Journal of Economics and Public Finance
ISSN 2377-1038 (Print) ISSN 2377-1046 (Online)
Vol. 6, No. 1, 2020
www.scholink.org/ojs/index.php/jepf

Original Paper

Demand Forecast in Retail Assortment Optimization—Based on an Empirical Analysis of Beverage Sales

Jun Chen[1], Xinyijing Zhang[1*] & Chenyang Zhao[1]

[1] Economic and Finance Department, SILC, Shanghai University, Shanghai, China

[*] Xinyijing Zhang, Economic and Finance Department, SILC, Shanghai University, Shanghai, China

Received: January 22, 2020 Accepted: February 4, 2020 Online Published: February 11, 2020

doi:10.22158/jepf.v6n1p57 URL: http://dx.doi.org/10.22158/jepf.v6n1p57

Abstract

This paper focus on establishing the demand forecasting model to optimize product assortments from a set of SKUs in the same category. The aim of the model is to achieve revenue maximization. Based on the attribute level, the demand model considers the consumers' preference and the possibility of substitution between different attributes. Then it divides the product's specific attributes and multiplies these attributes effects. Furthermore, one beverage case was applied to the demand model to do empirical analysis. Top beverage categories were selected and e-commerce sales data were collected to represent the pre-sale of whole categories. Moreover, a store named S with some beverage SKUs is assumed and applied to the model, which predicted sales volume of each existing SKU and the total revenue.

Keywords

Assortment Optimization, Demand Estimation

1. Introduction

In the planning of product assortment, retailers need to consider the demand estimation of each stock-keeping unit (SKU) in a category and the price to get the expected revenue. The goal of retail assortment optimization is to maximize the revenue, which is the profit or the net profit after minus the cost of the store (Robert, Harald, & Tammo, 2013). Many retailers are exploring how to improve assortment localization for stores or similar store groups. Zimmerman (2006), O'Connell (2008) and McGregor (2008) studied Wal Mart, Macy's and Best Buy respectively, they tried to lead a category change of each store according to the taste of local consumers, and the result suggested that all stores' revenue has been increased (Zimmerman, 2006; O'Connell, 2008; McGregor, 2008). Therefore, regional factors should be considered when designing product assortment for one store or similar store groups, and the structure of each existing category should be adjusted according to the local taste to tap the unmet demand. Moreover, the concept of geographical space is added to the local taste to consider the needs of regional consumers served by stores in a particular location. Assuming that the consumer's choice of goods is based on their preference for attributes of goods, thus, the local taste can be viewed as a preferred attribute of local consumers (Robert, Harald, & Tammo, 2013). At the same time, the possibility of substitution between similar goods does exist. Consumers' best preference for different attributes of goods generates the first choice of goods. It is assumed that consumers will buy the existing similar goods when their ideal goods are not available (Honhon, Gaur, & Seshadri, 2010).

This article established a demand forecast model to help retailers select products efficiently and optimize the existing assortment. It maximizes the diversity of categories under the constraint of limited shelf space and achieves profit maximization. The pre-sale of the whole categories is used to estimate the market share of attributes and consumer preference, then the demand for every single product is predicted. Meanwhile, taking into account the probability of choosing similar products without the first choice, the possibility of substitution between two single products is predicted. Finally, this literature also applies the demand forecast model into beverage sales, and found several difficulties in the actual application of the model.

2. Literature Review

When retailers optimize the product assortment, they need to consider many factors comprehensively. The basic factors include price, demand, shelf-space restriction, commodity characteristics, and dynamic factors include new product introduction, commodity similarity, sales promotion, and commodity combination. Generally, there are four challenges in assortment optimization.

2.1 New Product Demand Forecast

The introduction of new products will have an impact on the demand for existing product groups. This impact includes the negative effect on similar competitive products and the positive effect on complimentary products, which increases the instability of the existing product demand.

At the same time, from the perspective of the new product itself, its demand estimation also has high uncertainty. After classifying the attributes of a new product, the proportion of new products in the total demand for similar products in the same category could be foretasted. However, the consumers' acceptance of new products, promotional activities and the early marketing effect will cause the inaccuracy of new product demand prediction (Deza, Huang, & Metel, 2015).

