Mathematics in Industry and Technology (MIT) Challenge 2018
The Jolly Baker - SOLUTIONS

27th June 2018

Task 1

To find the optimal value of Q, we simply calculate the profits that would have been made during 2017 if the quantity bought was Q instead of 80. This will give us a result that is at least as reliable as any other method.

Since there was an oversupply of packs in 2017, we can treat the sales data from that year as the demand. Let x_i equal the demand on day i (day i's sales from the data), and let q_i equal the packs sold on day i in our hypothetical. Then q_i is either the supply Q or the demand x_i, whichever is smaller:

$$q_i = \min(x_i, Q).$$

Since each pack costs c and sells for p, the profit P_i from day i is

$$P_i = pq_i - cQ,$$

and the profit P over the whole year is

$$P = \sum_{i=1}^{365} P_i.$$

Putting these together gives

$$P = p\sum_{i=1}^{365}(\min(x_i, Q)) - 365cQ. \tag{1}$$

A second way to approach this is by looking at the histogram of the data. Then we can consider separately the days with a demand less than Q and the days with a demand more than Q. We make a frequency table of the data in Table 1, and plot the histogram in Figure 1.

If the frequency of a demand x is f_x, then the number n of packs sold over the year is

$$n = \sum_{x<Q} xf_x + \sum_{x\geq Q} Qf_x,$$

and the profit over the year is

$$\begin{aligned} P &= pn - 365cQ \\ &= p\left(\sum_{x<Q} xf_x + Q\sum_{x\geq Q} f_x\right) - 365cQ. \end{aligned} \tag{2}$$

Demand	Frequency	Demand	Frequency
43	0	62	11
44	1	63	33
45	1	64	20
46	1	65	16
47	3	66	16
48	3	67	13
49	5	68	11
50	3	69	3
51	6	70	8
52	6	71	1
53	9	72	3
54	13	73	2
55	17	74	1
56	22	75	0
57	20	76	0
58	30	77	0
59	34	78	0
60	30	79	1
61	22	80	0

Table 1: Demand frequency over 2017

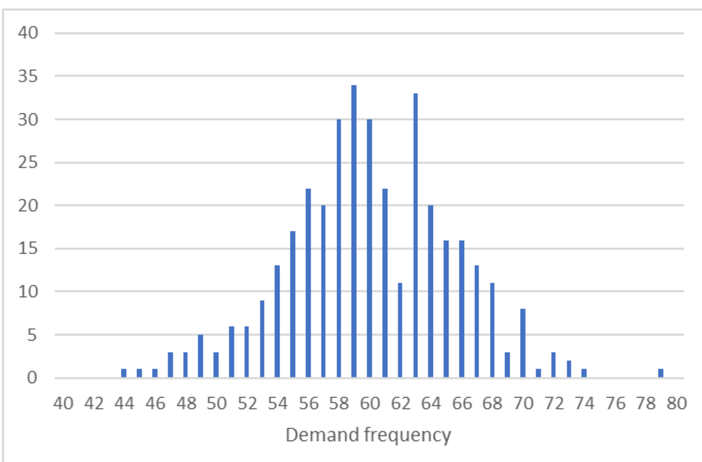

Figure 1: Histogram of demand frequency over 2017

This is the same result as (1) but leads to neater code.

Then we simply calculate P for different values of Q and choose the one that maximises P. It is clear that the optimal value of Q will lie between the smallest demand and the highest demand, so we only have to calculate the profits for the range $Q \in [\min(x_i), \max(x_i)]$. We give a subset of the values in Table 2.

Given the data we have, the maximum profit is \$69 827, acheived when $Q = 61$.

Q	P
...	...
55	67097
56	67930
57	68631
58	69212
59	69613
60	69810
61	69827
62	69712
63	69531
64	69152
...	...

Table 2: Profit P given quantity Q bought

Task 2a

Using $\bar{x}(p)$ to indicate the mean demand in a year where the price of a pack is p, we have

$$\bar{x}(5.8) = 64,$$
$$\bar{x}(6) = 60.044,$$

where to get $\bar{x}(6)$ we calculated the mean of the sales data from 2017.

To forecast the demand when $p = 6.2$, we are required to make a choice in extrapolating the data. As an approximation we assume that there is a linear relationship between the price and the mean demand, and using the above two values of $\bar{x}(p)$ we get

$$\bar{x}(p) = 178.724 - 19.780p,$$

which gives $\bar{x}(6.2) = 56.088$.

