M.Eng in Operations Research and Information Engineering

Homework 1

Subject: Pricing Analytics and Revenue Management

Professor: Omar El Housni **oe46**

Authors: Federico Tiberga **ft232**

Due date: February 10th

Problem 1

Single fare class overbooking
In class, we have seen a problem of finding optimal overbooking amount for single fare class. Suppose the capacity is C seats, the price per seat is fixed at p, the number of no-show customers is a random variable $Y \sim cdf\, F(\cdot)$ and the number of overbooked tickets is q. Assume we are able to sell all $(C+q)$ seats for any q. The profit or the revenue function is as follows.

$$\Pi(q) = p \cdot (C+q) - \gamma \cdot E_Y[(q-Y)_+]$$

where $\gamma > p$, is the penalty per customer who can't get a seat. Suppose Y is a Poisson random variable with mean 5, $p = 100, \gamma = 300, C = 100$.

(a) Compute the optimal overbooking limit, q^* that maximizes the expected profit.

Solution:
Let's compute the overage and underage costs

$$c_u = p$$
$$c_o = \gamma - p$$

We can now compute the critical fractal

$$CF = \frac{c_u}{c_u + c_o} = \frac{p}{\gamma - p + p} = \frac{p}{\gamma} = \frac{1}{3}$$

We know that $F_Y(q^*) = CF = \frac{1}{3}$, therefore

$$q^* = F_Y^{-1}(\frac{1}{3}) \implies 3 \leq q^* \leq 4$$

Therefore the optimal overbooking limit is either 3 or 4. Since

$$\begin{aligned}\Pi(3) &= 100 \cdot (100+3) - 300 \cdot E_Y[(3-Y)_+] \\ &= 10300 - 300 \cdot \sum_{k=0}^{3} \left[(3-k)\frac{e^{-5}5^k}{k!}\right] \\ &= 10248.45\end{aligned}$$

$$\begin{aligned}\Pi(4) &= 100 \cdot (100+4) - 300 \cdot E_Y[(4-Y)_+] \\ &= 10400 - 300 \cdot \sum_{k=0}^{4} \left[(4-k)\frac{e^{-5}5^k}{k!}\right] \\ &= 10268.95\end{aligned}$$

(b) for q^* computed in part (a), find the expected penalty paid due to overbooking.

Solution:
We simply need to compute

$$300 \cdot E_Y[(4-Y)_+] = 300 \cdot \sum_{k=0}^{4} \left[(4-k)\frac{e^{-5}5^k}{k!}\right] = 131.05$$

Problem 2

US Tennis Association (USTA) is trying to devise a strategy to sell tickets for the Men's Quarterfinals for US Open 2016 in New York. The stadium has a seating capacity of 20000. Since the quarterfinals are held on a weekday, the demand is quite uncertain (assume all quarter-finals are held on a single day and there is a single ticket that grants admission for all four quarter-finals). USTA plans to offer a low-price ticket of \$50 with a restriction that USTA can repurchase it from the customer anytime at the price of \$75. USTA estimates that there is enough demand for the low-price tickets and any number will be sold. USTA also plans to offer a high-price ticket of \$100 that does not have any restrictions. However, demand for high-price tickets is uncertain and USTA estimates that it is uniformly distributed between 8000 and 17000. Assume that all the low-price demand arrives before the first high-price demand arrives and there is enough demand for low-price tickets in the initial period that any number of low-price tickets can be sold. However, no low-price tickets can be sold after high-price demand arrives.

1. What is the optimal number of tickets that USTA should offer at high-price initially to maximize the expected revenue from sales?

   Solution:
   First, let's compute the underage and overage costs. Since we can repurchase the tickets for \$75 and then sell them for \$100, the costs are as follows:

$$c_u = 100 - (100 - 75 + 50) = 25$$
$$c_o = 50$$

   We can now compute the critical fractal

$$CF = \frac{c_u}{c_u + c_o} = \frac{25}{25+50} = \frac{1}{3}$$

   The demand for high-price tickets is $D \sim U(8000, 17000)$. We know that $F_D(q^*) = CF = \frac{1}{3}$, therefore

$$q^* = F_D^{-1}(\frac{1}{3}) = 11000$$

   We should initially offer **11000 tickets**.

2. What is the expected revenue if USTA offers high-price tickets as computed in part (1)?

   Solution:
   We initially sell $(20000 - q)$ tickets for \$50, then we sell D tickets for \$100 dollars, and, if necessary, we pay \$75 dollars when $q < D$. Therefore, we can compute the expected revenue in the following way:

$$\begin{aligned}
\Pi(q) &= 50 \cdot (20000 - q) + 100\mathrm{E}[D] - 75 \cdot E_Y[(D-q)_+] \\
\Pi(11000) &= 50 \cdot 9000 + 100 \cdot 12500 - 75 \cdot E_Y[(D - 11000)_+] \\
&= 450000 + 1250000 - \int_{x=11000}^{17000} \frac{(x - 11000)}{9000}\, dx \\
&= 1700000 - 150000 = \mathbf{1550000}
\end{aligned}$$

Problem 3

A hotel with 400 rooms is trying to optimize the protection levels for business travelers. Leisure travelers typically book in advance and we will assume that there are enough of them to take the room at the "leisure" price of \$160 per night. (To make things simple, we assume that everyone stays in this hotel just for one night.) The hotel can block some rooms (not make them available to leisure travelers who book in advance) and make them available only closer to the actual date at the business-traveler price of \$200. We have one year of business demand data in the file *HotelProtectionLevelData.xls*.