2.2 Determination of Substitution Possibility

Although many workers such as Parlar and Goyal (1984), Netessine and Rudi (2003) have studied the static substitution model, there are relatively little pieces of literature on dynamic substitution models. It is difficult to determine the possibility of substitution between products. Even if there are only two products in the substitution process, the change of substitution position will lead to a change in the possibility. Furthermore, the substitution possibility of different attributes of products is different, consequently, we should pay attention to the consumer preference as far as possible in the attribute level division. Meanwhile, the product attributes have an interactive relationship and joint attributes may come into existence, so it is harder to determine the substitution possibility between attributes. What's more, different types of consumers have different substitution possibilities when choosing the same similar product without the first choice (Dong & Tian, 2009). To get accurate substitution possibilities, consumer segmentation should be provided.

2.3 Promotional Effect on Revenue Caused by Category Localization

Continuous adjustment and planning of categories for local consumers not only cause higher operating costs but also request for an upgraded store information software. To meet the new demands of consumers, the overall operational risk will increase due to the cost of new product development and procurement. For most retailers, cost control is the priority, they need to evaluate the ratio of input and future output to decide whether or not to implement category localization. Alptekinoğlu and Grasas (2014) found that the optimal assortment always follows strict return policies that balance the risk and return (Aydın & Alex, 2014). At present, although many retailers have tried it, the degree of category localization still depends on the relationship between cost and revenue.

2.4 Joint Consideration of Category Optimization and Pricing

The price will affect consumer's preference and demand, nevertheless, the ultimate goal of category optimization is to maximize revenue. Since revenue is multiplied by sales volume and price, category optimization and pricing need to be considered simultaneously. It is worthwhile to mention that category planning and price will affect each other. For revenue maximization, category selection, demand, and price should be incorporated at the same time. Hopp and Xu (2008), Aydin and Heese (2014), Federgruen and Hu (2015) solved this problem in an aggregate method of the whole market where price and category are both optimized. However, as possible results will increase exponentially after considering the combination of the three factors, it is necessary to take measures to reduce the large computing load.

3. Demand Estimation Model

The aim of the model is maximizing revenue through optimizing retailassortment. Meanwhile, it allows a constraint on the number of categories because of the limited shelf space in stores.

We assume that SKUs in the same category may have different attributes and each attribute can be classified into several levels. Thus, for every product that contains A attributes, define$d$ one particular attribute, so $a \in \{1,2,\cdots,A\}$. And a single attribute contains N_a levels, define u as one particular level, so $u \in \{1,2,\cdots,N_a\}$.

One significant hypothesis in this model is that consumers prefer typical attribute levels before the selection process, regardless of the environmental effect, such as location and sales promotions. Under this hypothesis, we conduct a pre-sale for all categories and calculate the proportion of customer preference to each SKU. Meanwhile, after separating SKUs that in the same category by different attribute levels, we could get the proportion of customer preference to each attribute level.

For store S, we define customer preference to one particular attribute level as f_{au}^{s}, in order to signify the category of SKUi, use $f_{ai_a}^{s}$ to constitute for f_{au}^{s}, where i_a stands for$u$.

Finally, we consider the possibility of substitution, which means customers may choose other similar SKUs when their first choice is not available when he visits the store. Assume that π_{auv}^{s} is the probability of substituting attribute level u for v, π_{ij}^{s} is the possibility of substituting product j for i.

There are n possible SKUs in one particular category, the price of SKU is p_i. Define D^s as the demand for each product category in store S and D_i^s as the demand for SKU_i. In the case of full category pre-sale, the preference ratio of consumers for different attribute level f_{au}^{s} is obtained.