However we also need to make a choice about how the demand is distributed. There are two simple options, both of which involve manufacturing a new data set from the 2017 data set:

1. For each data point x_i in the year 2017, make a new data point y_i equal to

$$y_i = x_i \frac{\bar{x}(p)}{\bar{x}(6)}.$$

2. For each data point x_i in the year 2017, make a new data point y_i equal to

$$y_i = x_i - \bar{x}(6) + \bar{x}(p).$$

In both scenarios (by design), the mean of the y_i will be $\bar{x}(p)$ (since the mean of the x_i is $\bar{x}(6)$). Scenario 1 amounts to *scaling* the data according to the new mean, while scenario 2 amounts to *shifting* the data according to the new mean. This means that as the price goes up (pushing average demand down), there will be a wider range of demand in the shifted data scenario than in the scaled data scenario. Figure 2 shows the histogram of the 2018 forecast in both scenarios.

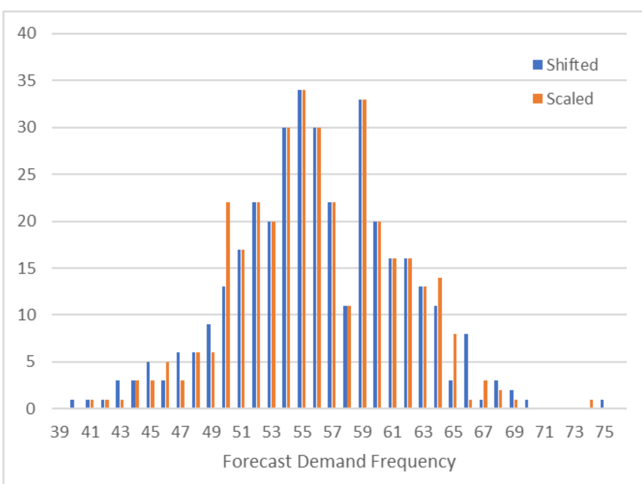

Figure 2: Histogram of forecasted demand frequency for 2018. Both shifted and scaled data scenarios are shown.

A more sophisticated method is to define the forecasted data as being linearly dependent on the old data,

$$y_i = ax_i + b,$$

then taking the mean and using the requirement that the average of the y_i be $\bar{x}(p)$,

$$\bar{x}(p) = a\,\bar{x}(6) + b,$$

we get

$$y_i = a(x_i - \bar{x}(6)) + \bar{x}(p),$$

which can be seen as a general case of both the shifted and scaled data scenarios. (The scaled data scenario is obtained by setting $a = \frac{\bar{x}(p)}{\bar{x}(6)}$, and the shifted data scenario by $a = 1$.) The parameter a can be used to choose a desired combination of shifting and scaling.

If more information is available an appropriate value of a can be found, however in our case a choice must be made.

Task 2b

For this task, we have to compare profits for different values of p and different values of Q, given a fixed value of $c = 2.6$.

Different values of p will affect the demand according to the rule we decided on in Task 2a. So we use our method from Task 2a to forecast demand data based on different values of p. Then we use our method from Task 1 to find the optimal value of Q for greatest profit, given the demand data we have generated. This can be coded up in a range of softwares, from Excel to Mathematica.

The results for different values of p are shown in Figure 3.

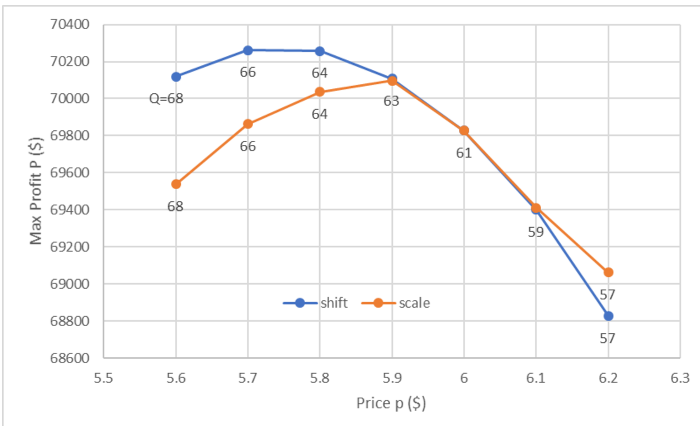

Figure 3: Maximum profit for given values of p. The value of Q that gives this profit is shown next to the data point. Both shifted and scaled data scenarios are shown.