1. Let us first treat all data the same (not adjusting the booking limit by days etc.). Suppose that the underlying distribution is normal with the empirical mean and empirical standard deviation. What is the protection level that you recommend for business travelers?

Solution:
Using the excel functions `AVERAGE` and `STDEV`, we find out that the business demand is distributed as a normal $\mathcal{N}(\mu = 299.34, \sigma = 34.46)$.
As always, we start by computing the underage and overage costs. Since we can repurchase the tickets for \$75 and then sell them for \$100, the costs are as follows:

$$\begin{aligned} c_u &= 200 - 160 = 40 \\ c_o &= 160 \\ CF &= \frac{c_u}{c_u + c_o} = \frac{40}{160+40} = \frac{1}{5} \end{aligned}$$

The demand for high-price tickets is $D \sim \mathcal{N}(\mu = 299.34, \sigma = 34.46)$. We know that $F_D(q^*) = CF = \frac{1}{5}$, therefore

$$q^* = F_D^{-1}(\frac{1}{5}) \implies 270 < q^* < 271$$

Therefore, the optimal protection level is either 270 or 271. Since

$$\begin{aligned} \Pi(q) &= 160 \cdot (400 - q) + 200 \cdot E[min(q, D)] \\ \Pi(270) &= 160 \cdot 130 + 200 \cdot E[min(270, D)] \\ &= 20800 + 200 \cdot \int_{x=0}^{\infty} min(270, x) \cdot f_D(x)\, dx \\ &= 20800 + 200 \cdot \left[\int_{x=0}^{270} x \cdot f_D(x)\, dx + \int_{x=270}^{\infty} 270 \cdot f_D(x)\, dx\right] \\ &= 20800 + 200 \cdot [49.48 + 216.64] \\ &= 74044 \end{aligned}$$

$$\begin{aligned} \Pi(271) &= 160 \cdot 129 + 200 \cdot E[min(271, D)] \\ &= 20640 + 200 \cdot \int_{x=0}^{\infty} min(271, x) \cdot f_D(x)\, dx \\ &= 20640 + 200 \cdot \left[\int_{x=0}^{271} x \cdot f_D(x)\, dx + \int_{x=271}^{\infty} 271 \cdot f_D(x)\, dx\right] \\ &= 20640 + 200 \cdot [51.69 + 215.33] \\ &= 74043.8 \end{aligned}$$

Thus, we should protect **270 rooms**

Problem 3 continued on next page...

2. What is the protection level for business travelers (still without distinguishing between days) but using the empirical distribution instead of assuming normality?

Solution:
First, let's set up the python file.

In [1]:
```
import pandas as pd
import numpy as np
from google.colab import files
uploaded = files.upload()
df = pd.read_excel(uploaded['HotelProtectionLevelData.xlsx'], usecols = ['
    Date', 'Demand', 'weekday', 'holiday name'])
```

Now, we create a function that, given the vector with all the demands, computes the probabilities of each demand.

In [2]:
```
def P_demand(v):
  d = {}
  m = min(v)
  M = max(v)
  for i in range(m, M+1):
    d[i] = 0
  for i in v:
    d[i] = v.count(i)
  sd = {}
  for key in sorted(d.keys()):
    sd[key] = d[key]/sum(d.values())
  return sd
```

The following function computes the cdf of the empirical distribution.

In [3]:
```
def F_emp(d):
  emp = {}
  summa = 0
  for k in d.keys():
    summa += d[k]
    emp[k] = min(summa,1)
  return emp
```

Now we can create a function that gives us the last value x such that $F(x) < CF$.

In [4]:
```
def find_CF(F, CF):
  m = min(F.keys())
  i = m
  while F[i] < CF:
    i+=1
  return i-1
```

Finally, we compute all the data we need

In [5]:
```
ED = P_demand(list(df['Demand']))
Femp = F_emp(ED)
CF = 0.2
find_CF(Femp, CF)
```

Out[5]:
```
267
```

Since $F(267) < 0.2$ and $F(268) > 0.2$, we know that the optimal protection level is either 267 or 268.

Problem 3 continued on next page. . .

We define a function that computes the profits for these two values to see which is actually the best option.

In [6]:
```python
pl = 160
pb = 200
C = 400
def compute_profit(q, d):
  leisure = pl*(C-q)
  summa = 0
  for i in d.keys():
    summa+=min(i, q)*d[i]
  business = pb*summa
  return leisure + business
```

In [7]:
```python
print(compute_profit(267, ED))
print(compute_profit(268, ED))
```

Out[7]:
```
74278.36065573765
74278.46994535517
```

Therefore, we should reserve **268 rooms**, but the difference in profit is not that relevant when we reserve 267 rooms.

3. Next, consider making different decisions for different days. What is your recommended protection level for business travelers for a non-holiday Monday? Use the empirical distribution for this question.

Solution:
We simply need to re-execute the same code with the filtered dataset

In [8]:
```python
df_monday = df[(df['weekday'] == 'Monday') & (df['holiday name'].isna())]
ED_monday = P_demand(list(df_monday['Demand']))
Femp_m = F_emp(ED_monday)
```

In [9]:
```python
idx = find_CF(Femp_m, CF)
print(idx)
print(compute_profit(idx, ED_monday))
print(compute_profit(idx+1, ED_monday))
```

Out[9]:
```
261
73948.33333333333
73950.83333333334
```

So we should reserve **262 rooms**, since the profit is greater.