For a selected SKUi in a specific category, the preference probability of consumers for the selected SKUi is obtained by the algorithm of multiplication, that is,

$$f_i^s = \prod_{a=1}^{a=A} f_{ai_a}^{s} \tag{1}$$

Then, when the store S does not sell product i, we calculate the possibility of using product j to replace product i. According to the attribute levels, we use the arithmetic of multiplication to get the equation:

$$\pi_{ij}^{s} = \prod_{a=q}^{a=A} \pi_{ai_a j_a}^{s} \tag{2}$$

The purchase probability of a SKU includes two parts: one is the consumer's preference when the store S has product j, the other is the sum of all other similar single product substitution possibilities existing in store S when store S does not have product j. Therefore, the purchase possibility equation of product j is:

$$F_j(S) = f_j + \sum_{i \notin s} f_i \, \pi_{ij} \tag{3}$$

Add up the purchase probability of all products in the same category in the store S, and the equation is:

$$F(S) = \sum_{j \in s} F_j \,(S) \tag{4}$$

Assume x is the sales volume of each existing product in store S, and x_j is the sales volume of SKUj, $j \in s$. Divide the sales volume of each SKU by the sum of the purchase possibility $F(S)$, the demand of products in a certain category is estimated, that is

$$D = \sum_{j \in s} \frac{x_j}{F(S)} \tag{5}$$

Then the demand of SKUj in store S is obtained:

$$D_j^s = f_j^s D \tag{6}$$

The revenue of store S is calculated as follows:

$$R_s = \left(\sum_{i \in s} p_i \, D_i^s + \sum_{i \notin s} p_i \, D_i^s \pi_{ij}^s \right) \tag{7}$$

The first part refers to the income earned from customers whose most preferred product is offered in the store and the second part is the substitution income from customers whose most preferred product was not in the store.

4. Empirical Analysis

4.1 General Description

We selected two best-selling beverage brands, Masterkong and Uni-president, and collected sales data of each beverage item of two brands in the Tmall supermarket, which is one of the biggest e-commerce platforms in China. These sales data were assumed as the data obtained from the pre-sale of the whole category. Based on the typical attributes of beverage, three key attributes can be divided: brand, taste, and package. The brand attribute contains two levels: Masterkong and Uni-president. There are 14 kinds of beverage tastes, including 9 fruit tastes and 5 tea tastes. The package attribute also includes two types: large and small. The price of the large package is more expensive and the price of the small package is normal. In the following part, we use High Price and Normal Price to express these two terms respectively.

In the case of whole category sales, we obtained the preference ratio of consumers for different taste levels. Then these data were divided into two beverage types to make a comparison. In the fruit beverage, the proportion of big packages is about 28%, while this figure is about 71% in a small package. However, this gap is relatively small in the tea beverage, which is about 54% and 46% respectively. Uni-president brand has the brand advantage in fruit beverage and accounts for 60% approximately, Nevertheless, in tea beverage, Masterkong has the brand advantage, and constitutes for 58% approximately. At the same time, the following table is obtained:

Table 1. Best Seller Price Comparison in Beverages

	Normal Price		High Price	
	Masterkong	Uni-president	Masterkong	Uni-president
Fruit beverage	26.18%	44.83%	13.93%	15.05%
Tea beverage	29.45%	16.36%	28.68%	25.51%

Since different attributes are not necessarily independent of each other, there is likely a certain interactive relationship between two attributes. Therefore, considering joint attributes in the establishment of attributes is necessary. In this case, the brand and package are combined to get a new joint attribute called brand-package and the consumer preference ratios are in the table above. Now the three attributes are reduced to two: taste and brand-package, resulting in $14\times4 = 56$ different possible SKUs. The preference ratio of fruit and tea beverage are calculated respectively because of the unique characteristics of the two types.

Now the substitution possibility between attributes is considered. Since different consumers have different preferences for tastes, we assumed that the substitution possibility is zero in each taste. Thus, substitution possibility exists in brand-package merely. Furthermore, consumers can’t substitute a big package for small package and vice versa. Consequently, we can only consider the substitution possibility between different brands with the same package and the possibility ration is stipulated as follows:

Table 2. Substitution Possibility between Different Brands with Same Package

Fruit beverage				
	Masterkong-Normal Price	Uni-president-Normal Price	Masterkong-High Price	Uni-president-High Price
Masterkong-Normal Price		75%	0	0
Uni-president-Normal Price	3%		0	0
Masterkong-High Price	0	0		65%
Uni-president-High Price	0	0	33%	
Tea beverage				
	Masterkong-Normal Price	Uni-president-Normal Price	Masterkong-High Price	Uni-president-High Price
Masterkong-Normal Price		13%	0	0
Uni-president-Normal Price	20%		0	0
Masterkong-High Price	0	0		38%
Uni-president-High Price	0	0	22%	

Note. Horizontal attributes substitute for vertical attributes.