We compare the optimal quantity and maximum profit in three cases — the current case ($p = 6$), the queried case where $p = 6.2$, and the optimal case — under the scaled data scenario:

	Current	Queried	Optimal
p	6	6.2	5.9
Q_{opt}	61	57	63
P_{max}	69 827	69 063.80	70 095.60

We do the same under the shifted data scenario:

	Current	Queried	Optimal
p	6	6.2	5.7
Q_{opt}	61	57	66
P_{max}	69 827	68 828.20	70 261.50

A combination of scaling and shifting would produce a result somewhere in between.

We conclude that setting a price of $p = 6.2$ produces a profit that is around \$1000 to \$1400 lower than the optimal choice of price, and around \$800 to \$1000 lower than the current price, which is a significant difference, so he should not increase his price to \$6.20.

Note that a simple choice of how to deal with the data means a difference in prices (\$5.70 to \$5.90) comparable with the difference from the actual price of \$6, and a similar statement can be made about the quantity Q. So our uncertainty is quite large. But what really matters is the profit: The best profit our calculation gives is only \$400 more than the maximum possible with the original set price (\$69 827), so the original price actually seems to have been a reasonably good choice. (Our consultant's fee is probably more than \$400 anyway!)

Task 3a

This task is similar to Task 2b, but with $c = 3$ instead of 2.6. Using the same method, we compare the maximum profit for different values of p. The results are shown in Figure 4.

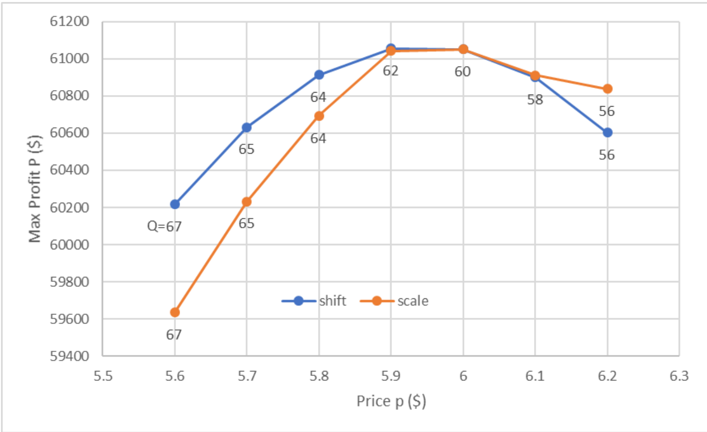

Figure 4: Maximum Profit for given values of p, when $c = 3$. The value of Q that gives this profit is shown next to the data point. Both shifted and scaled data scenarios are shown.

Under the scaled data scenario, we get for optimal price, quantity and profit:

$$p_{opt} = 6$$
$$Q_{opt} = 60$$
$$P_{max} = 61\ 050.00$$

Under the shifted data scenario, we get for optimal price, quantity and profit:

$$p_{opt} = 5.9$$
$$Q_{opt} = 62$$
$$P_{max} = 61\ 054.50$$

Task 3b

Using the same code again, we produce the optimal values of p and Q for supplier prices up to $c = 4$. The results are in Table 3.

We could also turn our excel spreadsheet into something user-friendly and simply provide it to JB.

	Shifted data			**Scaled data**		
c	p	Q	Max Profit	p	Q	Max Profit
2.5	5.9	63	72395.1	5.7	67	72675.7
2.6	5.9	63	70095.6	5.7	66	70261.5
2.7	5.9	62	67831.7	5.8	64	67921
2.8	5.9	62	65568.7	5.8	64	65585
2.9	5.9	62	63305.7	5.9	62	63317.5
3	6	60	61050	5.9	62	61054.5
3.1	6	60	58860	6	60	58860
3.2	6.2	56	56750.6	6	59	56692
3.3	6.2	55	54726.9	6.1	57	54587.9
3.4	6.3	53	52755.2	6.1	57	52507.4
3.5	6.3	53	50820.7	6.2	55	50476.3
3.6	6.4	51	48941.2	6.2	55	48468.8
3.7	6.4	51	47079.7	6.3	53	46510.7
3.8	6.6	47	45301.6	6.3	53	44576.2
3.9	6.6	47	43586.1	6.4	51	42691.1
4	6.6	47	41870.6	6.4	50	40849.6

Table 3: Optimal values of p and Q, and associated maximum profit P, for each value of c. Results for both shifted and scaled data scenarios.