Problem description

General information

Consider a large retailer that has 5 products in a specific small category. After several discussions with the suppliers, and meetings at the marketing department the firm decides to standardize prices among their different retail branches. Along the past year, each product has been sold at three different prices depending on the specific period and specific branch. Table 1 shows the three prices for each product. Table 2 shows the per unit cost for the retailer. The profit per product sold is the price at which the item is sold minus its cost.

Price level/Product	A	B	C	D	E
1	\$23	\$24	\$22	\$33	\$28
2	\$26	\$29	\$24	\$43	\$34
3	\$29	\$37	\$32	\$51	\$37

Table 1: Price per product $\{A, B, C, D, E\}$ and price level $\{1, 2, 3\}$.

Product	A	B	C	D	E
Wholesale price	\$10	\$11	\$9	\$14	\$12

Table 2: Costs per unit of product.

Historical data

Historical data of the monthly aggregate sales from different branches selling different assortments is available in the Excel file historical-data-2020.csv. In order to simplify this problem, we assume that the number of customers that arrive to each of the branches is the same. Without loss of generality will be assume that the total number of potential buyers per month is 1 (i.e. this is just a scaling). Thus, if 0.2 consumers bought product A at a price of 26, it means that 20 percent of the buyers chose that particular product at that price. Each entry of the file shows the assortment of products displayed (a 1 denotes that product was displayed and 0 otherwise). Next, follows the percentage of customers that bought each of the items displayed for that specific assortment. Note that the sum of the percentage of consumers that bought a product in the assortment is not necessarily 100%. This reflects the fact that some customers may decide not to purchase any of the products that were available (for example they may decide to buy from a competitor because the prices are too high for them). A

hypothetical example of such an entry is shown in Table 3. As there can't be more than one price for each product, observe that at each entry of the file, there is at most one price level per product that has a value 1 whereas all the other classes have a value of zero.

Consumer behavior

Each consumers that arrives to the store is interested in buying at most one product. Think for example as an arriving consumer who is interested in purchasing a mobile phone, a flight ticket, a soft drink, or a snack. The demand of any product depends on the complete set of offered products since customers potentially make substitutions to an available product if their most preferred product is not available.

Product	A1	A2	A3	B1	B2	B3	C1	C2	C3	D1	D2	D3	E1	E2	E3
Assortment	1	0	0	1	0	0	0	1	0	1	0	0	0	0	0
Percentage	0.45	0	0	0.13	0	0	0	0.045	0	0.36	0	0	0	0	0

Table 3: Hypothetical example of historical data. For assortment $\{A1, B1, C2, D1\}$ of products displayed, the percentage of customers that bought each product displayed is 45%, 13%, 4.5% and 36% respectively.

Such substitution behavior is captured by a customer choice model that can be thought of as a heterogeneous distribution over preference lists (or permutations of products). We explain: Consumers can be partitioned into customer types. Each customer type has associated a particular preference list and will purchase the most preferable product that is available. If none of her preferable products are available, the consumers of that specific type will buy nothing. For illustration, consider a potential consumer named Cecilia who belongs to a consumer type that is associated to the following preference list: [D1, A1, C1, D2, A2, B1, B2, C2]. For Cecilia, as for any other consumer that belongs to this customer type, her most preferred product is D1, if that product is not available he will buy A1 if available, otherwise if product A1 is not available she will buy the first available product in her preference list. If none of her preferred products are available, that is if the products displayed are contained in {A3, B3, C3, D3, E1, E2, E3} , Cecilia will leave the store without buying any product. Notice that Cecilia is rational and therefore, faced at the decision of buying the same product at different prices, she will prefer the cheapest price. For example, for product A, Cecilia prefers A1 over A2 and the same for all products and their price points. For simplification, we assume all consumers are rational and that the distribution of consumer preferences is the same among all branches. A market research study has found that there are between 40 to 50 different customer types in this market. Moreover, 20 percent of the customer types have a preference list that has a length of at most 3.

Methodology and result

Why Mixed Multinomial Logit model?