4.2 Demand and Revenue Forecast

In this section, a store that offers 36 beverage SKUs is assumed. And the total 56 possible SKUs can be divided into existing and unsold items of store. At first, use Equation (3) and combine with the preference ratio of each attribute level and substitution possibility, the purchase probability of all existing SKUs can be obtained. By adding up these data, the total purchase probability is 87.90%. Secondly, add the sales volume of each SKU in store S to get total demand and divide it by the total purchase probability. The total estimated demand is 2564. Then the estimated demand for each existing SKU in store S can be calculated using Equation (6). At last, the revenue of this store is estimated by Equation (7), which is 11176 yuan. The detailed data and calculation process are shown in Figure 1.

5. Model Limitations

5.1 Seasonality and Promotion Factors

This model assumed that consumers' preference to specific attribute level is constant regardless of the environmental effect. This hypothesis may violate the actual situation. Many factors could affect this figure, and typical examples are seasonality factor and promotion factor. Sales of products like beverage, cloth, refrigerator and air conditioner fluctuate with the season, which means that consumers' preference fluctuate with the season. Furthermore, the purchase intention and preference could be changed by promotional mechanisms. For instance, many retailers will implement a clearance price that is far lower than normal situation. Caused by this motivation, more consumers prefer the product with lower price, leading to inaccurate estimation results of the model.

5.2 Repeated Arithmetic Operation of an Optimal Decision

On the basis of the existing products, this model can quickly determine whether a particular item or combination of items should be added, subtracted or replaced. However, each time when making a decision, the total revenue of the store after the decision should be recalculated and compared with the previous revenue. The total number of possible SKUs in any category is very large, and the combination of individual items is more likely to be even huge. Therefore, it takes a lot of calculation to find the category decision that can achieve the maximum benefit after making decisions, the model still needs to be optimized. At the same time, the model can estimate the demand of each existing item, which is very meaningful to the operation of the retailer.

5.3 Model Difficulties

Consumers may not be able to form a stable attribute preference for any category, even if they are very familiar with the category. There are two possibilities for the change of consumer feature preferences. Firstly, the attribute preference will change with time, but there will be many short stable periods of feature preference. Secondly, consumers will return to their stable attribute preferences after changing their preferences at some time. Different types of consumers for different types of goods for the changes in the characteristics of preferences are different. Consequently, we need to identify and describe the time period of change. At present, the model can only be used to predict the final demand. In order to improve the accuracy of the prediction, we still need to introduce the time variable and find consumer preferences for product features that vary along the timeline. Moreover, since the purchase decision is random, Bayesian random wave distribution can be used. At the same time, consumers have different possibilities of substitution, so we need to classify the consumers and realize that the same type of consumers has the same possibility of substitution.

6. Conclusion

This paper establishes a commodity feature-based demand forecasting model with the goal of revenue maximization, with the final output of the demand driven forecast of the individual SKUs. In this paper, beverage sales data are taken to apply for the practical model both for on line and off line sales. In the empirical analysis, the interrelation between features have taken into considerations. Using the demand forecast model, the predicted demand is obtained and the revenue of the retail store is obtained. However, the forecast can not be 100% complete accurate due to the limitation of the model setup. Under most circumstances, the forecast of aggregated items is more accurate than individual items, aggregate estimation results are more reliable for decision-making. Therefore the beverages manufactures should focus on the differentiation of semi-finished products at the push-pull boundary and achieve production postponement. By cooperating these strategies with the product assortment selection, the competitiveness of products will be enhanced.

