

# IEOR 151 – Service Operations Design and Analysis

## Homework 1 (Due 09/23/2016)

October 9, 2016

1. A newsboy would like to select the optimal level of inventory of newspapers using a newsvendor model without production costs. Here, demand  $X \sim \mathcal{U}(15, 60)$  is in units of newspapers, holding costs are \$0.9 per newspaper, and the sale price is \$1.5 per newspaper. What is the optimal inventory level? (4 points)

Solution:

$$F(\delta^*) = \frac{p}{p+q} = \frac{1.5}{1.5+0.9} = 0.625$$

Given that  $X$  is uniformly distributed, we solve  $\delta^*$  as follows:  $\frac{\delta^* - 15}{45} = 0.625$ . Hence,  $\delta^* = 44$  units

2. A real estate developer would like to select the optimal level of residential apartments to build in her new building using a newsvendor model with production costs. Here, demand  $X \sim \mathcal{N}(200, 50)$  is in apartment units, fixed costs are \$90,000, variable costs are \$750 per unit, holding costs are \$400 per apartment unit, and the (expected) sale price is \$1,200 per apartment unit. What is the optimal inventory level? (3 points)

Solution:

$$F(\delta^*) = \frac{p-c_v}{p+q} = \frac{1,200-750}{1,200+400} = 0.28125. \text{ Hence, } \delta^* = 200 - .58\sqrt{(50)} = 454 \text{ units}$$

3. Suppose you are the manager of American Giant and would like to determine the number of hoodies that should be made for the upcoming month. Assume the selling price is \$62.25 ( $r = 62.25$ ), the per unit production cost is \$57.75 ( $c_v = 57.75$ ) and the holding cost is \$1.5 ( $q = 1.5$ ).

- Assume the monthly demand is distributed as a Gaussian random variable with mean 500 and variance 1000. Use the newsvendor model with production cost to choose how many hoodies to be produced. (4 points)

Solution:

$$F(\delta^*) = \frac{r-c_v}{r+q} = \frac{62.25-57.75}{62.25+1.5} = 0.07059. \text{ Hence, } z = -1.47 \text{ and } \delta^* = 500 + (-1.47)\sqrt{1000} = 454 \text{ units}$$

- Suppose you have measured the demand of hoodies for the past 20 months, and the values of the demand, sorted into ascending order, are: 113, 197, 221, 234, 371, 375, 421, 469, 482, 495, 503, 559, 570, 611, 645, 672, 731, 757, 791, 834. Use the nonparametric newsvendor model to choose how many hoodies the company should produce. Explain your reasoning. (4 points)

Hint: Consider the sample distribution:

$$\hat{F}(z) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(z \leq X_i)$$

Solution:

Considering the empirical CDF, we adopt the following approach -  $n \frac{r-c_v}{r+q} = 20 \frac{2.25-0.75}{2.25+0.5} = 1.412$ . Then the solution corresponds to the  $X_{\lceil n(\frac{r-c_v}{r+q}) \rceil} = X_2$ . Hence, we produce  $X_2 = 197$  units

4. A portion of the following question is open-ended. Think thoroughly and provide good reasoning as we would like to understand your rationale. Suppose you are the engineer in charged of building a spam filter for Gmail. The null hypothesis is that the incoming email is "ham" (the opposite of spam).

Using Logistic Regression (a classification algorithm), you utilize the following decision rule to build the classifier:

$$\delta(X) = \begin{cases} d_0 & \text{if } g(X) \leq \gamma^* \\ d_1 & \text{if } g(X) > \gamma^* \end{cases}$$

, where  $g$  represents the model built via a training set.

- Propose 3 predictors/statistics to collect in order to build a spam classifier  
Potential Predictors include: portion of capitalized letters in the subject, indicator variable which denotes whether sender's email address is among "trusted" addresses, and the portion of special characters in the body of the email
- Propose a reasonable loss function. Please Justify your answer.  
A 0-1 loss or classification error.
- Propose a metric to evaluate the degree of success achieved by the decision rule.  
Number of false positives and false negatives.
- Bonus Question: Graphically present the performance of such a classifier (as the threshold  $\gamma^*$  is increased)  
Hint: Analysts are often interested in the tradeoff between false positive rate and true positive rate Consider an (ROC) curve.
- The product manager reviewed user statistics and found the number of type-I errors abnormally high. Propose the simplest method to correct the tendency.  
We can simply adjust the threshold  $\gamma^*$  to make it harder to reject emails as spam in order to reduce false positives.

5. Suppose you are managing the IT operation at a hotel chain, and you would like to determine if the company should purchase a new electronic system to manage the sales/booking process. Purchasing the system will cost \$4600, and other hotel chains have found that using the electronic system leads to an average of 23 mistakes (e.g. lost sales opportunities) per day. If the current error rate is 27 mistakes per day, then purchasing the new system will have a net savings of \$5200 over the span of one year. You have decided to use a minimax hypothesis testing approach to answer this questions. As a first step, you record the number of mistakes made over 20 days: 19, 23, 22, 37, 28, 18, 41, 36, 21, 27, 22, 42, 20, 33, 22, 10, 37, 12, 27, 33.

- Assume that the number of mistakes per day is approximated by a Gaussian random variable with variance  $\sigma^2 = 12$ . Using a binary search and  $z$ -table, compute the threshold for this hypothesis test  $\gamma^*$  to within an accuracy of  $\pm 0.1$  (4 points)  
Hint: Use the following values for the minimax hypothesis test:  $n = 20$ ,  $\mu_0 = 23$ ,  $\mu_1 = 27$ ,  $\sigma^2 = 12$ ,  $L(\mu_0, \mu_0) = 0$ ,  $L(\mu_0, d_1) = a = 4600$ ,  $L(\mu_1, d_0) = b = 5200$ ,  $L(\mu_1, d_1) = 0$

Solution:

Consider the following comparison and recall that the goal is the select  $\gamma^*$  such that

$$a(1 - \Phi(\frac{\sqrt{n}(\gamma^* - \mu_0)}{\sigma})) = b\Phi(\frac{\sqrt{n}(\gamma^* - \mu_1)}{\sigma})$$

$$5100(1 - \Phi(\frac{\sqrt{10}(\gamma^* - 19)}{\sqrt{32.5}})) = 4300\Phi(\frac{\sqrt{10}(\gamma^* - 27)}{\sqrt{32.5}})$$

Since  $a < b$ , binary search should be conducted on [23, 25] and the best first guess is 24  
Note, the required accuracy concerns gamma rather than the difference between LHS and RHS.  
Using binary search, we obtain  $\gamma^* \in [24.94, 25]$

Step	$\gamma$	LHS	RHS
1	24	453.1	.28
2	24.5	120.5	3.2
3	24.75	54.7	9.6
4	24.88	34.5	16.1
5	24.94	28.5	20.3476

- Should you purchase the new electronic system? Explain your answer (2 points)  
Given the decision rule

$$\delta(X) = \begin{cases} d_0 & \text{if } \bar{X} \leq \gamma^* \\ d_1 & \text{if } \bar{X} > \gamma^* \end{cases}$$

$\bar{X} = 26.5 > \gamma^*$ . The manager chooses  $d_1$  and buys the new system.