This revenue management task is aiming to answer two questions, 1) which subset of products should we offer. 2) at what price should we sell. The key of solving this assortment problem is to predict the choices of individuals, who select a subset of the product that the company offered. Considering these facts from the task and the regularity axiom:

1. There are five products with three price levels with each product. This issue can be treated as 15 different products and a set of utilities can be assigned to each product. Also, notice that the customers in the system are rational. In short, A multinomial logit model is a foundation to describe the problem.
2. There are around 40 to 50 different customer types in the market. Hence, instead of a single set of utilities, several sets are needed so the different types of customer can be explained by the model. Each set of utility will contain a correspondent probability, which means the mixed multinomial logit model is the choice and each segment represents one type of customer.
3. According to the problem, the number of customers that arrive to each of the branches is the same, and the Markov chain choice Model is not suitable under this setting with only 150 data points.

The equation of probability that the customer chooses to buy a product can be illustrated below:

$$P(x,S) = \mathbb{E}[\frac{v_i}{1+\sum_{j\epsilon s} v_j}]$$

Giving a subset of products S, the probability P that an individual choose product x ∈ S can be calculated based on the vector Vi = (v1, v2…. vN) when there are N products and i ∈ M (random vector for realisation)

How can we do it?

The objective in this section is to use the historical data to build a Mixed Multinomial Logit model and test it through the R square metric. To address the overfitting and underfitting issue, the data will be divided into two sections, the training and the testing parts in a ratio of 130:20. The excel non- linear solver will be applied to measure the global minimum of the model in different realisation settings respectively.

The data set only contains 150 assortments and the probability that the customers would like to buy certain product under this assortment. Hence in reality it is hard to measure all heterogeneous distribution over the customer types. The number of M can be a hyperparameter that turned by testing set. To address the rational of the customer, there are three price levels for each product, say X1<X2<X3. There will be a constrain that the utility in one product meet V1>V2>V3 for every realisation. Final question is which Excel solver method will be used. There are two non-linear solvers, the GRG and the Evolutionary, GRG uses generalised reduced gradient, Evolutionary method works for non-smooth situation. The

starting points are critical for both methods. In this section, the two methods will be applied together to find the global minimum. R square between $P(x, S)$ and the real purchase probabilities will be presented to make a judgement on the model as well as all hyperparameters.

The model should aim to represent 20 presents of the customer type have a preference list that has a length of at most 3, which means 1) the utility variable should be allowed to zero. 2)The ideal scenario would be around 20% of the set of utilities contain at most three utilities.

The other constraints are all decision variables non-negative, the upper limit of utilities equal to 100 and the sum of fraction of realisations equal to 1.

Result and assortment optimisation

Number of zero P means how many the realisation percentage equal to zero, indicating this realisation does not exist under this model.

Number of Realisation	3	5	7	8	9
R^2 (train)	0.12237	0.09917	0.08624	0.07416	0.06465
R^2 (test)	0.02198	0.01656	0.01201	0.01015	0.01067
Number of zero P	0	0	0	0	0

From the table, three realisations are too simple, the model cannot capture the heterogeneity. On the other hands, 9 realisations seem over complex due to the size of the data been used to train the model, it contains lower training error but higher test error, indicating for overfitting. Therefore, considering the R square error in training and testing data set respectively, the optimal number of realisations is 8.

Noticed that the 20% of the customers have a preference list that has a length of at most 3 is hard to address, the solution would be a bigger data set so the number of realisation is bigger enough to allow the utility and fraction convergence to zero. Or we can use Markov model with proper assumption to describe the customer behaviour.

The final model from full data set can be found in Excel sheet, the prediction of unseen assortment can be found in Excel sheet as well.

The objective here is to maximise the potential revenue based on the utilities and realisations' distribution calculated before. Notice that in each product, only one price level will be used to build the assortment. This would work as a constraint during the optimisation.

There are mainly two ways for assortment management:

- Exact calculation (compute all possible combinations, then pick one with max profit)
- Heuristic method (Greedy or ordered approach)

Considering the number of possible combinations of the assortment, Exact calculation is more accuracy and feasible.

Objective:

$$\max\left(\sum P * profit\right)$$

when P is the probability of the customer purchase the selected product in a typical assortment.
The table illustrates the result.

Product	A1	A2	A3	B1	B2	B3	C1	C2	C3	D1	D2	D3	E1	E2	E3
Profit	13	16	19	13	18	26	13	15	23	19	29	37	16	22	25
Assortment selection	0	0	1	1	0	0	0	0	1	1	0	0	0	1	0
Probability of purchase	0	0	0.001675	0.357746	0	0	0	0	0.001263	0.507792	0	0	0	0.087362	0
Total profit	16.28159														