means there is a clear negative relationship between price and demand: as price increases, expected demand falls. For example, a 0.10 increase in price is associated with roughly 137 fewer units demanded on average. The R^2 value of 0.621 indicates that about 62% of the variation in demand across observations is explained by price alone, so the linear model captures a substantial, though not perfect, amount of the demand variability.



Part 2

This histogram shows the simulated demand distribution when the price is fixed at $p = 1$. Each bar counts how many scenarios fall into a given demand range. Most of the mass lies roughly between about 450 and 700 units, with relatively few scenarios below 300 or above 800, so extreme low or high demand is rare. The distribution is fairly spread out, indicating substantial uncertainty around the mean demand level (around the mid-500s to 600s). This distribution is what we use in the newsvendor model in Part 3 to evaluate how different order quantities perform under a wide range of possible demand realizations at this price.



Part 3

With rush and disposal costs included and price fixed at p = 1, this version of the LP model recommends producing about 629 newspapers per day ($q^* = 628.54$). This quantity balances the risk of ending up short and needing costly rush printing against the risk of printing too many copies and paying disposal costs on the leftovers, given the demand scenarios we use. Under this policy, the model predicts an expected average profit of about 178.99 dollars per day. Printing substantially more than this would increase expected disposal costs, while printing less would reduce sales and lead to more frequent shortages, lowering expected profit overall.

Part 4

Given the price-dependent demand model, the QP chooses both price p and quantity q to maximize expected profit. The solution we obtain is:

$$p^* = 0.95, \qquad q^* = 535, \qquad \text{Expected profit} \approx \$234.$$

The model recommends a slightly lower price than 1 (about 0.95). Because demand decreases with price, this small price reduction increases expected demand enough that total expected profit rises, even though the margin per unit is a bit smaller.

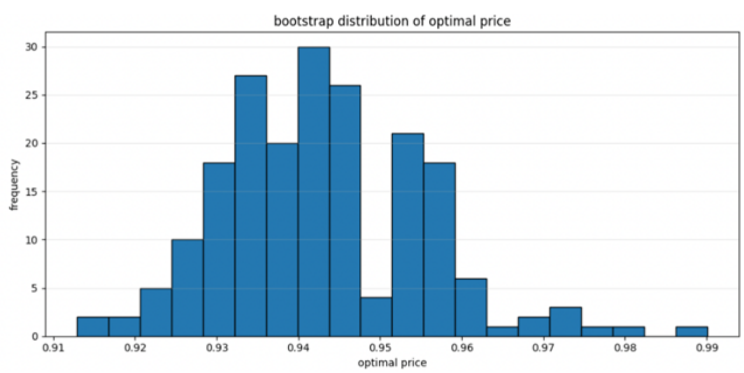
The optimal quantity $q^* = 535$ is the production level that best balances shortage vs. surplus risk at this new price: printing more would increase disposal risk; printing less would increase rush-printing costs.

The expected profit of about \$234 is higher than in the fixed-price LP from Part 3 (where p was fixed at 1 and profit was lower). This shows that jointly optimizing price and quantity creates additional value compared to only choosing the quantity at a given price.

Part 6 and 7

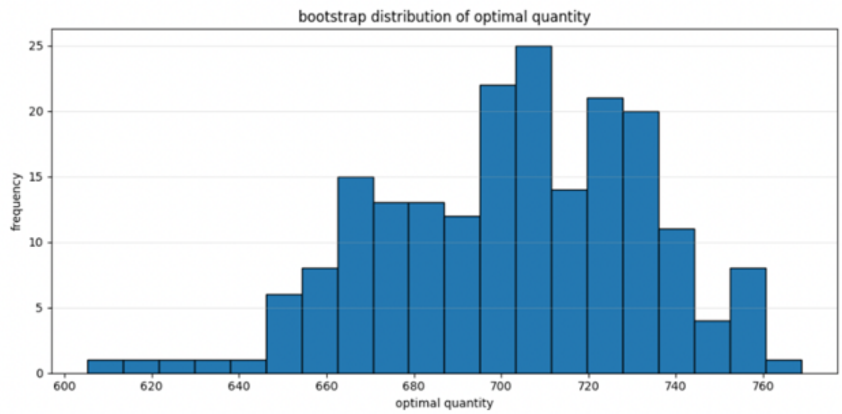
To understand how sensitive our optimal printing and pricing decisions are to sampling variation in the dataset, we implemented a nonparametric bootstrap. In each of 200 iterations, we resampled the original price–demand data with replacement, refit the linear regression, recomputed residuals, and solved the full QP for the optimal price and quantity. This procedure produces an empirical distribution of (p^*, q^*) that reflects how much our recommendations would change with new but similar data.