	Taste	Brand	Package	Preference ratio	Substitution possibility	Purchase possibility	Purchase possibility(Adjusted)	Sales volume	Sales(Sales volume/100)	Predicted demand
1	Snow pear	Masterkong	Normal Price	5.78%	3.00%	5.78%	6.58%	6218	62	169
2	Orange	Masterkong	Normal Price	1.07%	3.00%	1.07%	1.22%	1804	18	31
3	Mango	Masterkong	Normal Price	0.74%	3.00%					
4	Lenmon	Masterkong	Normal Price	0.41%	3.00%					
5	Grape	Masterkong	Normal Price	1.20%	3.00%	1.20%	1.37%	2028	20	35
6	Honey peach	Masterkong	Normal Price	0.77%	3.00%	0.77%	0.88%	1881	19	22
7	Plum	Masterkong	Normal Price	1.39%	3.00%	1.39%	1.58%	2347	23	41
8	Wild jujube	Masterkong	Normal Price	0.09%	3.00%					
9	Grapefruit	Masterkong	Normal Price	1.64%	3.00%	1.64%	1.87%	2766	28	48
10	Snow pear	Uni-president	Normal Price	9.90%	75.00%	9.90%	11.26%	16652	167	289
11	Orange	Uni-president	Normal Price	1.83%	75.00%	1.83%	2.08%	2251	23	53
12	Mango	Uni-president	Normal Price	1.26%	75.00%	1.81%	2.06%	1551	16	53
13	Lenmon	Uni-president	Normal Price	0.71%	75.00%	1.02%	1.16%	872	9	30
14	Grape	Uni-president	Normal Price	2.06%	75.00%	2.06%	2.34%	2530	25	60
15	Honey peach	Uni-president	Normal Price	1.32%	75.00%	1.32%	1.50%	2347	23	39
16	Plum	Uni-president	Normal Price	2.38%	75.00%	2.38%	2.71%	2929	29	69
17	Wild jujube	Uni-president	Normal Price	0.16%	75.00%					
18	Grapefruit	Uni-president	Normal Price	2.81%	75.00%	2.81%	3.20%	3452	35	82
19	Snow pear	Masterkong	High Price	3.08%	33.00%	3.08%	3.50%	4992	50	90
20	Orange	Masterkong	High Price	0.57%	33.00%					
21	Mango	Masterkong	High Price	0.39%	33.00%					
22	Lenmon	Masterkong	High Price	0.22%	33.00%					
23	Grape	Masterkong	High Price	0.64%	33.00%					
24	Honey peach	Masterkong	High Price	0.41%	33.00%					
25	Plum	Masterkong	High Price	0.74%	33.00%	0.74%	0.84%	1320	13	22
26	Wild jujube	Masterkong	High Price	0.05%	33.00%					
27	Grapefruit	Masterkong	High Price	0.87%	33.00%	1.18%	1.34%	528	5	34
28	Snow pear	Uni-president	High Price	3.32%	65.00%	3.32%	3.78%	4427	44	97
29	Orange	Uni-president	High Price	0.61%	65.00%	0.98%	1.11%	899	9	29
30	Mango	Uni-president	High Price	0.42%	65.00%					
31	Lenmon	Uni-president	High Price	0.24%	65.00%					
32	Grape	Uni-president	High Price	0.69%	65.00%	1.11%	1.26%	1011	10	32
33	Honey peach	Uni-president	High Price	0.44%	65.00%					
34	Plum	Uni-president	High Price	0.80%	65.00%	0.80%	0.91%	1170	12	23
35	Wild jujube	Uni-president	High Price	0.05%	65.00%					
36	Grapefruit	Uni-president	High Price	0.94%	65.00%					
37	Black tea	Masterkong	Normal Price	2.98%	20.00%	2.98%	3.39%	13474	135	87
38	Green tea	Masterkong	Normal Price	2.35%	20.00%					
39	Jasmine tea	Masterkong	Normal Price	4.90%	20.00%					
40	Oolong tea	Masterkong	Normal Price	0.17%	20.00%					
41	Milk tea	Masterkong	Normal Price	4.32%	20.00%	4.32%	4.91%	9110	91	126
42	Black tea	Uni-president	Normal Price	1.66%	13.00%	1.66%	1.89%	8743	87	48
43	Green tea	Uni-president	Normal Price	1.31%	13.00%	1.61%	1.83%	6787	68	47
44	Jasmine tea	Uni-president	Normal Price	2.72%	13.00%	3.36%	3.82%	8689	87	98
45	Oolong tea	Uni-president	Normal Price	0.09%	13.00%					
46	Milk tea	Uni-president	Normal Price	2.40%	13.00%	2.40%	2.73%	9364	94	70
47	Black tea	Masterkong	High Price	2.90%	22.00%	2.90%	3.30%	7699	77	85
48	Green tea	Masterkong	High Price	2.29%	22.00%	2.74%	3.12%	8963	90	80
49	Jasmine tea	Masterkong	High Price	4.77%	22.00%	4.77%	5.43%	18521	185	139
50	Oolong tea	Masterkong	High Price	0.16%	22.00%	0.20%	0.23%	645	6	6
51	Milk tea	Masterkong	High Price	4.21%	22.00%	4.21%	4.79%	23883	239	123
52	Black tea	Uni-president	High Price	2.58%	38.00%	2.58%	2.94%	12262	123	75
53	Green tea	Uni-president	High Price	2.04%	38.00%					
54	Jasmine tea	Uni-president	High Price	4.24%	38.00%	4.24%	4.82%	14552	146	124
55	Oolong tea	Uni-president	High Price	0.15%	38.00%					
56	Milk tea	Uni-president	High Price	3.74%	38.00%	3.74%	4.25%	18765	188	109
Total				99.98%		87.90%		225432	2254	2564
Total demand=Total sales/Total purchase possibility									2564	
Note:Existing SKUs in store S are in Red										