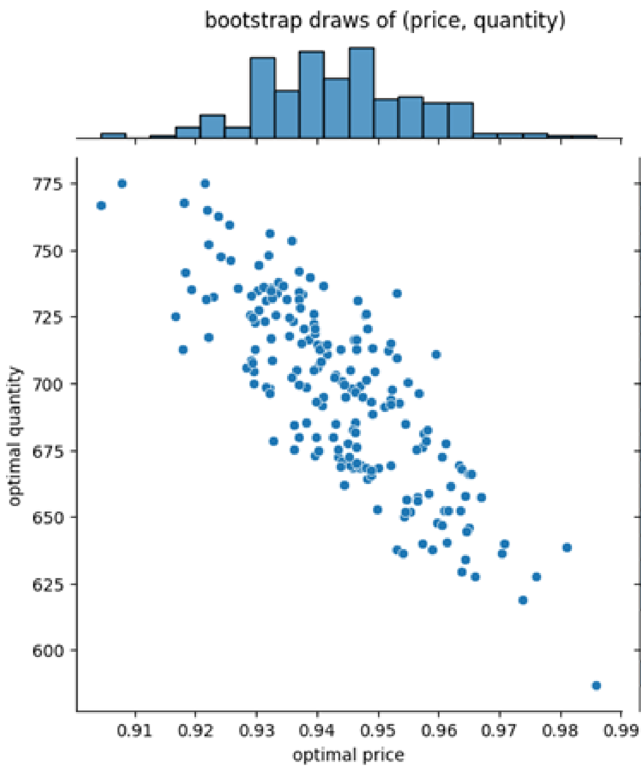
Out of 200 iterations, 198 solved successfully, indicating that the optimization is numerically stable. Across these replications, the optimal price is highly consistent: the distribution is centered at 0.9429 with a very small standard deviation (0.0124) and a 95% confidence interval of $[0.9219, 0.9714]$. The histogram shows a tight, unimodal cluster: almost all optimal prices fall between 0.93 and 0.96, which closely matches the original Gurobi solution of 0.9432. This suggests price recommendations are robust even when the dataset is perturbed.



Optimal quantities exhibit slightly more variability but remain stable overall. The mean bootstrap quantity is 702.29, very close to the Part 4 value of 704.18, with a standard deviation of 30.24 and a 95% confidence interval of $[646.13, 755.22]$. The spread is wider because quantity is more sensitive to small shifts in predicted demand, but the distribution is still centered tightly enough to give confidence in consistent production planning.

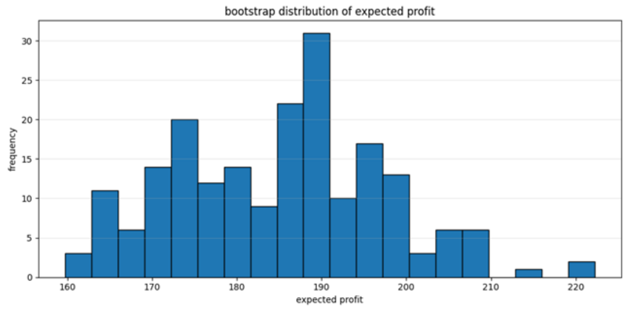
The joint scatterplot of price vs. quantity reveals a clear negative relationship: lower optimal prices are paired with higher quantities, and vice versa. This behavior aligns with economic intuition: lower prices lead to higher expected demand, making larger print runs optimal.





Finally, expected profits from the bootstrap replications average \$185.44 (sd = 13.62), compared to \$183.41 from the original optimization. The distribution remains well-behaved, with a 95% interval of [\$160.04, \$210.32]. This confirms that profitability estimates are also stable and that our model's recommendations do not hinge on any single unusual data point.

Overall, the bootstrap analysis shows that the extended newsvendor model produces reliable and resilient pricing and production decisions. The optimal price is especially robust, and the optimal quantity and expected profit vary only moderately across resampled datasets. The model is therefore appropriate for managerial use even with limited historical data.



6 Conclusion

Overall, the analysis demonstrates that the standard newsvendor model currently used by management is materially less effective than the extended price-dependent framework. The traditional model ignores how price influences demand, leading to a fixed price of 1 and a lower profit benchmark in Part 3. In contrast, the extended model jointly optimizes price and quantity, identifying an optimal price near \$0.95 and quantity around 535, which increases expected profit to approximately \$234 in Part 4. The bootstrap analysis further validates this improvement: across 200 resampled datasets, optimal prices remain tightly concentrated between \$0.93 and \$0.96, optimal quantities center around 700, and expected profits average \$185.44 with a stable 95% interval of \$160–\$210. These results show that the extended model not only captures rush and disposal costs more accurately but also aligns decisions with the observed negative price–demand relationship.

The main advantages of the extended model are its realism, ability to adapt pricing, and higher financial performance; its disadvantages are the added complexity, reliance on price-varying data, and need for an optimization solver. The standard model remains simple and easy to interpret but ultimately leaves money on the table. Switching to the extended framework would measurably improve production and pricing decisions for the company.