Figure 1. Application in Beverage Sales

References

Awi, F., & Ming Hu. (2015). Multi-product price and assortment competition. *Operations Research*, *63*(3), 572-584. https://doi.org/10.1287/opre.2015.1380

Aydın, A., & Alex, G. (2014). When to carry eccentric products? Optimal retail assortment under consumer returns. *Production and Operations Management*, *23*(5), 877-892. https://doi.org/10.1111/poms.12081

Deza, A., Huang, K., & Metel, M. R. (2015). Imperfect demand estimation for new product production planning. *Mathematics*.

Dong, L., & Tian, Z. (2009). *Dynamic Pricing and Inventory Control of Substitute Products* (Vol. 11, No. 2, pp. 317-339).

Goker, A., & Sebastian Heese, H. (2014). Bargaining for an assortment. *Management Science*, *61*(3), 542-559. https://doi.org/10.1287/mnsc.2013.1854

Honhon, D., Gaur, V., & Seshadri, S. (2010). Assortment planning and inventory decisions under stockout-based substitution. *Operations Research*, *58*(5), 1364-1379. https://doi.org/10.1287/opre.1090.0805

McGregor, J. (2008). At Best Buy, marketing goes micro. *Business Week*, May 14.

Netessine, S., & Rudi, N. (2003). Centralized and competitive inventory models with demand substitution. *Operations Research*, *51*(2), 329-335. https://doi.org/10.1287/opre.51.2.329.12788

O'Connell, V. (2008). Reversing field, Macy's goes local. *Wall Street Journal*, April 21.

Parlar, M., & Goyal, S. K. (1984). Optimal ordering decisions for two substitutable products with stochastic demands. *Operations Research*, *21*(1), 1-15.

Robert, P., Harald, J., & Tammo, H. (2013). Optimizing Retail Assortments. *Marketing Science*, *32*(5), 699-715. https://doi.org/10.1287/mksc.2013.0800

Wallace, J., Hopp, & Xiaowei Xu. (2008). A static approximation for dynamic demand substitution with applications in a competitive market. *Operations Research*, *56*(3), 630-645. https://doi.org/10.1287/opre.1080.0541

Zimmerman, A. (2006). To boost sales, Wal-Mart drops one-sizefits-all approach. *Wall Street Journal*, September 7.